Papers of Oscar Lanford III

Abstract

This is an attempt to build a collected works of Oscar Lanford III, 1940-2013.

Rigorous Derivation of the Phase Shift Formula for the Hilbert Space Scattering Operator of a Single Particle

T. A. GREEN AND O. E. LANFORD, III
Wesleyan University, Middletown, Connecticut

(Received January 5, 1960)

For a single nonrelativistic particle moving in a spherically symmetric potential, the existence of the Hilbert space wave operators and S operator is proved and phase shift formulas for these operators are deduced. The probability, $P(\Omega)$, for scattering into the solid angle Ω is obtained from the time dependent theory. The relation between $P(\Omega)$ and the R matrix of the standard plane wave formulation of scattering theory is established. For collimated incoming packets, it is shown that $P(\Omega)$ can be expressed as an energy average of the differential cross section.

I. INTRODUCTION

THE importance of the asymptotic behavior of the field operators in quantum field theories has recently motivated mathematically rigorous studies of the asymptotic behavior of the solutions of the nonrelativistic Schroedinger equation.[1-5] In these studies the Hamiltonian operators of the free and interacting particle are defined as Hilbert space operators following Von Neumann[6] and Kato,[7] so that the kind of convergence involved in the asymptotic limits can be precisely specified. Suitable restrictions are placed on the scattering potential $V(\mathbf{x})$; for example, that $V(\mathbf{x})$ be square integrable over any finite region of three-dimensional space, and that as $r \to \infty$ $V(\mathbf{x})$ be $0(r^{-1-\epsilon})$, where r is the radial variable in spherical coordinates and $\epsilon > 0$. It is then possible to prove that for every Hilbert space element u (i.e., for every normalizable wave function, $u(\mathbf{x})$), there are elements $u_\pm$ belonging to the continuum subspace of the total Hamiltonian H such that as the time t approaches $\mp\infty$,

$$\exp(-iH_0t)u \to \exp(-iHt)u_\pm \tag{1.1}$$

in the sense of strong convergence in Hilbert Space. In Eq. (1.1) H_0 is the kinetic energy operator and $H = H_0 + V(\mathbf{x})$. Wave operators $\Omega_\pm$ are defined by the relations $u_\pm = \Omega_\pm u$, and it is shown that they and their adjoints $\Omega_\pm^*$ obey the relations

$$\Omega_\pm^*\Omega_\pm = 1 \tag{1.2a}$$

and

$$\Omega_\pm\Omega_\pm^* = P_c, \tag{1.2b}$$

where 1 is the unit operator and P_c is the projection operator onto the continuum subspace of H. The S operator is defined as the operator, which connects the incoming and outgoing states associated through Eq. (1.1) with a given time-dependent continuum state. It follows that

$$S = \Omega_-^*\Omega_+ \tag{1.3}$$

and that S is unitary. Equations (1.1) to (1.3) thus provide a mathematically rigorous time-dependent basis for scattering theory.

The present paper adds to the foregoing considerations in three respects. First, Eq. (1.1) is proved for potentials which are effectively $0(r^{-2+\epsilon})$ rather than $0(r^{-\frac{1}{2}+\epsilon})$ as $r \to 0$. Second, explicit phase shift formulas for $\Omega_\pm$ and S are obtained. Third, the experimentally important formula for the scattering probability as an energy average over the usual differential cross section is deduced from the time-dependent Hilbert space formalism.

The material is presented as follows. In Sec. II a well-known eigenfunction expansion for the Schroedinger equation is stated so that it can be used to define the Hamiltonian operators. In Sec. III, the Hamiltonians are defined. In Sec. IV, Eq. (1.1) is proved and the formulas for $\Omega_\pm$ and S are obtained. In Sec. V the formula for the scattering probability is derived.

This section will be concluded with a statement of the precise conditions imposed on $V(r)$. It is assumed that $V(r)$ is Lebesgue integrable over any finite interval not including the origin, that for $0 < R < \infty$

$$\int_0^R rV(r)dr < \infty, \tag{1.4a}$$

$$\int_R^\infty V(r)dr < \infty, \tag{1.4b}$$

and that either

$$\int_r^\infty V(s)ds \text{ belongs to } L^2(R,\infty), \tag{1.5a}$$

or as $r \to \infty$,

$$V(r) = 0(r^{-1-\epsilon}). \tag{1.5b}$$

The notation $L^2(a,b)$ designates the class of functions, which are Lebesgue measurable and square integrable

[1] J. M. Cook, J. Math. Phys. **36**, 82 (1957).
[2] J. M. Jauch, Helv. Phys. Acta **31**, 127 and 661 (1958).
[3] J. M. Jauch and I. I. Zinnes, Nuovo cimento **11**, 553 (1959).
[4] M. N. Hack, Nuovo cimento **9**, 731 (1958).
[5] S. T. Kuroda, Nuovo cimento **12**, 431 (1959).
[6] J. Von Neumann, *Mathematical Foundations of Quantum mechanics*, translated by R. T. Beyer (Princeton University Press, Princeton, New Jersey, 1955).
[7] Tosio Kato, Trans. Am. Math. Soc. **70**, 195 (1951).

on the interval (a,b). Equations (1.4) are used to establish the eigenfunction expansion; one or the other of Eqs. (1.5) is joined to Eqs. (1.4) in the proof of Eq. (1.1).

II. EIGENFUNCTION EXPANSION

In this section, the bound state and continuum solutions, $Y_{ml}(\theta,\phi)r^{-1}\psi_l(r)$, of the Schroedinger equation are used to generate a mean-square eigenfunction expansion of the Hilbert space elements, u, which is used in Sec. III for the definition of H and H_0. The expansion theorem could be obtained as a special case of a general theorem of Titchmarsh[8] by adapting his proof to the conditions of Eq. (1.4). However, for the simple problem under discussion, the elementary approach used here serves its purpose in a direct way in terms of formulas which the physicist will find familiar. For ease in reference in later sections, the angular and radial parts of the expansion theorem are treated separately.

Let L^2 designate the Hilbert space of complex-valued Lebesgue measurable functions, $u(x_1,x_2,x_3)$, which are square integrable on $-\infty<x_i<\infty$, $i=1, 2, 3$. Let $u(r,\theta,\phi)$ be an abbreviation for $u(r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta)$. Then $r(\sin\theta)^{\frac{1}{2}}u(r,\theta,\phi)$ is measurable and square integrable on $(0\leqslant r<\infty, 0\leqslant\theta\leqslant\pi, 0\leqslant\phi<2\pi)$. Let $Y_{ml}(\theta,\phi)$ designate the normalized spherical harmonics. As is well known, it can be shown that[9]

$$r(\sin\theta)^{\frac{1}{2}}u(r,\theta,\phi)=\underset{L\to\infty}{\text{l.i.m.}}\Sigma_L(\sin\theta)^{\frac{1}{2}}Y_{ml}(\theta,\phi)\alpha_{ml}(r), \quad (2.1)$$

where

$$\alpha_{ml}(r)=\int_{4\pi}\bar{Y}_{ml}(\theta,\phi)ru(r,\theta,\phi)d\Omega. \quad (2.2)$$

In Eq. (2.1) the notation $\sum_L$ stands for

$$\sum_{l=0}^{L}\sum_{m=-l}^{l}.$$

The notation l.i.m. means the limit in mean square on the interval $(0\leqslant r<\infty,\ 0\leqslant\theta\leqslant\pi,\ 0\leqslant\phi<2\pi)$. In Eq. (2.2), $d\Omega$ stands for $\sin\theta d\theta d\phi$ and $\int_{4\pi}$ indicates integration over $(0\leqslant\theta\leqslant\pi, 0\leqslant\phi<2\pi)$. The functions $\alpha_{ml}(r)$ belong to $L^2(0,\infty)$ and have the property that

$$\|u\|^2\equiv\int_{-\infty}^{\infty}dx_1\int_{-\infty}^{\infty}dx_2\int_{-\infty}^{\infty}dx_3|u(x_1,x_2,x_3)|^2$$
$$=\sum_{\infty}\int_0^{\infty}|\alpha_{ml}(r)|^2dr. \quad (2.3)$$

Conversely, given any set $\{\beta_{ml}(r)\}$ of functions belonging to $L^2(0,\infty)$ and such that the right-hand side of Eq. (2.3) is finite, the right-hand side of Eq. (2.1) exists and defines a function $g(x_1,x_2,x_3)$ belonging to L^2. Moreover, if $\gamma_{ml}(r)$ is the function calculated for $g(x_1,x_2,x_3)$ from Eq. (2.2), $\gamma_{ml}(r)$ equals $\beta_{ml}(r)$ almost everywhere. Equations (2.1) and (2.2) thus establish a one to one correspondence between the elements of L^2 and the sets, $\{\alpha_{ml}(r)\}$, of functions for which the right-hand side of Eq. (2.3) is finite.

Since each $\alpha_{ml}(r)$ belongs to $L^2(0,\infty)$, it can itself be expanded in mean square on $(0,\infty)$ according to

$$\alpha_{ml}(r)=\underset{N\to\infty}{\text{l.i.m.}}\sum_{n=0}^{N}\alpha_{mln}\psi_{ln}(r)$$
$$+\underset{\omega\to\infty}{\text{l.i.m.}}\int_0^{\omega}\phi_{ml}(k)\psi_l(r,k)dk, \quad (2.4)$$

where

$$\alpha_{mln}=\int_0^{\infty}\alpha_{ml}(r)\psi_{ln}(r)dr, \quad (2.5a)$$

and

$$\phi_{ml}(k)=\underset{\omega\to\infty}{\text{l.i.m.}}\int_0^{\omega}\alpha_{ml}(r)\psi_l(r,k)dr. \quad (2.5b)$$

Furthermore, for each (ml),

$$\int_0^{\infty}|\alpha_{ml}(r)|^2dr=\sum_{n=0}^{\infty}|\alpha_{mln}|^2+\int_0^{\infty}|\phi_{ml}(k)|^2dk. \quad (2.6)$$

The $\psi_{ln}(r)$, $n=0, 1, \cdots$, are the normalized eigensolutions of the radial equation

$$-u''+(l(l+1)r^{-2}+2\mu V(r))u(r)=k^2u(r), \quad (2.7)$$

for $k^2\leqslant 0$. The function $\psi_l(r,k)$ is the solution for $k>0$, which is normalized so that

$$\psi_l(r,k)\longrightarrow(2/\pi)(\sin(kr-l\pi/2+\delta_l(k))$$

as $r\to\infty$; $\delta_l(k)$ is the phase shift. For all k the solutions are $0(r^{l+1})$ as $r\to 0$. The scattered particle's mass is μ; its total energy is $k^2/2\mu$.

With each $\alpha_{ml}(r)$ belonging to $L^2(0,\infty)$, Eqs. (2.4) and (2.5) associate a function $\phi_{ml}(k)$ belonging to $L^2(0,\infty)$ and a set of constants α_{mln} such that the right-hand side of Eq. (2.6) is finite. Conversely, given a function $x_{ml}(k)$ and a set of constants β_{mln} with the above properties, Eq. (2.4) defines a function $\beta_{ml}(r)$ belonging to $L^2(0,\infty)$. If $\xi_{ml}(k)$ and γ_{mln} are calculated for $\beta_{ml}(r)$ from Eqs. (2.5), $\beta_{mln}=\gamma_{mln}$ for all n and $\xi_{ml}(k)=x_{ml}(k)$, almost everywhere. Thus Eqs. (2.4) and (2.5) establish a one to one correspondence between the $\alpha_{ml}(r)$ belonging to $L^2(0,\infty)$ and the sets $\{\phi_{ml}(k),\alpha_{mln}\}$ for which the right hand side of Eq. (2.6) is finite.

A proof of the radial expansion theorem stated above has been given by Kodaira.[10] In this proof it was assumed that $V(r)$ is continuous on $(0,\infty)$, that $V(r)=0(r^{-2+\epsilon})$ as $r\to 0$, and that $V(r)=0(r^{-1-\epsilon})$ as $r\to\infty$. These conditions are equivalent to those of Eq. (1.4) for physical applications, except that Eq. (1.4) allows discontinuous potential wells of the kind which are

[8] E. C. Titchmarsh, *Eigenfunction Expansions, Part II* (Oxford University Press, New York, 1958), Chaps. 12 and 15.

[9] A proof is given in O. E. Lanford III, Thesis, Wesleyan University, 1959, Chap. II. This paper henceforth will be referred to as I.

[10] K. Kodaira, Am. J. Math. **71**, 921 (1949).

frequently convenient in practice. One of the authors (T.A.G.) has proved the expansion theorem using Eq. (1.4). The proof will be omitted.

Equations (2.1)–(2.6) jointly establish a one to one correspondence between functions $u(x_1,x_2,x_3)$ belonging to L^2 and the sets of functions and constants $\{\phi_{ml}(k),\alpha_{mln}\}$ such that

$$\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left\{\int_0^{\infty}|\phi_{ml}(k)|^2dk+\sum_{n=0}^{\infty}|\alpha_{mln}|^2\right\}<\infty. \quad (2.8)$$

Moreover, by Eqs. (2.3) and (2.6)

$$\|u\|^2=\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left\{\int_0^{\infty}|\phi_{ml}(k)|^2dk+\sum_{n=0}^{\infty}|\alpha_{mln}|^2\right\}. \quad (2.9)$$

The set $\{\phi_{ml}(k),\alpha_{mln}\}$ will be referred to as the transform, Fu, of the Hilbert space element, u. This element is then the inverse transform, $F^{-1}\{\phi_{ml}(k),\alpha_{mln}\}$, of $\{\phi_{ml}(k),\alpha_{mln}\}$. It is easy to verify that the elements $\{\phi_{ml}(k),\alpha_{mln}\}$ such that the right-hand side of Eq. (2.9) is finite constitute a Hilbert space with a norm given by the right-hand side of Eq. (2.9) and self-evident rules for addition, etc. The transform depends on the potential. It will be convenient to denote by F_0u the transform calculated with $V(r)\equiv 0$. In this case, there are no bound states so no coefficients α_{mln} appear.

III. OPERATORS H AND H_0

The transforms introduced in Sec. II are defined in terms of the solutions of the Schroedinger equation. Hence, it is physically clear that H must be the operator multiplication by $(k^2/2\mu)$ in the space of the transforms $\{\phi_{ml}(k),\alpha_{mln}\}$ and that H_0 must be the corresponding operator for $V(r)=0$, provided that the operators thus defined are unique and self-adjoint.

For a given $V(r)$ and $l=0$, however, it is well known that Eq. (2.7) belongs to the limit circle case at $r=0$. This implies that $\psi_0(r,k)$ (and, thus, the transform) is not unique; it also implies that $\psi_0(r,k)$ is not necessarily $0(r)$ as $r\to 0$. Hence, a boundary condition must be imposed to fix $\psi_0(r,k)$ uniquely. That the boundary condition $\psi_0(r,k)=0(r)$ as $r\to 0$ is the correct one is suggested by physical considerations. It is required by the physical interpretation of the quantum theory that the free particle Hamiltonian H_0 be the self-adjoint operator multiplication by $|\mathbf{k}|^2/2\mu$ in the space of Fourier-Plancherel transforms $\hat{u}(k_1,k_2,k_3)$ of the functions $u(x_1,x_2,x_3)$ belonging to L^2. This follows from the interpretation of $|\hat{u}(k_1,k_2,k_3)|^2$ as the probability density for momentum. It may be shown[11] that the operator multiplication by $k^2/2\mu$ in the space of transforms with $V(r)=0$ is identical with H_0 if and only if $\psi_0(r,k)=0(r)$ as $r\to 0$.

The boundary condition being thus determined, the operator H is defined as follows: The element, u, whose transform is $\{\phi_{ml}(k),\alpha_{mln}\}$ is in the domain $D(H)$ of H if and only if

$$\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left\{\int_0^{\infty}|k^2\phi_{ml}(k)|^2dk+\sum_{n=0}^{\infty}|k_{ln}^2\alpha_{mln}|^2\right\}<\infty. \quad (3.1)$$

Then, by definition,

$$Hu=F^{-1}\{(k^2/2\mu)\phi_{ml}(k),(k_{ln}^2/2\mu)\alpha_{mln}\}, \quad (3.2)$$

where k_{ln}^2 is the eigenvalue of the eigenfunction $\psi_{ln}(r)$ of Eq. (2.7). H_0 is defined analogously for $V(r)=0$. It is readily verified that H and H_0 are self-adjoint operators.[12]

Having defined H and H_0, it is a straightforward matter to define the unitary operators $\exp(-iHt)$ and $\exp(-iH_0t)$, which determine the time dependence of the scattered wave packet. This is done in Chap. III of I with the expected result that if $Fu=\{\phi_{ml}(k),\alpha_{mln}\}$,

$$\exp(-iHt)f=F^{-1}\{\phi_{ml}(k)\exp(-ik^2t/2\mu), \alpha_{mln}\exp(-ik_{ln}^2t/2\mu)\} \quad (3.3)$$

A corresponding formula is valid for H_0. Equation (3.3) is the starting point in the derivation of Eq. (1.1), which is carried out in the next section.

This section will be concluded with a few remarks about the use of the eigenfunction transform as a means of defining H. The method just presented can be generalized to non spherically symmetrical potentials and to an arbitrary number of particles. The essential steps in such a program have been carried out in Chapters XII and XIII of reference 8 where the existence of a unique[13] eigenfunction transform is established on the basis of physically reasonable assumptions. The transform established by Titchmarsh can be reduced in the problem under consideration to the one established directly in Sec. II.

The eigenfunction transform method of defining H differs from that used by Kato[7] although the two methods must of course lead to the same final result. In order to point up the difference, Kato's method will be briefly described.

The kinetic energy operator is defined as the closure of the differential operator T_1, which is defined to be $-\nabla^2/2\mu$ on a suitably chosen linear manifold D_1. It is then proved that H_0 is equal to the operator, multiplication by $|\mathbf{k}|^2/2\mu$, in the space of Fourier-Plancherel transforms. With H_0 thus defined, the potential $V(x_1,x_2,x_3)$ is restricted sufficiently that Vu is defined everywhere on the domain of H_0. The total Hamiltonian, H, is defined as the closure of an operator H_1, which itself is taken to be $-\nabla^2/2\mu+V$ for elements of D_1. It is proved that $H=H_0+V$, the domain of H being

[11] See Appendix A.

[12] See Chap. III of I.

[13] In footnote 8, the requirement that for $V(r)=0$ the Green's function $G_0(\mathbf{x},\mathbf{y},E)$ be singular only at $\mathbf{x}=\mathbf{y}$ accomplishes the same result as regards uniqueness as the kinetic energy argument used above.

the same as that of H_0. Kato's simple and elegant method, which he has formulated for the many particle problem, has the merit of guaranteeing a self-adjoint Hamiltonian without requiring the introduction of eigenfunction transforms.

Because in the problem under consideration $V(r)$ is more singular than the potentials envisaged in Kato's proof, and because for a partial wave analysis the existence of the eigenfunction transform is essential to begin with, the authors found it simplest to employ the definition of H given in Eqs. (3.1) and (3.2). When Kato's conditions on $V(r)$ are joined to those in Eq. (1.4), the two definitions of H yield the same operator.

IV. ASYMPTOTIC LIMITS

The purpose of this section is to prove Eq. (1.1). Let u belong to L^2 and be such that $Fu=\{\phi_{ml}(k),0\}$ so that u is orthogonal to the subspace spanned by the bound states.[14] Let $u_t=\exp(-iHt)u$. By the expansion theorems of Sec. II and Eq. (3.3),

$$r(\sin\theta)^{\frac{1}{2}}u_t(r,\theta,\phi)=\underset{L\to\infty}{\text{l.i.m.}}\sum_L(\sin\theta)^{\frac{1}{2}}Y_{ml}(\theta,\phi)u_{ml}(r,t), \quad (4.1)$$

where

$$u_{ml}(r,t)=\underset{\omega\to\infty}{\text{l.i.m.}}\int_0^\omega \exp(-ik^2t/2\mu)\phi_{ml}(k)\psi_l(r,k)dk. \quad (4.2)$$

The asymptotic behavior of $\psi_l(r,k)$ [see below Eq. (2.7)] now motivates the consideration of the function $\tilde{u}_t(r,\theta,\phi)$ defined by

$$r(\sin\theta)^{\frac{1}{2}}\tilde{u}_t(r,\theta,\phi)=\underset{L\to\infty}{\text{l.i.m.}}\sum_L(\sin\theta)^{\frac{1}{2}}Y_{ml}(\theta,\phi)\tilde{u}_{ml}(r,t), \quad (4.3)$$

where

$$\tilde{u}_{ml}(r,t)=\underset{\omega\to\infty}{\text{l.i.m.}}\int_0^\omega \exp(-ik^2t/2\mu)\phi_{ml}(k)x_l(r,k)dk, \quad (4.4)$$

and in Eq. (4.4), $x_l(r,k)=(2/\pi)^{\frac{1}{2}}\sin(kr-l\pi/2+\delta_l(k))$. It is easy to show using the theory of Fourier transforms in $L^2(-\infty,\infty)$ that $\tilde{u}_{ml}(r,t)$ belongs to $L^2(0,\infty)$ for all t and that

$$\int_0^\infty|\tilde{u}_{ml}(r,t)|^2dr\leqslant 2\int_0^\infty|\phi_{ml}(k)|^2dk. \quad (4.5)$$

As the first main step in the derivation of Eq. (1.1), it will now be shown that

$$\lim_{|t|\to\infty}\|u_t-\tilde{u}_t\|=0. \quad (4.6)$$

By Eq. (2.3)

$$\|u_t-\tilde{u}_t\|^2=\sum_{l=0}^\infty\sum_{m=-l}^{l}\int_0^\infty|u_{ml}(r,t)-\tilde{u}_{ml}(r,t)|^2dr. \quad (4.7)$$

Minkowski's inequality applies to the integrals of Eq. (4.7). Therefore, by using Eq. (4.5) and the corresponding equation for $\int_0^\infty|u_{ml}(r,t)|^2dr$, which follows from Eq. (2.6), it is seen that the convergence of the series on the right-hand side of Eq. (4.7) is uniform with respect to t for $-\infty<t<\infty$. Therefore, if

$$\lim_{|t|\to\infty}\int_0^\infty|u_{ml}(r,t)-\tilde{u}_{ml}(r,t)|^2dr=0 \quad (4.8)$$

for all (l,m), Eq. (4.6) is valid.

The rest of the discussion requires $k>0$. For this reason, functions $u_{mlN}(r,t)$ and $\tilde{u}_{mlN}(r,t)$ are defined by restricting the k integration in Eqs. (4.2) and (4.4) to the interval $[N^{-1},N]$, $(1<N<\infty)$. It is not hard to prove (see p. 55 of I) that

$$\int_0^\infty|u_{ml}(r,t)-\tilde{u}_{ml}(r,t)|^2dr\to 0 \quad \text{as} \quad |t|\to\infty$$

if

$$\lim_{|t|\to\infty}\int_0^\infty|u_{mlN}(r,t)-\tilde{u}_{mlN}(r,t)|^2dr=0 \quad (4.9)$$

for all N. Now

$$u_{mlN}(r,t)-\tilde{u}_{mlN}(r,t)=\int_{1/N}^{N}\exp(-ik^2t/2\mu)\phi_{ml}(k)\times[\psi_l(r,k)-x_l(r,k)]dk. \quad (4.10)$$

Also, for all r and N, $\phi_{ml}(k)(\psi_l(r,k)-x_l(r,k))$ is summable on $[1/N,N]$. Hence, the Riemann-Lebesgue lemma shows that

$$\lim_{|t|\to\infty}[u_{mlN}(r,t)-\tilde{u}_{mlN}(r,t)]=0 \quad (4.11)$$

for all $0\leqslant r<\infty$. Consequently, if in Eq. (4.9) the limit can be carried under the integral sign, the proof that $\|u_t-\tilde{u}_t\|\to 0$ will be accomplished. Consider first

$$\int_0^R|u_{mlN}(r,t)-\tilde{u}_{mlN}(r,t)|^2dr.$$

For $0\leqslant r\leqslant R$ and $1/N\leqslant k\leqslant N$, $\psi_l(r,k)-x_l(r,k)$ is bounded. Hence, by Eq. (4.10) $|u_{mlN}(r,t)-\tilde{u}_{mlN}(r,t|\leqslant K$ for all t and consequently, for all $1<N<\infty$ and all $0<R<\infty$

$$\lim_{|t|\to\infty}\int_0^R|u_{mlN}(r,t)-\tilde{u}_{mlN}(r,t)|^2dr=\int_0^R\lim_{|t|\to\infty}|u_{mlN}(r,t)-\tilde{u}_{mlN}(r,t)|^2dr=0. \quad (4.12)$$

It is therefore sufficient to show that

$$\lim_{R\to\infty}\int_R^\infty|u_{mlN}(r,t)-\tilde{u}_{mlN}(r,t)|^2dr=0, \quad (4.13)$$

uniformly with respect to t for $-\infty<t<\infty$.

[14] The elements, u, constitute what has been referred to as the continuum subspace of H in earlier sections.

One sufficient condition is readily obtained from the asymptotic formula

$$\psi_l(r,k)-x_l(r,k)=0\left[\int_r^\infty V(y)dy\right]+0(1/r), \quad (4.14)$$

for $k>0$ and $r\to\infty$. [Equation (4.14) is readily deduced from Eq. (4.16).] Suppose $\int_r^\infty |V(y)|dy$ belongs to $L^2(R,\infty)$ for sufficiently large R. Then the Schwarz inequality applied to Eq. (4.10) shows that for all t and sufficiently large r,

$$|u_{mlN}(r,t)-\tilde{u}_{mlN}(r,t)|^2\leqslant g(r), \quad (4.15)$$

where $g(r)$ belongs to $L(R,\infty)$. Thus the condition expressed by Eq. (4.13) is satisfied. Consequently, Eq. (4.6) is valid.

The above condition on $V(r)$ can be replaced by the condition, $V(r)=0(r^{-1-\epsilon})$ as $r\to\infty$, for some $\epsilon>0$. This proved as follows. For $k>N^{-1}$ and $r>R(N,\epsilon)$, $\psi_l(r,k)$ satisfies the integral equation,

$$\psi_l(r,k)=x_l(r,k)-1/k\int_r^\infty \sin k(r-s)\times q(s)\psi_l(s,k)ds, \quad (4.16)$$

where $q(s)=l(l+1)/s^2+2\mu V(s)$. It follows from the iteration of Eq. (4.16) that as $r\to\infty$,

$$\psi_l(r,k)=x_{nl}(r,k)+0(r^{-(n+1)\epsilon}), \quad (4.17)$$

where $x_{nl}(r,k)$ is the function obtained by iterating Eq. (4.16) n times. Given ϵ, n can be chosen so that $n\epsilon>1$. This suffices to make $\psi_l(r,k)-x_{nl}(r,k)$ belong to $L^2(R,\infty)$ so that the argument below Eq. (4.13) can be applied to $\psi_l(r,k)-x_{nl}(r,k)$. Furthermore,

$$x_{nl}(r,k)-x_l(r,k)=\int_r^\infty dr_1G_k(r,r_1)x_l(r_1)+\cdots+\int_r^\infty dr_1\int_{r_1}^\infty dr_2\cdots\int_{r_{n-1}}^\infty dr_nG_k(r,r_1)G_k(r_1,r_2)\cdots\times G_k(r_{n-1},r_n)x_l(r_n), \quad (4.18)$$

where $G_k(x,y)=-k^{-1}q(y)\sin k(x-y)$.

With reference to Eq. (4.10), now consider

$$\xi(r,t)\equiv\int_{1/N}^N \exp(-ik^2t/2\mu)\phi_{ml}(k)\times[x_{nl}(r,k)-x_l(r,k)]dk. \quad (4.19)$$

For all r and t, Eq. (4.18) can be substituted into Eq. (4.19) and the k integral carried out first in each of the terms of the resulting sum. Moreover, the products

$$k^{-p}\sin k(r-r_1)\cdots\sin k(r_{p-1}-r_p)\sin(kr_p-l\pi/2+\delta_l(k))$$

can be decomposed into a sum of 2^p terms of the form $\sin(kZ-l\pi/2+\delta_l(k))$, or $\cos(kZ-l\pi/2+\delta_l(k))$ where Z is of the form $2r_i-2r_j+-+\cdots\pm r$.[15] In the definition of Z, r_i, r_j, etc., are selected from $r_1, r_2, r_3, \cdots, r_p$, and each distinct combination of 0, 1, 2, $\cdots$, p of them appears exactly once. Let

$$g_p(Z,t)=\int_{1/N}^N \exp(-ik^2t/2\mu)\phi_{ml}(k)k^{-p}\times\sin[kZ-l\pi/2+\delta_l(k)]dk. \quad (4.20)$$

By the theory of Fourier transforms

$$\int_0^\infty dZ|g_p(Z,t)|^2\leqslant\pi\int_{1/N}^N k^{-2p}|\phi_{ml}(k)|^2dk. \quad (4.21)$$

If $h_p(Z,t)$ is defined by Eq. (4.20) with $\cos(kZ-l\pi/2+\delta_l(k))$ in place of $\sin(kZ-l\pi/2+\delta_l(k))$, Eq. (4.21) applies with $h_p(Z,t)$ in place of $g_p(Z,t)$. With the k integrations done, $\xi(r,t)$ is given in part by a sum of terms of the form

$$\int_r^\infty dr_1q(r_1)\int_{r_1}^\infty dr_2q(r_2)\cdots\int_{r_{i-1}}^\infty dr_iq(r_i)g_p(Z,t)\times\int_{r_i}^\infty dr_{i+1}q(r_{i+1})\cdots\int_{r_{p-1}}^\infty dr_pq(r_p), \quad (4.22)$$

where Z contains r_i but none of the r_l for $l>i$. In addition, there are analogous terms with $h_p(Z,t)$ in place of $g_p(Z,t)$. Finally, there are terms with $g_p(r,t)$ and $h_p(r,t)$ which factor out of the integrals over the r_i. By applying the Schwarz inequality and Eq. (4.21) to the integrals containing $g_p(Z,t)$ and $h_p(Z,t)$, and by noting that as $r\to\infty$ $q(r)=0(r^{-1-\epsilon})$, it is readily verified that for all t as $r\to\infty$,

$$\xi(r,t)=g_p(r,t)0(r^{-\epsilon})+h_p(r,t)0(r^{-\epsilon})+0(r^{-\frac{1}{2}(1+\epsilon)}), \quad (4.23)$$

for all fixed l, m, and N. In Eq. (4.10), $(\psi_l(r,k)-x_l(r,k))$ is now written as $(\psi_l(r,k)-x_{nl}(r,k))+(x_{nl}(r,k)-x_l(r,k))$. It then follows from Eq. (4.17) (with $n\epsilon>1$) and Eq. (4.19) that as $r\to\infty$,

$$u_{mlN}(r,t)-\tilde{u}_{mlN}(r,t)=\xi(r,t)+0(r^{-1-\epsilon}) \quad (4.24)$$

Finally, Eqs. (4.24), (4.23), and (4.21) show that as $R\to\infty$,

$$\int_R^\infty |u_{mlN}(r,t)-\tilde{u}_{mlN}(r,t)|^2=0(R^{-\epsilon}) \quad (4.25)$$

for all ϵ, l, m, N, and t. Therefore, Eq. (4.13) is satisfied and the validity of Eq. (4.6) is established.

The last step in the discussion is the proof that as $t\to\pm\infty$, $\tilde{u}(r,t)$ approaches its outgoing and incoming parts, respectively. Let $\phi_{ml}(k)$ be the function in Eq.

[15] If p is even, sine functions are obtained; if p is odd, cosine functions occur. If the number of factors r_i, r_j is even, r enters with a plus sign.

(4.2) and by definition let

$$r(\sin\theta)^{\frac{1}{2}}u_t^{\pm}(r,\theta,\phi) = \underset{L\to\infty}{\text{l.i.m.}} \sum_L (\sin\theta)^{\frac{1}{2}} Y_{ml}(\theta,\phi) u_{ml}^{\pm}(r,t), \quad (4.26)$$

where

$$u_{ml}^{\pm}(r,t) = \underset{\omega\to\infty}{\text{l.i.m.}} \int_0^{\omega} \exp(-ik^2t/2\mu)\phi_{ml}(k) \times (2\pi)^{-\frac{1}{2}} \exp[\pm i(kr-(l+1)\pi/2+\delta_l(k))]dk. \quad (4.27)$$

The $u_{ml}^{\pm}(r,t)$ belong to $L^2(0,\infty)$ for all t and their norms satisfy Eq. (4.5) without the factor of two. Moreover, by comparing Eqs. (4.3) and (4.4) with Eqs. (4.26) and (4.27) it is seen that

$$\tilde{u}_t = u_t^{+} + u_t^{-}. \quad (4.28)$$

It will be shown at the end of this section that

$$\lim_{t\to\mp\infty} \|u_t^{\pm}\| = 0 \quad (4.29)$$

Therefore, by Eqs. (4.28), (4.6), and the definition of u_t above Eq. (4.1), if $Fu=\{\phi_{ml}(k),0\}$,

$$\lim_{t\to\pm\infty} \|[\exp(-iHt)]u - u_t^{\pm}\| = 0, \quad (4.30)$$

where $u_t^{\pm}$ are defined by Eqs. (4.26) and (4.27).

The desired asymptotic limits follow directly from Eq. (4.30). Let g belong to L^2 and let $F_0 g=\{\chi_{ml}(k)\}$. Equation (4.30) applies to g in the form in which H is replaced by H_0 and the $u_t^{\pm}$ are replaced by functions $g_t^{\pm}$, which are defined by replacing $\phi_{ml}(k)$ by $\chi_{ml}(k)$ and setting $\delta_l(k)$ equal to zero in Eqs. (4.26) and (4.27). Now let $g_{\pm}=F^{-1}\{\chi_{ml}(k)\exp(\pm i\delta_l(k)),0\}$. The application of Eqs. (4.30), (4.26), and (4.27) to each of these functions shows that

$$\lim_{t\to\mp\infty} \|e^{-iHt}g_{\pm} - e^{-iH_0t}g\| = 0. \quad (4.31)$$

Thus Eq. (1.1) is established.

The phase shift formulas for the wave operators can be given concisely in terms of F and F_0. In order to do this, the element $\{\theta_{ml}(k)\}=F_0 u$ is identified with the element $\{\theta_{ml}(k),0\}$ of the Hilbert space Γ consisting of all $\{\phi_{ml}(k),\alpha_{mln}\}$ such that the right-hand side of Eq. (2.9) is finite. With this convention, F and F^{-1} establish a one to one correspondence between L^2, and Γ while F_0 and F_0^{-1} establish a one to one correspondence between L^2 and the continuum subspace of Γ. The formulas for $\Omega_{\pm}$ are now very simple. By the definition of $\Omega_{\pm}$ below Eq. (1.1) and the definition of $g_{\pm}$ below Eq. (4.30),

$$\Omega_{\pm} = F^{-1}\exp(\pm i\delta_l(k))F_0. \quad (4.32)$$

By using (4.32) and the norm-preserving properties of F and F_0, it is easy to show that

$$\Omega_{\pm}^{*} = F_0^{-1}[\exp(\mp i\delta_l)]\tilde{P}_c F, \quad (4.33)$$

where $\tilde{P}_c$ is the projection operator for the continuum subspace of Γ, $(\tilde{P}_c\{\phi_{ml}(k),\alpha_{mln}\}=\{\phi_{ml}(k),0\})$. Equations (1.2) follow directly from Eqs. (4.32) and (4.33). Finally, from Eqs. (1.3), (4.32) and (4.33) it is seen that

$$S = F_0^{-1}[\exp(2i\delta_l(k))]F_0. \quad (4.34)$$

The relation of the Hilbert space operator, S, to the R matrix of the plane wave formulation of scattering theory will be taken up in the next section. This section will be concluded with an outline of the proof of Eq. (4.29), the complete details of which are given in Chapter IV of I.

By Eqs. (4.26) and (2.3), Eq. (4.29) will hold as $t\to\infty$ if

$$\lim_{t\to\infty} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \int_0^{\infty} |u_{ml}^{-}(r,t)|^2 dr = 0. \quad (4.35)$$

The series in Eq. (4.35) converges uniformly with respect to t, so it remains to be shown that the integrals tend toward zero. This is done by approximating $\phi_{ml}(k)\exp(-i\delta_l(k))$ (Eq. (4.27)) in mean square by a step function zero near the origin and zero for large k. This reduces the problem to the consideration of integrals of the type

$$\int_0^{\infty} \left| \int_a^b dk \exp(-ik^2t/(2\mu) - ikr) \right|^2 dr, \quad (4.36)$$

where $0<a<b<\infty$. For sufficiently large t, and all r, it can be shown that

$$\left| \int_a^b dk \exp(-ik^2t/(2\mu) - ikr) \right|^2 < A(r^2+B)^{-1}, \quad (4.37)$$

where A and B are positive constants, Moreover, the integral over k tends toward zero by the Riemann-Lebesgue lemma. Thus the $\lim(t\to\infty)$ can be taken inside the integrals over r in Eq. (4.36) and the limit is zero. Therefore Eq. (4.29) is valid insofar as u_t^{-} is concerned. The proof for $t\to-\infty$ is obtained by an identical argument.

V. RELATION OF S TO THE R MATRIX OF THE PLANE WAVE THEORY AND TO THE SCATTERING CROSS SECTION

In this section, the probability $P(\Omega)$ for scattering into a given solid angle, Ω, is computed from the time dependent formalism. The conditions under which $P(\Omega)$ can be described in terms of the R matrix are then discussed. Finally, a mathematically nonrigorous, but physically convincing argument is given, which shows that for wave packets of the type used in conventional scattering experiments,

$$P(\Omega) = \sigma(\Omega)P(\mathbf{a}), \quad (5.1)$$

where $\sigma(\Omega)$ is the usual scattering cross section averaged over energy, and $P(\mathbf{a})$ is the two dimensional proba-

bility density for the incident particle to strike the point, **a**, where the scatterer is located in a plane perpendicular to the motion of the incident particle. This is the result which one would desire for it guarantees that when multiple scattering and interference effects can be neglected the average number of particles scattered into Ω for N incident particles is equal to $Nt\rho\sigma(\Omega)$, where t is the target thickness and ρ the number of scatterers per unit volume.

The formula for $P(\Omega)$ is obtained as follows. Let $V(\Omega; a,b)$ designate the region $(0\leqslant a\leqslant r\leqslant b\leqslant\infty$, $\theta_0\leqslant\theta\leqslant\theta_1$, $\phi_0\leqslant\phi\leqslant\phi_1)$. Let $u_t=\exp(-iHt)u$, where $Fu=\{\phi_{ml}(k),0\}$ as in Sec. IV and consider the probability

$$P_t(\Omega; a,b)=\int_{V(\Omega;a,b)}|u_t|^2d\mathbf{x} \tag{5.2}$$

that the scattered particle be in $V(\Omega; a,b)$ at time t. From Eqs. (4.26), (4.27) and (4.30) it is easy to see that

$$\lim_{t\to\pm\infty}\left(P_t(\Omega; a,b)-\int_{V(\Omega;a,b)}|u_t^{\pm}(r,\theta,\phi)|^2d\mathbf{x}\right)=0, \tag{5.3}$$

and that for all t

$$\int_{V(\Omega;a,b)}|u_t^{\pm}(r,\theta,\phi)|^2d\mathbf{x}=\lim_{L\to\infty}\sum_{l=0}^{L}\sum_{m=-l}^{l}\sum_{l'=0}^{L}\sum_{m'=-l'}^{l'}\times\int_{\Omega}Y_{ml}(\theta,\phi)\bar{Y}_{m'l'}(\theta,\phi)d\Omega\int_a^b u_{ml}^{\pm}(r,t)\bar{u}_{m'l'}^{\pm}(r,t)dr, \tag{5.4}$$

the convergence of the series being uniform with respect to t. From Eq. (4.27) and the theory of Fourier transforms

$$\int_{-\infty}^{\infty}u_{ml}^{\pm}(r,t)\bar{u}_{m'l'}^{\pm}(r,t)dr=\int_0^{\infty}\exp[\pm i(\delta_l(k)-\delta_{l'}(k)-(l-l')\pi/2)]\times\phi_{ml}(k)\bar{\phi}_{m'l'}(k)dk. \tag{5.5}$$

Furthermore, for any finite c

$$\int_{-\infty}^{c}u_{ml}^{+}(r,t)\bar{u}_{m'l'}^{+}(r,t)dr=\int_0^{\infty}u_{ml}^{+}(c-s,t)\bar{u}_{m'l'}^{+}(c-s,t)ds, \tag{5.6}$$

where, by Eq. (4.27),

$$u_{ml}^{+}(c-s,t)=\text{l.i.m.}_{\omega\to\infty}\int_0^{\omega}\exp(-ik^2t/2\mu)\phi_{ml}(k)(2\pi)^{-\frac{1}{2}}\exp(ikc)\times\exp[i(-ks-(l+1)\pi/2+\delta_l(k))]dk. \tag{5.7}$$

Now, to within a factor $\exp[i(kc-(l+1)\pi+2\delta_l(k))]$, $u_{ml}^{+}(c-s,t)$ has the same form as $u_{ml}^{-}(s,t)$. Hence, it is easily seen from the arguments below Eq. (4.35) and the Schwarz inequality that, for all finite c,

$$\lim_{t\to\infty}\int_{-\infty}^{c}u_{ml}^{+}(r,t)\bar{u}_{m'l'}^{+}(r,t)dr=0. \tag{5.8}$$

The same kind of argument applies to $u_{ml}^{-}(r,t)$ for $t\to-\infty$. Therefore, by Eqs. (5.2), (5.3), (5.4), and (5.8), for all $0\leqslant a<\infty$,

$$\begin{aligned}\lim_{t\to\pm\infty}P_t(\Omega; a,\infty)&=\lim_{L\to\infty}\sum_{l=0}^{L}\sum_{m=-l}^{l}\sum_{l'=0}^{L}\sum_{m'=-l'}^{l'}\int_{\Omega}Y_{ml}(\theta,\phi)\bar{Y}_{m'l'}(\theta,\phi)d\Omega\\&\times\int_0^{\infty}\exp[\pm i(\delta_l(k)-\delta_{l'}(k)-(l-l')\pi/2)]\times\phi_{ml}(k)\bar{\phi}_{m'l'}(k)dk\\&=\lim_{L\to\infty}\int_0^{\infty}dk\int_{\Omega}d\Omega\Big|\sum_{l=0}^{L}\sum_{m=-l}^{l}Y_{ml}(\theta,\phi)\phi_{ml}(k)\\&\times\exp[\pm i(\delta_l-l\pi/2)]\Big|^2.\end{aligned} \tag{5.9}$$

For finite a and b the limit is zero. Thus, the scattered particle is asymptotically outside of any sphere of finite radius a.

The probability, $P(\Omega)$, for scattering into the solid angle Ω should clearly be defined by the relation

$$P(\Omega)=\lim_{t\to\infty}P_t(\Omega,a,\infty). \tag{5.10}$$

Equation (5.9) then provides a formula for $P(\Omega)$ in terms of the phase shifts and the properties of the incident wave packet. The formula can be rendered more concise in terms of the Fourier transforms of the incoming and outgoing wave packets. As was shown in Sec. IV, as $t\to\infty$, $u_t\to\exp(-iH_0t)u^{\pm}$, where $F_0u^{\pm}=\{\phi_{ml}(k)\exp(\pm i\delta_l(k))\}$. Furthermore, as is proved in Appendix A, the Fourier-Plancherel transforms $\hat{u}^{\pm}(k,\theta,\phi)$ of $u^{\pm}$ satisfy the relation

$$k(\sin\theta)^{\frac{1}{2}}\hat{u}^{\pm}(k,\theta,\phi)=\lim_{L\to\infty}\sum_{l=0}^{L}\sum_{m=-l}^{l}(\sin\theta)^{\frac{1}{2}}Y_{ml}(\theta,\phi)\times(-i)^l\phi_{ml}(k)\exp(\pm i\delta_l(k)). \tag{5.11}$$

Consequently, by Eqs. (5.10), (5.9), and (5.11)

$$P(\Omega)=\int_0^{\infty}\int_{\Omega}|\hat{u}^{+}(k,\theta,\phi)|^2k^2dkd\Omega. \tag{5.12}$$

The physical interpretation of Eq. (5.12) is straightforward. The probability that the particle be scattered into Ω is equal to the probability that the momentum vector of the outgoing packet lie in Ω. This well-known

result, which has just been shown to be a rigorous consequence of the time-dependent formalism, is the basis for the physical interpretation of calculations in which $\hat{u}^+(k,\theta,\phi)$ is obtained from a time-independent formalism.

The connection of the Hilbert space formulas with the R matrix can now be readily deduced. Let $P_-(\Omega)$ designate the probability that the incident particle be scattered into Ω in the absence of the scatterer. [Use $\hat{u}^-(k,\theta,\phi)$ in place of $\hat{u}^+(k,\theta,\phi)$ in Eq. (5.12).] Let

$$P'(\Omega)=\int_0^\infty\int_\Omega|\hat{u}^+(k,\theta,\phi)-\hat{u}^-(k,\theta,\phi)|^2k^2dkd\Omega. \quad (5.13)$$

It is easy to prove that

$$|P'(\Omega)-P(\Omega)|\leqslant P_-(\Omega)+2(P_-(\Omega)P(\Omega))^{\frac{1}{2}}. \quad (5.14)$$

Therefore, if the incident beam is appropriately collimated, the scattering probability can be calculated accurately from $P'(\Omega)$ except near the forward direction. Now, by Eq. (5.11),

$$\begin{aligned}&k(\sin\theta)^{\frac{1}{2}}(\hat{u}^+(k,\theta,\phi)-\hat{u}^-(k,\theta,\phi))\\&=\text{l.i.m.}_{L\to\infty}\sum_{l=0}^{L}\sum_{m=-l}^{l}(\sin\theta)^{\frac{1}{2}}Y_{ml}(\theta,\phi)\\&\quad\times[\exp(2i\delta_l(k))-1](-i)^l\phi_{ml}(k)\exp(-i\delta_l(k))\\&=\text{l.i.m.}_{L\to\infty}\int_{4\pi}k(\sin\theta)^{\frac{1}{2}}R_L(\theta,\phi;\theta',\phi';k)\\&\qquad\times\hat{u}^-(k,\theta',\phi')d\Omega', \quad (5.15)\end{aligned}$$

where

$$\begin{aligned}&R_L(\theta,\phi;\theta',\phi';k)\\&=\sum_{l=0}^{L}\sum_{m=-l}^{l}Y_{ml}(\theta,\phi)\bar{Y}_{ml}(\theta',\phi')[\exp(2i\delta_l(k))-1],\\&=\sum_{l=0}^{L}(4\pi)^{-1}(2l+1)P_l(\cos\Theta)\\&\qquad\times[\exp(2i\delta_l(k))-1]. \quad (5.16)\end{aligned}$$

In obtaining Eq. (5.16), the addition theorem for spherical harmonics was used. The angle Θ is the angle between the vectors $\mathbf{k}'(k,\theta',\phi')$ and $\mathbf{k}(k,\theta,\phi)$. Aside from a multiplicative factor, $R_L(\theta,\phi;\theta',\phi';k)$ is just the sum of the first L terms of the series for the scattering amplitude which appears in the stationary-state formulation of scattering theory for monochromatic incident plane waves. Suppose that the series (5.16) converges to a function $R(\theta,\phi;\theta',\phi';k)$ in such a way that

$$\begin{aligned}&\lim_{L\to\infty}\int_{4\pi}k(\sin\theta)^{\frac{1}{2}}R_L(\theta,\phi;\theta',\phi';k)\hat{u}^-(k,\theta',\phi')d\Omega'\\&\quad=\int_{4\pi}k(\sin\theta)^{\frac{1}{2}}R(\theta,\phi;\theta',\phi';k)\hat{u}^-(k,\theta'\phi')d\Omega' \quad (5.17)\end{aligned}$$

for almost all (k,θ,ϕ). Then, the limit functions in Eqs. (5.15) and (5.17) are equal almost everywhere, and

$$\begin{aligned}&\hat{u}^+(k,\theta,\phi)-\hat{u}^-(k,\theta,\phi)\\&\quad=\int_{4\pi}R(\theta,\phi;\theta',\phi';k)\hat{u}^-(k,\theta',\phi')d\Omega'. \quad (5.18)\end{aligned}$$

In this case, the scattering probability can be calculated from the incoming wave packet through Eqs. (5.18) and (5.13). The relation of $R(\theta,\phi;\theta',\phi';k)$ to the R matrix is the following. The R matrix, $R(\mathbf{k},\mathbf{k}')$, is defined by the formal relation[16]

$$\begin{aligned}&\hat{u}^+(k,\theta,\phi)-\hat{u}^-(k,\theta,\phi)\\&\quad=\int\int\int(-2\pi i)R(\mathbf{k},\mathbf{k}')\delta(E-E')\hat{u}^-(\mathbf{k}')d\mathbf{k}', \quad (5.19)\end{aligned}$$

where $E=k^2/2\mu$ and $\delta(E-E')$ is the Dirac delta function. Equation (5.19) means

$$\begin{aligned}&\hat{u}^+(k,\theta,\phi)-\hat{u}^-(k,\theta,\phi)\\&\quad=-2\pi ik\mu\int_{4\pi}R(\mathbf{k},\mathbf{k}')\hat{u}^-(\mathbf{k}')d\Omega', \quad (5.20)\end{aligned}$$

where $|\mathbf{k}|=|\mathbf{k}'|$. By comparing Eqs. (5.20) and (5.18), it is seen that the R matrix is defined on the energy shell whenever the limit $R(\theta,\phi;\theta',\phi';k)$ of $R_L(\theta,\phi;\theta',\phi';k)$ exists and Eq. (5.17) is valid.

From the physical point of view, there is no point in discussing potentials for which Eq. (5.18) does not hold, because if the series in Eq. (5.16) does not converge fairly rapidly, the phase shift approach will be useless for computation anyway. It is possible, of course, to contemplate potentials for which the series (5.16) diverges for $\Theta=0$ since in practice the calculation of nonforward scattering using Eqs. (5.13) and (5.18) need not require integration over $\Theta=0$. The convergence of the series (5.16) and the validity of Eq. (5.17) can be tested by using the Born approximation for the phase shifts.[17] As is well known, it is sufficient for convergence for $\Theta\neq0$ that as $r\to\infty$ $V(r)=0(r^{-2-\epsilon})$, $\epsilon>0$.[18] The stronger condition $V(r)=0(r^{-3-\epsilon})$ is sufficient to guarantee absolute and uniform convergence for $0\leqslant\theta\leqslant\pi$.

The usual formula for $P'(\Omega)$ in terms of the differential scattering cross section is obtained by specializing $\hat{u}^-(k,\theta,\phi)$ so that it conforms to experimental conditions. This has been done by Ekstein,[16] Eisenbud,[19]

[16] See, for example, H. Eckstein, Phys. Rev. **101**, 880 (1956).
[17] D. S. Carter, Thesis, Princeton University, 1952 (unpublished).
[18] L. I. Schiff, *Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York, 1949), 1st ed., p. 187, problem 5.
[19] L. Eisenbud, Thesis, Princeton University, 1948 (unpublished).

and Jauch.[20] An alternative formulation, based on the same physical arguments, is presented below.

Suppose that the scatterer is located at the point whose cartesian coordinates are $(a_1,a_2,0)$ in the reference frame in which the scattered beam is directed along the positive x_3 axis. If the change of location of the scatterer from the origin to $(a_1,a_2,0)$ is taken into account in the usual way, it follows from Eqs. (5.13) and (5.18) that

$$P'(\Omega)=\int_\Omega d\Omega\int_0^\infty dk\left|\int_{4\pi}R(\theta,\phi;\theta',\phi';k)\times\exp(i\mathbf{k}'\cdot\mathbf{a})\hat{u}^-(k,\theta',\phi')d\Omega'\right|^2. \quad (5.21)$$

Now, with a typical beam (beam diam ~ 1 cm, momentum $\sim 10^8$ cm^{-1}), $\hat{u}^-(k,\theta',\phi')$ goes to zero strongly outside a forward cone of apex angle $\sim 10^{-8}$ rad centered on the x_3 axis. Thus, in cases of physical interest $R(\theta,\phi;\theta',\phi';k)$ can certainly be replaced by $R(\theta,\phi;0,0;k)$. This leads to

$$P'(\Omega)\approx\int_\Omega d\Omega\int_0^\infty dk\sigma_k(\theta,\phi)\times\left|(2\pi)^{-1}\int_0^\pi\int_0^{2\pi}\exp(i\mathbf{k}'\cdot\mathbf{a})\hat{u}^-(k,\theta',\phi')\times k^2\sin\theta'd\theta'd\phi'\right|^2, \quad (5.22)$$

where

$$\sigma_k(\theta,\phi)=|k^{-1}\sum_{l=0}^\infty(2l+1)P_l(\cos\theta)\exp(i\delta_l(k))\sin\delta_l(k)|^2. \quad (5.23)$$

It will be recognized that $\sigma_k(\theta,\phi)$ is the differential cross section as usually defined. Equation (5.22) can be further transformed by noting that with $k\sim 10^8$ and $\theta'\sim 10^{-8}$, it will be a very good approximation to write[21]

$$(2\pi)^{-1}\int_0^\pi\int_0^{2\pi}\exp(i\mathbf{k}'\cdot\mathbf{a})\hat{u}^-(k,\theta',\phi')k^2\sin\theta'd\theta'd\phi'$$
$$\approx(2\pi)^{-1}\int_{-\infty}^\infty\int_{-\infty}^\infty\exp(i\mathbf{k}'\cdot\mathbf{a})\hat{u}^-(k_1,k_2,k)dk_1dk_2$$
$$\approx(2\pi)^{-\frac{1}{2}}\int_{-\infty}^\infty\exp(-ikx_3)u^-(a_1,a_2,x_3)dx_3. \quad (5.24)$$

Finally, with conventional collimation, it should be possible to describe the beam in terms of packets of the form

$$u^-(x_1,x_2,x_3)=g(x_1,x_2)e(x_3). \quad (5.25)$$

In this event, Eq. (5.22) becomes

$$P'(\Omega)\approx P(\mathbf{a})\int_\Omega d\Omega\int_0^\infty dk\sigma_k(\theta,\phi)|\hat{e}(k)|^2, \quad (5.26)$$

where $\hat{e}(k)$ is the Fourier transform of $e(x_3)$[$\hat{e}(k)=0$ for $k<0$] and $P(\mathbf{a})=|g(a_1,a_2)|^2$. Since $u^-(x_1,x_2,x_3)$ is normalized, $\int_{-\infty}^\infty\int_{-\infty}^\infty P(\mathbf{a})da_1da_2=1$ and $\int_0^\infty|\hat{e}(k)|^2=1$. Equation (5.26), which because of Eq. (5.14) is practically equivalent to Eq. (5.1), is the final result. It shows how the cross section is to be averaged over the energy spectrum of the incoming beam, and shows explicity through $P(\mathbf{a})$ how the scattering decreases when the target is not in the center of the beam.

ACKNOWLEDGMENT

The authors wish to extend their thanks to Professor H. Tong for his clarification of numerous mathematical questions connected with this paper.

APPENDIX A

In this appendix, the relation between the Fourier-Plancherel transform, $\hat{u}(k_1,k_2,k_3)$, of $u(x_1,x_2,x_3)$ and the transform, $\{\phi_{ml}(k)\}$, for $V=0$ is established. Let $u_n(r,\theta,\phi)$ be equal to $u(r,\theta,\phi)$ for $0\leqslant r\leqslant n$ and zero otherwise. Because of the norm-preserving properties of both transforms and because $u_n\to u$ in mean square, $\hat{u}_n\to\hat{u}$, and $\phi_{mln}(k)\to\phi_{ml}(k)$ in mean square. ($\{\phi_{mln}(k)\}\equiv F_0u_n$.) Since $\hat{u}_n$ and $\hat{u}$ belong to L^2, they possess expansions of the form given in Eqs. (2.1)–(2.3). Let $\gamma_{mln}(k)$ and $\gamma_{ml}(k)$ correspond, respectively, to the quantity called $\alpha_{ml}(r)$ in these equations. Clearly, $\gamma_{mln}(k)\to\gamma_{ml}(k)$ in mean square. Furthermore,

$$\gamma_{mln}(k)=k\int_0^\pi\int_0^{2\pi}\bar{Y}_m(\theta,\phi)d\Omega(2\pi)^{-\frac{1}{2}}\times\int_0^n r'^2dr'\int_0^\pi\int_0^{2\pi}\exp(-i\mathbf{k}\cdot\mathbf{r}')u(r',\theta',\phi')d\Omega', \quad (A1)$$

where $\mathbf{k}$ is the radius vector to the point (k,θ,ϕ). In Eq. (A1), the order of integration can be reversed and the exponential can be expanded in terms of spherical harmonics and Bessel functions. In this way, there results

$$\gamma_{mln}(k)=(-i)^l\int_0^n\int_0^\pi\int_0^{2\pi}(kr)^{\frac{1}{2}}J_{l+\frac{1}{2}}(kr)\times\bar{Y}_{ml}(\theta,\phi)ru(r,\theta,\phi)drd\Omega$$
$$=(-i)^l\phi_{mln}(k). \quad (A2)$$

The last equality in Eq. (A2) follows from Eqs. (2.2) and (2.5) and the fact that for $V=0$ $\psi_l(r,k)$ is equal to $(kr)^{\frac{1}{2}}J_{l+\frac{1}{2}}(kr)$. It follows from Eq. (A2) and the con-

[20] See the first article of footnote 2. In this discussion the energy spread of the incoming packet is not considered.

[21] The final result in Eq. (5.24) is obtained from the theory of Fourier transforms and implies physically harmless mathematical restrictions on $u^-(x_1,x_2,x_3)$.

vergence of $\gamma_{mln}(k)$ and $\phi_{mln}(k)$ that

$$\gamma_{ml}(k)=(-i)^l\phi_{ml}(k) \tag{A3}$$

almost everywhere. Furthermore, from the definition of $\gamma_{ml}(k)$ it is easy to see that for any finite K and $p>0$

$$\int_0^K\int_0^\pi\int_0^{2\pi} k^p|\hat{u}(k,\theta,\phi)|^2k^2dkd\Omega = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\int_0^K k^p|\phi_{ml}(k)|^2dk. \tag{A4}$$

By taking $p=4$, it follows that $\|k^2\hat{u}\|<\infty$ if and only if Eq. (3.1) is satisfied. (Note that $V(r)=0$.) Furthermore, from Eq. (A3) and the bi-uniqueness of the transforms in question, it follows that when $\|k^2\hat{u}\|<\infty$, the function whose Fourier-Plancherel transform is $(k^2/2\mu)\hat{u}$ is identical with $F_0^{-1}\{(k^2/2\mu)\phi_{ml}(k)\}$. Hence, H_0, as defined by Eqs. (3.1) and (3.2), is equal to the operator multiplication by $k^2/2\mu$ in the space of Fourier-Plancherel transforms.

To see that the foregoing is true only with the boundary condition $\psi_0(x,k)=0(x)$ for $x\to 0$, consider the radial part of the transform for $l=0$ without this condition.[22] For any function $u(r)$ belonging to $L^2(0,\infty)$

$$u(r)=\lim_{\omega\to\infty}\int_0^\omega \psi^\alpha(r,k)\phi^\alpha(k)dk, \tag{A5}$$

where

$$\phi^\alpha(k)=\lim_{\omega\to\infty}\int_0^\omega \psi^\alpha(x,k)u(x)dx. \tag{A6}$$

[22] E. C. Titchmarsh, *Eigenfunction Expansions Associated with Second Order Differential Equations* (Oxford University Press, London, England, 1946), p. 59.

The function $\psi^\alpha(r,k)$ is given by

$$\psi^\alpha(x,k)=\frac{2^{\frac{1}{2}}}{\pi^{\frac{1}{2}}}\left\{\frac{\cos\alpha}{(\cos^2\alpha+k^2\sin^2\alpha)^{\frac{1}{4}}}\sin kx - \frac{k\sin\alpha}{(\cos^2\alpha+k^2\sin^2\alpha)^{\frac{1}{4}}}\cos kx\right\}. \tag{A7}$$

For $\alpha=0$, $\psi^\alpha(x,k)$ reduces to the function $\psi_0(x,k)$ which figures in Eqs. (2.4)–(2.6). Now consider the function $g(x_1,x_2,x_3)=\exp(-pr)$. It belongs to L^2 and it is readily verified that $\int\int\int|k^2\hat{g}(k_1,k_2,k_3)|^2d\mathbf{k}<\infty$. Therefore, g belongs to the domain of the operator, multiplication by $k^2/2\mu$ in the space of Fourier-Plancherel transforms.

Let $x^\alpha(k)$ be the transform of g defined by Eqs. (2.1)–(2.6) for $V=0$ using the $\psi^\alpha(x,k)$. (Only the term with $l=0$ contributes.) In this case the function $u(r)$ in Eqs. (A5) and (A6) is $2\pi^{\frac{1}{2}}r\exp(-pr)$. Direct calculation now shows that as $k\to\infty$,

$$\begin{aligned} x^\alpha(k)&=(2\sqrt{2}/k^2)(1+0(k^{-2})); && \sin\alpha\neq 0,\\ &=0(k^{-3}), && \sin\alpha=0. \end{aligned} \tag{A8}$$

From Eq. (A8) it is clear that $k^2x^\alpha(k)$ belongs to $L^2(0,\infty)$ if and only if $\sin\alpha=0$. Therefore, $g(x_1,x_2,x_3)$ is in the domain of the operator, multiplication by $k^2/2\mu$ in the space of the transform defined by Eqs. (2.1)–(2.6) if and only if $\sin\alpha=0$. Thus, the domains of the operators, multiplication by $k^2/2\mu$ in the space of Fourier-Plancherel transforms, $\hat{u}$, and multiplication by $k^2/2\mu$ in the space of the transforms F_0u are identical if and only if $\sin\alpha=0$.

A NOTE ON A PAPER OF GINSBURG

By Oscar E. Lanford, III

Introduction. Let P be a partially ordered system and let S and T be non-empty subsets of P. If, for every $p \in S$, there exists a $q \in T$ such that $q \geq p$, T is said to be confinal in S. For every $p \in P$, we denote the set of successors of p in P by $A_P(p)$. If two partially ordered systems P and Q are order isomorphic with confinal subsets of some partially ordered system, they are said to be confinally similar. A partially ordered system P without maximal elements is said to have sufficiently many non-cofinal subsets if, for any two distinct elements p and q of P, either $A_P(p)$ is not cofinal in $A_P(q)$ or $A_P(q)$ is not cofinal in $A_P(p)$. The properties of sets having sufficiently many non-cofinal subsets have been investigated by Ginsburg [1], who poses the following question: "If P has sufficiently many non-cofinal subsets and Q is cofinally similar to P, does Q contain a cofinal subset S which has sufficiently many non-cofinal subsets?" It will be shown by example that the answer to this question is negative.

A subset S of a partially ordered system P is said to be a residual subset if, for every $p \in S$, $A_P(p)$ is contained in S. A subset S of P is said to be maximal residual if S is a residual subset which is not a proper cofinal subset of any residual subset of P. The set of maximal residual subsets of P, ordered by the dual of set inclusion, is denoted by $F(P)$. Ginsburg proves the following theorem (Theorem 5 of [1]): *If P has sufficiently many non-cofinal subsets, P is cofinally similar to $F(P)$.* It is shown that the proof given for this theorem is invalid, and a counterexample to the theorem is given.

1. An example. An example is to be given of two cofinally similar partially ordered systems, one of which has sufficiently many non-cofinal subsets and the other of which contains no cofinal subset having sufficiently many non-cofinal subsets.

Let ω_1 be the first non-denumerable ordinal, and let $W(\omega_1)$ be the set of ordinals less than ω_1. Associate with each $x \in W(\omega_1)$ an infinite subset A_x of the set of integers in such a way that distinct ordinals are assigned distinct sets of integers. Now, for any finite set of integers A, one of the following two cases occurs:

i. For each $x \in W(\omega_1)$ there exists an $s \in W(\omega_1)$, $s \geq x$, such that $A \subset A_s$.

ii. For some $x \in W(\omega_1)$, A is not contained in A_s for any $s \geq x$, while, for all $y < x$, there exists a $z \in W(\omega_1)$, $z \geq y$, such that $A \subset A_z$.

We now consider the set of all x's associated with sets of integers in the second category. This is a denumerable set of denumerable ordinals; hence, there exists a denumerable ordinal ω' which is greater than any of the ordinals

Received March 21, 1962. This work was supported by the National Science Foundation.

in this set. Let W' denote the set of denumerable ordinals greater than ω'. From the definition of W' it follows that if A is a finite set of integers contained in some A_x with $x \in W'$, and if $y \in W'$, then there exists a $z \in W'$, $z \geq y$, such that $A \subset A_z$.

Define a partial ordering on the set of finite sequences of integers by $(a_1, \cdots, a_n) \geq (b_1, \cdots, b_m)$ if $n \geq m$ and $a_1 = b_1, \cdots, a_m = b_m$. For ease of writing we will write $(a_1, \cdots, a_n) \subset A$ if $a_i \in A$, $i = 1, \cdots, n$. We are not restricting ourselves to considering only sequences all of whose elements are distinct.

Consider the set P of elements of the form $(x, F, +)$ and $(x, F, -)$, where $x \in W'$ and F is a finite sequence of integers, $F \subset A_x$. We may define a partial ordering on P by

1. $(x, F, -) \geq (y, G, -)$ if $x \geq y$, $F \geq G$.
2. $(x, F, +) \geq (y, G, -)$ if $x \geq y$, $F \geq G$.
3. $p \geq (x, F, +)$ if either $p = (x, F, +)$ or $p \geq$ some $(x, G, -) \in P$, where $G > F$, $G \not\geq F$.

It is easy to show that this is indeed a partial ordering, i.e., that it is transitive, reflexive, and anti-symmetric. It is also evident that P has no maximal element.

We separate P into two subsets P_+ and P_- consisting of the elements of P with $+$ and $-$ signs respectively. Both are clearly cofinal subsets of P; hence, they are cofinally similar.

We shall first show that P_- contains no cofinal subset having sufficiently many non-cofinal subsets. Let S be a cofinal subset in P_-. Suppose S is denumerable. Then there exists a denumerable ordinal z such that $(x, F, -) \in S$ implies $x < z$. Then no successor of any element of P_- of the form $(z, G, -)$ belongs to S. This is a contradiction, since S is cofinal in P_-. Hence, S must be non-denumerable. Since the family of finite sequences of integers is denumerable, at least two of the elements of S must have the same sequence of integers and differ only in their ordinals. Let one such pair be $(x, F, -)$ and $(y, F, -)$ and suppose for definiteness that $x > y$. We will show that the sets of successors in S of these two elements are cofinal in each other.

Since $x > y$, $A_S((x, F, -))$ is contained in $A_S((y, F, -))$; hence, the latter set is cofinal in the former. Let $(z, G, -) \in A_S((y, F, -))$. Since $G \subset A_z$, it follows from the properties of W' that there exists a $v \in W'$, $v > \sup\{x, z\}$, such that $G \subset A_v$. Hence, $(v, G, -) \in P_-$. Since S is cofinal in P_-, there exists a successor p of $(v, G, -)$ belonging to S. It is easy to show that $(v, G, -) \geq (x, F, -)$ and $(v, G, -) \geq (z, G, -)$, so the same relations hold with $(v, G, -)$ replaced by p. Thus, $A_S((x, F, -))$ is cofinal in $A_S((y, F, -))$, so S does not have sufficiently many non-cofinal subsets.

Next, we consider P_+ and show that it has sufficiently many non-cofinal subsets. Let $(x, F, +)$ and $(y, G, +)$ be two distinct elements of P_+; we shall show that either $A_{P_+}((x, F, +))$ is not cofinal in $A_{P_+}((y, G, +))$ or $A_{P_+}((y, G, +))$

is not cofinal in $A_{P_+}((x, F, +))$. Assume first that neither $F \geq G$ nor $G \geq F$. It follows that F and G have no common successor in the set of finite sequences of integers, and hence that $A_{P_+}((x, F, +)) \cap A_{P_+}((y, G, +))$ is empty. Therefore, we need only consider the case in which F and G are comparable. Suppose that $F = G$. Then $x \neq y$, so $A_x \neq A_y$ and $D = (A_x - A_y) \cup (A_y - A_x)$ is non-empty. Let $b \in D$, and assume for definiteness that $b \in (A_y - A_x)$. If $G = (a_1, \cdots, a_n)$, let $G' = (a_1, \cdots, a_n, b)$. Then $(y, G', +) \in A_{P_+}((y, G, +))$ but $A_{P_+}((y, G', +)) \cap A_{P_+}((x, G, +))$ is empty, so $A_{P_+}((x, G, +))$ is not cofinal in $A_{P_+}((y, G, +))$. Now suppose $F \neq G$, and assume for definiteness that $F > G$. Let $F = (a_1, \cdots, a_n)$ and $G = (a_1, \cdots, a_m)$, $n > m$. Let $b \in A_y$, $b \neq a_{m+1}$. Such a b exists since A_y is infinite. Let $G' = (a_1, \cdots, a_m, b)$. Then $(y, G', +) \geq (y, G, +)$ but $A_{P_+}((y, G', +)) \cap A_{P_+}((x, F, +))$ is empty, so $A_{P_+}((x, F, +))$ is not cofinal in $A_{P_+}((y, G, +))$. This completes the proof that P_+ has sufficiently many non-cofinal subsets, so P_+ and P_- provide the desired example.

2. Cofinal similarity of P and $F(P)$. Ginsburg asserts that, if P has sufficiently many non-cofinal subsets, $F(P)$ is cofinally similar to P (Theorem 5 of [1]). The proof of this result is based on the assertion that the mapping f, which takes an element p of P into that maximal residual subset $f(p)$ which contains $A_P(p)$ as a cofinal subset, is an order isomorphism of P onto a cofinal subset of $F(P)$. This assertion is not correct. It may happen that $q \in f(p)$ (and hence $f(q) \subset f(p)$), even if $q \not\geq p$. This is in fact the case with certain pairs of elements of the set P_+ defined above. Indeed, it is not hard to show that $F(P_+)$ contains a denumerable cofinal subset and that consequently it cannot be cofinally similar to P_+. (See Appendix.)

However, the corollary to Theorem 5 of [1] is correct. Let P be a partially ordered system such that $F(P)$ has sufficiently many non-cofinal subsets; what is to be shown is that $F(P)$ is cofinally similar to $F(F(P))$. This is proved by observing that the proof of Theorem 5 is valid for $F(P)$ if it is shown that $F(P)$ has the property that $T \geq S$ if $A_{F(P)}(S)$ is cofinal in $A_{F(P)}(T)$. This suffices to guarantee that the mapping f constructed in the proof of Theorem 5 in [1] is an order isomorphism. Thus, let $S, T \in F(P)$ be such that $A_{F(P)}(S)$ is cofinal in $A_{F(P)}(T)$, and suppose that $T \not\geq S$. Then, by the definition of $F(P)$, there exists a $p \in P$ such that $p \in T - S$. Since $p \notin S$, and S is a maximal residual subset of P, S is not cofinal in $A_P(p)$. Let $q \geq p$ be such that $A_P(q) \cap S = \phi$, and let $f(q)$ denote the unique maximal residual subset of P in which $A_P(q)$ is cofinal. We shall show that $A_{F(P)}(f(q)) \cap A_{F(P)}(S) = \phi$, which contradicts the fact that $A_{F(P)}(S)$ is cofinal in $A_{F(P)}(T)$ as $f(q) \geq T$. If $A_{F(P)}(f(q)) \cap A_{F(P)}(S) \neq \phi$, there is a maximal residual subset of P contained in both $f(q)$ and S, so it suffices to show that $f(q) \cap S = \phi$. Hence, let $s \in f(q) \cap S$. Since $A_P(q)$ is cofinal in $f(q)$, there exists a $t \in A_P(q)$ such that $t \geq s$. But this implies that $t \in A_P(q) \cap S$. Since $A_P(q) \cap S = \phi$, this proves that $T \geq S$, and hence the corollary to Theorem 5 of [1].

The author would like to express his gratitude to Professor H. Tong for his help and encouragement in carrying out this work.

Appendix. Proof that P_+ is not cofinally similar to $F(P_+)$.

We begin by defining, for each finite sequence of integers F contained in some $A_x (x \in W')$, the set $B(F) = \{(x, G, +) \in P_+ \mid G \geq F\}$. It is easy to see that each $B(F)$ is a maximal residual subset of P_+. Next, we shall show that the set of such B's is cofinal in $F(P_+)$. To do this, let $S \in F(P_+)$, and let $(y, F, +) \in S$. Let $F' \subset A_y$, $F' > F$, $F' \neq F$. We shall show that $B(F')$ is contained in S. Because S is a maximal residual subset of P_+, it suffices to show that S is cofinal in $B(F')$. Let $(z, G, +) \in B(F')$ and let $G' > G$, $G' \neq G$, $G' \subset A_z$. From the properties of W' it follows that for some $v > \sup \{y, z\}$, $(v, G', +) \in P_+$. Now $(v, G', +) \geq (y, F, +)$ and consequently $(v, G', +) \in S$, since S is residual. Since $(v, G', +)$ is also a successor of $(z, G, +)$, we have shown that S is cofinal in $B(F')$. Hence, $B(F')$ is contained in S, and the set of B's is a cofinal subset of $F(P_+)$. Moreover, the set of B's is denumerable. It is easy to show, however, that any partially ordered system cofinally similar to a partially ordered system having a denumerable cofinal subset itself has a denumerable cofinal subset. Since P_+ contains no denumerable cofinal subset, $F(P_+)$ is not cofinally similar to P_+.

Reference

1. Seymour Ginsburg, *A class of everywhere branching sets*, this Journal, vol. 20(1953), pp. 521–526.

Wesleyan University

Integral Representations of Invariant States on B^* Algebras

O. LANFORD† AND D. RUELLE‡
State University of New York at Stony Brook, Stony Brook, New York

(Received 2 September 1966)

Let $\mathfrak{A}$ be a B^* algebra with a group G of automorphisms and K be the set of G-invariant states on $\mathfrak{A}$. We discuss conditions under which a G-invariant state has a unique integral representation in terms of extremal points of K, i.e., extremal invariant states.

1. INTRODUCTION AND NOTATIONS

LET $\mathfrak{A}$ be a B^* algebra, G a group, and τ a (group) homomorphism of G into the * automorphisms of $\mathfrak{A}$. If $\mathfrak{A}$ has an identity, the set of G-invariant states on $\mathfrak{A}$ is compact (for the w^* topology) and one may try to obtain an integral representation of G-invariant states in terms of extremal invariant states. If G is reduced to the identity, such an integral representation is unique if and only if $\mathfrak{A}$ is Abelian. It has, however, been remarked recently that uniqueness prevails under more general circumstances (see Refs. 1 and 2, and for further information, Refs. 3 and 4). The aim of this note is to discuss the general problem of existence and uniqueness of integral representations of invariant states, using Choquet's theory of integral representations on convex compact sets. While some of our results are best possible (in particular, the characterization of G-Abelian B^* algebras, Theorem 2.3), others could certainly be improved (see Sec. 4). Questions related to the existence of a topology on G are relevant for applications to physics, but are not discussed here.

If K is a metrizable compact (phase space) and G a group of homomorphisms of K (time evolution), it is known (see Ref. 5) that a measure on K, invariant under G, can be uniquely decomposed into ergodic measures, i.e., has an integral representation in terms of extremal invariant measures. In this note we obtain an extension of this result of ergodic theory to the noncommutative case (using an algebra of operators in Hilbert space instead of the algebra of continuous functions on a compact) and we weaken the metrizability requirement. The physical problem we have in mind is that of statistical mechanics of an infinite system. An equilibrium state of such a system can be represented by a state ρ on a B^* algebra (e.g., the algebra of canonical commutation relations for a system of bosons), and we may assume invariance of ρ under some natural group G (e.g., the product of the Euclidean group and of the particle number gauge group). One can see that a decomposition of ρ into extremal G-invariant states corresponds to a decomposition into pure thermodynamic phases. Such a decomposition should thus be unique and the problem arises to study the conditions on a non-Abelian algebra and a group of automorphisms such that the invariant states have a unique integral representation in terms of extremal invariant states.

Throughout this note we use the following notations: $\mathfrak{A}$, a B^* algebra; G, a group; $\tau: g \to \tau_g$ a representation of G into the * automorphisms of $\mathfrak{A}$; $\mathfrak{A}'$, the dual of $\mathfrak{A}$ with the w^* topology; $E \subset \mathfrak{A}'$, the set of states on $\mathfrak{A}$ (if $\mathfrak{A}$ has an identity, E is compact); $\mathfrak{L}_G$, the subspace of $\mathfrak{A}$ generated by the elements $A - \tau_g A$ with $A \in \mathfrak{A}$, $g \in G$; $\mathfrak{L}_G^{\perp}$, the orthogonal complement of $\mathfrak{L}_G$ in $\mathfrak{A}'$; $E \cap \mathfrak{L}_G^{\perp}$, the set of G-invariant states.

If $\rho \in E$, we denote by $\mathfrak{H}_\rho$, the Hilbert space of the Gel'fand–Segal construction; π_ρ, the corresponding * homomorphism of $\mathfrak{A}$ into the bounded operators on $\mathfrak{H}_\rho$; $\Omega_\rho \in \mathfrak{H}_\rho$, the normalized vector, cyclic with respect to $\pi_\rho(\mathfrak{A})$ and such that $\rho(A) = (\Omega_\rho, \pi_\rho(A)\Omega_\rho)$ for all $A \in \mathfrak{A}$.

If $\rho \in E \cap \mathfrak{L}_G^{\perp}$ we denote by U_ρ, the unitary representation of G in $\mathfrak{H}_\rho$ such that $U_\rho(g)\Omega_\rho = \Omega_\rho$, $U_\rho(g)\pi_\rho(A)U_\rho(g^{-1}) = \pi_\rho(\tau_g A)$ for all $g \in G$, $A \in \mathfrak{A}$; P_ρ, the projection on the subspace of $\mathfrak{H}_\rho$ formed by the vectors invariant under $U_\rho(G)$.

2. G-ABELIAN ALGEBRAS

In Refs. 1 and 2, the group G was taken to be R^ν and it was assumed that if $A_1, A_2 \in \mathfrak{A}$ the commutator $[A_1, \tau_g A_2]$ vanishes when $g \to \infty$. A suitable generalization of this condition is the basis of our analysis; we formulate it first in a different manner.

† Permanent address: Department of Mathematics, University of California, Berkeley, California.

‡ Permanent address: Institut des Hautes Etudes Scientifiques, 91 Bures-sur-Yvette, France.

[1] D. Ruelle, Commun. Math. Phys. 3, 133 (1966).

[2] D. Kastler and D. Robinson, Commun. Math. Phys. 3, 151 (1966).

[3] S. Doplicher, D. Kastler, and D. Robinson, Commun. Math. Phys. 3, 1, (1966).

[4] D. Robinson and D. Ruelle, "Extremal Invariant States," Institut des Hautes Etudes Scientifiques (1966).

[5] R. Phelps, *Lectures on Choquet's Theorem* (D. Van Nostrand Company, Inc. Princeton, New Jersey, 1966).

Definition 2.1: $\mathfrak{A}$ is said to be G-Abelian if for all $\rho \in E \cap \mathfrak{L}_G^{\perp}$ and $A_1, A_2 \in \mathfrak{A}$,

$$[P_\rho \pi_\rho(A_1) P_\rho, P_\rho \pi_\rho(A_2) P_\rho] = 0.$$

In other words the von Neumann algebra generated by $P_\rho \pi_\rho(\mathfrak{A}) P_\rho$ is Abelian.

Theorem 2.2 (Alaoglu-Birkhoff): Let $\{U_\alpha\}_{\alpha \in I}$ be a semigroup of contractions on a Hilbert space $\mathfrak{H}$, i.e., a collection of operators such that

(1) $\|U_\alpha\| \leq 1$ for all $\alpha \in I$

(2) For any $\alpha, \beta \in I$, $U_\alpha U_\beta = U_\gamma$ for some $\gamma \in I$.

Let P be the orthogonal projection onto the set of all vectors in $\mathfrak{H}$ left invariant by all the U_α's. Then P is in the strong closure at the convex hull of $\{U_\alpha\}_{\alpha \in I}$.

This theorem is proved in Riesz–Nagy.[6] The theorem stated by Riesz and Nagy is slightly different from the one given above; what they do is to construct a net of convex linear combinations of the U_α's and show that it converges strongly. Although the fact that P is the strong limit of this net is not included in the statement of the theorem, it appears in the course of the proof.

Theorem 2.3: In order that $\mathfrak{A}$ be G-Abelian it is necessary and sufficient that, for all Hermitian A_1, $A_2 \in \mathfrak{A}$ and all $\rho \in E \cap \mathfrak{L}_G^{\perp}$,

$$\inf_{A_1'} |\rho([A_1', A_2])| = 0,$$

where A_1' runs over the convex hull of $\{\tau_g A_1 : g \in G\}$.

In order that $\mathfrak{A}$ be G-Abelian, it is evidently necessary and sufficient that, for any $\rho \in E \cap \mathfrak{L}_G^{\perp}$, $\Psi \in P_\rho \mathfrak{H}_\rho$ with $\|\Psi\| = 1$, and A_1, A_2 Hermitian elements of the unit ball of $\mathfrak{A}$, we have

$$(\Psi, \pi_\rho(A_1) P_\rho \pi_\rho(A_2) \Psi) = (\Psi, \pi_\rho(A_2) P_\rho \pi_\rho(A_1) \Psi) \quad (*).$$

We prove first the sufficiency of the criterion stated in the proposition. Let $\epsilon > 0$; then by the preceding theorem, we can find positive numbers λ_i with $\sum_i \lambda_i = 1$ and elements g_i of G such that

$$\|(\sum \lambda_i U_\rho(g_i) - P_\rho) \pi_\rho(A_1) \Psi\| \leq \tfrac{1}{2}\epsilon.$$

If we define

$$A_1' = \sum \lambda_i \tau_{g_i} A_1,$$

then both sides of (*) are unchanged if we replace A_1 by A_1', and we have

$$\begin{aligned} &\|P_\rho \pi_\rho(A_1') \Psi - U_\rho(g) \pi_\rho(A_1') \Psi\| \\ &\quad = \|P_\rho \pi_\rho(A_1) \Psi - U_\rho(g) \pi_\rho(A_1') \Psi\| \\ &\quad = \|U_\rho(g)[P_\rho \pi_\rho(A_1) \Psi - \pi_\rho(A_1') \Psi]\| \leq \tfrac{1}{2}\epsilon \end{aligned}$$

for all $g \in G$.

Using this inequality, and the fact that A_1' is Hermitian, we get for any positive numbers λ_i' with $\sum_i \lambda_i' = 1$ and any $g_i' \in G$,

$$\begin{aligned} &|(\Psi, \pi_\rho(A_1) P_\rho \pi_\rho(A_2) \Psi) - (\Psi, \pi_\rho(A_2) P_\rho \pi_\rho(A_1) \Psi)| \\ &= |(\Psi, \pi_\rho(A_1') P_\rho \pi_\rho(A_2) \Psi) - (\Psi, \pi_\rho(A_2) P_\rho \pi_\rho(A_1') \Psi)| \\ &\leq 2 \cdot \sum_i \lambda_i' \|\pi_\rho(A_2) \Psi\| \cdot \|P_\rho \pi_\rho(A_1') \Psi - U_\rho(g_i') \pi_\rho(A_1') \Psi\| \\ &\qquad + |(\Psi, \pi_\rho([\textstyle\sum \lambda_i' \tau_{g_i'} A_1', A_2]) \Psi)| \\ &\leq \epsilon + |(\Psi, \pi_\rho([\textstyle\sum \lambda_i' \tau_{g_i'} A_1', A_2]) \Psi)|. \end{aligned}$$

But by hypothesis, $|(\Psi, \pi_\rho([\sum \lambda_i' \tau_{g_i'} A_1', A_2]) \Psi)|$ can be made arbitrarily small by an appropriate choice of λ_i' and g_i', so

$$|(\Psi, \pi_\rho(A_1) P_\rho \pi_\rho(A_2) \Psi) - (\Psi, \pi_\rho(A_2) P_\rho \pi_\rho(A_1) \Psi)| \leq \epsilon.$$

Thus, (*) holds, so $\mathfrak{A}$ is G-Abelian.

Now we suppose that $\mathfrak{A}$ is G-Abelian, so (*) holds, and we let λ_i, g_i be as above. Then

$$\begin{aligned} &\left|\left(\Psi, \pi_\rho\left(\left[\sum_i \lambda_i \tau_{g_i} A_1, A_2\right]\right)\Psi\right)\right| \\ &= \left|\left(\sum_i \lambda_i U_\rho(g_i) \pi_\rho(A_1) \Psi, \pi_\rho(A_2) \Psi\right)\right. \\ &\quad \left. - \left(\pi_\rho(A_2), \sum_i \lambda_i U_\rho(g_i) \pi_\rho(A_1) \Psi\right)\right| \\ &\leq 2 \cdot \|\pi_\rho(A_2) \Psi\| \cdot \left\|\left(\sum_i \lambda_i U_\rho(g_i) - P_\rho\right) \pi_\rho(A_1) \Psi\right\| \\ &\quad + |(\Psi, \pi_\rho(A_1) P_\rho \pi_\rho(A_2) \Psi) - (\Psi, \pi_\rho(A_2) P_\rho \pi_\rho(A_1) \Psi)| \\ &\leq \epsilon, \end{aligned}$$

so

$$\inf_{A_1' \in \text{convex hull of } \{\tau_g A_1\}} |(\Psi, \pi_\rho([A_1', A_2]) \Psi)| = 0,$$

so the criterion of the proposition holds.

Corollary 2.4: Let H be a subgroup of G. Then, if $\mathfrak{A}$ is H-Abelian, it is also G-Abelian.

We need only apply the criterion of the preceding proposition, observing that $\mathfrak{L}_G^{\perp}$ is contained in $\mathfrak{L}_H^{\perp}$ and that the convex hull of $\{\tau_g A_1 : g \in G\}$ contains the convex hull of $\{\tau_h A_1 : h \in H\}$.

Corollary 2.5: $\mathfrak{A}$ is G-Abelian whenever one of the following conditions is satisfied.

[6] F. Riesz and B. Sz.-Nagy, *Functional Analysis*, translated by L. Boron (Frederick Ungar Publishing Company, New York, 1955), Sec. 146.

(i) For all $\rho \in E \cap \mathfrak{L}_G^\perp$ and self-adjoint $A_1, A_2 \in \mathfrak{A}$,

$$\inf_{g \in G} |\rho([A_1, \tau_g A_2])| = 0.$$

(ii) $\mathfrak{A}$ is Abelian.

(iii) $E \cap \mathfrak{L}_G^\perp$ is empty.

The usefulness of Definition 2.1 appears in the next two sections; we indicate here, however, the following result.

Proposition 2.6: If $\rho \in E \cap \mathfrak{L}_G^\perp$ and the von Neumann algebra $[P_\rho \pi_\rho(\mathfrak{A}) P_\rho]''$ generated by $P_\rho \pi_\rho(\mathfrak{A}) P_\rho$ is Abelian, then

$$P_\rho [P_\rho \pi_\rho(\mathfrak{A}) P_\rho]' = P_\rho [P_\rho \pi_\rho(\mathfrak{A}) P_\rho]''.$$

The vector Ω_ρ is cyclic for the restriction to $P_\rho \mathfrak{H}_\rho$ of $P_\rho [P_\rho \pi_\rho(\mathfrak{A}) P_\rho]''$; hence, if this von Neumann algebra is commutative, it is equal to its commutant (see Ref. 7, p. 89, Corollaire 2), namely to

$$P_\rho [P_\rho \pi_\rho(\mathfrak{A}) P_\rho]'$$

restricted to $P_\rho \mathfrak{H}_\rho$.

3. INTEGRAL REPRESENTATION OF G-INVARIANT STATES

In this and the next section, we use the theory of integral representations on convex compact sets (see Ref. 8). Let K be a convex compact set in a locally convex topological vector space. The unit mass at $\kappa \in K$ is denoted by δ_κ. We remind the reader that an order relation is defined on the positive measures of norm 1 on K by $\mu \prec u' \Leftrightarrow \mu(f) \leq \mu'(f)$ for all convex continuous f on K. A measure is called maximal if it is maximal for the order $\prec$, and K is said to be a simplex if every $\kappa \in K$ is the resultant of a *unique* maximal measure on K. In what follows we take $K = E \cap \mathfrak{L}_G^\perp$, where $\mathfrak{A}$ is assumed to have an identity. If $A \in \mathfrak{A}$, we denote by $\hat{A}$ the function on $E \cap \mathfrak{L}_G^\perp$ defined by $\hat{A}(\rho) = \rho(A)$.

Theorem 3.1: Let $\mathfrak{A}$ have an identity, $\rho \in E \cap \mathfrak{L}_G^\perp$, and let the von Neumann algebra generated by $P_\rho \pi_\rho(\mathfrak{A}) P_\rho$ be Abelian. Then, there exists a unique maximal measure μ_ρ on $E \cap \mathfrak{L}_G^\perp$ such that $\mu_\rho \succ \delta_\rho$ (i.e., μ_ρ has resultant ρ). The measure μ_ρ is determined by

$$\mu_\rho(\hat{A}_1 \cdots \hat{A}_l) = (\Omega_\rho, \pi_\rho(A_1) P_\rho \pi_\rho(A_2) P_\rho \cdots P_\rho \pi_\rho(A_l) \Omega_l). \quad (1)$$

Take $A_1, \cdots, A_l$ self-adjoint. Since the operators $P_\rho \pi_\rho(A_1) P_\rho, \cdots, P_\rho \pi_\rho(A_l) P_\rho$ commute, there exists a projection-valued measure F on R^l such that

$$P_\rho \pi_\rho(A_i) P_\rho = \int t_i \, dF(t_1, \cdots, t_l).$$

If $\mathfrak{F}$ is a complex polynomial of l variables, we have

$$\begin{aligned}
&|(\Omega_\rho, \mathfrak{F}(P_\rho \pi_\rho(A_1) P_\rho, \cdots, P_\rho \pi_\rho(A_l) P_\rho) \Omega_\rho)| \\
&= \left| \left(\Omega_\rho, \int \mathfrak{F}(t_1, \cdots, t_l) \, dF(t_1, \cdots, t_l) \Omega_\rho \right) \right| \\
&\leq \sup_{\|\Phi\|=1, P_\rho \Phi = \Phi} |\mathfrak{F}((\Phi, \pi_\rho(A_1)\Phi), \cdots, (\Phi, \pi_\rho(A_l)\Phi))| \\
&\leq \sup_{\sigma \in E \cap \mathfrak{L}_G^\perp} |\mathfrak{F}(\sigma(A_1), \cdots, \sigma(A_l))| \\
&= \sup_{\sigma \in E \cap \mathfrak{L}_G^\perp} |\mathfrak{F}(\hat{A}_1(\sigma), \cdots, \hat{A}_l(\sigma))|.
\end{aligned}$$

This shows that Eq. (1) defines a linear functional on the polynomials in the $\hat{A}$, which is continuous for the topology of uniform convergence on $E \cap \mathfrak{L}_G^\perp$. By the Stone–Weierstrass theorem, this functional extends uniquely to a measure μ_ρ on $E \cap \mathfrak{L}_G^\perp$, which is ≥ 0 and of norm 1.

Let $\rho_1, \cdots, \rho_m \in E \cap \mathfrak{L}_G^\perp$, $\lambda_1, \cdots, \lambda_m > 0$, $\sum \lambda_i = 1$ and $\rho = \sum \lambda_i \rho_i$. There exist (see Ref. 9, 2.5.1.) uniquely defined self-adjoint operators $T_i \in [\pi_\rho(\mathfrak{A})]'$ such that $0 \leq T_i \leq 1$ and for all $A \in \mathfrak{A}$.

$$\lambda_i \rho_i(A) = (T_i \Omega_\rho, \pi_\rho(A) T_i \Omega_\rho).$$

The T_i satisfy $\sum T_i^2 = 1$. If $g \in G$, we have

$$U(g) T_i U(g^{-1}) \in [\pi_\rho(\mathfrak{A})]',$$

the uniqueness of T_i and the fact that $\lambda_i \rho_i \in \mathfrak{L}_G^\perp$ then shows that $U(g) T_i U(g^{-1}) = T_i$, hence,

$$T_i \in [U(G)]', \quad [T_i, P_\rho] = 0.$$

By the uniqueness of the Gel'fand–Segal construction, we may identify $\mathfrak{H}_{\rho_i}$ with the closure of $\pi_\rho(\mathfrak{A}) T_i \Omega_\rho$, π_{ρ_i} with the restriction of π_ρ to $\mathfrak{H}_{\rho_i}$, and Ω_{ρ_i} with $\lambda_i^{-\frac{1}{2}} T_i \Omega_\rho$. Then U_{ρ_i} is identified with the restriction of U_ρ to $\mathfrak{H}_{\rho_i}$ and P_{ρ_i} with the restriction of P_ρ to $\mathfrak{H}_{\rho_i}$. In particular, $[P_{\rho_i} \pi_{\rho_i}(\mathfrak{A}) P_{\rho_i}]''$ is Abelian and μ_{ρ_i} is thus defined. We have

$$\begin{aligned}
\mu_{\rho_i}(\hat{A}_1 \cdots \hat{A}_l) &= (\Omega_\rho, \pi_\rho(A_1) P_{\rho_i} \cdots P_{\rho_i} \pi_{\rho_i}(A_l) \Omega_{\rho_i}) \\
&= \lambda_i^{-1} (T_i \Omega_\rho, \pi_\rho(A_1) P_\rho \cdots P_\rho \pi_\rho(A_l) T_i \Omega_\rho) \\
&= \lambda_i^{-1} (\Omega_\rho, \pi_\rho(A_1) P_\rho \cdots P_\rho \pi_\rho(A_l) T_i^2 \Omega_\rho)
\end{aligned}$$

so that

$$\sum \lambda_i \mu_{\rho_i}(\hat{A}_1 \cdots \hat{A}_l) = \mu_\rho(\hat{A}_1 \cdots \hat{A}_l).$$

[7] J. Dixmier, *Les algèbres d'opérateurs dans l'Espace Hilbertien*, (Gauthier-Villars, Paris, 1957).

[8] G. Choquet and P. A. Meyer, Ann. Inst. Fourier **13**, 139 (1963).

[9] J. Dixmier, *Les C*-Algèbres et leurs Représentations* (Gauthier-Villars, Paris, 1964).

Now let μ be a measure on $E \cap \mathfrak{L}_G^\perp$ such that $\mu \succ \delta_\rho$. If $\phi \in \mathfrak{C}(E \cap \mathfrak{L}_G^\perp)$ and $\epsilon > 0$, one can find a measure μ' with finite support: $\mu' = \sum \lambda_i \delta_{\rho_i}$, $\lambda_i > 0$, $\rho_i \in E \cap \mathfrak{L}_G^\perp$, such that $|\mu(\phi) - \mu'(\phi)| < \epsilon$ and $\sum \lambda_i \rho_i = \rho$ (see Ref. 10, p. 217, Prop. 3). If ϕ is convex we thus have

$$\mu(\phi) - \epsilon \leq \mu'(\phi) = \sum \lambda_i \delta_{\rho_i}(\phi) \leq \sum \lambda_i \mu_{\rho_i}(\phi) = \mu_\rho(\phi),$$

hence $\mu_\rho \succ \mu$. Since μ is an arbitrary measure on $E \cap \mathfrak{L}_G^\perp$ such that $\mu \succ \delta_\rho$, we see that μ_ρ is the unique maximal measure on $E \cap \mathfrak{L}_G^\perp$ such that $\mu_\rho \succ \delta_\rho$ which concludes the proof of the theorem.

Corollary 3.2: If $\mathfrak{A}$ has an identity and is G-Abelian, then $E \cap \mathfrak{L}_G^\perp$ is a simplex.

Remark 3.3: If $\mathfrak{A}$ is Abelian, the problem considered in this section reduces to that of decomposing an invariant measure on a compact set into ergodic measures (see Ref. 5, Sec. 10).

4. EXTREMAL G-INVARIANT STATES

Let $\mathfrak{E}(E \cap \mathfrak{L}_G^\perp)$ be the set of extremal points of $E \cap \mathfrak{L}_G^\perp$, i.e., the extremal invariant states. The following statement characterizes the elements of $\mathfrak{E}(E \cap \mathfrak{L}_G^\perp)$.

Proposition 4.1: Let $\rho \in E \cap \mathfrak{L}_G^\perp$. If $\mathfrak{A}$ is G-Abelian, the following conditions are equivalent:

(i) $\rho \in \mathfrak{E}(E \cap \mathfrak{L}_G^\perp)$.
(ii) The set $\pi_\rho(\mathfrak{A}) \cup U_\rho(G)$ is irreducible in $\mathfrak{H}_\rho$.
(iii) P_ρ is one dimensional.

The simple proof is left to the reader. We remark only that the implications (i) $\Leftrightarrow$ (ii) $\Leftarrow$ (iii) do not make use of the assumption that $\mathfrak{A}$ is G-Abelian, and that (ii) $\Rightarrow$ (iii) follows from Proposition 2.6.

The measure μ_ρ of Theorem 3.1 is in the "good cases" carried by $\mathfrak{E}(E \cap \mathfrak{L}_G^\perp)$. This is so, for instance, if $\mathfrak{A}$ is (norm-)separable, because $E \cap \mathfrak{L}_G^\perp$ is then metrizable (see Ref. 8, Corr. 14). We indicate now without proofs some more results in this direction.

Proposition 4.2: Let $\mathfrak{A}$ have an identity and $\mathfrak{B}$ be a self-adjoint subalgebra of $\mathfrak{A}$; define

$$\mathcal{F} = \{\sigma \in E: \text{The restriction of } \rho \text{ to } \mathfrak{B} \text{ has norm } 1\}.$$

Then,

(i) $\mathcal{F}$ is a G_δ (a countable intersection of open subsets of E).
(ii) If μ is a measure on E such that $\mu \geq 0$, $\mu(E) = 1$, and μ has resultant ρ, then
$$\rho \in \mathcal{F} \leftrightarrow \mu \text{ is carried by } \mathcal{F},$$

cf. Ref. 1, Theorem, Part 4.

Proposition 4.3: Let $(\mathfrak{A}_\alpha)$ be a countable family of sub-B^* algebras of $\mathfrak{A}$ such that $\bigcup_\alpha \mathfrak{A}_\alpha$ is dense in $\mathfrak{A}$. Let $\mathfrak{C}_\alpha$ be a separable closed two-sided ideal of $\mathfrak{A}_\alpha$ for each α, and define

$$\mathcal{F}_\alpha = \{\sigma \in E: \text{the restriction of } \sigma \text{ to } \mathfrak{C}_\alpha \text{ has norm } 1\}, \quad \mathcal{F} = \bigcap_\alpha \mathcal{F}_\alpha.$$

Then,

(i) If $\rho \in \mathcal{F}$, then $\mathfrak{H}_\rho$ is separable.
(ii) There exists a sequence (A_i) of self-adjoint elements of $\mathfrak{A}$ such that if $\rho \in \mathcal{F}$ and $\sigma \in E$, then $\rho(A_i) \neq \sigma(A_i)$ for some i.
(iii) If $\mathfrak{A}$ has an identity and is G-Abelian and if μ is a measure on $E \cap \mathfrak{L}_G^\perp$ such that $\mu \geq 0$, $\mu(E \cap \mathfrak{L}_G^\perp) = 1$ and μ has resultant $\rho \in \mathcal{F}$, then
(μ maximal on $E \cap \mathfrak{L}_G^\perp$)
$$\Leftrightarrow (\mu \text{ carried by } \mathfrak{E}(E \cap \mathfrak{L}_G^\perp)).$$

(i) and (ii) are easy, the proof of (iii) uses (ii), Corollary 3.2 and an argument in Ref. 1, Theorem, Part 5.

The usefulness of (iii) appears in statistical mechanics, where $\mathfrak{A}$ may not be norm separable but the states of interest satisfy a condition of the type $\rho \in \mathcal{F}$. One has then a unique decomposition $\rho \to \mu_\rho$ of ρ into extremal invariant states and those states are again in $\mathcal{F}$. For an explicit treatment see Ref. 11, in particular, the Appendix.

[10] N. Bourbaki, *Intégration* (Hermann et Cie., Paris, 1965), 2nd ed., Chaps. 1–4.

[11] D. Ruelle, "The States of Classical Statistical Mechanics," J. Math. Phys. (to be published).

Commun. math. Phys. 9, 176—191 (1968)

The Classical Mechanics of One-Dimensional Systems of Infinitely Many Particles

I. An Existence Theorem

O. E. LANFORD III

I.H.E.S., 91-Bures-sur-Yvette

Received April 19, 1968

Abstract. We prove a global existence and uniqueness theorem for solutions of the classical equations of motion for a one-dimensional system of infinitely many particles interacting by finite-range two-body forces which satisfy a Lipschitz condition.

§ 1. Introduction

In this paper, we prove an existence and uniqueness theorem for solutions of the equations of motion of a system of infinitely many classical point particles, constrained to move in one dimension, interacting by two-body forces of finite range. Thus, let (q_i, p_i) be a sequence of pairs of real numbers representing the positions and velocities of an infinite set of particles. We assume that each bounded interval in $\boldsymbol{R}$ contains only finitely many particles, and we want to solve the differential equations:

$$\frac{dq_i(t)}{dt} = p_i(t) \qquad \frac{dp_i(t)}{dt} = \sum_{j \neq i} F(q_i(t) - q_j(t)) \tag{1.1}$$

with the initial conditions:

$$q_i(0) = q_i \; ; \qquad p_i(0) = p_i \; .$$

(For simplicity, we are taking the particles to be identical and to have mass one). The interparticle force F will be assumed to be bounded and to have compact support. As long as each bounded interval in $\boldsymbol{R}$ contains only finitely many $q_j(t)$'s, the sum on the right of the second equation has only finitely many non-zero terms for each i and the equations therefore make sense. It is clear, however, that for some initial configurations we must expect the Eq. (1.1) to lead in finite time to a catastrophic situation with infinitely many particles in some bounded interval. To take a trivial example, if there are no interparticle forces and if $p_i = -q_i$ for each i, then all the particles are at the origin at time one. The crux of the problem of proving an existence theorem is to find a set of initial configurations for which such catastrophies can be shown not to happen.

A first result in this direction was obtained by J. GINIBRE (unpublished) who proved that the Eq. (1.1) have a solution valid for all values of t if the initial configuration admits an upper bound on the absolute values of the velocities of the various particles and on the numbers of particles in the various intervals of unit length; furthermore, he proved a local existence theorem for systems of particles moving in $\boldsymbol{R}^\nu$ (instead of $\boldsymbol{R}$), with the analogous restrictions on the initial momenta and densities. Our theorem, which holds only in one dimension, gives existence for initial configurations in which, roughly speaking, the velocities increase at most logarithmically with distance from the origin and the number of particles in an interval of unit length increases at most logarithmically with the distance from that interval to the origin. (The precise condition we impose on the initial densities is actually a bit more restrictive; see § 2.)

The main interest of the existence and uniqueness theorem which we will prove lies in its application to the time-evolution problem in the classical statistical mechanics of infinite systems. We will discuss this application in detail in a subsequent publication. Nevertheless, we give here a brief sketch of how the application is made, as motivation for our theorem and to explain why it is important to be able to solve the equations of motion for a set of initial configurations more general than those with bounded velocities and densities.

We have first to explain what is meant by a state of classical statistical mechanics. By a *locally finite configuration of labelled particles* we will mean either:

a) An n-tuple $x = (q_1, p_1; q_2, p_2; \ldots; q_n, p_n)$ of pairs of real numbers or

b) A sequence $x = (q_i, p_i)$ of pairs of real numbers such that each bounded set in $\boldsymbol{R}$ contains only finitely many q_i's, i.e., such that $\lim_{i \to \infty} |q_i| = \infty$.

We will denote the set of all such configurations by $\mathscr{X}$. A *locally finite configuration of unlabelled particles* will mean an equivalence class of locally finite configurations of labelled particles, where two configurations are equivalent if they differ only by a permutation of the indexing set; we will denote by $[x]$ the equivalence class of x and by $[\mathscr{X}]$ the set of equivalence classes. Space translations act on an obvious way on $\mathscr{X}$ and on $[\mathscr{X}]$. The space $[\mathscr{X}]$ may be equipped with a topology in a natural way [1]; then a *state of classical statistical mechanics* is a Borel probability measure on $[\mathscr{X}]$ which is invariant under the action of space translations.

Now suppose that the equations of motion (1.1) have a unique solution for every initial configuration x in some subset $\hat{\mathscr{X}}$ of $\mathscr{X}$, and that the solution curve

$$\{x(t) = (q_i(t), p_i(t)) : -\infty < t < \infty\}$$

in contained in $\hat{\mathscr{X}}$. By the invariance of Eq. (1.1) under permutations of the indexing set, it is clear that $\hat{\mathscr{X}}$ can be taken to be a union of equivalence classes; let $[\hat{\mathscr{X}}]$ denote the corresponding set of unlabelled configurations. For any t, we can define a mapping T^t of $[\hat{\mathscr{X}}]$ into itself by setting $T^t[x]$ equal to the equivalence class of the value at time t of a solution of the equations of motion whose value at time zero belongs to the equivalence class $[x]$. By the uniqueness of the solution, this definition does not depend on the choices made, and $\{T^t\}$ is a one-parameter group of mappings of $[\hat{\mathscr{X}}]$ onto itself.

If ϱ is a measure on $[\mathscr{X}]$ which is concentrated on $[\hat{\mathscr{X}}]$, i.e. which has

$$\varrho([\mathscr{X}]\backslash[\hat{\mathscr{X}}]) = 0\,,$$

and if each T^t is a measurable mapping of $[\hat{\mathscr{X}}]$ onto itself, then we can define for each t a measure ϱ^t by

$$\varrho^t([\mathscr{X}]\backslash[\hat{\mathscr{X}}]) = 0\,,$$

$$\varrho^t(A) = \varrho(T^{-t}A) \quad \text{if} \quad A \subset [\hat{\mathscr{X}}]\,.$$

Thus, we get a satisfactory time evolution for those states which are concentrated on $[\hat{\mathscr{X}}]$, and the usefulness of an existence theorem depends on whether or not interesting states are concentrated on the set of allowed initial configurations. In the subsequent publication referred to above, we will give a criterion for states to be concentrated on our set of initial configurations which implies in particular that

a) states obtained by taking the infinite volume limit of thermodynamic ensembles at low activity [2], and

b) states obtained by taking an infinite volume limit of thermodynamic ensembles with non-negative potentials, at any activity have this property. On the other hand, since any state obtained by taking an infinite volume limit of canonical or grand-canonical ensembles has a Maxwellian velocity distribution[1], it is easy to see that no such state can be concentrated on the set of configurations with bounded velocities (unless it is the trivial state which is concentrated on the configuration with no particles).

In fact, the logarithmic rate of increase of density fluctuations is almost the slowest increase which can be allowed if we are to have a sufficient set of initial configurations for applications to statistical

[1] A state is said to have a Maxwellian velocity distribution if the velocity of any given particle is independent of the position of that particle and of the positions and velocities of the other particles, and if the velocity of a single particle is distributed with probability density $\sqrt{\frac{\beta}{2\pi}}e^{-\beta p^2/2}$, where β is some positive real number.

mechanics; a typical configuration of non-interacting particles has density fluctuations which increase like $\log(d)/\log(\log(d))$, where d is the distance to the origin. To make this statement precise, we define a function on $[\mathscr{X}]$ as follows: For any configuration $[x]$, take the number of particles in the interval $[n, n+1)$, divide by $\log(n)/\log(\log(n))$, and take the lim. sup. as n goes to infinity. This gives either a non-negative real number or $+\infty$. An elementary calculation shows that, for the state obtained by taking the infinite volume limit of the grand canonical ensemble for non-interacting particles, with any (non-zero) temperature and any chemical potential, this function takes on the value one on the complement of a set of measure zero.

Having made these remarks by way of motivation, we will devote our attention for the rest of this article to the problem of solving the differential Eq. (1.1), without considering further the application to statistical mechanics. In § 2, we give a precise definition of the set of initial configurations for which we can solve these equations, state the main result, and sketch the ideas underlying the proof. In § 3, we reduce the equations of motion to a non-linear evolution equation:

$$\frac{d\zeta(t)}{dt} = A(\zeta(t)), \tag{1.2}$$

on a Banach space isomorphic to l^∞ (where the derivative is to be taken in the sense of the product tolopogy), and we estimate the norm of $A(\zeta)$. In § 4, we show that, although the non-linear operator A does not satisfy a norm Lipschitz condition, it does satisfy a Lipschitz condition with respect to a family of semi-norms defining the product topology. We then use this Lipschitz condition, together with norm estimates on the operator A, to prove the existence and uniqueness of solutions of (1.2), and to show that these solutions can be obtained by solving the integral equation

$$\zeta(t) = \zeta(0) + \int_0^t d\tau A(\zeta(\tau))$$

by iteration.

§ 2. The Set of Allowed Initial Configurations

Before defining the set of initial configurations for which we can solve the equations of motion, we need some notation. First, to cut off the logarithm function for small values of its argument, we make the definition:

$$\log_+(q) = \log(|q| \vee e) \tag{2.1}$$

where the symbol $\vee$ denotes "supremum". We shall make frequent use of the elementary inequalities:

$$\begin{aligned} \log_+(a+b) &\leqq \log_+(a) + \log_+(b)\ ; \\ \log_+(a \cdot b) &\leqq \log_+(a) + \log_+(b)\ ; \\ \log_+(a) &\leqq |a| \quad \text{if} \quad |a| \geqq 1\ . \end{aligned} \tag{2.2}$$

Second, for any $x = (q_i, p_i)$ in $\mathscr{X}$, and any bounded set Λ in $\boldsymbol{R}$, we define $N_\Lambda(x)$ to be the number of i's such that $q_i \in \Lambda$, i.e., the number of particles in the region Λ for the configuration x.

The set of initial configurations $x = (q_i, p_i)$ which we want to consider will be those satisfying the following two conditions:

1) There is a constant K_1, such that, for each i,

$$|p_i| \leqq K_1 \log_+(q_i) . \tag{2.3}$$

2) There is a constant K_2 such that, if (α, β) is any bounded open interval whose length $\beta - \alpha$ is larger than $\log_+((\alpha + \beta)/2)$, then

$$N_{(\alpha,\beta)}(x) \leqq K_2(\beta - \alpha) . \tag{2.4}$$

Condition 2) may be reformulated by saying that there is an upper bound for the mean density of particles in any interval of length greater than one whose length is also greater than the logarithm of the distance from its center to the origin. It implies in particular that the number of particles in any interval of unit length is bounded by a constant times the logarithm of the distance from its center to the origin.

We will denote by $\hat{\mathscr{X}}$ the set of labelled configurations satisfying 1) and 2), and by $[\hat{\mathscr{X}}]$ the corresponding set of unlabelled configurations. The set $\hat{\mathscr{X}}$, although not defined in a manifestly translation-invariant way, is easily seen to be mapped into itself by translations.

For any $x \in \hat{\mathscr{X}}$, we will let $|x|$ denote the smallest number which will work for both K_1 and K_2 in (2.3) and (2.4) respectively, i.e.,

$$|x| = \left[\sup_i \left\{\frac{|p_i|}{\log_+(q_i)}\right\}\right] \vee \left[\sup\left\{\frac{N_{(\alpha,\beta)}(x)}{\beta - \alpha} : \beta - \alpha > \log_+\left(\frac{\beta + \alpha}{2}\right)\right\}\right] . \tag{2.5}$$

We can now formulate our main result:

Theorem 2.1. *Let F be a real-valued function with compact support, satisfying a Lipschitz condition:*

$$|F(q_1) - F(q_2)| \leqq K \cdot |q_1 - q_2| \tag{2.6}$$

and let $x = (q_i, p_i)$ belong to $\hat{\mathscr{X}}$. Then there is one and only one function $x(t) = (q_i(t), p_i(t))$, defined for $-\infty < t < \infty$, with values in $\mathscr{X}$, satisfying

1. $\dfrac{dq_i(t)}{dt} = p_i(t) \quad \dfrac{dp_i(t)}{dt} = \sum_{j \neq i} F(q_i(t) - q_j(t))$ (1.1)

2. $q_i(0) = q_i ; \quad p_i(0) = p_i$

3. *$|x(t)|$ is a locally bounded function of t, i.e., is bounded on any bounded interval.*

We will say that a solution of the Eq. (1.1) is *regular* if it satisfies condition 3. Note that the uniqueness statement of the theorem is not as strong as one might hope, since it does not rule out the possibility of non-regular solutions.

The formulation of the theorem supposes tacitly that we use the same labelling set for all t and in particular that the total number of particles does not change with time. The result is well-known if the initial configuration has only finitely many particles. We will therefore give the proof only for initial configurations with infinitely many particles; the argument can readily be adapted, at the expense of complicating the notation, to apply simultaneously to the two cases.

The proof of the theorem is obscured by technical problems and by straight forward but tedious estimates. It seems worthwhile, therefore, to give a brief and unencumbered description of the underlying idea. The central difficulty is that of showing that the differential equations cannot drive infinitely many particles into a finite region of space in finite time. One can convince oneself of this by showing that the differential Eq. (1.1) imply integral inequalities for the quantity $|x(t)|$ which prevent its going to infinity in finite time. These inequalities are gotten as follows: If we know $|x(\tau)|$ for $0 \leqq \tau \leqq t$, then we have in particular estimates on the density for that interval of time. Majorizing the force on the i^{th} particle by the maximum of $|F|$ times the number of particles within a distance R (the range of the force) of q_i, we can convert our density estimates to estimates on the forces and thus, by integrating the second differential equation, to estimates on the velocities at time t. Similarly, from knowing $|x(\tau)|$ for $0 \leqq \tau \leqq t$, we get bounds on the velocities and therefore on the distances travelled. Using these bounds we can find, for a given interval (α, β), a larger interval from which a particle has to start if it is to be in (α, β) at time t. Knowing $|x(0)|$ enables us to majorize the number of particles in this larger interval at time zero and therefore the number of particles in (α, β) at time t.

Combining all these estimates gives a bound for $|x(t)|$ in terms of $|x(\tau)|$ for $0 \leqq \tau \leqq t$, and this bound may be seen to imply that $|x(t)|$ cannot go to infinity in finite time. Unfortunately, the estimates are very tedious to write out since the density and velocity bounds vary with position as well with time. Furthermore, while these inequalities are convincing evidence that well-behaved solutions to the equations exist, there remains the problem of constructing a proof. We have found it convenient to bypass these inequalities and to organize the proof in a different way, by reducing the Eq. (1.1) to a non-linear evolution equation on a Banach space. The estimates described above then reappear in the form of norm estimates on the "infinitesimal generator".

§ 3. Reformulation of the Differential Equations

Given any configuration $x \in \hat{\mathscr{X}}$, we will introduce a space of "neighboring configurations" in which the evolution takes place. Let $\mathscr{Y}_x$ denote the Banach space of sequences of pairs of real numbers $\zeta = (\xi_i, \eta_i)$

such that

$$\|\zeta\|_x = \sup_i \frac{|\xi_i| \vee |\eta_i|}{\log_+(q_i)} < \infty . \tag{3.1}$$

Given ζ in $\mathscr{Y}_x$, we denote by $x + \zeta$ the sequence of pairs of numbers $(q_i + \xi_i, p_i + \eta_i)$. The following two lemmas motivate the introduction of the space $\mathscr{Y}_x$: The first shows that, for any ζ in $\mathscr{Y}_x$, $x + \zeta$ is in $\hat{\mathscr{X}}$, and the second shows that any regular solution of the equations of motion with x as initial value stays inside the set of configurations of this form.

Lemma 3.1. *Let $x \in \hat{\mathscr{X}}$ and $\zeta \in \mathscr{Y}_x$. Then $x + \zeta \in \hat{\mathscr{X}}$. Moreover, $|x + \zeta|$ is bounded on bounded sets in $\mathscr{Y}_x$.*

Lemma 3.2. *Let $x \in \hat{\mathscr{X}}$, and let $x(t) = (q_i(t), p_i(t))$ be defined for t in a bounded open interval I containing zero. Assume:*

1. $\frac{dq_i(t)}{dt} = p_i(t)$ *for all* $t \in I$.
2. $x(0) = x$.
3. $|x(t)|$ *is bounded on* I.

Then, for each $t \in I$, we can write

$$x(t) = x + \zeta(t)$$

with $\zeta(t)$ in $\mathscr{Y}_x$, and $\|\zeta(t)\|_x$ in bounded on I.

We postpone the proofs of these lemmas. By Lemma 3.2, to find regular solutions of the differential equations, we can concentrate our attention on ones of the form $x + \zeta(t)$, with $\zeta(t) \in \mathscr{Y}_x$. In terms of the new dependent variable $\zeta(t) = (\xi_i(t), \eta_i(t))$, the differential equations become:

$$\frac{d\xi_i(t)}{dt} = p_i + \eta_i(t) \; ; \quad \frac{d\eta_i(t)}{dt} = \sum_{j \neq i} F(q_i + \xi_i(t) - q_j - \xi_j(t))$$

or, schematically,

$$\frac{d\zeta(t)}{dt} = A_x(\zeta(t)) , \tag{3.2}$$

where the derivative is to be taken a co-ordinate at a time and $A_x(\zeta)$ is the sequence of pairs of real numbers

$$A_x(\zeta) = \left(p_i + \eta_i, \sum_{j \neq i} F(q_i + \xi_i - q_j - \xi_j)\right) . \tag{3.3}$$

The following proposition shows that A_x defines a bounded non-linear operator on $\mathscr{Y}_x$:

Proposition 3.3. *Let F be a bounded function vanishing outside $(-R, R)$, and let $\zeta \in \mathscr{Y}_x$. Let $A_x(\zeta)$ be defined by* (3.3). *Then $A_x(\zeta) \in \mathscr{Y}_x$; moreover, there exist constants C, D (depending on $|x|$) such that*

$$\|A_x(\zeta)\|_x \leqq C + D\|\zeta\|_x \log_+(\|\zeta\|_x) \tag{3.4}$$

for all $\zeta \in \mathscr{Y}_x$.

Again we postpone the proof. Combining Lemma 3.1, Lemma 3.2, and Proposition 3.3, we see that Theorem 2.1 is equivalent to:

Theorem 2.1′. *Let F, x be as in Theorem 2.1, and let A_x be defined by (3.3). Then there is one and only one function $\zeta(t) = (\xi_i(t), \eta_i(t))$, defined for $-\infty < t < \infty$, with values in $\mathscr{Y}_x$, satisfying:*

1. $\dfrac{d\zeta(t)}{dt} = A_x(\zeta(t))$.
2. $\zeta(0) = 0$.
3. $\|\zeta(t)\|_x$ *is a locally bounded function of* t.

In 1., *the derivative is to be understood in the sense of the product topology on* $\mathscr{Y}_x$.

We will now give the proofs of Lemma 3.1, Lemma 3.2, and Proposition 3.3. Let us begin with Lemma 3.2. We have to show that

$$\frac{|q_i(t) - q_i|}{\log_+(q_i)} \quad \text{and} \quad \frac{|p_i(t) - p_i|}{\log_+(q_i)}$$

are bounded with respect to i and t. The differential equation and the boundedness of $|x(t)|$ imply that, for some K,

$$\left|\frac{d}{dt} q_i(t)\right| \leqq K \log_+(q_i(t)) .$$

Therefore,

$$|q_i(t) - q_i| \leqq K \int_I dt \log_+(q_i(t)) \leqq K' \log_+(|q_i| + \Delta Q_i) ,$$

where

$$\Delta Q_i = \sup_{t \in I} |q_i(t) - q_i| ,$$

and K' is K times the length of I.

Thus

$$\Delta Q_i \leqq K' \log_+(|q_i| + \Delta Q_i) \leqq K' [\log_+(q_i) + \log_+(\Delta Q_i)] ,$$

where, to get the second inequality, we have used (2.2). Rearranging this inequality, we get

$$\frac{\Delta Q_i}{\log_+(q_i)} \leqq K' \left[1 + \frac{\log_+(\Delta Q_i)}{\log_+(q_i)}\right] ,$$

which implies that $\dfrac{\Delta Q_i}{\log_+(q_i)}$ is bounded, i.e., that $\dfrac{|q_i(t) - q_i|}{\log_+(q_i)}$ is bounded with respect to i and t. The boundedness of $\dfrac{|p_i(t) - p_i|}{\log_+(q_i)}$ follows at once, since

$$\begin{aligned} \frac{|p_i(t) - p_i|}{\log_+(q_i)} &\leqq \frac{|p_i(t)|}{\log_+(q_i(t))} \cdot \frac{\log_+(|q_i| + \Delta Q_i)}{\log_+(q_i)} + \frac{|p_i|}{\log_+(q_i)} \\ &\leqq |x(t)| \cdot \sup_i \frac{\log_+(|q_i| + \Delta Q_i)}{\log_+(q_i)} + |x(0)| . \end{aligned}$$

This completes the proof of Lemma 3.2.

The proofs of Lemma 3.1 and Proposition 3.3 both involve some tedious particle-tracing estimates which are isolated in the following lemma.

Lemma 3.4. *There exists a constant K such that, for all $x \in \mathscr{X}$, all $\beta > \alpha$, all $\lambda \geqq 1$, and all sequences (ξ_i) of real numbers such that*

$$\sup_i \{|\xi_i|/\log_+(q_i)\} \leqq \lambda ,$$

the inequality

$$\#\{j : q_j + \xi_j \in [\alpha, \beta]\} \leqq |x|[\beta - \alpha + K\lambda(\log_+(\lambda) + \log_+(|\alpha| \vee |\beta|))] \quad (3.5)$$

holds. (The notation $\#X$, X a set, denotes the number of elements in the set).

Proof. It is enough to prove the lemma with the added restriction that α and β have the same sign, since we can prove the general result from this more restricted one by breaking up any interval into a piece to the right of the origin and a piece to the left. By symmetry, we can assume $\alpha \geqq 0$.

Let

$$W = \{q \in \boldsymbol{R} : q + \xi \in [\alpha, \beta] \text{ for some } \xi \text{ with } |\xi| \leqq \lambda \log_+(q)\} .$$

We will proceed by estimating the length of W and then applying the definition of $|x|$ to estimate the number of q_j's in W.

The first remark we need is the following: If

$$a \geqq 1 + 2\lambda , \quad (3.6)$$

then

$$q - \lambda \log_+(q) \leqq a \quad \text{implies} \quad q - a < 2\lambda \log_+(a) . \quad (3.7)$$

To prove this remark, we first observe that $q - a \leqq \lambda \log_+(q)$, so it suffices to prove that $q < a^2$. But $q - \lambda \log_+(q)$ is a strictly increasing function of q for $q \geqq \lambda$, and, by the hypotheses of (3.7), $a \geqq q - \lambda \log_+(q)$, so we have only to prove that

$$a^2 - \lambda \log_+(a^2) > a .$$

Dividing this inequality by a, using the fact that $\log_+(a^2) = 2\log_+(a)$, and transposing, we see that it suffices to show:

$$a > 1 + \frac{2\lambda \log_+(a)}{a} .$$

But $1 + \frac{2\lambda \log_+(a)}{a} < 1 + 2\lambda \leqq a$ [by (3.6)], so this inequality holds and our remark is proved.

Now let $q \leqq 0$ and suppose

$$q + \lambda \log_+(q) \geqq \alpha , \quad \text{i.e.,} \quad |q| - \lambda \log_+(|q|) \leqq -\alpha \leqq 0 .$$

Applying the above remark, with $a = 1 + 2\lambda$ and $|q|$ replacing q, we see that

$$|q| < (1 + 2\lambda) + 2\lambda \log_+(1 + 2\lambda) . \quad (3.8)$$

To save writing, we denote the right-hand side of this inequality by h.

Similarly, if $q - \lambda \log_+(q) \leqq \beta$, then

a) If $\beta < 1 + 2\lambda, q < h$. (3.9)

b) If $\beta \geqq 1 + 2\lambda, q < \beta + 2\lambda \log_+(\beta)$. (3.10)

Combining (3.8), (3.9), and (3.10), we see that

$$W \subset (-h, h) \cup (0, \beta + 2\lambda \log_+(\beta)) .$$

We can reduce the second interval as follows: If $q \in (0, \beta) \cap W$, then

$$\alpha \leqq q + \lambda \log_+(q) < q + \lambda \log_+(\beta) ,$$

i. e.

$$q > \alpha - \lambda \log_+(\beta) .$$

Thus,

$$W \subset (-h, h) \cup (\alpha - \lambda \log_+(\beta), \beta + 2\lambda \log_+(\beta)) .$$

Applying the definition of $|x|$, we see that

$$\# \{j : q_j \in W\} \leqq |x| [\beta - \alpha + 2h + 3\lambda \log_+(\beta)] ;$$

inserting the value of h and making some elementary re-arrangements completes the proof of the lemma.

We can now give the proofs of Lemma 3.1 and Proposition 3.3. To prove Lemma 3.1, we have to find bounds on

$$\sup_i \left\{ \frac{|p_i + \eta_i|}{\log_+(q_i)} \right\} \text{ and } \sup \left\{ \frac{N_{(\alpha,\beta)}(x + \zeta)}{\beta - \alpha} : \beta - \alpha > \log_+ \frac{(\alpha + \beta)}{2} \right\}$$

valid for all ζ with $\|\zeta\|_x \leqq \lambda$, where we can assume $\lambda \geqq 1$.

The momentum bound is immediate since

$$\frac{|p_i + \eta_i|}{\log_+(q_i)} \leqq \frac{|p_i|}{\log_+(q_i)} + \frac{|\eta_i|}{\log_+(q_i)} \leqq |x| + \|\zeta\|_x . \tag{3.11}$$

To get the density bounds, we apply Lemma 3.4 to show that for any $\beta > \alpha$,

$$\frac{N_{(\alpha,\beta)}(x + \zeta)}{\beta - \alpha} \leqq |x| \left[1 + K\lambda \left(\frac{\log_+(\lambda)}{\beta - \alpha} + \frac{\log_+(|\alpha| \vee |\beta|)}{\beta - \alpha} \right) \right] .$$

Using the equation

$$|\alpha| \vee |\beta| = \left| \frac{\alpha + \beta}{2} \right| + \left| \frac{\alpha - \beta}{2} \right|$$

and the sub-additivity of $\log_+$, we see that, if $\beta - \alpha > \log_+ \left(\frac{\beta + \alpha}{2} \right)$, we have

$$\frac{N_{(\alpha,\beta)}(x + \zeta)}{\beta - \alpha} \leqq |x| \left[1 + K\lambda \left(\log_+(\lambda) + \frac{\log_+ \left(\frac{\beta - \alpha}{2} \right)}{\beta - \alpha} + 1 \right) \right]$$

which gives the desired bound on the density and completes the proof of Lemma 3.1.

To prove Proposition 3.3, it suffices to find constants C and D such that

a) $\sup_i \frac{|p_i + \eta_i|}{\log_+(q_i)} \leqq C + D\lambda \log_+(\lambda)$,

b) $\sup_i \frac{|\sum_{j \neq i} F(q_i + \xi_i - q_j - \xi_j)|}{\log_+(q_i)} \leqq C + D\lambda \log_+(\lambda)$

whenever $\|\zeta\|_x \leqq \lambda$ and $\lambda \geqq 1$. (We have introduced λ just to avoid having to discuss separately what happens for $\|\zeta\|_x$ small).

We have already made the necessary momentum estimate (3.11). To get the estimate on the forces, we first write:

$$\left|\sum_{j \neq i} F(q_i + \xi_i - q_j - \xi_j)\right| \leqq M \# \{j : q_j + \xi_j \in [q_i + \xi_i - R, q_i + \xi_i + R]\}.$$

But

$$\log_+(|q_i + \xi_i| + R) \leqq \log_+(q_i) + \log_+(\lambda \log_+(q_i)) + \log_+(R) \leqq 2 \log_+(q_i) + \log_+(\lambda) + \log_+(R).$$

Hence, here are constants C and D such that

$$\left|\sum_{j \neq i} F(q_i + \xi_i - q_j - \xi_j)\right| \leqq [C + D\lambda \log_+(\lambda)] \log_+(q_i),$$

so inequality b) is satisfied and the proposition is proved.

We will now make a brief digression to show in a heuristic way how the norm estimates of Proposition 3.3 imply that the differential equations cannot drive infinitely many particles into a finite region in finite time. Although the argument we will give is not a necessary part of the proof of the main theorem, it illuminates the role played by the choice of a logarithmic rate of growth of velocities and densities in the proof of a global existence theorem; we will also obtain an intermediate result needed in § 4.

From the differential equation $\frac{d\zeta(t)}{dt} = A_x(\zeta(t))$ and the initial condition $\zeta(0) = 0$, it is at least plausible that the inequality

$$\|\zeta(t)\|_x \leqq \int_0^t d\tau \, \|A_x(\zeta(\tau))\|_x$$

holds for $t \geqq 0$. Using the estimate

$$\|A_x(\zeta)\|_x \leqq C + D \|\zeta\|_x \log_+(\|\zeta\|_x)$$

we get:

$$\|\zeta(t)\|_x \leqq \int_0^t d\tau [C + D\|\zeta(\tau)\|_x \log_+(\|\zeta(\tau)\|_x)].$$

Hence, if $h(t)$ is the solution of the integral equation:

$$h(t) = \int_0^t d\tau [C + Dh(\tau) \log_+(h(\tau))], \tag{3.12}$$

it is again at least plausible that

$$\|\zeta(t)\|_x \leqq h(t)\,,$$

for all $t \geqq 0$ for which $h(t)$ is defined. Thus, to show that $\|\zeta(t)\|_x$ cannot go to infinity in a finite time, it suffices to prove that $h(t)$ is defined for all t, i.e., that the solution of (3.12) does not go to infinity in a finite time. But by elementary calculus it is easily seen that $h(t)$ is given implicitly by

$$t = \int_0^h \frac{ds}{C + Ds\log_+(s)}$$

and that, since $1/s \log s$ is not integrable at infinity, $h(t)$ does not go to infinity unless t does also. Thus, we have an *a priori* estimate on the norm of $\zeta(t)$ valid for all t, so we can expect to be able to prove a global existence theorem if we can prove a local one.

If, instead of allowing density and velocity fluctuations to increase like the logarithm of the distance from the origin, we allow a faster increase (e.g., like some power of the distance), we can carry through most of the constructions and estimates of this section. However, $\|A(\zeta)\|$ increases more rapidly with $\|\zeta\|$ in this case; the reciprocal of the bound is no longer non-integrable at infinity; and our technique for proving a global existence theorem fails. The choice of a logarithmic growth rate is thus to a large extent determined by two conflicting requirements: on the one hand, if we take a growth rate which is significantly faster, we are unable to prove a global existence theorem; on the other hand, if we take a growth rate which is significantly slower, we do not get enough allowed configurations for our intended applications to statistical mechanics.

§ 4. Proof of the Main Theorem

If the non-linear operator A_x satisfied a norm Lipschitz condition on each bounded set in $\mathscr{Y}_x$, standard theorems would enable us to conclude the existence and uniqueness of solutions of the equation:

$$\frac{d\zeta(t)}{dt} = A_x(\zeta(t))\,.$$

Unfortunately, the operator A_x almost never satisfies a norm Lipschitz condition (no matter how regular the potential is assumed to be), and it is not even norm-continuous in general. We will show, however, that A_x satisfies a Lipschitz condition of a very special kind in the product topology on $\mathscr{Y}_x$, and that this Lipschitz condition allows the standard existence proofs to be carried out almost exactly as in the Banach-space case.

To simplify the notation in this section, we will assume that we are dealing with a definite initial configuration $x = (q_i, p_i)$, and we will therefore write $\|\zeta\|$ instead of $\|\zeta\|_x$; $\mathscr{Y}$ instead of $\mathscr{Y}_x$, and A instead of A_x.

For any positive real number m, define a semi-norm ${}_m\| \ \|$ on $\mathscr{Y}$ by

$$\left.\begin{aligned} {}_m\|\zeta\| &= \sup\left\{\frac{|\xi_i| \vee |\eta_i|}{\log_+(q_i)} : |q_i| \leqq m\right\} \text{ if this set is non-empty} \\ &= 0 \text{ otherwise .} \end{aligned}\right\} \tag{4.1}$$

Evidently, the set of semi-norms $\{{}_m\| \ \|\}$ defines the product topology on $\mathscr{Y}$; furthermore,

$$\|\zeta\| = \sup {}_m\|\zeta\| .$$

The following lemma gives the Lipschitz condition satisfied by A:

Lemma 4.1. *Let F satisfy the hypotheses of Theorem 2.1, and let a real number d be given. Then there exists a constant B such that, for any $\alpha > 1$, there exists an m_0 such that, for all $m \geqq m_0$ and all ζ, ζ' with $\|\zeta\| \leqq d$, $\|\zeta'\| \leqq d$, we have:*

$${}_m\|A(\zeta) - A(\zeta')\| \leqq B \log_+(m) \, {}_{\alpha m}\|\zeta - \zeta'\| . \tag{4.2}$$

Some interpretation of this lemma may by helpful. What is asserted is that, on any norm-bounded set in $\mathscr{Y}$ (the ball of radius d), the m-norm of $A(\zeta) - A(\zeta')$ may be majorized by a constant multiple of the larger αm-norm of $\zeta - \zeta'$, provided that m is large enough. The "constant" can be taken to increase no faster than logarithmically with m, and to be independent of α.

Proof. By the definition of A and ${}_m\|\zeta\|$,

$${}_m\|A(\zeta) - A(\zeta')\| = \sup\left\{\frac{|\eta_i - \eta_i'| \vee |\Delta F_i|}{\log_+(q_i)} : |q_i| \leqq m\right\},$$

where

$$\Delta F_i = \sum_{j \neq i} [F(q_i + \xi_i - q_j - \xi_j) - F(q_i + \xi_i' - q_j - \xi_j')] . \tag{4.3}$$

Since

$$\frac{|\eta_i - \eta_i'|}{\log_+(q_i)} \leqq {}_m\|\zeta - \zeta'\| \quad (\text{for } |q_i| \leqq m) ,$$

we have only to estimate ΔF_i.

We first choose m_0 so that, if $m \geqq m_0$, and if $|q_j| \geqq \alpha m$, $|q_i| \leqq m$, $\|\zeta\| \leqq d$, then

$$|q_i + \xi_i - q_j - \xi_j| > R .$$

(Here, R is some number such that $F(q) = 0$ for $|q| \geqq R$.) This can be done by choosing m_0 so that

$$2d \cdot \log_+(\alpha m_0) + R < (\alpha - 1) \, m_0 .$$

Next, using Lemma 3.1, we see that there is an E such that

$$\#\{j : |q_i + \xi_i - q_j - \xi_j| \leqq R\} \leqq E \log_+(q_i) \tag{4.5}$$

whenever $\|\zeta\| \leqq d$.

In the sum defining ΔF_i, we can evidently omit all j's such that $F(q_i + \xi_i - q_j - \xi_j)$ and $F(q_i + \xi_i' - q_j - \xi_j')$ are both zero. By (4.5), the number of terms left is no greater than $2E \log_+(q_i)$. We estimate each remaining term using the Lipschitz condition satisfied by F; thus:

$$|F(q_i + \xi_i - q_j - \xi_j) - F(q_i + \xi_i' - q_j - \xi_j')| \leqq K\,[|\xi_i - \xi_i'| + |\xi_j - \xi_j'|]\,.$$

But now, if $m \geqq m_0$ and $|q_i| \leqq m$, (4.4) implies that $|q_j| \leqq \alpha m$ (or else the term in question would have been zero). Hence

$$|\xi_j - \xi_j'| \leqq \log_+(\alpha m)\, {}_{\alpha m}\|\zeta - \zeta'\|\,,$$

and

$$|\Delta F_i| \leqq 2E \log_+(q_i) \cdot K \cdot [\log_+(m) + \log_+(\alpha m)] \cdot {}_{\alpha m}\|\zeta - \zeta'\|\,.$$

This inequality immediately implies the lemma.

We can now proceed to construct the solution of the equations of motion. It is easy to see that the differential Eq. (3.2) and the boundary condition $\zeta(0) = 0$ are equivalent to the integral equation

$$\zeta(t) = \int_0^t A(\zeta(\tau))\, d\tau\,, \tag{4.6}$$

where the integral is to be evaluated co-ordinate by co-ordinate. We will solve this equation by sucessive approximations: Let

$$\zeta_0(t) = 0\,; \quad \zeta_{n+1}(t) = \int_0^t A(\zeta_n(\tau))\, d\tau \quad \text{for} \quad n \geqq 0\,. \tag{4.7}$$

Proposition 4.2. *Let F satisfy the hypotheses of Theorem* 2.1, *and let $\zeta_n(t)$ be defined by* (4.7). *Then:*

1. *For each t, $\zeta_n(t)$ converges in the product topology on $\mathscr{Y}$ to a limit $\zeta(t)$.*
2. *For any m, ${}_m\|\zeta_n(t) - \zeta(t)\|$ converges to zero as n goes to infinity; the convergence is uniform in t on any bounded set.*
3. *The function $\zeta(t)$ is a solution of* (4.6).
4. *$\|\zeta(t)\|$ is a locally bounded function of t; moreover $\|\zeta_n(t)\|$ is bounded in t on any bounded interval, uniformly in n.*

Proof. We will consider only $t \geqq 0$; the proof for $t < 0$ is obtained from the argument we give here by changing some signs. Let C, D be as in Proposition 3.3, i.e., such that

$$\|A(\zeta)\| \leqq C + D\,\|\zeta\| \log_+(\|\zeta\|)\,.$$

Let h be the solution of the integral equation

$$h(t) = \int_0^t d\tau\,[C + Dh(\tau) \log_+(h(\tau))]\,;$$

we saw in § 3 that $h(t)$ is defined for all positive t.

We now claim that

$$\|\zeta_n(t)\| \leqq h(t)$$

for all $t \geqq 0$. The assertion is clearly true for $n = 0$. On the other hand, it is easy to see that

$$\begin{aligned}\|\zeta_{n+1}(t)\| &\leqq \int_0^t d\tau \, \|A(\zeta_n(\tau))\| \\ &\leqq \int_0^t d\tau [C + D \, \|\zeta_n(\tau)\| \log_+(\|\zeta_n(\tau)\|)] \, .\end{aligned}$$

Hence, if $\|\zeta_n(\tau)\| \leqq h(\tau)$, we have

$$\|\zeta_{n+1}(t)\| \leqq \int_0^t d\tau [C + Dh(\tau) \log_+(h(\tau))] = h(t) \, ,$$

and (4.8) follows by induction.

Now for any $T > 0$, apply Lemma 4.1 to get B such that

$${}_m\|A(\zeta) - A(\zeta')\| \leqq B \log_+(m) \, {}_{\alpha m}\|\zeta - \zeta'\| \, ,$$

whenever $\|\zeta\|$ and $\|\zeta'\|$ are not larger than $h(T)$ and m is large enough. Choose $\alpha > 1$ so that $B \log(\alpha) T < 1$. If m is large enough and if $0 \leqq t \leqq T$, we have:

$$\begin{aligned}{}_m\|\zeta_{n+1}(t) - \zeta_n(t)\| &\leqq \int_0^t d\tau \, {}_m\|A(\zeta_n(\tau)) - A(\zeta_{n-1}(\tau))\| \\ &\leqq B \log_+(m) \int_0^t d\tau \, {}_{\alpha m}\|\zeta_n(\tau) - \zeta_{n-1}(\tau)\| \, .\end{aligned}$$

Repeating this argument n times, we get:

$$\begin{aligned}{}_m\|\zeta_{n+1}(t) - \zeta_n(t)\| &\leqq B^n \log_+(m) \ldots \log_+(\alpha^{n-1} m) \\ &\quad \cdot \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \ldots \int_0^{\tau_{n-1}} d\tau_n \, {}_{\alpha^n m}\|\zeta_1(\tau)\| \, .\end{aligned}$$

Since

$${}_{\alpha^n m}\|\zeta_1(t)\| \leqq \|\zeta_1(t)\| \leqq h(T) \, ,$$

we get finally

$${}_m\|\zeta_{n+1}(t) - \zeta_n(t)\| \leqq \frac{B^n \cdot \log_+(m) \ldots \log_+(\alpha^{n-1} m) \cdot h(T) \cdot T^n}{n!} \, .$$

The ratio of succeeding terms on the right is

$$\frac{B \log_+(\alpha^n T) \cdot T}{n+1} \, ;$$

as n goes to infinity, this ratio approaches $B \log(\alpha) T$ which, by the choice of α, is less than one. Hence

$$\sum_{n=0}^{\infty} {}_m\|\zeta_{n+1}(t) - \zeta_n(t)\| \, ,$$

converges uniformly in t on $[0, T]$. This proves statements 1., 2., and 4. of Proposition 4.2. Statement 3. follows from statements 2. and 4., and the Lipschitz condition.

Remark 4.3. All the above estimates have been made for a definite choice of the initial configuration x. It is easy to see that the convergence of ${}_m\|\zeta_n(t) - \zeta(t)\|$ to zero and the bound on $\|\zeta_n(t)\|$ can be taken to be uniform in x on $\{x : |x| \leqq \delta\}$ for any real number δ.

Proposition 4.4. *Let $\zeta(t)$ and $\zeta'(t)$ be solutions of the integral Eq. (4.6); suppose $\|\zeta(t)\| \leqq M$ and $\|\zeta'(t)\| \leqq M$ for $|t| \leqq T$. Then $\zeta(t) = \zeta'(t)$ for $|t| \leqq T$.*

Proof. Again we consider only $t \geqq 0$. Choose B so that the Lipschitz condition (4.2) holds for $\|\zeta\| \leqq M$; $\|\zeta'\| \leqq M$. Choose $\alpha > 1$ so that $BT\log(\alpha) < 1$; then for m large enough and $0 \leqq t \leqq T$

$$
\begin{aligned}
{}_m\|\zeta(t) - \zeta'(t)\| &\leqq \int_0^t d\tau \, {}_m\|A(\zeta(\tau)) - A(\zeta'(\tau))\| \\
&\leqq B \log_+(m) \int_0^t d\tau \, {}_{\alpha m}\|\zeta(\tau) - \zeta'(\tau)\| \, .
\end{aligned}
$$

Iterating n times and using the fact that $\|\zeta(\tau) - \zeta'(\tau)\| \leqq 2M$ for $0 \leqq \tau \leqq T$, we get:

$$
{}_m\|\zeta(t) - \zeta'(t)\| \leqq \frac{B^n T^n \log_+(m) \ldots \log_+(\alpha^{n-1} m) \cdot 2M}{n!} \, .
$$

As before, the ratio of succeeding terms on the right approaches a limit which is less than one, so

$$
{}_m\|\zeta(t) - \zeta'(t)\| = 0 \, .
$$

This is true for any large m, so $\zeta(t) = \zeta'(t)$.

Combining Propositions 4.2 and 4.4 gives Theorem 2.1′ which, by the discussion in § 3, is equivalent to Theorem 2.1.

Acknowledgements. I am grateful to G. Gallavotti, S. Miracle, and D. Ruelle for valuable discussions in the course of this work and to M. L. Motchane for his hospitality at the Institut des Hautes Etudes Scientifiques.

References

1. Ruelle, D.: J. Math. Phys. 8, 1657 (1967).
2. — Ann. of Phys. **25**, 109 (1963).

O. E. Lanford III
Department of Mathematics
University of California
Berkeley, California

Commun. math. Phys. 9, 327—338 (1968)

Statistical Mechanics of Quantum Spin Systems. III

Oscar E. Lanford III
I. H. E. S., Bures-sur-Yvette
and
Derek W. Robinson
CERN — Geneva

Received May 10, 1968

Abstract. In the algebraic formulation the thermodynamic pressure, or free energy, of a spin system is a convex continuous function P defined on a Banach space $\mathfrak{B}$ of translationally invariant interactions. We prove that each tangent functional to the graph of P defines a set of translationally invariant thermodynamic expectation values. More precisely each tangent functional defines a translationally invariant state over a suitably chosen algebra $\mathfrak{A}$ of observables, i. e., an equilibrium state. Properties of the set of equilibrium states are analysed and it is shown that they form a dense set in the set of all invariant states over $\mathfrak{A}$. With suitable restrictions on the interactions, each equilibrium state is invariant under time-translations and satisfies the Kubo-Martin-Schwinger boundary condition. Finally we demonstrate that the mean entropy is invariant under time-translations.

1. Introduction

The purpose of this paper is to continue the general analysis of quantum spin systems which was presented in [1, 2] and [3]. In [2] we gave an algebraic formulation of the mathematical framework of quantum spin systems and showed that the thermodynamic pressure, or free energy, P could be considered as a convex continuous function defined on a Banach space of translationally invariant interactions. Further it was shown that the pressure also served as a generating functional of equilibrium states in the sense that the functional derivatives, i.e., the tangent functionals to the graph of P, determined translationally invariant states over a suitably chosen C^* algebra $\mathfrak{A}$ of observables. The states introduced in this manner play the same role as the more conventionally used correlation functions or thermodynamic expectation values. The results of [2] were, however, incomplete in the sense that we could only rigorously establish that P generated equilibrium states under certain restrictive conditions. In particular it was shown that if the interaction Φ were such that the tangent functional to the graph of P at Φ was unique then this tangent functional determined an equilibrium state. It was further shown that the equilibrium states obtained under such conditions described pure thermodynamic phases. This latter result

was derived by establishing and using a variational principle for the pressure which involves the mean entropy introduced in [1]. In the following we complete the results of [2] by proving that each tangent functional to the graph of P determines an equilibrium state, thus covering the situation when mixtures of phases can occur. Further we establish a variational principle for the mean entropy which involves the pressure and also show that every translationally invariant state over $\mathfrak{A}$ can be approximated by physical equilibrium states. Next we extend the results of [3] by proving that if the interactions are such that time translations correspond to a one-parameter group of automorphisms of $\mathfrak{A}$ then the corresponding equilibrium states are invariant under such translations and satisfy the Kubo-Martin-Schwinger boundary condition. Finally, we demonstrate that the mean entropy is invariant under time-translations.

It should perhaps be pointed out that whilst we work in an essentially quantum mechanical setting the results we derive also have relevance for classical spin systems and lattice gases. In fact the analysis of [1, 2] was based on earlier works [4, 5, 6] in a classical framework; many of our present results can be directly transcribed to this framework.

2. Convexity Theorems

The aim of this Section is to derive two mathematical theorems concerning the tangent planes to the graph of a convex function; the physical application of these results will be dealt with in the following Section.

Lemma 1. *Let X and Y be complete metric spaces and let Y be separable. If $Z \subset X \times Y$ is a residual set, i.e., the complement of a set of first category, then there is a residual set $X_1 \subset X$ such that for all $x \in X_1$ the set $Z \cap (\{x\} \times Y)$ is a residual set in $\{x\} \times Y$.*

Proof. We may assume that Z is open and dense and then it is sufficient to find $X_1 \subset X$ such that $Z \cap (\{x\} \times Y)$ is dense in $\{x\} \times Y$ for all $x \in X_1$. Let $a_1, a_2 \ldots$ be a denumerable dense set in Y and define W_i by

$$W_i = \Pi_1 \left\{ z \in Z;\, d(\Pi_2(z); a_i) < \frac{1}{i} \right\}$$

where $\Pi_1(z), \Pi_2(z)$ denote the co-ordinates of z and $d(.;.)$ the metric in Y. Clearly W_i is open and dense. If $x_0 \in \bigcap_i W_i$ it follows that for each i there is a $y_i \in Y$ such that $(x_0, y_i) \in Z$ and $d(y_i; a_i) < \frac{1}{i}$. Then $\{y_i\}$ is dense in Y.

Corollary. *Let $\mathfrak{X}$ be a Banach space and Y a subset of the closed unit ball in $\mathfrak{X}$ which is a residual set. Let $\omega \in \mathfrak{X}$ be a unit vector. It follows that for $\varepsilon > 0$ there is a unit vector ω' with $\|\omega - \omega'\| < \varepsilon$ such that*

$$\{\lambda;\, \lambda\omega' \in Y, -1 \leqq \lambda \leqq 1\}$$

is a residual set in $[-1, 1]$.

In the following we will need the notion of the tangent functional to the graph of a convex function; a tangent functional is essentially a tangent plane normalised suitably. If f is a convex continuous function defined on a Banach space $\mathfrak{X}$ an element $y_x \in \mathfrak{X}'$ is said to be a tangent functional to the graph of f at x if

$$f(x+\omega) \geqq f(x) + y_x(\omega)\,, \qquad \omega \in \mathfrak{X}\,.$$

If f is differentiable at x the only tangent functional at x is the derivative Df_x.

Theorem 1. *Let f be a convex function defined and continuous on a neighbourhood of zero in a separable Banach space $\mathfrak{X}$. Let $y \in \mathfrak{X}'$ be a tangent functional at zero to the graph of f. It follows that y is contained in the weak * closed convex hull of the set of tangent functionals $\mathfrak{z}$ defined by $\mathfrak{z} = \{z \in \mathfrak{X}'$; there exist $x_\alpha \to 0$ (in norm) such that f is differentiable at each x_α and weak * $\lim_\alpha Df x_\alpha = z\}$.*

Proof. From convexity we may directly deduce that for a sufficiently small neighbourhood $\mathscr{V}$ of zero there is an $M > 0$ such that $|f(x) - f(y)| \leqq M\,\|x - y\|$ for $x, y \in \mathscr{V}$. In particular it follows that $\|y\| \leqq M$ and $\|z\| \leqq M$ for all $z \in \mathfrak{z}$. Now assume the theorem is false; then there exists a weak * continuous linear functional on $\mathfrak{X}'$, i.e., an element of $\mathfrak{X}$, which strongly separates y from $\mathfrak{z}$. In particular there exists a unit vector $\omega \in \mathfrak{X}$ and a real number m such that $y(\omega) > m$ and $z(\omega) \leqq m$ for all $z \in \mathfrak{z}$. Since $\mathfrak{z}$ is bounded we can replace ω by any ω' sufficiently close to it and still obtain separation. But as f is convex it is differentiable on a residual set and hence, using the preceding corollary, we see that we may assume that f is differentiable at $\lambda\omega$ for all λ in a residual subset of $[-1, 1]$. By weak * compactness we can choose a net $\lambda_\alpha \to 0$ and $\lambda_\alpha \geqq 0$ such that f is differentiable at each $\lambda_\alpha\omega$ and $Df_{\lambda_\alpha\omega}$ converges in the weak * topology on $\mathfrak{X}'$. Since $\lim_\alpha Df_{\lambda_\alpha\omega} \in \mathfrak{z}$ we have $\lim_\alpha Df_{\lambda_\alpha\omega}(\omega) \leqq m$, i.e.,

$$\lim_\alpha \left[\frac{d}{d\lambda} f(\lambda\omega)\right]_{\lambda=\lambda_\alpha} \leqq m\,. \tag{1}$$

However, since $\lambda_\alpha \geqq 0$ the slope of any tangent to $f(\lambda\omega)$ at zero must be majorised by the left-hand side of (1). But, since y is a tangent functional to the graph of f at zero, there is a tangent line to the function $\lambda \to f(\lambda\omega)$ at $\lambda = 0$ with slope $y(\omega)$. Hence $y(\omega) \leqq m$. But this contradicts our assumption $y(\omega) > m$, and thus the theorem is proved.

Lemma 2[1]. *Let f be a non-negative C^∞ function defined on R^n; then the derivative Df of f satisfies the inequality*

$$\min_{\|x\| < a} (1 + \|x\|)\,\|Df\|\,(x) \leqq \frac{f(0)}{\log(1+a)}\,, \quad a \in R^+$$

[1] The proofs of this and the following lemma are based upon suggestions by D. RUELLE.

and hence

$$\min_{x \in R^n} (1 + \|x\|) \|Df\|(x) = 0$$

where the $\|.\|$ *refers to the usual Euclidean norm on* R^n *which is also identified with its dual.*

Proof. We may assume $\|Df\| > 0$ for $\|x\| \leqq a$ because the contrary assumption leads trivially to the desired result. Now, let $x(t)$ be an arc in R^n with $x(0) = 0$ and such that

$$\frac{d}{dt} x(t) = - \frac{Df}{\|Df\|}(x(t)) .$$

We note that for $t > 0$ we have $\|x(t)\| \leqq t$ and

$$0 \leqq f(x(a)) = f(0) + \int_0^a dt \frac{d}{dt} f(x(t))$$

$$= f(0) - \int_0^a dt \|Df\|(x(t))$$

$$\leqq f(0) - \min_{\|x\| \leqq a} (1 + \|x\|) \|Df\|(x) \int_0^a \frac{dt}{1+t} .$$

A simple rearrangement yields the desired result.

Lemma 3. *Let* f *be a convex continuous non-negative function defined on* R^n *and let* $a > 0$ *be given. There is an* $x \in R^n$, *with* $\|x\| \leqq a$ *and a tangent functional* h_x *to the graph of* f *at* x *such that* $(1 + \|x\|) \|h_x\| < 2f(0)/\log(1+a)$.

Proof. Let ϱ_n be a sequence of positive C^∞ functions of compact support with the following properties

1. $\int dx \varrho_n(x) = 1$
2. $\varrho_n * f \to f$ uniformly on compact sets
3. $(\varrho_n * f)(0) \leqq 2f(0)$.

Now $\varrho_n * f$ is non-negative, C^∞, and convex; therefore, there exists an x_n with $\|x_n\| \leqq a$ such that

$$(1 + \|x_n\|) \|D_{x_n}(\varrho_n * f)\| \leqq \frac{2f(0)}{\log(1+a)}$$

by lemma 2. Next, possibly passing to a subsequence, we can assume $x_n \to x$ and $h_n = D_{x_n}(\varrho_n * f) \to h_x$. We then have

$$(1 + \|x\|) \|h_x\| \leqq \frac{2f(0)}{\log(1+a)} .$$

But, by convexity, we also have

$$(\varrho_n * f)(x_n + \delta) \geqq (\varrho_n * f)(x_n) + h_n(\delta) , \qquad \delta \in R^n$$

and therefore

$$f(x+\delta) \geqq f(x) + h_x(\delta)$$

i.e., h_x is a tangent functional to the graph of f at x. This completes the proof of the lemma.

Theorem 2. *Let f be a convex continuous function defined on a separable Banach space $\mathfrak{X}$ and let $h \in \mathfrak{X}'$ have the properties that $h(x) \leqq f(x)$ for all $x \in \mathfrak{X}$. It follows that h is contained in the weak $*$ closure of the set of tangent functionals to the graph of f.*

Proof. We can suppose, without loss of generality, that $h = 0$. Now let $\omega_1, \omega_2, \ldots, \omega_n \in \mathfrak{X}$ and $\varepsilon > 0$ be given. We have to find an $x \in \mathfrak{X}$ and a tangent functional y_x to the graph of f at x such that $|y_x(\omega_i)| < \varepsilon$ for $i = 1, 2, \ldots, n$. Now by the Hahn-Banach extension theorem, it suffices to find an x in the linear subspace $\tilde{\mathfrak{X}}$ of $\mathfrak{X}$ spanned by $\omega_1, \ldots, \omega_n$ and a tangent functional $\tilde{y}_x \in \tilde{\mathfrak{X}}'$ such that $|\tilde{y}_x(\omega_i)| < \varepsilon$ for $i = 1, 2, \ldots, n$, i.e., we can, effectively, assume that $\mathfrak{X}$ is finite dimensional. The proof of the theorem is thus immediately given by lemma 3.

Note that x and the tangent functional y_x can be chosen such that we not only have $|y_x(\omega_i)| < \varepsilon$ for $i = 1, 2, \ldots, n$ but also $|y_x(x)| < \varepsilon$. This remark, which will be of importance in the next Section, follows from the estimate given in lemma 3.

3. Equilibrium States

In this Section we apply the foregoing results to the characterization of the equilibrium states of a quantum spin system and to the derivation of certain properties of these states. The characterization we obtain completes earlier results obtained in [2] and [3]. We begin by recalling the mathematical framework associated with a quantum spin system.

A quantum spin system is described in terms of a simple separable C^* algebra $\mathfrak{A}$ of quasi-local observables and a collection $\{\mathfrak{A}(\Lambda)\}$ of C^* subalgebras of $\mathfrak{A}$, where Λ takes values on the finite subsets of Z^ν. Elements of the $\mathfrak{A}(\Lambda)$ are called strictly local observables. The algebras $\mathfrak{A}$ and $\mathfrak{A}(\Lambda)$, $\Lambda \subset Z^\nu$, satisfy the following properties

1. $\mathfrak{A}(\Lambda_1) \subset \mathfrak{A}(\Lambda_2)$ if $\Lambda_1 \subset \Lambda_2$
2. $\mathfrak{A}$ is the norm closure of $\bigcup_{\Lambda \in Z^\nu} \mathfrak{A}(\Lambda)$
3. $[\mathfrak{A}(\Lambda_1), \mathfrak{A}(\Lambda_2)] = 0$ if $\Lambda_1 \cap \Lambda_2 = \emptyset$
4. the group Z^ν of space translations is a subgroup of the automorphism group of $\mathfrak{A}$ and the action of these automorphisms is such that

$$A \in \mathfrak{A}(\Lambda) \to \tau_x A \in \mathfrak{A}(\Lambda + x), \quad x \in Z^\nu$$

and

$$\|[A, \tau_x B]\| \xrightarrow[|x| = \infty]{} 0, \quad A, B \in \mathfrak{A} \text{ and } x \in Z^\nu$$

5. for each $\Lambda \subset Z^\nu$, $\mathfrak{A}(\Lambda)$ is isomorphic to the matrix algebra of bounded operators $\mathfrak{B}(\mathfrak{H}_\Lambda)$ on a finite dimensional Hilbert space $\mathfrak{H}_\Lambda$.

The states, i.e., the normalized positive linear functionals over $\mathfrak{A}$, form a weakly compact convex subset E of $\mathfrak{A}'$ and the translationally invariant states, i.e., the states such that

$$\varrho(\tau_x A) = \varrho(A), \quad A \in \mathfrak{A}, \quad x \in Z^\nu$$

form a weakly compact convex subset $E \cap L_{Z^\nu}^{\downarrow}$ of E. The extremal elements $\mathscr{E}(E \cap L_{Z^\nu}^{\downarrow})$ of this latter subset enjoy many remarkable properties of an ergodic nature (see for example [7] and [8]) which allow the physical interpretation that they describe single thermodynamic phases. If we consider a state ϱ restricted to any subalgebra $\mathfrak{A}(\Lambda)$ then, by property 5. above, the state defines a positive operator ϱ_Λ on $\mathfrak{H}_\Lambda$ such that

$$\mathrm{Tr}_{\mathfrak{H}_\Lambda}(\varrho_\Lambda) = 1 \quad \text{and} \quad \mathrm{Tr}_{\mathfrak{H}_\Lambda}(\varrho_\Lambda A) = \varrho(A)$$

for $A \in \mathfrak{A}(\Lambda)$ [here and in the sequel, we tacitly identify $\mathfrak{A}(\Lambda)$ and $\mathfrak{B}(\mathfrak{H}_\Lambda)$]. The density matrices ϱ_Λ are related by certain compatibility conditions, but for our present purposes it suffices to note that we can define a local entropy $S_\varrho(\Lambda)$ of a state via

$$S_\varrho(\Lambda) = -\mathrm{Tr}_{\mathfrak{H}_\Lambda}(\varrho_\Lambda \log \varrho_\Lambda)$$

and, if ϱ is an invariant state, i.e., $\varrho \in E \cap L_{Z^\nu}^{\downarrow}$, a mean entropy via

$$S(\varrho) = \lim_{\Lambda \to \infty} \frac{S_\varrho(\Lambda)}{N(\Lambda)} = \inf_{\Lambda} \frac{S_\varrho(\Lambda)}{N(\Lambda)}$$

where $N(\Lambda)$ is the number of points in the set $\Lambda \subset Z^\nu$ and, for simplicity, here, and in the following, we take the limits over parallelepipeds whose sides each tend to infinity. The mean entropy defined in this manner is a non-negative affine upper semi-continuous function on $E \cap L_{Z^\nu}^{\downarrow}$ (for details, and proofs of these statements, see [1]).

Physically we consider the points $x \in Z^\nu$ as sites of particles or "spins", which interact together. In our rather abstract setting we introduce an interaction Φ as a function from the finite sets $X \subset Z^\nu$ to $\mathfrak{A}$ with values $\Phi(X) \in \mathfrak{A}(X)$. We assume

1. $\Phi(X)$ is Hermitian

2. $\Phi(X + a) = \tau_a \Phi(X)$ for $a \in Z^\nu$

and 3. $\|\Phi\| = \sum_{X \ni 0} \frac{\|\Phi(X)\|}{N(X)} < +\infty$.

With respect to the norm introduced in the last conditions the interactions Φ form a separable Banach space $\mathfrak{B}$. The finite range interactions, i.e., those interactions such that for $X \ni 0$ $\Phi(X) = 0$ unless $X \subset \Lambda$ for some finite Λ, form a dense subset $\mathfrak{B}_0 \subset \mathfrak{B}$. It is convenient to

introduce an auxiliary Banach space $\mathfrak{B}_1$, which we leave arbitrary up to the assumption that $\mathfrak{B}_0 \subset \mathfrak{B}_1 \subset \mathfrak{B}$ and $\mathfrak{B}_0$ is dense in $\mathfrak{B}_1$. The interaction energy of a spin system confined to the finite set Λ is defined for $\Phi \in \mathfrak{B}_1$ by

$$U_\Phi(\Lambda) = \sum_{X \subset \Lambda} \Phi(X) .$$

We also introduce the "interaction energy" at the origin by

$$A_\Phi = \sum_{X \ni 0} \frac{\Phi(X)}{N(X)} .$$

The following theorem gives information concerning the equilibrium states of spin systems with interactions $\Phi \in \mathfrak{B}_1$; in part the theorem summarises results already derived in [2].

Theorem 3. 1. *If $\Phi \in \mathfrak{B}_1$ then the thermodynamic pressure*

$$P(\Phi) = \lim_{\Lambda \to \infty} \frac{1}{N(\Lambda)} \log \mathrm{Tr}_{\mathfrak{H}_\Lambda}(e^{-U_\Phi(\Lambda)})$$

exists. The function $\Phi \to P(\Phi)$ is convex continuous on the Banach space $\mathfrak{B}_1$ and

$$|P(\Phi) - P(\Psi)| \leqq \|\Phi - \Psi\| , \quad \Phi, \Psi \varepsilon \mathfrak{B}_1 .$$

2. *If $\alpha_\Phi \in \mathfrak{B}_1'$ is a tangent functional to the graph of P at Φ, i.e.,*

$$P(\Phi + \Psi) \geqq P(\Phi) - \alpha_\Phi(\Psi) \quad \text{for all} \quad \Psi \varepsilon \mathfrak{B}_1$$

then α_Φ determines a state $\varrho_\Phi \in E \cap L_{\mathfrak{Z}^\nu}^{\perp}$ through the relation

$$\alpha_\Phi(\Psi) = \varrho_\Phi(A_\Psi) .$$

The states ϱ_Φ defined in this way will be called equilibrium states.

3. *If $T \subset \mathfrak{B}_1$ is the set of Φ such that the graph of P has a unique tangent functional at Φ then T is a residual set in $\mathfrak{B}_1$ and for $\Phi \varepsilon T$ the equilibrium state ϱ_Φ determined by the tangent functional α_Φ is ergodic i.e., $\varrho_\Phi \varepsilon \mathscr{E}(E \cap L_{\mathfrak{Z}^\nu}^{\perp})$. Further we have for $\Phi \varepsilon T$ the relation*

$$\alpha_\Phi(\Psi) = \varrho_\Phi(A_\Psi) = \lim_{\Lambda \to \infty} \frac{1}{Tr_{\mathfrak{H}_\Lambda}(e^{-U_\Phi(\Lambda)})} \mathrm{Tr}_{\mathfrak{H}_\Lambda}\left(e^{-U_\Phi(\Lambda)} \frac{U_\Psi(\Lambda)}{N(\Lambda)}\right) . \quad (2)$$

4. *The pressure P, the mean entropy S, and the set of equilibrium states are related as follows*

$$P(\Phi) = S(\varrho_\Phi) - \varrho_\Phi(A_\Phi) = \sup_{\varrho \varepsilon E \cap L_{\mathfrak{Z}^\nu}^{+}} \{S(\varrho) - \varrho(A_\Phi)\} , \quad \Phi \in \mathfrak{B}_1 , \quad (3)$$

where ϱ_Φ is any equilibrium state associated with Φ. The supremum in the last expression is reached by a unique state ϱ_Φ if, and only if, $\Phi \varepsilon T$.

5. *The pressure P, the mean entropy S, and the space $\mathfrak{B}_1$ of interactions are related as follows*

$$S(\varrho) = \inf_{\Phi \in \mathfrak{B}_1} \{P(\Phi) + \varrho(A_\Phi)\} \quad \text{for} \quad \varrho \varepsilon E \cap L_{\mathfrak{Z}^\nu}^{\perp} .$$

6. *The equilibrium states are weak * dense in the set* $E \cap L_{\mathcal{Z}^\nu}^{\perp}$ *of all translationally invariant states over* $\mathfrak{A}$.

Proof. Statements 1. and 3. together with parts of statement 4. are proved in [2]. In particular it is shown in this reference that the maximum principle (3) holds and that, for $\Phi \varepsilon T$, the tangent functional α_Φ determines an ergodic equilibrium state ϱ_Φ, the relation (2) is valid, and ϱ_Φ gives the unique supremum in (3). However it now follows directly from theorem 1 that a general tangent functional α_Φ determined an equilibrium state ϱ_Φ; in the present context theorem 1 states that a tangent functional α_Φ with $\Phi \notin T$ can be approximated weakly by convex combinations of tangent functionals α_Ψ with $\Psi \varepsilon T$. The facts that in general ϱ_Φ gives the maximum in (3) and that this maximum is unique only if $\Phi \varepsilon T$ follow from considerations reproduced in [2] and [3]. It remains to prove statements 5. and 6.; we begin with the latter.

Let $\varrho \varepsilon E \cap L_{\mathcal{Z}^\nu}^{\perp}$ be any invariant state; then from (3) we see that

$$P(\Phi) \geqq S(\varrho) - \varrho(A_\Phi) \geqq -\varrho(A_\Phi)$$

where we have used the non-negativity of S to obtain the second inequality. Thus the function $\Phi \to \alpha(\Phi) = \varrho(A_\Phi)$ is linear and its graph lies below the graph of P. Hence by theorem 2 α lies in the weak * closure of the set of tangent functionals to P and thus by statement 2. of the above theorem we obtain the desired result.

To prove statement 5. we note that by (3)

$$P(\Phi) + \varrho(A_\Phi) - S(\varrho) \geqq 0 \tag{4}$$

for $\Phi \in \mathfrak{B}_1$ and $\varrho \varepsilon E \cap L_{\mathcal{Z}^\nu}^{\perp}$. However, given $\varepsilon > 0$ we can choose $\Phi \in \mathfrak{B}_1$ and ϱ_Φ such that

$$S(\varrho) + \frac{\varepsilon}{2} > S(\varrho_\Phi) = P(\Phi) + \varrho_\Phi(A_\Phi) \tag{5}$$

and

$$|\varrho_\Phi(A_\Phi) - \varrho(A_\Phi)| < \frac{\varepsilon}{2}\,. \tag{6}$$

Here we have used the upper semi-continuity of S and the remark at the end of the proof of theorem 2. Combining (4), (5) and (6) we find with this choice of Φ

$$\varepsilon > P(\Phi) + \varrho(A_\Phi) - S(\varrho) \geqq 0\,.$$

This establishes the desired property and completes the proof of the theorem.

In the foregoing we have left a certain arbitrariness in the definition of the Banach space $\mathfrak{B}_1$. In the following, however, we will consider one specific Banach space which we define as the set of interactions $\Phi \in \mathfrak{B}$ which have the property that

$$\|\Phi\|_1 = \sum_{X \ni 0} \|\Phi(X)\| \exp\{N(X)\} < +\infty\,. \tag{7}$$

For this space of interactions it is possible to discuss the time development of the spin system. In particular, for each $\Phi \varepsilon \mathfrak{B}_1$ there exists a one-parameter group of automorphisms of the algebra $\mathfrak{A}$ of quasilocal observables corresponding to time translations. We denote the action of this group by $A \varepsilon \mathfrak{A} \to \tau_t^\Phi A \varepsilon \mathfrak{A}$ for $t \varepsilon R$; the action is defined by

$$\tau_t^\Phi A = \lim_{\Lambda \to \infty} e^{itU_\Phi(\Lambda)} A e^{-itU_\Phi(\Lambda)} \quad t \varepsilon R, \quad A \varepsilon \mathfrak{A}, \quad \Phi \varepsilon \mathfrak{B}_1 .$$

(The existence of this limit was established in [3] for a dense subset of $\mathfrak{B}_1$; RUELLE [9] has shown that the arguments of [3] can be improved to establish the existence for all $\Phi \varepsilon \mathfrak{B}_1$.)

Theorem 4. *If $\Phi \varepsilon \mathfrak{B}_1$, the space of interactions whose norm is given by (7), then any equilibrium state ϱ_Φ, defined by a tangent functional to the graph of the pressure P at Φ, has the following properties;*

1. *ϱ_Φ is invariant under time-translations, i.e.*

$$\varrho_\Phi(\tau_t^\Phi A) = \varrho_\Phi(A) \quad \textit{for all} \quad A \varepsilon \mathfrak{A}, \; t \varepsilon R .$$

2. *ϱ_Φ satisfies the Kubo-Martin-Schwinger boundary condition. Explicitly, for $A, B \varepsilon \mathfrak{A}$, the function $t \to \varrho_\Phi(A(\tau_t^\Phi B))$ extends to a bounded continuous function on the strip $0 \leqq \operatorname{Im}\{t\} \leqq 1$ which is analytic on the interior of the strip, and we have*

$$\varrho_\Phi(A(\tau_{t+i}^\Phi B)) = \varrho_\Phi((\tau_t^\Phi B)A) .$$

Proof. Let $T \subset \mathfrak{B}_1$ be the set of interactions at which the graph of P has a unique tangent plane. For Φ in T the properties stated in the theorem have already been proved in [3]; we will obtain the general statement from this result by an approximation argument using theorem 1. It is easy to see that weak limits of convex combinations of states satisfying 1. and 2. again satisfy 1. and 2.; hence, by theorem 1, it will suffice to prove the theorem in the special case in which

$$\varrho_\Phi = \lim_\alpha \varrho_{\Phi_\alpha}$$

where Φ_α is a net in T converging in norm to Φ and ϱ_{Φ_α} is the state determined by the unique tangent plane to the graph of P at Φ_α. Moreover, we can assume that A and B are strictly local; the assertions for general elements of $\mathfrak{A}$ are then obtained by a straightforward limiting argument.

It follows easily from the estimates in [9] that

$$\lim_\alpha \|\tau_t^\Phi A - \tau_t^{\Phi_\alpha} A\| = 0 .$$

uniformly for t in any bounded interval. Hence, using the invariance of ϱ_{Φ_α} under $\tau_t^{\Phi_\alpha}$, we get

$$\begin{aligned} |\varrho_\Phi(\tau_t^\Phi A) - \varrho_\Phi(A)| &\leqq |\varrho_\Phi(\tau_t^\Phi A) - \varrho_{\Phi_\alpha}(\tau_t^\Phi A)| \\ &\quad + \|\tau_t^\Phi A - \tau_t^{\Phi_\alpha} A\| + |\varrho_{\Phi_\alpha}(A) - \varrho_\Phi(A)| \end{aligned}$$

and the right-hand side goes to zero as $\alpha \to \infty$. This proves 1.

To prove 2., we first remark that

$$\frac{d}{dt}\tau_t^{\Phi_\alpha} B = \tau_t^{\Phi_\alpha}(B'_\alpha)$$

where

$$B'_\alpha = i \lim_{\Lambda\to\infty} [U_{\Phi_\alpha}(\Lambda), B]$$

and that $\|B'_\alpha\|$ is bounded with respect to α for any fixed B. Hence,

$$\varrho_{\Phi_\alpha}(A(\tau_t^{\Phi_\alpha} B))$$

is a net of continuous functions on the strip $0 \leqq \mathrm{Im}\{t\} \leqq 1$ which are holomorphic on the interior of the strip and whose derivatives are bounded uniformly in α and in t. Since

$$\lim_\alpha \varrho_{\Phi_\alpha}(A(\tau_t^{\Phi_\alpha} B)) = \varrho_\Phi(A(\tau_t^\Phi B))$$

$$\lim_\alpha \varrho_{\Phi_\alpha}(A(\tau_{t+i}^{\Phi_\alpha} B)) = \varrho_\Phi((\tau_t^\Phi B)A)$$

for all real t, this net converges pointwise to a function continuous and bounded in the closed strip, holomorphic on the interior of the strip, with the right boundary values, so 2. is proved.

4. Conservation of Entropy

Theorem 5. *Let $\Phi \varepsilon \mathfrak{B}_1$ and let ϱ be a translation-invariant state over $\mathfrak{A}$. For any $t \varepsilon R$, let the state ϱ_t over $\mathfrak{A}$ be defined by*

$$\varrho_t(A) = \varrho(\tau_t^\Phi A)\,.$$

Then, $S(\varrho_t) = S(\varrho)$ for all t.

Proof. By reversibility, it will be sufficient to show that $S(\varrho_t) \geqq S(\varrho)$, and, since S is upper semi-continuous, this will follow if we can show that ϱ_t can be approximated arbitrarily well by states with the same entropy as ϱ.

If a is a strictly positive integer, we let

$$\Lambda(a) = \{(n_1, \ldots, n_\nu) \in Z^\nu;\ -a < n_i \leqq a\}$$

$$N(a) = N(\Lambda(a)) = (2a)^\nu$$

$$\Gamma_a \quad = \{(2n_1 a, \ldots, 2n_\nu a);\ n_1, \ldots, n_\nu \in Z\}$$

and we let $x_1, x_2, \ldots$ be an enumeration of the elements of Γ_a. Define a one-parameter group of automorphisms ${}^a\tau_t^\Phi$ of $\mathfrak{A}$ by

$${}^a\tau_t^\Phi(A) = \lim_{N\to\infty} \exp\left\{\mathrm{it} \sum_{j=1}^N \tau_{x_j} U_\Phi(\Lambda(a))\right\} A \exp\left\{-\mathrm{it} \sum_{j=1}^N \tau_{x_j} U_\Phi(\Lambda(a))\right\}.$$

This one-parameter group of automorphisms corresponds to an interaction which differs from that defined by Φ only in that all interactions between translates of $\Lambda(a)$ by different elements of Γ_a are suppressed. Note that:

1. If $A \in \mathfrak{A}(\Lambda(a))$, ${}^a\tau_t^\Phi(A) = \exp\{\mathrm{i}t U_\Phi(\Lambda(a))\} A \exp\{-\mathrm{i}t U_\Phi(\Lambda(a))\}$.
2. If $x \in \Gamma_a$, $\quad \tau_x {}^a\tau_t^\Phi = {}^a\tau_t^\Phi \tau_x$.

Let

$$ {}^a\varrho_t(A) = \varrho({}^a\tau_t^\Phi(A)) $$

then ${}^a\varrho_t$ is a state over $\mathfrak{A}$ invariant under the subgroup Γ_a of Z^ν and its entropy is equal to that of ϱ. Therefore, if we define

$$ {}^a\bar{\varrho}_t(A) = \frac{1}{N(a)} \sum_{x \in \Lambda(a)} {}^a\varrho_t(\tau_x A) , $$

${}^a\bar{\varrho}_t$ is invariant under Z^ν and has the same entropy as ϱ. Taking into account the remarks at the beginning of the proof we see that all we have to prove is that

$$ \lim_{a \to \infty} {}^a\bar{\varrho}_t(A) = \varrho_t(A) $$

for all strictly local A in $\mathfrak{A}$.

By the translation invariance of ϱ,

$$ {}^a\bar{\varrho}_t(A) = \varrho\left(\frac{1}{N(a)} \sum_{x \in \Lambda(a)} \tau_{-x} {}^a\tau_t^\Phi \tau_x(A)\right) $$

so it will suffice to prove

$$ \lim_{a \to \infty} \frac{1}{N(a)} \sum_{x \in \Lambda(a)} \tau_{-x} {}^a\tau_t^\Phi \tau_x(A) = \tau_t^\Phi(A) . $$

Since A is strictly local, the terms in the sum on the left with $\tau_x(A) \notin \mathfrak{A}(\Lambda(a))$ become negligible as $a \to \infty$, so we can replace the left-hand side by:

$$ \lim_{a \to \infty} \frac{1}{N(a)} \sum_{x \in \Lambda(a)} \exp\{\mathrm{i}t U_\Phi(\Lambda(a) - x)\} A \exp\{-\mathrm{i}t U_\Phi(\Lambda(a) - x)\} . $$

Thus, to complete the proof it will suffice to prove the following assertion: For any $A \in \mathfrak{A}$, any t, and any $\varepsilon > 0$, there is a finite subset Λ of Z^ν such that, whenever $\Lambda' \supset \Lambda$,

$$ \|\exp\{\mathrm{i}t U_\Phi(\Lambda')\} A \exp\{-\mathrm{i}t U_\Phi(\Lambda')\} - \tau_t^\Phi(A)\| < \varepsilon . $$

This assertion is equivalent to the assertion that, for any t, any A, and any increasing sequence Λ_n of finite subsets of Z^ν whose union is all of Z^ν,

$$ \lim_{n \to \infty} \exp\{\mathrm{i}t U_\Phi(\Lambda_n)\} A \exp\{-\mathrm{i}t U_\Phi(\Lambda_n)\} = \tau_t^\Phi A . $$

For t small and A strictly local, this follows from the power series expansion for $\tau_t^\Phi(A)$. For t small and general A, the assertion follows since a sequence of isometries on a Banach space which converges strongly on a dense subset converges strongly everywhere. Finally, the assertion for general t is proved by remarking that, if a sequence of isometries on a Banach space converges strongly, the sequence of n^{th} powers converges strongly to the n^{th} power of the limit.

Acknowledgements. The authors thank Drs. G. Gallavotti, S. Miracle-Sole and D. Ruelle for helpful and stimulating discussions and Monsieur L. Motchane for his kind hospitality at the I. H. E. S.

References

1. Lanford, O. E., and D. W. Robinson: CERN preprint TH. 783 (J. Math. Phys. to appear).
2. Robinson, D. W.: Commun. Math. Phys. **6**, 151 (1967).
3. — Commun. Math. Phys. **7**, 337 (1968).
4. —, and D. Ruelle: Commun. Math. Phys. **5**, 288 (1967).
5. Gallavotti, G., and S. Miracle-Sole: Commun. Math. Phys. **5**, 317 (1967).
6. Ruelle, D.: Commun. Math. Phys. **5**, 324 (1967).
7. Kastler, D., and D. W. Robinson: Commun. Math. Phys. **3**, 151 (1966).
8. Ruelle, D.: Commun. Math. Phys. **3**, 133 (1966).
9. — Statistical mechanics, rigorous results. New York: Benjamin (to appear).

O. E. Lanford,
Department of Mathematics,
University of California
Berkeley, California (USA)

D. W. Robinson,
Theoretical Physics Division
CERN
CH 1211 Genf 23

Commun. math. Phys. 11, 257—292 (1969)

The Classical Mechanics of One-Dimensional Systems of Infinitely Many Particles

II. Kinetic Theory

O. E. Lanford III*
I.H.E.S., 91 — Bures-sur-Yvette, France

Received August 1, 1968

Abstract. We apply the existence theorem for solutions of the equations of motion for infinite systems to study the time evolution of measures on the set of locally finite configurations of particles. The set of allowed initial configurations and the time evolution mappings are shown to be measurable. It is shown that infinite volume limit states of thermodynamic ensembles at low activity or for positive potentials are concentrated on the set of allowed initial configurations and are invariant under the time evolution. The total entropy per unit volume is shown to be constant in time for a large class of states, if the potential satisfies a stability condition.

§ 1. Introduction

In [1], we proved an existence and uniqueness theorem for solutions of the equations of motion for systems of infinitely many particles. In this article, we will apply this theorem to the study of the time-evolution of states of classical statistical mechanics. Let us recall briefly the notation and results of [1]. We denote by $\mathscr{X}$ the set of locally finite configurations of labelled particles and by $[\mathscr{X}]$ the corresponding set of configurations of unlabelled particles. A state of classical statistical mechanics is a probability measure on $[\mathscr{X}]$ invariant under space translations. Let $\hat{\mathscr{X}}$ denote the set of labelled configurations satisfying conditions 1) and 2) of [1]. Theorem 2.1 of [1] asserts the existence of a solution of the equations

$$\frac{dq_i(t)}{dt} = p_i(t)$$

$$\frac{dp_i(t)}{dt} = \sum_{j \neq i} F(q_i(t) - q_j(t))$$

and the initial conditions

$$q_i(0) = q_i\,, \qquad p_i(0) = p_i\,,$$

provided that F has compact support and satisfies a Lipschitz condition and that the initial configuration (q_i, p_i) is in $\hat{\mathscr{X}}$; it also asserts the uni-

* On leave from: Department of Mathematics, University of California, Berkeley, California.

queness of the solution in the class of trajectories satisfying a certain regularity condition. This theorem enables us to define a one-parameter group T^t of evolution operators mapping $[\hat{\mathscr{X}}]$ (the set of unlabelled configurations corresponding to the set $\hat{\mathscr{X}}$ of labelled configurations) onto itself. If the mappings T^t are measurable, they define a time evolution of measures on $[\hat{\mathscr{X}}]$ and in particular of states of classical statistical mechanics which are concentrated on $[\hat{\mathscr{X}}]$ (i.e., for which $[\mathscr{X}]\backslash[\hat{\mathscr{X}}]$ has measure zero). This time evolution will be the object of our investigations.

In § 2, we develop some notation and tools which will be needed in the course of this article, and we restate in a convenient form the results we will need from [1]. Section 3 is devoted to some measurability questions which are technically important if not very interesting; we prove that $[\hat{\mathscr{X}}]$ is a Borel subset of $[\mathscr{X}]$ and that the time-evolution mapping $(t, x) \longmapsto T^t x$ is a Borel mapping from $\boldsymbol{R} \times [\hat{\mathscr{X}}]$ to $[\hat{\mathscr{X}}]$. In § 4 we show that for a state ϱ of classical statistical mechanics to be concentrated on $[\hat{\mathscr{X}}]$ it is sufficient that ϱ:

i) has a Maxwellian velocity distribution,

ii) has correlation functions $\bar{\varrho}_n(q_1, \ldots, q_n)$ of all orders (see [2]) admitting a majorization of the form

$$\bar{\varrho}_n(q_1, \ldots, q_n) \leqq \lambda^n$$

with λ independent of $n, q_1, \ldots, q_n$.

These two conditions are satisfied if ϱ is an infinite volume limit state obtained from the grand canonical ensemble at small activity and, for a non-negative potential, at arbitrarily large activity.

In § 5, we prove an approximation theorem which will be our main technical device for the rest of this article and which asserts that the time evolution of the part of an infinite system contained in a bounded region can be arbitrarily well approximated by the evolution of the corresponding part of a system with a large but finite number of particles. We apply this approximation theorem in § 6 to show that states obtained by taking infinite volume limits of grand canonical ensembles with a given twice-continuously-differentiable stable[1] potential of compact support are invariant under the time-evolution defined by that potential, provided again either that the activity is small or the potential non-negative.

In § 7, we show that the entropy per unit volume is conserved by the time evolution. Here, for the first time in our investigations, thermodynamic stability properties of the potential defining the interparticle

[1] A function Φ on $\boldsymbol{R}$ is a *stable* potential if there exists B such that, for all n and all $q_1, \ldots, q_n$,

$$\sum_{1 \leqq i < j \leqq n} \Phi(q_i - q_j) \geqq -nB .$$

force play an essential role. We assume that the interparticle force is the derivative of a potential Φ with compact support which may be written in the form

$$\Phi = \Phi_1 + \Phi_2$$

with Φ_1 stable and of compact support and Φ_2 non-negative, continuous, and strictly positive at the origin. We consider a state ϱ which is concentrated on $[\mathscr{X}]$ and which has:

i) finite mean kinetic energy in (0, 1),

ii) finite mean square number of particles in (0, 1),

and we let ϱ^t denote the time-evolved measure $\varrho \circ T^{-t}$.
Under all these hypotheses we show that, for any t, the entropy per unit volume of ϱ^t is equal to that of ϱ.

§ 2. Preliminaries

We will need, unfortunately, a rather complicated set of tools; these are developed in this section. Most of the results given here are of limited originality. In particular, Sections 2.1 and 2.2 draw heavily on [2], and Section 2.3 on [3].

2.1. The Space of Locally Finite Configurations

Recall that the set $\mathscr{X}$ of locally finite labelled configurations is defined as the set of all mappings (q_i, p_i) from an index set [which is either $(1, 2, 3, \ldots, n)$ or $(1, 2, 3, \ldots)$] to $\boldsymbol{R} \times \boldsymbol{R}$, subject to the restriction that $\lim_{i \to \infty} |q_i| = \infty$ if the index set is not finite, and that $[\mathscr{X}]$ denotes the set of all equivalence classes of such mappings, two mappings being equivalent if they differ only by a permutation of the index set. We will define a topology on $[\mathscr{X}]$ by specifying a class of functions on $[\mathscr{X}]$ and giving $[\mathscr{X}]$ the weakest topology making each function in this class continuous.

Let $\mathscr{K}_1$ denote the set of all continuous real-valued functions f on $\boldsymbol{R} \times \boldsymbol{R}$ whose supports have bounded projections onto the first factor. In other words, a continuous real-valued function f is in $\mathscr{K}_1$ if and only if there is a bounded set $\Lambda \subset \boldsymbol{R}$ such that $f(q, p) = 0$ whenever $q \notin \Lambda$. For f in $\mathscr{K}_1$, we define a function Sf on $[\mathscr{X}]$

$$Sf(x) = \sum_i f(q_i, p_i)$$

if (q_i, p_i) is a representative of x. The sum has only finitely many non-zero terms because of the support properties of f and the local finiteness of x; moreover, $Sf(x)$ evidently does not depend on the choice of the representative (q_i, p_i) of x. We give $[\mathscr{X}]$ the weakest topology making Sf continuous for every f in $\mathscr{K}_1$.

We can give another description of this topology. For any x in $[\mathscr{X}]$, let (q_i, p_i) be a representative of x, and define a measure μ_x on $\boldsymbol{R} \times \boldsymbol{R}$ by

$$\mu_x = \sum_i \delta_{q_i} \otimes \delta_{p_i} .$$

In other words, μ_x is the measure which assigns, to every subset of $\boldsymbol{R} \times \boldsymbol{R}$, the number of particles whose position and momentum lie in that set. It is easy to see that μ_x determines x, so $[\mathscr{X}]$ may be thought of as a set of measures on $\boldsymbol{R} \times \boldsymbol{R}$. From the formula

$$Sf(x) = \int f \, d\mu_x$$

it follows that the topology on $[\mathscr{X}]$ is just the weak topology on measures defined by the space $\mathscr{K}_1$ of functions. Using measure theory, one proves the following:

Lemma 2.1. *The set of positive linear functionals on $\mathscr{K}_1$ of the form $f \longmapsto Sf(x)$ is closed in the weak topology in the algebraic dual of $\mathscr{K}_1$.*

By TYCHONOV'S theorem, this implies the following compactness criterion:

Proposition 2.2. *A closed subset X of $[\mathscr{X}]$ is compact if and only if each Sf is bounded on X.*

We want to transform this criterion into one which is more directly applicable. Before doing this, we will define some notation. For any bounded set $\Lambda \subset \boldsymbol{R}$, we define three functions on $[\mathscr{X}]$:

$$N_\Lambda(x) = \# \{i : q_i \in \Lambda\} ,$$

$$K.E._\Lambda(x) = \frac{1}{2} \sum \{p_i^2 : q_i \in \Lambda\} ,$$

$$\overline{P}_\Lambda(x) = 0 \vee \sup \{|p_i| : q_i \in \Lambda\} .$$

Here, as usual, (q_i, p_i) is a representative of x. N_Λ, $K.E._\Lambda$, and $\overline{P}_\Lambda$ are respectively the number of particles in Λ, the total kinetic energy of the particles in Λ, and the maximum velocity of any particle in Λ. The following proposition is easily obtained from Proposition 2.2.

Proposition 2.3. *A closed subset X of $[\mathscr{X}]$ is compact if and only if, for every bounded open set Λ, N_Λ and $\overline{P}_\Lambda$ are bounded on X.*

We also have:

Proposition 2.4. *Let Λ be a bounded open subset of $\boldsymbol{R}$. Then N_Λ, $K.E._\Lambda$, and $\overline{P}_\Lambda$ are lower semi-continuous functions on $[\mathscr{X}]$.*

Proof. Let ψ_n be an increasing sequence of non-negative continuous functions on $\boldsymbol{R} \times \boldsymbol{R}$ converging pointwise to the characteristic function of $\Lambda \times \boldsymbol{R}$. Then $N_\Lambda = \operatorname{Sup}_n S\psi_n$; since each $S\psi_n$ is continuous by definition, N_Λ is lower semi-continuous. A similar argument shows that $K.E._\Lambda$ and that

$$\overline{P}_\Lambda^n = [\sum \{|p_i|^n : q_i \in \Lambda\}]^{1/n}$$

are lower semi-continuous. Every point of $[\mathscr{X}]$ has a neighborhood on which N_Λ is bounded; on such a neighborhood, $\overline{P}_\Lambda^n$ converges uniformly to $\overline{P}_\Lambda$ as n goes to infinity. Hence, $\overline{P}_\Lambda$ is lower semi-continuous.

The following proposition is proved by showing that the topology of $[\mathscr{X}]$ may be defined by a suitably chosen countable subset of the functions of the form Sf, f in $\mathscr{K}_1$.

Proposition 2.5. *The space $[\mathscr{X}]$ is a Polish space, i.e. it is separable and its topology is compatible with a metric with respect to which it is complete.*

2.2. Borel Measures on $[\mathscr{X}]$

For any bounded non-empty set $\Lambda \subset \boldsymbol{R}$, we let $[\mathscr{X}](\Lambda)$ denote the set of configurations of finitely many unlabelled particles in Λ:

$$[\mathscr{X}](\Lambda) = \coprod_{n=0}^{\infty} (\Lambda \times \boldsymbol{R})^n_{\text{symm}}$$

where $\coprod$ denotes disjoint union and $(\Lambda \times \boldsymbol{R})^n_{\text{symm}}$ the symmetric product of n copies of $\Lambda \times \boldsymbol{R}$, i.e., the set of all equivalence classes of n-tuples of points in $\Lambda \times \boldsymbol{R}$, two n-tuples being equivalent if they differ only by a permutation of their labels. Since $[\mathscr{X}](\Lambda)$ is a disjoint union of quotients of products of copies of $\Lambda \times \boldsymbol{R}$, it has a natural topology. This topology in itself is not very useful, and we will use it only to define the Borel subsets of $[\mathscr{X}](\Lambda)$.

Given any bounded non-empty subset Λ of $\boldsymbol{R}$, there is a natural mapping from $[\mathscr{X}]$ to $[\mathscr{X}](\Lambda)$ which simply forgets about all particles outside of Λ. We will refer to this mapping as the *restriction* from $\boldsymbol{R}$ to Λ; it is a Borel mapping[2] if Λ is a Borel set[3].

Let $[a, b)$ be a non-trivial bounded interval in $\boldsymbol{R}$; then we can decompose $\boldsymbol{R}$ into a countable union of disjoint translates of $[a, b)$. It is easy to see that this decomposition gives a bijective mapping from $[\mathscr{X}]$ to $\coprod_{n=-\infty}^{\infty} [\mathscr{X}]([a_n, b_n))$ $(a_n = a + n(b-a), b_n = b + n(b-a))$ and that this mapping is in fact a Borel isomorphism if $\coprod_{n=-\infty}^{\infty} [\mathscr{X}]([a_n, b_n))$ is given the product topology. Thus, we have a fairly simple description of the Borel structure on $[\mathscr{X}]$.

[2] A mapping from one topological space to another is *Borel* if the inverse image of every Borel set is Borel.

[3] The restriction mapping is not continuous: A continuous trajectory in $[\mathscr{X}]$ can have a varying number of particles in Λ (since particles can move in and out of Λ), whereas the number of particles is constant on continuous trajectories in $[\mathscr{X}](\Lambda)$. This is why the topology we have defined on $[\mathscr{X}](\Lambda)$ is not very useful. If Λ is open, a better topology can be defined on $[\mathscr{X}](\Lambda)$ by imitating the definition of the topology on $[\mathscr{X}]$; this topology makes the restriction mapping continuous and has the same Borel sets as the topology we are using.

The representation of $[\mathscr{X}]$ as a product space gives a useful technique for constructing measures on $[\mathscr{X}]$. Let μ be a probability measure on $[\mathscr{X}]([a, b))$. For each n, translation by $n(b-a)$ gives a Borel isomorphism of $[\mathscr{X}]([a_n, b_n))$ with $[\mathscr{X}]([a, b))$, and we get therefore a probability measure on each $[\mathscr{X}]([a_n, b_n))$. Taking the product of all these measures gives a probability measure on $[\mathscr{X}]$. We will refer to this procedure for passing from a measure on $[\mathscr{X}]([a, b))$ to a measure on $[\mathscr{X}]$ as the *product measure construction*. It gives a measure which is periodic under translations, with period $b-a$.

We can apply this construction in particular to thermodynamic ensembles. Thus, let a stable two-body potential Φ, an inverse temperature β, and a chemical potential μ be given. For any interval $[a, b)$, we define a measure on $([a, b) \times \boldsymbol{R})^n$ as:

$$\frac{1}{n!} \exp\left\{\beta\left[\mu - \frac{1}{2}\sum_i p_i^2 - \sum_{i<j} \Phi(q_i - q_j)\right]\right\} dq_1, \ldots, dq_n\, dp_1, \ldots, dp_n$$

(where $dq_i\, dp_i$ is Lebesgue measure on $[a, b) \times \boldsymbol{R}$). Because of the stability of the potential Φ, this collection of measures defines a finite measure on $[\mathscr{X}]([a, b))$, and by normalizing we get a probability measure, which we will call the grand canonical ensemble on $[a, b)$. Applying the product measure construction to this measure gives a probability measure on $[\mathscr{X}]$ which, physically, corresponds to the grand canonical ensemble for the infinite system with insulating walls at the points a_n.

Let $\mathscr{M}^1[\mathscr{X}]$ denote the set of Borel probability measures on $[\mathscr{X}]$. We will introduce two topologies on $\mathscr{M}^1[\mathscr{X}]$, each of which is a weak topology defined by regarding $\mathscr{M}^1[\mathscr{X}]$ as a subset of the dual of a space of bounded measurable functions on $[\mathscr{X}]$. Thus, let $\mathfrak{A}$ be the C^* algebra of functions on $[\mathscr{X}]$ generated by the set of all functions of the form $\varphi(Sf_1, \ldots, Sf_k)$, where $f_1, \ldots, f_k$ belong to $\mathscr{K}_1$ and φ is a bounded continuous function on $\boldsymbol{R}^k$. The C^* algebra $\mathfrak{A}$ defines a topology on $\mathscr{M}^1[\mathscr{X}]$ which we will refer to as the $\mathfrak{A}$ topology. When we speak of convergence in $\mathscr{M}^1[\mathscr{X}]$ without specifying a topology, we will always mean convergence with respect to the $\mathfrak{A}$ topology. It is sometimes useful to consider the topology defined by the C^* algebra $\mathfrak{A}_\infty$ generated by all functions obtained by composing a bounded Borel function on $[\mathscr{X}](\Lambda)$ (Λ some bounded Borel set) with the restriction mapping from $[\mathscr{X}]$ to $[\mathscr{X}](\Lambda)$. We will refer to this topology as the $\mathfrak{A}_\infty$ topology; it is evidently strictly stronger than the $\mathfrak{A}$ topology.

We may regard $[\mathscr{X}]$ as a subset of the spectrum of $\mathfrak{A}$, and it may be seen that the $\mathfrak{A}$ topology on $[\mathscr{X}]$ coincides with the initial topology. Since $\mathfrak{A}$ is defined as an algebra of functions on $[\mathscr{X}]$, it is clear that $[\mathscr{X}]$ is dense in the spectrum of $\mathfrak{A}$. The following proposition further clarifies the way $[\mathscr{X}]$ lies in the spectrum of $\mathfrak{A}$.

Proposition 2.6. *There exists a family* $(\varphi_{m,n})$ *of elements of* $\mathfrak{A}$ $0 \leqq \varphi_{m,n} \leqq \varphi_{m+1,n} \leqq 1$, *such that, if the* $\varphi_{m,n}$ *are regarded as functions on the spectrum of* $\mathfrak{A}$, *then the characteristic function of* $[\mathscr{X}]$ *is*

$$\inf_n \sup_m \varphi_{m,n} .$$

In particular $[\mathscr{X}]$ *is a Baire set*[4] *in the spectrum of* $\mathfrak{A}$.

This proposition is due to RUELLE ([2], Proposition 4.2 and Corollary 4.4). The proof in this reference is inseparable from other and more complicated considerations; for the convenience of the reader we will give here a direct proof. The idea of the proof is simple: A point of the spectrum of $\mathfrak{A}$ which does not belong to $[\mathscr{X}]$ heuristically represents a situation in which some bounded interval, which we can take to be of the form $(-n, n)$, contains either infinitely many particles or a particle with infinite velocity; we will therefore construct $\varphi_{m,n}$ so that $\lim_{m\to\infty} \varphi_{m,n} = 1$ on $[\mathscr{X}]$ but such that $\varphi_{m,n} = 0$ for all m if there are infinitely many particles, or a particle with infinite velocity, in $(-n, n)$.

Let χ be a continuous non-increasing function on $\boldsymbol{R}$ such that $\chi(t) = 0$ for $t \geqq 1$ and $\chi(t) = 1$ for $t \leqq 0$. Let ψ_n be a continuous non-negative function on $\boldsymbol{R}$ which has compact support and which is equal to one on $(-n, n)$. We will prove that we may take:

$$\varphi_{m,n}(x) = \chi(Sf_n(x) - m) ,$$

with $f_n(q, p) = \psi_n(q)\,(1 + p^2)$. Since f_n is in $\mathscr{K}_1$ and χ is bounded and continuous, it follows from the definition that $\varphi_{m,n}$ belongs to $\mathfrak{A}$. Since $0 \leqq \chi \leqq 1$, and since χ is non-decreasing, we have

$$0 \leqq \varphi_{m,n} \leqq \varphi_{m+1,n} \leqq 1 .$$

Clearly, $\lim_{m\to\infty} \varphi_{m,n}(x) = 1$ for all x in $[\mathscr{X}]$.

It remains to be shown that the $\varphi_{m,n}$'s have the desired property of separating $[\mathscr{X}]$ from the rest of the spectrum of $\mathfrak{A}$. Thus, let x be a point of the spectrum of $\mathfrak{A}$ which does not belong to $[\mathscr{X}]$, and let x_α be a net in $[\mathscr{X}]$ converging to x. We claim that, for some n, $\limsup_\alpha Sf_n(x_\alpha) = \infty$.

If this were not the case, it would follow from Proposition 2.3 that the net (x_α) has a cluster point in $[\mathscr{X}]$, and this would contradict the assumption that $\lim_\alpha x_\alpha \notin [\mathscr{X}]$. Thus, for that value of n, we must have

$$\varphi_{m,n}(x) = \lim_\alpha \varphi_{m,n}(x_\alpha) = 0$$

[4] On any topological space, we define the set of Baire functions to be the smallest set of functions containing the continuous functions and closed under pointwise limits, and the Baire sets to be those sets whose characteristic functions are Baire functions. Every Baire set is also a Borel set; the converse is true if the topological space in question is metrizable.

for all m. Hence, $\inf_n \sup_m \varphi_{m,n} = 0$ on $\mathbb{C}[\mathscr{X}]$, whereas we have seen that $\inf_n \sup_m \varphi_{m,n} = 1$ on $[\mathscr{X}]$.

Let $E(\mathfrak{A})$ denote the set of states of $\mathfrak{A}$; $E(\mathfrak{A})$ may be identified with the set of regular Borel probability measures on the spectrum of $\mathfrak{A}$. Any Borel probability measure on $[\mathscr{X}]$ may be regarded as a Borel probability measure on the spectrum of $\mathfrak{A}$ which assigns measure zero to the complement of $[\mathscr{X}]$. Moreover, such a measure is automatically regular. To see this we remark first that, since any finite Borel measure on a compact space is regular on the Baire sets [4], it suffices to show that any Borel set in $[\mathscr{X}]$ is a Baire set in the spectrum of $\mathfrak{A}$. This last assertion follows from the fact that $[\mathscr{X}]$ is a Baire set in the spectrum of $\mathfrak{A}$ and the fact, easily verified, that there is a countable set in $\mathfrak{A}$ which generates the topology of $[\mathscr{X}]$. We may therefore, when convenient, regard $\mathscr{M}^1[\mathscr{X}]$ as a subset of $E(\mathfrak{A})$. It is easy to construct integrals of functions with values in $E(\mathfrak{A})$; we will need some technical results which assure us that, if we consider such a function whose values actually lie in $\mathscr{M}^1[\mathscr{X}]$, then the integral is also a measure on $[\mathscr{X}]$.

Proposition 2.7. *Let (X, ν) be a probability measure space and $x \longmapsto \varrho_x$ a mapping from X to $E(\mathfrak{A})$ such that $x \longmapsto \varrho_x(\psi)$ is measurable for every ψ in $\mathfrak{A}$. Define a state ϱ of $\mathfrak{A}$ by*

$$\varrho(\psi) = \int d\nu(x)\, \varrho_x(\psi)\,. \tag{2.1}$$

Then, for all bounded Baire functions f on the spectrum of $\mathfrak{A}$, $x \longmapsto \int f d\varrho_x$ is measurable and

$$\int f\, d\varrho = \int d\nu(x) \int f\, d\varrho_x\,.$$

Proof. Let $\mathscr{F}$ denote the class of all bounded Baire functions on the spectrum of $\mathfrak{A}$ for which the proposition holds. Since the continuous functions on the spectrum of $\mathfrak{A}$ are just the elements of $\mathfrak{A}$, $\mathscr{F}$ contains the continuous functions by the definition of ϱ. On the other hand, if (f_n) is a uniformly bounded sequence of functions in $\mathscr{F}$ converging pointwise to f, then a double application of the dominated convergence theorem shows that f is in $\mathscr{F}$. Hence, $\mathscr{F}$ contains all bounded Baire functions.

Corollary 2.8. *Let the notation be as in Proposition 2.7. Then ϱ belongs to $\mathscr{M}^1[\mathscr{X}]$ if and only if ϱ_x belongs to $\mathscr{M}^1[\mathscr{X}]$ for almost all x.*

Proof. By Proposition 2.6, the characteristic function $\chi_{[\mathscr{X}]}$ of $[\mathscr{X}]$ is a Baire function. The assertion therefore follows from (2.1) and the remark that ϱ belongs to $\mathscr{M}^1[\mathscr{X}]$ if and only if

$$\int d\varrho\,[1 - \chi_{[\mathscr{X}]}] = 0\,.$$

Corollary 2.8a. *Let the notation be as in Proposition 2.7, but assume that ϱ_x is in $\mathscr{M}^1[\mathscr{X}]$ for almost all x. Let f be a bounded or non-negative*

Borel function on $[\mathscr{X}]$. *Then*

$$\int f\, d\varrho = \int d\nu(x) \int f\, d\varrho_x \,. \tag{2.2}$$

Proof. The assertion for non-negative functions follows from the assertion for bounded functions by the monotone convergence theorem. The assertion for bounded functions follows from the corresponding assertion for characteristic functions of sets, by a standard approximation argument. The assertion for characteristics functions of sets follows from Proposition 2.7 and the fact that every Borel set in $[\mathscr{X}]$ is a Baire set in the spectrum of $\mathfrak{A}$.

If ϱ belongs to $\mathscr{M}^1[\mathscr{X}]$, we define a translated measure $\tau_s\varrho$ by

$$(\tau_s\varrho)(\varphi) = \varrho(\varphi \circ \tau_s)$$

for φ in $\mathfrak{A}$. It is easy to see that, for any φ in $\mathfrak{A}$ and any x in $[\mathscr{X}]$, $s \longmapsto \varphi(\tau_s x)$ is continuous. Hence, by the dominated convergence theorem, $s \longmapsto \tau_s\varrho$ is continuous in the $\mathfrak{A}$ topology. We may therefore construct $1/a \int_0^a ds\, \tau_s\varrho$, which belongs to $\mathscr{M}^1[\mathscr{X}]$. If we temporarily denote this measure by $\bar{\varrho}$, then by Corollary 2.9 we have

$$\int f\, d\bar{\varrho} = \frac{1}{a} \int_0^a ds \left[\int d\varrho\, (f \circ \tau_s) \right] \tag{2.3}$$

for any semi-bounded Borel function f on $[\mathscr{X}]$. If ϱ is periodic with period a, then $\bar{\varrho}$ is translation invariant. We will refer to this measure as the *average of* ϱ *over translations*. We can apply this construction in particular when ϱ is obtained by the product measure construction from a measure ϱ_1 on $[\mathscr{X}]([a, b))$; we will refer to the operation of passing from ϱ_1 to this invariant measure as the *averaged product measure construction*.

2.3. Entropy

We will need the notion of the total entropy per unit volume of a measure on $[\mathscr{X}]$ which is periodic under translations. In our discussion we will follow, roughly, the work of Robinson and Ruelle [3]. The infinite volume of momentum space, however, introduces some complications not present in the theory of the configurational entropy.

We will start by defining, abstractly, the entropy of a probability measure with respect to an arbitrary σ-finite measure. Let X be a set and $\mathscr{S}$ a σ-algebra of subsets of X. We will consider the σ-algebra $\mathscr{S}$ to be fixed and suppress it from our notation, i.e., we will speak of measures on X rather than of measures defined on $\mathscr{S}$ and of measurable functions on X rather than of functions measurable with respect to $\mathscr{S}$. Let σ be a σ-finite measure on X. If ϱ is a probability measure on X, we want to

define the entropy of ϱ with respect to σ as $-\infty$ if ϱ is not absolutely continuous with respect to σ and as

$$-\int \left(\frac{d\varrho}{d\sigma}\right) \log \left(\frac{d\varrho}{d\sigma}\right) d\sigma$$

if ϱ is absolutely continuous with respect to σ. Unfortunately, this integral need not make any sense; the positive and negative parts of the integrand may both have infinite integrals. If σ is a finite measure, this difficulty does not arise because $x \log x$ is bounded below. If σ is not a finite measure, we must restrict the class of probability measures ϱ that we consider. This we do by choosing a non-negative measurable function ϕ which is rapidly increasing in the sense that $\int e^{-\phi}\, d\sigma < \infty$, and considering only ϱ's such that $\int \phi\, d\varrho < \infty$.

Lemma 2.9. *Let X, σ, ϕ be as above; let f be a non-negative measurable function on X such that $\int f\, d\sigma = 1$ and $\int f \phi\, d\sigma < \infty$. Then the positive part of $-f \log f$ has finite σ-integral, and*

$$-\int f \log f\, d\sigma \leqq \int f.\phi\, d\sigma + \log\left(\int e^{-\phi}\, d\sigma\right). \tag{2.4}$$

Proof. The first statement of the lemma follows at once from the identity

$$-f \log f = f.\phi - e^{-\phi}\,(f/e^{-\phi}) \log (f/e^{-\phi}) \tag{2.5}$$

and the integrability of $f.\phi$ and $e^{-\phi}$. The inequality (2.4) is proved by integrating this identity and using the concavity of $-x \log x$.

We now make the following definition: Let ϱ be a probability measure on X such that $\int \phi\, d\varrho < \infty$. Then we define $s(\varrho, \sigma)$, the *entropy of ϱ relative to σ*, by

$$s(\varrho, \sigma) = -\int \left(\frac{d\varrho}{d\sigma}\right) \log \left(\frac{d\varrho}{d\sigma}\right) d\sigma$$

if ϱ is absolutely continuous with respect to σ,

$$= -\infty \text{ otherwise.}$$

(We could have given a more general definition by defining the entropy for any probability measure ϱ for which there exists a ϕ with the desired properties. For our purposes, it will be convenient to work with a fixed ϕ.)

We will denote by $\mathscr{M}^1(X)$ the set of probability measures on X. There is an obvious pairing between $\mathscr{M}^1(X)$ and the space of all bounded measurable functions on X; we will refer to the weak topology induced on the set of measures by this pairing as the $\mathscr{L}^\infty$ topology. Any statements implying a topology on $\mathscr{M}^1(X)$ are to be understood in the $\mathscr{L}^\infty$ topology. If we identify probability measures which are absolutely continuous with respect to σ with elements of $L^1(\sigma)$, then the $\mathscr{L}^\infty$ topology corresponds to the weak topology on $L^1(\sigma)$.

Proposition 2.10. *Let σ be a finite measure, and let N be a real number. Then $s(\varrho, \sigma)$ is an upper semi-continuous function of ϱ and*

$$\{\varrho : s(\varrho, \sigma) \geqq -N\}$$

is compact.

Proof. Using the concavity of $-x \log x$, one checks easily that

$$s(\varrho, \sigma) = \inf \left\{ -\sum_i \varrho(A_i) \log \frac{\varrho(A_i)}{\sigma(A_i)} \right\}$$

where the infimum is to be taken over all partitions of X into a finite number of disjoint measurable sets $\{A_1, \ldots, A_n\}$ each of which has strictly positive σ-measure. Since $\varrho \longmapsto \varrho(A_i)$ is continuous, $s(\varrho, \sigma)$ is the infimum of a collection of continuous functions and is therefore upper semi-continuous. In particular, $\{\varrho : s(\varrho, \sigma) \geqq -N\}$ is closed in $\mathscr{M}^1(X)$.

Let $\mathscr{P}_1(\sigma)$ be the set of non-negative elements of $L^1(\sigma)$ with integral one; to complete the proof of the proposition it will suffice to prove that

$$\mathscr{K} = \{f \in \mathscr{P}_1(\sigma) : \int f \log f \, d\sigma \leqq N\}$$

is relatively compact for the weak topology on $L^1(\sigma)$.

To prove this, it is enough to show that

$$\lim_{\sigma(E) \to 0} \int_E f \, d\sigma = 0$$

uniformly for f in $\mathscr{K}$ [5].

Let λ be a real number greater than one. Then

$$\int_E f \, d\sigma = \int_{E \cap \{f \leqq \lambda\}} f \, d\sigma + \int_{E \cap \{f > \lambda\}} f \, d\sigma \leqq \lambda \, \sigma(E) + \int_{\{f > \lambda\}} f \, d\sigma \, .$$

We therefore want to show that

$$\lim_{\lambda \to \infty} \int_{\{f > \lambda\}} f \, d\sigma = 0$$

uniformly for f in $\mathscr{K}$. But

$$N \geqq \int f \log f \, d\sigma \geqq \log(\lambda) \int_{\{f > \lambda\}} f \, d\sigma - \frac{1}{e} \int_X d\sigma \, ,$$

or

$$N + \frac{\sigma(X)}{e} \geqq \log(\lambda) \int_{\{f > \lambda\}} f \, d\sigma \, ,$$

which completes the proof of the proposition.

Proposition 2.11. *Let (X, σ) be a measure space, and let ϕ be a non-negative measurable function on X such that $\int e^{-\alpha\phi} \, d\sigma < \infty$ for all $\alpha > 0$. Let M, N be real numbers. Then $\varrho \longmapsto s(\varrho, \sigma)$ is an upper semi-continuous function on $\{\varrho \in \mathscr{M}^1(X) : \int \phi \, d\varrho \leqq M\}$, and $\{\varrho \in \mathscr{M}^1(X) : \int \phi \, d\varrho \leqq M$ and $s(\varrho, \sigma) \geqq -N\}$ is compact.*

Proof. It follows from the hypotheses on ϕ that there exists a non-negative measurable function $\hat{\phi}$ on X such that $\int e^{-\hat{\phi}} \, d\sigma < \infty$ and such

that

$$\lim_{\phi(X)\to\infty} \frac{\hat{\phi}(X)}{\phi(X)} = 0 . \tag{2.6}$$

Let σ' denote the finite measure $e^{-\hat{\phi}}\sigma$. Then we have

$$s(\varrho, \sigma) = s(\varrho, \sigma') + \int \hat{\phi}\, d\varrho$$

by (2.5). Proposition 2.10 asserts that $s(\varrho, \sigma')$ is an upper semi-continuous function on $\mathscr{M}^1(X)$. We will prove the upper semi-continuity of $s(\varrho, \sigma)$ by proving that $\varrho \longmapsto \int \hat{\phi}\, d\varrho$ is continuous on

$$\{\varrho \in \mathscr{M}^1(X) : \int \phi\, d\varrho \leqq M\} .$$

Let $\hat{\phi}_n = \hat{\phi} \wedge n$; then $\hat{\phi}_n$ is a bounded measurable function on X, so $\varrho \longmapsto \int \hat{\phi}_n\, d\varrho$ is continuous. On the other hand,

$$\left| \int \hat{\phi}\, d\varrho - \int \hat{\phi}_n\, d\varrho \right| \leqq \int\limits_{\{\phi \geqq n\}} \hat{\phi}\, d\varrho \leqq \sup_{\hat{\phi}(X) \geqq n} \left\{ \frac{\hat{\phi}(X)}{\phi(X)} \right\} \cdot \int \phi\, d\varrho ,$$

and, by (2.6), the right-hand side goes to zero as n goes to infinity uniformly for ϱ in $\{\varrho : \int \phi\, d\varrho \leqq M\}$. Hence, $\varrho \longmapsto \int \hat{\phi}\, d\varrho$ is a uniform limit of continuous functions and is therefore continuous.

It remains to prove the compactness of

$$\{\varrho \in \mathscr{M}^1(X) : \int \phi\, d\varrho \leqq M, s(\varrho, \sigma) \geqq -N\} .$$

Since $\int \phi\, d\varrho = \sup_n \int (\phi \wedge n)\, d\sigma$, $\varrho \longmapsto \int \phi\, d\varrho$ is lower semi-continuous, so $\{\varrho \in \mathscr{M}^1(X) : \int \phi\, d\varrho \leqq M, s(\varrho, \sigma) \geqq -N\}$ is closed. Let $\sigma'' = e^{-\phi}\sigma$; then

$$s(\varrho, \sigma) = s(\varrho, \sigma'') + \int \phi\, d\sigma .$$

It therefore suffices to prove that $\{\varrho \in \mathscr{M}^1(X) : s(\varrho, \sigma'') \geqq -(M + N)\}$ is compact; this follows from Proposition 2.10 since σ'' is a finite measure.

Proposition 2.12. *Let $\varrho_1, \varrho_2 \in \mathscr{M}^1(X)$, and suppose $\int \phi\, d\varrho_1 < \infty$ and $\int \phi\, d\varrho_2 < \infty$. Let $0 \leqq \alpha \leqq 1$. Then*

$$\begin{aligned} \alpha\, s(\varrho_1, \sigma) + (1-\alpha)\, s(\varrho_2, \sigma) &\leqq s(\alpha\, \varrho_1 + (1-\alpha)\, \varrho_2, \sigma) \\ &\leqq \alpha\, s(\varrho_1, \sigma) + (1-\alpha)\, s(\varrho_2, \sigma) + \log 2 . \end{aligned} \tag{2.7}$$

Proof. This follows, just as in [3], from the concavity of $-x \log x$ and the monotonicity of $\log x$.

Proposition 2.13. *Let T be a one-one measurable mapping of X onto itself with a measurable inverse. Suppose that σ is invariant under T, i.e., that $\sigma(T^{-1}E) = \sigma(E)$ for every measurable subset E of X. Let ϱ be a probability measure on X such that*

$$\int \phi\, d\varrho < \infty , \qquad \int \phi\, d(\varrho \circ T^{-1}) < \infty .$$

Then

$$s(\varrho, \sigma) = s(\varrho \circ T^{-1}, \sigma) .$$

Proof. This proposition follows at once from the definitions.

We now apply this general construction to statistical mechanics. For any bounded Borel set Λ, let σ_Λ be the Borel measure on $[\mathscr{X}](\Lambda)$ whose restriction to each $(\Lambda \times \boldsymbol{R})^n_{\text{symm}}$ is given by the measure

$$\frac{1}{n!} dq_1, \ldots, dq_n\, dp_1, \ldots, dp_n$$

on $(\Lambda \times \boldsymbol{R})^n$. If ϱ is a probability measure on $[\mathscr{X}]$, the image of ϱ under the restriction mapping is a probability measure ϱ_Λ on $[\mathscr{X}](\Lambda)$. The role of the function ϕ of the preceeding propositions will be played by the kinetic energy in Λ; note that, for any $\alpha > 0$, we have

$$\int e^{-\alpha K.E._\Lambda}\, d\sigma_\Lambda = e^{\sqrt{\frac{2\pi}{\alpha}}\, V(\Lambda)} \tag{2.8}$$

where $V(\Lambda)$ is the Lebesgue measure of Λ. If $\int K.E._\Lambda\, d\varrho < \infty$, we will define $s_\Lambda(\varrho)$, the *entropy of* ϱ in Λ, to be $s(\varrho_\Lambda, \sigma_\Lambda)$.

Proposition 2.14. *Let ϱ be a probability measure on $[\mathscr{X}]$ such that $\int K.E._\Lambda\, d\varrho < \infty$ for all bounded Borel sets Λ. Then $s_\Lambda(\varrho)$ is a subadditive set function, i.e.,*

$$s_{\Lambda_1 \cup \Lambda_2}(\varrho) \leqq s_{\Lambda_1}(\varrho) + s_{\Lambda_2}(\varrho) \tag{2.9}$$

for all pairs Λ_1, Λ_2 of disjoint bounded Borel sets, and we have

$$s_\Lambda(\varrho) \leqq 3\left(\frac{\pi}{2}\right)^{1/3} V(\Lambda)\left[\frac{1}{V(\Lambda)} \int K.E._\Lambda\, d\varrho\right]^{1/3} \tag{2.10}$$

for all bounded Borel sets Λ.

Proof. The subadditivity is proved in exactly the same way as in Proposition 1 of [3]; the inequality (2.10) is obtained by applying (2.4) of Lemma 2.9 and (2.8) to get

$$s_\Lambda(\varrho) \leqq \alpha \int K.E._\Lambda\, d\varrho + \sqrt{\frac{2\pi}{\alpha}}\, V(\Lambda)$$

for any $\alpha > 0$, then minimizing with respect to α.

Let a be a positive real number, and let $\mathscr{M}^1_a[\mathscr{X}]$ denote the set of probability measures on $[\mathscr{X}]$ which are periodic under space translations with period a. If $\varrho \in \mathscr{M}^1_a[\mathscr{X}]$, and if $\int d\varrho\, K.E._{[0,a)} < \infty$, then $\int d\varrho\, K.E._\Lambda < \infty$ for every bounded Borel set Λ. Thus, we can define $s_\Lambda(\varrho)$ for all such Λ.

Proposition 2.15. *Let $\varrho \in \mathscr{M}^1_a[\mathscr{X}]$, and suppose that $\int K.E._{[0,a)}\, d\varrho < \infty$. Then:*

1. $\lim_{\beta - \alpha \to \infty} \frac{1}{\beta - \alpha} s_{[\alpha,\beta)}(\varrho)$ *exists. We denote this limit by $\bar{s}(\varrho)$; it is the entropy of ϱ per unit volume.*
2. $\bar{s}(\varrho) = \inf_n \frac{1}{na} s_{[0,na)}(\varrho)$.
3. *$\bar{s}(\varrho)$ is an affine function of ϱ.*

4. $\bar{s}(\varrho) \leqq 3\left(\frac{\pi}{2}\right)^{1/3}\left[\frac{1}{a}\int K.E._{\cdot[0,a)}\, d\varrho\right]^{1/3}$.

5. *For any pair of real numbers* M, N,

$$\mathscr{K} = \{\varrho \in \mathscr{M}^1_a[\mathscr{X}] : \int K.E._{\cdot[0,a)}\, d\varrho \leqq M, \bar{s}(\varrho) \geqq -N\}$$

is compact for the $\mathfrak{A}_\infty$ *topology, and the* $\mathfrak{A}$ *topology agrees with the* $\mathfrak{A}_\infty$ *topology on* $\mathscr{K}$.

6. *For any real number* M, $\bar{s}(\varrho)$ *is an upper-semi continuous function for the* $\mathfrak{A}$ *topology on* $\{\varrho \in \mathscr{M}^1_a[\mathscr{X}] : \int K.E._{\cdot[0,a)}\, d\varrho \leqq M\}$.

Proof. By Proposition 2.14, $s_{[0,na)}(\varrho)$ is a sub-additive function of n, so $\lim_{n\to\infty}\frac{1}{na}s_{[0,na)}(\varrho)$ exists and is equal to $\inf_n \frac{1}{na}s_{[0,na)}(\varrho)$. Now let $\beta > \alpha$ be given and let n, n' be chosen so that $na \leqq \alpha < (n+1)a$, $n'a \leqq \beta < (n'+1)a$. Then by subadditivity

$$s_{[\alpha,\beta)}(\varrho) \leqq s_{[(n+1)a,n'a)}(\varrho) + s_{[\alpha,(n+1)a)}(\varrho) + s_{[n'a,\beta)}(\varrho)\,.$$

By (2.10) the right-hand side is not greater than

$$s_{[(n+1)a,n'a)}(\varrho) + 6\left(\frac{\pi}{2}\right)^{1/3} a^{2/3}\left[\int K.E._{\cdot[0,a)}\, d\varrho\right]^{1/3},$$

so

$$\limsup_{\beta-\alpha\to\infty}\frac{1}{\beta-\alpha}s_{[\alpha,\beta)}(\varrho) \leqq \lim_{n\to\infty}\frac{1}{na}s_{[0,na)}(\varrho)\,.$$

Similarly,

$$\liminf_{\beta-\alpha\to\infty}\frac{1}{\beta-\alpha}s_{[\alpha,\beta)}(\varrho) \geqq \lim_{n\to\infty}\frac{1}{na}s_{[0,na)}(\varrho)\,,$$

so statements 1. and 2. are proved.

The fact that $s(\varrho)$ is an affine function of ϱ follows from inequality (2.7) of Proposition 2.12. The bound 4. follows from 2. and inequality (2.10) of Proposition 2.14.

To prove statement 5., let ϱ_α be a universal net[5] in $\mathscr{K}$. Then since we have:

$$s(\varrho_{\alpha,[-na,na)}, \sigma_{[-na,na)}) \geqq -2naN$$
$$\int K.E._{\cdot[-na,na)}\, d\varrho_\alpha \leqq 2nM$$

it follows from Proposition 2.11 that $\varrho_{\alpha,[-na,na)}$ converges in the $\mathscr{L}^\infty$ topology for measures on $[\mathscr{X}]([-na, na))$. Let $\hat{\varrho}_{[-na,na)}$ denote the limiting measure. Then the collection of measures $(\hat{\varrho}_{[-na,na)})$ is consistent and therefore defines a unique measure $\hat{\varrho}$ on $[\mathscr{X}]$; evidently ϱ_α converges to $\hat{\varrho}$ in the $\mathfrak{A}_\infty$ topology. It remains to be checked that $\hat{\varrho}$ belongs to $\mathscr{K}$. It is clear that $\hat{\varrho}$ is in $\mathscr{M}^1_a[\mathscr{X}]$. Since for any $\varrho \in \mathscr{M}^1[\mathscr{X}]$ $\int K.E._{\cdot[0,a)}\, d\varrho = \sup \int [K.E._{\cdot[0,a)} \wedge n]\, d\varrho$, $\varrho \longmapsto \int K.E._{\cdot[0,a)}\, d\varrho$ is a lower semi-contin-

[5] See [6], for the definition of a universal net. A universal net is roughly one which is a refined as possible. Every net has a universal subnet; the image of a universal net under any mapping is universal; a Hausdorff topological space is compact if and only if every universal net in it converges.

uous function on $\mathscr{M}^1[\mathscr{X}]$ with the $\mathfrak{A}_\infty$ topology; hence, $\int K.E._{[0,a)} \cdot d\hat{\varrho} \leqq M$.

By Proposition 2.11, each $s_{[-na,na)}(\varrho)$ is an $\mathfrak{A}_\infty$ upper semi-continuous function on $\{\varrho \in \mathscr{M}_a^1[\mathscr{X}] : \int K.E._{[0,a)}\, d\varrho \leqq M\}$ and hence, by 2., $\bar{s}(\varrho)$ is also $\mathfrak{A}_\infty$ upper semi-continuous on this set. Hence, in particular, $\bar{s}(\hat{\varrho}) \geqq -N$, so $\hat{\varrho}$ belongs to $\mathscr{K}$ and $\mathscr{K}$ is $\mathfrak{A}_\infty$-compact. From the compactness of $\mathscr{K}$ in the $\mathfrak{A}_\infty$ topology, and the fact that the $\mathfrak{A}$ topology is a Hausdorff topology which is no finer than the $\mathfrak{A}_\infty$ topology, it follows that the $\mathfrak{A}$ topology coincides with the $\mathfrak{A}_\infty$ topology on $\mathscr{K}$.

We still have to prove 5., and we know already that $\bar{s}(\varrho)$ is $\mathfrak{A}_\infty$ upper semi-continuous on $\mathscr{K}' = \{\varrho \in \mathscr{M}_a^1[\mathscr{X}] : \int K.E._{[0,a)}\, d\varrho \leqq M\}$. To prove $\mathfrak{A}$ upper semi-continuity, it is enough to prove that, if ϱ_α is a net in $\mathscr{K}'$ which converges to ϱ, then $\bar{s}(\varrho) \geqq \limsup_\alpha \bar{s}(\varrho_\alpha)$. There is evidently nothing to be proved if $\limsup_\alpha \bar{s}(\varrho_\alpha) = -\infty$. If this is not the case, then we can pass to a subnet for which $\bar{s}(\varrho_\alpha)$ is bounded below, without changing the lim sup. In other words, we can assume that all the ϱ_α are contained in

$$\{\varrho \in \mathscr{M}_a^1[\mathscr{X}] : \int K.E._{[0,a)}\, d\varrho \leqq M, \bar{s}(\varrho) \geqq -N\},$$

for some choice of N. But, on this set, the $\mathfrak{A}$ topology coincides with the $\mathfrak{A}_\infty$ topology, so ϱ_α converges to ϱ in the $\mathfrak{A}_\infty$ topology, and

$$\bar{s}(\varrho) \geqq \limsup_\alpha s(\varrho_\alpha)$$

follows from the $\mathfrak{A}_\infty$ upper semi-continuity of $\bar{s}$.

Proposition 2.16. *Let X be a compact topological space, $x \longmapsto \varrho_x$ a continuous mapping from X to $\mathscr{M}_a^1[\mathscr{X}]$, and ν a Radon probability measure on X. Suppose $\int K.E._{[0,a)}\, d\varrho_x$ is bounded with respect to x. Let $\bar{\varrho} = \int d\nu(x)\, \varrho_x$. Then*

$$\int K.E._{[0,a)}\, d\bar{\varrho} = \int d\nu(x) \int K.E._{[0,a)}\, d\varrho_x$$

and

$$\bar{s}(\bar{\varrho}) = \int d\nu(x)\, \bar{s}(\varrho_x).$$

Proof. The first formula follows immediately from Corollary 2.9. Replacing X, ν by their images under $x \longmapsto \varrho_x$ we can suppose that $X \subset \mathscr{M}_a^1[\mathscr{X}]$ and that $\bar{\varrho}$ is the barycenter of ν. But $\bar{s}(\varrho)$ is affine and upper semi-continuous on any set on which $\int K.E._{[0,a)}\, d\varrho$ is bounded; hence, by the theorem of the barycenter[6]

$$\bar{s}(\bar{\varrho}) = \int d\nu(x)\, \bar{s}(\varrho_x).$$

[6] The theorem of the barycenter asserts that, if $\mathscr{K}$ is a compact convex set in a locally convex topological vector space, ν a probability measure on $\mathscr{K}$ with barycenter $r(\nu)$, and f an affine upper semi-continuous function on $\mathscr{K}$, then $f(r(\nu)) = \int f(x)\, d\nu(x)$. See [7].

(To justify the application of the theorem of the barycenter, we have to know that the closed convex hull of the image of X in $\mathscr{M}^1[\mathscr{X}]$ is compact, or equivalently, that the closed convex hull of the image of X in $E(\mathfrak{A})$ is contained in $\mathscr{M}^1[\mathscr{X}]$. This follows easily from Corollary 2.8.)

Corollary 2.17. *Let* $\varrho \in \mathscr{M}_a^1[\mathscr{X}]$, *and assume* $\int K.E._{[0,a)}\, d\varrho < \infty$. *Let* $\bar{\varrho} = \frac{1}{a}\int_0^a ds(\tau_s\varrho)$. *Then*

$$\int K.E._{[0,a)}\, d\bar{\varrho} = \int K.E._{[0,a)}\, d\varrho < \infty, \quad \textit{and} \quad \bar{s}(\bar{\varrho}) = \bar{s}(\varrho)\,.$$

Proof. By Proposition 2.16, we have only to prove

$$\int K.E._{[0,a)} \circ \tau_s\, d\varrho = \int K.E._{[0,a)}\, d\varrho \quad \text{for} \quad 0 \leqq s \leqq a\,.$$

Now $K.E._{[0,a)} \circ \tau_s = K.E._{[-s,s-a)} = K.E._{[-s,0)} + K.E._{[0,a-s)}$. By the periodicity of ϱ,

$$\int K.E._{[-s,0)}\, d\varrho = \int K.E._{[a-s,a)}\, d\varrho\,.$$

Reassembling gives:

$$\int K.E._{[0,a)} \circ \tau_s\, d\varrho = \int [K.E._{[0,a-s)} + K.E._{[a-s,a)}]\, d\varrho = \int K.E._{[0,a)}\, d\varrho$$

2.4. The Existence Theorem

In this section we summarize and reformulate the main results of [1] in a form which will be convenient for our purposes in this article.

For any $x = (q_i, p_i)$ in $\mathscr{X}$, we define

$$|x| = \sup_i \left(\frac{|p_i|}{\log_+(q_i)}\right) \vee \sup \left\{\frac{N_{(\alpha,\beta)}(x)}{\beta-\alpha} : \beta - \alpha > \log_+\left(\frac{\beta+\alpha}{2}\right)\right\},$$

where $\log_+(q) = \log(|q| \vee e)$. The quantity $|x|$ is either a non-negative real number or $+\infty$. We will regard $|\ |$ either as a function on $\mathscr{X}$ or on $[\mathscr{X}]$ as convenient.

The set $\hat{\mathscr{X}}$ is $\{x \in \mathscr{X} : |x| < \infty\}$. For any non-negative real number δ, we define $\hat{\mathscr{X}}_\delta = \{x \in \mathscr{X} : |x| \leqq \delta\}$. We denote by $[\hat{\mathscr{X}}]$ and by $[\hat{\mathscr{X}}_\delta]$ the corresponding sets of equivalence classes.

We want to solve the equations of motion

$$\frac{dq_i(t)}{dt} = p_i(t); \quad \frac{dq_i(t)}{dt} = \sum_{j \neq i} F(q_i(t) - q_j(t))$$

with initial data in $\hat{\mathscr{X}}$. Throughout this article, we will assume that F has compact support and satisfies a Lipschitz condition. We will always use R to denote the range of F, i.e., the smallest number such that $F(q) = 0$ whenever $|q| \geqq R$.

To solve the equations of motion, we introduce for each initial configuration $x = (q_i, p_i)$ the Banach space $\mathscr{Y}_x$ of sequences $\zeta = (\xi_i, \eta_i)$ of pairs of real numbers such that

$$\|\zeta\|_x = \sup_i \frac{|\xi_i| \vee |\eta_i|}{\log_+(q_i)} < \infty\,.$$

For ζ in $\mathscr{Y}_x$, we let $x+\zeta$ denote the configuration $(q_i+\xi_i, p_i+\eta_i)$. The equations of motion with initial configuration x can be reformulated as an evolution equation in $\mathscr{Y}_x$:

$$\frac{d\zeta_x(t)}{dt} = A_x(\zeta_x(t));$$

the solution of the original equations is then obtained as

$$x(t) = x + \zeta_x(t) .$$

We may obtain the solution of the evolution equation by an iterative procedure. Define

$$\zeta_{0,x}(t) \equiv 0$$
$$\zeta_{n,x}(t) \equiv \int_0^t d\tau\, A_x(\zeta_{n-1,x}(\tau)) \quad \text{for} \quad n = 1, 2, 3, \ldots,$$

i. e.,

$$\begin{aligned} \xi_{i,n,x}(t) &= \int_0^t d\tau\, [p_i + \eta_{i,n-1,x}(\tau)] \\ \eta_{i,n,x}(t) &= \int_0^t d\tau \left[\sum_{j \neq i} F(q_i + \xi_{i,n-1,x}(\tau) - q_j - \xi_{j,n-1,x}(\tau)) \right]. \end{aligned} \tag{2.12}$$

Let $x_n(t) = x + \zeta_{n,x}(t)$. For each positive real number m, define a seminorm ${}_m\| \ \|_x$ on $\mathscr{Y}_x$ by

$${}_m\|\zeta\|_x = 0 \vee \sup \left\{ \frac{|\zeta_i| \vee |\eta_i|}{\log_+(q_i)} \ : |q_i| \leqq m \right\}.$$

The following proposition is a more explicit version of Remark 4.3 of [1]:

Proposition 2.18. *There exist functions $h(\delta, T)$ and $\varepsilon(n, m, \delta, T)$ such that*

i) $\|\zeta_{n,x}(t)\|_x \leqq h(\delta, T)$ *for all* n *whenever* $|x| \leqq \delta$ *and* $|t| \leqq T$.

ii) $\lim_{n\to\infty} \varepsilon(n, m, \delta, T) = 0$ *for all* m, δ, T

and

$${}_m\|\zeta_{n,x}(t) - \zeta_x(t)\|_x \leqq \varepsilon(n, m, \delta, T)$$

whenever $|x| \leqq \delta$ *and* $|t| \leqq T$.

We define T^t to be the mapping of $[\mathscr{X}]$ into itself which takes the equivalence class of x to the equivalence class of $x(t)$. The mappings T^t form a one-parameter group of transformations on $[\hat{\mathscr{X}}]$. We let T_n^t denote the mapping of $[\mathscr{X}]$ into itself which takes the equivalence class of x to the equivalence class of $x_n(t)$. From Proposition 2.18, one easily obtains the following:

Proposition 2.19. *For any pair of positive numbers δ, T, and any ψ in $\mathfrak{A}$,*

$$\lim_{n\to\infty} \psi(T_n^t x) = \psi(T^t x)$$

uniformly for x in $[\hat{\mathscr{X}}_\delta]$ and $|t| \leqq T$.

2.5. *Space-Periodized Systems*

We will have occasion to consider systems of a finite number of particles moving in a finite interval $[-a, b)$ "with periodic boundary conditions". From a fundamental point of view, this is a matter of studying a second order differential equation on torus. For our purposes, it is convenient to formulate the equations in a slightly different way.

For any function f defined on $\boldsymbol{R}$, and any positive real number α, we let f_α denote the function on $\boldsymbol{R}$ which is periodic with period α and which agrees with f on $[-\alpha/2, \alpha/2)$. To find the motion of a system of n particles moving on $[-a, b)$ with interparticle force F and with periodic boundary conditions, it is enough to solve the system of ordinary differential equations:

$$\begin{aligned} \frac{dq_i(t)}{dt} &= p_i(t); \\ \frac{dp_i(t)}{dt} &= \sum_{j \neq i} \tilde{F}_{a+b}(q_i(t) - q_j(t)) . \end{aligned} \tag{2.13}$$

(Note that, if $a + b \geqq 2R$, then $\tilde{F}_{a+b}$ satisfies a Lipschitz condition since F does.) The $p_i(t)$'s are the correct velocities, but the $q_i(t)$'s are not necessarily the correct positions; these latter are obtained from the $q_i(t)$'s by subtracting appropriate integral multiples of $a + b$ to give values in $[-a, b)$.

The solution of this system of equations gives a one-parameter group of mappings of $[\mathscr{X}]$ ($[-a, b)$) onto itself; we denote these mappings by $\tilde{T}^t_{[-a,b)}$. If we identify

$$[\mathscr{X}] = \prod_{n=-\infty}^{\infty} [\mathscr{X}]\big([-a + n(a+b), b + n(a+b))\big)$$

and if we consider the separate evolution of each factor, we get a one-parameter group of mappings of $[\mathscr{X}]$ onto itself which we will also denote by $\tilde{T}^t_{[-a,b)}$.

We also need to adapt some ideas from statistical mechanics to the framework of periodized systems. A two-body potential Φ will be said to be *P-stable* if there exist constants B and D such that

$$\sum_{1 \leqq i < j \leqq n} \tilde{\Phi}_d(q_i - q_j) \geqq -Bn \tag{2.14}$$

for all $n, q_1, \ldots, q_n$, whenever $d \geqq D$. Passing to the limit $d \to \infty$ in (2.14) with the q_i held fixed gives

$$\sum_{1 \leqq i < j \leqq n} \Phi(q_i - q_j) \geqq -Bn ,$$

i.e., any P-stable potential is stable.

Although the notion of P-stability is a useful tool, it is not a very pleasing hypothesis from an aesthetic point of view. Fortunately, for potentials of compact support, it reduces to the ordinary notion of stability.

Proposition 2.20.[7] *Any stable potential of compact support is P-stable.*

Proof. Suppose $\Phi(q) = 0$ for $|q| \geqq R$, and suppose that

$$\sum_{1 \leqq i < j \leqq n} \Phi(q_i - q_j) \geqq -Bn$$

for all $n, q_1, \ldots, q_n$. We will show that

$$\sum_{1 \leqq i < j \leqq n} \tilde{\Phi}_d(q_i - q_j) \geqq -Bn$$

for all $n, q_1, \ldots, q_n$, if $d \geqq 2R$. We can assume that $q_1, \ldots, q_n \in [0, d)$. For any positive integer N, define $q_{Kn+i} = Kd + q_i$ for $K = 0, 1, \ldots, N-1$ and $i = 1, 2, \ldots, n$.

Then

$$\sum_{1 \leqq i < j \leqq n} \tilde{\Phi}_d(q_i - q_j) = \lim_{N \to \infty} \frac{1}{N} \sum_{1 \leqq i < j \leqq Nn} \Phi(q_i - q_j)$$

$$\geqq \frac{1}{N}(-BNn) = -Bn\,.$$

If Φ is P-stable and if $a + b$ is large enough, we can construct the grand canonical ensemble on $[-a, b)$ for the potential $\tilde{\Phi}_{a+b}$. We will refer to this measure on $[\mathcal{X}]$ ($[-a, b)$) as *the periodized grand canonical ensemble* on $[-a, b)$. If $F(q) = -\frac{d}{dq}\Phi(q)$, where Φ is an even P-stable potential, and if $a + b$ is larger than $2R$, then the measure on $[\mathcal{X}]$ obtained by the product measure construction from the periodized grand canonical ensemble on $[-a, b)$ is invariant under the one-parameter group $\tilde{T}^t_{[-a,b)}$.

§ 3. Measurability of $[\hat{\mathcal{X}}]$ and T^t

Proposition 3.1. *Each $[\hat{\mathcal{X}}_\delta]$ is a compact subset of $[\mathcal{X}]$.*

Proof. By definition, $[\hat{\mathcal{X}}_\delta]$ is the set of all x in $[\mathcal{X}]$ such that

i) $\overline{P}_{(-q,q)}(x) \leqq \delta \log_+(q)$ for all positive real numbers q.

ii) $N_{(\alpha,\beta)}(x) \leqq \delta(\beta - \alpha)$ for all α, β with $\beta - \alpha > \log_+\left(\frac{\alpha+\beta}{2}\right)$.

By Proposition 2.5, for any q, α, β,

$$\{x : \overline{P}_{(-q,q)}(x) \leqq \delta \log_+(q)\}$$

and

$$\{x : N_{(\alpha,\beta)}(x) \leqq \delta(\beta - \alpha)\}$$

[7] This proposition is due to D. Ruelle (unpublished).

are closed in $[\mathscr{X}]$. Hence, $[\hat{\mathscr{X}}_\delta]$ is the intersection of a collection of closed sets and is therefore closed. On the other hand, for any bounded open set A, $\overline{P}_A$ and N_A are bounded on $[\hat{\mathscr{X}}_\delta]$. Hence, by Proposition 2.3, $[\hat{\mathscr{X}}_\delta]$ is compact.

Corollary 3.2. *$[\hat{\mathscr{X}}]$ is a Borel subset of $[\mathscr{X}]$.*

Next we investigate the measurability of the time-evolution mappings T^t.

Proposition 3.3. *For each δ, the mapping $(t, x) \longmapsto T^t x$ is continuous from $\boldsymbol{R} \times [\hat{\mathscr{X}}_\delta]$ to $[\mathscr{X}]$.*

Proof. It suffices to prove that $(t, x) \longmapsto \psi(T^t x)$ is continuous for every ψ in $\mathfrak{A}$. We know from Proposition 2.19 that

$$\lim_{n\to\infty} \psi(T_n^t x) = \psi(T^t x)$$

and that the convergence is uniform as t runs over any bounded interval and x runs over $[\hat{\mathscr{X}}_\delta]$. Hence, it will be enough to prove that $(t, x) \longmapsto \psi(T_n^t x)$ is continuous. Furthermore, we can suppose that ψ depends only on the co-ordinates of the particles in some bounded interval $[-\lambda, \lambda]$. By Proposition 2.18 there is, for any $T > 0$, a constant H such that

$$|\xi_{i,n,x}(t)| \leqq H \log_+(q_i)$$

whenever $|t| \leqq T$ and $x = (q_i, p_i) \in \hat{\mathscr{X}}_\delta$.

Now choose λ_0 so that $|q| - H \log_+(q) \leqq \lambda$ implies $|q| < \lambda_0$; if $x \in \hat{\mathscr{X}}_\delta$, if $|t| \leqq T$, and if $|q_i + \xi_{i,n,x}(t)| \leqq \lambda$ for some n, then $|q_i| < \lambda_0$. Choose successively $\lambda_1, \lambda_2, \ldots$ so that $|q| \leqq \lambda_j$ and $|q'| \geqq \lambda_{j+1}$ implies

$$|q| + H \log_+(q) + R < |q'| - H \log_+(q') \qquad j = 0, 1, 2, \ldots$$

Then if $x \in \hat{\mathscr{X}}_\delta$ and if $|q_i| \leqq \lambda_k$, $|q_j| \geqq \lambda_{k+1}$, we have

$$|q_{i,n,x}(t) - q_{j,n,x}(t)| > R$$

for all n and all $|t| \leqq T$. [We have introduced the notation $q_{i,n,x}(t) = q_i + \xi_{i,n,x}(t)$.]

From the formula (2.12) for $\xi_{i,n,x}(t)$, $\eta_{i,n,x}(t)$, we see that, if $|t| \leqq T$ and if $|q_i| \leqq \lambda_k$, then $q_{i,n,x}(t)$ and $p_{i,n,x}(t)$ depend only on t, on the values of $p_{i,n-1,x}(\tau)$, τ between 0 and t, and on the values of $q_{j,n-1,x}(\tau)$, τ between 0 and t, for those j's with $|q_j| \leqq \lambda_{k+1}$. By induction, and using the fact that $q_{i,0,x}(t) = q_i$, $p_{i,0,x}(t) = p_i$, we see that, if $|q_i| \leqq \lambda_0$, and if $|t| \leqq T$, then $q_{i,n,x}(t)$ and $p_{i,n,x}(t)$ depend only on t and on those q_j's and p_j's with $|q_j| \leqq \lambda_n$. Furthermore, from the continuity of F we see that the $(q_{i,n,x}(t), p_{i,n,x}(t))$, and hence $\psi(x_n(t))$, are continuous functions of these variables. (Here, we regard ψ as a function on $\mathscr{X}$ rather than on $[\mathscr{X}]$).

It is now an elementary exercise in the topology of $[\mathscr{X}]$ to conclude from these statements that $\psi(T_n^t x)$ varies continuously with t and x for $|t| \leqq T$ and $x \in [\hat{\mathscr{X}}_\delta]$.

Corollary 3.4. *The mapping $(t, x) \longmapsto T^t x$ is a Borel mapping from $\boldsymbol{R} \times [\hat{\mathscr{X}}]$ to $[\hat{\mathscr{X}}]$.*

§ 4. Measures Concentrated on $[\hat{\mathscr{X}}]$

If ϱ is a probability measure on $[\mathscr{X}]$, we let ϱ_Λ denote the probability measure on $[\mathscr{X}]$ (Λ) which is the image of ϱ under the restriction mapping. Specifying a probability measure on $[\mathscr{X}]$ (Λ) is equivalent to specifying, for each n, a finite permutation-invariant measure on $(\Lambda \times \boldsymbol{R})^n$ such that the sum of the total masses of these measures is one. We say that ϱ has a Maxwellian velocity distribution with inverse temperature β if, for each Λ, the component of ϱ_Λ on $(\Lambda \times \boldsymbol{R})^n$ has the form

$$d\hat{\varrho}_\Lambda^n(q_1, \ldots, q_n) \exp\left\{-\frac{\beta}{2} \sum_i p_i^2\right\} dp_1, \ldots, dp_n$$

where $\hat{\varrho}_\Lambda^n$ is a permutation-invariant measure on Λ^n.

Proposition 4.1. *Let $\varrho \in \mathscr{M}^1[\mathscr{X}]$. For ϱ to be concentrated on $[\hat{\mathscr{X}}]$ it is sufficient (but not necessary) that the following two conditions both hold:*

A. *ϱ has a Maxwellian velocity distribution (with some inverse temperature β).*

B. *There exists a real number λ such that, for any interval $[a, b)$ of length at least one, and any $n = 0, 1, 2, \ldots$,*

$$\int d\varrho\, N_{[a,b)}(N_{[a,b)} - 1) \ldots (N_{[a,b)} - n) \leqq [\lambda(b - a)]^{n+1}. \quad (4.1)$$

Moreover, we can put the estimates in a more quantitative form: There exists a function $\varepsilon(\delta, \beta, \lambda)$ such that $\lim_{\delta \to \infty} \varepsilon(\delta, \beta, \lambda) = 0$ for all (β, λ) and such that

$$\varrho(\complement[\hat{\mathscr{X}}_\delta]) \leqq \varepsilon(\delta, \beta, \lambda)$$

whenever ϱ satisfies A. and B.

Proof. Let P be a real number, and let

$Y_{1,P} = \{x \in [\mathscr{X}]$: for some integer m there is a particle with

$$m \leqq q_i < m + 1 \text{ and } |p_i| \geqq P \log_+(m)\};$$

$Y_{2,P} = \{x \in [\mathscr{X}]$: for some integers m, j with $2j \geqq \log_+(m)$, the interval

$$[m - j, m + j) \text{ contains more than } 2Pj \text{ particles}\}.$$

We will estimate $\varrho(Y_{1,P})$ and $\varrho(Y_{2,P})$ and show that they both go to zero as P goes to infinity; this will prove the proposition.

Let $\varrho_{m,n}$ be the ϱ measure of the set of configurations with precisely n particles in $[m, m+1)$, and let

$$\phi(\xi) = \sqrt{\frac{2}{\pi}} \int_{\xi}^{\infty} e^{-p^2/2}\, dp\,.$$

An elementary calculation shows that the ϱ measure of the set of configurations having at least one particle in $[m, m+1)$ with velocity at least $P \log_+(m)$ is

$$\begin{aligned}
&\sum_{n=0}^{\infty} \varrho_{m,n}\left[1 - \left(1 - \phi\left(\sqrt{\beta}\, P \log_+(m)\right)\right)^n\right] \\
&\quad \leqq \phi\left(\sqrt{\beta}\, P \log_+(m)\right) \sum_{n=0}^{\infty} \varrho_{m,n} \cdot n \\
&\quad = \phi\left(\sqrt{\beta}\, P \log_+(m)\right) \int d\varrho\, N_{[m,m+1)} \\
&\quad \leqq \phi\left(\sqrt{\beta}\, P \log_+(m)\right) \cdot \lambda\,.
\end{aligned}$$

Hence,

$$\begin{aligned}
\varrho(Y_{1,P}) &\leqq \sum_{m=-\infty}^{\infty} \varrho\{x : \text{There is a particle in } [m, m+1) \text{ with velocity} \\
&\qquad \text{at least } P \log_+(m)\} \\
&\leqq \lambda \sum_{m} \phi\left(\sqrt{\beta}\, P \log_+(m)\right).
\end{aligned}$$

Using the fact that $\phi(\xi)$ decreases more rapidly at infinity than $e^{-\xi^2/2}$ it is easy to verify that the right-hand side of this inequality goes to zero as P goes to infinity.

To estimate the measure of $Y_{2,P}$, we let $\varrho_{m,j,n}$ denote the ϱ measure of the set of configurations having exactly n particles in the interval $[m-j, m+j)$, and we let

$$\begin{aligned}
\sigma_{m,j,k} &= \int d\varrho\, N_{[m-j,m+j)}\,,\ (N_{[m-j,m+j)} - 1) \cdots (N_{[m-j,m+j)} - k + 1) \\
&= \sum_{n=k}^{\infty} \frac{n!}{(n-k)!}\, \varrho_{m,j,n}\,.
\end{aligned}$$

By condition B.,

$$\sigma_{m,j,k} \leqq (2\lambda j)^k\,.$$

It is straightforward to verify, using this inequality, that[8]

$$\varrho_{m,j,n} = \frac{1}{n!} \sum_{l=0}^{\infty} \frac{(-1)^l}{l!}\, \sigma_{m,j,l+n} \leqq \frac{1}{n!}\, e^{2\lambda j} (2\lambda j)^n\,.$$

[8] The possibility of using such an identity to estimate particle number probabilities was suggested to me by D. RUELLE.

Thus, the measure of the set of configurations with more than $2Pj$ particles in $[m-j, m+j)$ is not greater than

$$e^{2\lambda j} \sum_{n>2Pj} \frac{1}{n!} (2\lambda j)^n .$$

If $P \geqq 2\lambda$, the ratio of succeeding terms in this sum is not greater than $1/2$, so the sum is not greater than twice its first term. Using STIRLING'S formula, we see the sum is majorized by $C \cdot \left(\frac{e^{P+\lambda} \cdot \lambda^P}{P^P}\right)$. Letting $f(P, \lambda) = \left(\frac{e^{P+\lambda}\, \lambda^P}{P^P}\right)$, we have

$$\varrho(Y_{2,P}) \leqq \sum_{m=-\infty}^{\infty} \sum_{2j \geqq \log_+(m)} C[f(P, \lambda)]^{2j}$$

Since $\lim_{P\to\infty} f(P, \lambda) = 0$, the right-hand side goes to zero as P goes to infinity, so the proof of the proposition is complete.

Condition B. holds, in particular, if ϱ has correlation functions of all orders $\bar{\varrho}_n(q_1, \ldots, q_n)$ and if there exists a constant λ such that

$$\bar{\varrho}_n(q_1, \ldots, q_n) \leqq \lambda^n \tag{4.2}$$

for all n and all $q_1, \ldots, q_n$. In fact, we have:

$$\int d\varrho\, N_{(a,b]} (N_{(a,b]} - 1) \ldots (N_{(a,b]} - n + 1)$$
$$= \int_{[a,b)^n} dq_1, \ldots, dq_n\, \bar{\varrho}_n(q_1, \ldots, q_n) .$$

There are two cases in which inequalities of the form (4.2) are known to hold:

i) ϱ is a state obtained by taking the infinite-volume limit of the grand-canonical ensemble at low activity[9] [8].

ii) ϱ is a state obtained by taking any infinite volume cluster point of the grand-canonical ensemble for a system with a non-negative potential, at any value of the temperature and chemical potential.

In both these cases, ϱ also has a Maxwellian velocity distribution and is therefore concentrated on $[\hat{\mathscr{X}}]$.

Case ii) requires some explanation. Suppose we have fixed a temperature and a chemical potential, and suppose Φ is a non-negative two-body potential. For any positive number m, let $\bar{\varrho}_m$ be the measure on $[\mathscr{X}]$ obtained from the grand canonical ensemble on $[-m, m)$ by the averaged product measure construction, and let $\bar{\bar{\varrho}}_m$ be the corresponding measure obtained from the periodized grand canonical ensemble. The measures $\bar{\varrho}_m$ and $\bar{\bar{\varrho}}_m$ are translation invariant, and it may be seen that they both have correlation functions satisfying (4.2) with λ equal to the activity $\mathfrak{z}$. Furthermore, $\int K.E._{[0,1)}\, d\bar{\varrho}_m$ and $-s(\bar{\varrho}_m)$, and the corre-

[9] The *activity* $\mathfrak{z}$ corresponding to the inverse temperature β and chemical potential μ is $e^{\beta\mu}\sqrt{2\pi/\beta}$.

sponding quantities for $\bar{\varrho}_m$, are bounded if m stays away from zero. Hence, by statement 5. of Proposition 2.15, $\{\bar{\varrho}_m : m \geqq 1\}$ and $\{\bar{\bar{\varrho}}_m : m \geqq 1\}$ have compact closures in $\mathscr{M}^1[\mathscr{X}]$. Any cluster point of the net $(\bar{\varrho}_m)$, or of the net $(\bar{\bar{\varrho}}_m)$, has a Maxwellian velocity distribution and satisfies (4.2) with $\lambda = \mathfrak{z}$, and is therefore concentrated on $[\hat{\mathscr{X}}]$. It is to such cluster points that ii) refers. We will see in § 6 that, if Φ has compact support and has a first derivative satisfying a Lipschitz condition, so that we can solve the equations of motion with $F(q) = -\frac{d}{dq}\Phi(q)$, then any cluster point of the net $(\bar{\bar{\varrho}}_m)$ is invariant under the time evolution given by this F.

§ 5. Approximation by Space-Periodized Systems

In this section we show that, if we consider the time-evolution of an infinite configuration x, but look only at those particles in some bounded interval, we find a motion which is well approximated by the evolution of a corresponding system which is space-periodized with respect to some much larger interval. To be more explicit, we will show that, for any ψ in $\mathfrak{A}$, any x in $[\hat{\mathscr{X}}]$ and any t, $\psi(T^t x) = \lim \psi(\tilde{T}^t_{[-a,b)} x)$ as a, b go to infinity in an appropriate way. Since the space-periodized evolution is constructed by putting together infinitely many independent finite systems, this result enables us to approximate infinite system by finite ones. It therefore makes possible the use of the classical mechanics of finite systems, notably of LIOUVILLE'S theorem and energy conservation, to obtain information about the infinite system.

This approximation theorem will be proved as follows: we start from the equations for a finite periodized system in the form given in § 2.5, Eq. (2.13). These equations are formally identical with the equations for a non-periodized system, and we will study them in the same way. Thus, we convert these equations to a non-linear evolution equation, which we solve by successive approximations. By keeping track of the way the estimates depend on the interval $[-a, b)$ of periodization, we find that the convergence of the solution by successive approximations is uniform in a, b, provided that they are large enough and that one is not too much larger than the other. Thus, all we have to do is to show that the nth approximation for the periodized system is close to the nth approximation for the original infinite system. Now the evolutions of the two systems differ only by "boundary effects" having to do with the behavior of particles near the ends of the periodicity interval. Using the finite range of the forces, and bounds on the distances particles can travel, we can control the propagation of these boundary effects and show that, for any n and any finite interval (α, β), the n-th approximations to the motion of the particles inside (α, β) for the periodized system and the

non-periodized system are identical, provided that the ends of the periodicity interval are far enough from the origin.

The details of the proof will consist primarily of the rewriting of the estimates of [1] in the slightly different context of finite periodized systems. We will frequently have to impose some restrictions on the periodicity intervals we consider, and it is convenient to have an abbreviation for this set of restrictions. We will say that an interval $[-a, b)$ is *allowable* if $a \geqq e^e$, $b \geqq e^e$, $a + b \geqq 2R$, and $1/2 \leqq \log(a)/\log(b) \leqq 2$. (No special importance should be assigned to the number e^e; it is simply a conveniently large number. A similar remark holds for the bounds on $\log(a)/\log(b)$.)

Let $[-a, b)$ be a finite interval, and let $x = (p_1, p_1; \ldots; q_n, p_n)$ be a configuration of particles in $[-a, b)$. We may regard x as a configuration in $\boldsymbol{R}$ which happens to be finite and to be contained in $[-a, b)$; with this convention we will use the definitions given in § 2.5 for $|x|$, $\mathscr{Y}_x$, $\| \ \|_x$, $\|\ \|_{x,m}$, etc. If $\tilde{x}(t)$ denotes the solution of the Eqs. (2.13) with $\tilde{x}(0) = x$, we write $\tilde{x}(t) = x + \tilde{\xi}_x(t)$, with $\tilde{\xi}_x(t)$ in $\mathscr{Y}_x$. The differential equations become

$$\frac{d\tilde{\xi}(t)}{dt} = \tilde{A}_{x,a+b}(\tilde{\xi}_x(t)) \tag{5.1}$$

where

$$\tilde{A}_{x,a+b}(\zeta) = \left(p_i + \eta_i, \sum_{j \neq i} \tilde{F}_{a+b}(q_i + \xi_i - q_j - \xi_j)\right). \tag{5.2}$$

The following lemma generalizes Lemma 3.4 of [1]:

Lemma 5.1. *There exists a constant K such that, for all allowable intervals $[-a, b)$, for all finite configurations $x = (q_1, p_1; \ldots; q_n, p_n)$ in $[-a, b)$, for all closed intervals $[\alpha, \beta] \subset [-a, b]$, for all $\lambda \geqq 1$, and for all n-tuples of numbers (ξ_i) with $\sup_i \frac{|\xi_i|}{\log_+(q_i)} \leqq \lambda$, we have the inequality:*

$$\begin{aligned} &\#\left\{j : q_j + \xi_j \in \bigcup_{k=-\infty}^{\infty} \{[\alpha, \beta] + k(a + b)\}\right\} \\ &\leqq |x| \left\{\beta - \alpha + K\lambda[\log_+(\lambda) + \log_+(|\alpha| \vee |\beta|)]\right\}. \end{aligned}$$

Proof. By Lemma 3.4 of [1], we have an estimate of the desired form on $\#\{j : q_j + \xi_j \in [\alpha, \beta]\}$. Thus, we want to consider j's such that $q_j + \xi_j \in [\alpha, \beta] + k(a + b)$ for some $k \neq 0$. Since $[\alpha, \beta] \subset [-a, b]$, we must at least have $q_j + \xi_j \notin (-a, b)$. But $|q_j| \leqq a \vee b$, so q_j must be within a distance $\lambda \log_+(a \vee b)$ of the boundary of $(-a, b)$, and the number of such q_j's is not greater than $2|x| \lambda \log_+(a \vee b)$. On the other hand, the interval $[\alpha, \beta] + k(a + b)$ must also come within a distance $\lambda \log_+(a \vee b)$ of the boundary of $(-a, b)$, and this implies that

$$|\alpha| \vee |\beta| \geqq a \wedge b - \lambda \log_+(a \vee b).$$

Using the inequalities

$$\log_+(a \vee b) \leqq 2 \log_+(a \wedge b)$$

and

$$\frac{1}{2}\log_+(a \vee b) - \log_+(\log_+(a \vee b)) \geqq \frac{1}{10}\log_+(a \vee b)$$

(which follows from $a \vee b \geqq e^e$), and also using the sub-additivity of $\log_+$, we get

$$\begin{aligned}\log_+(|\alpha| \vee |\beta|) &\geqq \log_+(a \wedge b) - \log_+(\lambda \log_+(a \vee b)) \\ &\geqq \frac{1}{2}\log_+(a \vee b) - \log_+(\lambda) - \log_+(\log_+(a \vee b)) \\ &\geqq \frac{1}{10}\log_+(a \vee b) - \log_+(\lambda)\,.\end{aligned}$$

Hence,

$$\begin{aligned}&\#\left\{j : q_j + \xi_j \in \bigcup_{k \neq 0} ([\alpha, \beta] + k(a+b))\right\} \\ &\leqq 20\,|x|\,\lambda\,[\log_+(|\alpha| \vee |\beta|) + \log_+(\lambda)]\end{aligned}$$

so the proof of the lemma is complete.

Proposition 5.2. *There exist constants C, D such that for all allowable $[-a, b)$, all configurations x in $[-a, b)$, and all ζ in $\mathscr{Y}_x$, we have*

$$\|\tilde{A}_{x, a+b}(\zeta)\| \leqq (1 + |x|)\,[C + D\,\|\zeta\|_x \log_+(\|\zeta\|_x)]\,.$$

The proof of this proposition, using Lemma 5.1, is nearly identical with the proof of Proposition 3.3 of [1]; we omit the details.

Lemma 5.3. *Let a real number d be given. Then there exists a constant B such that, for all $\alpha > 0$ there exists an m_0 such that, for all $m \geqq m_0$, all allowable $[-a, b)$, all configurations x in $[-a, b)$, and all ζ, ζ' in $\mathscr{Y}_x$ with $\|\zeta\| \leqq d$, $\|\zeta'\| \leqq d$, we have:*

$${}_m\|\tilde{A}_{x, a+b}(\zeta) - \tilde{A}_{x, a+b}(\zeta')\|_x \leqq B\,|x|\,\log_+(m)_{\alpha m}\|\zeta - \zeta'\|_x\,.$$

The proof is essentially the same as that of Lemma 4.1 of [1].

We now define:

$$\tilde{\xi}_{0, x}(t) = 0\,,$$

$$\tilde{\xi}_{n, x}(t) = \int_0^t d\tau\; \tilde{A}_{x, a+b}(\tilde{\xi}_{n-1, x}(\tau)) \quad \text{for} \quad n = 1, 2, 3, \ldots$$

Proposition 5.4. *There exist functions $\tilde{h}(\delta, T)$ and $\tilde{\varepsilon}(n, m, \delta, T)$, with*

$$\lim_{n \to \infty} \tilde{\varepsilon}(n, m, \delta, T) = 0$$

for all m, δ, T, such that:

i) $\|\tilde{\xi}_{n, x}(t)\|_x \leqq \tilde{h}(\delta, T)$

ii) ${}_m\|\tilde{\xi}_{n, x}(t) - \tilde{\xi}_x(t)\|_x \leqq \tilde{\varepsilon}(n, m, \delta, T)$

for all n, m, whenever $[-a, b)$ is an allowable interval, x is a configuration in $[-a, b)$ with $|x| \leqq \delta$, and $|t| \leqq T$.

The proof is essentially the same as that of Proposition 4.2 of [1].

We now state the principal result of this section:

Proposition 5.5. *Let $\gamma > 0$, δ, and T be given, and let ψ belong to $\mathfrak{A}$. Then there exists a real number A such that, whenever $a \geqq A$, $b \geqq A$, $1/2 \leqq \log(a)/\log(b) \leqq 2$, we have*

$$|\psi(T^t x) - \psi(\tilde{T}^t_{[-a,b)} x)| \leqq \gamma$$

if $x \in [\hat{\mathscr{X}}_\delta]$ and $|t| \leqq T$.

Proof. It suffices to prove the proposition for ψ of the form $\phi(Sf_1, \ldots, Sf_k)$, with ϕ a bounded continuous function on $\boldsymbol{R}^n$ and $f_1, \ldots, f_k$ in $\mathscr{K}_1$. Choose m so that $f_i(q, p) = 0$ for all i if $|q| \geqq m$.

For any labelled configuration x belonging to $\hat{\mathscr{X}}$ and any finite interval $[-a, b)$, let $\tilde{x}$ be the part of x in $[-a, b)$. [The index set for the finite labelled configuration $\tilde{x}$ may not be of the form $(1, 2, 3, \ldots, n)$, but this is inessential.] Note that $|\tilde{x}| \leqq |x|$. Using Propositions 2.18 and 5.4, we see that there is a constant H such that

$$\|\zeta_{n,x}(t)\|_x \leqq H, \quad \|\tilde{\zeta}_{n,\tilde{x}}(t)\|_{\tilde{x}} \leqq H \tag{5.3}$$

whenever $[-a, b)$ is allowable, $|x| \leqq \delta$, and $|t| \leqq T$.

Choose m_0 so that

$$|q| - 2H \log_+(q) \leqq m \quad \text{implies} \quad |q| \leqq m_0 . \tag{5.4}$$

Now these inequalities imply that, if $[-a, b)$ is allowable, if $|x| \leqq \delta$, if $|t| \leqq T$, and if $a \geqq m$, $b \geqq m$, the sums defining $Sf_j(T^t x)$ and $Sf_j(\tilde{T}^t_{[-a,b)} x)$ can be restricted to those i's with $|q_i| \leqq m_0$. The number of terms in such a sum is not greater than $2\delta m_0$, and (5.3) enables us to put an upper bound on the velocities of the particles entering the sum. Hence, the f_j, and also ϕ, are uniformly continuous on the relevant ranges of variables. To prove the proposition, then, it will suffice to prove the following assertion: For all m_0, δ, T and $\varepsilon > 0$, there exists a real number $A > m_0$ such that, if $1/2 \leqq \log(a)/\log(b) \leqq 2$, if $a \geqq A$, $b \geqq A$, if $|x| \leqq \delta$, and if $|t| \leqq T$, then

$$\sup \left\{ \frac{|\tilde{\xi}_{i,\tilde{x}}(t) - \xi_{i,x}(t)| \vee |\tilde{\eta}_{i,\tilde{x}}(t) - \eta_{i,x}(t)|}{\log_+(q_i)} : |q_i| \leqq m_0 \right\} \leqq \varepsilon . \tag{5.5}$$

Again using Propositions 2.18 and 5.4, we see that there exists n such that

$$m_0 \|\zeta_{n,x}(t) - \zeta_x(t)\|_x \leqq \varepsilon/2$$

$$m_0 \|\tilde{\zeta}_{n,x}(t) - \tilde{\zeta}_x(t)\|_x \leqq \varepsilon/2$$

whenever $|t| \leqq T$, $|x| \leqq \delta$, and $[-a, b)$ is allowable. Comparing these inequalities with (5.5) shows that it suffices to find, for any given n, a constant $A \geqq m_0$ such that, whenever $[-a, b)$, t, x are as above, with

$a \geqq A$, $b \geqq A$, we have

$$\tilde{\xi}_{i,n,\tilde{x}}(t) = \xi_{i,n,x}(t), \qquad \tilde{\eta}_{i,n,\tilde{x}}(t) = \eta_{i,n,x}(t) \tag{5.6}$$

for all i with $|q_i| \leqq m_0$.

We choose successively $m_1, m_2, \ldots, m_n$ so that, if $|q| \leqq m_i$, $|q'| \geqq m_{i+1}$,

$$|q| + H \log(q) + R < |q'| - 2H \log_+(q') . \tag{5.7}$$

We assert that we can take $A = m_n$. We will prove (5.6) by showing by induction that, for $[-a, b)$, t, x as above, for $a \geqq m_n$ and $b \geqq m_n$, and for $0 \leqq k \leqq n$, if $|q_i| \leqq m_{n-k}$, then

$$\tilde{\xi}_{i,k,\tilde{x}}(t) = \xi_{i,k,x}(t) , \qquad \tilde{\eta}_{i,k,\tilde{x}}(t) = \eta_{i,k,x}(t) . \tag{5.8}$$

This is clearly true for $k = 0$ since everything is then identically zero. Suppose it is true for k; we will prove it for $k+1$. Now

$$\tilde{\xi}_{i,k+1,x}(t) = \int_0^t d\tau [p_i + \tilde{\eta}_{i,k,x}(\tau)] , \tag{5.9}$$

$$\tilde{\eta}_{i,k+1,\tilde{x}}(t) = \int_0^t d\tau \left[\sum_{j \neq i} \tilde{F}_{a+b}(q_i + \tilde{\xi}_{i,k,\tilde{x}}(\tau) - q_j - \tilde{\xi}_{j,k,\tilde{x}}(\tau)) \right], \tag{5.10}$$

and corresponding Eq. (2.12) hold for $\xi_{i,k+1,x}(t)$ and $\eta_{i,k+1,x}(t)$. If $|q_i| \leqq m_{n-k-1}$, then $|q_i| \leqq m_{n-k}$, so $\tilde{\eta}_{i,k,\tilde{x}}(\tau) = \eta_{i,k,x}(\tau)$ by the induction hypothesis. Hence, by (5.9) and the first part of (2.12), $\tilde{\xi}_{i,k+1,\tilde{x}}(t) = \xi_{i,k+1,x}(t)$.

From the inequalities (5.3) and (5.7) it follows that, if $|q_i| \leqq m_{n-k-1}$, the sums over j in (5.10) and in the second part of (2.12) may be restricted to those j's with $|q_j| \leqq m_{n-k}$, and $\tilde{F}_{a+b}$ may be replaced by F. But from the induction hypothesis

$$\tilde{\xi}_{j,k,\tilde{x}}(\tau) = \xi_{j,k,x}(\tau)$$

for all j with $|q_j| \leqq m_{n-k}$. This proves the induction step (5.8) and therefore the proposition.

§ 6. Equilibrium States

In this section, we prove two propositions which imply that many infinite-volume limits of thermodynamic ensembles are invariant under the time-evolution defined by the corresponding potentials.

Proposition 6.1. *Let (a_α) be a net of positive numbers, with $\lim_\alpha a_\alpha = \infty$. For each α, let ϱ_α be a probability measure on $[\hat{\mathcal{X}}]$ which is invariant under $\tilde{T}^t_{[-a_\alpha, a_\alpha)}$. Suppose that*

$$\lim_{\delta \to \infty} \varrho_\alpha(\mathbb{C}[\hat{\mathcal{X}}_\delta]) = 0$$

uniformly in α, and that

$$\lim_\alpha \int d\varrho_\alpha \psi = \varrho(\psi)$$

for all ψ in $\mathfrak{A}$. *Then the state* ϱ *of* $\mathfrak{A}$ *is a probability measure concentrated on* $[\mathscr{X}]$ *and is invariant under* T^t.

Proposition 6.2. *Let* (a_α) *be a net of positive numbers, with* $\lim_\alpha a_\alpha = \infty$. *For each* α, *let* ϱ_α *be a probability measure on* $[\hat{\mathscr{X}}]$ *which is invariant under* $\tilde{T}^t_{[-a_\alpha, a_\alpha)}$. *Let* $\bar{\varrho}_\alpha = \frac{1}{2a_\alpha} \int_{-a_\alpha}^{a_\alpha} ds\,(\tau_s \varrho_\alpha)$ *and suppose that*

$$\lim_{\delta \to \infty} \bar{\varrho}_\alpha(\complement[\hat{\mathscr{X}}_\delta]) = 0$$

uniformly in α *and that*

$$\lim_\alpha \int d\bar{\varrho}_\alpha \psi = \varrho(\psi)$$

for every ψ *in* $\mathfrak{A}$. *Then the state* ϱ *of* $\mathfrak{A}$ *is a probability measure concentrated on* $[\hat{\mathscr{X}}]$ *and is invariant under* T^t.

We will give the details only for Proposition 6.2; the proof of Proposition 6.1 is similar but less complicated. Let us first dispose of showing that ϱ is concentrated on $[\hat{\mathscr{X}}]$. Choose δ so that $\bar{\varrho}_\alpha(\complement[\hat{\mathscr{X}}_\delta]) \geqq 1 - \varepsilon$ for all α. Since ϱ is a measure on the spectrum of $\mathfrak{A}$, and since, by Proposition 3.1, $[\hat{\mathscr{X}}_\delta]$ is a compact subset of the spectrum of $\mathfrak{A}$, we have:

$$\varrho([\hat{\mathscr{X}}_\delta]) = \inf\{\varrho(\psi) : \psi \in \mathfrak{A},\ \psi \geqq 0,\ \psi \geqq 1 \text{ on } [\hat{\mathscr{X}}_\delta]\}\,.$$

But for any such ψ, $\bar{\varrho}_\alpha(\psi) \geqq 1 - \varepsilon$ for all α, so $\varrho(\psi) \geqq 1 - \varepsilon$; hence, $\varrho([\hat{\mathscr{X}}_\delta]) \geqq 1 - \varepsilon$. We can make this argument for any $\varepsilon > 0$, so $\varrho([\hat{\mathscr{X}}]) = 1$.

To prove the rest of the proposition, it suffices to show that

$$\int d\varrho\ \psi \circ T^t = \int d\varrho\ \psi \tag{6.1}$$

for all ψ in $\mathfrak{A}$ with $\|\psi\| \leqq 1$. The first thing we want to show is that

$$\int d\varrho\ \psi \circ T^t = \lim_\alpha \int d\bar{\varrho}_\alpha\ \psi \circ T^t\,. \tag{6.2}$$

This is not immediate since $\psi \circ T^t$ is not in $\mathfrak{A}$. However, we can argue as follows: Let δ be chosen large enough so that $\bar{\varrho}_\alpha(\complement[\hat{\mathscr{X}}_\delta]) \leqq \varepsilon$ for all α. By Proposition 3.3, $\psi \circ T^t$ is continuous on $[\hat{\mathscr{X}}_\delta]$. Since $[\hat{\mathscr{X}}_\delta]$ is compact in the spectrum of $\mathfrak{A}$, the Tietze extension theorem asserts that there exists $\hat{\psi}$ in $\mathfrak{A}$ with $\|\hat{\psi}\| \leqq 1$, such that $\hat{\psi} = \psi \circ T^t$ on $[\hat{\mathscr{X}}_\delta]$. Then

$$|\int d\bar{\varrho}_\alpha(\psi \circ T^t - \hat{\psi})| \leqq 2\,\varepsilon$$

for all α, and similarly

$$|\int d\varrho(\psi \circ T^t - \hat{\psi})| \leqq 2\,\varepsilon\,.$$

Hence, whenever α is large enough so that $|\varrho(\hat{\psi}) - \bar{\varrho}_\alpha(\hat{\psi})| \leqq \varepsilon$, we have

$$\begin{aligned}|\int d\varrho(\psi \circ T^t) - \int d\varrho_\alpha(\psi \circ T^t)| &\leqq |\int d\varrho(\psi \circ T^t - \hat{\psi})| \\ &\quad + |\varrho(\hat{\psi}) - \bar{\varrho}_\alpha(\hat{\psi})| + |\int d\bar{\varrho}_\alpha(\psi \circ T^t - \hat{\psi})| \leqq 5\,\varepsilon\,,\end{aligned}$$

so (6.2) is proved.

Next observe that Eq. (2.3), applied with f equal to the characteristic function of $[\hat{\mathscr{X}}]$, implies that $\tau_s(\varrho_\alpha)$ is concentrated on $[\hat{\mathscr{X}}]$ for almost all (and therefore for all) s between $-a_\alpha$ and a_α. Also notice that:

$$\begin{aligned}\int d\bar{\varrho}_\alpha\,\psi &= \frac{1}{2a_\alpha}\int_{-a_\alpha}^{a_\alpha} ds \int d\varrho_\alpha\,\psi\circ\tau_s \\ &= \frac{1}{2a_\alpha}\int_{-a_\alpha}^{a_\alpha} ds \int d\varrho_\alpha\,\psi\circ\tau_s\circ\tilde{T}^t_{[-a_\alpha,a_\alpha)} \\ &= \frac{1}{2a_\alpha}\int_{-a_\alpha}^{a_\alpha} ds \int d\varrho\;\psi\circ\tilde{T}^t_{[-a_\alpha+s,a_\alpha+s)}\circ\tau_s \\ &= \int d\bar{\varrho}_\alpha(\psi\circ T^t) + \frac{1}{2a_\alpha}\int_{-a_\alpha}^{a_\alpha} ds\int d\varrho_\alpha \\ &\quad\cdot\left[\psi\circ\tilde{T}^t_{[-a_\alpha+s,a_\alpha+s)} - \psi\circ T^t\right]\circ\tau_s .\end{aligned}$$

[The first equality is just the definition of $\bar{\varrho}_\alpha$; the second follows from the invariance of ϱ_α under $\tilde{T}^t_{[-a_\alpha,a_\alpha)}$; the third follows from:

$$\tau_s\circ\tilde{T}^t_{[-a_\alpha,a_\alpha)} = \tilde{T}^t_{[-a_\alpha+s,a_\alpha+s)}\circ\tau_s\ ;$$

and the fourth uses Eq. (2.3) with $f=\psi\circ T^t$.]

Let $b_\alpha = \sup\{s<a_\alpha : 2\log(a_\alpha-s)\geqq\log(a_\alpha+s)\}$; then $\lim_\alpha \frac{b_\alpha}{a_\alpha}=1$, and proving (6.1) reduces to proving:

$$\lim_\alpha \frac{1}{2a_\alpha}\int_{-b_\alpha}^{b_\alpha} ds\int d\varrho_\alpha\left[\psi\circ\tilde{T}^t_{[-a_\alpha+s,a_\alpha+s)} - \psi\circ T^t\right]\circ\tau_s = 0\,. \tag{6.3}$$

Also, $\lim_\alpha(a_\alpha-b_\alpha)=\infty$, and, if $|s|\leqq b_\alpha$, $1/2\leqq\log(a_\alpha-s)/\log(a_\alpha+s)\leqq 2$.

We are therefore in a position to apply Proposition 5.5. For any δ, define χ_δ on $[\mathscr{X}]$ by

$$\begin{aligned}\chi_\delta &= \frac{1}{\delta}\text{ on }[\hat{\mathscr{X}}_\delta] \\ &= 2\text{ on }\complement[\hat{\mathscr{X}}_\delta]\,.\end{aligned}$$

Then for sufficiently large α

$$\sup_{|s|\leqq b_\alpha}\left|\psi\circ T^t_{[-a_\alpha+s,a_\alpha+s)} - \psi\circ T^t\right| \leqq \chi_\delta\,.$$

Hence,

$$\begin{aligned}&\limsup_\alpha \frac{1}{2a_\alpha}\left|\int_{-b_\alpha}^{b_\alpha} ds\int d\varrho_\alpha\left[\psi\circ\tilde{T}^t_{[-a_\alpha+s,a_\alpha+s)} - \psi\circ T^t\right]\circ\tau_s\right| \\ &\leqq \limsup_\alpha \int d\bar{\varrho}_\alpha\,\chi_\delta\end{aligned}$$

for all δ. Since $\lim_{\delta \to \infty} \int d\bar{\varrho}_\alpha \, \chi_\delta = 0$ uniformly in α, (6.3) is proved and the proof of the proposition is complete.

We now describe two applications of these propositions to proving that thermodynamic limit states are invariant under the time evolution. Assume that $F(q) = -\frac{d}{dq}\Phi(q)$, where Φ is even, of compact support, and stable. We choose an inverse temperature and a chemical potential, and we construct, for every real number $m \geqq R$ which is large enough so that the P-stability inequality holds, the periodized grand canonical ensemble on $[-m, m]$. Recall that, in our terminology, this is a probability measure on $[\mathscr{X}]$ ($[-m, m)$). Let $\tilde{\varrho}_m$ be the measure on $[\mathscr{X}]$ obtained from this measure by the product measure construction. By LIOUVILLE's theorem and energy conservation, $\tilde{\varrho}_m$ is invariant under $\tilde{T}^t_{[-m,m)}$.

A straightforward adaptation of the arguments in [8] shows that, if the activity is sufficiently small, the measures $\tilde{\varrho}_m$:

i) have correlation functions satisfying a bound of the form (4.2), where λ may be taken to be independent of m;

ii) converge as m goes to infinity to a translation-invariant measure ϱ on $[\mathscr{X}]$.

Moreover, ϱ is the same state as is obtained in [8] as the infinite volume limit of non-periodized grand canonical ensembles. The discussion in this reference is in fact formulated in terms of correlation functions and not in terms of measures, but, because of the bound (4.2), statements about correlation functions may easily be translated into statements about measures[10].

Taking into account Proposition 4.1, we see that

$$\lim_{\delta \to \infty} \tilde{\varrho}_m(\mathbb{C}[\hat{\mathscr{X}}_\delta]) = 0$$

uniformly in m. Hence, Proposition 6.1 applies and shows that ϱ is invariant under the time evolution.

Now assume that Φ is non-negative, rather than merely stable, and consider $\bar{\tilde{\varrho}}_m$, the average of $\tilde{\varrho}_m$ over translations, for any fixed values of the inverse temperature and chemical potential (i.e., not necessarily for

[10] If a measure $\hat{\varrho}$ on the space of locally finite position configurations has correlation functions $\bar{\varrho}$ satisfying (4.2), then the measure may be recovered from the correlation functions as follows: Let Λ be a bounded Borel subset in $\boldsymbol{R}$, and let E be a symmetric Borel subset of Λ^n. The $\hat{\varrho}$-measure of the set of all configurations having precisely n particles in Λ and those n particles distributed so that their co-ordinates form a point of E is

$$\frac{1}{n!} \sum_{j=0}^{\infty} \frac{(-1)^j}{j!} \int_E dq_1, \ldots, dq_n \int_{\Lambda^j} dq'_1, \ldots, dq'_j \, \bar{\varrho}_{n+j}(q_1, \ldots, q_n, q'_1, \ldots, q'_j) \,.$$

This formula, which does not seem to appear in the literature, was pointed out to me by D. RUELLE.

small activity). By the last paragraph of § 4, $\{\bar{\bar{\varrho}}_m\}$ has compact closure in $\mathscr{M}^1[\mathscr{X}]$ and

$$\lim_{\delta\to\infty} \bar{\bar{\varrho}}_m(\mathbb{C}[\hat{\mathscr{X}}_\delta]) = 0$$

uniformly in m. Hence, Proposition 6.2 implies that any cluster point of the net $(\bar{\bar{\varrho}}_m)$ is invariant under the time evolution.

§ 7. Conservation of Entropy

In this section, we will assume that the interparticle force F is of the form $-\frac{d}{dq}\Phi(q)$, where Φ is even, of compact support, and stable, and we will consider probability measures ϱ on $[\mathscr{X}]$ which are translation invariant and concentrated on $[\hat{\mathscr{X}}]$, and which have, in addition,

$$\int d\varrho\, N^2_{[0,1)} < \infty\,, \qquad \int d\varrho\, K.E._{[0,1)} < \infty\,.$$

For such a measure ϱ, we denote by ϱ^t the measure $\varrho \circ T^t$. The main result of this section is the following:

Proposition 7.1. *Let ϱ, F be as above. Then, for any t,*

$$\int d\varrho^t\, K.E._{[0,1)} < \infty$$

so $\bar{s}(\varrho^t)$ is defined, and we have:

$$\bar{s}(\varrho^t) \geqq \bar{s}(\varrho)\,.$$

This would immediately imply that $s(\varrho^t) = s(\varrho)$ if we knew that

$$\int d\varrho^t\, N^2_{[0,1)} < \infty\,;$$

we will also prove this, but under more restrictive assumptions on the potential Φ.

In outline, the proof of Proposition 7.1 goes as follows:

1. We consider, instead of ϱ^t, the measure $\varrho^t_n = \varrho \circ \tilde{T}^t_{[-n,n)}$, which should be a good approximation to ϱ^t for large n, by Proposition 5.5. Although it need not be translation invariant, ϱ^t_n is periodic with period 2 n.

2. Using conservation of energy for the periodized system. together with the P-stability of the potential, we obtain a bound

$$\frac{1}{2n}\int d\varrho^t_n\, K.E._{[-n,n)} \leqq M$$

valid for large n. In particular $\bar{s}(\varrho^t_n)$ is defined for such n.

3. Using LIOUVILLE's theorem, we show that

$$\bar{s}(\varrho^t_n) = \bar{s}(\varrho)\,.$$

4. Denoting by $\bar{\varrho}_n^t$ the average of ϱ_n^t over translations

$$\bar{\varrho}_n^t = \frac{1}{2n} \int_{-n}^{n} ds\, \tau_s \varrho_n^t ,$$

we have by Corollary 2.17 $\bar{s}(\bar{\varrho}_n^t) = \bar{s}(\varrho)$

$$\int d\bar{\varrho}_n^t\, K.E._{\cdot[0,1)} = \frac{1}{2n} \int K.E._{\cdot[-n,n)}\, d\varrho_n^t \leqq M .$$

5. Finally, using Proposition 5.5, we prove

$$\lim_{n\to\infty} \bar{\varrho}_n^t = \varrho^t$$

which implies by a semi-continuity argument:

$$\int K.E._{\cdot[0,1)}\, d\varrho^t \leqq \liminf_n \int K.E._{\cdot[0,1)}\, d\bar{\varrho}_n^t \leqq M$$

$$\bar{s}(\varrho^t) \geqq \limsup_n \bar{s}(\bar{\varrho}_n^t) = \bar{s}(\varrho) .$$

We now proceed to fill in the details. Step 1. merely defines the notation. For step 2., we define the energy in $[-n, n)$ for the periodized interaction as:

$$\tilde{E}_{[-n,n)} = K.E._{\cdot[-n,n)} + \widetilde{P.E.}_{\cdot[-n,n)}$$

$$\widetilde{P.E.}_{\cdot[-n,n)}(x) = \frac{1}{2} \sum_{\substack{i \neq j \\ q_i, q_j \in [-n,n)}} \tilde{\Phi}_{2n}(q_i - q_j)$$

where (q_i, p_i) is a representative of x. By conservation of energy for the periodized system,

$$\tilde{E}_{[-n,n)} \circ \tilde{T}^t_{[-n,n)} = \tilde{E}_{[-n,n)} .$$

We will prove that $\tilde{E}_{[-n,n)}$ is ϱ-integrable and estimate its integral. By the translation invariance of ϱ,

$$\int d\varrho\, K.E._{\cdot[-n,n)} = 2n \int d\varrho\, K.E._{\cdot[0,1)} ;$$

on the other hand, if K is an upper bound for $|\Phi|$, and if J is an integer not smaller than R, then

$$|\widetilde{P.E.}_{\cdot[-n,n)}| \leqq \frac{K}{2} \sum_{i=-n}^{n-1} \sum_{j=-J}^{J} N_{[i,i+1)}\, N_{[i \dot{+} j,\, i \dot{+} j+1)} ,$$

where $i \dot{+} j$ means that integer in $[-n, n)$ which is equal to $i + j$ modulo $2n$. Applying the Schwarz inequality and translation invariance, we get

$$\int d\varrho\, \widetilde{P.E.}_{\cdot[-n,n)} \leqq K(2J+1)\, n \int d\varrho\, N^2_{[0,1)} .$$

Hence,

$$\int d\varrho_n^t\, \tilde{E}_{[-n,n)} = \int d\varrho\, \tilde{E}_{[-n,n)} \circ \tilde{T}^t_{[-n,n)} = \int d\varrho\, \tilde{E}_{[-n,n)} \leqq 2n\, M' ,$$

where M' does not depend on n.

Now suppose that n is large enough so that

$$\sum_{1 \leqq i < j \leqq m} \tilde{\Phi}_{2n}(q_i - q_j) \geqq -Bm$$

for all $m, q_1, \ldots, q_m$. Then

$$\widetilde{P.E.}_{[-n,n)} + B\, N_{[-n,n)} \geqq 0\,.$$

But we also have

$$\begin{aligned}\int d\varrho_n^t\, N_{[-n,n)} &= \int d\varrho\, N_{[-n,n)} \circ \tilde{T}^t_{[-n,n)} = \int d\varrho\, N_{[-n,n)} \\ &= 2n \int d\varrho\, N_{[0,1)}\end{aligned}$$

and therefore

$$\int d\varrho_n^t\, K.E._{[-n,n)} \leqq \int d\varrho_n^t\, [\tilde{E}_{[-n,n)} + B\, N_{[-n,n)}] \leqq 2n\, M$$

where again M does not depend on n. This finishes the proof of 2.

To prove step 3., it suffices, by statement 2. of Proposition 2.15, to prove that

$$s_{[-jn,jn)}(\varrho_n^t) = s_{[-jn,jn)}(\varrho)$$

for all odd integers j. LIOUVILLE'S theorem asserts that the measure $\sigma_{[-jn,jn)}$ is invariant under $\tilde{T}^t_{[-n,n)}$; our statement now follows from Proposition 2.13.

Step 4. is a straightforward application of Corollary 2.17. For step 5., we have first to prove

$$\lim_{n\to\infty} \frac{1}{2n} \int_{-n}^{n} ds \int d\varrho\, \psi \circ \tau_s \circ T^t_{[-n,n)} = \int d\varrho\, \psi \circ T^t \qquad (7.1)$$

for all ψ in $\mathfrak{A}$. Using the equation

$$\tau_s \circ \tilde{T}^t_{[-n,n)} = \tilde{T}^t_{[-n+s,n+s)} \circ \tau_s$$

and the translation invariance of ϱ, we have

$$\begin{aligned}&\frac{1}{2n} \int_{-n}^{n} ds \int d\varrho\, \psi \circ \tau_s \circ \tilde{T}^t_{[-n,n)} \\ &= \frac{1}{2} \int_{-1}^{1} da \int d\varrho\, \psi \circ \tilde{T}^t_{[-n(1-a),n(1+a))}\,.\end{aligned}$$

By Proposition 5.5, if $-1 < a < 1$,

$$\lim_{n\to\infty} \psi \circ \tilde{T}^t_{[-n(1-a),n(1+a))} = \psi \circ T^t$$

on $[\hat{\mathfrak{X}}]$; hence, applying the dominated convergence theorem twice gives (7.1).

To finish the proof of step 5., and hence of the proposition, we have only to show that

$$\int K.E._{[0,1)}\, d\varrho^t \leqq M\ ;$$

then the statement about the mean entropy will follow from statement 6 of Proposition 2.15. Because ϱ^t is translation invariant, $K.E._{(0,1)} = K.E._{[0,1)}$ almost everywhere, so it will suffice to prove that

$$\varrho \longmapsto \int K.E._{(0,1)}\, d\varrho$$

is a lower semi-continuous function on $\mathscr{M}^1[\mathscr{X}]$. Let (χ_j) be an increasing sequence of non-negative continuous functions on $\boldsymbol{R}$ converging pointwise to the characteristic function of (0,1), and let $f_j(q,p) = \chi_j(q)\frac{p^2}{2}$. Then $\psi_j = (Sf_j) \wedge j$ is an increasing sequence in $\mathfrak{A}$ such that

$$\int K.E._{(0,1)}\, d\varrho = \sup_j \varrho(\psi_j)$$

for all ϱ in $\mathscr{M}^1[\mathscr{X}]$. Since $\varrho \longmapsto \varrho(\psi_j)$ is continuous, our assertion is proved.

Next, we take up the question of the integrability of $N^2_{(0,1)}$ with respect to ϱ^t.

Proposition 7.2. *Let ϱ be as in the first paragraph of this section; assume that $F(q) = -\frac{d}{dq}\Phi(q)$, where Φ has compact support and is of the form $\Phi_1 + \Phi_2$, with Φ_1 and Φ_2 both even, Φ_1 P-stable, and Φ_2 non-negative and bounded away from zero on a neighborhood of the origin. Then*

$$\int d\varrho^t\, N^2_{[0,1)} < \infty\,.$$

Proof. We will keep the notation of the proof of the preceding proposition. Since ϱ^t is translation invariant,

$$\int d\varrho^t\, N^2_{[0,1)} = \int d\varrho^t\, N^2_{(0,1)}\,.$$

Arguing as in the proof of step 5. of the preceding proposition, we see that $\varrho \longmapsto \int d\varrho\, N^2_{(0,1)}$ is lower semi-continuous on $\mathscr{M}^1[\mathscr{X}]$ and therefore that it suffices to obtain an upper bound on

$$\frac{1}{2n}\sum_{j=-n}^{n-1} \int d\bar{\varrho}^t_n\, N^2_{(j,j+1)} \text{ which is uniform in } n \text{ for large } n.$$

By Eq. (2.3),

$$\begin{aligned}\sum_{j=-n}^{n-1}\int d\bar{\varrho}^t_n\; N^2_{(j,j+1)} &= \frac{1}{2n}\int_0^{2n} ds \int d\varrho^t_n \sum_{j=-n}^{n-1} N^2_{(j,j+1)} \circ \tau_s \\ &= \frac{1}{2n}\int_0^{2n} ds \int d\varrho^t_n \sum_{j=-n}^{n-1} N^2_{(j-s,j-s+1)}\end{aligned}$$

For any value of s between 0 and $2n$, some of the intervals $(j-s, j-s+1)$ will be contained in $(-n,n)$ and some in $(-3n,-n)$. One, at most, will contain $-n$. If $j - s < -n < j - s + 1$, we can estimate

$$\begin{aligned}\int d\varrho^t_n\, N^2_{(j-s,j-s+1)} &= \int d\varrho^t_n\, (N_{(j-s,-n)} + N_{[-n,j-s+1)})^2 \\ &\leqq 2\int d\varrho^t_n\, [N^2_{(-n-1,-n)} + N^2_{[-n,-n+1)}]\,.\end{aligned}$$

From these two remarks, and the periodicity of ϱ_n^t, we see that to prove the proposition it will be sufficient to show that there is a constant K such that, whenever n is large enough,

$$\frac{1}{2n}\int d\varrho_n^t \sum_j N_{I_j}^2 \leqq K$$

for all pairwise-disjoint collections $\{I_1, I_2, \ldots\}$ of intervals of unit length contained in $[-n, n)$. Using the conservation of energy and particle number for the space-periodized evolution, we can get such an estimate if we can find constants C, C' such that

$$\sum_j N_{I_j}^2 \leqq C\,\tilde{E}_{[-n,n)} + C'\,N_{[-n,n)}$$

whenever n is large enough and $\{I_1, I_2, \ldots\}$ is as above.

From the P-stability of Φ_1, we have

$$2\,\tilde{E}_{[-n,n)}(x) \geqq -B\,N_{[-n,n)}(x) + \sum_{\substack{i \neq j \\ q_i, q_j \in [-n,n)}} \tilde{\Phi}_{2,2n}(q_i - q_j)$$

(where (q_i, p_i) is a representative of x).

Since Φ_2 is non-negative, we can omit any terms we like from the sum on the right; we keep only those pairs with q_i, q_j both belonging to the same one of the I_k's. The proposition now follows from the fact [9] that there exist B', B'', with $B'' > 0$, such that, for all m and all $q_1, \ldots, q_m$ in $[0,1]$,

$$\sum_{i \neq j} \Phi_2(q_i - q_j) \geqq -B'm + B''m^2 .$$

Acknowledgements. I am grateful to D. RUELLE for several helpful suggestions and for reading the manuscript, and to G. GALLAVOTTI and S. MIRACLE for useful discussions. I also thank Monsieur L. MOTCHANE for his hospitality at the Institut des Hautes Etudes Scientifiques.

References

1. LANFORD, O. E.: Commun. Math. Phys. **9**, 169—191 (1968).
2. RUELLE, D.: J. Math. Phys. **8**, 1657 (1967).
3. ROBINSON, D. W., and D. RUELLE: Commun. Math. Phys. **5**, 288 (1967).
4. HALMOS, P. R.: Measure theory, Section 52, Theorem G. New York: Van Nostrand 1950.
5. DUNFORD, N., and J. T. SCHWARTZ: Linear operators, Part I, Corollary IV.8.11 and Theorem V.6.1. New York: Interscience 1958.
6. KELLEY, J. L.: General topology, p. 81. New York: Van Nostrand 1955.
7. CHOQUET, G., and P. A. MEYER: Ann. Inst. Fourier **13**, 139 (1963), Lemme 10.
8. RUELLE, D.: Ann. Phys. **25**, 109 (1963).
9. — Helv. Phys. Acta **36**, 183 (1963), Sect. 1, Eq. (16).

O. E. LANFORD III
Department of Mathematics
University of California
Berkeley, California 94720, U.S.A.

Commun. math. Phys. 15, 47—68 (1969)

A General Class of Cut-Off Model Field Theories

ARTHUR M. JAFFE*
Lyman Laboratory of Physics, Harvard University, Cambridge, Massachusetts

OSCAR E. LANFORD III
Department of Mathematics, University of California, Berkeley

ARTHUR S. WIGHTMAN**
Institut des Hautes Etudes Scientifiques, Bures-sur-Yvette, Essonne

Received June 5, 1969

Abstract. We show that Heisenberg picture fields and their vacuum expectation values exist for a wide class of cut-off interactions among fermions and bosons.

I. Introduction

The quantum field theories studied in the present paper include cutoff versions of many standard relativistic quantum field theories. They have some interest of their own as examples of non-trivial dynamics. However, the main point of studying them is to obtain information about the relativistic theories that are their putative limits as the cutoffs are removed. For this purpose, it is desirable to show

1) that the knowledge of a suitable set of matrix elements of the Hamiltonian of the cutoff theory uniquely determines the one parameter group e^{iHt}, $-\infty < t < \infty$, describing the time evolution of the system,

2) that H has a reasonable spectrum,

3) that the Green's functions of the theory are uniquely determined.

The results of the present paper partially satisfy these requirements. It is shown that for the models considered

1') there is a dense set, D_0, of vectors in the Hilbert space of states in which the Hamiltonian is essentially self-adjoint.

2') that H has a purely discrete spectrum with finite multiplicity, bounded below and is such that its eigen functions lie in D_0.

3') that D_0 is invariant under the smeared fields and that for certain values of the coupling constants the ground state is non-degenerate.

* Alfred P. Sloan Foundation Fellow.

** On leave from Princeton University.

1′) completely settles 1). 2′) settles 2), 3′) does not completely settle 3); when the ground state is degenerate the definition of the Green's functions is ambiguous. For a class of models involving only Bose fields it is shown that the ground state is non-degenerate. On the other hand, a counter-example is given in which Fermions are present and the ground state *is* degenerate.

These results generalize those obtained by two of the present authors for a self-interacting Bose field [1] and for a Yukawa interaction of a spinor and a Bose field [2].

The paper deals with strong cutoffs in which only a finite number of boson modes are coupled. The vacuum expectation values are shown to be continuous in the times so that the Green's functions are unambiguously defined when the ground state is non-degenerate.

The essential mathematical idea of the proofs can be illustrated on the anharmonic oscillator

$$H = -\frac{d^2}{dx^2} + \alpha x^2 + \beta x^4, \quad \alpha \text{ real}, \quad \beta > 0.$$

One treats αx^2 as a perturbation on $-\frac{d^2}{dx^2} + \beta x^4$. This puts no restriction on α because, whatever the size of α, αx^2 is infinitely small compared to $-\frac{d^2}{dx^2} + \beta x^4$ in the sense of T. Kato. The class of models considered is restricted by the requirement that appropriate formally positive dominating boson self-couplings (analogues of βx^4 in the anharmonic oscillator) be present.

Most of the remaining technical difficulties of the paper arise because of the necessity of treating not only the Hamiltonian H but all its powers, in order to establish the properties of the invariant domain of vectors, D_0, and from the fact that we want to allow different species of bosons to have different order dominant self-interactions. The basic idea of the proof is to show that for every positive integer n

$$(H_0^f + H_1 + H_2)^n - (H_0^f + H_1)^n \tag{1.1}$$

is infinitely small compared

$$(H_0^f + H_1)^n \tag{1.2}$$

in the sense of T. Kato. Here H_0^f is the free Hamiltonian of the fermion fields: H_1 is the free Hamiltonian of the boson fields plus the dominant self-interactions and H_2 is the rest of the interaction. When one expands (1.1) by the binomial theorem and estimates the resulting terms relative to $(H_0^f + H_1)^n$ one arrives at the detailed conditions on the interaction stated in the next section.

II. Notation and Preliminaries

If B is an operator, $D(B)$ denotes its domain and $C^\infty(B) = \bigcap_{n=1}^{\infty} D(B^n)$. $(B)^-$ denotes the closure of B. If B and C are operators, then

$$\mathrm{Ad}\, C(B) = [C, B]$$

defines an operator. With this notation the multiple commutator $[C, [C, \ldots [C, B]]]$ with $n C's$ is $(\mathrm{Ad}\, C)^n(B)$.

We put all fields in a space box of volume V with periodic boundary conditions (3-torus!). In the strongly cutoff case, boson fields have a sharp ultra-violet cutoff. For example, if ϕ is a scalar boson field, the interaction term would depend on the cutoff field

$$\phi_{K,V}(x) = \frac{1}{(2V)^{1/2}} \sum_{k \in \Gamma_{K,V}} \exp[-ikx]\{a_V^*(k) + a_V(-k)\}\, \omega(k)^{-1/2}, \tag{2.1}$$

where

$$\omega(k) = [k^2 + m^2]^{1/2},\ \Gamma_V = \left\{k : k = \frac{2\pi}{V^{1/3}} \nu, \nu \in Z^3\right\},\ \Gamma_{K,V} = \Gamma_V \cap \{k : |k| \leqq K\},$$

and the $a_V(k)$, $a_V^*(k)$ are the standard Bose annihilation and creation operators normalized so that

$$[a_V(k), a_V^*(l)]_- = \delta_{kl}. \tag{2.2}$$

Analogous formulae hold for boson fields of other tensor characters. The fermion fields will be assumed to enter the interaction in regularized form

$$\psi_\varrho(x) = \int \varrho(x - y)\, dy\, \psi(y) \tag{2.3}$$

where ϱ is a smooth function of fast decrease i.e. belongs to the Schwartz space $\mathscr{S}$. In contrast to the case of boson fields $\psi_\varrho(x)$ is a bounded operator for each x.

The Hamiltonian is assumed to be a sum of a *free Hamiltonian*, $H_{0,V}$, and an *interaction Hamiltonian*, $H_{I,V}$. The different fields contribute additively to the free Hamiltonian. The contribution of a boson field like those described above is

$$\sum_{k \in \Gamma_V} a_V^*(k)\, a_V(k)\, \omega(k); \tag{2.4}$$

the contribution of a species of fermion to $H_{0,V}$ has precisely this form but the $a_V(k)$ and $a_V^*(k)$ are solutions of the anti-commutation relations instead of (2.2).

The interaction Hamiltonian is an integral over the box of an interaction Hamiltonian density:

$$H_{I,V} = \int_V d\boldsymbol{x}\, \mathfrak{H}_{I,V}(\boldsymbol{x}) . \tag{2.5}$$

$\mathfrak{H}_{I,V}(\boldsymbol{x})$ is a formally hermitean polynomial in the cutoff boson fields. For example, when a single hermitean scalar boson field is present

$$\mathfrak{H}_{I,V}(\boldsymbol{x}) = J^{(0)}(\boldsymbol{x}) + \sum_{\alpha=1}^{2n} J^{(\alpha)}(\boldsymbol{x})\, \phi_{K,V}(\boldsymbol{x})^{\alpha} \tag{2.6}$$

where the $J^{(\alpha)}(\boldsymbol{x})$ are polynomial expressions in the fermion fields. When N cutoff hermitean scalar boson fields are present

$$\mathfrak{H}_{I,V}(\boldsymbol{x}) = J^{(0)}(\boldsymbol{x}) + \sum_{\alpha} J^{(\alpha)}(\boldsymbol{x}) \prod_{j=1}^{N} (\phi^{(j)}_{K,V}(\boldsymbol{x}))^{\alpha_j} \tag{2.7}$$

Here $\alpha = (\alpha_1, \alpha_2, \ldots \alpha_n)$ is a multi-index, and the summation is taken over all multi-indices for which $0 \leqq \alpha_j \leqq 2n_j$. For convenience, the $J^{(\alpha)}$ will be referred to as *fermion currents*.

Now we come to the restrictions on the interaction Hamiltonian imposed by the requirement already mentioned above, that it should be dominated by boson self-interaction terms. In the case (2.6) of a single scalar boson field, it is fairly natural that that should be taken to mean

$$J^{(2n)}(\boldsymbol{x}) = \lambda 1 \tag{2.8}$$

λ a real number > 0. For the case of several boson fields the natural requirement is not obvious. It turns out to be sufficient to assume

a) that the only term in (2.7) in which $\phi^{(j)}_{K,V}(\boldsymbol{x})$ occurs with maximal degree $2n_j$ is of the form

$$\lambda_j (\phi^{(j)}_{K,V}(\boldsymbol{x}))^{2n_j}, \qquad \lambda_j > 0, j = 1, \ldots N .$$

b) For each remaining interaction term

$$\sum_{j=1}^{N} \frac{\alpha_j}{2n_j} < 1 . \tag{2.9}$$

In addition, it turns out that we need a boundedness property of the fermion currents.

c) For all α occuring in $\mathfrak{H}_{I,V}$

$$\| (\mathrm{ad}\, H_{0,V})^n (J^{(\alpha)}(\boldsymbol{x})) \| \leqq M_{n,\alpha} < \infty \tag{2.10}$$

for some constants $M_{n,\alpha}$ and all $n = 0, 1, 2, \ldots$

Evidently, for $N = 1$, a) and b) reduce to the above mentioned restriction for a single boson field. As they stand these requirements exclude

the possibility that the interaction is only linear in a boson field. We admit this possibility with the special rule that n_j is to be set equal to 1 in this case. The reason for this rule is that there are always quadratic terms in the free boson Hamiltonian.

Conditions a), b) and c) give one precise form of the idea of dominant boson self interaction but by no means the only one. For example, it would also suffice to replace a) and b) by

a') that there is a term of highest degree $2n$ in the boson fields of the form

$$\lambda\left[\sum_{j=1}^{N}(\phi^{(j)}_{K,V}(x))^2\right]^n \tag{2.11}$$

b') that all remaining terms in the interaction Hamiltonian are of the form indicated in (2.7) with

$$\sum_{j=1}^{N}\frac{\alpha_j}{2n}<1\,. \tag{2.12}$$

There are cases covered by a') b') not included under a), b). For example, the cross term $\phi^{(1)2}\phi^{(2)2}$ which arises from $(\phi^{(1)2}+\phi^{(2)2})^2$ is not admissible under b) because $\alpha_1=\alpha_2=2$, $n_1=n_2=2$ so $\sum_{j=1}^{2}\frac{\alpha_i}{2n_j}=1$. The reader will be able to invent still other sufficient conditions and also to generalize to an arbitrary set of tensor fields after having read the next section. We content ourselves with listing some examples

Examples of Admissible Interaction Hamiltonian Densities $\mathfrak{H}_{I,V}(x)$

1) Neutral pseudo-scalar meson theory

$$g\psi^*_\varrho(x)\,\gamma^5\psi_\varrho(x)\,\phi_{K,V}(x)+\lambda(\phi_{K,V}(x))^4,\qquad \lambda>0\,. \tag{2.13}$$

2) The σ-model of pions

$$g(\sigma^2_{K,V}(x)+\boldsymbol{\phi}_{K,V}(x)^2)^2+g_1\boldsymbol{\phi}_{K,V}(x)^2\sigma_{K,V}(x)+g_2\sigma_{K,V}(x),\qquad g>0\,. \tag{2.14}$$

Here $\boldsymbol{\phi}_{K,V}$ has three components (isospin).

3) Two hermitean scalar boson fields $\phi^{(1)}$, $\phi^{(2)}$

$$g[\phi^{(1)}_{K,V}(x)]^2\phi^{(2)}_{K,V}(x)+\lambda(\phi^{(1)}_{K,V}(x))^6,\qquad \lambda>0\,. \tag{2.15}$$

Here the condition (2.9) is very restrictive. Replacing the sixth power by the fourth would *not* yield an admissible interaction as long as $g\neq 0$.

4) Quantum Electrodynamics of spin $\frac{1}{2}$ particles in the Coulomb gauge

$$e\psi^*_\varrho(x)\,\psi_\varrho(x)\int_V dy\,\mathscr{D}_V(x-y)\,\psi^*_\varrho(y)\,\psi_\varrho(y)-e\psi^*_\varrho(x)\,\boldsymbol{\alpha}\psi_\varrho(x)\cdot A_{K,V}(x)\,. \tag{2.16}$$

Here $A_{K,V}(x)$ is constructed so as to satisfy $\nabla \cdot A_{K,V}(x) = 0$, and $\mathscr{D}_V(x-y)$ is the analogue on the torus, V, of the Coulomb potential.

It is also worth pointing out that there are physically interesting cases to which our methods in their present form do *not* apply. For example, in the theory of Yang-Mills fields there occur a pair of three component boson fields $\boldsymbol{g}$ and $\boldsymbol{h}$ with a dominant coupling term

$$\lambda(\boldsymbol{g} \times \boldsymbol{h})^2, \quad \lambda > 0. \tag{2.17}$$

In the "direction" $\boldsymbol{g} \times \boldsymbol{h}$ this does not grow, and therefore it does not come under the present theory. It is in fact rather unstable since a term $-\varepsilon^2(\boldsymbol{g}^2 + \boldsymbol{h}^2)$ can make the interaction unbounded below, no matter how small ε is. It would be of some interest to extend our results to this case with positivity imposed on lower order terms.

The preceding statements about the Hamiltonians of the models under discussion have been put in a form in which it is obvious what relativistic theories one might hope to obtain by removal of the cutoffs. For the calculations that follow it is much more convenient to have the boson Hamiltonians in the form of partial differential operators. This is done by introducing appropriate hermitean linear combinations $Q_{\boldsymbol{k}}^{(j)}, P_{\boldsymbol{k}}^{(j)}, j = 1, \ldots N, \boldsymbol{k} \in \Gamma_{KV}$, of the annihilation and creation operators occuring in equation (2.1) so chosen that they satisfy the canonical commutation relations. The cutoff field $\phi_{K,V}^{(j)}(x)$ is then a finite linear combination of the $Q_{\boldsymbol{k}}^{(j)}, \boldsymbol{k} \in \Gamma_{K,V}$ alone. The free Hamiltonian of the boson fields splits into two parts, one involving the uncoupled modes, the other, the coupled modes. The part involving the coupled modes together with the dominant self-interaction can be written as a partial differential operator. Altogether, we have

$$H_{K,V} = H_0^f + H_1 + H_2 \tag{2.18}$$

where H_0^f is the free fermion Hamiltonian,

$$H_1 = \sum_{j=1}^{N} \sum_{\substack{\text{uncoupled} \\ \text{modes of } \phi^{(j)}}} \omega_j(\boldsymbol{k})\, a^{(j)*}(\boldsymbol{k})\, a^{(j)}(\boldsymbol{k}) \tag{2.19}$$

$$+ \sum_{j=1}^{N} \sum_{\substack{\text{coupled} \\ \text{modes of } \phi^{(j)}}} [(-\Delta)_{Q^{(j)}} + V^{(j)}(Q^{(j)})], \tag{2.20}$$

and

$$H_2 = \sum_{j=1}^{N} \sum_{\alpha} M^{(j)\alpha} \Pi (Q^{(j)})^{\alpha}. \tag{2.21}$$

The product in the expression for H_2 runs over all coupled modes. The $M^{(j)\alpha}$ are bounded operators acting on the fermion variables; they are integrals over the box of the product of fermion currents with the sines and cosines occuring in the expansion of the $\phi^{(j)}$ in terms of $Q^{(j)}$.

III. Strongly Cut-Off Theories

Our first main result is

Theorem 3.1. *Let $H_{K,V} = H_0^f + H_1 + H_2$ where H_0^f is the free field Hamiltonian of the fermions H_1 is the free field boson Hamiltonian plus the cutoff dominant boson self-interactions and H_2 is the rest of the strongly cuttoff interaction as given by* (2.21). *$H_{0,V}$ is then the sum of the free field Hamiltonians for bosons and fermions. Let n be any positive integer. Then*

a) $H_{K,V}^n$ is essentially self-adjoint on

$$D_0 = C^\infty(H_{0,V})$$

b) $D(H_{K,V}^n) = D((H_0^f + H_1)^n)$

c) $C^\infty(H_{K,V}) = C^\infty(H_{0,V})$

d) The spectrum of $H_{K,V}$ is bounded below and consists of isolated eigen values with finite multiplicity.

e) Let $A(f)$ be one of the fields at $t=0$ and $A(f,t) = \exp iH_{K,V}t\, A(f) \exp -iH_{K,V}t$. Then for any vector $\Omega \in D_0$, and all positive integers k

$$\frac{d^k}{dt^k} A(f,t)\,\Omega \in D_0\,.$$

The proof of the theorem will be arrived at in stages. The first is a general Hilbert space argument which will later enable us to pass from the essential self-adjointness of (H_0^f) and $(H_1)^n$ on their respective domains to the essential self-adjointness of $(H_0^f + H_1)^n$.

Lemma 3.2. *Let A_1 and A_2 be hermitean operators in Hilbert spaces $\mathscr{H}_1$ and $\mathscr{H}_2$ respectively. Suppose that for some integer $n \geqq 1$, A_1^n and A_2^n are essentially self-adjoint on domains D_1 and D_2 respectively. Then $C = (A_1 \otimes 1 + 1 \otimes A_2)^j$ is essentially self-adjoint on $D_1 \otimes D_2$ for $1 \leqq j \leqq n$.*

Remarks. 1) By convention, $D_1 \otimes D_2$ stands for the (algebraic) tensor product of D_1 and D_2,i.e. the subset of the full tensor product $\mathscr{H}_1 \otimes \mathscr{H}_2$ consisting of finite linear combinations of vectors of the form $v_1 \otimes v_2$ with $v_1 \in D_1, v_2 \in D_2$.

2) The lemma has an easy generalization to the case of k hermitean operators $A_1 \ldots A_k$ acting on domains $D_1 \ldots D_k$ contained respectively in Hilbert spaces $\mathscr{H}_1 \ldots \mathscr{H}_k$. From the hypothesis $A_1^n, \ldots, A_k^n$ essentially

self-adjoint for some positive integer n, one concludes

$$C=(A_1\otimes 1\otimes\cdots\otimes 1+1\otimes A_2\otimes\cdots\otimes 1+\cdots 1\otimes\cdots\otimes A_n)^j$$

is essentially self-adjoint on $D_1\otimes\cdots\otimes D_k$ for $1\leqq j\leqq n$. We give the proof for the case $k=2$ for simplicity. We will have occasion to use the case $k=3$ also.

3) The hypotheses of the lemma do not include the requirement that A_1 and A_2 be essentially self-adjoint. That property follows from the following elementary argument. Suppose A is a hermitean operator and that for some integer $n\geqq 2$, at least one of the deficiency indices of A^n is zero. Then, A^n-i, say, has dense range. Now if a_k are the n^{th} roots of i, the polynomial identity $z^n-i=\prod_{k=1}^{n}(z-a_k)$ holds and therefore for each integer k, $1\leqq k\leqq n$:

$$A^n-i=(A-a_k)\prod_{j\neq k}(A-a_j)$$

This shows that $A-a_k$ has dense range. Since for $n\geqq 2$ at least one of the a_k is in the upper half-plane and at least one is in the lower, A is essentially self-adjoint. Thus in the proof of the lemma we may assume A_1 and A_2 essentially self-adjoint.

Proof. We will prove first that

$$D([(A_1\otimes 1+1\otimes A_2)^j]^-)\supset D((A_1^n)^-)\otimes D((A_2^n)^-)\,. \tag{3.1}$$

To this end, let Φ and Ψ be any two elements of $D((A_1^n)^-)$ and $D((A_2^n)^-)$ respectively. Then there exist sequences $\Phi_k\in D(A_1^n)$ and $\psi_k\in D(A_2^n)$ such that $\Phi_k\to\Phi$, $A_1^n\Phi_k\to(A_1^n)^-\Phi$, $\Psi_k\to\Psi$, and $A_2^n\Psi_k\to(A_2^n)^-\Psi$. If $\hat{A}_1$ and $\hat{A}_2$ are respectively the self-adjoint extensions of A_1 and A_2, $\hat{A}_1^n$ (resp. $\hat{A}_2^n$) is a self-adjoint extension of A_1^n (resp. A_2^n) and therefore $\hat{A}_1^n=(A_1^n)^-$ and $\hat{A}_2^n=(A_2^n)^-$. Now for $0\leqq j\leqq n$

$$\|\hat{A}_1^j\chi\|\leqq\|\hat{A}_1^n\chi\|+\|\chi\| \tag{3.2}$$

for all $\chi\in D((A_1^n)^-)$ and similarly for $\hat{A}_2^j$. (This is a consequence of the inequality $\lambda^{2j}\leqq\lambda^{2n}+1$, which obviously holds for all real λ, $1\leqq j\leqq n$, and the fact that by diagonalizing $\hat{A}_1$, one can convert (3.2) into the equivalent form

$$[\int\lambda^{2j}\mu(d\lambda)]^{1/2}\leqq[\int(\lambda^{2n}+1)\,\mu(d\lambda)]^{1/2}\leqq[\int\lambda^{2n}\mu(d\lambda)]^{1/2}+[\int\mu(d\lambda)]^{1/2}.)$$

Since $\|\Phi_k-\Phi\|\to 0$ and $\|\hat{A}_1^n(\Phi_k-\Phi)\|\to 0$, we must have $\|\hat{A}^j(\Phi_n-\Phi)\|\to 0$ i.e. $\hat{A}_1^j\Phi_n=A_1^j\Phi_n\to\hat{A}_1^j\Phi$. Similarly $A_2^j\Psi_n\to\hat{A}_2^j\Psi$. It follows therefore that $(A_1\otimes 1+1\otimes A_2)^j\,\Phi_n\otimes\Psi_n\to(\hat{A}_1\otimes 1+1\otimes\hat{A}_2)^j\,\Phi\otimes\Psi$. Hence $\Phi\otimes\Psi\in D([(A_1\otimes 1+1\otimes A_2)^j]^-)$ which completes the proof of (3.1). Further-

more, we have

$$[(A_1 \otimes 1 + 1 \otimes A_2)^j]^- \Phi \otimes \Psi = (\hat{A}_1 \otimes 1 + 1 \otimes \hat{A}_2)^j \Phi \otimes \Psi \quad 1 \leqq j \leqq n. \quad (3.3)$$

We can now easily show that $[(A_1 \otimes 1 + 1 \otimes A_2)^j]^-$ is self-adjoint. Choose a vector $\Phi \in D((A_1^n)^-)$ and a vector $\Psi \in D((A_2^n)^-)$ such that their supports are compact with respect to the spectral resolution of the operators $(A_1^n)^-$ and $(A_2^n)^-$ respectively. Then Φ is an analytic vector for $(A_1^n)^-$, Ψ is an analytic vector for $(A_2^n)^-$ and $\Phi \otimes \Psi$ is an analytic vector for $[(A_1 \otimes 1 + 1 \otimes A_2)^j]^-$. (The last statement is clear, as in the case of (3.2) when $\hat{A}_1$ and $\hat{A}_2$ are diagonalized.) Since the linear span of vectors of the form $\Phi \otimes \Psi$ is dense, $[(A_1 \otimes 1 + 1 \otimes A_2)^j]^-$ has a dense set of analytic vectors and is therefore self-adjoint by Nelson's theorem [6]. (3.3) shows, in fact, that

$$[(A_1 \otimes 1 + 1 \otimes A_2)^j]^- = (\hat{A}_1 \otimes 1 + 1 \otimes \hat{A}_2)^j \quad 1 \leqq j \leqq n, \quad (3.4)$$

on the domain of the right handside. The proof is complete.

The next four lemmas yield inequalities enabling us to estimate the terms described in the introduction just after equation (1.2). It is convenient to introduce some terminology.

Definition 3.3. Let f and g be two complex-valued functions defined on the same set S. Then *f dominates g* if there exist constants A and B such that

$$|g(x)| \leqq A|f(x)| + B \quad (3.5)$$

for all $x \in S$.

f strongly dominates g if for every $A > 0$ there is a $B > 0$ such that (3.5) holds for all $x \in S$.

Each measurable function f on $\mathbb{R}^s$ defines an operator of pointwise multiplication on some subset of $L^2(\mathbb{R}^s; dx): \Phi \to f\Phi$ with $(f\Phi)(x) = f(x)\Phi(x)$. In the four lemmas that follow all scalar products and norms are to be taken in $L^2(\mathbb{R}^s; dx)$.

Lemma 3.4. *Let f and g be infinitely differentiable polynomially bounded functions on $\mathbb{R}^n$. Suppose*

1) *f dominates each of its derivatives,*

2) *f dominates g and every derivative of g.*

Then for any multi-index $\sigma = (\sigma_1, \ldots, \sigma_s)$, there are constants A, B, C, D such that

$$\|D^\sigma g \Psi\| \leqq A\|(-\Delta)^{\frac{|\sigma|}{2}} f\Psi\| + B\|(-\Delta)^{\frac{|\sigma|}{2}} \Psi\| + C\|f\Psi\| + D\|\Psi\| \quad (3.6)$$

for all $\Psi \in \mathscr{S}(\mathbb{R}^s)$, the space of infinitely differentiable functions of fast decrease.

Moreover, if we replace 2) by

2') f strongly dominates g and dominates every derivative of g,

then A can be taken as any strictly positive real number when B, C, D, are suitably chosen.

Proof. We proceed by induction on $|\sigma|$. The lemma is clearly true when $|\sigma| = 0$, because, from $|g| \leqq A|f| + B$ there follows by an elementary calculation

$$\|g\Psi\| \leqq A\|f\Psi\| + B\|\Psi\|\,, \tag{3.7}$$

which is (3.6) for this case. To carry the induction from $|\sigma| - 1$ to $|\sigma|$, we distinguish the cases $|\sigma|$ even and odd. For $|\sigma|$ odd, we can write $D^\sigma = D_j D^{\sigma'}$ for some j where $|\sigma'| = |\sigma| - 1$ is even. Then

$$\begin{aligned}\|D^\sigma g\Psi\| &\leqq \|(-\Delta)^{(|\sigma|-1)\frac{1}{2}} D_j g\Psi\| \\ &\leqq \|g(-\Delta)^{(|\sigma|-1)\frac{1}{2}} D_j\Psi\| + \|[g,(-\Delta)^{(|\sigma|-1)\frac{1}{2}} D_j]\,\Psi\|\end{aligned} \tag{3.8}$$

The first term on the right hand side is

$$\leqq C_1\|f(-\Delta)^{(|\sigma|-1)\frac{1}{2}} D_j\Psi\| + C_2\|(-\Delta)^{(|\sigma|-1)\frac{1}{2}} D_j\Psi\| \tag{3.9}$$

by the same argument that led to (3.7). The first term on the right hand side of (3.9) in turn becomes

$$\leqq C_1[\|D_j(-\Delta)^{(|\sigma|-1)\frac{1}{2}} f\Psi\| + \|[f,(-\Delta)^{(|\sigma|-1)\frac{1}{2}} D_j]\,\Psi\|]\,. \tag{3.10}$$

Using

$$\|D_j(-\Delta)^{(|\sigma|-1)\frac{1}{2}} f\Psi\| \leqq \|(-\Delta)^{|\sigma|/2} f\Psi\| \tag{3.11}$$

to estimate the first term on the right hand side of (3.10), and similarly for the second term of (3.9) we have altogether

$$\begin{aligned}\|D^\sigma g\Psi\| \leqq\; & C_1\|(-\Delta)^{|\sigma|/2} f\Psi\| + C_2\|(-\Delta)^{|\sigma|/2}\Psi\| \\ & + C_1\|[f,(-\Delta)^{(|\sigma|-1)\frac{1}{2}} D_j]\,\Psi\| \\ & + \|[g,(-\Delta)^{(|\sigma|-1)\frac{1}{2}} D_j]\,\Psi\|\,.\end{aligned} \tag{3.12}$$

Here the constant C_1 can be taken arbitrarily small if 2') holds. (In the course of this argument we have several terms replaced derivatives with the appropriate powers of $-\Delta$. An elementary justification for this is obtained by passing to the Fourier transform. The required inequalities then follow from the polynomial inequalities $|x^\sigma| \leqq \left(\sum_j x_j^2\right)^{|\sigma|/2}$.)

We have still to deal with the commutator terms in (3.12). Here we use the fact that

$$[f,(-\Delta)^{(|\sigma|-1)\frac{1}{2}} D_j] = \sum_{|\tau|<|\sigma|} D^\tau h_\tau \tag{3.13}$$

where the h_τ are sums of derivatives of f and hence by assumption 1) dominated by f. We can therefore apply the induction hypothesis to show

$$\|[f,(-\Delta)^{(|\sigma|-1)\frac{1}{2}}D_j]\,\Psi\| \leqq \sum_{|\tau|<|\sigma|}\{A_\tau\|(-\Delta)^{|\tau|/2}f\Psi\| + B_\tau\|(-\Delta)^{|\tau|/2}\Psi\| + C_\tau\|f\Psi\| + D_\tau\|\Psi\|\}\,.$$

Now it is easy to show that

$$\|(-\Delta)^{|\tau|/2}f\Psi\| \leqq K\|(-\Delta)^{|\sigma|/2}f\Psi\| + K'\|f\Psi\| \tag{3.14}$$

and that K may be chosen as small as desired. (Again pass to the Fourier transform and use elementary inequalities for polynomials). Similarly, $\|(-\Delta)^{|\tau|/2}\Psi\|$ is majorized by $\|(-\Delta)^{|\sigma|/2}\Psi\| + \|\Psi\|$. Thus, we get

$$\begin{aligned}\|[f,(-\Delta)^{(|\sigma|-1)\frac{1}{2}}D_j]\,\Psi\| \leqq{}& A'\|(-\Delta)^{|\sigma|/2}f\Psi\| \\ &+ B'\|(-\Delta)^{|\sigma|/2}\Psi\| + C'\|f\Psi\| + D'\|\Psi\|\end{aligned} \tag{3.15}$$

where A' may be taken as small as desired. In a similar way, we get a majorization for

$$\|[g,(-\Delta)^{(|\sigma|-1)\frac{1}{2}}D_j]\,\Psi\|$$

whose right hand side is of the form of the right hand side of (3.15). Combining this, (3.15) and (3.12) we have an estimate of the desired form (3.6).

When $|\sigma|$ is even the induction is slightly easier because instead of (3.8) one can write

$$\|D^\sigma g\Psi\| \leqq \|(-\Delta)^{|\sigma|/2}g\Psi\| \leqq \|g(-\Delta)^{|\sigma|/2}\Psi\| + \|[g,(-\Delta)^{|\sigma|/2}]\,\Psi\|$$

while (3.9) becomes

$$\leqq C_1\|f(-\Delta)^{|\sigma|/2}\Psi\| + C_2\|(-\Delta)^{|\sigma|/2}\Psi\|$$

and (3.10)

$$\leqq C_1[\|(-\Delta)^{|\sigma|/2}f\Psi\| + \|[f,(-\Delta)^{|\sigma|/2}]\,\Psi\|]\,,$$

which yields directly the analogue of (3.12)

$$\begin{aligned}\|D^\sigma g\Psi\| \leqq{}& C_1\|(-\Delta)^{|\sigma|/2}f\Psi\| + C_2\|(-\Delta)^{|\sigma|/2}\Psi\| \\ &+ C_1\|[f,(-\Delta)^{|\sigma|/2}]\,\Psi\| + \|[g,(-\Delta)^{|\sigma|/2}]\,\Psi\|\,.\end{aligned}$$

Because $|\sigma|$ is even, we may replace (3.13) with

$$[f,(-\Delta)^{|\sigma|/2}] = \sum_{|\tau|<|\sigma|} D^\tau h_\tau \tag{3.16}$$

and the argument goes just as before. (It is just this equality which would not be valid if $|\sigma|$ were odd.) This completes the induction and the proof of the lemma.

What kinds of functions dominate all their derivatives? It is easy to see that not every polynomial does. Take, for example, $P = x_1^2 - x_2^2$. It vanishes on the line $x_1 = x_2$ while $D_1 P = 2x_1$ grows there. Thus, it does not dominate its derivatives. On the other hand, $P = x_1^2 + x_2^2 + 1$ is easily seen to dominate its derivatives. More generally, we have

Lemma 3.5. *Let $P_1, \dots P_J$ be polynomials respectively of degree $2n_1, \dots, 2n_J$ on $\mathbb{R}^{m_1}, \mathbb{R}^{m_2}, \dots \mathbb{R}^{m_J}$. Suppose each P_j is everywhere $\geqq 1$ and that there is a constant, $\varrho > 0$, such that*

$$P_j(x) \geqq \varrho |x|^{2n_j}, \quad j = 1, \dots J. \tag{3.17}$$

Suppose V is the polynomial on $\mathbb{R}^m = \mathbb{R}^{m_1 + m_2 + \dots m_J} = \mathbb{R}^{m_1} \oplus \mathbb{R}^{m_2} \oplus \cdots \oplus \mathbb{R}^{m_J}$ defined by

$$V(x_1 \dots x_J) = P_1(x_1) + \cdots + P_J(x_J). \tag{3.18}$$

If $\beta > 0$ then $W(x) = [V(x)]^\beta$ strongly dominates all its derivatives.

Proof. One easily shows by induction that

$$D^\sigma W(x) = \sum_{k=1}^{|\sigma|} [V(x)]^{\beta - k} Q_k(x)$$

where $Q_k(x)$ is a product of k factors each of which is a derivative of $V(x)$ of order greater than zero. Thus

$$D^\sigma W(x) = W(x) \sum_{k=1}^{|\sigma|} \sum_{\tau_1 \dots \tau_k} \alpha_k(\tau_1 \dots \tau_k) \left(\frac{D^{\tau_1} V}{V}\right) \cdots \left(\frac{D^{\tau_k} V}{V}\right)$$

where the α's are real numbers. Now $D^{\tau_j} V = 0$ if the multi-index τ_j has a non-zero value for two different groups of variables. Thus a non-zero contribution is always of the form

$$\frac{D^{\tau_j} P_k}{V} \leqq \frac{D^{\tau_j} P_k}{\varrho \sum_{l=1}^{J} |x_l|^{2n_l}}.$$

Since the numerator always has degree less than $2n_k$ this approaches zero as $|x| \to \infty$. Thus $\dfrac{D^{\tau_j} V}{V} \to 0$ as $|x| \to \infty$ if $|\tau_j| \neq 0$, and therefore W strongly dominates $D^\sigma W$.

It is worth remarking in passing that the assumptions of this Lemma 3.5 imply $W \geqq 1$, and our principal application of Lemma 3.4 will be to

a case in which $f = W$. The statement and proof of Lemma 3.4 can be somewhat simplified in this case by using $\|\Psi\| \leqq \|f\Psi\|$. In particular, the inequality (3.6) is replaced by

$$\|D^\sigma g\Psi\| \leqq A\|(-\Delta)^{|\sigma|/2} f\Psi\| + B\|(-\Delta)^{|\sigma|/2}\Psi\| + C\|f\Psi\| . \quad (3.19)$$

The fact established in Lemma 3.5, that W strongly dominates all its derivatives will now be used to bound expressions of the form $\|(-\Delta)^{j/2} V^{n-j/2}\Psi\|^2$ by a multiple of $\|(-\Delta + V)^n\Psi\|^2$.

Lemma 3.6 [1]. *For each positive integer n, there is a constant b (depending on n) such that for all* $\Psi \in \mathscr{S}(\mathbb{R}^m)$

$$\sum_{j=0}^{2n} \binom{2n}{j} \|(-\Delta)^{j/2} V^{n-j/2}\Psi\|^2 \leqq b\|(-\Delta + V)^n\Psi\|^2 . \quad (3.20)$$

Proof. Denote the left hand side of (3.20) by Σ. Since

$$\Sigma = \sum_{j=0}^{2n} \binom{2n}{j} (\Psi, V^{n-j/2}(-\Delta)^j V^{n-j/2}\Psi) \quad (3.21)$$

and

$$\begin{aligned}\|(-\Delta + V)^n\Psi\|^2 &= (\Psi, (-\Delta + V)^{2n}\Psi) \\ &= \sum_{j=1}^{2n} \binom{2n}{j} (\Psi, V^{n-j/2}(-\Delta)^j V^{n-j/2}\Psi) \\ &\quad + \text{commutator terms},\end{aligned}$$

to prove the lemma it suffices to show that the commutator terms can be majorized in absolute value by

$$\varepsilon \sum_{j=0}^{2n} \binom{2n}{j} \|(-\Delta)^{j/2} V^{n-j/2}\Psi\|^2 + d\|\Psi\|^2$$

with $\varepsilon < 1$ because then

$$\|(-\Delta + V)^n\Psi\|^2 \geqq (1-\varepsilon)\Sigma - d\|\Psi\|^2$$

and, since $(-\Delta + V) \geqq 1$

$$\Sigma \leqq (1-\varepsilon)^{-1}(1+d)\|(-\Delta + V)^n\Psi\|^2 .$$

In fact, we will show that ε can be taken arbitrarily small for sufficiently large d.

The commutator terms can be expressed as a sum of expressions of the form $(\Psi, V_1 D^\tau V_2 \Psi)$ where V_1, V_2 are each products of $2n$-j factors each of which is either $V^{1/2}$ or a derivative of $V^{1/2}$, and where $|\tau| < 2j$, $1 \leqq j \leqq 2n-1$. By Schwarz's inequality, the above expression is majorized

in absolute value by

$$\|D^{\tau_1} V_1 \Psi\| \, \|D^{\tau_2} V_2 \Psi\| \quad \text{where} \quad D^{\tau} = D^{\tau_1} D^{\tau_2}, \quad \text{and} \quad |\tau_1| < j, |\tau_2| \leqq j .$$

This expression in turn is majorized by

$$\frac{1}{2\eta} \|D^{\tau_1} V_1 \Psi\|^2 + 2\eta \|D^{\tau_2} V_2 \Psi\|^2 \tag{3.22}$$

where η is any number > 0. It therefore suffices to prove that there are constants A and B such that

$$\|D^{\tau_2} V_2 \Psi\|^2 \leqq A \, \Sigma + B \|\Psi\|^2 \tag{3.23}$$

and that, for any $A' > 0$ there exists a B' such that

$$\|D^{\tau_1} V_1 \Psi\|^2 \leqq A' \Sigma + B' \|\Psi\| \tag{3.24}$$

for all $\Psi \in \mathscr{S}$.

Let us prove (3.24) and then indicate the changes necessary to obtain a proof of (3.23). Lemma 3.5 tells us that $V^{n-j/2}$ majorizes all its derivatives. Furthermore, it majorizes V_1 and every derivative of V_1. (The proof of this last runs exactly parallel to that of Lemma 3.5 itself.) Now we apply Lemma 3.4 to $D^{\tau_1} V_1$ and $V^{n-j/2}$. It asserts that there exist constants C, D, E, F such that

$$\begin{aligned} \|D^{\tau_1} V_1 \Psi\| \leqq C &\|(-\Delta)^{|\tau_1|/2} V^{n-j/2} \Psi\| \\ &+ D \|(-\Delta)^{|\tau_1|/2} \Psi\| \\ &+ E \|V^{n-j/2} \Psi\| \\ &+ F \|\Psi\| . \end{aligned}$$

Squaring this inequality and applying $|ab| \leqq \frac{1}{2}((a|^2 + (b|^2)$ to the cross terms appearing on the right hand side we get the same inequality with new set of constants C, D, E, F and $\| \, \|$ everywhere replaced by $\| \, \|^2$:

$$\begin{aligned} \|D^{\tau_1} V_1 \Psi\|^2 \leqq C &\|(-\Delta)^{|\tau_1|/2} V^{n-j/2} \Psi\|^2 \\ &+ D \|(-\Delta)^{|\tau_1|/2} \Psi\|^2 \\ &+ E \|V^{n-j/2} \Psi\|^2 \\ &+ F \|\Psi\|^2 . \end{aligned} \tag{3.25}$$

Since $|\tau_1|/2 < j/2$, for any $A' > 0$ we can find E', F', and F'' such that

$$\|(-\Delta)^{|\tau_1|/2} V^{n-j/2} \Psi\|^2 \leqq A' \|(-\Delta)^{j/2} V^{n-j/2} \Psi\|^2 + E' \|V^{n-j/2} \Psi\|^2 \tag{3.26}$$

$$\|(-\Delta)^{|\tau_1|/2} \Psi\|^2 \leqq A' \|(-\Delta)^n \Psi\|^2 + F' \|\Psi\|^2 \tag{3.27}$$

$$(E + E') \|V^{n-j/2} \Psi\|^2 \leqq A' \|V^n \Psi\|^2 + F'' \|\Psi\|^2 \tag{3.28}$$

Therefore

$$\begin{aligned}\|D^{\tau_1} V_1 \Psi\|^2 &\leqq A'\{\|(-\Delta)^{j/2} V^{n-j/2}\Psi\|^2 \\ &\quad + \|(-\Delta)^n \Psi\|^2 + \|V^n \Psi\|^2\} \\ &\quad + (F + F' + F'')\,\|\Psi\|^2 \\ &\leqq A'\Sigma + (F + F' + F'')\,\|\Psi\|^2 .\end{aligned}$$

The proof of (3.23) differs only in that because $|\tau_2| \leqq j$, A' cannot necessarily be taken arbitrarily small in (3.26), (3.27), and (3.28).

Lemma 3.7. *Let $\sigma = (\sigma_1 \ldots \sigma_s)$ and $\tau = (\tau_1 \ldots \tau_s)$ be multi-indices. Let x^τ be a monomial on $\mathbb{R}^s = \mathbb{R}^{s_1} \oplus \mathbb{R}^{s_2} \oplus \ldots \mathbb{R}^{s_J}$, whose degree in the variables of $\mathbb{R}^{s_j}$ is α_j, $l_j = 1, \ldots, J$.*

Suppose

$$\frac{|\sigma|}{2} + \sum_{j=1}^{J} \frac{\alpha_j}{2n_j} < n \tag{3.29}$$

for some fixed integer n.

Then the operator $D^\sigma x^\tau$ regarded as defined on $\mathscr{S}(\mathbb{R}^s)$ is infinitely small in the sense of T. Kato with respect to $(-\Delta + V)^n$, V being defined by (3.18).

Proof. The first step will be to prove that x^τ is strongly dominated by $V^{n-\frac{|\sigma|}{2}}$ whenever the condition (3.29) holds. In fact, if we set $r = \left[\sum_{j=1}^{N} |x_j|^{2n_j}\right]^{1/2}$, $|x_j|$ being the norm in $\mathbb{R}^{s_j}$ we have by assumption $V(x) \geqq \text{const}\, r^2$ and because the individual components x_j, belonging to $\mathbb{R}^{s_j}$ satisfy $|x_j| \leqq r^{1/n_j}$ we have also $|x^\tau| \leqq r^{\sum_{j=1}^{J} (\alpha_j/n_j)}$. Thus as $|x| \to \infty$, $|x^\tau|$ grows at most like $r^{\sum_{j=1}^{J}(\alpha_j/n_j)}$ while $V^{n-\frac{|\sigma|}{2}}$ grows at least like $r^{2n-|\sigma|}$. Therefore, $x^\tau [V]^{\frac{\sigma}{2}-n} \to 0$ as $|x| \to \infty$, so x^τ is strongly dominated by $V^{n-|\sigma|/2}$.

Since differentiating x^τ simply gives a multiple of a lower power of x, every derivative of x^τ is also dominated by $V^{n-|\sigma|/2}$.

To complete the proof we apply the preceding lemmas. Suppose $\varepsilon > 0$ is given. By Lemma 3.4 there exist constants B, C, D such that

$$\begin{aligned}\|D^\sigma x^\tau \Psi\| \leqq \varepsilon \|(-\Delta)^{|\sigma|/2} V^{n-|\sigma|/2}\Psi\| + B\|(-\Delta)^{|\sigma|/2}\Psi\| \\ + C\|V^{n-|\sigma|/2}\Psi\| + D\|\Psi\|\end{aligned}$$

for all $\Psi \in \mathscr{S}(\mathbb{R}^m)$. We can also evidently find D' and D'' such that

$$\begin{aligned}B\|(-\Delta)^{|\sigma|/2}\Psi\| &\leqq \varepsilon\|(-\Delta)^n \Psi\| + D'\|\Psi\| \\ C\|V^{n-|\sigma|/2}\Psi\| &\leqq \varepsilon\|V^n\Psi\| + D''\|\Psi\| .\end{aligned}$$

Thus,

$$\|D^\sigma x^\tau \Psi\| \leqq \varepsilon[\|(-\Delta)^{|\sigma|/2}\, V^{n-|\sigma|/2}\, \Psi\| + \|(-\Delta)^n \Psi\| + \|V^n \Psi\|] + (D + D' + D'')\, \|\Psi\| .$$

But from Lemma 3.6 it follows that there exist constants A, D'' such that

$$\|(-\Delta)^{|\sigma|/2}\, V^{n-|\sigma|/2}\, \Psi\| + \|(-\Delta)^n \Psi\| + \|V^n \Psi\| \leqq A\, \|(-\Delta + V)^n \Psi\| + D'''\, \|\Psi\|$$

thus,

$$\|D^\sigma x^\tau \Psi\| \leqq \varepsilon\, \|(-\Delta + N)^n \Psi\| + (D + D' + D'' + \varepsilon D''')\, \|\Psi\| .$$

Since ε is any positive number and, the choice of A does not depend on ε, this proves the lemma.

This completes the proof of the preliminary lemmas. Their application is based on a theorem of Kato.

Theorem 3.8 (Kato) *Let A be a linear operator on a Hilbert space $\mathscr{H}$. Suppose that A is essentially self-adjoint on the domain $D(A)$ and B is a hermitean operator such that*

a) $D(B) \supset D(A)$.

b) *For each $\varepsilon > 0$ there exists a b such that*

$$\|B\Phi\| \leqq \varepsilon\, \|A\Phi\| + b\, \|\Phi\| \tag{3.30}$$

for all $\Phi \in D(A)$. Then the closures B^- and A^- satisfy

$$\|B^-\Phi\| \leqq \varepsilon\, \|A^-\Phi\| + b\, \|\Phi\| \tag{3.31}$$

for all $\Phi \in D(A^-)$. Furthermore, $A^- + B^-$ is self-adjoint and

$$D(A^- + B^-) = D(A^-) . \tag{3.32}$$

(For a proof see [3], Chapter V, § 4.)

For the application to the present case, we follow a line of argument similar to that developed for the $\lambda\phi^4$ theory in [1], in an elegant form due to J. Cannon [4].

For brevity, we introduce the following notation. If A and B are two linear operators, $B < A$ if $D(A) \subset D(B)$ and (3.30) holds for all $\Phi \in D(A)$. With this notation we have the following two lemmas.

Lemma 3.9 (Cannon). *The set of all B such that $B < A$ is a complex vector space*

$$A < A^2 \quad \text{implies} \quad A < A^2 < A^3 < \cdots .$$

If A is self-adjoint $A < A^2$.

Lemma 3.10 (Cannon). *Suppose that for some operator B*

$$(\mathrm{Ad}\, A)^k (B) < A^{k+1} \qquad k = 0, 1, \ldots \tag{3.33}$$

Then if $l_1, l_2, \ldots$ *and* $l_1', l_2', \ldots$ *are non-negative integers such that* $\Sigma l_i + \Sigma l_i' < l$,

$$A^{l_1} B^{l_1'} A^{l_2} B^{l_2'} \cdots < A^l . \tag{3.34}$$

(For proofs, see [4]).

Proof of Theorem 3.1. We take $B = H_2$ and $A = H_0^f + H_1$. Recall that H_0^f is the free fermion Hamiltonian and the Hamiltonian H_1 is a sum of the free boson Hamiltonian, H_0^b and the dominant boson self-coupling. For each positive integer n, $(H_0^f)^n$ is essentially self-adjoint on the domain $C^\infty(H_0^f)$ in the $\Phi o \kappa$ space of the fermions, while H_1^n is essentially self-adjoint on $C^\infty(H_0^b)$ in the $\Phi o \kappa$ space of the bosons. The first of these statements is elementary; the second is a basic result of [1]. It follows from Lemma 3.2 that, in the notation of that lemma, $(H_0^f \otimes 1 + 1 \otimes H_1)^n$ is essentially self-adjoint on $C^\infty(H_0^f) \otimes C^\infty(H_0^b)$ and therefore certainly on $C^\infty(H_{0,V})$ which includes it. (The $\otimes 1$ and $1 \otimes$ will be suppressed whenever it is convenient so the operator under consideration may also be denoted $(H_0^f + H_1)^n$.)

Next we write

$$H_{K,V}^n = (H_0^f + H_1)^n + [H_{K,V}^n - (H_0^f + H_1)^n] \tag{3.35}$$

and study the term in square brackets on the right hand side. It is a sum of monomials in $(H_0^f + H_1)$ and H_2 of precisely the form of the left hand side of (3.34), when $l = n$. Thus to apply Lemma 3.10, we have only to verify

$$[\mathrm{Ad}(H_0^f + H_1)]^k (H_2) < (H_0^f + H_1)^{k+1} . \tag{3.36}$$

At this stage, for clarity, we indicate explicitly the action of operators as tensor products $A \otimes A' \otimes A''$, where the first factor, A, acts on the coupled boson modes, the second, A', on the uncoupled boson modes and the third, A'', on the fermion modes. With this notation H_1 is re-written as

$$H_3 \otimes 1 \otimes 1 + 1 \otimes H_4 \otimes 1 \tag{3.37}$$

where H_4 is the free Hamiltonian of the uncoupled boson modes, H_3 is the free Hamiltonian of the coupled boson modes plus the dominant boson self-interaction. Similarly, the notation, H_0^f, for the free fermion Hamiltonian is replaced by $(1 \otimes 1 \otimes H_0^f)$. Finally, H_2 is a sum of terms of the form $Q^\tau \otimes 1 \otimes F$ where F is a bounded operator.

Thus $[\mathrm{Ad}(H_0^f + H_1)]^k (H_2)$ is a sum of terms of the form

$$[\mathrm{Ad}(-\Delta)]^k (Q^\tau) \otimes 1 \otimes [\mathrm{Ad}(H_0^f)]^k (F) . \tag{3.38}$$

By assumption (2.11), the last factor is a bounded operator, F_k. The first is a sum of terms of the form $D^\sigma Q^{\tau'}$ where $|\tau'| < |\tau|$ and $|\sigma| \leqq 2k$. Now by

assumption τ is such that the criterion (2.10) is satisfied. That implies $Q^\tau < H_3$ by Lemma 3.7. (Set $n=1, |\sigma|=0$ there.) Then using $|\sigma| \leqq 2k, |\tau'| < |\tau|$ we get that (3.29) is satisfied with $n = k+1$ so $D^{\tau'} Q^\sigma < H_3^{k+1}$ again by Lemma 3.7. It remains only to show from this that (3.38) $< (H_0^f + H_1)^{k+1}$. The argument goes as follows. Vectors of $C^\infty(H_3) \otimes C^\infty(H_4) \otimes C^\infty(H_0^f)$ (algebraic tensor product) can be wirtten as finite sums $\chi = \sum_i \Phi_i \otimes \Psi_i$ where $\Phi_i \in C^\infty(H_3)$ and $\Psi_i \in C^\infty(H_4) \otimes C^\infty(H_0^f)$. Without loss of generality the set $\{\Psi_i\}$ may be assumed orthonormal. Noting $D^{\sigma'} Q^{\tau'} \otimes 1 \otimes F_k = (1 \otimes 1 \otimes F_k)(D^\sigma Q^{\tau'} \otimes 1 \otimes 1)$ and $\|1 \otimes 1 \otimes F_k\| = \|F_k\|$, we have therefore

$$\begin{aligned} \|(D^\sigma Q^{\tau'} \otimes 1 \otimes F_k)\chi\| &\leqq \|F_k\| \left\| \sum_i D^\sigma Q^{\tau'} \Phi_i \otimes \Psi_i \right\| \\ &\leqq \|F_k\| \left[\varepsilon \left\| (H_3^{k+1} \otimes 1) \left(\sum_i \Phi_i \otimes \Psi_i \right) \right\| + b\|\chi\| \right] \qquad (3.39) \\ &\leqq \varepsilon \|F_k\| \, \|(H_0^f + H_1)^{k+1} \chi\| + b\|F_k\| \, \|\chi\| . \end{aligned}$$

All but the last of these steps are elementary. Putting aside its justification for a moment, one can pass from (3.39) to the desired inequality valid everywhere on the domain of $(H_0^f + H_1)^{k+1}$ by closure. Since $\|F_k\|$ is independent of χ and ε can be chosen arbitrarily small, this shows (3.38) $< (H_0^f + H_1)^{k+1}$.

The last step of (3.39) follows from the fact that if A and B are essentially self-adjoint positive operators, commuting on $C^\infty(A) \cap C^\infty(B)$, then

$$\|A\Phi\| \leqq \|(A+B)\Phi\| \qquad (3.40)$$

for all $\Phi \in C^\infty(A) \cap C^\infty(B)$. To see this one notes that under the same assumptions AB is positive $((\Phi, AB\Phi) = \|A^{1/2} B^{1/2} \Phi\|^2)$ so

$$(\Phi, A^2 \Phi) \leqq (\Phi, (A^2 + B^2 + 2AB)\Phi) . \qquad (3.41)$$

The square root of this inequality is (3.40).

This completes the proof that all monomials appearing when $H_{K,V}^n - (H_0^f + H_1)^n$ is expanded satisfy the hypotheses of Lemma 3.10. Thus each such monomial is $< (H_0^f + H_1)^n$. Since, by Lemma 3.9, the operators $< (H_0^f + H_1)^n$ form a vector space, we have

$$H_{K,V}^n - (H_0^f + H_1)^n < (H_0^f + H_1)^n$$

and therefore, by Kato's Theorem 3.8, $H_{K,V}^n$ is essentially self-adjoint on $D((H_0^f + H_1)^n)$ and $D(H_{K,V}^n) = D((H_0^f + H_1)^n)$. This proves b) of Theorem 3.1. a) follows from Kato's Theorem 3.8 and the fact proved above that $(H_0^f + H_1)^n$ is essentially self-adjoint on $C^\infty(H_{0,V})$.

c) As a consequence of b) $C^\infty(H_{K,V}) = C^\infty(H_0^f + H_1)$. Thus, to prove c) it suffices to establish $C^\infty(H_0^f + H_1) = C^\infty(H_{0,V})$, that is, the addition

of the dominant boson self-interaction to the free Hamiltonian $H_0^f + H_0^b = H_{0,V}$ does not change the intersection of the domains of all the powers of the free Hamiltonian. Now one of the basic results of [1] was that $C^\infty(H_1) = C^\infty(H_0^b)$. Thus, the problem of proving c) is reduced to showing that the addition of H_0^f to H_1 and H_0^b does not affect this relation. That in turn is a consequence of

$$D((H_0^f \otimes 1 + 1 \otimes H_1)^n) = D((H_0^f \otimes 1)^n) \cap D((1 \otimes H_1)^n) \tag{3.42}$$

and similarly

$$D((H_0^f \otimes 1 + 1 \otimes H_0^b)^n) = D((H_0^f \otimes 1)^n) \cap D((1 \otimes H_0^b)^n). \tag{3.43}$$

(We have reinstated the tensor product notation for clarity). (3.42) and (3.43) follow from the positivity of H_0^f, H_0^b, and H_1. That may be seen as follows. We have, surely, that the left hand sides of (3.42) and (3.43) include the right hand sides. Furthermore,

$$D((H_0^f \otimes 1 + 1 \otimes H_1)^n) \supset D((H_0^f)^n) \otimes D((H_1)^n) \tag{3.44}$$

(again algebraic tensor product). On vectors velonging to the right hand side

$$(H_0^f \otimes 1 + 1 \otimes H_1)^n = \sum_{k=0}^{n} \binom{n}{k} (H_0^f)^k \otimes (H_1)^{n-k} \tag{3.45}$$

and because all operators occuring are positive and commute

$$\left\| \binom{n}{k} (H_0^f)^k \otimes (H_1)^{n-k} (\Phi_l - \Phi_m) \right\| \leqq \|(H_0^f \otimes 1 + 1 \otimes H_1)^n (\Phi_l - \Phi_m)\|$$

for any Cauchy sequence (Φ_j) of vectors belonging to the right hand side of (3.44). Thus, passing to the closure, we find that $\Phi \in D((H_0^f \otimes 1 + 1 \otimes H_1)^n)$ implies $\Phi \in D((H_0^f \otimes 1)^n) \cap D((1 \otimes H_1)^n)$. There is an analogous argument with H_1 replaced by H_0^b so the proof of c) is complete.

d) Both H_0^f and H_1 have compact resolvents [2, 1]. Thus, their eigenvalues E_i^f and E_j respectively have finite multiplicity and cluster only to infinity. Their eigen functions Φ_i^f and Φ_j respectively are complete in the respective $\Phi \circ \kappa$ spaces. $H_0^f + H_1$ has eigen values $E_i^f + E_j$ with corresponding eigen functions $\Phi_i^f \otimes \Phi_j$. Since they are complete $H_0^f + H_1$ also has compact resolvent. Now as we have seen in the proof of a), H_2 is infinitely small relative to $H_0^f + H_1$. It is a general result of perturbation theory that the addition of such a perturbation preserves the compactness of the resolvent. (See [3], p. 214, Theorem 3.17). Thus d) is proved.

Now we prove e). b) tells us that $C^\infty(H_{0,V})$ is invariant under $H_{K,V}$ and $e^{iH_{K,V}t}$. (For $H_{K,V}$ this is obvious; for $e^{iH_{K,V}t}$, it is an easy consequence

of the spectral theorem.) Since, from the definition of $A(f, t)$,

$$\frac{d^n}{dt^n} A(f,t)\,\Omega = \left[iH_{K,V}, \frac{d^{n-1}}{dt^{n-1}} A(f,t) \right] \Omega, \tag{3.46}$$

we see that to prove e) it suffices to prove that $A(f)\Omega$ is in $C^\infty(H_{0,V})$ if Ω is. There are two things to verify: First that $D(A(f)) \supset C^\infty(H_{0,V})$ and second that $A(f)\, C^\infty(H_{0,V}) \subset C^\infty(H_{0,V})$. Both these statements require only elementary calculations in $\Phi\, o\, \kappa$ space. For all fields $A(f)$, boson or fermion, one finds that for each $\varepsilon > 0$ there is a b such that

$$\|A(f)\Phi\| \leqq [\varepsilon \|H_{0,V}\Phi\| + b\|\Phi\|]\, |f|_s \tag{3.47}$$

for all $\Phi \in D(H_{0,V})$. Here $|\ |_s$ is some norm on the test function space. For example, for fermion fields of spin $\frac{1}{2}$ it is the ordinary L^2 norm, while for the conjugate scalar boson field π,

$$|f|_s = [\int |\hat{f}(\boldsymbol{p})|^2\, [m^2 + p^2]^{1/2}\, d\boldsymbol{p}]^{1/2} \tag{3.48}$$

will do. From this inequality (3.47), one gets immediately $D(A(f)) \supset D(H_{0,V})$ and therefore a fortiori $D(A(f)) \supset C^\infty(H_{0,V})$. To see the invariance of $C^\infty(H_{0,V})$ under $A(f)$, note that the identity

$$H_{0,V} A(f) = A(f)\, H_{0,V} + A(f') \tag{3.49}$$

holds on, say, the vectors with only a finite number of non-zero components in $\Phi\, o\, \kappa$ space, all of which are infinitely differentiable and of compact support. Here f' is obtained from f by multiplying its Fourier transform by kinematical factors (e.g., $\omega(\boldsymbol{p})$ to some power) that depend on which field is under consideration. From (3.47) and (3.49) one gets, passing to the closure, $A(f)\, D(H_{0,V}^n) \subset D(H_{0,V}^{n-1})$ as desired.

Theorem 3.1 together with some of the information obtained in the course of its proof permit one to construct the basic objects, vacuum expectation values and Green's functions, in terms of which the content of a field theory is customarily expressed. Notice first that any eigenfunction Ω of H is certainly in $C^\infty(H_{K,V})$ and therefore any product of the fields smeared with test functions in $\mathscr{S}$ in the space variables is certainly applicable to it. Thus the expectation value

$$\left(\Omega, \prod_j A_j(f_j, t_j)\, \Omega\right) \tag{3.50}$$

is well defined. From c) we see that it is infinitely differentiable in the t's. It is clearly a multilinear functional of the f's. Theorem 3.1 e) implies that it is in fact a tempered distribution in each of these variables and thus by the nuclear theorem in all the space variables together. The time ordering operation carried out on the expressions (3.50) is unambigous

because they are smooth in the times; that shows the Green's functions exist as piecewise C^∞ functions of the times and tempered distributions of the space variables.

The preceding discussion is valid for any eigenfunction Ω of H. Of course, the traditional definition of vacuum expectation values and Green's functions uses the ground state for Ω. It should be noticed that nothing said up to this point prevents the ground state from being degenerate. This does not cause any trouble with the existence proof for the vacuum expectation values and Green's functions; it just means that unless further physical requirements are imposed there is, in general, a family of equally admissible vacuum expectation values and Green's functions.

These general remarks can be supplemented and made more precise in special cases. When only bosons are present, the ground state is non-degenerate by virtue of the same argument that shows that the ground state of the non-relativistic Schrödinger equation for spin-less particles has no nodes. On the other hand, the following example shows that when fermions are present no such general argument can exist.

Example 3.11. Consider a single species of fermion interacting with a single species of boson. Suppose $H_2 = g A^* A \tilde{Q}$ where A^* is the creation operator for a fermion mode with corresponding contribution to the free Hamiltonian $M A^* A$, and $\tilde{Q}$ is some linear combination of $Q_1 \ldots Q_m$. Then the Hilbert space may be written as a direct sum $\mathscr{H} = \mathscr{H}_0 \oplus \mathscr{H}_1$ where $\mathscr{H}_0$ is the subspace on which $A^* A = 0$ and $\mathscr{H}_1$ that on which $A^* A = 1$. Each of these subspaces is mapped into itself by the full Hamiltonian H. On $\mathscr{H}_0$ the lowest eigen-value E_0 of H is that of $H_1 = -\Delta_Q + V(Q)$ while on $\mathscr{H}_1$ the lowest eigenvalue of H is equal to the lowest eigenvalue of $-\Delta_Q + V(Q) + M + g\tilde{Q}$. By adjusting g one can make this eigenvalue take any value between $E_0 + M$ and $-\infty$, so in particular one can make it equal to E_0. With this choice of g the ground state of H is at least two-fold degenerate.

Using Theorem 3.1 and the above remarks it is completely straight forward to verify, following the pattern of [1, 2] that the standard Heisenberg equations of motion for the operators $A(f, t)$ hold as equalities on vectors of $C^\infty(H_{0,\nu})$ and that the usual differential equations for the Green's functions hold as equations in tempered distributions. We summarize in

Theorem 3.12. *For quantum field theories satisfying the hypotheses of Theorem 3.1, vacuum expectation values exist as tempered distributions in the space variables and infinitely differentiable functions of the time variables. The Green's functions are tempered distributions and piecewise infinitely differentiable in the times. The standard Heisenberg equations of motion for the field operators are valid, when smeared with a test*

function in the space variable, as operator equalities valid on the vectors of $C^\infty(H_{0,V})$. The standard differential equations for the Green's functions are valid as relations between tempered distributions.

References

1. Jaffe, A. M.: Dynamics of a cut-off $\lambda\varphi^4$ theory. Princeton Thesis 1965 and Amer. Math. Soc. Memoir to appear.
2. Lanford, O. E.: Construction of quantum fields interacting by a cut-off Yukawa coupling. Princeton Thesis 1966 and Amer. Math. Soc. Memoir to appear.
3. Kato, T.: Perturbation theory for linear operators. Berlin-Heidelberg-New York: Springer 1966.
4. Cannon, J.: Princeton Thesis 1967. To appear.
5. See theorems 8.3 and 8.4 proved by Lanford, O., A. Jaffe, and S. Doplicher respectively quoted in Cargèse lectures 1964 edited by M. Levy Gordon and Breach N. Y. 1967. See also I. Segal: Non-linear functions of stochastic processes I. Jour. Funct. Anal. to appear.
6. Nelson, E.: Ann. of Math. **70**, 572 (1959).

Arthur M. Jaffe
Lyman Laboratory of Physics
Harvard University
Cambridge, Massachusetts, USA

Oscar E. Lanford III
Department of Mathematics
University of California
Berkeley, California, USA

Arthur S. Wightman
Dept. of Physics
Princeton University
Princeton, New Jersey 08540, USA

Commun. math. Phys. 13, 194—215 (1969)

Observables at Infinity and States with Short Range Correlations in Statistical Mechanics

O. E. LANFORD III* and D. RUELLE
I. H. E. S., Bures-Sur-Yvette, France

Received April 28, 1969

Abstract. We say that a representation of an algebra of local observables has short-range correlations if any observable which can be measured outside all bounded sets is a multiple of the identity, and that a state has finite range correlations if the corresponding cyclic representation does. We characterize states with short-range correlations by a cluster property. For classical lattice systems and continuous systems with hard cores, we give a definition of equilibrium state for a specific interaction, based on a local version of the grand canonical prescription; an equilibrium state need not be translation invariant. We show that every equilibrium state has a unique decomposition into equilibrium states with short-range correlations. We use the properties of equilibrium states to prove some negative results about the existence of metastable states. We show that the correlation functions for an equilibrium state satisfy the Kirkwood-Salsburg equations; thus, at low activity, there is only one equilibrium state for a given interaction, temperature, and chemical potential. Finally, we argue heuristically that equilibrium states are invariant under time-evolution.

1. Introduction

The aim of equilibrium statistical mechanics is to describe the equilibrium states of a system, once the interaction between its microscopic components are known. These interactions are usually invariant under a large group G of transformations (the Euclidean group, say, or a translation group) and one may thus assume that an equilibrium state ϱ of an infinite system is invariant under G. We say that ϱ is G-ergodic if there is no decomposition $\varrho = \frac{1}{2}\varrho_1 + \frac{1}{2}\varrho_2$ where ϱ_1 and ϱ_2 are distinct states invariant under G. It can be argued that ϱ is ergodic if it describes a pure thermodynamic phase and non-ergodic if it describes a mixture[1]. The decomposition of G-invariant states into G-ergodic states has received much attention recently[2].

If one thinks now of an equilibrium state ϱ corresponding to a crystal, it appears that the crystal has a symmetry group H_α smaller than the group G under which ϱ is invariant. This *spontaneous symmetry*

* Supported in part by NSF research grant GP-7176.

[1] For a discussion of this point, see RUELLE [25].

[2] See for instance DOPLICHER, KASTLER and ROBINSON [5], RUELLE [22], LANFORD and RUELLE [17], STØRMER [27].

breakdown can be understood by writing ϱ as a superposition

$$\varrho = \int d\alpha \, \varrho_\alpha \tag{1.1}$$

where ϱ_α is a state describing a crystal with fixed orientation and lattice position, and $d\alpha$ is a measure on G/H_α. Given an equilibrium state ϱ, one may now ask what the prescription is, to find a physically meaningful decomposition like (1.1)[3]. This problem has been considered by a number of authors[4] mostly from a group-theoretical viewpoint, and it was suggested by HAAG that the decomposition (1.1) should be into ergodic states for time-evolution[5].

In the present paper we adopt the point of view that the decomposition (1.1) should distinguish states ϱ_α only if they differ "far away" in space; one could say that we look for a decomposition into states ϱ_α which differ macroscopically and not just by local fluctuations. We shall say that such states have *short-range correlations* and give them a precise definition in Section 2. In Section 3 we restrict ourselves to classical systems and establish equations which must be satisfied by any equilibrium state. If Δ is the set of states satisfying these equations, it turns out that the states with short-range correlations are just the extremal points of Δ. The results of Sections 3 are used in Section 4 to derive several negative statements about the existence of metastable states in statistical mechanics. In Section 5 we exhibit a case where the invariant equilibrium states already have short range correlations. In Section 6 we give a heuristic argument to show that, for continuous systems, equilibrium states are invariant under time evolution.

Note. After the manuscript of the present article was completed (summer 68), J. LASCOUX pointed out to us that results along the same lines had been obtained by R. L. DOBRUSHIN [see Teorija Verojatn. i ee Prim. **13**, 201—229 (1968); Funkts. Analiz i ego Pril. **2**, 31—43 (1968); **2**, 44—57 (1968); **3**, 27—35 (1969)]. We have not modified our manuscript to take DOBRUSHIN'S work into account, but we urge the reader to consult the articles quoted above. It is of particular interest that DOBRUSHIN could prove the existence of a symmetry breakdown for some non-trivial models of a lattice gas [Funkts. Analiz i ego Pril. **2**, 44—57 (1968)].

2. Observables at Infinity and States with Short Range Correlations

For the purposes of this section, let $\mathfrak{A}$ be a C^* algebra and let $\{\mathfrak{A}_\Lambda\}$ be a collection of sub C^*-algebras of $\mathfrak{A}$ labelled by the bounded open

[3] This decomposition may in some cases (liquid crystals) be into "almost periodic" rather than periodic states ϱ_α. Notice that we look for a "natural" decomposition of ϱ, not a finest possible decomposition. For a classical system one can decompose ϱ into pure states ϱ_α where all the positions (and possibly momenta) of all the particles are fixed; this decomposition is too fine to be of interest to us here.

[4] See in particular KASTLER and ROBINSON [13], ROBINSON and RUELLE [19], DOPLICHER, GALLAVOTTI and RUELLE [4], HAAG, KASTLER, and MICHEL [11].

[5] R. HAAG, private communication.

subsets of $\boldsymbol{R}^\nu$ (continuous systems) or $\boldsymbol{Z}^\nu$ (lattice systems). These objects are subject to the restrictions:

QLA 1. $\bigcup_\Lambda \mathfrak{A}_\Lambda$ is norm-dense in $\mathfrak{A}$.

QLA 2. If $\Lambda \cap M = \emptyset$, and if $A \in \mathfrak{A}_\Lambda$, $B \in \mathfrak{A}_M$ then $[A, B] = 0$.

For any bounded open Λ, let $\tilde{\mathfrak{A}}_\Lambda$ denote the sub C^*-algebra of $\mathfrak{A}$ generated by $\{\mathfrak{A}_M : M \cap \Lambda = \emptyset\}$. If $\mathfrak{A}_\Lambda$ is interpreted as the algebra of observables measurable inside Λ, then $\tilde{\mathfrak{A}}_\Lambda$ is to be interpreted as the algebra of observables measurable *outside* Λ. Note that, by QLA 2., $\mathfrak{A}_\Lambda$ and $\tilde{\mathfrak{A}}_\Lambda$ commute.

Now let π be a *-representation of $\mathfrak{A}$ on a Hilbert space $\mathfrak{H}_\pi$, and define

$$\mathfrak{B}_\pi = \bigcap_\Lambda \overline{\pi(\tilde{\mathfrak{A}}_\Lambda)}\,,$$

where $\overline{}$ denotes weak-operator closure. Since $\overline{\pi(\tilde{\mathfrak{A}}_\Lambda)}$ may be interpreted as the algebra of observables (in a generalized sense) measurable outside Λ, $\mathfrak{B}_\pi$ may be interpreted as the algebra of observables measurable outside any given bounded open set; we will therefore refer to $\mathfrak{B}_\pi$ as the *algebra of observables at infinity*. We will say that the representation π has *short range correlations* if the corresponding algebra $\mathfrak{B}_\pi$ contains only the scalars, and that a state ϱ on $\mathfrak{A}$ has *short range correlations* if the corresponding cyclic representation does.

2. 1. Proposition. *For any *-representation π of $\mathfrak{A}$, the algebra $\mathfrak{B}_\pi$ is contained in the center of $\overline{\pi(\mathfrak{A})}$.*

Since $\mathfrak{B}_\pi$ is evidently contained in $\overline{\pi(\mathfrak{A})}$, it suffices by QLA 1. to show that, for any $B \in \mathfrak{B}_\pi$ and any $A \in \mathfrak{A}_\Lambda$ for some bounded open Λ, $[B, A] = 0$. But since $B \in \mathfrak{B}_\pi$, $B \in \overline{\pi(\tilde{\mathfrak{A}}_\Lambda)}$; since $\tilde{\mathfrak{A}}_\Lambda$ commutes with $\mathfrak{A}_\Lambda$, the proposition is proved.

It follows at once from this proposition that any factor representation of $\mathfrak{A}$ has short range correlations.

One is most interested in the case in which $\mathfrak{A}$ is one of the C^*-algebras used to describe statistical mechanics. Consider first a one-dimensional classical lattice gas. For such a system, the requirement that a translation-invariant state (i.e., an invariant measure on the space of configurations) have short range correlations is analogous to the requirement that the dynamical system defined by the translation mappings and the invariant measure be a K-system (see SINAI [26] or JACOB [12], Section 10.9). Indeed, ϱ defines a K-system if and only if:

$$\bigcap_m \overline{\bigcup_{\Lambda \subset (-\infty, -m)} \pi_\varrho(\mathfrak{A}_\Lambda)} = \{\lambda \mathbf{1}\}$$

or if and only if

$$\bigcap_m \overline{\bigcup_{\Lambda \subset (m, \infty)} \pi_\varrho(\mathfrak{A}_\Lambda)} = \{\lambda \mathbf{1}\};$$

on the other hand, ϱ has finite range correlations if and only if

$$\bigcap_m \overline{\bigcup_{\Lambda \subset (-\infty,-m)\cup(m,\infty)} \pi_\varrho(\mathfrak{A}_\Lambda)} = \{\lambda \mathbf{1}\}\,.$$

Thus, if ϱ has short range correlations, it defines a K-system, and it seems a plausible conjecture that the converse is also true. In any case, states of classical statistical mechanics having short range correlations may be thought of roughly as multi-dimensional generalizations of K-systems. The following proposition shows, however, that the interpretation is quite different in quantum statistical mechanics: for quantum spin systems, the states with short range correlations are precisely the factor states.

2.2. Proposition. *Let $\mathfrak{A}$ be the quasi-local algebra describing a quantum spin system*[6]. *Then for any *-representation π of $\mathfrak{A}$, $\mathfrak{B}_\pi$ coincides with the center of $\overline{\pi(\mathfrak{A})}$.*

By Proposition 2.1, all we have to show is that any B in the center of $\overline{\pi(\mathfrak{A})}$ belongs to $\overline{\pi(\tilde{\mathfrak{A}}_\Lambda)}$ for each bounded Λ. Thus, let B_α be a net of elements of $\mathfrak{A}$ such that $\pi(B_\alpha)$ converges strongly to B. We can suppose that each B_α belongs to some $\mathfrak{A}_{M_\alpha}$, where $M_\alpha \supset \Lambda$. Now $\mathfrak{A}_\Lambda$ is a finite matrix algebra; let (e_{ij}) be a set of matrix units for it. Since B commutes with $\pi(\mathfrak{A}_\Lambda)$,

$$B = \sum_i \pi(e_{i1})\, B \pi(e_{1i}) = \text{st.-}\lim_\alpha \pi\left(\sum_i e_{i1} B_\alpha e_{1i}\right).$$

But $\sum_i e_{i1} B_\alpha e_{1i}$ belongs to $\mathfrak{A}_{M_\alpha}$ and commutes with $\mathfrak{A}_\Lambda$; hence, belongs to $\mathfrak{A}_{M_\alpha} \subset \tilde{\mathfrak{A}}_\Lambda$, so $B \in \overline{\pi(\tilde{\mathfrak{A}}_\Lambda)}$ and the proposition is proved.

A similar argument shows that, if $\mathfrak{A}$ is the quasi-local algebra describing a boson lattice gas or a continuous boson system, $\mathfrak{B}_\pi$ coincides with the center of $\overline{\pi(\mathfrak{A})}$ provided that the restriction of π to each $\mathfrak{A}_\Lambda$ is quasi-equivalent to the Fock representation; this will be the case for representations of physical interest (see RUELLE [21], DELL'ANTONIO, DOPLICHER, and RUELLE [3]).

The following proposition shows that, as the terminology suggests, states with short range correlations are characterized by cluster properties. It contains as special cases known results about K-systems (SINAI [26]) and uniformly hyperfinite C^* algebras (POWERS [18], Theorem 2.5);

[6] By the quasi-local algebra describing a quantum spin system we mean a system $\{\mathfrak{A}, \mathfrak{A}_\Lambda\}$ constructed as follows: Let $\mathfrak{H}$ be a finite-dimensional Hilbert space, $\mathfrak{H}_x$ a copy of $\mathfrak{H}$ for every x in $\mathbf{Z}^\nu$, and $\mathfrak{H}_\Lambda = \bigotimes_{x\in\Lambda} \mathfrak{H}_x$ for every finite $\Lambda \subset \mathbf{Z}^\nu$. Let $\mathfrak{A}_\Lambda$ be the algebra of bounded operators on $\mathfrak{H}_\Lambda$. If $\Lambda \subset M$, the natural isomorphism $\mathfrak{H}_M = \mathfrak{H}_\Lambda \otimes \mathfrak{H}_{M/\Lambda}$ identifies $\mathfrak{A}_\Lambda$ with a subalgebra of $\mathfrak{A}_M$. Then $\mathfrak{A}$ is the norm closure of the union of the $\mathfrak{A}_\Lambda$'s (i.e., the inductive limit of the $\mathfrak{A}_\Lambda$'s). See LANFORD and ROBINSON [15].

the method of proof is a straightforward adaptation of that used in the latter reference.

2.3. Proposition. *Let $\{\mathfrak{A}, \mathfrak{A}_\Lambda\}$ be as above, and let ϱ be a state on $\mathfrak{A}$. Then the following are equivalent:*

1. *ϱ has short range correlations.*
2. *For every $A \in \mathfrak{A}$, there is a bounded open set Λ such that*

$$|\varrho(AB) - \varrho(A)\,\varrho(B)| \leqq \|B\|$$

whenever $B \in \tilde{\mathfrak{A}}_\Lambda$.

Assume that 1. holds but that 2. does not. Then there exists $A \in \mathfrak{A}$, an increasing net M_α of bounded open sets whose union is the whole space, and operators $B_\alpha \in \tilde{\mathfrak{A}}_{M_\alpha}$; $\|B_\alpha\| \leqq 1$, such that

$$\lim_\alpha \varrho(AB_\alpha) - \varrho(A)\,\varrho(B) \neq 0\,.$$

By passing to a subnet, we can assume that $\pi_\varrho(B_\alpha)$ converges in the weak operator topology; since the limit is in $\bigcap_\alpha \overline{\pi\,(\tilde{\mathfrak{A}}_{M_\alpha})}$ it must, by 1., be of the form $b\mathbf{1}$. Then

$$\begin{aligned}\lim_\alpha \varrho(AB_\alpha) - \varrho(A)\,\varrho(B_\alpha) &= \lim_\alpha [(\Omega_\varrho, \pi_\varrho(A)\,\pi_\varrho(B_\alpha)\,\Omega_\varrho) - (\Omega_\varrho, \pi_\alpha(A)\Omega_\varrho) \\ &\quad \times (\Omega_\varrho, \pi_\varrho(B_\alpha)\Omega_\varrho)] \\ &= b(\Omega_\varrho, \pi_\varrho(A)\Omega_\varrho) - b(\Omega_\varrho, \pi_\varrho(A)\Omega_\varrho) = 0\,,\end{aligned}$$

contradicting our earlier assumption and proving that 1. implies 2.

Now suppose that 2. holds, and let $B \in \mathfrak{B}_{\pi_\varrho}$. Then

$$|(\Omega_\varrho, \pi_\varrho(A)\,B\Omega_\varrho) - (\Omega_\varrho, \pi_\varrho(A)\Omega_\varrho)\,(\Omega_\varrho, B\Omega_\varrho)| \leqq \|B\|$$

for all $A \in \mathfrak{A}$. Replacing A by λA multiplies the left-hand side by $|\lambda|$ and leaves the right-hand side unchanged, so the left-hand side must be zero. Letting $b = (\Omega_\varrho, B\Omega_\varrho)$, and using Proposition 2.1, we get therefore:

$$(\pi_\varrho(A_1)\Omega_\varrho, B\pi_\varrho(A_2)\Omega_\varrho) = b(\pi_\varrho(A_1)\Omega_\varrho, \pi_\varrho(A_2)\Omega_\varrho)$$

for all $A_1, A_2 \in \mathfrak{A}$, and hence

$$B = b\mathbf{1}\,.$$

2.4. Corollary. *Let $\{\mathfrak{A}, \mathfrak{A}_\Lambda\}$ be as above and let τ be a representation of the translation group in the automorphism group of $\mathfrak{A}$ such that*

$$\tau_x(\mathfrak{A}_\Lambda) = \mathfrak{A}_{\Lambda+x}\,.$$

Let ϱ be a state of $\mathfrak{A}$ which is invariant under τ and which has short range correlations. Let $A_1, \ldots, A_n \in \mathfrak{A}$. Then

$$\lim_{\min|x_i - x_j| \to \infty} \varrho(\tau_{x_1} A_1 \ldots \tau_{x_n} A_n) = \varrho(A_1) \ldots \varrho(A_n)\,.$$

We can assume that $A_1, \ldots, A_n \in \mathfrak{A}_\Lambda$ for some Λ. Then translation invariance and Proposition 2.3 gives

$$\lim_{\substack{\min |x_i - x_j| \to \infty \\ i \neq j}} \varrho(\tau_{x_1} A_1 \ldots \tau_{x_n} A_n) = \varrho(A_1) \lim_{\substack{\min |x_i' - x_j'| \to \infty \\ i \neq j}} \varrho(\tau_{x_2'} A_2 \ldots \tau_{x_n'} A_n)$$

where $x_i' = x_i - x_1$, $2 \leqq i \leqq n$. The corollary now follows by induction on n.

Because the algebra $\mathfrak{B}_{\pi_\varrho}$ is abelian, it gives a decomposition of the state ϱ. Heuristically, one expects this decomposition to be the coarsest possible decomposition into states with short range correlations. We will not study this decomposition in general. Instead, we will concentrate on the study of the decomposition of equilibrium states of classical statistical mechanics, using special methods to be developed in the next section.

3. Equilibrium Equations for Classical Systems

We shall consider, in this and the following section, only classical lattice gases; in Appendix B we show how our results may be extended to classical hard core continuous systems.

For a lattice gas, $\mathfrak{A} = \mathscr{C}(K)$ is the algebra of continuous complex functions on the compact set[7]

$$K = \{0,1\}^{\mathbf{Z}^\nu} = \mathscr{P}(\mathbf{Z}^\nu)\,. \tag{3.1}$$

An element $X : \mathbf{Z}^\nu \to \{0,1\}$ of $\{0,1\}^{\mathbf{Z}^\nu}$ is here identified with the set $\{X \in \mathbf{Z}^\nu : X(x) = 1\} \in \mathscr{P}(\mathbf{Z}^\nu)$; K is compact as product of the sets $\{0,1\}$ (which are compact with the discrete topology). If Λ is a finite subset of $\mathbf{Z}^\nu$, $\mathfrak{A}_\Lambda$ is the algebra of "cylindrical functions" A such that for some $\varphi \in \mathscr{C}(\mathscr{P}(\Lambda))$,

$$A(X) = \varphi(X \cap \Lambda) \quad \text{for all} \quad X \in K.$$

If $x \in \mathbf{Z}^\nu$, τ_x is the automorphism of $\mathfrak{A}$ defined by

$$\tau_x A(X) = A(X - x) \quad \text{for all} \quad X \in \mathscr{P}(\mathbf{Z}^\nu)$$

where $X - x$ is the set X translated by $-x$.

A state ϱ on $\mathfrak{A}$ is the same thing as a probability measure on K. If Λ is a finite subset of $\mathbf{Z}^\nu$, we shall define, for every $X \subset \Lambda$, a measure $\varrho_\Lambda(X, dY)$ on $\mathscr{P}(\mathbf{Z}^\nu \backslash \Lambda)$ by

$$\varrho(A) = \sum_{X \subset \Lambda} \int A(X \cup Y)\, \varrho_\Lambda(X, dY)\,. \tag{3.2}$$

We shall say that ϱ is a $\mathbf{Z}^\nu$-invariant state, or simply an *invariant state* if

$$\varrho(\tau_x A) = \varrho(A) \quad \text{for all} \quad x \in \mathbf{Z}^\nu, A \in \mathfrak{A}\,.$$

[7] We denote by $\mathscr{P}(E)$ the set of all subsets of E.

An interaction Φ of the lattice gas is a real function on the finite subsets of $\boldsymbol{Z}^\nu$ satisfying

1. $\Phi(\emptyset) = 0$,

2. translation invariance: $\Phi(X + x) = \Phi(X)$,

3. $\|\Phi\| = \sum_{X \ni 0} |\Phi(X)| < +\infty$. (3.3)

The interactions form a Banach space $\mathscr{B}$ with respect to the norm (3.3). Let $\mathscr{B}_0$ consist of the finite range interactions, i.e. of the interactions Φ such that $\Phi(X) \neq 0$ for only a finite number of sets $X \ni 0$; the space $\mathscr{B}_0$ is dense in $\mathscr{B}$. For finite $X, \Lambda \subset \boldsymbol{Z}^\nu$ we let

$$U_\Phi(X) = \sum_{Y \subset X} \Phi(Y), \tag{3.4}$$

$$P_\Lambda(\Phi) = N(\Lambda)^{-1} \log \sum_{X \subset \Lambda} \exp[-U_\Phi(X)] \tag{3.5}$$

(where $N(\Lambda)$ is the number of elements in Λ); then one can show that the following limit exists for all $\Phi \in \mathscr{B}$:

$$P(\Phi) = \lim_{\Lambda \to \infty} P_\Lambda(\Phi) \tag{3.6}$$

when Λ tends to infinity in an appropriate sense (see Appendix A). The function P is convex and continuous on $\mathscr{B}$.

The definition of an invariant equilibrium state corresponding to the interaction Φ is a somewhat delicate question which has been considered in detail in the literature[8]. A description of the problem is given in Appendix A, which contains also the proof of Theorem 3.2 below. Here it is convenient to accept provisionally the following somewhat untransparent definition.

3.1. Definition. *If $\Psi \in \mathscr{B}$, let $A_\Psi \in \mathfrak{A}$ be defined by*

$$A_\Psi(X) = \sum_{Y \subset X : Y \ni 0} \frac{\Psi(Y)}{N(Y)}. \tag{3.7}$$

An invariant state ϱ on $\mathfrak{A}$ is an invariant equilibrium state for the interaction Φ if the linear functional $\Psi \to -\varrho(A_\Psi)$ is tangent to the graph of $P(\cdot)$ at $(\Phi, P(\Phi))$, i.e. if

$$P(\Phi + \Psi) \geqq P(\Phi) - \varrho(A_\Psi) \quad \text{for all} \quad \Psi \in \mathscr{B}. \tag{3.8}$$

3.2. Theorem. *For finite $\Lambda \subset \boldsymbol{Z}^\nu$, let $f_\Lambda \in \mathscr{C}(\mathscr{P}(\Lambda) \times \mathscr{P}(\boldsymbol{Z}^\nu \setminus \Lambda))$ be defined by*

$$f_\Lambda(X, Y) = \exp\left[-\sum_{S \subset X \cup Y : S \cap X \neq \emptyset} \Phi(S)\right]. \tag{3.9}$$

[8] See GALLAVOTTI and MIRACLE [7], RUELLE [24], LANFORD and ROBINSON [16]; for a review see RUELLE [25].

An invariant state ϱ is an invariant equilibrium state if and only if, for all Λ and X,

$$\varrho_\Lambda(X, dY) = f_\Lambda(X, Y)\, \varrho_\Lambda(\emptyset, dY) \tag{3.10}$$

where the notation (3.2) *has been used.*

This theorem is proved in Appendix A. The Eqs. (3.10) can be understood as follows. Instead of an infinite system, consider a system enclosed in the finite region $M \subset \mathbf{Z}^\nu$. The equilibrium state of the latter system is described by a measure μ on $\mathscr{P}(M)$ such that[9], if $X \subset M$,

$$\mu(\{X\}) = \left\{ \sum_{Y \subset M} \exp[-U_\Phi(Y)] \right\}^{-1} \exp[-U_\Phi(X)]\,. \tag{3.11}$$

Now, if $\Lambda \subset M$ and $X \subset \Lambda$, $Y \cap \Lambda = \emptyset$, we have

$$\mu(\{X \cup Y\}) = \exp\left[- \sum_{S \subset X \cup Y : S \cap X \neq \emptyset} \Phi(S) \right] \mu(\{Y\})\,. \tag{3.12}$$

If we formally let $M \to \infty$ in (3.12) we obtain (3.10).

It is known (see Appendix A) that for each $\Phi \in \mathscr{B}$ there is at least one invariant equilibrium state and therefore an invariant state satisfying (3.10).

3.3. Definition. *A state ϱ on $\mathfrak{A}$ is an equilibrium state for the interaction Φ if it satisfies the equations* (3.10). *We denote by Δ_Φ or Δ the set of equilibrium states for Φ.*

By Theorem 3.2, an invariant equilibrium state is an equilibrium state, and in particular Δ is not empty; Δ is convex and compact for the weak topology[10].

3.4. Theorem. *A state $\varrho \in \Delta$ has short range correlations if and only if it is an extremal point of Δ.*

The non-extremality of ϱ in Δ is equivalent to the existence of $h \in L^\infty(\varrho)$, $0 \leqq h \leqq 1$, h not a multiple of 1, such that $h\varrho$ satisfies (3.10), i.e.

$$h(X, Y)\, \varrho_\Lambda(X, dY) = f_\Lambda(X, Y)\, h(\emptyset, Y)\, \varrho_\Lambda(\emptyset, dY)\,. \tag{3.13}$$

Since ϱ satisfies (3.10), (3.13) is equivalent to

$$\varrho_\Lambda(X, dY)\, [h(X, Y) - h(\emptyset, Y)] = 0$$

i.e. to $h(X, Y) = h(\emptyset, Y)$ ϱ-almost everywhere. This means that $h \in \overline{\pi_\varrho(\mathfrak{A}_\Lambda)}$ and therefore $h \in \mathscr{B}_{\pi_\varrho}$. But the existence of $h \in \mathscr{B}_{\pi_\varrho}$, $0 \leqq h \leqq 1$, h not a multiple of 1, is equivalent to ϱ not having short range correlations, proving the theorem.

The following result shows that every equilibrium state has a unique decomposition into equilibrium states with short range correlations.

---

[9] This is the "grand canonical" prescription of GIBBS, with the factor $1/kT$ and the chemical potential term absorbed in the definition of the interaction Φ.

[10] The weak topology of the dual of $\mathscr{C}(K)$, also called the vague topology.

3.5. Proposition. *The set Δ_Φ is a simplex in the sense of Choquet; hence, every $\varrho \in \Delta_\Phi$ is the resultant of a unique measure m_ϱ on Δ_Φ carried by the extremal points of Δ_Φ:*

$$\varrho(A) = \int \sigma(A)\, dm_\varrho(\sigma) \tag{3.14}$$

for all $A \in \mathfrak{A}$.

Let $\mathfrak{B}_\Phi$ be the vector space of measures on K which satisfy (3.10). If $\mu \in \mathfrak{B}_\Phi$, then $|\mu| \in \mathfrak{B}_\Phi$; therefore, $\mathfrak{B}_\Phi$ is a lattice[11] with respect to the usual order relation for measures. Since Δ_Φ is a basis of the cone of positive elements of $\mathfrak{B}_\Phi$, Δ_Φ is a simplex in the sense of Choquet[12]. The fact that every $\varrho \in \Delta_\Phi$ has a unique integral representation in terms of extremal points of Δ_Φ follows then from the metrizability of Δ_Φ[12].

4. Non Existence of Metastable States

It is known that if water is heated above its boiling point at a certain pressure, it does not necessarily undergo the expected phase transition to water vapor but may stay in the liquid phase in a so-called *metastable state*. A great variety of such metastable states are known experimentally.

One may think that metastable states are truly unstable but, due to the finite size of systems, decay only very slowly in time[13]. Another possibility is that a metastable state for an infinite systems has an infinite lifetime and is very similar to a true equilibrium state except that it does not obey the usual variational principle (maximum entropy at fixed energy and density or maximum pressure at fixed temperature and chemical potential). In support of the second alternative comes the fact that a metastable branch occurs in the Van der Waals theory, suggesting that metastable states are in general analytic continuations of stable equilibrium states.

In this section we give a certain number of negative results, tending to prove that metastable states, as close analogues or analytic continuations of stable equilibrium states, cannot exist.

We consider the case of lattice gases[14]; then the proof of Theorem 3.2 (see Appendix A) gives in particular.

4.1. Proposition. *An invariant state ϱ satisfying the Eqs.* (3.10) *and metastable in the sense that*

$$s(\varrho) - \varrho(A_\Phi) < P(\Phi)$$

cannot exist (the entropy $s(\cdot)$ is defined by (A.9)).

---

[11] I.e. every finite set of elements of $\mathfrak{B}_\Phi$ has a g.l.b. and a l.u.b.

[12] See CHOQUET and MEYER [2].

[13] Slowly provided that the effect of impurities and other disturbances is adequately eliminated.

[14] An extension to hard core continuous systems is immediate.

In Fig. 1 we draw a typical pressure versus activity isotherm with a kink at z_0 corresponding to a first order phase transition. We have also drawn a "metastable branch" (dotted) below the equilibrium curve. From Proposition 4.1, it follows that there cannot exist an invariant state ϱ_m satisfying the Eqs. (3.10) and corresponding to the point (z_m, P_m).

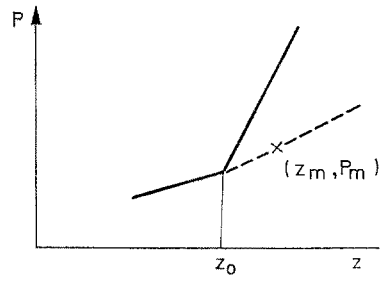


Fig. 1

4.2. Proposition. *Suppose that Φ has finite range and that P is not analytic with respect to z at z_0. It is impossible that a state ϱ_z be defined for z in a neighborhood of z_0 so that*

1. *The correlation functions*[15] $\varrho_z(X)$ *are real analytic in* z.

2. ϱ_z *is the stable equilibrium state for* $z \leqq z_0$.

Since Φ has finite range, the Eqs. (3.10) may be written in the form (A.6) and therefore expressed in terms of the correlation functions (the $\varrho_\Lambda(\{X\})$ are finite linear combinations of the $\varrho(Y)$). The Eqs. (A.6), which are satisfied by ϱ_z for $z \leqq z_0$, remain satisfied for $z > z_0$ by analytic continuation. Therefore (by Theorem 3.2) ϱ_z corresponds to a tangent to the graph of P for all z in a neighborhood of z_0 and in particular the one-point correlation function (density) is given by

$$\varrho_z(\{0\}) = z \frac{d}{dz} P(\Phi(z)) .$$

The analyticity in z of the left-hand side contradicts the assumed existence of a singularity of the right-hand side at z_0, proving the proposition.

Remark. In the same direction of excluding the existence of metastable states, it has been conjectured by FISCHER [6] that, as an analytic function of z, P must exhibit a singularity at the point z_0 of a first order phase transition; a proof of this fact has been announced for the Ising model by a group of Russian workers [1].

[15] We define $\varrho(X) = \int \varrho_X(X, dY)$; see Section 5.

5. Application: Derivation of the Kirkwood-Salsburg Equations

In this section we show that one can, from the Eqs. (3.10), derive the Kirkwood-Salsburg equations for the correlation functions. Since it is known that under suitable conditions (sufficiently low activity), the Kirkwood-Salsburg equations have a unique solution[16], it follows that under these conditions the set Δ_Φ of equilibrium states is reduced to a point.

We assume that Φ is a pair interaction, i.e. $\Phi(X) = 0$ if $N(X) > 2$; we may then write

$$f_\Lambda(X, Y) = z^{N(X)} \exp\left[-\sum_{\{x,x'\}\subset X} \varphi(x' - x) - \sum_{x\in X}\sum_{y\in Y} \varphi(y - x)\right] \tag{5.1}$$

where z is the activity and φ the pair potential associated with the pair interaction Φ. We have

$$D = \sum_{x\neq 0} |\varphi(x)| < +\infty\,. \tag{5.2}$$

We define also $\varphi(0) = +\infty$.

The correlation function associated with the state ϱ is a function $X \mapsto \varrho(X)$ of finite subsets of $\boldsymbol{Z}^\nu$ defined by

$$\varrho(X) = \int \varrho_X(X, dY)\,. \tag{5.3}$$

If $x_1 \in X$ and $X_1 = X\backslash\{x_1\}$, we have

$$\begin{aligned}
\varrho_X(X, dY) &= f_X(X, Y)\, \varrho_X(\emptyset, dY) \\
&= z\exp\left[-\sum_{x\in X_1} \varphi(x - x_1)\right] \prod_{y\in Y} [1 + (e^{-\varphi(y-x_1)} - 1)] \\
&\quad \cdot f_X(X_1, Y)\, \varrho_X(\emptyset, dY) \\
&= z\exp\left[-\sum_{x\in X_1} \varphi(x - x_1)\right] \sum_{S\subset Y} \prod_{y\in S} (e^{-\varphi(y-x_1)} - 1)\, \varrho_X(X_1, dY).
\end{aligned}$$

Therefore

$$\begin{aligned}
\varrho(X) &= \int \varrho_X(X, dY) \\
&= z\exp\left[-\sum_{x\in X_1} \varphi(x - x_1)\right] \int \sum_{S\subset Y} \prod_{y\in S} (e^{-\varphi(y-x_1)} - 1)\, \varrho_X(X_1, dY) \\
&= z\exp\left[-\sum_{x\in X_1} \varphi(x - x_1)\right] \sum_{S\subset \boldsymbol{Z}^\nu\backslash X} \int \prod_{y\in S} (e^{-\varphi(y-x_1)} - 1) \\
&\quad \cdot \varrho_{X\cup S}(X_1 \cup S, dZ) \\
&= z\exp\left[-\sum_{x\in X_1} \varphi(x - x_1)\right] \sum_{S\subset \boldsymbol{Z}^\nu\backslash X} \prod_{y\in S} (e^{-\varphi(y-x_1)} - 1) \\
&\quad \cdot \int \varrho_{X\cup S}(X_1 \cup S, dZ) \\
&= z\exp\left[-\sum_{x\in X_1} \varphi(x - x_1)\right] \sum_{S\subset \boldsymbol{Z}^\nu\backslash X} \prod_{y\in S} (e^{-\varphi(y-x_1)} - 1) \\
&\quad \cdot [\varrho(X_1 \cup S) - \varrho(X \cup S)]\,.
\end{aligned}$$

[16] See RUELLE [20].

Therefore

$$\varrho(X) = z \exp\left[-\sum_{x \in X_1} \varphi(x - x_1)\right] \sum_{S \subset \mathbf{Z}^\nu \setminus X_1} \prod_{y \in S} (e^{-\varphi(y - x_1)} - 1)\, \varrho(X_1 \cup S) \quad (5.4)$$

These relations are the Kirkwood-Salsburg equations; they determine uniquely the correlation function and therefore the state ϱ if

$$z \left[\exp \sum_{x \neq 0} |\varphi(x)|\right] \left[\exp \sum_{x} |e^{-\varphi(x)} - 1|\right] < 1 \quad (5.5)$$

(in particular if $z < \exp[-D - e^D]$).

Instead of the Kirkwood-Salsburg equations one could obtain other equations, due to GALLAVOTTI and MIRACLE[17] and for which it is not necessary to assume that Φ is a pair interaction; we write

$$f_\Lambda(X, Y) = z^{N(X)} \exp\left[-\sum_{S \subset X \cup Y : S \cap X \neq \emptyset} \Phi'(X)\right] \quad (5.6)$$

where $\Phi'(X) = 0$ when $N(X) = 1$. One finds here that the set Δ_Φ of equilibrium states is reduced to a point if

$$\frac{z e^{D-C}}{1 + z e^{D-C}} \cdot [2 \exp(e^D - 1) - 1] < 1 \quad (5.7)$$

where

$$C = \sum_{X \ni 0} \Phi'(X), \; D = \sum_{X \ni 0} |\Phi'(X)| = \|\Phi'\| . \quad (5.8)$$

Using the fact that the extremal points of Δ have short range correlations (Theorem 3.4), we see that, when the correlation functions are uniquely determined by the Kirkwood-Salsburg equations or the equations of GALLAVOTTI and MIRACLE the (unique) equilibrium state has short range correlations and therefore, by Proposition 2.3 and Corollary 2.4, has strong cluster properties.

6. Time-Invariance of Equilibrium States

In this section, we give a heuristic argument indicating that states of continuous classical-mechanical systems satisfying the analogue of (3.10) should be invariant under time evolution. We proceed in the following way: Consider first a finite system in a region M, and the part of that system contained in a smaller region Λ. Using LIOUVILLE's equation for the time-evolution of density distributions in M, we obtain an integrodifferential equation giving the time derivative of the density distributions in Λ in terms of the density distributions in a larger region Λ'. Since M no longer appears in this equation, we can take this system of equations as describing the time-evolution of the part of an infinite system which is contained in the bounded

[17] See GALLAVOTTI and MIRACLE [8], GALLAVOTTI, MIRACLE, and ROBINSON [10] and RUELLE [25] (Theorem 4.2.7.).

region Λ. We then show using these equations that any state which satisfies the continuous analogue of (3.10) has zero time derivative. We emphasize that the argument is only heuristic: For infinite systems in more than one dimension, no satisfactory theory of time-evolution exists, and, even for one-dimensional systems for which such a theory does exist [14], we have not shown that our formal condition for invariance under infinitesimal time translations rigorously implies time-invariance. Since our argument is only formal we will not worry about differentiability questions and interchanges of order of limits. We will assume that the finite systems we consider interact by interparticle forces defined by potentials with finite range R and with conservative external forces defining the walls of the system.

Before looking at the time-evolution problem, we outline the description of a state for an infinite continuous system in terms of local density distributions. Consider first a system in a bounded region M. The state of the system is specified by giving the density distributions $f_M^{(n)}(q_1, \ldots, q_n; p_1, \ldots, p_n)$ such that the probability of finding precisely n particles in M, and these particles with positions and velocities defining a point of $E \subset (M \times \boldsymbol{R}^\nu)^n$ is

$$\frac{1}{n!} \int_E dq_1 \ldots dq_n \, dp_1 \ldots dp_n \quad f_M^{(n)}(q_1, \ldots q_n; p_1, \ldots p_n) .$$

(We are using the functional notation for measures, so the formulas we write will be strictly valid only for measures absolutely continuous with respect to Lebesgue measure; it is not hard to rewrite them in a way that allows general probability measures). The function $f_M^{(n)}(q_1, \ldots, q_n; p_1, \ldots, p_n)$ is symmetric in the variables (q_i, p_i) and normalized by

$$\sum_{n=0}^{\infty} \frac{1}{n!} \int_{(M \times \boldsymbol{R}^\nu)^n} dq_1 \ldots dp_n \, f_M^{(n)}(q_1, \ldots, q_n; p_1, \ldots, p_n) = 1 . \tag{6.1}$$

If $\Lambda \subset M$, then the density distributions in Λ are given by

$$f_\Lambda^{(n)}(q_1, \ldots, q_n, p_1, \ldots, p_n) = \sum_{l=0}^{\infty} \frac{1}{l!} \int_{(M \setminus \Lambda \times \boldsymbol{R}^\nu)^l} dq_{n+1} \ldots dq_{n+l} dp_{n+1} \ldots dp_{n+l}$$
$$\cdot f_M^{(n+l)}(q_1, \ldots, q_{n+l}, p_1, \ldots, p_{n+l}) . \tag{6.2}$$

We can abbreviate the notation by letting x denote $(n; q_1, \ldots, q_n; p_1, \ldots, p_n)$, letting $\mathscr{X}(M)$ denote the set of all such configurations of particles in M, and letting

$$\int_{\mathscr{X}(M)} dx = \sum_{n=0}^{\infty} \frac{1}{n!} \int dq_1 \ldots dq_n \, dp_1 \ldots dp_n .$$

Then formulas (6.1) and (6.2) can be rewritten as:

$$\int_{\mathscr{X}(M)} dx\, f_M(x) = 1\,, \tag{6.1'}$$

$$f_\Lambda(x) = \int_{\mathscr{X}(M\setminus\Lambda)} dy\, f_M(x, y)\,. \tag{6.2'}$$

Suppose now that we have, for every bounded open M, a non-negative symmetric function f_M on $\mathscr{X}(M)$ satisfying (6.1), and that this system of functions satisfies (6.2) for every pair Λ, M of bounded open sets with $\Lambda \subset M$. The system of functions then determines a state of the infinite system as defined in [23]. We will refer to f_Λ as the *system of local density distributions* defining the state in question.

We return to the consideration of a system contained in the bounded open set M, with time evolution defined by a Hamiltonian H. We have, by Liouville's Theorem,

$$\frac{\partial f_M(x, y, t)}{\partial t} = \sum_{i=1}^{n+l} \left[\frac{\partial H}{\partial q_i}\frac{\partial f_M}{\partial p_i} - \frac{\partial H}{\partial p_i}\frac{\partial f_M}{\partial q_i}\right].$$

(Here, $x = (n; q_1, \ldots, q_n; p_1, \ldots, p_n) \in \mathscr{X}(\Lambda)$; and $y = (l; q_{n+1}, \ldots, q_{n+l}$ $p_{n+1}, \ldots, p_{n+l}) \in \mathscr{X}(M\setminus\Lambda)$). Integrating over y gives:

$$\frac{\partial f_\Lambda(x, t)}{\partial t} = \int_{\mathscr{X}(M\setminus\Lambda)} dy \sum_{i=1}^{n+l} \left[\frac{\partial H}{\partial q_i}\frac{\partial f_M}{\partial p_i} - \frac{\partial H}{\partial p_i}\frac{\partial f_M}{\partial q_i}\right].$$

We will assume that f_M is even in each p_i separately; this will be the case, for example, if the momentum distribution is Maxwellian. (This assumption is not necessary, but it permits considerable simplifications; it implies that there is no net flow of particles into Λ). Then $\frac{\partial H}{\partial q_i}\frac{\partial f_M}{\partial p_i}$ and $\frac{\partial H}{\partial p_i}\frac{\partial f_M}{\partial q_i}$ are both odd in p_i, so the terms with $n+1 \leqq i \leqq n+l$ in the above equation give zero when integrated over y. Also, for $1 \leqq i \leqq n$, $\frac{\partial H(x, y)}{\partial p_i}$ does not depend on y and may therefore be taken outside the integral. Finally, if we let Λ' denote the set of points of M which are at a distance less than R from Λ, and if we use x' to denote the variables in $\Lambda'\setminus\Lambda$ and y' to denote the variables in $M\setminus\Lambda'$, then, for $1 \leqq i \leqq n$, $\frac{\partial H(x, x', y')}{\partial q_i}$ does not depend on y' so

$$\int_{\mathscr{X}(M\setminus\Lambda)} dy \frac{\partial H}{\partial q_i}\frac{\partial f_M}{\partial p_i} = \int_{\mathscr{X}(\Lambda'\setminus\Lambda)} dx' \frac{\partial H(x, x')}{\partial q_i}\frac{\partial}{\partial p_i}\int_{\mathscr{X}(M\setminus\Lambda')} dy'\, f_M(x, x', y', t)$$

$$= \int_{\mathscr{X}(\Lambda'\setminus\Lambda)} dx' \frac{\partial H(x, x')}{\partial q_i}\frac{\partial f_{\Lambda'}(x, x', t)}{\partial p_i}\,.$$

Thus, we obtain the integro-differential equation:

$$\frac{\partial f_\Lambda(x,t)}{\partial t} = -\sum_{i=1}^{n} \frac{\partial H(x)}{\partial p_i}\frac{\partial f_\Lambda(x,t)}{\partial q_i} - \int_{\mathscr{X}(\Lambda'\setminus\Lambda)} dx' \frac{\partial H(x,x')}{\partial q_i}\frac{\partial f_{\Lambda'}(x,x',t)}{\partial p_i}. \quad (6.3)$$

If, now, M contains all points within a distance R of Λ, and if the external forces defining the walls of M do not affect particles inside Λ, Eq. (6.3) is independent of M and we can let $M \to \infty$. We will therefore take the system of Eqs. (6.3), with Λ running over all bounded open sets, to describe the time evolution of the state of the infinite system defined by the system of density distributions $\{f_\Lambda\}$.

We next show, using these equations, that an equilibrium state has zero time derivative. As above, for any bounded open set Λ, let Λ' be the set of points at a distance less than R from Λ. Let $\Lambda'' = (\Lambda')'$, let x denote a variable in $\mathscr{X}(\Lambda)$, x' a variable in $\mathscr{X}(\Lambda'\setminus\Lambda)$, and x'' a variable in $\mathscr{X}(\Lambda''\setminus\Lambda')$. Let y be a configuration of particles in $R^\nu\setminus\Lambda''$, and let $W(x'', y)$ denote the energy of interaction between the configuration (x, x', x'') in Λ'' and the configuration y. We have built into our notation the fact that, because the range of the potentials is R, this interaction energy depends only on x'' and y. All we will use of the definition of equilibrium state is that an equilibrium state is defined by a system of local density distributions with $f_{\Lambda''}(x, x', x'')$ a linear superposition of functions of the form $e^{-\beta H(x,x',x'')-\beta W(x'',y)}$. It will thus suffice to prove that, if

$$f_{\Lambda',y}(x,x') = \int_{\mathscr{X}(\Lambda''\setminus\Lambda')} dx'' e^{-\beta H(x,x',x'')-\beta W(x'',y)}$$

and if

$$f_{\Lambda,y}(x) = \int_{\mathscr{X}(\Lambda'\setminus\Lambda)} dx' f_{\Lambda',y}(x,x'),$$

then, for $1 \leqq i \leqq n$

$$\frac{\partial f_{\Lambda,y}(x)}{\partial q_i}\frac{\partial H(x)}{\partial p_i} = \int_{\mathscr{X}(\Lambda'\setminus\Lambda)} dx' \frac{\partial H(x,x')}{\partial q_i}\frac{\partial f_{\Lambda',y}(x,x')}{\partial p_i}.$$

This formula is proved by a straightforward calculation, which we omit.

Appendix A. Equilibrium States

Before coming to the proof of Theorem 3.2, we mention a certain number of facts connected with the definition of invariant equilibrium states.

First, when we write $\Lambda \to \infty$ for finite $\Lambda \in \boldsymbol{Z}^\nu$ we mean convergence in the sense of van Hove, i.e. $N(\Lambda) \to \infty$ and for every finite set $X \subset \boldsymbol{Z}^\nu$:

$$N(\{x : x + X \subset \Lambda\})/N(\Lambda) \to 1.$$

If Λ, M are finite subsets of $\mathbf{Z}^\nu$ and $A \in \mathfrak{A}_\Lambda$, the finite system equilibrium state is given, according to (3.11) by

$$\mu_M(A) = \left\{ \sum_{X \subset M} \exp[-U_\Phi(X)] \right\}^{-1} \sum_{X \subset M} A(X) \exp[-U_\Phi(X)] \tag{A.1}$$

where we have assumed $\Lambda \subset M$. Without this assumption, we define another linear functional $\bar{\mu}_{M\Lambda}$ on $\mathfrak{A}_\Lambda$, obtained by averaging $\mu_{M\Lambda}$ over translations:

$$\bar{\mu}_{M\Lambda}(A) = N(M)^{-1} \sum_{x : x + \Lambda \subset M} \mu_M(\tau_x A) . \tag{A.2}$$

A.1. Theorem[18]. *Let Γ_Φ be the set of all invariant equilibrium states for Φ. Then:*

a) *The set*

$$D = \{\Phi \in \mathscr{B} : \Gamma_\Phi \quad \text{consists of a single point} \quad \varrho^\Phi\}$$

is dense in $\mathscr{B}$.

b) *Let $\Phi \in \mathscr{B}$. Given a sequence $M_n \to \infty$ there exists a subsequence M_n' and $\varrho \in \Gamma_\Phi$ such that for every finite $\Lambda \subset \mathbf{Z}^\nu$ and $A \in \mathfrak{A}_\Lambda$,*

$$\lim_{n \to \infty} \bar{\mu}_{M_n' \Lambda}(A) = \varrho(A) . \tag{A.3}$$

In particular, if $\Phi \in D$,

$$\lim_{M \to \infty} \bar{\mu}_{M\Lambda}(A) = \varrho^\Phi(A) . \tag{A.4}$$

c) *Let (Φ_i, ϱ_i) be any sequence such that $\Phi_i \in \mathscr{B}$, $\varrho_i \in \Gamma_{\Phi_i}$, $\Phi_i \to \Phi$ and (ϱ_i) has the (weak) limit ϱ; then $\varrho \in \Gamma_\Phi$.*

d) *Let $\Phi \in \mathscr{B}$; then Γ_Φ is the closed convex hull of the set of all ϱ obtained in the manner of (c) with sequences such that $\Phi_i \in D$.*

We come now to the proof Theorem 3.2. First, let Φ be a finite range interaction ($\Phi \in \mathscr{B}_0$). There is then a finite set $Q \subset \mathbf{Z}^\nu$, $Q \ni 0$, such that $X \neq \emptyset$ and $\Phi(X \cup Y) \neq 0$ imply $Y \subset X + Q$. In particular, Eq. (3.19) shows that $f_\Lambda(X, Y)$ depends on Y only through $Y \cap [\Lambda + Q]$. Let $\Lambda' \supset \Lambda + Q$, (3.12) and (A.2) yield

$$\bar{\mu}_{M\Lambda'}(\{X \cup Y\}) = f_\Lambda(X, Y)\, \bar{\mu}_{M\Lambda'}(\{Y\}) . \tag{A.5}$$

Using part (b) of Theorem A.1 shows then that, for some state $\varrho \in \Gamma_\Phi$,

$$\varrho_{\Lambda'}(\{X \cup Y\}) = f_\Lambda(X, Y)\, \varrho_{\Lambda'}(Y) \tag{A.6}$$

where

$$\varrho_\Lambda(\{X\}) = \int \varrho_\Lambda(X, dY) \tag{A.7}$$

Therefore (3.10) is satisfied by ϱ. Using part (c) of the theorem and the density of $\mathscr{B}_0$ one concludes that the Eqs. (3.10) are satisfied by ϱ^Φ when $\Phi \in D$. Finally, using part (a) and part (d), one sees that Eqs. (3.10)

[18] See Gallavotti and Miracle [7] and Ruelle [24] for (a), (b) and (c), Lanford and Robinson [16] for (d).

hold for all $\Phi \in \mathscr{B}$ and $\varrho \in \Gamma_\Phi$. This proves the first part of Theorem 3.2, namely that an invariant equilibrium state satisfies (3.10). To finish the proof, we show that every invariant state satisfying (3.10) is an invariant equilibrium state. An invariant state ϱ is an invariant equilibrium state if[19]

$$P(\Phi) = s(\varrho) - \varrho(A_\Phi) . \tag{A.8}$$

where

$$s(\varrho) = \lim_{\Lambda \to \infty} \frac{1}{N(\Lambda)} S(\varrho_\Lambda) , \tag{A.9}$$

ϱ_Λ is the measure on $\mathscr{P}(\Lambda)$ defined by (A.7), and

$$S(\varrho_\Lambda) = - \sum_{X \subset \Lambda} \varrho_\Lambda(\{X\}) \log \varrho_\Lambda(\{X\}) . \tag{A.10}$$

Moreover, for *any* invariant state ϱ, $\varrho(A_\Phi) = \lim_{\Lambda \to \infty} \frac{1}{N(\Lambda)} \varrho_\Lambda(U_\Phi)$; also,

$$P_\Lambda(\Phi) = \sup\left\{\frac{1}{N(\Lambda)}[S(\mu) - \mu(U_\Phi)] : \mu \text{ a probability measure on } \mathscr{P}(\Lambda)\right\}. \tag{A.11}$$

It will therefore suffice to prove the following assertion (which is a bit stronger than the statement of Theorem 3.2 since the requirement of translation invariance has been dropped).

If ϱ satisfies (3.10), *then*

$$\liminf_{\Lambda \to \infty} \frac{1}{N(\Lambda)} [S(\varrho_\Lambda) - \varrho_\Lambda(U_\Phi)] \geqq P(\Phi) . \tag{A.12}$$

Proof. By (3.10),

$$\varrho_\Lambda(\{X\}) = \int_{Y \subset \mathbf{Z}^\nu \setminus \Lambda} f_\Lambda(X, Y)\, \varrho_\Lambda(\emptyset, dY)$$

and, therefore, ϱ_Λ may be expressed as a (generalized) convex linear combination of the probability measures $\mu_{\Lambda, Y}$ defined by

$$\mu_{\Lambda, Y}(\{X\}) = f_\Lambda(X, Y) \Big/ \left(\sum_{X \subset \Lambda} f_\Lambda(X, Y)\right) .$$

Introducing:

$$W_\Phi(X, Y) = \sum_{\substack{S \subset X \cup Y \\ S \cap X \neq \emptyset \neq S \cap Y}} \Phi(S)$$

we get:

$$f_\Lambda(X, Y) = \exp[- U_\Phi(X) - W_\Phi(X, Y)]$$

[19] For any invariant state σ one has

$$P(\Phi) \geqq s(\sigma) - \sigma(A_\Phi) .$$

Therefore, for all Ψ,

$$P(\Phi + \Psi) \geqq s(\varrho) - \varrho(A_{\Phi+\Psi}) = s(\varrho) - \varrho(A_\Phi) - \varrho(A_\Psi) = P(\Phi) - \varrho(A_\Psi) ,$$

so $\Psi \mapsto - \varrho(A_\Psi)$ is a tangent plane to the graph of P. See RUELLE [24].

and therefore:

$$S(\mu_{\Lambda,Y}) - \mu_{\Lambda,Y}(U_\Phi) - N(\Lambda) P_\Lambda(\Phi) = \mu_{\Lambda,Y}(W_\Phi(\cdot, Y)) + \log [\mu_{\Lambda,\phi} e^{-W_\Phi(\cdot,Y)}]$$
$$\geqq -2 \sup_{\substack{X \subset \Lambda \\ Y \subset \mathbf{Z}^\nu \setminus \Lambda}} |W_\Phi(X, Y)| .$$

By the concavity of S,

$$\frac{1}{N(\Lambda)} [S(\varrho_\Lambda) - \varrho_\Lambda(U_\Phi)] - P_\Lambda(\Phi) \geqq \frac{-2}{N(\Lambda)} \sup_{\substack{X \subset \Lambda \\ Y \subset \mathbf{Z}^\nu \setminus \Lambda}} |W_\Phi(X, Y)| .$$

An elementary calculation, using (3.3), shows that the right-hand side of this inequality goes to zero as $\Lambda \to \infty$, so our assertion is proved.

Appendix B. Hard Core Continuous Systems

We turn now to the case of hard core continuous systems[20]. The diameter of the hard core is a fixed number $a > 0$. We define K to be the set of subsets X of $\boldsymbol{R}^\nu$ such that if $x, x' \in X$, $x \neq x'$ then $|x - x'| \geqq a$ where $|\cdot|$ is the Euclidean distance. Given a bounded open set $\Lambda \subset \boldsymbol{R}^\nu$ and an integer $n \geqq 0$ we define

$$O_{\Lambda n} = \{X \in K : N(X \cap \Lambda) \geqq n\} \tag{B.1}$$

Similarly for a compact $F \subset \boldsymbol{R}^\nu$ we let

$$O_{Fn} = \{X \in K : N(X \cap K) \leqq n\} . \tag{B.2}$$

The sets $O_{\Lambda n}$ and O_{Fn} generate a topology for which K is compact. We let $\mathfrak{A} = \mathscr{C}(K)$; for a bounded open set $\Lambda \subset \boldsymbol{R}^\nu$ we define $\mathfrak{A}_\Lambda$ to be the subalgebra of $\mathfrak{A}$ constituted by the functions which depend upon X only through $X \cap \Lambda$. The translations of $\boldsymbol{R}^\nu$ define automorphisms τ_x of $\mathfrak{A}$ in an obvious manner.

A state ϱ on $\mathfrak{A}$ is a measure on K. Given a bounded open set $\Lambda \subset \boldsymbol{R}^\nu$ we write

$$\oint_\Lambda dX F(X) = \sum_n \frac{1}{n!} \int_\Lambda dx_1 \ldots \int_\Lambda dx_n F(\{x_1, \ldots, x_n\}) \tag{B.3}$$

where the integrations are with respect to Lebesgue measure and are restricted by $|x_j - x_i| \geqq a$ if $i \neq j$. We write also, as in (3.2)[21],

$$\varrho(A) = \oint_\Lambda dX \int A(X \cup Y) \varrho_\Lambda(X, dY) . \tag{B.4}$$

[20] See Gallavotti and Miracle [9].

[21] Equation (B.4) is imprecise in two respects. First, there need not exist a *function* $X \mapsto \varrho_\Lambda(X, \cdot)$ from configurations in Λ to measures on the set of configurations with no particle in Λ making (B.4) true. This difficulty can easily be remedied by replacing "function" by "measure" in the obvious way; however, we shall be interested only in the case where such a function does exist. Second, even formally, the equation defines for a given X only a measure on the set of configurations Y with no particles in Λ and such that $X \cup Y \in K$. We remedy this defect by defining the measure on the set of configurations Y such that $X \cup Y \notin K$ to be zero.

Let K_n be the subspace of K consisting of sets X such that $N(X) = n$ and let K_F be the topological sum of the K_n. Let $\mathscr{B}^*$ be the space of real continuous functions Φ on K_F satisfying

1. $\Phi(\emptyset) = 0$,

2. translation invariance: $\Phi(X + x) = \Phi(X)$.

We say that $\Phi \in \mathscr{B}^*$ is a *finite range interaction* if there exists $C_\Phi > 0$ such that $\Phi(X) = 0$ whenever the Euclidean diameter of X is larger than C_Φ. We let $B_0 \subset \mathscr{B}^*$ be the space of finite range interactions. Let also $\mathscr{B}^{**}$ be the subspace of $\mathscr{B}^*$ constituted by those Φ such that

$$\|\Phi\| = \sup_{X \in K,\, X \ni 0} \sum_{Y \subset X,\, N(Y) < +\infty} |\Phi(Y)| < +\infty . \tag{B.5}$$

Finally let $\mathscr{B}$ be the closure of $\mathscr{B}_0$ in $\mathscr{B}^{**}$ with respect to the norm (B.5). The elements of $\mathscr{B}$ are taken as the *interactions* of hard core continuous systems.

If $X \in K_F$ we retain the definition (3.4) of U_Φ. If Λ is bounded open in $\boldsymbol{R}^\nu$, we define P_Λ by

$$P_\Lambda(\Phi) = V(\Lambda)^{-1} \log \oint_\Lambda dX \exp[-U_\Phi(X)] \tag{B.6}$$

where $V(\Lambda)$ is the Lebesgue measure of Λ and the notation (B.3) has been used. We define P by (3.6) where Λ tends to infinity in the sense of van Hove, i.e. $V(\Lambda) \to \infty$ and, for all $\delta > 0$, $V(\Lambda)^{-1} V_\delta(\Lambda) \to 0$ where $V_\delta(\Lambda)$ is the Lebesgue measure of the set of points of $\boldsymbol{R}^\nu$ with Euclidean distance to the boundary of Λ less than δ.

We choose a continuous function $\varphi \geqq 0$ on $\boldsymbol{R}^\nu$ such that $\int \varphi(x)\, dx = 1$ and $\varphi(x) = 0$ for $|x| > \frac{1}{2} a$. We let also $\varphi(\{x_1, \ldots, x_n\}) = \sum_{i=1}^{n} \varphi(x_i)$ and, by analogy with (3.7) we define

$$A_\Psi(X) = \sum_{Y \subset X} \varphi(Y) \frac{\Psi(Y)}{N(Y)} . \tag{B.7}$$

With this modification we accept Definition 3.1. for an *invariant equilibrium state*. Theorem 3.2 is then replaced by the following result.

B.1. Theorem. *Let Λ be a bounded open subset of $\boldsymbol{Z}^\nu$; let $X, Y \in K$, with $X \subset \Lambda$, $Y \subset \boldsymbol{Z}^\nu \setminus \Lambda$, and define*

$$f_\Lambda(X, Y) = \exp\left[-\sum_{S \subset X \cup Y;\, S \cap X \neq \emptyset} \Phi(S)\right] \quad \text{if} \quad X \cup Y \in K$$
$$= 0 \quad \text{if} \quad X \cup Y \notin K . \tag{B.8}$$

An invariant state ϱ is an invariant equilibrium state if and only if, for all Λ and all $X \subset \Lambda$,

$$\varrho_\Lambda(X, dY) = f_\Lambda(X, Y)\, \varrho_\Lambda(\emptyset, dY) \tag{B.9}$$

where the notation (B.4) *has been used.*

The second part of the proof of Theorem 3.2 may be adapted with only minor changes in notation to apply to the case at hand. A similar modification can be carried out on the first part of the proof, using the following lemma and an analogous lemma for sequences of states on a fixed $\mathfrak{A}_{\Lambda'}$ satisfying the analogue of (A.5) for a fixed finite-range interaction.

B.2. Lemma. *Let Φ_n be a sequence in $\mathscr{B}$ converging to Φ, and for each n let ϱ_n be a state satisfying* (B.9) *for the interaction Φ_n. Assume that ϱ_n converges weakly to ϱ. Then ϱ satisfies* (B.9) *with the interaction Φ.*

We let $\mathfrak{A}_\infty$ denote the C^* algebra of all (bounded) Borel functions on K which are uniform limits of sequences of bounded Borel functions each of which depends on X only through $X \cap \Lambda$ for some bounded open Λ. Now (B.9) may be re-expressed in the following way: For any $A \in \mathscr{C}(K)$,

$$\begin{aligned} \varrho(A) &= \oint_\Lambda dX \int_{Y \cap \Lambda = \emptyset} f_\Lambda(X, Y)\, A(X \cup Y)\, \varrho(dY) \\ &= \int \varrho(dY)\, A_{\Phi,\Lambda}(Y) \end{aligned} \tag{B.10}$$

where

$$\begin{aligned} A_{\Phi,\Lambda}(Y) &= \oint_\Lambda dX\, f_\Lambda(X, Y)\, A(X \cup Y) \quad \text{if} \quad Y \cap \Lambda = \emptyset \\ &= 0 \quad \text{if} \quad Y \cap \Lambda \neq \emptyset\,. \end{aligned} \tag{B.11}$$

(In the above equations, $A(X \cup Y)$ is defined arbitrarily for $X \cup Y \notin K$). The function $A_{\Phi,\Lambda}$ is easily seen to belong to $\mathfrak{A}_\infty$; moreover, $\lim_{n\to\infty} \|A_{\Phi_n,\Lambda} - A_{\Phi,\Lambda}\| = 0$ if $\Phi_n \to \Phi$ in the Banach space $\mathscr{B}$ of interactions. Hence, it will suffice to prove that

$$\lim_{n\to\infty} \varrho_n(B) = \varrho(B) \tag{B.12}$$

for every $B \in \mathfrak{A}_\infty$; we already know that this is true for $B \in \mathscr{C}(K)$.

Now let

$$\begin{aligned} \varrho_{n,\Lambda}(X) &= \int \varrho_{n,\Lambda}(X, dY) \\ &= \int f_\Lambda^{\Phi_n}(X, Y)\, \varrho_{n,\Lambda}(\emptyset, dY); \end{aligned} \tag{B.13}$$

if B is any bounded Borel function on K which depends only on $X \cap \Lambda$, we have:

$$\varrho_n(B) = \oint_\Lambda dX\, \varrho_{n,\Lambda}(X)\, B(X)\,. \tag{B.14}$$

It is not hard to verify from the definition of the space of interactions that there is a constant C_Λ such that

$$f_\Lambda^{\Phi_n}(X, Y) \leqq C_\Lambda$$

for all n, all $X \subset \Lambda$, and all $Y \subset \mathbf{Z}^\nu \backslash \Lambda$. Since

$$\int \varrho_{n,\Lambda}(\emptyset, dY) < 1\,,$$

we get

$$\varrho_{n,\Lambda}(X) < C_\Lambda \tag{B.15}$$

for all n, X.

From (B.15), (B.14), and the fact that (B.12) holds for all B in $\mathfrak{A}_\Lambda$, it follows that (B.12) holds for all bounded Borel functions B depending only on $X \cap \Lambda$ and therefore for all $B \in \mathfrak{A}_\infty$.

From Theorem B.1, it follows that (3.3) is again a reasonable definition of an equilibrium state; with this definition, the set Δ of equilibrium states is again convex and compact. Theorem 3.4 and Proposition 3.5 remain true, their proofs being left unchanged.

Acknowledgement. The first-named author wishes to express his gratitude to Monsieur L. MOTCHANE for his hospitality at the I.H.E.S.

Bibliography

1. BEREZIN, F. A., R. L. DOBRUSHIN, R. A. MINLOS, A. YA. POVZNER, and YA. G. SINAI. Some facts about general properties of thermodynamical potentials and phase transitions of the first kind by low temperatures. Unpublished Report (1966).
2. CHOQUET, G., and P.-A. MEYER: Existence et unicité des représentations intégrales dans les convexes compacts quelconques. Ann. Inst. Fourier **13**, 139—154 (1953).
3. DELL'ANTONIO, G.-F., S. DOPLICHER, and D. RUELLE: A theorem on canonical commutation and anticommutation relations. Commun. Math. Phys. **2**, 223—230 (1966).
4. DOPLICHER, S., G. GALLOVOTTI, and D. RUELLE: Almost periodic states on C*-algebras. Unpublished Report (1966).
5. — D. KASTLER, and D. W. ROBINSON: Covariance algebras in field theory and statistical mechanics. Commun. Math. Phys. **3**, 1—28 (1966).
6. FISHER, M. E.: The theory of condensation and the critical point. Physics **3**, 255—283 (1967).
7. GALLAVOTTI, G., and S. MIRACLE: Statistical mechanics of lattice systems. Commun. Math. Phys. **5**, 317—323 (1967).
8. —, — Correlation functions of a lattice gas. Commun. Math. Phys. **7**, 274—288 (1968).
9. —, — A variational principle for the equilibrium of hard sphere systems. Ann. Inst. Henri Poincaré 8, 287—299 (1968).
10. — —, and D. W. ROBINSON: Analyticity properties of a lattice gas. Phys. Letters 25 A, 493—494 (1967).
11. HAAG, R., D. KASTLER, and L. MICHEL. Central decomposition of ergodic states. Unpublished Report (1968).
12. JACOBS, K.: Lecture notes on ergodic theory. Aarhus: Aarhus Universitet 1963.
13. KASTLER, D., and D. W. ROBINSON: Invariant states in statistical mechanics. Commun. Math. Phys. **3**, 151—180 (1966).
14. LANFORD, O.: The classical mechanics of one-dimensional systems of infinitely many particles. I. An existence theorem. Commun. Math. Phys. **9**, 179—191 (1968). II. Kinetic theory. Commun. Math. Phys. **11**, 257—292 (1969).
15. —, and D. W. ROBINSON: Mean entropy of states in quantum statistical mechanics. J. Math. Phys. **9**, 1120—1125 (1968).

16. Landford, O., and D. W. Robinson: Statistical mechanics of quantum spin systems III. Commun. Math. Phys. **9**, 327—338 (1968).
17. —, and D. Ruelle: Integral Representations of Invariant States on B*-Algebras. J. Math. Phys. **8**, 1460—1463 (1967).
18. Powers, R. T.: Representations of uniformly hyperfinite algebras and their associated von Neumann rings. Ann. Math. **86**, 138—171 (1967).
19. Robinson, D. W., and D. Ruelle: Extremal invariant states. Ann. Inst. Henri Poincaré **6**, 299—310 (1967).
20. Ruelle, D.: Correlation functions of classical gases. Ann. Phys. **25**, 109—120 (1963).
21. — Quantum statistical mechanics and canonical commutation relations. Cargèse lectures in theoretical physics 1965. New York: Gordon & Breach 1967.
22. — States of physical systems. Commun. Math. Phys. **3**, 133—150 (1966).
23. — States of classical statistical mechanics. J. Math. Phys. **8**, 1657—1668 (1967).
24. — A variational formulation of equilibrium statistical mechanics and the Gibbs phase rule. Commun. Math. Phys. **5**, 324—329 (1967).
25. — Statistical mechanics, rigorous results. New York: Benjamin (1969).
26. Sinai: Ya. G.: Probabilistic ideas in ergodic theory. International Congress of Mathematicians, Stockholm 1962, 540—559, and Amer. Math. Soc. Transl. (2) **31**, 62—81 (1963).
27. Størmer, E.: Large groups of automorphisms of C*-algebras. Commun. Math. Phys. **5**, 1—22 (1967).

O. E. Lanford
Department of Mathematics
University of California
Berkeley, California, USA

D. Ruelle
Institut des Hautes
Etudes Scientifiques
F 91 Bures-sur-Yvette

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Mean Entropy of States in Quantum-Statistical Mechanics

OSCAR E. LANFORD III
University of California, Berkeley, California
AND
DEREK W. ROBINSON
CERN, Geneva

(Received 6 September 1967)

The equilibrium states for an infinite system of quantum mechanics may be represented by states over suitably chosen C^* algebras. We consider the problem of associating an entropy with these states and finding its properties, such as positivity, subadditivity, etc. For the states of a quantum-spin system, a mean entropy is defined and it is shown that this entropy is affine and upper semicontinuous.

1. INTRODUCTION

In the algebraic theory of statistical mechanics the class of possible equilibrium states is defined as the subclass K of states ρ, over the C^* algebra $\mathcal{A}$ of local observables, which satisfy certain subsidiary conditions of a physical origin. Firstly, it is assumed that the theory is invariant under a symmetry group G (the translation group R^ν, for example), and the states $\rho \in K$ considered are taken to be G invariant. Secondly, as one wishes to describe only systems with a finite number of particles in each finite subsystem, extra conditions must be introduced. The consequence of these latter "finite mean density" conditions can be described as follows: If $\Lambda \subset R^\nu$ is an open set of compact closure and $\mathcal{A}(\Lambda) \subset \mathcal{A}$ is the corresponding subalgebra of strictly local observables, then a state $\rho \in K$ must be such that its restriction to each $\mathcal{A}(\Lambda)$ is described by a density matrix ρ_Λ acting on a Hilbert space $\mathcal{H}_\Lambda$. As a direct result of this last property we may introduce, for each $\rho \in K$, a family of entropies $S(\rho_\Lambda)$ by the definition

$$S(\rho_\Lambda) = -\mathrm{Tr}_{\mathcal{H}_\Lambda}(\rho_\Lambda \log \rho_\Lambda).$$

Consequently, we may study properties of $S(\rho_\Lambda)$, attempt to introduce for each $\rho \in K$ an entropy per unit volume $S(\rho)$, and, subsequently, analyse the linearity and continuity properties, etc., of $S(\rho)$.

The program outlined above was recently completed by Ruelle, in collaboration with one of the present authors (D. W. R.), in the framework of classical statistical mechanics.[1] The purpose of the present paper is to attempt the same program for quantum-statistical mechanics. In this latter setting many difficulties arise due to noncommutativity, and our results are complete only for quantum-spin systems. In the general case many interesting problems remain open.

[1] D. W. Robinson and D. Ruelle, Commun. Math. Phys. **5**, 288 (1967).

2. GENERAL FORMULATION

We want to investigate both continuous infinite quantum-statistical systems and lattice systems. Thus, we consider a C^* algebra $\mathcal{A}$ and a collection $\{\mathcal{A}(\Lambda)\}$ of C^* subalgebras of $\mathcal{A}$, where Λ runs over:

(i) the bounded open sets in R^ν for continuous systems;

(ii) the finite subsets of Z^ν for lattice systems.

We suppose that these subalgebras satisfy the following axioms:

(1) $\mathcal{A}(\Lambda_1) \subset \mathcal{A}(\Lambda_2)$ if $\Lambda_1 \subset \Lambda_2$.

(2) For each Λ, $\mathcal{A}(\Lambda)$ is isomorphic to $\mathfrak{L}(\mathcal{H}_\Lambda)$ for some Hilbert space $\mathcal{H}_\Lambda$. We will usually identify $\mathcal{A}(\Lambda)$ with $\mathfrak{L}(\mathcal{H}_\Lambda)$, although this is not strictly compatible with axiom (1).

(3) $\mathcal{A}$ is the norm closure of $\cup_\Lambda \mathcal{A}(\Lambda)$.

(4) $\mathcal{A}(\Lambda_1 \cup \Lambda_2)$ is generated by $\mathcal{A}(\Lambda_1) \cup \mathcal{A}(\Lambda_2)$ in the weak operator topology on $\mathfrak{L}(\mathcal{H}_{\Lambda_1 \cup \Lambda_2})$.

(5) Let G denote the group of translations, i.e., $G = Z^\nu$ for lattice systems and $G = R^\nu$ for continuous systems. Then G acts on $\mathcal{A}$ by automorphisms τ_x in such a way that $\tau_x(\mathcal{A}(\Lambda)) = \mathcal{A}(\Lambda + x)$ for all regions Λ and translations x.

Finally, we need a condition expressing the independence of observables belonging to disjoint regions. This condition may take one of two forms, depending on whether we are considering bosons or fermions:

(6) Either

(a) If Λ_1 and Λ_2 are any two disjoint regions, then $\mathcal{A}(\Lambda_1)$ commutes with $\mathcal{A}(\Lambda_2)$; or

(b) Each $\mathcal{A}(\Lambda)$ is generated by a set of creation and annihilation operators satisfying the canonical anticommutation relations, and, if Λ_1 and Λ_2 are disjoint regions, the creation and annihilation operators for Λ_1 anticommute with those for Λ_2.

These axioms describe several kinds of physical systems:

(1) Ordinary continuous quantum systems, either bosons or fermions.

(2) Quantum lattice systems, again either bosons or fermions, with finitely many creation and annihilation operators associated with each lattice site. For fermion lattice systems, $\mathcal{H}_\Lambda$ is finite-dimensional for each finite set Λ, but for boson systems this is, of course, not true.

(3) Quantum-spin systems. In this case, $\mathcal{H}_\Lambda$ is finite-dimensional for each bounded region Λ, but the different unit rays in $\mathcal{H}_{\{x\}}$, where x is a lattice point, are interpreted as describing different polarization states of a particle localized at x rather than varying occupation numbers for that point. We will assume that such systems satisfy axiom (6a).

The statistical-mechanical states of $\mathcal{A}$ are those which, when restricted to an $\mathcal{A}(\Lambda)$, are given by a density matrix. In other words, such a state ρ defines, for each region Λ, a positive operator ρ_Λ on $\mathcal{H}_\Lambda$, with $\mathrm{Tr}_{\mathcal{H}_\Lambda}(\rho_\Lambda) = 1$, such that

$$\rho(A) = \mathrm{Tr}_{\mathcal{H}_\Lambda}(\rho_\Lambda A),$$

if $A \in \mathcal{A}(\Lambda) = \mathfrak{L}(\mathcal{H}_\Lambda)$. This statement imposes no restriction on ρ if $\mathcal{H}_\Lambda$ is finite-dimensional; otherwise, it corresponds to the requirement that there be only finitely many particles in the region Λ.

Every statistical-mechanical state ρ defines a family $\{\rho_\Lambda\}$ of density matrices. Conversely, the assignment of a density matrix to each bounded region defines a statistical-mechanical state on $\mathcal{A}$, provided that the assignment satisfies the obvious compatibility condition that, if $\Lambda_1 \subset \Lambda_2$ and if $A \in \mathcal{A}(\Lambda_1)$, then

$$\mathrm{Tr}_{\mathcal{H}_{\Lambda_1}}(\rho_{\Lambda_1} A) = \mathrm{Tr}_{\mathcal{H}_{\Lambda_2}}(\rho_{\Lambda_2} A).$$

We can reformulate the compatibility condition as follows: If $\Lambda_1 \subset \Lambda_2$, then $\mathcal{A}(\Lambda_1)$ is a type I factor contained in $\mathcal{A}(\Lambda_2) = \mathfrak{L}(\mathcal{H}_{\Lambda_2})$. Hence, we may factorize

$$\mathcal{H}_{\Lambda_2} = \mathcal{H}_{\Lambda_1} \otimes \mathcal{H}'$$

in such a way that an operator A in $\mathcal{A}(\Lambda_1) = \mathfrak{L}(\mathcal{H}_{\Lambda_1})$ is identified with the operator $A \otimes 1$ on $\mathcal{H}_{\Lambda_1} \otimes \mathcal{H}'$. [The space $\mathcal{H}'$ may be identified with $\mathcal{H}_{\Lambda_2 - \Lambda_1}$, but operators in $\mathcal{A}(\Lambda_2 - \Lambda_1)$ do not factorize as nicely as those in $\mathcal{A}(\Lambda_1)$ unless algebras for disjoint regions commute. See below.] The compatibility condition may now be formulated as

$$\rho_{\Lambda_1} = \mathrm{Tr}_{\mathcal{H}'}(\rho_{\Lambda_2}),$$

where $\mathrm{Tr}_{\mathcal{H}'}$ means the partial trace with respect to $\mathcal{H}'$, i.e., if $\{\varphi_i\}$ is an orthonormal basis for $\mathcal{H}_{\Lambda_1}$ and $\{\psi_i\}$ is an orthonormal basis for $\mathcal{H}'$, then

$$(\rho_{\Lambda_1}\varphi_i, \varphi_k) = \sum_{j=1}^{\infty} [\rho_{\Lambda_2}(\varphi_i \otimes \psi_j), \varphi_k \otimes \psi_j].$$

The condition that a state be translation-invariant may easily be formulated in terms of the corresponding system of density matrices. For any region Λ and any translation x, τ_x is an isomorphism of $\mathcal{A}(\Lambda)$ onto $\mathcal{A}(\Lambda + x)$. Since $\mathcal{A}(\Lambda)$ is identified with $\mathfrak{L}(\mathcal{H}_\Lambda)$ and $\mathcal{A}(\Lambda + x)$ with $\mathfrak{L}(\mathcal{H}_{\Lambda+x})$, there is a unitary operator $U_{\Lambda,x}$ from $\mathcal{H}_\Lambda$ to $\mathcal{H}_{\Lambda+x}$ which induces this isomorphism, and $U_{\Lambda,x}$ is determined up to a multiplicative constant. Then the state defined by the system $\{\rho_\Lambda\}$ of density matrices is translation-invariant if and only if

$$\rho_{\Lambda+x} = U_{\Lambda,x}\rho_\Lambda U_{\Lambda,x}^{-1},$$

for all regions Λ and translations x.

We now want to make a more careful analysis of the relation of $\rho_{\Lambda_1 \cup \Lambda_2}$ to ρ_{Λ_1} and ρ_{Λ_2} when Λ_1 and Λ_2 are disjoint regions. We have already remarked that the inclusion of $\mathcal{A}(\Lambda_1)$ in $\mathcal{A}(\Lambda_1 \cup \Lambda_2)$ gives a factorization of $\mathcal{H}_{\Lambda_1 \cup \Lambda_2}$ as $\mathcal{H}_{\Lambda_1} \otimes \mathcal{H}'$, where operators in $\mathcal{A}(\Lambda_1)$ go over into operators of the form $A \otimes 1$. If we are considering a boson system or a spin system, then the commutant of $\mathcal{A}(\Lambda_1)$ in $\mathcal{A}(\Lambda_1 \cup \Lambda_2)$ is precisely $\mathcal{A}(\Lambda_2)$. In this case, there is an essentially unique way to identify $\mathcal{H}'$ with $\mathcal{H}_{\Lambda_2}$, and operators in $\mathcal{A}(\Lambda_2)$ take the form $1 \otimes A$ on $\mathcal{H}_{\Lambda_1} \otimes \mathcal{H}_{\Lambda_2}$. Hence we have

$$\rho_{\Lambda_1} = \mathrm{Tr}_{\mathcal{H}_{\Lambda_2}}(\rho_{\Lambda_1 \cup \Lambda_2}), \quad \rho_{\Lambda_2} = \mathrm{Tr}_{\mathcal{H}_{\Lambda_1}}(\rho_{\Lambda_1 \cup \Lambda_2}).$$

For fermion systems, although $\mathcal{H}'$ has the same dimension as $\mathcal{H}_{\Lambda_2}$, there is no unique sensible way to identify $\mathcal{H}'$ with $\mathcal{H}_{\Lambda_2}$. Nevertheless, by using the special structure of fermion systems, we can construct a useful identification of $\mathcal{H}'$ with $\mathcal{H}_{\Lambda_2}$. This we do as follows: Let N_1 and N_2 denote the number operators for the regions Λ_1 and Λ_2, respectively. Then a simple calculation with the anticommutation relations shows that the commutant of $\mathcal{A}(\Lambda_1)$ in $\mathcal{A}(\Lambda_1 \cup \Lambda_2)$ is precisely $(-1)^{N_1 N_2}\mathcal{A}(\Lambda_2)(-1)^{N_1 N_2}$. Therefore, we can identify $\mathcal{H}'$ with $\mathcal{H}_{\Lambda_2}$ in such a way that if A is in $\mathcal{A}(\Lambda_2) = \mathfrak{L}(\mathcal{H}_{\Lambda_2})$, then A goes over into

$$(-1)^{N_1 N_2}(1 \otimes A)(-1)^{N_1 N_2},$$

in $\mathfrak{L}(\mathcal{H}_{\Lambda_1} \otimes \mathcal{H}_{\Lambda_2})$. With this identification we have

$$\rho_{\Lambda_1} = \mathrm{Tr}_{\mathcal{H}_{\Lambda_2}}(\rho_{\Lambda_1 \cup \Lambda_2}),$$
$$\rho_{\Lambda_2} = \mathrm{Tr}_{\mathcal{H}_{\Lambda 1}}[(-1)^{N_1 N_2}\rho_{\Lambda_1 \cup \Lambda_2}(-1)^{N_1 N_2}].$$

The second of these equations can be simplified if we assume that ρ is an even state of $\mathcal{A}$. By definition, an element A of some $\mathcal{A}(\Lambda)$ is *odd* if

$$(-1)^{N(\Lambda)} A (-1)^{N(\Lambda)} = -A,$$

where $N(\Lambda)$ is the number operator for the region Λ. A state ρ of $\mathcal{A}$ is *even* if ρ vanishes on every odd element of every $\mathcal{A}(\Lambda)$. If ρ is now an even state and A is an element of $\mathcal{A}(\Lambda_2)$, then

$$\rho(A) = \mathrm{Tr}_{\mathcal{H}_{\Lambda_1} \otimes \mathcal{H}_{\Lambda_2}}[\rho_{\Lambda_1 \cup \Lambda_2}(-1)^{N_1 N_2}(1 \otimes A)(-1)^{N_1 N_2}]$$
$$= \mathrm{Tr}_{\mathcal{H}_{\Lambda_1} \otimes \mathcal{H}_{\Lambda_2}}[\rho_{\Lambda_1 \cup \Lambda_2}(1 \otimes A)].$$

To prove this equation, note that we can write A as the sum of an even part and an odd part, that the odd part contributes nothing to $\rho(A)$, and that the even part commutes with $(-1)^{N_1 N_2}$. Collecting these results we have:

Proposition 1: Let ρ be a statistical-mechanical state of the C^* algebra $\mathcal{A}$, and let $\{\rho_\Lambda\}$ be the corresponding system of density matrices. If $\mathcal{A}$ is the algebra for a fermion system, we further assume that ρ is even. Then, if Λ_1 and Λ_2 are disjoint regions, we may identify $\mathcal{H}_{\Lambda_1 \cup \Lambda_2}$ with $\mathcal{H}_{\Lambda_1} \otimes \mathcal{H}_{\Lambda_2}$ in such a way that

$$\rho_{\Lambda_1} = \mathrm{Tr}_{\mathcal{H}_{\Lambda_2}}(\rho_{\Lambda_1 \cup \Lambda_2}) \quad \rho_{\Lambda_2} = \mathrm{Tr}_{\mathcal{H}_{\Lambda_1}}(\rho_{\Lambda_1 \cup \Lambda_2}).$$

Note that if we are dealing with a fermion system, then translation invariance implies that the state ρ is even. To show this[2] let A be an odd element of some $\mathcal{A}(\Lambda)$ and let x be a translation large enough so that $\Lambda + nx$ does not intersect Λ for $n = 1, 2, 3, \cdots$ Let

$$A_N = \frac{1}{N} \sum_{n=0}^{N-1} \tau_{nx} A.$$

Now

$$\begin{aligned} \|A_N\|^2 = \|A_N^* A_N\| &\leq \|\{A_N^*, A_N\}\| \\ &\leq \frac{1}{N^2} \sum_{n,m=0}^{N-1} \|\{(\tau_{nx}A)^*, \tau_{mx}A\}\| \\ &\leq \frac{2}{N} \|A\|^2, \end{aligned}$$

where the last inequality is a consequence of

$$\{(\tau_{nx}A)^*, \tau_{mx}A\} = 0, \quad \text{for} \quad n \neq m.$$

However, by translation invariance,

$$\rho(A) = \rho(A_N) = \lim_{N \to \infty} \rho(A_N).$$

But as $\lim_{N \to \infty} \|A_N\| = 0$, then $\lim_{N \to \infty} \rho(A_N) = 0$ and thus $\rho(A) = 0$.

Given a statistical-mechanical state ρ and a region Λ, we can define the entropy of the region Λ as follows:

$S(\rho_\Lambda) = +\infty$, if $\rho_\Lambda \log \rho_\Lambda$ is not of trace class on $\mathcal{H}_\Lambda$

$= -\mathrm{Tr}_{\mathcal{H}_\Lambda}(\rho_\Lambda \log \rho_\Lambda)$, otherwise.

In defining the operator $\rho_\Lambda \log \rho_\Lambda$ we use the usual convention $x \log x = 0$, for $x = 0$.

3. BASIC INEQUALITIES FOR THE ENTROPY

Lemma 1[3]*:* If A and B are positive, self-adjoint, trace-class operators on a Hilbert space $\mathcal{H}$, then

$$\mathrm{Tr}_{\mathcal{H}}[A \log A - A \log B - A + B] \geq 0.$$

Proof: Let $\psi_i(\varphi_i)$ be a complete orthonormal set of eigenfunctions of $A(B)$ and let $a_i(b_i)$ be the corresponding eigenvalues. Let $U = (U_{ij})$ be a unitary mapping defined by

$$\psi_i = \sum_j U_{ij} \varphi_j.$$

Now

$$\begin{aligned} (\psi_i | A \log A - A \log B | \psi_i) &= a_i \left\{ \log a_i - \sum_j |U_{ij}|^2 \log b_j \right\} \\ &\geq a_i \left\{ \log a_i - \log \sum_j |U_{ij}|^2 b_j \right\} \\ &\geq a_i - \sum_j |U_{ij}|^2 b_j \\ &= (\psi_i | A - B | \psi_i), \end{aligned}$$

where, to obtain the first inequality, we have used the concavity of the logarithm and, to obtain the second inequality, we have used

$$\log x \geq 1 - 1/x \quad (x > 0).$$

The result follows by summation.

Lemma 2: If A and B be positive, self-adjoint operators on a Hilbert space $\mathcal{H}$, then, for $1 \geq \alpha \geq 0$,

$$[\alpha A + (1 - \alpha)B] \log [\alpha A + (1 - \alpha)B] \leq \alpha A \log A + (1 - \alpha) B \log B,$$

and furthermore

$$A \geq B \geq 0 \quad \text{implies} \quad \log A \geq \log B.$$

The statements of the lemma are special consequences of the theory of convex and monotone operator functions initially developed by Löwner.[4] For further results, the reader may consult Ref. 5. The details of the application of the general theory to the case at hand are worked out in Refs. 6 and 7. Moreover, we do not need the full force of the first inequality of the lemma, but only the inequality obtained by taking its trace; this latter inequality can be proved without use of the general theory of convex operator functions.[7,8]

We remark that Lemma 1 may be deduced from the first statement of Lemma 2. We preferred, however, to give the simple straightforward proof reproduced above.

---

[2] This proof was independently discovered by R. T. Powers.

[3] This lemma, together with its proof, was communicated to one of us (D. W. R.) by Professor R. Jost, who attributed it to O. Klein. If $\mathrm{Tr}(A) = \mathrm{Tr}(B) = 1$, this lemma is a particular case of Theorem 1 of H. Umegaki, Kodai Math. Sem. Rep. **14**, 59 (1962).

[4] C. Löwner, Math. Z. **38**, 177 (1934).

[5] J. Bendat and S. Sherman, Trans. Am. Math. Soc. **79**, 58 (1955).

[6] M. Nakamura and H. Umegaki, Proc. Japan Acad. **37**, 149 (1961).

[7] C. Davis, Proc. Japan Acad. **37**, 533 (1961).

[8] I. E. Segal, J. Math. & Mech. **9**, 623 (1960).

The preceding lemmas may now be used to deduce the following results for the quantum entropy, specializations of which appear in Refs. 9 and 10.

Theorem 1: Let ρ be a statistical-mechanical state of the C^* algebra $\mathcal{A}$, and let $\{\rho_\Lambda\}$ be the corresponding system of density matrices. If $\mathcal{A}$ is the algebra for a fermion system, we assume further that ρ is an even state. Then the associated entropy $S(\rho_\Lambda)$ is a positive set function, i.e.,

$$S(\rho_\Lambda) \geq 0,$$

which is subadditive, i.e.,

$$S(\rho_{\Lambda_1 \cup \Lambda_2}) \leq S(\rho_{\Lambda_1}) + S(\rho_{\Lambda_2}),$$

if

$$\Lambda_1 \cap \Lambda_2 = \phi.$$

Further, if $\{\rho_\Lambda^{(1)}\}$ and $\{\rho_\Lambda^{(2)}\}$ are two families of density matrices and $1 \geq \alpha \geq 0$, then

$$\begin{aligned}\alpha S(\rho_\Lambda^{(1)}) + (1-\alpha) S(\rho_\Lambda^{(2)}) &\leq S[\alpha\rho_\Lambda^{(1)} + (1-\alpha)\rho_\Lambda^{(2)}] \\ &\leq \alpha S(\rho_\Lambda^{(1)}) + (1-\alpha) S(\rho_\Lambda^{(2)}) \\ &\quad - \alpha \log \alpha - (1-\alpha)\log(1-\alpha).\end{aligned}$$

Proof: The positivity of $S(\rho_\Lambda)$ is an immediate consequence of the normalization of ρ_Λ and the fact that

$$-\lambda \log \lambda \geq 0, \quad 1 \geq \lambda \geq 0.$$

The subadditivity property follows from Lemma 1, Proposition 1, and the identification $\mathcal{H} = \mathcal{H}_{\Lambda_1} \otimes \mathcal{H}_{\Lambda_2}$, $A = \rho_{\Lambda_1 \cup \Lambda_2}$, and $B = \rho_{\Lambda_1} \otimes \rho_{\Lambda_2}$. The final inequality is a consequence of Lemma 2. The lower bound is immediately obtained from the first statement of that lemma, while the upper bound is obtained from the second statement as follows: We have

$$\alpha\rho_\Lambda^{(1)} + (1-\alpha)\rho_\Lambda^{(2)} \geq \alpha\rho_\Lambda^{(1)} \geq 0,$$

and hence

$$\log[\alpha\rho_\Lambda^{(1)} + (1-\alpha)\rho_\Lambda^{(2)}] \geq \log \alpha\rho_\Lambda^{(1)}.$$

Thus

$$\begin{aligned}-\alpha \operatorname{Tr}[\rho_\Lambda^{(1)} \log(\alpha\rho_\Lambda^{(1)} + (1-\alpha)\rho_\Lambda^{(2)})] &\leq -\alpha \operatorname{Tr}[\rho_\Lambda^{(1)} \log \alpha\rho_\Lambda^{(1)}] \\ &= -\alpha \operatorname{Tr}[\rho_\Lambda^{(1)} \log \rho_\Lambda^{(1)}] - \alpha \log \alpha.\end{aligned}$$

Similarly,

$$\begin{aligned}&-(1-\alpha)\operatorname{Tr}[\rho_\Lambda^{(2)} \log(\alpha\rho_\Lambda^{(1)} + (1-\alpha)\rho_\Lambda^{(2)})] \\ &\quad \leq -(1-\alpha)\operatorname{Tr}[\rho_\Lambda^{(2)} \log(1-\alpha)\rho_\Lambda^{(2)}] \\ &\quad = -(1-\alpha)\operatorname{Tr}[\rho_\Lambda^{(2)} \log \rho_\Lambda^{(2)}] - (1-\alpha)\log(1-\alpha).\end{aligned}$$

Adding the last two inequalities yields the desired result.

4. MEAN ENTROPY—THE QUANTUM LATTICE SYSTEM

The next desirable aim would be to define an entropy per unit volume, or mean entropy, by establishing, under suitable restrictions, the existence of $S(\rho_\Lambda)/V(\Lambda)$ in the limit of Λ increasing such that the volume $V(\Lambda) \to \infty$. Unfortunately, we are at present able to do this only for a quantum lattice system, and even then it is not possible to establish the existence of the limit in the most desirable generality. A possible means of improving our results is discussed in the concluding section.

Let us define for $a = (a_1, \cdots, a_\nu) \in Z^\nu$ and $a_1 > 0, \cdots, a_\nu > 0$ the parallelepiped $\Lambda(a)$ by

$$\Lambda(a) = \{x \in Z^\nu;\ 0 < x_i \leq a_i \quad \text{for} \quad i = 1, 2, \cdots, \nu\}.$$

The measure (volume) $V[\Lambda(a)]$ of $\Lambda(a)$ is then given by $\prod_{i=1}^{\nu} a_i$.

Theorem 2: If the family $\rho = \{\rho_\Lambda\}$ of density matrices of a quantum lattice system satisfies the conditions of normalization, compatibility, and translation invariance, then the mean entropy

$$S(\rho) = \lim_{a_1, \ldots, a_\nu \to \infty} \frac{S(\rho_{\Lambda(a)})}{V(\Lambda(a))}$$

exists, and in fact

$$S(\rho) = \inf_{a_1, \ldots, a_\nu} \frac{S(\rho_{\Lambda(a)})}{V(\Lambda(a))}.$$

Further, $S(\rho)$ is an affine function. Explicitly, if $\rho^{(1)} = \{\rho_\Lambda^{(1)}\}$ and $\rho^{(2)} = \{\rho_\Lambda^{(2)}\}$ are two appropriate families of density matrices and $1 \geq \alpha \geq 0$, then

$$S[\alpha\rho^{(1)} + (1-\alpha)\rho^{(2)}] = \alpha S(\rho^{(1)}) + (1-\alpha) S(\rho^{(2)}).$$

Proof: By translation invariance, $S[\Lambda(a)]$ is a function of $a_1, \cdots, a_\nu$ only. Moreover, if we are dealing with a fermion system, translation invariance implies that the state ρ is even (see Sec. 2). But if we introduce a function $S(a_1, \cdots, a_\nu)$ through the definition

$$S(a_1, \cdots, a_\nu) = S[\Lambda(a)],$$

the subadditivity of $S(\Lambda)$ implies that $S(a_1, \cdots, a_\nu)$ is subadditive in each argument a_i $(1 \leq i \leq \nu)$ separately, i.e.,

$$\begin{aligned}S(a_1, \cdots, a_i^{(1)} + a_{iv}^{(2)}, \cdots, a_\nu) &= S(a_1, \cdots, a_i^{(1)}, \cdots a_\nu) \\ &\quad + S(a_i, \cdots, a_i^{(2)}, \cdots, a_\nu).\end{aligned}$$

[9] M. Delbrück and G. Molière, Abhandl. Preuss. Akad. P. 1 (1937).
[10] R. Jost, Helv. Phys. Acta **20**, 491 (1947).

A standard argument [cf. Ref. 11, Lemma A1] establishes the existence of

$$S(\rho) = \lim_{a_1,\cdots,a_\nu\to\infty} \frac{S(a_1,\cdots,a_\nu)}{a_1a_2,\cdots,a_\nu} = \inf_{a_1,\cdots,a\nu} \frac{S(a_1,\cdots,a_\nu)}{a_1a_2,\cdots,a_\nu}.$$

The affine property of $S(\rho)$ follows from the last statement of Theorem 1, if one takes $\Lambda = \Lambda(a)$, divides by $V[\Lambda(a)]$, and takes the appropriate limit.

5. PROPERTIES OF THE MEAN ENTROPY

For fermion lattice systems and spin systems we can exploit the finite dimensionality of the $\mathcal{H}_\Lambda$'s to prove some additional properties of the mean entropy.

Theorem 3: Let $\mathcal{A}$ be the C^* algebra for a fermion lattice system or a spin lattice system. If x is a lattice point, let N denote the dimension of $\mathcal{H}_{\{x\}}$. Equip the set of states of $\mathcal{A}$ with the weak * topology. Then

(1) For any invariant state ρ of $\mathcal{A}$, $0 \leq S(\rho) \leq \text{Log } N$.

(2) The mean entropy is an upper semicontinuous function on the set of invariant states of $\mathcal{A}$. If F is any closed subset of the set of invariant states of $\mathcal{A}$, then the restriction of the mean entropy to F attains its maximum.

(3) If ρ is an invariant state of $\mathcal{A}$, and if μ_ρ is the unique probability measure with barycenter ρ concentrated on the extremal invariant states of $\mathcal{A}$, then

$$S(\rho) = \int d\mu_\rho(\rho')S(\rho').$$

In physical language, statement 3 says that if ρ is an average of pure phases, then the mean entropy of ρ is the same average of the entropies of the pure phases making up ρ. For the existence and uniqueness of the measure μ_ρ, see Ref. 12, Theorem 2, or Ref. 13, Theorem 3. We remark that, under the hypotheses of this theorem, $\mathcal{A}$ is separable so there are no technical difficulties about the sense in which the measure is concentrated on the extremal invariant states.

Proof: For any finite set Λ of lattice points, the dimension of $\mathcal{H}_\Lambda$ is $N^{V(\Lambda)}$. Now

$$S(\rho_\Lambda) = -\sum_{i=1}^{N^{V(\Lambda)}} \lambda_i \log \lambda_i,$$

where the λ_i are the eigenvalues of ρ_Λ. By elementary estimates, if $\mu_1,\cdots,\mu_n$ are positive real numbers with $\sum_i \mu_i = 1$, then

$$-\sum_{i=1}^{n} \mu_i \log \mu_i \leq \log n.$$

Hence,

$$S(\rho_\Lambda) \leq \log (N^{V(\Lambda)}) = V(\Lambda) \log N.$$

Dividing by $V(\Lambda)$ and taking the limit of infinite volume gives

$$S(\rho) \leq \log N.$$

Since $S(\rho)$ is nonnegative by definition, we have proved part (1) of Theorem 3.

To prove part (2), observe first that the ρ_Λ's are continuous functions of ρ and that the eigenvalues of ρ_Λ vary continuously with ρ_Λ by perturbation theory. Since $-\lambda \log \lambda$ is a continuous function of λ,

$$S(\rho_\Lambda) = -\sum_i \lambda_i \log \lambda_i$$

is a continuous function of ρ. But

$$S(\rho) = \inf_\Lambda \left\{\frac{S(\rho_\Lambda)}{V(\Lambda)}\right\},$$

where the infimum is to be taken over all rectangles. Thus, $S(\rho)$ is the infimum of a family of continuous functions and is therefore upper semicontinuous. In particular, if F is any closed set of invariant states on $\mathcal{A}$, then the restriction of S to F takes on its maximum, since any upper semicontinuous function on a compact set takes on its maximum.

Furthermore, since $S(\rho)$ is both affine and upper semicontinuous, it respects barycentric decompositions. More precisely, if μ is any probability measure on the set of invariant states of $\mathcal{A}$, and if the barycenter of μ is ρ, then

$$S(\rho) = \int d\mu(\rho')S(\rho').$$

(This follows from Lemma 10 of Ref. 14.) In particular, if μ_ρ is the unique decomposition of ρ into extremal invariant states,[15] then the above equation holds. This proves part (3) and completes the proof of the theorem.

6. CONCLUSION

While we have been able to obtain most of the desired results concerning the entropy of a quantum spin system, the position is less satisfactory in other cases. The main gap in the development is the failure to establish the existence of the mean entropy $S(\rho)$ under general circumstances. In classical statistical

[11] D. Ruelle, J. Math. Phys. **6**, 201 (1965).
[12] D. Kastler and D. W. Robinson, Commun. Math. Phys. **3**, 51 (1966).
[13] O. E. Lanford and D. Ruelle, J. Math. Phys. **8**, 1460 (1967).
[14] G. Choquet and P. A. Mayer, Ann. Inst. Fourier **13**, 139 (1963).
[15] Note that the uniqueness proofs given in Refs. 12 and 13 for such decompositions are valid even for Fermi systems. In the Fermi case, $\mathcal{A}$ is $R\nu$ (or $Z\nu$) Abelian, in the sense of Ref. 13, as an argument similar to that appearing after Proposition 1 readily establishes.

mechanics[1] these existence problems were solved by showing that the entropy satisfied a condition of strong subadditivity. One could believe, and even support one's belief by heuristic physical arguments, that the same condition holds for the quantum entropy.

Conjecture: The quantum entropy $S(\rho_\Lambda)$ satisfies the inequality

$$S(\rho_{\Lambda_1\cup\Lambda_2}) + S(\rho_{\Lambda_1\cap\Lambda_2}) \leq S(\rho_{\Lambda_1}) + S(\rho_{\Lambda_2}).$$

A satisfactory discussion of the existence of the mean entropy would ensue, if this conjecture were proved. There would, however, still exist a problem in establishing a barycentric decomposition of the mean entropy in the general case because, although it would clearly be an affine function, it could not be expected to be upper semicontinuous.

We have not discussed in any detail the physical relevance of the mean entropy which we have introduced but postpone this to a forthcoming publication.[16]

ACKNOWLEDGMENTS

The authors are grateful to Professor D. Ruelle for many helpful and stimulating conversations. One of us (D. W. R.) would like to thank Professor R. Jost for a number of informative discussions. Finally, we would like to thank Professor P. Porcelli for his hospitality at the University of Louisiana during the Baton Rouge Symposium on C^* algebras where the present collaboration began.

[16] D. W. Robinson, Commun. Math. Phys. **6**, 151 (1967).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 9, NUMBER 7 JULY 1968

Parameter Differentiation of Exponential Operators and the Baker–Campbell–Hausdorff Formula

M. LUTZKY
U.S. Naval Ordnance Laboratory, White Oak, Silver Spring, Maryland

(Received 8 November 1967)

Parameter differentiation of exponential operators is used to derive a method for obtaining the Baker–Campbell–Hausdorff coefficients in a more explicit form than is available from the standard Hausdorff recursion formula. In passing, a derivation of the Hausdorff recursion formula is given which is simpler than the proof usually presented.

I. INTRODUCTION

The problem of solving the equation $e^z = e^x e^y$ for z, where x and y are noncommuting operators, frequently arises in physics. For instance, Weiss and Maradudin[1] discussed this problem in a study of x-ray scattering in crystals, and Snider[2] encountered it in the course of an investigation involving a linearization of the Boltzmann equation. The classical solution, given by Hausdorff,[3] involves an expansion of z into an infinite series of terms homogeneous in y, the successive terms being found from a recursion formula which utilizes the Hausdorff polarization operator. The recursion formula is known as the Baker–Campbell–Hausdorff formula (BCH), and has been discussed recently by several authors.[4] Unfortunately, the BCH formula is somewhat difficult to use for the computation of higher-order terms. In this paper, we use the method of parameter differentiation of exponential operators[5] to derive in a rather simple way a more explicit form for the BCH coefficients. In passing, we also show how the method may be used to provide a derivation of the BCH recursion formula that is somewhat simpler than the proof usually given.

II. PRELIMINARY DEFINITIONS AND FORMULAS

Our formulas will be considerably simplified by the use of the curly commutator bracket, recursively defined by

$$\{y, x^0\} = y, \tag{1}$$

$$\{y, x^{k+1}\} = [\{y, x^k\}, x]. \tag{2}$$

If $f(x)$ has a power-series expansion

$$f(x) = \sum_{k=0}^{\infty} a_k x^k,$$

[1] G. H. Weiss and A. A. Maradudin, J. Math. Phys. **3**, 771 (1962).
[2] R. F. Snider, J. Math. Phys. **5**, 1586 (1964).
[3] F. Hausdorff, Ber. Verhandl. Saechs. Akad. Wiss. Leipzig, Math.-Naturw. Kl. **58**, 19 (1906).
[4] W. Magnus, Commun. Pure Appl. Math. **7**, 649 (1954); J. Wei, J. Math. Phys. **4**, 1337 (1963); and Ref. 1.
[5] R. M. Wilcox, J. Math. Phys. **8**, 962 (1967).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 11, NUMBER 9 SEPTEMBER 1970

Thermodynamic Limit of Time-Dependent Correlation Functions for One-Dimensional Systems

G. GALLAVOTTI
The Rockefeller University, New York, New York 10021

O. E. LANFORD III*
Department of Mathematics, University of California, Berkeley, California 94720

AND

JOEL L. LEBOWITZ†
Belfer Graduate School of Science, Yeshiva University, New York, New York 10033

(Received 9 March 1970)

We investigate the time evolution of the correlation functions of a nonequilibrium system when the size of the system becomes very large. At the initial time $t = 0$, the system is represented by an equilibrium grand canonical ensemble with a Hamiltonian consisting of a kinetic energy part, a pairwise interaction potential energy between the particles, and an external potential. At time $t = 0$ the external field is turned off and the system is permitted to evolve under its internal Hamiltonian alone. Using the "time-evolution theorem" for a 1-dimensional system with bounded finite-range pair forces, we prove the existence of infinite-volume time-dependent correlation functions for such systems, $\lim \rho_\Lambda(t; q_1, p_1; \cdots; q_n, p_n)$, as $\Lambda \to \infty$, where Λ is the size of the finite system. We also show that these infinite-volume correlation functions satisfy the infinite BBGKY hierarchy in the sense of distributions.

1. INTRODUCTION

The rigorous mathematical study of equilibrium statistical mechanics during the last decade has achieved many successes. This study concerns itself primarily with the properties of equilibrium systems in the thermodynamic limit, i.e., as the size of the system becomes infinite at fixed temperature and activity (or density). In particular, the existence and analyticity of the correlation functions at small values of the activity z has been proven for a wide class of interacting systems.[1] The existence and convexity of the free energy has been proven for an even larger class of systems at all values of the activity.[2]

The comparable mathematical investigation of the infinite-volume limit of nonequilibrium systems is much more difficult and has begun only recently. The results are restricted to 1-dimensional systems of particles interacting by smooth, finite-range pair forces, and they prove the existence for all times of a "regular" solution of Newton's equations of motion for a "regular" initial configuration. A regular configuration is, roughly speaking, one in which the number of particles in a unit interval and the magnitude of the momentum of any particle in that interval have a bound of the form $\delta \log R$, where R denotes the distance of the interval from the origin. It is shown in Ref. 3 that, at equilibrium, if either the activity is small or the interparticle potential is positive, the set of nonregular configuration has probability zero.

A question left open by these results is whether a state which at time $t = 0$ is described by a set of correlation functions can still be described by a set of correlation functions when $t \neq 0$.

In this paper we investigate this question and prove that, for certain classes of initial states, the time-evolving state is described by correlation functions and that these correlation functions satisfy the BBGKY hierarchy in the sense of distributions [see (2.9)].

The initial states we consider can be described as follows: Suppose that the system is in equilibrium at temperature β^{-1} and activity z under the influence of a pair potential and an external potential h which is localized in a finite region I_h. At time $t = 0$, we switch off the external field and the system begins to evolve. We prove that, if the activity is sufficiently small (i.e., if we are deep inside the gaseous phase), the system can always be described by a set of correlation functions which vary in time according to the BBGKY hierarchy. We are unable to prove even that the time-averaged correlation functions evolve toward the correlation functions which correspond to the equilibrium state at temperature β^{-1} and activity z (in absence of external field), as would be expected. We are, however, able to prove that the time-averaged correlation functions converge to a limit satisfying the stationary BBGKY hierarchy.

We note that initial states of the kind just described suffice, in principle, for the study of transport properties at low activity.

2. DESCRIPTION OF INITIAL STATE AND SUMMARY OF RESULTS

We consider a 1-dimensional system of identical particles of unit mass, interacting through a stable pair-potential $\Phi(q)$ which has finite range and is

twice continuously differentiable. We denote

$$C = -\inf_{n,q_1,\cdots,q_n}\left(n^{-1}\sum_{i<j}\Phi(q_i - q_j)\right). \tag{2.1}$$

The condition of stability says precisely that $C < \infty$; it guarantees that the thermodynamic functions are well defined.[1]

The initial states we consider will be equilibrium states, at inverse temperature β and chemical potential μ, for an interaction coming from the pair potential Φ and an external potential $h(q)$ which is continuous and which vanish outside some bounded interval I_h. We will assume that the activity z [$= e^{\mu\beta}(2\pi/\beta)^{\frac{1}{2}}$, in units where Planck's constant is unity] is small enough so that the Mayer series converges, i.e.,

$$z < B(\beta)^{-1}\exp(-\beta C - 1), \tag{2.2}$$

where

$$B(\beta) = \int dq\,|\exp[-\beta\Phi(q)] - 1|. \tag{2.3}$$

Under these conditions, the thermodynamic limit for the correlation functions is known to exist for $h = 0$.[1,4]

Now let Λ denote a finite interval centered at the origin and containing I_h, and let

$$H_\Lambda(q_1, p_1; \cdots; q_n, p_n) = \tfrac{1}{2}\sum_{i=1}^{n} p_i^2 + \sum_{i<j}\Phi_\Lambda(q_i - q_j) \tag{2.4}$$

denote the Hamiltonian for a system of n particles, in the box Λ with periodic boundary conditions, interacting by the periodized 2-body interaction Φ_Λ. (By Φ_Λ we mean the potential obtained by periodizing Φ with respect to Λ. In order that Φ_Λ be unambiguous, we will assume that the length of Λ is at least twice the range of Φ.[3]) We let $\tilde{T}_\Lambda^t$ denote the time-evolution mapping in the periodic box Λ determined by the Hamiltonian H_Λ. Finally, we let

$$h(q_1, \cdots, q_n) = \sum_{i=1}^{n} h(q_i). \tag{2.5}$$

Here it is convenient to introduce a piece of notation. Instead of writing $(q_1, p_1; \cdots; q_n, p_n)$ for a point of $(R^2)^n$, we will write $(x)_n$. If

$$(x)_n = (q_1, p_1; \cdots; q_n, p_n)$$

and if

$$(y)_m = (q_1', p_1'; \cdots; q_m', p_m'),$$

we use $(x)_n \cup (y)_m$ to denote

$$(q_1, p_1; \cdots; q_n, p_n; q_1', p_1'; \cdots; q_m', p_m') \in (R^2)^{n+m}.$$

We also write $d(x)_n$ for $dq_1\,dp_1 \cdots dq_n\,dp_n$.

We now want to consider the following situation: We start at time $t = 0$ with the equilibrium state in the box Λ with inverse temperature β and chemical potential μ (and the external potential h as well as the interparticle potential Φ). We let the state evolve in time with $\tilde{T}_\Lambda^t$ (without the external potential); we write down the correlation functions for the time-evolved state; and we study their behavior as $\Lambda \to \infty$. Physically, this situation corresponds to having a system in equilibrium in the presence of an external potential h, turning off the external potential at $t = 0$, and watching the evolution of the correlations functions as $\Lambda \to \infty$.

Thus, we want to examine the correlation functions[5]:

$$\rho_\Lambda(t; (x)_n) = \frac{1}{\Xi_\Lambda}\sum_{m=0}^{\infty}\frac{e^{\beta\mu(m+n)}}{m!}\int_{(\Lambda\times R)^m} d(x')_m \times \exp\{-\beta(H_\Lambda + h)[\tilde{T}_\Lambda^{-t}((x)_n \cup (x')_m)]\}, \tag{2.6}$$

where

$$\Xi_\Lambda = \sum_{m=0}^{\infty}\frac{e^{\beta\mu m}}{m!}\int_{(\Lambda\times R)^m} d(x')_m \exp\{-\beta(H_\Lambda + h)[(x')_m]\}. \tag{2.7}$$

Our main results can be stated as follows:

(i) If h is nonnegative, the limit as $\Lambda \to \infty$ of $\rho_\Lambda(t; (x)_n)$ exists for all t and $(x)_n$.

(ii) If h is not assumed to be nonnegative, the limit as $\Lambda \to \infty$ of $\rho_\Lambda(t; (x)_n)$ exists in the sense of distributions in $(x)_n$ for each t; the limiting distribution is actually a locally square-integrable function of $(x)_n$. In either case, we will denote the limit by $\rho(t; (x)_n)$.

(iii) The infinite-volume correlation functions $\rho(t; (x)_n)$ satisfy the BBGKY hierarchy in the following form: For any C^1 function $f(x)_n$ of compact support on $(R^2)^n$, we let

$$\rho_t(f) = \int_{(R^2)^n}\rho(t; (x)_n)f(x)_n\,d(x)_n. \tag{2.8}$$

Then $\rho_t(f)$ is a differentiable function of t and

$$\frac{d}{dt}\rho_t(f) = \rho_t(\{H, f\}) - \rho_t(f_1), \tag{2.9}$$

where

$$\{H, f\}(q_1, p_1; \cdots; q_n, p_n) = \sum_{i=1}^{n}\left(\frac{\partial H}{\partial p_i}\frac{\partial f}{\partial q_i} - \frac{\partial H}{\partial q_i}\frac{\partial f}{\partial p_i}\right), \tag{2.10}$$

$$f_1(q_1, p_1; \cdots; q_{n+1}, p_{n+1}) = \sum_{i=1}^{n}\frac{\partial\Phi(q_i - q_{n+1})}{\partial q_i}\frac{\partial f}{\partial p_i}, \tag{2.11}$$

$$H(q_1, p_1; \cdots; q_n, p_n) = \tfrac{1}{2}\sum_{i=1}^{n} p_i^2 + \sum_{i<j}\Phi(q_i - q_j). \tag{2.12}$$

Equation (2.9) may be obtained from the standard formal BBGKY hierarchy[5] by multiplying by the test function $f(x)_n$, integrating over (x_n), and putting the q and p derivatives on the test function by integration by parts.

(iv) If $\rho_0(x)_n$ denotes the equilibrium correlations with no external potential, then for all $n, m > 0$

$$\lim_{a\to\infty} \rho(t; q_1, p_1; \cdots; q_n, p_n; q_{n+1} + a, p_{n+1}; \cdots; q_{n+m} + a, p_{n+m}) = \rho(t; q_1, p_1; \cdots; q_n, p_n) \times \rho_0(q_{n+1}, p_{n+1}; \cdots; q_{n+m}, p_{n+m}). \quad (2.13)$$

3. INFINITE SYSTEMS

Although we did not need the theory of actually infinite systems to formulate our results, the proofs depend on this theory. We will summarize in this section the main results that we need. For more details, see Refs. 1 or 3.

A locally finite configuration of particles is defined by giving a sequence (possibly finite) of positions and momenta (q_i, p_i) such that each bounded interval in R contains only finitely many q_i. However, since the particles are supposed to be identical, we identify configurations which differ only by the labeling of the particles. Thus, a configuration may be thought of as subset of phase space R^2 with multiplicity, where the subset is just the set of occupied points and the multiplicity of each point is the number of particles at the point. We will let $\mathfrak{X}$ denote the set of all such configurations. If X and Y are configurations belonging to $\mathfrak{X}$, we let $X \cup Y$ be the configuration obtained by adding the multiplicities for X and Y. Also, if $\Lambda \subset R$, and if $X \in \mathfrak{X}$, we let $X \cap \Lambda$ denote the configuration obtained from $\mathfrak{X}$ by omitting all particles whose positions are not in Λ. The set of configurations with all particles in Λ will be denoted by $\mathfrak{X}(\Lambda)$.

We will say that a function f on $\mathfrak{X}$ is *measurable in* Λ if

$$f(X) = f(X \cap \Lambda)$$

for all $X \in \mathfrak{X}$. There is a simple way to construct such functions. Let ψ be a function on R^2 such that $\psi(q, p) = 0$ for $q \notin \Lambda$. Then define

$$(\sum \psi)(X) = \sum_i \psi(q_i, p_i),$$

where X is determined by (q_i, p_i). If Λ is bounded, there are only finitely many nonzero terms in this sum. Clearly, $\sum \psi$ is measurable in Λ. We give $\mathfrak{X}$ the weakest topology such that $\sum \psi$ is continuous for all continuous $\psi(q, p)$ whose support in q is bounded. It can be convincingly argued that states of classical statistical mechanics should be identified with Borel probability measures on $\mathfrak{X}$ (see Ref. 6).

If Λ is a bounded open subset of R, the mapping $X \to X \cap \Lambda$ is Borel from $\mathfrak{X}$ to $\mathfrak{X}(\Lambda)$. A Borel measure γ on $\mathfrak{X}$ defines therefore a measure γ_Λ on $\mathfrak{X}(\Lambda)$, i.e., a sequence $\gamma_{n,\Lambda}$ of symmetric Borel measures on $(\Lambda \times R)^n$, $n = 0, 1, 2, \cdots$. If each $\gamma_{n,\Lambda}$ is absolutely continuous with respect to Lebesgue measure, we define *density distributions* $f_\Lambda(x)_n$ by

$$d\gamma_{n,\Lambda} = f_\Lambda(x)_n \, d(x)_n / n!, \quad (3.1)$$

where $f_\Lambda(x)_n/n!$ is the probability density of finding precisely n particles with position $q_1, \cdots, q_n$ in Λ and momenta $p_1, \cdots, p_n$.

For any symmetric continuous function ψ on $(R^2)^n$, with compact support, we define a function $\sum \psi$ on $\mathfrak{X}$ by

$$\sum \psi(X) = \sum_{i_1 < i_2 < \cdots < i_n} \psi(q_{i_1}, p_{i_1}; \cdots; q_{i_n}, p_{i_n}), \quad (3.2)$$

where the configuration X is defined by (q_i, p_i). If γ is a measure on $\mathfrak{X}$ such that $\sum \psi$ is γ-integrable for all such $\psi(x)_n$, then

$$\psi \to \int d\gamma(\sum \psi)$$

is a positive linear functional on the space of continuous symmetric functions on $(R^2)^n$ of compact support, i.e., a symmetric measure on $(R^2)^n$. When this measure exists and is absolutely continuous with respect to Lebesgue measure, so that it can be written $\rho(x)_n \, d(x)_n/n!$, we say that $\rho(x)_n$ is the nth correlation function of γ. To recapitulate: The correlation function $\rho(x)_n$ is defined by the relation

$$\int \psi(x)_n \rho(x)_n \frac{d(x)_n}{n!} = \int \sum \psi(X) \, d\gamma(X). \quad (3.3)$$

It is not hard to see that, if γ has correlation functions of all orders, then density distributions exist and, for $q_1, \cdots, q_n \in \Lambda$,

$$\rho(q_1, p_1; \cdots; q_n, p_n) = \sum_{m=0}^{\infty} \int_{(\Lambda\times R)^m} \frac{dq_1' \, dp_1' \cdots dq_m' \, dp_m'}{m!} \times f_{\Lambda, n+m}(q_1, p_1; \cdots; q_m', p_m'). \quad (3.4)$$

Conversely, if there exists a constant η such that, for all Λ and n,

$$\int_{(\Lambda\times R)^n} d(x)_n \rho(x)_n \leq [\eta V(\Lambda)]^n, \quad (3.5)$$

$V(\Lambda)$ being the length of the interval Λ, the density distributions can be reexpressed in terms of the

correlation functions by

$$f_\Lambda(x)_n = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} \int_{(\Lambda\times R)^m} d(y)_m \rho((x)_n \cup (y)_m). \quad (3.6)$$

It has been shown[1] that, for activities satisfying (2.2), the infinite-volume limits of the correlation functions for finite-volume equilibrium ensembles exist; they have the form

$$\rho(q_1, p_1; \cdots; q_n, p_n) = \rho(q_1, \cdots, q_n) \exp\left[-\tfrac{1}{2}\beta(p_1^2 + \cdots + p_n^2)\right], \quad (3.7)$$

where, for some real number ξ,

$$\rho(q_1, \cdots, q_n) \leq \xi^n \quad (3.8)$$

for all n, $q_1, \cdots, q_n$, and hence they satisfy an estimate of the form (3.5). Thus, a measure on $\mathfrak{X}$ may be reconstructed from this set of correlation functions; we denote this measure γ_0 and call it the infinite-volume equilibrium state. It follows easily from the estimates in Ref. 7 that, if ψ is a bounded Borel function on $\mathfrak{X}$ measurable in some bounded set, then

$$\int d\gamma_0 \psi = \lim_{\Lambda\to\infty} \int_{\chi(\Lambda)} d\gamma_{(\Lambda)} \psi, \quad (3.9)$$

where $\gamma_{(\Lambda)}$ is the finite-volume grand canonical ensemble density [regarded as a probability measure on $\mathfrak{X}(\Lambda)$].

4. THE EVOLUTION THEOREM

In this section we summarize the results of Ref. 3 in a form convenient for our purposes.

(i) The existence of a solution of the equations of motion has not been established for arbitrary initial data in $\mathfrak{X}$, but only for a special set $\hat{\mathfrak{X}}$ of configurations. This set $\hat{\mathfrak{X}}$ may be written as the union of a family of subsets $\hat{\mathfrak{X}}_\delta$, where δ runs through the positive real numbers. We have

$$\hat{\mathfrak{X}}_\delta \supset \hat{\mathfrak{X}}_{\delta'}, \quad \delta \geq \delta'.$$

Each $\hat{\mathfrak{X}}_\delta$ is compact in $\mathfrak{X}$. The sets $\hat{\mathfrak{X}}_\delta$ are large in the sense that

$$\gamma_0(\hat{\mathfrak{X}}) = 1, \quad \text{i.e.,} \quad \lim_{\delta\to\infty} \gamma_0(\mathfrak{X} \setminus \hat{\mathfrak{X}}_\delta) = 0.$$

In fact, a slightly stronger statement is true

$$\lim_{\delta\to\infty} \gamma_{(\Lambda)}(\mathfrak{X}(\Lambda) \cap \hat{\mathfrak{X}}_\delta) = 1$$

uniformly in Λ for large Λ.

(ii) The crux of the existence of time evolution is contained in the following statement: There is a 1-parameter group of mappings T^t of $\hat{\mathfrak{X}}$ onto itself such that, for any continuous function ψ on $\mathfrak{X}$ which is measurable in some bounded interval,

$$\lim_{\Lambda\to\infty} \psi(\tilde{T}_\Lambda^t (X \cap \Lambda)) = \psi(T^t X)$$

for $X \in \hat{\mathfrak{X}}$. The convergence is uniform for X in any fixed $\hat{\mathfrak{X}}_\delta$ and t in any bounded interval.[8] For any fixed δ, $(t, X) \to T^t X$ is continuous from $R \times \hat{\mathfrak{X}}_\delta$ to $\mathfrak{X}$. If an appropriate labeling of the particles in

$$T^t X = (q_1(t), p_1(t), q_2(t), \cdots)$$

is chosen, then the $(q_i(t), p_i(t))$ solve the differential equations

$$\frac{dq_i(t)}{dt} = p_i(t), \quad \frac{dp_i(t)}{dt} = \sum_{j\neq i} \Phi'(q_i(t) - q_j(t)).$$

(iii) The mapping $(X, X') \to X \cup X'$, sending $\mathfrak{X} \times \mathfrak{X}$ to $\mathfrak{X}$, is continuous. If $X \in \hat{\mathfrak{X}}_\delta$ and $X' \in \hat{\mathfrak{X}}_{\delta'}$, then $X \cup X' \in \hat{\mathfrak{X}}_{\delta+\delta'}$.

(iv) Every point in $(R^2)^n$ determines a point in $\mathfrak{X}$ representing a configuration of exactly n particles. We will usually fail to distinguish between $(x)_n$ as a point of $(R^2)^n$ and the corresponding point in $\mathfrak{X}$. The mapping from $(R^2)^n$ to $\mathfrak{X}$ so defined is continuous and the image of each bounded set is contained in some $\hat{\mathfrak{X}}_\delta$.

(v) The equilibrium measure γ_0 on $\hat{\mathfrak{X}}$ (more precisely, the measure obtained by restricting γ_0 to $\hat{\mathfrak{X}}$) is invariant under T^t for all t; i.e., if $E \subset \hat{\mathfrak{X}}$ is a Borel set, then

$$\gamma_0(E) = \gamma_0(T^{-t}E).$$

5. INFINITE-VOLUME LIMITS OF TIME-DEPENDENT QUANTITIES

Proposition 1: Let ϕ and ψ be functions on $\mathfrak{X}$, both measurable in some bounded interval I. We assume ϕ to be continuous and ψ to be a Borel function. We also assume that, for some real number α,

$$|\phi(X)| \leq \exp[\alpha N_I(X)], \quad (5.1)$$

$$|\psi(X)| \leq \exp[\alpha N_I(X)] \quad (5.2)$$

[where $N_I(X)$ is the number of particles in the interval I for the configuration X]. Then

(i) $\psi(Y)\phi(T^t Y)$ is γ_0 integrable and

$$\int_\chi d\gamma_0(Y)\psi(Y)\phi(T^t Y) = \lim_{\Lambda\to\infty} \int_{\chi(\Lambda)} d\gamma_{(\Lambda)}(Y)\psi(Y)\phi(\tilde{T}_\Lambda^t Y). \quad (5.3)$$

(ii) If ϕ is further assumed to be bounded, then, for any $(x)_n \in (R^2)^n$, $\psi(Y)\phi(T^t(Y \cup (x)_n))$ is γ_0 integrable, and

$$\int_{\mathfrak{X}} d\gamma_0(Y)\psi(Y)\phi(T^t(Y \cup (x)_n)) = \lim_{\Lambda\to\infty} \int_{\mathfrak{X}(\Lambda)} d\gamma_{(\Lambda)}(Y)\psi(Y)\phi(\tilde{T}^t_\Lambda(Y \cup (x)_n)). \quad (5.4)$$

The integral varies continuously with $(x)_n$.

Proof: There exists a real number ξ such that the nth correlation function of $\gamma_{(\Lambda)}$, $\rho_\Lambda(x)_n$ satisfies

$$\rho_\Lambda(x)_n \le \xi^n \exp\left(-\tfrac{1}{2}\beta \sum_{i=1}^{n} p_i^2\right), \quad (x)_n = (q_1, p_1; \cdots; q_n, p_n) \quad (5.5)$$

for all n, $(x)_n$, and all sufficiently large Λ. This inequality persists for the infinite-volume correlation functions. By (3.6), the probability of finding precisely N particles in I, with respect to any $\gamma_{(\Lambda)}$ or with respect to γ_0, is majorized by

$$(n!)^{-1}[\xi(2\pi/\beta)^{\frac{1}{2}}V(I)]^n \exp\,[\xi(2\pi/\beta)^{\frac{1}{2}}V(I)]. \quad (5.6)$$

It follows that $\exp\,[\alpha N_I(Y)]$ is square-integrable with respect to each $\gamma_{(\Lambda)}$ and with respect to γ_0 and that its square-integral has an upper bound which is independent of Λ. By (5.1) and (5.2), ϕ and ψ are both γ_0 square-integrable and, since γ_0 is invariant under T^t, $\psi \circ T^t$ is also γ_0 square-integrable. By the Schwarz inequality, then, $\psi(Y)\phi(T^tY)$ is γ_0-integrable. Similarly, if ϕ is bounded, $\psi(Y)\phi(T^t(Y \cup (x)_n))$ is γ_0 integrable. By 4(ii)–(iv), $T^t(Y \cup (x)_n)$ varies continuously with $(x)_n$; hence,

$$\int d\gamma_0(Y)\psi(Y)\phi(T^t(Y \cup (x)_n))$$

is a continuous function of $(x)_n$, by the Lebesgue dominated-convergence theorem.

Because of the boundedness of the square-integrals, replacing ϕ by $-\lambda$ if $\phi \le \lambda$, ϕ if $-\lambda \le \phi \le \lambda$, and λ if $\phi \ge \lambda$ with λ large, makes a change in

$$\int d\gamma_{(\Lambda)}\psi(\phi \circ T^t)$$

which is small uniformly in Λ. Hence, in proving (5.3), we can assume that ϕ is bounded. In this case, (5.3) is a special case of (5.4). In a similar way, we see that, in proving (5.4), ψ may also be assumed to be bounded.

To prove (5.4), assuming ψ bounded, we choose $\epsilon > 0$ and then choose δ large enough so that

$$\gamma_{(\Lambda)}(\mathfrak{X}(\Lambda) \setminus \hat{\mathfrak{X}}_\delta) < \epsilon, \quad \text{for all sufficiently large } \delta\Lambda, \quad (5.7)$$

and

$$\gamma_0(\mathfrak{X} \setminus \mathfrak{X}_\delta) < \epsilon;$$

this is possible by (i). Now, by 4(ii)–(iv), the mapping

$$Y \to \Phi(T^t(Y \cup (x)_n))$$

is continuous on $\hat{\mathfrak{X}}_\delta$, and $\hat{\mathfrak{X}}_\delta$ is compact in $\mathfrak{X}$. The collection of all functions on $\hat{\mathfrak{X}}_\delta$, which are restrictions of continuous functions on $\mathfrak{X}$ measurable in bounded intervals (the interval may vary with the function), is an algebra of continuous functions on $\hat{\mathfrak{X}}_\delta$ containing the constants and separating points. Hence, by the Stone–Weierstrass theorem,[9] there is a continuous function Φ_1 on $\mathfrak{X}$, measurable in some bounded interval, such that

$$|\phi_1(Y) - \phi(T^t(Y \cup (x)_n))| < \epsilon \quad (5.8)$$

for all $Y \in \hat{\mathfrak{X}}_\delta$. We can also assume

$$\|\phi_1\|_\infty \le \|\phi\|_\infty. \quad (5.9)$$

Because

$$\lim_{\Lambda\to\infty} \phi(\tilde{T}^t_\Lambda([Y \cup (x)_n] \cap \Lambda)) = \phi(T^t(Y \cup (x)_n))$$

uniformly for $Y \in \hat{\mathfrak{X}}_\delta$ (by 4(ii)), we have

$$|\phi_1(Y) - \phi(\tilde{T}^t_\Lambda(Y \cup (x)_n))| < \epsilon \quad (5.10)$$

for all sufficiently large Λ and all $Y \in \hat{\mathfrak{X}}_\delta \cap \mathfrak{X}(\Lambda)$.

Now

$$\begin{aligned}
&\left|\int_{\mathfrak{X}} d\gamma_0(Y)\psi(Y)\phi(T^t(Y \cup (x)_n)) - \int_{\mathfrak{X}(\Lambda)} d\gamma_{(\Lambda)}(Y)\psi(Y)\phi(\tilde{T}^t_\Lambda(Y \cup (x)_n))\right| \\
&\le \left|\int_{\mathfrak{X}} d\gamma_0(y)\psi(Y)[\phi(T^t(Y \cup (x)_n)) - \phi_1(Y)]\right| \\
&\quad + \left|\int_{\mathfrak{X}} d\gamma_0(Y)\psi(Y)\phi_1(Y) - \int_{\mathfrak{X}(\Lambda)} d\gamma_{(\Lambda)}(Y)\psi(Y)\phi_1(Y)\right| \\
&\quad + \left|\int_{\mathfrak{X}(\Lambda)} d\gamma_{(\Lambda)}(Y)\psi(Y)[\phi_1(Y) - \phi(\tilde{T}^t_\Lambda(Y \cup (x)_n))]\right|.
\end{aligned} \quad (5.11)$$

The first term on the right of (5.11) is majorized by

$$\begin{aligned}
&\left|\int_{\hat{\mathfrak{X}}_\delta} d\gamma_0(Y)\psi(Y)[\phi(T^t(Y \cup (x)_n)) - \phi_1(Y)]\right| \\
&\quad + \left|\int_{\mathfrak{X} \setminus \hat{\mathfrak{X}}_\delta} d\gamma_0(Y)\psi(Y)[\phi(T^t(Y \cup (x)_n)) - \phi_1(Y)]\right| \\
&\le \|\psi\|_\infty \epsilon + \epsilon \|\psi\|_\infty (2\|\phi\|_\infty) = \epsilon \|\psi\|_\infty (1 + 2\|\phi\|_\infty).
\end{aligned}$$

[We have used (5.7), (5.8), and (5.9).] Similar arguments show that the same quantity majorizes the the third term on the right of (5.11) provided that Λ

is large enough so that (5.10) holds. Finally, the middle term on the right of (5.11) approaches zero as $\Lambda \to \infty$ by (3.9). Hence, for Λ sufficiently large,

$$\left| \int_{\mathfrak{X}} d\gamma_0(Y)\psi(Y)\phi(T^t(Y \cup (x)_n)) - \int_{\mathfrak{X}_{(\Lambda)}} d\gamma_{(\Lambda)}(Y)\psi(Y)\phi(\tilde{T}_\Lambda^t(Y \cup (x)_n)) \right| \leq 3\epsilon \|\psi\|_\infty (1 + 2\|\phi\|_\infty).$$

Since $\epsilon > 0$ is arbitrary, (5.4) follows.

Corollary 1: Assume the external potential h is nonnegative, and let the time-dependent finite-volume correlation functions be defined as in (2.6). Then

$$\lim_{\Lambda \to \infty} \rho_\Lambda(t; (x)_n)$$

exists for all t and $(x)_n$ and the convergence is locally uniform in $(x)_n$. Furthermore,

$$\lim_{\Lambda \to \infty} \rho_\Lambda(t; (x)_n) = e^{\beta\mu n} \int_{\mathfrak{X}} d\gamma_0(Y) \times \exp[-\beta \sum h(T^{-t}(Y \cup (x)_n))] \times \exp\{-\beta[H(x)_n + W((x)_n, Y)]\} \times \left(\int_{\mathfrak{X}} d\gamma_0(Y) e^{-\beta \Sigma h(Y)} \right)^{-1}, \quad (5.12)$$

where

$$W((x)_n, Y) = \sum_{i,j} \phi(q_i - q_j');$$

$Y = (q_i', p_i')$, and the limit is a continuous function of $(x)_n$.

Proof: From the definition of $\rho_\Lambda(t; (x)_n)$ and the fact that

$$H_\Lambda \circ \tilde{T}_\Lambda^t = H_\Lambda$$

(conservation of energy), it follows that

$$\rho_\Lambda(t; (x)_n) = e^{\beta\mu n} \int_{\mathfrak{X}_{(\Lambda)}} d\gamma_{(\Lambda)}(Y) \times \exp[-\beta \sum h(\tilde{T}_\Lambda^t(Y \cup (x)_n))] \times \exp\{-\beta[H(x)_n + W((x)_n, Y)]\} \times \left(\int_{\mathfrak{X}_{(\Lambda)}} d\gamma_{(\Lambda)}(Y) e^{-\beta \Sigma h(Y)} \right)^{-1}; \quad (5.13)$$

the corollary then follows by straightforward application of Proposition 1.

Corollary 2: Let the time-dependent finite-volume correlation functions be defined as in (2.6), and let the external potential h be any continuous function of compact support. Let $f(x)_n$ be any continuous function of compact support on $(R^2)^n$. Then

$$\lim_{\Lambda \to \infty} \int \frac{d(x)_n}{n!} \rho_\Lambda(t; (x)_n) f(x)_n = \int d\gamma_0(Y) \exp[-\beta(\sum h)(T^{-t}Y)](\sum f)(Y) \times \left(\int d\gamma_0(Y) e^{-\beta \Sigma h(Y)} \right)^{-1}; \quad (5.14)$$

moreover, there exist locally square-integrable functions $\rho(t; (x)_n)$ such that

$$\lim_{\Lambda \to \infty} \int \frac{d(x)_n}{n!} \rho_\Lambda(t; (x)_n) f(x)_n = \int \frac{d(x)_n}{n!} \rho(t; (x)_n) f(x)_n, \quad (5.15)$$

for all continuous f of compact support.

Proof: Again, by the definition of the finite-volume correlation functions and the conservation of energy, we have

$$\int \frac{d(x)_n}{n!} \rho_\Lambda(t; (x)_n) f(x)_n = \int d\gamma_{(\Lambda)}(Y) \exp[-\beta \sum h(\tilde{T}_\Lambda^{-t} Y)] \sum f(Y) \times \left(\int d\gamma_{(\Lambda)}(Y) e^{-\beta \Sigma h(Y)} \right)^{-1}; \quad (5.16)$$

thus (5.14) follows from Proposition 1. The existence of the infinite-volume correlation functions $\rho(t; (x)_n)$ as locally square-integrable functions (and not merely as measures) follows from the fact, easily verified, that, if Ω is any bounded open set in $(R^2)^n$, the mapping $f \to \sum f$ from the space of continuous functions with support in Ω to the space of continuous functions on $\mathfrak{X}$ extends to a continuous mapping from $L^2(\Omega, d(x)_n)$ to $L^2(\mathfrak{X}, d\gamma_0)$. Hence, since $e^{-\Sigma h(\cdot)} \in L^2(\mathfrak{X}, d\gamma_0)$, the mapping

$$f \to \lim_{\Lambda \to \infty} \int \frac{d(x)_n}{n!} f(x)_n \rho_\Lambda(t; (x)_n)$$

extends to a continuous linear functional on $L^2(\Omega)$ and is therefore given by a function square-integrable on Ω.

6. THE BBGKY HIERARCHY

Theorem 1: Let ψ be a nonnegative function on $\mathfrak{X}$ with

$$\int \psi(Y)\, d\gamma_0(Y) = 1 \quad \text{and} \quad \int \psi(Y)^2\, d\gamma_0(Y) < \infty.$$

Let γ denote the probability measure $\psi(Y)\,d\gamma_0(Y)$ on $\mathfrak{X}$, and let γ^t be the time-evolved measure defined by

$$\int \phi(Y)\,d\gamma^t(Y) = \int \phi \circ T^t(Y)\,d\gamma(Y). \tag{6.1}$$

Then γ^t has correlation functions of all orders, and these correlation functions are locally square-integrable. Moreover, for any function f which is infinitely differentiable and of compact support on $(R^2)^n$, $\int d\gamma^t \sum f$ is a differentiable function of t, and

$$\frac{d}{dt}\int \sum f\,d\gamma^t = \int \sum (\{H,f\})\,d\gamma^t - \int \sum f_1\,d\gamma^t, \tag{6.2}$$

where the notation is defined in (2.10)–(2.12).

Proof: By a simple calculation, using the invariance of γ_0 under T^t, we have

$$d\gamma^t = (\psi \circ T^t)\,d\gamma_0,$$
$$\int d\gamma_0\,|\psi \circ T^t|^2 = \int d\gamma_0\,|\psi|^2 < \infty.$$

Thus γ^t is obtained from γ_0 by multiplication by a square-integrable function. The arguments used in the proof of Corollary 2 show that this implies that γ^t has locally square-integrable correlation functions of all orders. On the other hand,

$$\int \sum f\,d\gamma^t = \int \sum [f \circ T^t(Y)]\psi(Y)\,d\gamma_0(Y). \tag{6.3}$$

It follows readily from 4(ii) that, for any infinitely differentiable f and $Y \in \hat{\mathfrak{X}}$,

$$\frac{d}{dt}\sum f(T^tY) = \sum (\{H,f\})(T^tY) - \sum f_1(T^tY). \tag{6.4}$$

The right-hand side of this expression may be verified to be γ_0-square-integrable; hence, its absolute value is γ integrable, and the integral is a bounded function of t. The complement of $\hat{\mathfrak{X}}$ has γ-measure zero. Hence, by standard theorems about differentiation under the integral sign,

$$\frac{d}{dt}\int \sum f\,d\gamma^t = \int [\sum (\{H,f\}) - \sum f_1]\,d\gamma^t, \tag{6.5}$$

which is just Eq. (6.2).

Corollary 3: Equation (2.9) holds.

Proof: The $\rho(t;(x)_n)$ are the correlation functions of the measure obtained by evolving in time the measure

$$e^{-\beta\sum h(Y)}\,d\gamma_0(Y)\Big/\int d\gamma_0(Y')e^{-\beta\sum h(Y')}$$

[by Sec. 4(iii)], and $e^{-\beta\sum h(Y)}$ is γ_0-square-integrable.

7. CLUSTER PROPERTIES

Let τ_a denote the operation of translation by a, acting on $\mathfrak{X}$, i.e., $\tau_a(q_i,p_i) = (q_i + a, p_i)$. The equilibrium state γ_0 is invariant under τ_a and has strong cluster properties under the action of τ_a. For a detailed discussion of these cluster properties, see Ref. 1; we will need the following fact, easily deduced from the results in this reference: If ψ and ϕ are functions on $\mathfrak{X}$ which are γ_0-square-integrable, then

$$\lim_{|a|\to\infty}\int d\gamma_0(Y)\psi(Y)\phi(\tau_aY) = \int d\gamma_0(Y)\psi(Y)\int d\gamma_0(Y)\psi(Y). \tag{7.1}$$

Using this result, we will prove the following:

Theorem 2: Let the $\rho(t;(x)_n)$ be defined as in Corollary 2, and let $\rho_0(x)_n$ denote the nth correlation function of the equilibrium measure γ_0. Then, for continuous symmetric functions $f(x)_n$, $g(y)_m$ of compact support,

$$\lim_{|a|\to\infty}\int d(x)_n\,d(y)_m f(x)_n g(y)_m \rho\,(t;(x)_n \cup \tau_a(y)_m) = \int d(x)_n f(x)_n \rho(t;(x)_n)\int d(y)_m g(y)_m \rho_0(y)_m. \tag{7.2}$$

Proof: Let $g_a(y)_m = g(\tau_a(y)_m)$, and let $\mathfrak{X}_a$ be the function on $\mathfrak{X}$ defined by

$$\mathfrak{X}_a((q_i,p_i)) = \frac{1}{(n+m)!}\sum_{i_1,\cdots,i_{n+m}}{}' f(q_{i_1},p_{i_1};\cdots;q_{i_n},p_{i_n}) \times g_a(q_{i_{n+1}},\cdots,p_{i_{n+m}}), \tag{7.3}$$

where the sum $\sum'$ is to be taken over all $(n+m)$-tuples of distinct indices. Since f and g both have compact supports, for sufficiently large a,

$$f(q_{i_1},\cdots,p_{i_n})g_a(q_{i_{n+1}},\cdots,p_{i_{n+m}}) = 0$$

if any i_k, $1 \le k \le n$, is equal to any i_e, $n+1 \le e \le n+m$. Hence, for such a,

$$\chi_a((q_i,p_i)) = \frac{1}{(n+m)!}\sum_{i_1,\cdots,i_n}{}' f(q_i,\cdots,p_{i_n}) \times \sum_{i_{n+1}\cdots,i_{n+m}}{}' g_a(q_{i_{n+1}},\cdots,p_{i_{n+m}}),$$

$$\mathfrak{X}_a(Y) = \frac{n!\,m!}{(n+m)!}(\sum f)(Y)(\sum g)(\tau_aY). \tag{7.4}$$

If we let

$$\psi(Y) = e^{-\beta\sum h(Y)}\left(\int d\gamma_0(Y')e^{-\beta\sum h(Y')}\right)^{-1}, \tag{7.5}$$

then Corollary 2 and the definition of correlation functions gives

$$\frac{1}{(n+m)!}\int d(x)_n\,d(y)_m f(x)_n g(y)_m \rho\,(t;(x)_n \cup \tau_a(y)_m) = \int d\gamma_0(Y)\psi(T^{-t}Y)\mathfrak{X}_a(Y) = \frac{n!\,m!}{(n+m)!}\int d\gamma_0(Y)\psi(T^{-t}Y)(\sum f)(Y)(\sum g)(\tau_aY) \tag{7.6}$$

for large a. Since any power of ψ, $\sum f$, or $\sum g$ is γ_0-integrable, it follows from (7.1) that

$$\lim_{|a|\to\infty} \int d(x)_n\, d(y)_m f(x)_n g(y)_m \rho\,(t;(x)_n \cup \tau_a(y)_m)$$
$$= n!\, m! \left(\int d\gamma_0(Y)\psi(T^{-t}Y)\sum f(Y)\right) \times \left(\int d\gamma_0(Y)\sum g(Y)\right)$$
$$= \left(\int d(x)_n f(x)_n \rho(t;(x)_n)\right) \times \left(\int d(y)_m g(y)_m \rho_0(y)_m\right),$$

which is just (7.2).

8. REMARKS

We have seen that the infinite-volume correlation functions $\rho(t;(x)_n)$ are the correlation functions for a measure obtained by multiplying the equilibrium measure γ_0 by $\psi \circ T^{-t}$, where ψ is defined in (7.5). We may define time-averaged correlation functions

$$\bar{\rho}(T,(x)_n) = T^{-1}\int_0^T dt\rho(t;(x)_n); \tag{8.1}$$

these are the correlation functions of the measure obtained by multiplying γ_0 by

$$\bar{\psi}_T = T^{-1}\int_0^T dt\psi \circ T^{-t}. \tag{8.2}$$

By the mean ergodic theorem,[10] $\bar{\psi}_T$ converges in $L^2(\chi, d\gamma_0)$ to some limiting function $\bar{\psi}_\infty$ which has γ_0-integral unity and is invariant under T^t. As in the proof of Corollary 2, the measure $\bar{\gamma}_\infty$ obtained by multiplying γ_0 by $\bar{\psi}_\infty$ has locally square-integrable correlation functions $\bar{\rho}(x)_n$. Trivially, we have

$$\int \frac{d(x)_n}{n!} f(x)_n \bar{\rho}_\infty(x)_n =$$
$$\int d\gamma_0 \bar{\psi}_\infty \sum f = \lim_{T\to\infty}\int d\gamma_0 \bar{\psi}_T \sum f$$
$$= \lim_{T\to\infty}\int \frac{d(x)_n}{n!} f(x)_n \bar{\rho}(T;(x)_n), \tag{8.3}$$

i.e.,

$$\bar{\rho}_\infty(x)_n = \lim_{T\to\infty}\bar{\rho}(T;(x)_n) \tag{8.4}$$

in the sense of distributions.

Moreover, the measure $\bar{\gamma}_\infty$ is time invariant and is obtained by multiplying γ_0 by a square-integrable function. Hence (Theorem 1) its correlation functions must satisfy the stationary BBGKY hierarchy. We have thus shown that the time-averaged correlation functions tend, as $T\to\infty$, to stationary correlation functions. Unfortunately, we do not know that these stationary correlation functions are the equilibrium ones.[11] This would follow if it could be proved that the equilibrium measure γ_0 is ergodic with respect to T^t. The ergodicity of the low-activity equilibrium measures is the outstanding problem in the theory of the time evolution of 1-dimensional systems, and no serious attack has yet been made on it.

Recently, Ruelle[12] has shown that, for a large class of potentials (the so-called superstable potentials) and for arbitrary temperature and activity, the finite-volume correlation functions satisfy an inequality of the form (5.5), where ξ may be chosen to be independent of Λ. One can construct infinite-volume equilibrium measures by taking limits along subsequences of boxes converging to infinity; the equilibrium measures obtained in this way need not be unique (i.e., they may depend on the particular sequence of boxes chosen), but any one of them is concentrated on $\hat{\mathfrak{X}}$, invariant under T^t and has correlation functions satisfying (5.5). It is easy to see that all our results, except those in Sec. 7, extend with appropriate modifications to apply to states obtained by making local perturbations on these equilibrium states.

* Alfred E. Sloan Foundation Fellow, also supported in part by U.S. Office Naval Research, Contract N 00014-69-A-0200-1002.

† Supported in part by the U.S. Air Force Office Special Research, under Grant 68-1416.

[1] D. Ruelle, *Statistical Mechanics* (Benjamin, New York, 1969).

[2] J. L. Lebowitz and E. Lieb, Phys. Rev. Letters **22**, 631 (1969).

[3] O. E. Lanford, Commun. Math. Phys. **9**, 126 (1968); **11**, 257 (1969).

[4] The existence of the equilibrium correlation functions for $h \neq 0$ for the systems considered here is a consequence of our general results. It may also be proven independently for all dimensions.

[5] See, for example, N. N. Bogoliubov, J. Phys. USSR **10**, 265 (1946) ("Problems of a Dynamical Theory in Statistical Physics," transl. E. Gora, Providence College, Providence, R.I., 1959.)

[6] D. Ruelle, J. Math. Phys. **8**, 1657 (1967).

[7] D. Ruelle, Ann. Phys. (N.Y.) **25**, 109 (1963).

[8] The last statement, which is the heart of the evolution theorem, means in effect that, if we concentrate our attention on the motion (during a finite time t) of the particles of a given configuration which are initially in a certain finite interval, their motion will not be much affected by the particles initially very far away (the actual size of the "region of influence" will, of course, depend on t and δ). It is this intuitively reasonable statement which provides the key to our ability of controlling the dynamics of our system at least to the extent of proving the rather primitive results of this paper.

[9] L. H. Loomis, *Abstract Harmonic Analysis* (Van Nostrand, Princeton, N.J., 1953).

[10] See P. R. Halmos, *Lectures on Ergodic Theory* (Chelsea, New York, 1956).

[11] The fact that the equilibrium correlation functions, for low activity, satisfy the BBGKY hierarchy even in higher dimensions and for more general potentials has been proven by G. Gallavotti, Nuovo Cimento **52b**, 208 (1968).

[12] D. Ruelle, Commun. Math. Phys., to be published.

Commun. math. Phys. 24, 193—210 (1972)

Approach to Equilibrium of Free Quantum Systems

O. E. LANFORD, III* and DEREK W. ROBINSON
Université d'Aix-Marseille, Centre de Luminy, France

Received October 12, 1971

Abstract. It is proved for fermi systems that each translationally invariant state ω with square integrable correlation functions approaches a limit under the free time evolution. The limit state is the gauge invariant quasi-free state with the same two-point function as ω and it is characterized by a maximum entropy principle. Various properties of the limit are discussed, and the extension of the results to bose systems is also given.

1. Introduction

The study of the structural properties of infinitely extended systems has played a useful role in the understanding of the nature and properties of equilibrium states. Similarly one would expect that analysis of the time development of such systems would aid the understanding of non-equilibrium phenomena. Very little work has however been done in this direction because it is notoriously difficult even to define the time-development of systems with an infinite number of degrees of freedom. In fact the only interacting systems for which one has a satisfactory definition are quantum spin systems [1] and a class of one-dimensional classical systems [2]. If, however, one turns to non-interacting systems the definition of time-development is relatively simple, and it is possible to analyse the properties of states of the system as they change with time. For example it has recently been shown that the equilibrium states of non-interacting classical systems have strong ergodic properties with respect to time and in fact provide examples of K-systems [3, 4]. Alternatively, for free systems of fermi particles, it has been argued that many states, possibly differing globally from equilibrium states, approach limiting equilibrium states as time progresses [5]. In this paper we will characterize a class of states of quantum systems which converge to a limit state as they evolve freely and examine properties of the convergence and the limiting "equilibrium states".

In the first part of this paper we consider exclusively fermions and in Sect. 2 we define the set of states whose time-development towards

* Alfred P. Sloan Foundation Fellow; on leave from Department of Mathematics, University of California, Berkeley, California (USA).

an equilibrium state is analyzed in Sect. 3. Properties of the approach to equilibrium are discussed in Sect. 4 and the method of extension of our results to bosons is presented in Sect. 5.

2. Square Integrable States

A system of fermi particles moving in the configuration space R^ν can be described in a well-known way by the C^*-algebra $\mathscr{A}$ associated with the canonical anti-commutation relations. We will adopt the standard notation and terminology used, for example, in Chapter VII of [6]. In particular, we denote by

$$f \in L^2(R^\nu) \to a(f) \in \mathscr{A}, \qquad g \in L^2(R^\nu) \to a^*(g) \in \mathscr{A}$$

the generating elements of $\mathscr{A}$ which satisfy the anti-commutation relations.

$$\{a(f), a^*(g)\} = (f, g) \qquad \text{etc.}$$

The group R^ν of space translations is represented as a group of strongly continuous automorphisms α of $\mathscr{A}$ whose action is defined by

$$\alpha_x(a(f)) = a(U_x f) \qquad x \in R^\nu$$

where

$$(U_x f)(y) = f(y - x) \qquad \text{etc.}$$

The one-parameter group of free time translations is also represented as a group of strongly continuous automorphisms τ of $\mathscr{A}$ and the action of this group is defined by

$$\tau_t(a(f)) = a(V_t f) \qquad t \in R$$

where

$$(V_t f)(x) = \frac{1}{(2\pi)^{\nu/2}} \int dp\, \hat{f}(p)\, e^{ip^2 t - ipx}$$

[$\hat{f}$ is the Fourier transform of f].

As $\mathscr{A}$ is generated by the $a(f)$, $a^*(g)$ each state ω over $\mathscr{A}$ is determined by the set of values

$$\{\omega(a^*(f_1) \ldots a^*(f_n)\, a(g_1) \ldots a(g_m)) ;\, f_1, \ldots, f_n, g_1, \ldots, g_m \in L^2(R^\nu)\} .$$

We introduce

$$W_{nm}(f_1, \ldots, f_n ; g_1, \ldots, g_m) = \omega(a^*(f_1) \ldots a^*(f_n)\, a(g_1) \ldots a(g_m)) .$$

The state ω is defined to be even if

$$W_{nm}(f_1, \ldots, f_n ; g_1, \ldots, g_m) = 0$$

whenever $n + m$ is odd.

An even state ω over $\mathscr{A}$ is also completely determined by the truncated functions ω^T_{nm} which are defined recursively by the following formulae

$$\omega_{nm}(f_1, \ldots, f_n; g_1, \ldots, g_m) = \sum_{\pi} (-1)^{\sigma(\pi)} \omega^T_{r_1 s_1}(f_{i_1}, \ldots, f_{i_{r_1}}; g_{j_1}, \ldots, g_{j_{s_1}}) \ldots$$
$$\ldots \omega^T_{r_p s_p}(f_{k_1}, \ldots, f_{k_{r_p}}; g_{l_1}, \ldots, g_{l_{s_p}})$$

where the sum is taken over all partitions π of $\{f_1, \ldots, f_n; g_1, \ldots, g_m\}$ into disjoint subsets $\{f_{i_1}, \ldots, f_{i_{r_1}}; g_{j_1}, \ldots, g_{j_{s_1}}\}, \ldots, \{f_{k_1}, \ldots, f_{k_{r_1}}, g_{l_1}, \ldots, g_{l_{s_p}}\}$. The f's and g's appear within each subset of a partition π in the same order that they appear in the original set and $\sigma(\pi)$ is the permutation of the set required to rearrange the f's and g's into the order in which they appear in the partitioned subsets. The requirement of evenness of ω ensures that this latter convention is unambiguous.

The state ω over $\mathscr{A}$ is said to be translationally invariant if

$$\omega(\alpha_x A) = \omega(A), \quad A \in \mathscr{A}, \; x \in R^\nu.$$

Each translationally invariant state is automatically even [7]. Such a state ω is completely determined by the set of functions

$$\omega^T_{\{f_n\}\{g_m\}}(x_2 - x_1, \ldots, x_{n+m} - x_{n+m-1})$$
$$= \omega^T_{nm}(U_{x_1} f_1, \ldots, U_{x_n} f_n; U_{x_{n+1}} g_1, \ldots, U_{x_{n+m}} g_m).$$

Definition 1. *The translationally invariant state ω over $\mathscr{A}$ is defined to be square integrable if*

$$\int d\xi_1 \ldots d\xi_{n+m-1} \, |\omega_{\{f_n\}\{g_m\}}(\xi_1, \ldots, \xi_{n+m-1})|^2 < +\infty \quad \textit{for} \quad n+m>2$$

and for all f_i, g_j such that $\tilde{f}_i, \tilde{g}_j \in \mathscr{D}$.

Remark 1. Using the continuity and linearity properties of $f \to a^*(f)$ etc., one can associate with each translationally invariant state ω a set of tempered distributions W^T_{nm} such that

$$\omega^T_{nm}(f_1, \ldots, f_n; g_1, \ldots, g_m)$$
$$= \int dx_1 \ldots dx_{n+m} W^T_{nm}(x_2 - x_1, \ldots, x_{n+m} - x_{n+m-1}) f_1(x_1) \ldots$$
$$\ldots f_n(x_n) \, g_1(x_{n+1}), \ldots, g_{n+m}(x_{n+m}).$$

If the W^T_{nm} are in fact square integrable then ω is a square integrable state in the sense of Definition 1. This definition actually allows the W^T_{nm} to have local singularities and corresponds to a condition of square integrability at infinity. It might seem natural to demand the square integrability condition to be satisfied for all $f_i, g_j \in \mathscr{D}$ but this is in fact a stronger condition than the specified requirement $\hat{f}_i, \tilde{g}_j \in \mathscr{D}$.

Remark 2. If ω is square integrable then it is automatically strongly mixing of all orders, i.e. for all $A_1, \ldots, A_n \in \mathscr{A}$

$$\lim_{\substack{\min|x_i - x_j| \to \infty \\ i \neq j}} \omega(\alpha_{x_1} A_1 \ldots \alpha_{x_n} A_n) = \omega(A_1) \ldots \omega(A_n),$$

and in particular R^ν ergodic. The existence of square integrable states is assured by the work of Ginibre [8] who shows that the equilibrium states of a large class of interacting systems have this property at low density. we suspect that the set of such states is weak*-dense among the set of all translationally invariant states over $\mathscr{A}$.

Before proceeding we recall that an even state ω over $\mathscr{A}$ is called quasi-free if

$$\omega_{nm}^T = 0 \quad \text{for} \quad n + m > 2$$

and is called a gauge invariant quasi-free state if

$$\omega_{nm}^T = 0 \quad \text{for} \quad (n, m) \neq (1, 1).$$

Further, if ω is an arbitrary state over $\mathscr{A}$ we can define the associated gauge invariant quasi-free state $\hat{\omega}$ by

$$\hat{\omega}(a^*(f)\, a(g)) = \omega(a^*(f)\, a(g)), \qquad f, g \in L^2(R^\nu)$$

and $\hat{\omega}_{nm}^T = 0$ for $(n, m) \neq (1, 1)$. It is well known that this definition actually does determine a state.

3. Approach to Equilibrium

We next wish to analyse the time development of square integrable states. Before this, however, we examine properties of two point functions of a general translationally invariant state.

Theorem 1. *If ω is a translationally invariant state over $\mathscr{A}$ then*

$$\omega(\tau_t(a^*(f)\, a(g))) = \omega(a^*(f)\, a(g)),$$

$$\lim_{t \to \infty} \omega(\tau_t(a(f)\, a(g))) = 0,$$

and

$$\lim_{t \to \infty} \omega(\tau_t(a^*(f)\, a^*(g))) = 0$$

for all $f, g \in L^2(R^\nu)$.

Proof. First, note that as $\|a(f)\| = \|a^*(f)\| = |f|_2$ one has

$$|\omega(a^*(f)\, a(g))| \leqq |f|_2\, |g|_2.$$

Thus there exists an operator A on $L^2(R^\nu)$ with $\|A\| \leqq 1$ such that

$$\omega(a^*(f)\, a(g)) = (f, Ag).$$

Now as ω is translationally invariant A commutes with the group of translation operators $f \to U_x f$ on $L^2(R^\nu)$ and hence is a multiplication operator in momentum space. Thus there is a function $\tilde{\varrho}$, with $|\tilde{\varrho}| \leqq 1$, such that

$$\begin{aligned}\omega(a^*(f)\, a(g)) &= \int dp\, \tilde{\varrho}(p)\, \bar{\tilde{f}}(p)\, \tilde{g}(p) \\ &= \omega(\tau_t(a^*(f)\, a(g)))\end{aligned}$$

and the first part of the theorem is proved.

Similarly there is an operator B with $\|B\| \leqq 1$, such that

$$\omega(a(f)\, a(g)) = (\bar{f}, Bg)\,.$$

Again B commutes with translations and hence there is a function $\tilde{\sigma}$ with $|\tilde{\sigma}| \leqq 1$, such that

$$\omega(a(f)\, a(g)) = \int dp\, \tilde{\sigma}(p)\, \tilde{f}(-p)\, \tilde{g}(p)\,.$$

But then we have

$$\omega(\tau_t(a(f)\, a(g))) = \int dp\, \tilde{\sigma}(p)\, \tilde{f}(-p)\, \tilde{g}(p)\, e^{2ip^2 t}$$

which goes to zero as $t \to \infty$ by the Riemann-Lebesgue Lemma.

Corollary 1. *If ω is a translationally invariant quasi-free state and $\hat{\omega}$ the associated gauge invariant quasi-free state then*

$$\lim_{t \to \infty} \omega(\tau_t A) = \hat{\omega}(A) \qquad A \in \mathscr{A}.$$

This follows simply by noting that the value of ω on each monomial $a^*(f_1) \ldots a^*(f_n)\, a(g_1) \ldots a(g_m)$ is a sum of products of the two point functions $\omega(a^*(f_i)\, a^*(f_j))$, $\omega(a^*(f_k)\, a(g_l))$, $\omega(a(g_p)\, a(g_q))$.

Remark 3. As positivity implies that

$$\omega((\lambda a^*(f) + a(g))^* (\lambda a^*(f) + a(g))) \geqq 0$$

one can deduce actually that

$$\tilde{\varrho}(p) \geqq 0 \quad \text{and} \quad \tilde{\varrho}(p)\,(1 - \tilde{\varrho}(-p)) \geqq |\tilde{\sigma}(p)|^2\,.$$

Further, one knows that if

$$\tilde{\varrho}(p)\,(1 - \tilde{\varrho}(-p)) = |\tilde{\sigma}(p)|^2$$

then ω is automatically a pure quasi-free state, but if this equality is not valid then the quasi-free state determined by $\tilde{\varrho}$ and $\tilde{\sigma}$ is a mixed state and in fact primary.

Theorem 2. *If ω is a square integrable state over $\mathscr{A}$ and $\hat{\omega}$ the associated gauge invariant quasi-free state then*

$$\lim_{t \to \infty} \omega(\tau_t A) = \hat{\omega}(A)\,, \qquad A \in \mathscr{A}.$$

Proof. The proof proceeds in four steps:

1. It suffices to prove that the truncated functions ω_{nm}^T satisfy the convergence property

$$\lim_{t\to\infty} \omega_{nm}^T(V_t f_1, \ldots, V_t f_n; V_t g_1, \ldots, V_t g_m) = 0$$

for all $(n, m) \neq (1, 1)$ and all f_i, g_i with Fourier transforms in $\mathscr{D}$. This follows because convergence of the state is equivalent to the convergence of the functions

$$\omega_{nm}(V_t f_1, \ldots, V_t f_n; V_t g_1, \ldots, V_t g_m)$$

for all $f, g \in L^2(R^\nu)$ which is in turn equivalent to the convergence of the truncated functions. To obtain this convergence it is sufficient to prove convergence for the f_i, g_i in a dense set of $L^2(R^\nu)$, for example $\tilde{f}_i, \tilde{g}_i \in \mathscr{D}$. Finally $\hat{\omega}$ is obtained from ω by setting the truncated functions with $(n, m) \neq (1, 1)$ equal to zero and hence the necessity that ω_{nm}^T must converge to zero.

2. The appropriate convergence of the two point functions has been dealt with in Theorem 1.

3. Now with $\tilde{f}_1, \ldots, \tilde{f}_n, \tilde{g}_1, \ldots, \tilde{g}_m \in \mathscr{D}$ we can choose $\tilde{h} \in \mathscr{D}$ such that

$$\tilde{h}\tilde{f}_i = f_i, \qquad i = 1, \ldots, n,$$
$$\tilde{h}\tilde{g}_j = \tilde{g}_j, \qquad j = 1, \ldots, m.$$

But then we note that

$$\begin{aligned} V_t f_i &= V_t(f_i * h) \\ &= f_i * V_t h \end{aligned}$$

where the star denotes the convolution product. Thus

$$\begin{aligned} \omega_{nm}^T(V_t f_1, \ldots, V_t g_m) &= \int dx_1 \ldots dx_{n+m} V_t h(x_1) \ldots V_t h(x_{n+m}) \\ &\quad \cdot \omega_{nm}^T(U_{x_1} f_1 \ldots U_{x_{n+m}} g_m) \\ &= \int dx_1 \ldots dx_{n+m} V_t h(x_1) \ldots V_t h(x_{n+m}) \\ &\quad \cdot \omega_{\{f_n\}\{g_m\}}^T(x_2 - x_1, \ldots, x_{n+m} - x_{n+m-1}). \end{aligned}$$

Alternatively this last equation can be written

$$\omega_{nm}^T(V_t f_1 \ldots V_t g_m) = \int d\xi_1 \ldots d\xi_{n+m-1} H_t(\xi_1 \ldots \xi_{n+m-1}) \omega_{\{f_n\}\{g_m\}}^T(\xi_1 \ldots \xi_{n+m-1})$$

where

$$\tilde{H}_t(p_1, \ldots, p_{n+m-1}) = \exp\{itE(p)\}\, \tilde{H}_0(p_1, \ldots, p_{n+m-1})$$

and

$$E(p) = p_1^2 + \sum_{i=2}^{n} (p_i - p_{i-1})^2 - \sum_{i=n+1}^{n+m-1} (p_i - p_{i-2})^2 - p_{n+m-1}^2 .$$

Now $\omega_{\{f_n\}\{g_m\}}^T$ is square integrable by assumption and H_t by construction.

Thus the proof of the theorem is complete if we can show that H_t tends L_2-weakly to zero as $t \to \infty$. This is however accomplished by the following version of the Riemann-Lebesgue Lemma which constitutes the fourth and final step of the proof.

Lemma 1. *Let Φ be a Lebesgue integrable function on R^k and F a non-zero bilinear form then it follows that*

$$\lim_{t\to\infty} \int dr_1 \dots dr_k \, \Phi(r_1 \dots r_k) \exp\{itF(r_1 \dots r_k)\} = 0 .$$

Proof. It suffices to show that for any $r = (r_1 \dots r_k) \neq (0, 0, \dots, 0)$ there is a neighbourhood N_r such that the lemma holds for any continuous Φ with support in N_r. Since $V_r F \neq 0$ we may choose a non-singular system of local coordinates $u = (u_1, \dots, u_r)$ at r with $u_1 = F$. If the support of Φ is contained in the coordinate neighbourhood we then have

$$\int dr \, \Phi(r) \, e^{itF(r)} = \int du \, \Phi(r(u)) \, |J(u)| \, e^{itu_1}$$

where J is the Jacobian of the transformation $(r_1, \dots, r_j) \to (u_1, \dots, u_j)$. But this latter expression converges to zero as $t \to \infty$ by the Riemann-Lebesgue Lemma.

Remark 4. If we had considered a more general time evolution

$$(V_t f)(x) = \frac{1}{(2\pi)^{\nu/2}} \int dp \, f(p) \, e^{i\omega(p)t - ipx}$$

then the foregoing results would be valid as long as the Jacobian of the transformation $(p_1, \dots, p_{n+m-1}) \to (u_1, \dots, u_{n+m-1})$ where

$$u_1 = \omega(p_1) + \sum_{i=2}^{n} \omega(p_i - p_{i-1}) - \sum_{i=n+1}^{n+m-1} \omega(p_i - p_{i-1}) - \omega(p_{n+m-1})$$

$$u_j = p_j , \qquad j = 2, 3, \dots, n+m-1$$

is non-singular for all n, m.

A slightly stronger result than that given in Theorem 2 can be established with a slight elaboration of the above proof.

Theorem 3. *Let ω be a square integrable state over $\mathcal{A}$ and $\hat{\omega}$ the associated gauge invariant quasi-free state. Denote by π_ω the representation of $\mathcal{A}$, on the Hilbert space $\mathcal{H}_\omega$, associated with ω by the Gelfand-Segal construction. It follows that*

$$\lim_{t\to\infty} \omega(A(\tau_t B)\, C) = \omega(AC)\, \hat{\omega}(B) , \qquad A, B, C \in \mathcal{A},$$

i.e.

$$\operatorname*{weak\,lim}_{t\to\infty} \pi_\omega(\tau_t B) = \hat{\omega}(B)\, 1_\omega , \qquad B \in \mathcal{A}$$

where 1_ω is the identity operator on $\mathcal{H}_\omega$.

It further follows that the strong limit of $\pi_\omega(\tau_t B)$ does not exist for all $B \in \mathscr{A}$.

Proof. The proof of the first statement proceeds again in several steps.

1. It suffices to prove the statement with A, B, C monomials in $a^*(f_i)$, $a(g_i)$, where $f_i, g_i \in \mathscr{D}$, ordered with the a^* at the left and the a at the right. This follows because the linear hull of such monomials is uniformly dense in $\mathscr{A}$.

2. Next use the anti-commutation relations to order the monomial $A\tau_t(B)\,C$ with the a^* at the left and the a at the right. This process gives an ordered monomial of the same order as $A\tau_t(B)\,C$ plus lower order terms each of which is proportional to an anti-commutator of an a from A (or an a^* from C) with an a^* from B (an a from B), i.e. proportional to a factor of the form $(V_t f, g)$. As $t \to \infty$ this latter factor tends to zero and consequently the lower order terms do not contribute to the limit. Thus we need only study the highest order monomial, the ordered form of $A\tau_t(B)\,C$.

3. If we now write the value of this latter monomial in the state ω in terms of truncated functions we obtain, using the first statement of Theorem 1, a sum of terms exactly equal to $\omega(AC)\,\hat{\omega}(B)$ plus a number of t-dependent terms. It remains to prove that these latter terms tend to zero.

4. Introduce h as in step 3 of the proof of Theorem 2. Each of the relevant terms has a factor expressible in the form

$$\int d\xi_1 \ldots d\xi_{n+m-1}\, W^T_{\{f_n\}\{g_n\}}(\xi_1 \ldots \xi_{n+m-1})\, H_t(\xi_1 \ldots \xi_{n+m-1})$$

where

$$\tilde{H}_t(p_1 \ldots p_{n+m-1}) = \exp\{itE'(p)\}\, \tilde{H}_0(p_1 \ldots p_{n+m-1})$$

and E' is a non-zero bilinear form in a subset of the variables $p_1 \ldots p_{n+m-1}$. The proof of the desired result is then obtained from Lemma 1.

The last statement of the theorem is easily proved by contradiction. Assume $\pi_\omega(\tau_t a(f))$ converges strongly as $t \to \infty$; then it must tend strongly to zero because we have established that it converges weakly to zero. Thus $\pi_\omega(\tau_t(a^*(f)\,a(f)))$ converges weakly to zero. Similarly $\pi_\omega(\tau_t(a(f)\,a^*(f)))$ converges weakly to zero. But

$$\pi_\omega(\tau_t(a^*(f)\,a(f))) + \pi_\omega(\tau_t(a(f)\,a^*(f))) = |f|_2^2\, 1_\omega$$

by the anti-commutation relations which is a contradiction.

Remark 5. In the above theorems we have established the existence of pointwise limits of states or expectation values. One can also give general conditions under which ergodic averages exist. Firstly, if ω is

τ-invariant then the limits

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T dt\, \omega(A\tau_t(B)\, C) \qquad A, B, C \in \mathscr{A}$$

exist because $\mathscr{A}$ is R-abelian [7, 9]. The limit can be identified and in particular if ω is extremal τ-invariant then it gives $\omega(AC)\,\omega(B)$. This circumstance is of interest in examing the problem of return to equilibrium. Secondly, if the truncated functions are not square integrable in the sense of Definition 2 but are simply the Fourier transforms of measures then the limits

$$\lim_{t\to\infty} \frac{1}{T} \int_0^T dt\, \omega(\tau_t A), \qquad A \in \mathscr{A}$$

exist but are not necessarily equal to $\omega(A)$.

4. Properties of Approach to Equilibrium

In the previous section we have shown that a class of states approach a limit state, or equilibrium state, in the limit of infinite time as the system evolves freely. A priori one might be tempted to argue that the limit state should be identifiable with the Gibbs equilibrium state compatible with the given energy and particle densities; this is clearly not the case. The reason for this discrepancy is however immediate. The free evolution is pathological in the sense that it allows many too many constants of the motion. In particular Theorem 1 establishes that the gauge invariant two point function remains constant in time. Thus the information that could possibly be infered about the more realistic situation of an interacting system is limited. Nevertheless one can use the information gathered for the freely evolving system to check a number of general principles and provide examples and counter examples to conjectured behaviour. We now examine a few such points but it is first necessary to recall a few facts and definitions.

A state ω over $\mathscr{A}$ is called *locally normal* if its restriction to each $\mathscr{A}_\Lambda$ [the subalgebra generated by $\{a(f), a^*(g); f, g \in L^2(\Lambda)\}$] is normal with respect to the Fock representation of $\mathscr{A}_\Lambda$, i.e. a locally normal state is determined by a family of density matrices $\{\varrho_\Lambda\}$ on the Fock spaces $\{\mathscr{H}(\Lambda)\}$ in the following manner

$$\omega(A) = \mathrm{Tr}_{\mathscr{H}(\Lambda)}(\varrho_\Lambda A) \qquad A \in \mathscr{A}_\Lambda .$$

If ω is a translationally invariant locally normal state over $\mathscr{A}$ its *mean entropy* $S(\omega)$ is defined by

$$S(\omega) = \lim_{\Lambda\to\infty} -V(\Lambda)^{-1}\, \mathrm{Tr}_{\mathscr{H}(\Lambda)}(\varrho_\Lambda \log \varrho_\Lambda)$$

where $V(\Lambda)$ is the volume of Λ and the limit over the net of increasing parallelepipeds is known to exist [10].

A state ω is said to have *finite density* if for all $\Lambda \subset R^\nu$

$$N_\Lambda(\omega) = V(\Lambda)^{-1} \sum_{i \geqq 1} \omega(a^*(f_i)\, a(f_i)) < +\infty$$

where $\{f_i\}_{i \geqq 1}$ is an orthonormal basis of $L^2(\Lambda)$. A finite density state is locally normal, and, further,

$$N_\Lambda(\omega) = \mathrm{Tr}_{\mathscr{H}(\Lambda)} \left(\varrho_\Lambda \frac{N_\Lambda}{V(\Lambda)} \right)$$

where N_Λ is the unbounded operator which measures particle number on $\mathscr{H}(\Lambda)$. Clearly, the restriction of finite density is solely a condition on the gauge invariant two point function; if ω is translationally invariant it is readily shown [11] that $N_\Lambda(\omega)$ is independent of $\Lambda (= N(\omega))$ and

$$N(\omega) = \int dp\, \tilde{\varrho}(p)$$

in the notation of Theorem 1.

Next let us define the state ω to be *N-entire (N-analytic)* whenever ω is locally normal and

$$\mathrm{Tr}_{\mathscr{H}(\Lambda)} (\varrho_\Lambda\, e^{\alpha N_\Lambda}) < +\infty$$

for all $\Lambda \subset R^\nu$ and all $\alpha \in R$ (for some $\alpha > 0$). This definition is motivated by the following facts [12]. If ω is invariant under space translations and N-entire then the following quantity exists and is finite

$$e_N(\omega) = \sum_{r \geqq 0} \frac{|\alpha|^r}{r!} \lim_{\Lambda \to \infty} \mathrm{Tr}_{\mathscr{H}(\Lambda)} \left(\varrho_\Lambda \left(\frac{N_\Lambda}{V(\Lambda)} \right)^r \right), \qquad \alpha \in R$$

where the limit is again over increasing parallelepipeds; further, if ω is α-ergodic, i.e. extremal among the translationally invariant states, then $e_N(\omega)$ is identifiable as follows:

$$e_N(\omega) = \sum_{r \geqq 0} \frac{|\alpha|^r}{r!} \lim_{\Lambda \to \infty} \left[\mathrm{Tr}_{\mathscr{H}(\Lambda)} \left(\varrho_\Lambda \frac{N_\Lambda}{V(\Lambda)} \right) \right]^r .$$

Thus an α-ergodic state which is N-entire has small density fluctuations of all orders.

Finally recall that there are two notions of faithfulness of a state over a C^*-algebra. A state ω over $\mathscr{A}$ is *weakly faithful* if

$$\omega(A^* A) = 0$$

implies $A = 0$. The state ω is *strongly faithful* if the cyclic vector Ω_ω associated with it by the Gelfand-Segal construction is separating for

the von Neumann algebra π''_ω generated by the corresponding representation π_ω (equivalently Ω_ω is cyclic for the commutant π'_ω of the representation π_ω).

We now examine how these properties are affected by the approach to equilibrium.

A. Maximum Entropy Principle

The Gibbs equilibrium state of a free fermi gas can be characterized as the translationally invariant state which maximises the mean entropy at fixed energy and particle density. Although the states we are considering do not necessarily approach the Gibbs state in the equilibrium limit we will show that the state which they do approach has the maximum mean entropy compatible with the constants of the motion.

Proposition 1a. *Let ω be a translationally invariant state of finite mean density which has the property that*

$$\lim_{t\to\infty} \omega(\tau_t A) = \hat{\omega}(A), \qquad A \in \mathscr{A}$$

where $\hat{\omega}$ is the gauge invariant quasi-free state associated with ω. Further let K_ω denote the set of all translationally invariant states with the same two point function $(f, g) \to \omega(a^*(f)\, a(g))$ as ω and $\hat{\omega}$. It follows that

$$S(\hat{\omega}) = \sup_{\omega' \in K_\omega} S(\omega').$$

Proof. It suffices to prove, for an arbitrary $\omega' \in K_\omega$, that $S(\omega') \geqq S(\hat{\omega})$. Now as ω has finite mean density

$$\begin{aligned} N(\omega) = N(\omega') &= V(\Lambda)^{-1} \sum_{i \geqq 1} \omega(a^*(f_i)\, a(f_i)) \\ &= V(\Lambda)^{-1} \sum_{i \geqq 1} (f_i, A f_i) \\ &= V(\Lambda)^{-1} \operatorname{Tr}_{L^2(\Lambda)}(A) < +\infty \end{aligned}$$

where A is the operator which determines the two-point function of ω (cf. proof of Theorem 1) restricted to $L^2(\Lambda)$. Thus A is of trace-class. Let $\{\lambda_i\}_{i \geqq 1}$ and $\{f_i\}_{i \geqq 1}$ be the eigenvalues and a complete orthonormal set of eigenfunctions of A respectively. Further let $\mathscr{H}_i$ be the Fock space (2-dimensional) associated with the algebra generated by $a(f_i)$ and $a^*(f_i)$. $\mathscr{H}(\Lambda)$ has a tensor decomposition of the form $\mathscr{H}_i \otimes R_i$ and in fact

$$\mathscr{H}(\Lambda) = \prod_{i \geqq 1}^{\otimes} \mathscr{H}_i .$$

Next define σ_i as the matrix on $\mathscr{H}_i$ given by

$$\begin{aligned}\sigma_i &= \mathrm{Tr}_{R_i}(\varrho'_\Lambda) \\ &= \lambda_i\, a^*(f_i)\, a(f_i) + (1-\lambda_i)\, a(f_i)\, a^*(f_i)\end{aligned}$$

where ϱ'_Λ is the density matrix associated with ω'. It is easily checked that the density matrix associated with $\hat{\omega}$ is given by

$$\hat{\varrho}_\Lambda = \prod_{i\geqq 1}^{\otimes} \sigma_i\,.$$

Now using Lemma 1 of [7] we have

$$\begin{aligned}-\mathrm{Tr}_{\mathscr{H}(\Lambda)}(\varrho'_\Lambda \log \varrho'_\Lambda) &\leqq -\mathrm{Tr}_{\mathscr{H}(\Lambda)}(\varrho'_\Lambda \log \hat{\varrho}_\Lambda) \\ &= -\sum_{i\geqq 1} \mathrm{Tr}_{\mathscr{H}(\Lambda)}(\varrho'_\Lambda \log \sigma_i) \\ &= -\sum_{i\geqq 1} \mathrm{Tr}_{\mathscr{H}(\Lambda)}(\hat{\varrho}_\Lambda \log \sigma_i) \\ &= -\mathrm{Tr}_{\mathscr{H}(\Lambda)}(\hat{\varrho}_\Lambda \log \hat{\varrho}_\Lambda)\,.\end{aligned}$$

Dividing by $V(\Lambda)$ and taking the limit $\Lambda \to \infty$ gives the desired result.

Remark 6. If ω is translationally invariant and ω_t is defined by $\omega_t(A) = \omega(\tau_t A)$, $A \in \mathscr{A}$ then in general $S(\omega_t) = S(\omega)$, i.e. the entropy is a constant of the motion. This however does not rule out the increase of the entropy in the limit $t \to \infty$ as S is usually only a semi-continuous function. An example in which the mean entropy increases strictly is given by a non-gauge invariant quasi-free state which in the limit approachs the associated gauge invariant quasi-free state [13]. We suspect that this strict increase is a general property, i.e. the supremum in Proposition 1a is attained by only one state namely $\hat{\omega}$.

B. Density Fluctuations and Local Normality

Proposition 1b. *Let ω be a locally normal translationally invariant state which has the property that*

$$\lim_{t\to\infty} \omega(\tau_t A) = \hat{\omega}(A)\,, \qquad A \in \mathscr{A}$$

where $\hat{\omega}$ is the gauge invariant quasi-free state associated with ω.

It follows that $\hat{\omega}$ is locally normal if, and only if, ω has finite mean density and in the case $\hat{\omega}$ is N-entire.

Proof. That a quasi-free state is locally normal if and only if it has finite mean density is demonstrated for example in [14]. Assuming ω and

hence $\hat{\omega}$ has finite mean density then we have with the notation of Sect. 4a.

$$\begin{aligned}\mathrm{Tr}_{\mathscr{H}(\Lambda)}\left(\hat{\varrho}_\Lambda\, e^{\alpha N_\Lambda}\right) &= \prod_i \mathrm{Tr}_{\mathscr{H}_i}\left(\sigma_i\, e^{\alpha a^*(f_i)\, a(f_i)}\right)\\ &= \prod_{i \geqq 1} \left(1 + (e^\alpha - 1)\, \lambda_i\right)\\ &\leqq \exp\left\{(e^\alpha - 1) \sum_{i \geqq 1} \lambda_i\right\}\\ &= \exp\left\{(e^\alpha - 1)\, N(\omega)\, V(\Lambda)\right\}\end{aligned}$$

and hence $\hat{\omega}$ is N-entire.

C. Faithfulness

Proposition 1c. *If ω is a weakly faithful translationally invariant state which has the property that*

$$\lim_{t \to \infty} \omega(\tau_t A) = \hat{\omega}(A), \qquad A \in \mathscr{A}$$

where $\hat{\omega}$ is the gauge invariant quasi-free state associated with ω then it follows that $\hat{\omega}$ is strongly faithful.

Proof. The weak faithfulness of ω implies

$$\omega(a^*(f)\, a(f)) > 0, \qquad \omega(a(f)\, a^*(f)) > 0, \qquad f \in L^2(R^\nu)$$

and hence

$$1 > \tilde{\varrho} > 0$$

almost everywhere. Now one can define a one-parameter group of automorphisms of $\mathscr{A}$ by the following action on the generating elements

$$s \to \sigma_s(a(f)) = a(X_s f)$$

where

$$(X_s f)(x) = \int dp \left(\frac{\tilde{\varrho}(p)}{1 - \tilde{\varrho}(p)}\right)^{is} \tilde{f}(p)\, e^{-ipx}.$$

It is easily checked that the state $\hat{\omega}$ satisfies the K.M.S. boundary condition [16] with respect to this group of automorphisms, for example

$$\begin{aligned}\hat{\omega}(a^*(f)\, \sigma_i(a(g))) &= \int dp\, \tilde{f}(p)\, \tilde{g}(p)\, \tilde{\varrho}(p) \left(\frac{\tilde{\varrho}(p)}{1 - \tilde{\varrho}(p)}\right)^{-1}\\ &= \hat{\omega}(a(g)\, a^*(f)).\end{aligned}$$

It then follows from [16] that $\hat{\omega}$ is strongly faithful.

D. Purity

Mathematically it is natural to ask if starting with a pure state ω one always obtains a limiting state $\hat{\omega}$ which is also pure. We will now demonstrate that this is not generally the case by citing an example.

Consider the quasi-free state whose two point functions are given by

$$\tilde{\varrho}(p) = e^{-p^2}/1 + e^{-p^2}$$

and

$$\tilde{\sigma}(p) = e^{-p^2/2}/1 + e^{-p^2}.$$

As $|\tilde{\sigma}(p)|^2 = \tilde{\varrho}(p)(1 - \tilde{\varrho}(-p))$ this state is a pure translationally invariant state. Nevertheless the state attained in the equilibrium limit, the gauge invariant quasi-free state determined by $\tilde{\varrho}(p)$, is a type III factor state (and incidentally a Gibbs equilibrium state).

5. Bosons

Examination of the time development of free bose systems is technically more complicated because it is impossible to give an algebraic description in which the time evolution is realised as a group of strongly continuous automorphisms. By suitable choice of the underlying C^*-algebra one can retain the automorphism property but the continuity is lost. Nevertheless we will show that the time evolution can be defined as a continuous mapping of a subset of states and use this formalism to extend the foregoing results.

We will work with a C^*-algebra $\mathscr{A}$ defined in the following manner. On each Fock space $\mathscr{H}(\Lambda)$, $\Lambda \subset R^\nu$, we define $\mathscr{A}_\Lambda$ to be the C^*-algebra generated by the set of Weyl operators

$$\{U(f), V(g); f, g \in \mathscr{D} \cap L^2(\Lambda)\}.$$

Using the canonical identification of $\mathscr{A}_\Lambda$ as a subalgebra of $\mathscr{A}_{\Lambda'}$ whenever $\Lambda \subset \Lambda'$ we can construct the algebra $\mathscr{A}$ as the uniform closure of the union (over $\Lambda \subset R^\nu$) of the $\mathscr{A}_\Lambda$.

Note that the group R^ν is represented as a group of automorphisms α of $\mathscr{A}$ whose action is defined by

$$\alpha_x(U(f)) = U(U_x f) \quad \text{etc.}$$

where again

$$(U_x f)(y) = f(y - x)$$

but this group is not strongly continuous because of the easily established relation

$$\|U(f) - U(g)\| = \|U(f - g) - 1\| = 2 \quad \text{if} \quad f \neq g.$$

This latter relation also makes it impossible to translate the free evolution of the test functions

$$f \in \mathscr{D} \to V_t f \in \mathscr{S}, \qquad t \in R$$

into a group of strongly continuous automorphisms. This difficulty can however be dealt with as follows. First for convenience introduce (on Fock space)

$$W(f,g) = U(f)\, V(g)\, e^{-i(f,g)/2} = e^{i[\Phi(f)+\pi(g)]}$$

where $\Phi(f)$ and $\pi(g)$ denote the infinitesimal generators of $U(f)$ and $V(g)$ respectively.

Definition 2. *If ω is a state over $\mathscr{A}$ such that for each pair $f, g \in \mathscr{D}$ the function*

$$(f,g) \to \omega(W(f,g))$$

is continuous on $\mathscr{S} \times \mathscr{S}$ we define the time evolved state ω_t by

$$\omega_t(W(f,g)) = \omega(W(V_t f, V_t g)) \qquad t \in R\,.$$

We will now prove that if ω is a translationally invariant state with finite mean density then it has the continuity property demanded by the definition and in fact ω_t also has finite mean density.

Lemma 2. *If ω is a state of finite density and $f, g \in \mathscr{D} \cap L^2(\Lambda)$ then*

$$\|(W_\omega(f,g) - 1)\,\Omega_\omega\|^2 \leqq (2N_\Lambda(\omega)\, V(\Lambda) + 1)\,[|f|_2 + |g|_2]^2$$

where Ω_ω is the cyclic vector associated with ω and W_ω the representative of W.

Proof. As ω is locally normal we can write

$$\begin{aligned}\|(W_\omega(f,g)-1)\,\Omega_\omega\|^2 &= \omega((W(f,g)-1)^*\,(W(f,g)-1)) \\ &= \mathrm{Tr}_{\mathscr{H}(\Lambda)}\,(\varrho_\Lambda (e^{i(\Phi(f)+\pi(g))}-1)^*\,(e^{i(\Phi(f)+\pi(g))}-1)) \\ &\leqq \mathrm{Tr}_{\mathscr{H}(\Lambda)}\,(\varrho_\Lambda(\Phi(f)+\Phi(g))^2)\end{aligned}$$

because for a self-adjoint operator A on a Hilbert space

$$\|(e^{iA}-1)\,\Psi\| \leqq \|A\,\Psi\|\,, \qquad \Psi \in D(A)\,.$$

Hence as on the Fock space $\mathscr{H}(\Lambda)$ one has

$$\Phi(f)^2 \leqq (2N_\Lambda + 1)\,|f|_2^2 \quad \text{etc.}$$

we have

$$\|(W_\omega(f,g)-1)\,\Omega_\omega\|^2 \leqq \mathrm{Tr}_{\mathscr{H}(\Lambda)}\,(\varrho_\Lambda(2N_\Lambda+1))\,[|f|_2^2 + |g|_2^2 + 2|f|_2\,|g|_2]$$

where we have used the Schwarz inequality. This proves the Lemma.

Next let us introduce a norm $|||\,.\,|||$ on $\mathscr{S}$ as follows. Consider R^ν divided into cubic cells of unit size $\Lambda_1, \Lambda_2, \ldots$. Then the norm is defined by

$$|||f||| = \sum_{i \geqq 1} \left[\int_{\Lambda_i} dx\, |f(x)|^2\right]^{\frac{1}{2}}, \qquad f \in \mathscr{S}\,.$$

It is easily to verify that this norm is continuous. We now have

Lemma 3. *If ω is a translationally invariant state of finite mean density then for $f, g \in \mathscr{D}$,*

$$\|(\Phi_\omega(f) + \pi_\omega(g))\,\Omega_\omega\| \leqq \sqrt{2N_{\Lambda_1}(\omega)+1}\,[|||f||| + |||g|||]$$

and hence

$$\|(W_\omega(f,g) - 1)\,\Omega_\omega\| \leqq \sqrt{2N_{\Lambda_1}(\omega)+1}\,[|||f||| + |||g|||]\,.$$

Proof. We have immediately that

$$\begin{aligned}\|(\Phi_\omega(f) + \pi_\omega(g))\,\Omega_\omega\| &= \Big\|\sum_i (\Phi_\omega(f_i) + \pi_\omega(g_i))\,\Omega_\omega\Big\| \\ &\leqq \sum_i \|(\Phi_\omega(f_i) + \pi_\omega(g_i))\,\Omega_\omega\|\end{aligned}$$

where $f_i = f\chi_i$, $g_i = g\chi_i$ and χ_i is the characteristic function of Λ_i. The inequalities now follow as in Lemma 2 with the use of translational invariance to identify the various $N_{\Lambda_i}(\omega)$.

Lemma 4. *If ω is a translationally invariant state of finite mean density then the associated representation $W_\omega(f,g)$, which is defined for $f, g \in \mathscr{D}$, is strongly continuous with respect to the topology induced by $|||\,.\,|||$. Hence the representation W_ω of the commutation relations extends by continuity to a representation defined and continuous on $\mathscr{S}$. Further for $f, g \in \mathscr{S}$ the vector Ω_ω is in the domain of $\Phi_\omega(f) + \pi_\omega(g)$ and*

$$\|(\Phi_\omega(f) + \pi_\omega(g))\,\Omega_\omega\| \leqq \sqrt{2N_{\Lambda_1}(\omega)+1}\,[|||f||| + |||g|||]\,.$$

Proof. By the commutation relations and the cyclicity of Ω_ω the continuity of W_ω follows from its continuity on Ω_ω. It also follows from the commutation relations that W_ω on Ω_ω is continuous if it is continuous at the origin. This however follows from Lemma 3. The second inequality of Lemma 3 remains valid for $f, g \in \mathscr{S}$ by continuity; hence

$$\left\|\left(\frac{W_\omega(tf, tg) - 1}{t}\right)\Omega_\omega\right\| \leqq \sqrt{2N_{\Lambda_i}(\omega)+1}\,[|||f||| + |||g|||]\,.$$

But it follows by spectral theory (for an explicit proof see [15]) that if A is a self-adjoint operator on a Hilbert space then $\Psi \in D(A)$ if, and only if

$$\left\|\left(\frac{e^{iAt} - 1}{t}\right)\Psi\right\| \leqq C_\Psi$$

where C_Ψ is a constant independent of t. Further in such a case

$$\|A\Psi\| = \lim_{t \to 0}\left\|\left(\frac{e^{iAt} - 1}{t}\right)\Psi\right\| \leqq C_\Psi\,.$$

The remaining statements of the Lemma are then established by applying this result with $A = \Phi_\omega(f) + \pi_\omega(g)$.

Thus we have at this point established that each translationally invariant state with finite mean density satisfies the continuity conditions required to define its time evolution. But it is also possible to deduce the following result.

Theorem 4. *If ω is a translationally invariant state with finite mean density and ω_t its time translate, then $t \to \omega_t$ is a weak* continuous one parameter family of translationally invariant states with constant (finite) density.*

Proof. It is clear that ω_t is translationally invariant and the continuity follows from Lemma 3 but it remains to prove that the mean density is constant. Since however Ω_ω is in the domain of $\Phi_\omega(f) + \pi_\omega(g)$ for all $f, g \in \mathscr{S}$ we can introduce the bilinear form

$$f, g \in \mathscr{S} \to \left((\Phi_\omega(f) + i\pi_\omega(f))\,\Omega_\omega, (\Phi_\omega(g) + i\pi_\omega(g))\,\Omega_\omega\right).$$

By Lemma 3 this form is separately continuous in f and g and hence is determined by a tempered distribution in two variables. But using translation invariance we see in fact that

$$\left((\Phi_\omega(f) + i\pi_\omega(f))\,\Omega_\omega, (\Phi_\omega(g) + i\pi_\omega(g))\,\Omega_\omega\right) = \int dp\, \tilde{\varrho}(p) \tilde{f}(p)\, g(\tilde{p})$$

where $\tilde{\varrho}$ is a tempered distribution. (As the form is positive one has $\tilde{\varrho} \geqq 0$.) From this formula it now follows that

$$\begin{aligned} \left\|(\Phi_\omega(V_t f) + i\pi_\omega(V_t f))\,\Omega_\omega\right\| &= \left\|(\Phi_\omega(f) + i\pi_\omega(f))\,\Omega_\omega\right\| \\ &= \left\|(\Phi_{\omega_t}(f) + i\pi_{\omega_t}(f))\,\Omega_{\omega_t}\right\|. \end{aligned}$$

Consequently Ω_t is a state of finite density and

$$N_\Lambda(\omega_t) = N_\Lambda(\omega)$$

which completes the proof of the theorem.

If we wish to derive theorems similar to Theorems 3 and 2 for fermions it is necessary to consider a more restricted class of states, namely those possessing Wightman functions of all orders. Thus we now consider the translationally invariant C_∞ states, i.e. the states ω for which Ω_ω is in the domain of all polynomials of $\Phi_\omega(f)$ and $\pi_\omega(g)$ and

$$(\Phi_\omega(f_1) + \pi_\omega(g_1)) \ldots (\Phi_\omega(f_n) + \pi_\omega(g_n))\,\Omega_\omega$$

is continuous in f_i, g_i, $i = 1, 2, \ldots, n$.

Remark 7. If ω is translationally invariant and a C_∞ state for the local number operators, i.e. if ω is locally normal and

$$\mathrm{Tr}_{\mathscr{H}(\Lambda_1)}(\varrho_\Lambda \mathrm{N}_\Lambda^m) < +\infty$$

for all $\Lambda \subset R^\nu$ and $m = 1, 2, \ldots$ then it follows that ω is C_∞ because one can show that

$$\|\Phi_\omega(f_1) \ldots \Phi_\omega(f_n)\, \Omega_\omega\|^2 \leqq \mathrm{Tr}_{\mathscr{H}(\Lambda)}\, (\varrho_{\Lambda_1}(2N_{\Lambda_1} + 2n - 1)^n) \prod_{i=1}^{n} |||f_i|||$$

and a similar inequality if some of the $\Phi_\omega(f_i)$ are replaced by $\pi_\omega(f_i)$.

Finally for C_∞ states of finite mean density an analysis of the approach to equilibrium can be carried out in a manner parallel to that of the previous sections but the following changes should be noted.

1. The one point functions $\omega(\Phi(f))$, $\omega(\pi(g))$ are not automatically zero but due to translation invariance do not change with time.

2. The truncated part of the non-gauge invariant two point functions

$$\omega((\Phi(f) - i\pi(f))\,(\Phi(g) - i\pi(g)))$$

must be assumed square integrable, in the sense of Definition 1, together with the higher functions.

3. The convergence of a square integrable state to the associated gauge invariant quasi-free state as $t \to \infty$ is no longer weak* convergence but convergence of the Wightman functions.

References

1. Robinson, D. W.: Commun. math. Phys. **7**, 337 (1968).
2. Lanford, O. E.: Commun. math. Phys. **9**, 176 (1968).
3. Sinai, Y., Volkoviskij, K.: (Unpublished).
4. de Pazzis, O.: (To be published).
5. Haag, R., Kadison, R., Kastler, D.: (Unpublished).
   — Systèmes à un nombre infini de degrés de Liberté CNRS. C.N.R.S. (Paris) (1970).
6. Ruelle, D.: Statistical mechanics. New York: Benjamin 1969.
7. Lanford, O. E., Robinson, D. W.: J. Math. Phys. **9**, 1120 (1968).
8. Ginibre, J.: J. Math. Phys. **6**, 252 (1965).
9. Lanford, O. E., Ruelle, D.: J. Math. Phys. **8**, 1460 (1967).
10. Araki, H., Lieb, E.: Commun. math. Phys.
11. Miracle-Sole, S., Robinson, D. W.: Commun. math. Phys. **14**, 235 (1969).
12. Robinson, D. W.: (Unpublished).
13. Lanford, O. E., Robinson, D. W.: (To be published).
14. Verbeure, A.: Cargèse lectures in physics (1969). New York: Gordon & Breach 1970.
15. Courbage, M., Miracle-Sole, S., Robinson, D. W.: Ann. Henri Poincaré, vol. XIV. **2**, 171 (1971).
16. Haag, R., Hugenholtz, N. M., Winnink, M.: Commun. math. Phys. **5**, 215 (1967).

D. W. Robinson
Centre de Physique Théorique
C.N.R.S.
31, chemin J. Aiguier
F-13 Marseille 9^0, France

Université de Grenoble - Summer School of Theoretical Physics
Les Houches 1970
Supported by NATO
and the Commissariat à l'Energie Atomique

STATISTICAL MECHANICS AND QUANTUM FIELD THEORY

MECANIQUE STATISTIQUE ET THEORIE QUANTIQUE DES CHAMPS

edited by C. DeWitt
Faculté des Sciences, Grenoble and
University of North Carolina, Chapel Hill
R. Stora
Centre d'Etudes Nucléaires de Saclay
Service de Physique théorique

1971

GORDON AND BREACH SCIENCE PUBLISHERS
New York London Paris

Selected Topics in Functional Analysis

Oscar E. Lanford III

Dept. of Mathematics
University of California, Berkeley, California

I Introduction

In this set of lectures I will attempt to present some of the mathematical tools which are currently of importance in the investigation of the statistical mechanics of infinite systems. It is my intention to give a reasonably self-contained *introduction* to the rudiments of the theory of integral representations on compact convex sets (Choquet theory), C^*-algebra theory, and von Neumann algebra theory. To do this I will need to make frequent use of the results of more "elementary" branches of mathematics – general topology, measure theory, Hilbert space theory, and the theory of topological vector spaces. While it is out of the question to try to make a systematic presentation of these preliminaries, I will attempt to summarize some of the most important points. With this summary, a little elementary knowledge of real analysis, and some good will, I hope it will be possible to make at least some sense out of the lectures.

A General Topology

We start with a very elementary definition: A *topological space* is a set X together with a collection $\mathcal{G}$ of subsets of X, called open sets, such that:

1) X and ϕ (the null set) both belong to $\mathcal{G}$.

2) The intersection of two elements of $\mathcal{G}$ again belongs to $\mathcal{G}$.

3) The union of any family of elements of $\mathcal{G}$ belongs to $\mathcal{G}$.

Intuitively, one thinks of an open set as one which, if it contains x, also contains all other points "sufficiently near" to x. Thus, specifying a topology (a family of open sets) may be viewed as introducing a notion of approximation; from this point of view, the motivation for studying general topology, as opposed to the theory of metric spaces, may be seen as providing tools to investigate notions of approximation more subtle than those where one has a numerical measure of closeness. We will see plenty of examples later on.

If X is a topological space, a subset F of X is said to be *closed* if its complement $X \setminus F$ is open. In metric spaces, one has another way of describing closed sets—a set is closed if no sequence in it converges to a point outside it. We can give a corresponding characterization in general topology, at the expense of introducing the notion of net, which generalizes that of sequence. First, we recall that a sequence (x_n) in a set X is an indexed family of elements of X, the index set being $\{1, 2, 3, \ldots\}$. We say that x_n converges to x if for every open set W containing x, x_n is eventually in W, i.e., there is a N such that, for $n \geq N$, $x_n \in W$. A net in X is, like a sequence, an indexed family of elements of X, but the index set is more general than $\{1, 2, 3, \ldots\}$. We say

that a binary relation $\geqq$ on a set A is a *pre-order* if we have $\alpha \geqq \alpha$ for all $\alpha \in A$, $\alpha \geqq \beta$ and $\beta \geqq \gamma$ implies $\alpha \geqq \gamma$ for $\alpha, \beta, \gamma \in A$. A pre-ordered set A is said to be *directed* (*filtrant*) if, for any $\alpha, \beta, \in A$, there exists $\gamma \in A$ such that $\gamma \geqq \alpha$ and $\gamma \geqq \beta$. A *net* in a set X is an indexed family (x_α) of elements of X, indexed by a directed set A (i.e., a net in X is a mapping $\alpha \mapsto x_\alpha$ from a directed set A to X). A net (x_α) *converges to* x if, for any open set W containing x, there exists an $\alpha_0 \in A$ such that $x_\alpha \in W$ for $\alpha \geqq \alpha_0$. (We express this by saying that x is *eventually* in W.) We can now provide the promised characterization of closed set: A set F in a topological space is closed if, whenever a net (x_α) in F converges to a point x in X, the point x is actually in F.

It is perfectly possible (perhaps even desirable) to develop general topology without ever mentioning nets. The advantage of talking about nets is purely heuristic; it permits the use of intuition about sequences in Euclidean space to suggest arguments valid in general topological spaces. This tool must, however, be used cautiously; not *every* argument about sequences has a counterpart valid for nets. We shall soon see some examples.

It follows from the above characterization of closed set that, if one knows which nets converge to which limits, one knows what all the closed sets are; hence, what all the open sets are, i.e., one knows the topology. This suggests that one can specify a topology by specifying convergent nets. This can indeed be done, but of course not every collection of nets is precisely the set of convergent nets for some topology. (For example, a subnet (to be defined shortly) of a convergent net, like a subsequence of a convergent sequence, must converge.) For an axiomatization of topological spaces in terms of convergent nets, rather than open sets, see Kelley [1], p. 73. We will frequently define topologies by specifying what nets coverge to what limits, leaving it as an exercise to verify that such a topology exists. (It is usually easier to construct the topology directly than to verify that the conditions in Kelley's book hold.) As an example, consider the product topology: Let $(X_i)_{i \in I}$ be an indexed family of topological spaces; we define a topology on the product set $\prod_i X_i$ by requiring that a net (x^i_α) converges to (x^i) if and only if $\lim x^i_\alpha = x^i$ for all i. It is simple to verify that this description is equivalent to the usual construction of a product topology on $\prod_i X_i$.

In metric spaces, there are many equivalent characterizations of compact spaces: A space is compact if every open cover has a finite subcover, or if every sequence has a cluster point, or if every sequence has a convergent subsequence. These conditions cease to agree in general topological spaces. The first condition turns out to be the most useful one and is therefore taken as the general definition of compactness; examples can then be made of

subsequence. (It remains true that every sequence in a compact space has a cluster point.) However, by replacing "sequence" by "net", the alternative descriptions of compact sets can be saved. Thus, one says that a point x is a *cluster point* of a net (x_α) if, for every open set W containing x, and every $\alpha \in A$, there is an $\alpha' \geqq \alpha$ such that $x_{\alpha'} \in W$ (We express this by saying that the net (x_α) is *frequently* in W). A topological space is then compact if and only if every net has a cluster point. A *subnet* of a net $(x_\alpha)_{\alpha \in A}$ is a net of the form $\beta \mapsto x_{\phi(\beta)}$ where ϕ is a mapping of the directed set B to the directed set A such that, for any $\alpha \in A$ there exists $\beta \in B$ with the property that $\phi(\beta') \geqq \alpha$ if $\beta' \geqq \beta$. (In other words, $\phi(\beta)$ becomes "large" in A as β becomes "large" in B.) Note that a subnet of $(x_\alpha)_{\alpha \in A}$ may have an index set which is not merely a subset of A. In particular, sequences may have subnets which are not sequences. Now it turns out that, with this definition, a point x is a cluster point of a net (x_α) if and only if the net has a subnet converging to x. Thus, a topological space X is compact if and only if every net has a convergent subnet.

There is yet another characterization of compact spaces, having no sequence analogue, corresponding to the fact that one can construct nets which cannot really be refined any further, and which hence converge if they have any cluster point at all. To be precise: A net (x_α) is said to be a *universal net* if, for *any* subset Y of X, (x_α) is either eventually in Y or eventually in $X \setminus Y$. It is nearly a tautology that, if (x_α) is a universal net and x is a cluster point of (x_α), then $\lim x_\alpha = x$. It is also not hard to see that, if (x_α) is a universal net in X and if f is *any* mapping $X \to X'$, then $(f(x_\alpha))$ is a universal net in X'. What is not at all obvious, but true nevertheless, is that every net has a universal subnet (See Kelley[1], Chapter 2, Ex. J., p. 81.) Hence, a topological space is compact if and only if every universal net converges. This result leads immediately to Tychonoff's Theorem:

THEOREM *Let $(X_i)_{i \in I}$ be an indexed family of compact spaces. Then $\prod_i X_i$ is compact.*

Proof Let (x_i^α) be a universal net in $\prod_i X_i$. Then, for each i, x_i^α is a universal net in X_i; hence, converges. By the definition of the product topology (x_i^α) converges in $\prod_i X_i$.

Let us look at another, related, example. Let $\mathfrak{X}$ be a Banach space, $\mathfrak{X}^*$ the dual space of $\mathfrak{X}$ (the space of continuous linear functionals on $\mathfrak{X}$). The *weak-* topology* on $\mathfrak{X}^*$ is the topology for which ϕ_α converges to ϕ means $\phi_\alpha(\xi)$ converges to $\phi(\xi)$ for all $\xi \in \mathfrak{X}$. We now have the following theorem, which is usually derived as a corollary of the Tychonoff Theorem:

Theorem. The unit ball in the dual of a Banach space is compact in the weak *-top

Proof Let ϕ_α be a universal net in the unit ball of $\mathcal{X}^*$. For any $\xi \in \mathcal{X}$, $\phi_\alpha(\xi)$ is a universal net of complex numbers of absolute value $\leqq \|\xi\|$; since the set of such numbers is compact (ξ fixed!) $\phi_\alpha(\xi)$ converges for all ξ. Call the limit $\phi(\xi)$. Then $\xi \mapsto \phi(\xi)$ is linear, and $|\phi(\xi)| = \lim |\phi_\alpha(\xi)| \leqq \|\xi\|$ for all ξ, so ϕ is an element of the unit ball of $\mathcal{X}^*$, and the net ϕ_α converges in the weak-* topology to ϕ. Thus, every universal net in the unit ball of $\mathcal{X}^*$ converges, so the unit ball of $\mathcal{X}^*$ is compact.

In many cases, we find ourselves forced to consider several different topologies on a single set. (For example, on the dual of a Banach space, one may consider the norm topology as well as the weak-* topology.) The comparison of different topologies is usually confusing although in principle simple: We say that a topology $\mathcal{S}_1$ is stronger, or finer, than a topology $\mathcal{S}_2$ if every $\mathcal{S}_2$-open set is also $\mathcal{S}_1$-open (i.e., if $\mathcal{S}_1$ has more open sets than $\mathcal{S}_2$). Since a set is closed if and only if its complement is open, $\mathcal{S}_1$ also has more closed sets than $\mathcal{S}_2$. Since the closure of a set $E \subset X$ is the intersection of all closed sets containing E, the closure in a stronger topology is smaller than in a weaker topology. If x_α is a net converging to x in a topology $\mathcal{T}$, it also converges to x in all weaker topologies, but not necessarily in stronger topologies. If one starts with a continuous map between two spaces, and weakens the topology on the range space or strengthens the topology on the domain space, the map remains continuous. If on the other hand, one strengthens the topology on the range space or weakens the topology on the domain space the map need not remain continuous. On the dual of a Banach space, the weak-* topology is, as one would hope, weaker than the topology defined by the norm.

B Measure Theory

We continue with our impressionistic survey of the elements of analysis, by looking, first at abstract measure theory, then at measure theory on locally compact spaces. To start with: Let X be a set, Σ a collection of subsets of X. We say that Σ is a *ring* if $\phi \in \Sigma$ and if, for all $E, F \in \Sigma$, $E \cup F$, $E \cap F$, and $E \setminus F$ belong to Σ, and we say that Σ is a *σ-ring* if, in addition, $\bigcup_{i=1}^{\infty} E_i \in \Sigma$ whenever $E_1, E_2, \ldots$ all belong to Σ. If the set X itself belongs to Σ we say that Σ is an *algebra* or *σ-algebra*. Given any collection of subsets, there is a unique smallest σ-ring containing all of them; we refer to this as the σ-ring *generated by* the collection of sets in question. Given a σ-ring Σ, we define a measure μ on Σ to be a mapping from Σ to the positive real numbers and $+\infty$ such that if $E_1, E_2, \ldots$ are in Σ with $E_i \cap E_j = \phi$ for

If E_j then $\mu(E_1 \cup E_2 \cup \ldots) = \sum \mu(E_j)$ countable additive) (To avoid

pathologies, we also require $\mu(\phi) = 0$; this eliminates the possibility $\mu(E) = \infty$ for all $E \in \Sigma$.)

With this general set-up (a set, a σ-ring, and a measure), one can define an integral, and one gets an integration theory with most of the nice formal properties of the Lebesgue-integral. In particular the monotone convergence theorem and the dominated convergence theorem hold. (It should be remarked that these theorems hold for *sequences*; the analogous statements for nets are false. For example, every positive real-valued function on the real line is the pointwise limit of an increasing net of functions each of which is zero except at a finite number of points.)

Now let X be a locally compact topological space (*i.e.*, one in which each point belongs to some open set with a compact closure.) We want to investigate integration theory which will permit us to integrate continuous functions of compact support. As a minimal sort of condition to permit us to do this, we must have that, for any continuous non-negative function f of compact support, and every $\alpha > 0$, $\{x \in X : f(x) \geqq \alpha\}$ belongs to the σ-ring Σ. This set is closed and contained in the support of f; hence, is compact. However, it is a rather special kind of compact set: It can be written as the intersection of countably many open sets by $\{x : f(x) \geqq \alpha\} = \bigcap_{n=1}^{\infty} \{x : f(x) > \alpha - 1/n\}$. A set which can be obtained as a countable intersection of open sets is called a G_δ. (In a metric space, every closed set is a G_δ, so the distinction between general compact sets and compact G_δ's becomes irrelevant.)

We consider, therefore, measures on the σ-ring generated by the compact G_δ's. Elements of this σ-ring are called *Baire sets*. In order to ensure that the integral of every continuous function of compact support is finite, we also impose the condition that the measure of every compact G_δ is finite. Thus: A *Baire measure* on a locally compact space is a measure on the σ-ring of Baire sets which is finite on compact G_δ's. It turns out that every continuous function of compact support is integrable with respect to every Baire measure μ. The mapping $f \mapsto \int f\,d\mu$ is linear and positive (positive means that $\int f\,d\mu \geqq 0$ if $f(x) \geqq 0$ everywhere). Thus each Baire measure determines a positive linear functional on the vector space of continuous functions of compact support. Conversely, every such functional comes from a measure, and the measure is uniquely determined by the functional. Thus we have:

THEOREM I. B. 1 (*Preliminary form of the Riesz Representation Theorem*): *The positive linear functionals on the space of continuous function of compact support on a locally compact space X are in one-one correspondence with the Baire measures on X, the correspondence being defined by associating with a*

measure μ the functional $f \rightsquigarrow \int f d\mu$

This version of the Riesz Theorem is not fully satisfactory, since it does not permit us to speak of the measure of a general compact set. Hence, we want to extend Baire measures from the σ-ring of Baire sets to the larger σ-ring of *Borel* sets, defined as the σ-ring generated by the compact sets. It is always possible to do this, but the extension need not be unique. There is, however, a useful way of picking out a unique extension. To describe how this is done, we first mention a remarkable continuity property of Baire measures: Let μ be a Baire measure, E a Baire set. Then $\mu(E) = \sup\{\mu(K): K \text{ a compact } G_\delta \text{ contained in } E\} = \inf\{\mu(0): 0 \text{ an open Baire set containing } E\}$.

In other words, Baire sets can be approximated arbitrarily well from the inside by compact G_δ's and from the outside by open Baire sets. We say that a Borel measure μ (measure on the σ-ring of Borel sets assigning finite values to compact sets) is inner regular if, for each Borel set E, $\mu(E) = \sup\{\mu(K): K \text{ compact}; K \subset E\}$; outer regular if, for each Borel set E, $\mu(E) = \inf\{\mu(0): 0 \text{ an open Borel set}, E \subset 0\}$; regular if both inner regular and outer regular. It can then be shown that every Baire measure has a unique extension to a regular Borel measure. Hence:

THEOREM I, B. 2 (*Riesz Representation Theorem*): *Let X be a locally compact topological space. Then the set of positive linear functionals on the space of continuous functions of compact support on X is in one-one correspondence with the set of regular Borel measures on X, the correspondence being that associating with a measure μ the functional $f \mapsto \int f d\mu$.*

It should be made clear that the terminology we have adopted is far from universal. We have followed that used by Halmos [1], Chapter X. In particular the term "Borel set" is frequently taken to mean an element of the σ-algebra generated by the closed sets. A regular measure defined on the σ-ring generated by the compact sets may always be extended to a measure on the σ-algebra generated by the closed sets. The extension, however, is not in general unique. Furthermore, it cannot in general be taken to be both inner and outer regular. It is, however, always possible to extend a Borel measure to an inner regular measure on the σ-algebra generated by the closed sets, and the extension is evidently unique.

Once this extension has been made, we may formulate a slight modification of the monotone convergence theorem which will be useful later on. A real-valued function f is said to be *lower semi-continuous* if, for all real γ, $\{x: f(x) > \gamma\}$ is open. (Alternatively, f is lower semi-continuous if it is the supremum of a family of continuous functions.) f is *upper semi-continuous* if $-f$ is lower semi-continuous. If μ is a regular Borel measure extended as above to an inner-regular measure on the σ-algebra generated by

the closed sets and if f is a non-negative lower semi-continuous function then $\int f\,d\mu$ makes sense (but may be $+\infty$). We now have the following result:

THEOREM I. B. 3 *Let f_α be an increasing net of non-negative lower semi-continuous functions, and let $f = \lim_\alpha f_\alpha$. Then f is lower semi-continuous and $\int f\,d\mu = \lim_\alpha \int f_\alpha\,d\mu$. (In other words, the monotone convergence theorem is valid for increasing nets, rather than sequences, provided that the functions involved are lower semi-continuous.)*

The proof of the theorem is not difficult; it uses the inner regularity of μ and the fact that, if K is a compact set and $f(x) > \gamma$ for $x \in K$ then $f_\alpha(x) > \gamma$ for $x \in K$ for all sufficiently large α.

C Topological Vector Spaces

A *semi-norm* on a vector space E is a function $\xi \mapsto \|\xi\|$ from E to the non-negative real numbers such that:

a) $\|\lambda\xi\| = |\lambda| \cdot \|\xi\|$

b) $\|\xi + \eta\| \leq \|\xi\| + \|\eta\|$.

The first condition evidently implies that $\|0\| = 0$; if in addition we have

c) $\|\xi\| = 0$ implies $\xi = 0$ we say that $\|\cdot\|$ is a *norm* on E.

Given a family $(\|\cdot\|_i)_{i \in I}$ of semi-norms on a vector space E, we may define a topology on E by requiring that $\xi_\alpha \to \xi$ if and only if $\|\xi_\alpha - \xi\|_i \to 0$ for all i. This makes E into a topological vector space (i.e., a vector space, with a topology, such that the algebraic operations are continuous); a topological vector space whose topology obtained in this way from a family of semi-norms is called a *locally convex* topological vector space. We will normally want to consider only *Hausdorff* locally convex spaces, i.e., those for which, for any $\xi \neq 0$, there exists i such that $\|\xi\|_i \neq 0$. If the topology of a locally convex space is defined by a single norm, we speak of a *normed* vector space; a complete normed vector space is called a *Banach* space.

If E is any vector space, and ϕ is a linear functional on E, then $\ker(\phi) = \{\xi \in E: \phi(\xi) = 0\}$ is a linear subspace of E of codimension 1 (i.e., $E/\ker(\phi)$ is one-dimensional). (A linear subspace of codimension 1 in a vector space, or, more generally, a subset obtained by translating such a subset, is called a *hyperplane*.) Conversely, if E is a vector space, and F is a hyperplane in E passing through 0, then there is a linear functional on E with F as its kernel, and this functional is unique up to multiplication by a scalar.

If now E is a topological vector space, and ϕ is a continuous linear

functional, then $\ker \phi$ is close. Conversely if $\ker \phi$ is closed then ϕ is

continuous. It is worth remarking that, since the closure of a linear subspace is again a linear subspace, the closure of a hyperplane is either the hyperplane itself or all of E, i.e., a hyperplane is either closed or dense. Hence, if we have a hyperplane and a non-empty open set which does not intersect it, then the hyperplane must be closed, i.e., must be a translate of the kernel of a continuous linear functional. If E is a locally convex space, with a topology defined by a family of semi-norms $(\|\cdot\|_\iota)$, then a linear functional ϕ is continuous if and only if there exists a positive real number λ and finitely many indices $i_1, \ldots, i_n$ such that

$$\|\phi(\xi)\| \leqq \lambda(\|\xi\|_{\iota_1} + \cdots + \|\xi\|_{\iota_n}).$$

This characterization generalizes the familiar fact that a linear functional on a normed vector space is continuous if and only if it is bounded.

Perhaps the most powerful tool in the study of locally convex spaces is the Hahn-Banach Theorem, in its many different forms. We will state an algebraic version of the theorem which has all the other versions as more or less straightforward consequences. First we need some terminology. A set K in a vector space is *convex* if, for $\xi_1, \xi_2 \in K$ and $0 \leqq \alpha \leqq 1$, $\alpha\xi_1 + (1 - \alpha) \times \xi_2 \in K$. Geometrically, this means that the line segment from ξ_1 to ξ_2 is contained in K. A point ξ of a convex set K in a vector space E is said to be an *algebraic interior point* of K if, for all $\eta \in E$, $\xi + \lambda\eta \in K$ for sufficiently small positive λ. In other words, ξ is an algebraic interior point of K if one can move from ξ a finite distance in any direction without getting outside of K (but the distance may depend upon the direction chosen.) If E is a locally convex topological vector space, then any topological interior point of K (i.e., a point contained in an open set contained in K) is an algebraic interior point, but the converse need not be true. (But in a finite dimensional space, an algebraic interior point is the same thing as a topological interior point...).

The key theorem is now the following:

THEOREM I. C. 1 *Let E be a vector space over the real numbers, X a convex subset of E with at least one algebraic interior point, η a point of E not belonging to X. Then there exists a hyperplane F in E separating η from X in the sense that η is on one side of the hyperplane or on it and X is on the other side of the hyperplane (but possibly intersecting it). In other words, there is a non-zero linear functional ϕ on E and a real number λ such that $\phi(\xi) \leqq \lambda$ for $\xi \in X$ and $\phi(\eta) \geqq \lambda$.*

Note that the condition $\phi(\xi) \leqq \lambda$ for all $\xi \in X$ implies that $\phi(\xi) < \lambda$ for all algebraic interior points of X. Indeed, if ξ is an algebraic interior point and $\phi(\xi) = \lambda$, then, choosing $\zeta \in E$ such that $\phi(\zeta) > 0$, and choosing $\gamma > 0$ such that $\xi - \gamma\zeta \in X$ we get $\phi(\xi + \gamma\zeta) > \lambda$ contra —

dicting the requirement that $p \leq \lambda$ on X.

We can now read off a geometric form of the Hahn Banach Theorem:

THEOREM I. C. 2 (*Geometric Form of the Hahn Banach Theorem*): *Let E be a locally convex topological vector space over the real numbers; if X and Y are disjoint convex subsets of E at least one of which has a non-empty interior, then there is a closed hyperplane F in E separating X and Y. In other words, there is a continuous linear functional ϕ and a real number λ such that $\phi(\cdot) \geqq \lambda$ on X and $\phi(\cdot) \leqq \lambda$ on Y.*

If we drop the assumption that one of the sets X, Y has an interior point, but assume instead that there is an open set U containing 0 such that $(X + U) \cap Y = \phi$, then there is a closed hyperplane strictly separating X and Y in the sense that $\phi(\cdot) \geqq \lambda + \epsilon$ on X, and $\phi(\cdot) \leqq \lambda - \epsilon$ on Y, for some $\epsilon > 0$. This condition holds in particular if X is compact, Y is closed, and $X \cap Y = \phi$.

Proof For the first assertion, note that the set $X - Y = \{\xi - \eta : \xi \in X, \eta \in Y\}$ is convex, has a non-empty interior, and does not contain 0. Hence, by the preceding theorem there is a non-zero linear functional ϕ such that $\phi(\cdot) \geqq 0$ on $X - Y$. Let $\lambda = \inf\{\phi(\xi) : \xi \in X\}$. Then $\lambda \geqq \sup\{\phi(\chi) : \chi \in Y\}$, so the hyperplane $\{\xi : \phi(\xi) = \lambda\}$ separates X and Y. It remains only to prove that ϕ is continuous, but this is immediate since, by the argument given just above, the hyperplane $\phi(\xi) = \lambda$ cannot intersect the interior of either X or Y; hence, cannot be dense; hence, must be closed. To prove the second assertion: since E is locally convex, we may assume that U is convex; replacing U by $U \cap (-U)$, we may also assume that U is symmetric. Apply the first assertion to the sets $X + \frac{1}{2}U$, $Y + \frac{1}{2}U$. Then we get $\phi(\cdot) \geqq \lambda$ on $X + \frac{1}{2}U$ and $\phi(\cdot) \leqq \lambda$ on $Y + \frac{1}{2}U$. But $\phi(U)$ is a symmetric open interval; if we choose $\epsilon > 0$ so that $\phi(U) \supset (-2\epsilon, 2\epsilon)$, then $\phi(X + \frac{1}{2}U) \supseteq \phi(X) + (-\epsilon, \epsilon)$; since $\phi(\cdot) \geqq \lambda$ on $X + \frac{1}{2}U$, $\phi(\cdot) \geqq \lambda + \epsilon$ on X. Similarly, $\phi(\cdot) \leqq \lambda - \epsilon$ on Y. To prove the final assertion, we must first show that $X - Y$ is closed. Let $\xi_\alpha - \eta_\alpha$ be a convergent net in $X - Y$. Since X is compact we can, by passing to a subnet, assume that ξ_α converges to $\xi \in X$. Since $\xi_\alpha - \eta_\alpha$ and ξ_α converge separately, η_α must also converge, and the limit η must belong to Y since Y is closed. Hence, $\xi_\alpha - \eta_\alpha$ converges to $\xi - \eta \in X - Y$, so $X - Y$ is closed. Since $0 \notin X - Y$, there is a neighborhood U of 0, which may be taken to be convex and symmetric, such that $(X - Y) \cap U = \phi$. Then $(X + U) \cap Y = \phi$.

COROLLARY I. C.3 *Let E be a Hausdorff locally convex topological vector space over the real numbers, and let $\xi \in E$, $\xi \neq 0$. Then there is a continuous linear functional ϕ on E with $\phi(\xi) \neq 0$.*

Proof Let the topology on E be defined by the family of seminorms

$(\|\cdot\|_t)_{t\in I}$: Then $\{0\} = \bigcap \{\eta \in E : \|\eta\|_t = 0\}$ is closed similarly $\{\xi\}$ is closed.

Clearly, $\{\xi\}$ is compact. Hence, there exists a continuous linear functional ϕ strictly separating $\{0\}$ from $\{\xi\}$. Since $\phi(0) = 0$, we must have $\phi(\xi) \neq 0$.

Next, we derive the so-called analytic form of the Hahn-Banach Theorem.

THEOREM I. C. 4 (*Analytic Form of the Hahn Banach Theorem*). *Let E be a vector space over the real numbers, F a subspace of E, $\|\cdot\|$ a semi-norm on E, and ϕ a linear functional on F such that $|\phi(\xi)| \leqq \|\xi\|$ for $\xi \in F$. Then there exists a linear functional $\tilde{\phi}$ on E which extends ϕ and satisfies $|\tilde{\phi}(\xi)| \leqq \|\xi\|$ for all $\xi \in E$.*

Proof Consider the vector space $E \times \mathbb{R}$ with the product topology and the two convex sets

$$X = \text{graph of } \phi, \quad Y = \{(\xi, \lambda) : \lambda > \|\xi\|\}.$$

Then $X \cap Y = \phi$, and every point of Y is an algebraic interior point. Hence, we can find a hyperplane separating $X - Y$ from 0, i.e., separating X from Y. Let the hyperplane come from a linear functional ψ. Then, since X is a linear subspace $\psi(X) = \{0\}$ or $\mathbb{R}$, and the second alternative is ruled out by the condition that ψ separate X from Y. Hence, $\psi(X) = \{0\}$. Since every point of Y is an algebraic interior point, we must have $\psi(Y) \cap \psi(X) = \phi$. Thus, if $\psi((\zeta, \mu)) = 0$, then $(\zeta, \mu) \notin Y$, so $\mu \leqq \|\zeta\|$. Since $\ker(\psi)$ is a linear subspace, we get $-\mu \leqq \|-\zeta\| = \|\zeta\|$, so $|\mu| \leqq \|\zeta\|$. Thus, $\ker(\psi)$ is the graph of a linear functional $\tilde{\phi}$ satisfying $|\tilde{\phi}(\zeta)| \leqq \|\zeta\|$. Since $\ker(\psi) \supset X$ = graph of ϕ, $\tilde{\phi}$ extends ϕ.

COROLLARY I. C. 5 *Let E be a locally convex topological vector space over the real numbers, F a subspace of E, ϕ a continuous linear functional on F. Then ϕ may be extended to a continuous linear functional on E.*

The analytic form of the Hahn-Banach Theorem has a generalization which is occasionally useful: A sublinear functional on a vector space E is a mapping $p : E \to \mathbb{R}$ such that

a) $p(\lambda\xi) = \lambda p(\xi)$ for all $\lambda \geqq 0$ and all $\xi \in E$ (p is *positively homogeneous*).

b) $p(\xi + \eta) \leqq p(\xi) + p(\eta)$ for all $\xi, \eta \in E$ (p is *subadditive*).

Any semi-norm is sublinear, as is any linear functional. Essentially the same argument as that used in proving the analytic form of the Hahn-Banach Theorem proves:

THEOREM I. C. 6 *Let E be a vector space over the real numbers, p a sublinear functional on E, F a subspace of E, ϕ a linear functional on F such that $\phi(\xi) \leqq p(\xi)$ for all $\xi \in F$. Then ϕ may be extended to a linear functional $\tilde{\phi}$ on E such that $\tilde{\phi}(\xi) \leqq p(\xi)$ for all $\xi \in E$.*

We will also have occasion to use the following:

THEOREM I. C. 7 (*Extension Theorem for Positive Functionals*): *Let E be a vector space over the real numbers, K a convex cone in E (i.e., if $\xi, \eta \in K$, $\lambda > 0$, then $\xi + \eta$ and $\lambda\xi$ are in K). Let F be a linear subspace of E containing an algebraic interior point of K, and let ϕ be a linear functional on E which is*

non-negative on $E \setminus K$. Then there is an extension $\tilde{\Phi}$ of Φ to E which is non-

negative on K. If E is a locally convex space and F contains a point of the interior of K, then $\tilde{\phi}$ is continuous.

Proof We can assume $\phi \neq 0$. Then we must have $\phi(\xi) > 0$ for ξ any algebraic interior point of K. The set of algebraic interior points of K is convex, and does not intersect $\ker(\phi)$. Hence, there exists a linear functional ψ on E separating K from $\ker(\phi)$. Since $\ker(\phi)$ is a linear subspace of E, $\psi(\ker(\phi)) = \{0\}$ or $\mathbb{R}$; because of the separation property, $\psi(\ker(\phi)) = \{0\}$, i.e., $\ker(\psi) \supset \ker(\phi)$. Again by the separation property, ψ must take on only one sign on the set of algebraic interior points of K; we can assume, then, that $\psi(\xi) \geq 0$ for any algebraic interior point ξ of K. As usual, this implies $\psi(\xi) > 0$ for any algebraic interior point of K. If ξ is any algebraic interior point of K which is actually in F, we let $\tilde{\phi}$ be a multiple of ψ such that $\tilde{\phi}(\xi) = \phi(\xi)\ (>0)$. Now the restriction of $\tilde{\phi}$ to F is a linear functional on F whose kernel contains that of ϕ and which agrees with ϕ on a vector not in the kernel of ϕ; this implies that $\tilde{\phi}$ is an extension of ϕ. It remains to check that $\tilde{\phi}(\cdot) \geq 0$ on K; we know that this is true on the set of algebraic interior points of K. But if ξ' is an algebraic interior point of K, and ξ any point of K, then $\lambda\xi + (1-\lambda)\xi'$ is an algebraic interior point of K for $0 \leq \lambda < 1$. Thus, $\lambda\tilde{\phi}(\xi) + (1-\lambda)\tilde{\phi}(\xi') \geq 0$ for $0 \leq \lambda < 1$, so $\tilde{\phi}(\xi) \geq 0$.

For simplicity, we have limited the above discussion to real vector spaces. For complex vector spaces, the analytic form of the Hahn Banach Theorem remains valid as stated; the separation theorems have only to be changed by replacing "linear functional" by "real part of a complex linear functional" and "hyperplane" by "set of the form: $\{\xi\colon \mathrm{Re}\,\{\phi(\xi)\} = \lambda\}$, where ϕ is a non-zero complex-linear functional."

We can now describe a procedure for contructing new topologies for locally convex spaces. Let E be such a space, ϕ a continuous linear functional on E. Then $\xi \mapsto |\phi(\xi)|$ is easily verified to be a semi-norm on E. We consider the topology defined by all such semi-norms. In other words, we consider the topology such that $\xi_\alpha \to \xi$ if and only if $\phi(\xi_\alpha) \to \phi(\xi)$ for all continuous linear functionals ϕ. This topology is clearly weaker (i.e., not stronger) than the initial topology. (It may of course coincide with the initial topology.) It is, however, not too much weaker; every linear functional continuous for the initial topology is continuous for this new topology. In fact, the new topology can be uniquely described as the weakest topology with the same continuous linear functionals as the initial topology. It is called the *weak* (or, more properly, *weakened*) topology of E. The corollary to the geometric form of the Hahn-Banach Theorem ensures that the weak topology associated with a Hausdorff topology is again Hausdorff. One of the main reasons why weak topologies are useful is that they frequently have many more compact sets than the initial topology. For example, the unit ball of a Hilbert space is always compact for the weak topology, but is not compact

for the norm topology unless the space is finite dimensional

As a prelude to Choquet theory, we will show how to derive the Krein-Milman theorem from the Hahn-Banach Theorem. The Krein-Milman Theorem says that a compact convex set is generated by its extremal points. A point of a convex set K is said to be an *extremal point* if it is not an internal point of any line segment contained in K, i.e., if it cannot be written as $\alpha\xi_1 + (1 - \alpha)\xi_2$ for $\xi_1 \neq \xi_2 \in K$ and $0 < \alpha < 1$. (A little thought shows that ξ is an extremal point of K if it cannot be written as $\frac{1}{2}(\xi_1 + \xi_2)$ with $\xi_1, \xi_2 \in K$ and $\xi_1 \neq \xi_2$.) The extremal points of a triangle (or of any convex polygon) are the corners; the extremal points of the circle $\{(x, y): x^2 + y^2 \leq 1\}$ are all the boundary points. We let $\epsilon(K)$ denote the set of extremal points of the convex set K.

THEOREM I. C. 8 (Krein-Milman Theorem) *Any convex compact set in a locally convex topological vector space is the closed convex hull of its set of extremal points.*

In other words: We start with a convex compact set K. We form the set of its extremal points. We then form the set of all convex combinations of these extremal points, and finally take the closure of this set. This gives us a compact convex set contained in K; the Krein-Milman Theorem asserts that it is all of K. Note that it is not at all obvious that a convex compact set has any extremal points at all; indeed, the proof of the Krein-Milman Theorem reduces essentially to proving that extremal points do exist.

The proof of the Krein-Milman Theorem uses as a technical device the notion of a support of a compact convex set. Let K be a compact convex set in a locally convex topological vector space, and let $A \subset K$. A is said to be a *support* of K if A is non-empty, convex, and closed (hence compact), and if, whenever $\xi \in A$ can be written as $\frac{\xi_1 + \xi_2}{2}$, with $\xi_1, \xi_2 \in K$, then ξ_1 and ξ_2 both belong to A. For example, if ϕ is a continuous linear functional on E, and if $\lambda = \sup\{\phi(\xi): \xi \in K\}$, then $\{\xi \in K: \phi(\xi) = \lambda\}$ is a support of K. A set consisting of a single point is a support of K if and only if is an extremal point of K. The key technical lemma in the proof of the Krein-Milman Theorem is the following:

LEMMA I. C. 9 *Every support of K contains an extremal point of K.*

Using this lemma, we can easily prove the Krein-Milman Theorem. Let K be a compact convex set, and let K_1 be the closed convex hull of the set of extremal points of K. Then K_1 is a compact convex set contained in K; we want to show that it is all of K. Suppose not; let $\xi \in K \setminus K_1$. By the geometric form of the Hahn-Banach Theorem, there is a continuous linear functional ϕ on E such that $\phi(\xi) > \sup\{\phi(\eta): \eta \in K_1\}$. Let $\lambda = \sup\{\phi(\zeta): \zeta \in K\}$. Then $\{\zeta \in K: \phi(\zeta) = \lambda\}$ is a support of K; hence, contains an extremal point of K But by assumption K_1 contains all the extremal points of K. and

$\lambda > \sup\{\phi(n) : n \in K_1\}$ This is a contradiction and proves

It remains to prove the lemma. This is an elementary exercise in the use of Zorn's Lemma.

ZORN'S LEMMA *Let A be a partially ordered set. Suppose that every linearly ordered subset B of A has an upper bound (B is linearly ordered if, for all $\alpha, \beta \in B$, either $\alpha \geqq \beta$ or $\beta \geqq \alpha$. An upper bound for a subset B is an element γ of A such that $\gamma \geqq \beta$ for all $\beta \in B$). Then if α is any element of A, there is a maximal element γ of A with $\gamma \geqq \alpha$ (an element $\gamma \in A$ is maximal if $\gamma' \geqq \gamma$ implies $\gamma' = \gamma$).*

We consider the set of supports of K, ordered by $A \geqq B$ if $A \subset B$. (Note that, somewhat confusingly, supports are "large" in the sense of the ordering if they are "small" sets, and that a maximal element of the partially ordered set is a minimal support, i.e., a support which contains no strictly smaller support.) We will use Zorn's Lemma to prove that every support contains a minimal support, and then argue that a minimal support must reduce to a single point. To prove that every support contains a minimal support, it suffices to show that any linearly ordered family of supports has an upper bound in the ordering considered. We do this by arguing that the intersection of a linearly ordered family of supports is again a support. Let (A_α) be such a family. Since the A_α's are all compact, and since the intersection of any finite set of A_α's contains some A_α; hence, is non-empty, it follows that $\bigcap_\alpha A_\alpha$ is non-empty. Since $\bigcap_\alpha A_\alpha$ is the intersection of a family of closed sets, it is closed; hence, compact. Finally, if $\xi_1, \xi_2 \in K$, and if $(\frac{1}{2})\xi_1 + (\frac{1}{2})\xi_2 \in \bigcap_\alpha A_\alpha$, then $(\frac{1}{2})\xi_1 + (\frac{1}{2})\xi_2 \in A_\alpha$ for all α, so ξ_1 and ξ_2 are in A_α for all α, so ξ_1 and ξ_2 are in $\bigcap_\alpha A_\alpha$. Thus, $\bigcap_\alpha A_\alpha$ is a support; it is evidently an upper bound for $\{A_\alpha\}$ in the ordering we have considered. Thus, applying Zorn's Lemma, every support contains a support which contains no strictly smaller support. Let B be such a minimal support. We want to show that B consists of a single point. Suppose not; then there is a continuous linear functional ϕ which is not constant on B (i.e., take two points of B and apply the geometric form of the Hahn-Banach Theorem to get a functional which separates them). Let $\lambda = \sup\{\phi(\xi) : \xi \in B\}$ and let $B' = \{\xi \in B : \phi(\xi) = \lambda\}$. Then $B' \subsetneq B$, and it is easy to check that B' is a support of K. This contradicts the assumed minimality of B and hence proves the lemma.

II Integral Representations on Compact Convex Sets

A Introduction

We will be concerned, in this chapter, with the problem of representing a general point of a compact convex set [illegible]

general point of a compact convex set is a convex combination of extremal points
for orientation, let us consider convex polygons in the plane. It is

then easy to see that the extremal points are just the corners of the polygon, and that, if $\xi_1, \ldots, \xi_n$ are the corners, then any point of the polygon may be written $\xi = \alpha_1\xi_1 + \cdots + \alpha_n\xi_n$, with $\alpha_1, \ldots, \alpha_n \geqq 0$ and $\sum_{i=1}^{n} \alpha_i = 1$. Indeed, since any point of a convex polygon is contained in at least one triangle with vertices at the corners of the polygon, we may choose the α's so that only three of them are non-zero. Finally, if the polygon is a triangle, then the α's are unique, and if the polygon is not a triangle, then at least some of the points of the polygon have more than one representation as convex combinations of corners.

The above remarks generalize easily to compact convex sets in finite dimensional spaces. A classical theorem, due to Minkowski, states that if X is a compact convex set in a p-dimensional vector space, and if $\xi \in X$, then ξ may be written as a convex combination of some set of $p + 1$ extremal points of X.

The generalization to infinitely many dimensions is not quite so simple. We are at least assured by the Krein-Milman Theorem that any compact convex set X in a locally convex topological vector space has enough extremal points so that every points in X can be approximated arbitrarily closely by convex combinations of extremal points. On the other hand, we should not be too surprised if, in passing to the infinite dimensional case, we had to consider integrals of extremal points rather than ordinary convex combinations. The concept needed is that of the *resultant*, or *barycenter*, of a probability measure (measure with total mass one) on a compact convex set X. The resultant of μ, denoted by $r(\mu)$, is just the integral $r(\mu) = \int \xi \, d\mu(\xi)$, where the integral is to be understood in the weak sense, i.e., $\int \xi \, d\mu(\xi)$ is an element of E such that $\phi\left(\int \xi \, d\mu(\xi)\right) = \int \phi(\xi) \, d\mu(\xi)$ for all continuous linear functionals ϕ on E. While it is easy to see (using the Hahn-Banach Theorem) that this condition uniquely specifies $\int \xi \, d\mu(\xi)$ if the integral exists, the existence is less obvious. We will prove it shortly. The equation $\xi = \sum_{i=1}^{n} \alpha_i\xi_i$ can be transcribed to $\xi = r\left(\sum_i \alpha_i \delta_{\xi_i}\right)$, where δ_{ξ_i} is the unit point-mass at ξ_i, i.e., the Dirac measure at ξ_i. We now want to investigate generalizations of Minkowski's Theorem asserting that every point of a compact convex set X may be represented as the resultant of a probability measure concentrated on the set of extremal points of X. From another point of view, the results we will discuss are a more precise version of the Krein-Milman Theorem: The Krein-Milman Theorem asserts that every point of X may be *approximated by* resultants of measures with finite support concentrated on the set of extremal points of X; the results we will discuss say that every point of X is *equal to* the resultant of a measure (not necessarily with finite support) concentrated on the set of extremal points of X.

The prototype of such results in the following:

THEOREM (Choquet) *Let X be a compact convex set in a Hausdorff locally convex topological vector space. Let ε denote the set of extremal points of X. Assume that X is metrizable, i.e., that its topology has a countable base. Then:*

i) *ε is a Borel subset of X and*
ii) *Every point ξ of X is the resultant of a Borel probability measure μ on X which is concentrated on ε in the sense that $\mu(X \setminus \varepsilon) = 0$.*

This theorem has been refined in two directions. In the first place, the requirement that X be metrizable can almost be eliminated. The difficulty in doing this lies in the fact that the set of extremal points of X may not be a Borel set, i.e., may be pathological from the point of view of measure theory, and the sense in which the measure μ should be concentrated of the set of extremal points is therefore somewhat complicated. In the second place, an algebraic condition on X is given which is necessary and sufficient for every point of X to be the resultant of *a unique* measure concentrated on the set of extremal points of X.

The following notational conventions will be used throughout this chapter.

1) All vector spaces will be vector spaces over the real numbers, and all numerical functions will be real-valued.

2) X will denote a compact convex set in a locally convex topological vector space E.

3) ε will denote the set of extremal points of X.

4) $C(X)$ will denote the set of continuous, real valued functions on X.

5) S will denote the set of continuous *convex* functions on X; $S = \{f \in C(X) : f(\alpha\xi + (1-\alpha)\eta) \leqq \alpha f(\xi) + (1-\alpha) f(\eta)$ for all $\xi, \eta \in X, 0 \leqq \alpha \leqq 1\}$.

6) $A = S \cap (-S)$ will denote the set of continuous affine functions on X. If ϕ is a continuous linear functional on E, and if α is a real number, then $\phi(\cdot) + \alpha$ is an element of A (but not every element of A is necessarily of this form).

7) M^+ will denote the set of positive Borel measures on X, and M_1 the set of Borel probability measures on X. M^+ is contained in the dual of $C(X)$; we equip it with the weak-* topology, thus making M_1 compact.

We now have to prove a few preliminary results.

PROPOSITION II. A.1 *Let $\mu \in M^+$. Then there exists exactly one element $r(\mu)$ of E such that*

$$\phi(r(\mu)) = \int \phi(\xi)\, d\mu(\xi)$$

for all continuous linear functionals ϕ on E. If $\mu \in M_1$, $r(\mu) \in X$.

Proof We first prove uniqueness: If

$$\phi(\xi_1) = \int \phi(\xi)\, d\mu(\xi) \quad \text{and} \quad \phi(\xi_2) = \int \phi(\xi)\, d\mu(\xi)$$

for all continuous linear functionals ϕ on E, then

$$\phi(\xi_1 - \xi_2) = 0$$

for all continuous linear functionals ϕ on E, so $\xi_1 - \xi_2 = 0$. Also, if $\mu \in M_1$, and if $\phi(\xi) \leqq \alpha$ for all $\xi \in X$, then

$$\int \phi(\xi)\, d\mu(\xi) \leqq \alpha;$$

by the geometric form of the Hahn-Banach Theorem, this implies $r(\mu) \in X$. It remains to prove the existence of $r(\mu)$. It suffices to consider $\mu \in M_1$. If μ is a measure with finite support $\left(\mu = \sum_{i=1}^{n} \alpha_i \delta_{\xi_i}\right)$, then $r(\mu)$ exists and is equal to $\sum_i \alpha_i \xi_i$. Now let μ be a general element of M_1. Then there exists a net $\mu^{(\alpha)}$ of measures with finite support converging in the weak-* topology to μ. (This is just the approximability of integrals of continuous functions by Riemann sums.) Each $r(\mu^{(\alpha)})$ exists and belongs to X. Since X is compact, we can suppose (passing to a subnet if necessary) that the net $r(\mu^{(\alpha)})$ converges. Then for all continuous linear functionals ϕ on E

$$\phi\left(\lim_\alpha r(\mu^{(\alpha)})\right) = \lim_\alpha \phi(r(\mu^{(\alpha)})) = \lim_\alpha \int \phi(\xi)\, d\mu^{(\alpha)}(\xi)$$
$$= \int \phi(\xi)\, d\mu(\xi) \text{ (by the definition of weak-*convergence),}$$

so $\phi\left(\lim_\alpha r(\mu^{(\alpha)})\right) = \int \phi(\xi)\, d\mu(\xi)$, so we can take $\lim_\alpha r(\mu^{(\alpha)}) = r(\mu)$.

PROPOSITION II. A. 2 *$S - S$ is dense in $C(X)$, i.e., any continuous function on X may be approximated uniformly by differences of continuous convex functions.*

Proof By the Hahn-Banach Theorem, the continuous affine functions A on X separate points. Therefore, by the Stone-Weierstrass Theorem, polynomials in elements of A are dense in $C(X)$. Thus, it will suffice to prove that, if $P(Z_1, \ldots, Z_n)$ is a polynomial in n variables, and if $f_1, \ldots, f_n \in A$, then $\xi \mapsto P(f_1(\xi), \ldots, f_n(\xi))$ is the difference of two convex functions. Let $Y = \{(f_1(\xi), \ldots, f_n(\xi)) : \xi \in X\} \subset \mathbb{R}^n$. Then Y is convex and compact. If we can show

$$P(Z_1, \ldots, Z_n) = P_1(Z_1, \ldots, Z_n) - P_2(Z_1, \ldots, Z_n),$$

where P_1 and P_2 are polynomials convex on Y, we will be done. Since a polynomial $Q(Z_1, \ldots Z_n)$ is convex on Y if the matrix $\frac{\partial^2 Q}{\partial Z_i \partial Z_j}$

is positive semi-definite on Y, we get a decomposition of the desired form by writing

$$P(Z_1, \ldots, Z_n) = (P(Z_1, \ldots, Z_n) + \lambda(Z_1^2 + \cdots + Z_n^2)) - \lambda(Z_1^2 + \cdots + Z_n^2),$$

with λ sufficiently large.

PROPOSITION II. A. 3 *Any $\mu \in M_1$ can be approximated arbitrarily well in the weak-* topology by measures with finite support having the same resultant as μ.*

Proof Let $f_1, \ldots, f_m \in C(X)$ and let $\varepsilon > 0$. We have to find $\lambda_1, \ldots, \lambda_m \geqq 0$, with $\sum_{i=1}^{n} \lambda_i = 1$, and $\xi_1, \ldots, \xi_n \in X$, such that $\sum_{i=1}^{n} \lambda_i \xi_i = r(\mu)$ and

$$\left| \mu(f_j) - \sum_{i=1}^{n} \lambda_i f_j(\xi_i) \right| \leqq \varepsilon \quad \text{for} \quad 1 \leqq j \leqq m.$$

Since X is compact, there exists a finite set $U_1, \ldots, U_n$ of open convex sets in X, with $\bigcup_{i=1}^{n} U_i = X$, such that each f_j varies by less than ε on each U_i. Next, we can write $\mu = \sum_{i=1}^{n} \lambda_i \mu_i$, where each $\mu_i \in M_1$ and has support in U_i. (For example, if ϕ_i is the characteristic function of U_i, we can take $\lambda_i \mu_i = \dfrac{\phi_i \mu}{\sum_{i'} \phi_{i'}}$.)

Let $\xi_i = r(\mu_i)$; then evidently $r(\mu) = \sum_i \lambda_i \xi_i$. Since U_i is compact and convex, $r(\mu_i) \in U_i$, so

$$\left| \int d\mu_i f_j - f_j(\xi_i) \right| \leqq \varepsilon \quad \text{for all } i, j.$$

Therefore, for $1 \leqq j \leqq m$,

$$\left| \int d\mu f_j - \sum_{i=1}^{n} \lambda_i f_j(\xi_i) \right| \leqq \sum_{i=1}^{n} \lambda_i \left| \int d\mu_i f_j - f_j(\xi_i) \right| \leqq \varepsilon.$$

B The Existence Theorem

We will first outline the strategy for proving that every $\xi \in X$ is the resultant of a measure concentrated on the set of extremal points. Given ξ, we want to find a measure with resultant ξ which is pushed out as much as possible to the "corners" of X. One way to tell how much a measure μ is pushed toward the "corners" is to evaluate $\mu(f)$ for convex f; the larger $\mu(f)$ is for fixed convex f and fixed $r(\mu)$, the more we expect μ to be concentrated near the extremal points of X.

With this in mind, we define an order on M^+ by $\mu \succ \nu$ if $\mu(f) \geqq \nu(f)$ for

all continuous convex f. It is evident that this defines a preorder, i.e. that

$\succ$ is reflexive and transitive. The relation $\succ$ is also anti-symmetric and hence an order; if $\mu \succ \nu$ and $\nu \succ \mu$, then $\mu(f) = \nu(f)$ for all $f \in S$, therefore, for all $f \in S - S$; therefore, since $S - S$ is dense in $C(X)$, for all $f \in C(X)$.

Now let $\mu \succ \nu$ and let f be a continuous affine function. Since both f and $-f$ are convex, we have

$$\mu(f) \geqq \nu(f) \quad \text{and} \quad \mu(-f) \geqq \nu(-f), \text{ so } \quad \mu(f) = \nu(f).$$

In particular, $\|\mu\| = \mu(1) = \nu(1) = \|\nu\|$, and $r(\mu) = r(\nu)$, so two measures which are comparable in the sense of $\succ$ have the same total mass and the same resultant. It is thus easy to see that $\mu \succ \delta_\xi$ if and only if μ is a probability measure and $r(\mu) = \xi$.

We now have good candidates for measures with resultant ξ which are concentrated on the set of extremal points of X; they are the measures μ satisfying $\mu \succ \delta_\xi$ which are maximal in the sense of $\succ$. An easy argument using Zorn's Lemma shows that such maximal measures exist. It remains, however, to determine in what sense maximal measures are concentrated on the set of extremal points. For this purpose, we consider upper envelopes of functions: If f is a bounded function on X, we define the upper envelope of f, denoted $\hat{f}$, to be the smallest *concave* function which is everywhere greater than or equal to f or, more precisely, as the infimum of all continuous concave functions $\geqq f$. (Recall that the infimum of any family of concave functions is concave.)

Unfortunately, $\hat{f}$ need not be continuous even if f is. If we forget this fact for the moment, but remember that $\hat{f}$ is concave, we see that, to make μ maximal, we want to make $\mu(\hat{f})$ as small as possible (keeping $r(\mu)$ fixed). On the other hand, $\hat{f} \geqq f$, so $\mu(\hat{f}) \geqq \mu(f)$ for all $\mu \in M^+$. Thus, we might guess that, in order to have μ maximal we should have $\mu(f) = \mu(\hat{f})$ and it turns out, in fact, that μ is maximal if and only if $\mu(f) = \mu(\hat{f})$ for all continuous convex functions f. Since $\hat{f} \geqq f$, this means that μ is maximal if and only if

$$\mu(\{\xi : \hat{f}(\xi) > f(\xi)\}) = 0$$

for all continuous convex f, i.e., if and only if μ is carried by $\{\xi : \hat{f}(\xi) = f(\xi)\}$ for all continuous convex f. Finally, it turns out that the set of extremal points of X is precisely the set of points on which every continuous convex function is equal to its upper envelope, i.e.,

$$\mathcal{E} = \bigcap_{f \in S} \{\xi : \hat{f}(\xi) = f(\xi)\}.$$

The intersection is, in general, over a non-denumerable set of f's, so we cannot conclude that $\mu(X \setminus \mathcal{E}) = 0$ for μ maximal, or even that $\mu(X \setminus \mathcal{E})$ is defined. Thus, it is reasonable to say that a maximal measure is concentrated on $\mathcal{E}$, but the sense in which the measure is concentrated on $\mathcal{E}$ is a bit subtle. If X is metrizable all these difficulties diappear: There is a *single* continuous convex function f such that $\mathcal{E} = \{\xi : \hat{f}(\xi) = f(\xi)\}$

so in this case every maximal measure is concentrated on e in the sense that $\mu(X \setminus e) = 0$. Combining all these results gives the theorem of Choquet quoted in the introduction.

To give the proof of the existence theorem, then, we must prove four things:

a) Every measure on X is majorized in the sense of $\succ$ by a maximal measure.

b) A measure μ is maximal if and only if $\mu(f) = \mu(\hat{f})$ for all continuous convex f.

c) $e = \bigcap_{f \in S} \{\xi : \hat{f}(\xi) = f(\xi)\}$.

d) If X is metrizable, there exists $f \in S$ such that

$$e = \{\xi : \hat{f}(\xi) = f(\xi)\}.$$

Points a), c), and d) are more or less routine; the subtle part of the argument comes in the proof of b).

PROPOSITION II. B. 1. *Every $\mu \in M^+$ is majorized, in the sense of the ordering $\succ$, by a maximal measure.*

Proof By Zorn's Lemma, it suffices to show that any family $(\mu_i)_{i \in I} \subset M^+$ which is totally ordered by $\succ$ admits an upper bound. For $f \in S$, $(\mu_i(f))$ is monotonic, by the definition of $\succ$, and bounded since $\|\mu_i\|$ is independent of i. Therefore, $\mu_i(f)$ converges to something for all $f \in S$; hence, for all $f \in S - S$; hence, since $S - S$ is dense in $C(X)$, for all $f \in C(X)$. The limit defines a positive measure (i.e., a positive linear functional on $C(X)$) which evidently majorizes all the μ_i's.

As we indicated above, we define the upper envelope $\hat{f}$ of a bounded real-valued function f by $\hat{f}(\xi) = \inf \{g(\xi) : g$ concave and continuous, $g \geqq f$ everywhere$\}$. In other words, we define $\hat{f}$ to be the *infimum* of the set of all concave continuous functions $\geqq f$. Since the infimum of an arbitrary family of concave functions is concave, $\hat{f}$ is concave. Although $\hat{f}$ need not be continuous, it is the *infimum* of a family of continuous functions and is therefore upper semi-continuous. The mapping $f \mapsto \hat{f}$ is:

i) increasing (If $f_1 \geqq f_2$, then $\hat{f}_1 \geqq \hat{f}_2$) and

ii) sublinear ($\widehat{\lambda f} = \lambda \cdot \hat{f}$ if $\lambda \geqq 0$; $\widehat{f_1 + f_2} \leq \hat{f}_1 + \hat{f}_2$)

For any positive measure μ, we define $\hat{\mu}(f) = \mu(\hat{f})$. Then $\hat{\mu}$ is an increasing sublinear functional on $C(X)$, and $\hat{\mu} \geq \mu$, i.e., $\hat{\mu}(f) \geq \mu(f)$ for all $f \in C(X)$. Furthermore, the functional $\hat{\mu}$ is continuous. First to prove that $\hat{\mu}$ is continuous at zero note that

where $\|g\| = \sup_{\xi \in X} |g(\xi)|$. This is true since the constant function $\|g\|$ is concave and $\geqq g$. Thus,

$$|\hat{\mu}(g)| = |\mu(\hat{g})| \leqq \|g\| \mu(1).$$

The continuity of $\hat{\mu}$ at a general point of $C(X)$ follows easily from continuity at zero and subadditivity: For $f, g \in C(X)$,

$$\hat{\mu}(f + g) \leqq \hat{\mu}(f) + \hat{\mu}(g),$$

$$\hat{\mu}(f) \leqq \hat{\mu}(f + g) + \hat{\mu}(-g),$$

so

$$-\hat{\mu}(-g) \leqq \hat{\mu}(f + g) - \hat{\mu}(f) \leqq \hat{\mu}(g),$$

so

$$|\hat{\mu}(f + g) - \hat{\mu}(f)| \leqq \|g\| \mu(1).$$

Recall now that we are trying to show that μ is maximal if and only if $\mu(f) = \mu(\hat{f})$ for all convex continuous f, i.e., if and only if $\hat{\mu} = \mu$ on S.

PROPOSITION II. B. 2 *If $\mu \in M^+$, the following are equivalent*

i) $\hat{\mu} = \mu$ on S,

ii) $\hat{\mu} = \mu$ on $C(X)$,

iii) $\hat{\mu}$ is linear on $C(X)$.

Proof It is immediate that ii) implies iii) and that ii) implies i). To see that iii) implies ii), recall that $\hat{\mu}(f) \geqq \mu(f)$ for all f in $C(X)$. But we can equally well apply this inequality with f replaced by $-f$; then, since both μ and $\hat{\mu}$ are linear, $-\hat{\mu}(f) = \hat{\mu}(-f) \geqq \mu(-f) = -\mu(f)$.

So $\hat{\mu}(f) \leqq \mu(f)$, so $\hat{\mu}(f) = \mu(f)$ for all f in $C(X)$. We have therefore only to prove that i) implies ii). Note first that, if $g \in S$, $\widehat{-g} = -g$ (since $-g$ is already concave), so $\hat{\mu}(-g) = \mu(-g)$.

Now assume i), and let $f, g \in S$. By subadditivity.

$$\hat{\mu}(f) \leqq \hat{\mu}(f - g) + \hat{\mu}(+g) \quad \text{and} \quad \hat{\mu}(f - g) \leqq \hat{\mu}(f) + \hat{\mu}(-g),$$

so

$$|\hat{\mu}(f) - \hat{\mu}(g) \leqq \hat{\mu}(f - g) \leqq \hat{\mu}(f) + \hat{\mu}(-g).$$

Using

$$\hat{\mu}(f) = \mu(f); \quad \hat{\mu}(g) = \mu(g); \quad \hat{\mu}(-g) = \mu(-g);$$

and the linearity of μ, we get

$$\mu(f - g) \leqq \hat{\mu}(f - g) \leqq \mu(f - g),$$

so $\hat{\mu} = \mu$ on $S - S$. Since $S - S$ is dense in $C(X)$ and since μ and $\hat{\mu}$ are both continuous, $\hat{\mu} = \mu$ on all of $C(X)$.

The key remark is the following lemma, which will allow us to use the

Hahn-Banach theorem effectively

LEMMA II. B. 3 *Let $\mu \in M^+$ and let ν be a linear functional on $C(X)$. Then the following are equivalent:*

i) $\nu(f) \leqq \hat{\mu}(f)$ for all $f \in C(X)$.

ii) $\nu \in M^+$ and $\nu \succ \mu$.

Proof Let us assume i) and prove ii). Note that, since $\hat{f} = f$ for f concave, $\nu(f) \leqq \hat{\mu}(f)$ means $\nu(f) \leqq \mu(f)$ for f concave, so $\nu(f) \leqq \mu(f)$ for f convex, so $\nu \succ \mu$. It still has to by verified that, if $f \geqq 0$, $\nu(f) \geqq 0$. It is simpler to verify that $f \leqq 0$ implies $\nu(f) \leqq 0$; indeed, this is immediate since $f \leqq 0$ implies $\hat{f} \leqq 0$ (the constant function 0 is concave, continuous, and $\geqq f$), so $\hat{\mu}(f) = \mu(\hat{f}) \leqq 0$, so $\nu(f) \leqq \hat{\mu}(f) \leqq 0$.

Now assume ii). The fact we want to exploit is that $\nu(g) \leqq \mu(g)$ for all concave continuous g. We will use this fact by remarking that $\hat{f} = \inf\{g : g$ concave, continuous, and $\geqq f\}$. The set of concave, continuous functions $\geqq f$ is a decreasing net of continuous functions converging pointwise to $\hat{f}$; hence, by the generalized monotone convergence theorem for semi-continuous functions (Theorem I. B. 3);

$$\mu(\hat{f}) = \inf\{\mu(g) : g \text{ concave, continuous, and } \geqq f\},$$

$$\nu(\hat{f}) = \inf\{\nu(g) : g \text{ concave, continuous, and } \geqq f\}, \text{ so}$$

$$\nu(f) \leqq \nu(\hat{f}) \leqq \mu(\hat{f}) = \hat{\mu}(f) \text{ for all } f \text{ in } C(X).$$

We are now ready to prove:

THEOREM II. B. 4 *Let $\mu \in M^+$. Then μ is maximal if and only if $\mu(f) = \mu(\hat{f})$ for all $f \in S$.*

Proof Suppose first that $\mu(f) = \mu(\hat{f})$ for all $f \in S$. Then $\hat{\mu} = \mu$ on $C(X)$ by Proposition II. B. 2. If $\nu \succ \mu$, then $\nu(f) \leqq \hat{\mu}(f) = \mu(f)$ for all f in $C(X)$; since μ and ν are both linear, this implies $\nu(f) = \mu(f)$ for all f in $C(X)$, i.e., $\mu = \nu$. Therefore, μ is maximal. (It is majorized only by itself.)

Now assume $\mu(g) \neq \mu(\hat{g})$ for some $g \in S$; we will prove that μ is not maximal by producing $\nu \succ \mu$ such that $\nu(g) = \mu(\hat{g}) > \mu(g)$. Thus, we have to construct a measure, i.e., a positive linear functional on $C(X)$, so we will use the Hahn-Banach Theorem. By Lemma II. B. 3, we have only to construct a linear functional ν such that $\nu(f) \leqq \hat{\mu}(f)$ for all f. We start by defining $\nu(\lambda g) = \lambda\hat{\mu}(g)$; this defines ν on the one-dimensional subspace generated by g, and $\nu(\lambda g) \leqq \hat{\mu}(\lambda g)$ since $\hat{\mu}(\lambda g) \geqq \lambda\hat{\mu}(g)$ by the sublinearity of $\hat{\mu}$. Now, applying the generalization of the analytic form of the Hahn-Banach Theorem (Theorem I. C. 6), we extend ν to all of $C(X)$, preserving the relation $\nu(f) \leqq \hat{\mu}(f)$ for all f. This is possible since $\hat{\mu}$ is sublinear. We are now finished; since $\nu(f) \leqq \hat{\mu}(f)$ for all f, $\nu \succ \mu$ by Lemma II. B. 3, but $\nu \neq \mu$

since $\nu(g) = \hat{\mu}(g) \neq \mu(g)$

Now let f be a continuous convex function on X. We define $B(f) = \{\xi \in X : \hat{f}(\xi) = f(\xi)\}$, and any set of the form $B(f)$, $f \in S$, is called a *boundary set* of X. Since $\hat{f}$ is upper semi-continuous and f is continuous, $\hat{f} - f$ is upper semi-continuous so, for all α,

$$\{\xi \in X : \hat{f}(\xi) - f(\xi) < \alpha\}$$

is open. Hence

$$B(f) = \bigcap_{n=1}^{\infty} \{\xi \in X : \hat{f}(\xi) - f(\xi) < 1/n\}$$

is a countable intersection of open sets and is therefore a Borel set. Since $\hat{f} - f \geqq 0$, we have $\mu(\hat{f}) = \mu(f)$ if and only if

$$\mu(\{\xi \in X : \hat{f}(\xi) > f(\xi)\}) = 0.$$

If Y is a Borel subset of X, and if $\mu \in M^+$, we will say that μ is *carried by* Y if $\mu(X \setminus Y) = 0$. Thus we have:

PROPOSITION II. B. 5 *A measure $\mu \in M^+$ is maximal in the sense of the order $\succ$ if and only if it is carried by every boundary set of X.*

COROLLARY II. B. 6 *Every point $\xi \in X$ is the resultant of a probability measure μ on X carried by every boundary set of X.*

COROLLARY II. B. 7 *Any sum, or, more generally, any integral, of maximal measures is maximal. The greatest lower bound or least upper bound of a finite family of maximal measures is maximal.*

We now must investigate the relation between boundary sets and extremal points.

PROPOSITION II. B. 8 *The set of extremal points of X is the intersection of all the boundary sets of X.*

Proof We have to prove:

i) Any extremal point of X belongs to every boundary set.

ii) Every non-extremal point of X lies outside some boundary set.

i) Let ξ be an extremal point of X. Then the only probability measure with finite support having ξ as resultant is δ_ξ. But any probability measure having ξ as resultant can be approximated in the weak-* topology by measures with finite support having ξ as resultant (Proposition II. A. 3), so the only probability measure having ξ as resultant is δ_ξ. In other words, if $\mu \succ \delta_\xi$, then $\mu = \delta_\xi$. Hence, the measure δ_ξ is maximal, i.e., is carried by every boundary set of X, so ξ belongs to every boundary set of X.

ii) Let ξ be a non-extremal point of X. Then

$\xi = \frac{1}{2}\xi_1 + \frac{1}{2}\xi_2$, with $\xi_1 \neq \xi_2$

Choose a continuous affine f such that $f(\xi_1) \neq f(\xi_2)$. Then f^2 is convex, and

$$\begin{aligned}(f((1/2)\,\xi_1 + (1/2)\,\xi_2))^2 &= (1/4)\,(f(\xi_1))^2 + (1/4)\,(f(\xi_2))^2 + (1/2)f(\xi_1)f(\xi_2)\\ &= (1/2)\,(f(\xi_1))^2 + (1/2)\,(f(\xi_2))^2 - (1/4)\,(f(\xi_1) - f(\xi_2))^2.\end{aligned}$$

Since $(\widehat{f^2})$ is concave,

$$(\widehat{f^2})(\xi) = (\widehat{f^2})((1/2)\,\xi_1 + (1/2)\,\xi_2) \geqq (1/2)\,(\widehat{f^2})(\xi_1) + (1/2)\,(\widehat{f^2})(\xi_2)$$

$$\geqq (1/2)f^2(\xi_1) + (1/2)f^2(\xi_2) > f^2(\xi);$$

$$\text{so} \quad \widehat{f^2}(\xi) > f^2(\xi), \quad \text{so} \quad \xi \notin B(f^2).$$

We can prove a somewhat more precise result if X is metrizable.

PROPOSITION II. B. 9 *If X is metrizable, then ε is a boundary set of X.*

Proof We will use only the following consequence of the metrizability of X: There exists a sequence f_n of continuous affine functions separating the points of X (This means that, if $\xi_1 \neq \xi_2$, then $f_n(\xi_1) \neq f_n(\xi_2)$ for some n). It is left as an exercise in general topology to verify that the existence of such a sequence is equivalent to the metrizability of X. By scaling the f_n's, we can assume $|f_n| \leqq 1/n$ for each n. Since each f_n is affine, f_n^2 is convex for each n. Let $f = \sum_{n=1}^{\infty} f_n^2$;

then f is continuous and convex. We claim that f is strictly convex, i.e., that if

$$\xi_1 \neq \xi_2, \quad \text{then} \quad f((1/2)\,\xi_1 + (1/2)\,\xi_2) < (1/2)f(\xi_1) + (1/2)f(\xi_2).$$

Since

$$\hat{f}((1/2)\,\xi_1 + (1/2)\,\xi_2) \geqq (1/2)\hat{f}(\xi_1) + (1/2)\hat{f}(\xi_2) \geqq (1/2)f(\xi_1) + (1/2)f(\xi_2),$$

this will imply

$$\hat{f}((1/2)\,\xi_1 + (1/2)\,\xi_2) > f((1/2)\,\xi_1 + (1/2)\,\xi_2)$$

for all $\xi_1 \neq \xi_2$, i.e., that $\hat{f}(\xi) > f(\xi)$ if ξ is not an extremal point of X. Thus, $B(f) \subset \varepsilon$, and since we know that, in general, $B(f) \supset \varepsilon$ for all continuous convex f, we must have $B(f) = \varepsilon$.

It remains to show that f is strictly convex. Let $\xi_1 \neq \xi_2$, and choose n so that $f_n(\xi_1) \neq f_n(\xi_2)$. By the calculation in the proof of the preceding proposition,

$$f_n^2((1/2)\,\xi_1 + (1/2)\,\xi_2) < (1/2)f_n^2(\xi_1) + (1/2)f_n^2(\xi_2).$$

Since $f = f_n^2 + \sum_{j \neq n} f_j^2$ and since $\sum_{j \neq n} f_j^2$ is convex, f is strictly convex.

Putting together the preceding results, we get the theorem quoted in the introduction:

THEOREM II. B. 10 *Let X be a metrizable compact convex set in a Hausdorff locally convex topological vector space. Then*

i) *ε is a Borel subset of X and*

ii) *Each $\xi \in X$ is the resultant of a probability measure on X carried by ε.*

If X is not metrizable, the situation is a little more cloudy. We have shown that every point of X is the resultant of a maximal measure, and that a measure is maximal if and only if it is carried by every boundary set of X. Since the intersection of all boundary sets is the set of extremal points of X, a maximal measure is in some weak sense concentrated on the set of extremal points of X. The sense in which a maximal measure is concentrated on the set of extremal points may be clarified a bit; we quote by way of illustration the following result, which we will not prove. (It is Lemma 18 of Choquet and Mayer[1].)

PROPOSITION II. B. 11. *Let $K \subset X$ be a countable union of compact sets and contain the set of extremal points of X. Then every maximal measure on X is carried by K.*

C The Uniqueness Theorem

We now turn to an investigation of the conditions under which each $\xi \in X$ is the resultant of a unique maximal measure. We have first to prove several preliminary results. The first group of results concerns the upper envelope of a function f and provides some new ways to compute it.

LEMMA II. C. 1 *Let f be a bounded real-valued function on X. Then*

$$\Gamma = \{(\xi, t) \in E \times \mathbb{R}: \xi \in X, t \leqq \hat{f}(\xi)\}$$

is the closed convex hull of

$$\{(\xi, t) \in E \times \mathbb{R}: \xi \in X; t \leqq f(\xi)\}.$$

(i.e., the region below the graph of $\hat{f}$ is the closed convex hull of the region below the graph of f.)

Proof It follows easily from the concavity of $\hat{f}$ that Γ is convex and from the upper semi-continuity of $\hat{f}$ that Γ is closed. Thus, Γ is a closed convex set containing

$$\{(\xi, t) \in E \times \mathbb{R}: \xi \in X; t \leqq f(\xi)\}.$$

If we let Γ' denote the closed convex hull of the latter set, then $\Gamma' \subset \Gamma$, and what we have to prove is that $\Gamma' = \Gamma$. Thus, assume there is a $(\xi_0, t_0) \in \Gamma \setminus \Gamma'$

~~By the geometric form of the Hahn-Banach theorem...~~

By the geometric form of Hahn-Banach theorem there is a

hyperplane in $E \times \mathbb{R}$ strictly separating (ξ_0, t_0) from Γ. In other words, there is a linear functional ϕ on $E \times \mathbb{R}$ and a constant c such that

$$\phi(\xi_0, t_0) < c; \quad \phi(\xi, t) \geq c \quad \text{for} \quad t \leq f(\xi).$$

We claim that the hyperplane $\phi(\xi, t) = c$ is the graph of a continuous affine functional on X, which is greater than f everywhere, but less than t_0 at ξ_0, thus contradicting the asumption that $t_0 \leq \hat{f}(\xi_0)$. To see this, note that $\phi(\xi, t)$ must have the form $\phi(\xi) - \alpha t$, where ϕ is a continuous linear functional on E and $\alpha \in \mathbb{R}$. Since

$$\phi(\xi_0) - \alpha t_0 < c; \quad \phi(\xi_0) - \alpha t > c \quad \text{for} \quad t \leq f(\xi_0),$$

we must have $\alpha > 0$. Letting

$$\psi(\xi) = \frac{\phi(\xi) - c}{\alpha}$$

we have $\psi(\xi_0) < t_0$ but $\psi(\xi) > t$ for all $t \leq f(\xi)$. Thus, ψ is a continuous affine function everywhere greater than f; it is therefore impossible to have $\psi(\xi_0) < t_0 \leq \hat{f}(\xi_0)$, so we have a contradiction and $\Gamma' = \Gamma$ as desired.

In the course of the proof, we have also proved the following:

LEMMA II. C. 2 *If h is any concave upper semi-continuous function on X, then $h = \inf \{g \in A : g \geq h\}$.*

LEMMA II. C. 3 *Let $f \in C(X)$ and let $\xi \in X$. Then*

$$\hat{f}(\xi) = \sup \{\mu(f) : \mu \in M_1; r(\mu) = \xi\}$$

$$= \sup \{\mu(f) : \mu \in M_1; r(\mu) = \xi, \mu \text{ has finite support}\}.$$

Proof Since $\hat{f}$ is concave,

$$\hat{f}(\xi) \geq \sum_{i=1}^{n} \lambda_i \hat{f}(\xi_i) \geq \sum_{i=1}^{n} \lambda_i f(\xi_i) \quad \text{if} \quad \sum_{i=1}^{n} \lambda_i \xi_i = \xi.$$

Hence, $\hat{f}(\xi)$ is not smaller than the second supremum. On the other hand, Lemma II. C. 1, translated into the language of measures with finite supports, says that there exists a net of measures with finite supports

$$\mu^{(\alpha)} = \sum_{i=1}^{n(\alpha)} \lambda_i^{(\alpha)} \delta_{\xi_i^{(\alpha)}}$$

with $r(\mu^{(\alpha)}) \to \xi$ and $\liminf \mu^{(\alpha)}(f) \geq \hat{f}(\xi)$. If μ is any limit point of this net, then $r(\mu) = \xi$ and $\mu(f) \geq \hat{f}(\xi)$, so the first supremum is not less than $\hat{f}(\xi)$. But since any measure with resultant ξ may be approximated arbitrarily well in the weak-* topology by measures with finite supports and resultant μ (Proposition II. A. 3) the two suprema are equal and so $\hat{f}(\xi)$ is equal to ~~the~~ their [illegible] value.

PROPOSITION II. C. 4 *Let f be an affine upper semi-continuous function on X, and let $\mu \in M_1$. Then*

$$\int f(\xi)\, d\mu(\xi) = f(r(\mu)).$$

Proof In more picturesque form, we want to show that, if f is affine and upper semi-continuous, then

$$\int f(\xi)\, d\mu(\xi) = f\left(\int \xi\, d\mu(\xi)\right).$$

This is immediate if f is continuous, or if μ has finite support. We get the general case by a two-fold approximation. We first approximate μ by measures with resultant $r(\mu)$ and finite support. If ν is such a measure, and if g is continuous, *concave* and nowhere less than f, we have

$$f(r(\mu)) = \nu(f) \leqq \nu(g) \leqq g(r(\mu)).$$

Since μ is a weak limit of measures with finite support and resultant $r(\mu)$, we have

$$f(r(\mu)) \leqq \mu(g) \leqq g(r(\mu)) \qquad (*)$$

for any continuous concave function $g \geqq f$. Now f is concave and upper semi-continuous, so by Lemma II. C. 2, $f = \inf\{g : g \text{ concave and continuous}, g \geqq f\}$. Thus, $\{g : g \text{ concave and continuous}, g \geqq f\}$ is a decreasing family of continuous functions with infimum f; by the generalized Monotone Convergence Theorem for semi-continuous functions (Theorem I. B. 3),

$$\int f(\xi)\, d\mu(\xi) = \inf\{\mu(g) : g \text{ concave and continuous}; g \geqq f\}.$$

By (•), then,

$$f(r(\mu)) \leqq \int f(\xi)\, d\mu(\xi) \leqq f(r(\mu)),$$

so the proposition is proved.

We are now ready to formulate a condition on X which is equivalent to the statement that every point of X is majorized by a *unique* maximal measure. To do this, it is convenient to suppose that X is the base of a convex cone K. By this we mean that there is a closed hyperplane in E which intersects K in X, which does not pass through 0, and which intersects each ray in K exactly once. We can always arrange this by replacing E by $E \times \mathbf{R}$, identifying X with $X \times \{1\}$, and taking for K the cone generated by $X \times \{1\}$. In general, there may be many ways of realizing X as the base of a cone, but any two cones obtained in this way are linearly isomorphic, since any point of such a cone is uniquely specified as $\lambda\xi$, where λ is a positive real number and $\xi \in X$. Thus, the algebraic properties of the cone K depend only on the algebraic properties of X, and not on how X is imbedded in a topological vector space. It is convenient to extend every numerical function on X to a positively homogeneous function on the cone K denoted by $\hat{f}$.

same symbol. Thus, for example, the condition that f be convex (on X) translates into the condition

$$f(\xi + \eta) \leqq f(\xi) + f(\eta) \qquad (\text{on } K).$$

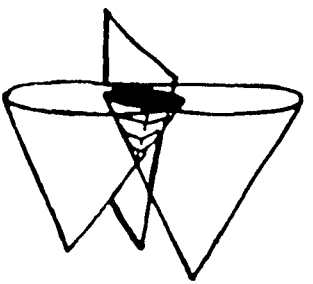
By definition, the *natural ordering* of a cone K in a vector space is the order in which $\xi \geqq \eta$ means $\xi - \eta \in K$. Also, a partially ordered set Y is a *lattice* if, for any pair ξ, η of elements of Y, there exists a least upper bound $\xi \vee \eta$ and a greatest lower bound $\xi \wedge \eta$ for ξ and η in Y. By definition, $\xi \vee \eta$ is an element of Y such that

$$\xi \vee \eta \geqq \xi, \quad \xi \vee \eta \geqq \eta, \quad \text{and} \quad \zeta \geqq \xi \quad \text{and} \quad \zeta \geqq \eta \quad \text{implies} \quad \zeta \geqq \xi \vee \eta.$$

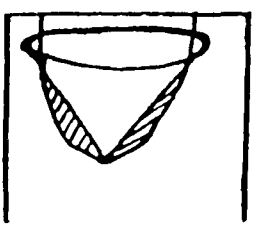
The element $\xi \vee \eta$ is uniquely determined if it exists. The definition of greatest lower bound is obtained by replacing " $\geqq$ " by " $\leqq$ " in the definition of least upper bound.

We now say that a compact convex set X is a *simplex* if it is the base of a cone which is a lattice with its natural ordering.

It does not seem to be altogether evident that this definition reduces to the usual one when X is finite-dimensional. To clarify the situation a little, let us see why a circle is not a simplex. Let ξ and η belong to the cone generated by a circle, and suppose that they are not comparable in the ordering defined by the cone. We will show that no least upper bound of ξ and η exists.



$\zeta \geqq \xi$ and $\zeta \geqq \eta$ if and only if ζ belongs to the intersections of the two cones with vertices ξ and η. A section of the figure by a vertical plane, perpendicular to the plane containing the axes of the cones, and passing through the lowest intersection point, is a hyperbola.



The only possible candidate for the least upper bound of ξ and η is the vertex of this hyperbola. A cone drawn at this vertex, however, does not contain all of the hyperbola, so there are points which are greater than ξ and greater than η, but not greater than the vertex of the hyperbola. Therefore, ξ and η have no least upper bound.

On the other hand, the usual n-dimensional simplex $\left\{(\lambda_1, \ldots, \lambda_{n+1}): \sum_i \lambda_i = 1\right\}$ is a base of the cone $\{(\lambda_1, \ldots, \lambda_{n+1}): \lambda_i \geqq 0 \text{ for all } i\}$ which is certainly a lattice with its natural ordering. The real justification for the definition, however, is contained in the following theorem:

THEOREM II. C. 5 *Let X be a compact convex set. Then the following are equivalent:*

i) *X is a simplex*

ii) *For any $f \in S$, $\hat{f}$ is affine*

iii) *Every $\xi \in X$ is the resultant of a unique maximal measure on X.*

Proof We will show that ii) $\Rightarrow$ iii) $\Rightarrow$ i) $\Rightarrow$ ii). Everything except i) $\Rightarrow$ ii) is easy.

ii) $\Rightarrow$ iii). Let ii) hold, let μ_1, μ_2 be two maximal measures on X both with resultant ξ, and let $f \in S$. Since μ_1 and μ_2 are maximal,

$$\mu_1(f) = \mu_1(\hat{f}); \quad \mu_2(f) = \mu_2(\hat{f}).$$

By hypothesis, $\hat{f}$ is affine, and $\hat{f}$ is always upper semi-continuous so by Proposition II. C. 3

$$\mu_1(\hat{f}) = \hat{f}(r(\mu_1)) = \hat{f}(\xi) = \mu_2(\hat{f}),$$

so

$$\mu_1(f) = \mu_1(\hat{f}) = \mu_2(\hat{f}) = \mu_2(f),$$

so

$$\mu_1 = \mu_2 \text{ on S, so } \mu_1 = \mu_2 \text{ on S-S}$$

so $\mu_1 = \mu_2$ on $C(X)$ (since S-S is dense in $C(X)$ by Proposition II. A. 2). Thus, there is only one maximal measure with resultant ξ. iii) $\Rightarrow$ i). By Corollary II. B. 7, the cone of maximal measures is a lattice; the mapping $\mu \leftrightarrow r(\mu)$ is a linear isomorphism from this cone to the cone generated by X.

i) $\Rightarrow$ ii). Let f be convex. We want to show that $\hat{f}$ is affine. Regarding f as defined and homogeneous on K, we have

$$\hat{f}(\xi + \eta) = \sup\left\{\sum_{i=1}^{n} f(\zeta_i): \ \zeta_i \in K; \ \sum_{i=1}^{n} \zeta_i = \xi + \eta\right\}$$

by Lemma III. C. 3. We will prove shortly that whenever we have $\xi + \eta = \zeta_1 + \cdots + \zeta_n$, with the ζ_i in K, we can decompose

$$\zeta_i = \xi_i + \eta_i, \quad \text{where} \quad \xi_i, \eta_i \in K; \ \sum_j \xi_j = \xi; \ \sum_j \eta_j = \eta.$$

Assuming this for the moment, we have:

$$\hat{f}(\xi + \eta) = \sup\left\{\sum_i f(\zeta_i): \ \zeta_i \in K; \ \sum_i \zeta_i = \xi + \eta\right\}$$

$$= \sup\left\{\sum_i f(\xi_i + \eta_i): \ \xi_i, \eta_i \in K; \ \sum_i \xi_i = \xi; \ \sum_i \eta_i = \eta\right\}.$$

(This is just the decomposition lemma stated above.)

$$\leqq \sup\left\{\sum_i f(\xi_i): \sum_i \xi_i = \xi\right\} + \sup\left\{\sum_i f(\eta_i): \sum_i \eta_i = \eta\right\} = \hat{f}(\xi) + \hat{f}(\eta).$$

(Since f is convex.)
Thus,

$$\hat{f}(\xi + \eta) \leqq \hat{f}(\xi) + \hat{f}(\eta).$$

On the other hand,

$$\hat{f}(\xi + \eta) \geqq \hat{f}(\xi) + \hat{f}(\eta)$$

since $\hat{f}$ is concave. Thus, for $\xi, \eta \in K$,

$$\hat{f}(\xi + \eta) = \hat{f}(\xi) + \hat{f}(\eta),$$

i.e., $\hat{f}$ is affine.

We still have to prove the following lemma.

LEMMA II. C. 6 *Let K be a cone which is a lattice with its natural order, let $\xi, \eta \in K$, and suppose $\xi + \eta = \zeta_1 + \ldots \zeta_n$ where $\zeta_1, \ldots, \zeta_n \in K$. Then there exist $\xi_1, \ldots, \xi_n, \eta_1, \ldots, \eta_n \in K$ such that $\zeta_i = \xi_i + \eta_i$ for $1 \leqq i \leqq n$, and*

$$\xi_1 + \cdots + \xi_n = \xi; \qquad \eta_1 + \cdots + \eta_n = \eta.$$

Proof It is easy to prove by induction that the lemma with $n = 2$ implies the lemma for general n (Write $\xi + \eta = \zeta_1 + (\zeta_2 + \cdots + \zeta_n)$; use the lemma with $n = 2$ to decompose

$$\zeta_1 = \xi_1 + \eta_1; \qquad (\zeta_2 + \cdots + \zeta_n) = \xi' + \eta',$$

where

$$\xi_1 + \xi' = \xi; \qquad \eta_1 + \eta' = \eta.$$

Then iterate.) For $n = 2$, it is trivial to verify that

$$\xi_1 = \zeta_1 \wedge \xi; \qquad \eta_1 = \zeta_1 - \zeta_1 \wedge \xi$$

$$\xi_2 = \xi - \zeta_1 \wedge \xi; \qquad \eta_2 = \zeta_2 - \xi + \zeta_1 \wedge \xi$$

satisfy all the conditions except, possibly, $\eta_2 \in K$. But we can rewrite:

$$\eta_2 = (\zeta_2 - \xi) + (\zeta_1 \wedge \xi)$$

$$= (\zeta_1 + (\zeta_2 - \xi)) \wedge (\xi + (\zeta_2 - \xi))$$

$$= \eta \wedge \xi_2 \in K$$

In many applications to statistical mechanics, the following theorem which bypasses the above characterization of simplices can be used.

THEOREM II. C. 7 *Let X be a compact convex set. Suppose there exists an affine mapping $\xi \mapsto \mu_\xi$ from X to M_1, such that $r(\mu_\xi) = \xi$ for all. Then each ξ on X is the resultant of a unique maximal measure, and μ_ξ is that maximal measure.*

Proof We will show that, if $r(\nu) = \xi$, then $\nu \prec \mu_\xi$, i.e., $\nu(f) \leqq \mu_\xi(f)$ for all continuous convex f. Since ν may be approximated by measures with finite support and resultant ξ it is enough to show that $\nu(f) \leqq \mu_\xi(f)$ for all ν with finite support and resultant ξ and all $f \in S$. But:

$$\nu(f) = \sum_i \lambda_i f(\xi_i) \leqq \sum_i \lambda_i \mu_{\xi_i}(f)$$

(Since f is convex.)

$$= \mu_{\sum_i \lambda_i \xi_i}(f)$$

(Since $\xi \mapsto \mu_\xi$ is affine.)

$$= \mu_\xi(f).$$

III C* Algebras

A Definitions and Algebraic Preliminaries

An *algebra* is a vector space equipped with a law of composition (multiplication), usually written $A \cdot B$ or AB, which is:

Associative: $A \cdot (B \cdot C) = (A \cdot B) \cdot C$ for all $A, B, C \in \mathfrak{A}$, and

Distributive: $(\alpha A + \beta B) \cdot C = \alpha(A \cdot B) + \beta(B \cdot C)$ for all $A, B, C \in \mathfrak{A}$

$C \cdot (\alpha A + \beta B) = \alpha(C \cdot A) + \beta(C \cdot B)$ and all scalars α, β.

We will consider only algebras for which the field of scalars in the complex numbers $\mathbb{C}$. $\mathfrak{A}$ is *commutative* if $A \cdot B = B \cdot A$ for all $A, B \in \mathfrak{A}$. An element r of $\mathfrak{A}$ is an *identity* if $\mathbf{I} \cdot A = A \cdot \mathbf{I} = A$ for all $A \in \mathfrak{A}$. $\mathfrak{A}$ has at most one identity element; if $\mathbf{I}, \mathbf{I}'$ are both identities, then $\mathbf{I} = \mathbf{I} \cdot \mathbf{I}' = \mathbf{I}'$.

A vector subspace $\mathscr{B}$ of $\mathfrak{A}$ is said to be a *subalgebra* if $A \cdot B \in \mathscr{B}$ for all A, $B \in \mathscr{B}$,

a *left ideal* if $A \cdot B \in \mathscr{B}$ for all $A \in \mathfrak{A}$, $B \in \mathscr{B}$ and if $\mathscr{B} \neq \mathfrak{A}$,

a *right ideal* if $B \cdot A \in \mathscr{B}$ for all $A \in \mathfrak{A}$, $B \in \mathscr{B}$ and if $\mathscr{B} \neq \mathfrak{A}$,

a *two-sided ideal* if it is both a left-ideal and a right ideal.

If $\mathscr{C}$ is a two-sided ideal in $\mathfrak{A}$, then the quotient vector space $\mathfrak{A}/\mathscr{C}$ may be made into an algebra by defining $\dot{A} \cdot \dot{B} = (A \cdot B)\dot{}$, where A is any element of $\dot{A}$ and B any element of $\dot{B}$ (Here, we are regarding elements of $\mathfrak{A}/\mathscr{C}$ as subsets of $\mathfrak{A}$ of the form $A + \mathscr{C}$; such a subset is denoted by $\dot{A}$). The requirement

that $\mathscr{C}$ be a two-sided ideal is just what is required to guarantee that the product be well-defined (i.e., not depend on the choice of representatives A, B of $\dot{A}$, $\dot{B}$).

An *involution* of an algebra $\mathfrak{A}$ is a mapping $A \mapsto A^*$ of $\mathfrak{A}$ into itself such that:

1) * is conjugate linear: $(\alpha A + \beta B)^* = \bar{\alpha} A^* + \bar{\beta} B^*$,

2) reverses the order of products: $(A \cdot B)^* = B^* \cdot A^*$,

3) $(A^*)^* = A$.

If $\mathfrak{A}$, $\mathscr{B}$ are algebras, a *morphism* $\phi : \mathfrak{A} \to \mathscr{B}$ is a linear mapping from $\mathfrak{A}$ to $\mathscr{B}$ such that

$$\phi(AB) = \phi(A)\,\phi(B)$$

for all A, B in $\mathfrak{A}$. If $\mathfrak{A}$ and $\mathscr{B}$ are algebras with involution, we will assume that any morphisms we consider preserve the involution, i.e.,

$$\phi(A^*) = \phi(A)^*$$

for all A in $\mathfrak{A}$.

A *normed algebra* is an algebra $\mathfrak{A}$ equipped with a norm $\| \ \|$ such that

$$\|A \cdot B\| \leq \|A\| \cdot \|B\|.$$

If the algebra $\mathfrak{A}$ is complete in the norm, we speak of a *Banach algebra*. If $\mathfrak{A}$ is equipped with an involution, and if $\|A^*\| = \|A\|$ for all $A \in \mathfrak{A}$, we speak of a *normed algebra with involution* or (if $\mathfrak{A}$ is complete) *Banach algebra with involution*. If $\mathfrak{A}$ has an identity, we will always assume $\|\mathbf{1}\| = 1$. This is not really necessary, but it simplifies things, and, if it is not true, $\mathfrak{A}$ can always be given an equivalent norm which makes it true.

We can now define: A C^* *algebra* is a Banach algebra with involution such that $\|A^*A\| = \|A\|^2$ for all $A \in \mathfrak{A}$. We shall see that this rather innocent-looking condition is very restrictive.

Examples

1. Let X be a locally compact space; consider the set $C_0(X)$ of continuous functions vanishing at ∞ on X. (A continuous function is said to vanish at ∞ if, for every $\varepsilon > 0$, there is a compact set K such that $|f(x)| < \varepsilon$ for $x \in X \setminus K$.) Define:

$$\|f\| = \sup_{x \in X} |f(x)|$$

$$(f + g)(x) = f(x) + g(x)$$

$$(\alpha f)(x) = \alpha f(x)$$

$$(f \cdot g)(x) = f(x)\, g(x)$$

$$f^*(x) = \overline{f(x)}.$$

Then $C_0(X)$ is a commutative C^* algebra. The equality: $\|f^* f\| = \|f\|^2$ holds since $\|f^* f\| = \sup_{x \in X} |f(x)|^2 = \|f\|^2$. $C_0(X)$ has an identity if and

only if X is compact: We will see shortly that every commutative C^* algebra is isomorphic to $C_0(X)$ for some locally compact space X, and that, moreover, X is determined up to homeomorphism by the C^*algebra $C_0(X)$.

2. Let $D = \{z \in \mathbb{C} : |z| < 1\}$, and let $\mathscr{H}^\infty(D)$ denote the set of bounded analytic functions on D. Define addition and multiplication on $\mathscr{H}^\infty(D)$ pointwise (i.e., as for $C_0(X)$); define

$$\|f\| = \sup_{z \in D} |f(z)|$$

$$f^*(z) = \overline{f(\bar{z})}.$$

Then $\mathscr{H}^\infty(D)$ is a commutative Banach algebra with involution ($\mathscr{H}^\infty(D)$ is complete since a uniform limit of analytic functions is analytic), but is *not* a C^* algebra; for example, if $f(z) = e^{iz}$, then $\|f\| = e$ but

$$f(z)f^*(z) = e^{iz}e^{-iz} = 1, \quad \text{so} \quad \|f^*f\| = 1 \neq \|f\|^2.$$

3. Let $\mathscr{H}$ be a Hilbert space, and let $\mathscr{L}(\mathscr{H})$ be the set of all bounded operators on $\mathscr{H}$. Define sums and products in the usual way, and let the norm be the usual operator norm:

$$\|A\| = \sup \{\|A\xi\| : \xi \in \mathscr{H}, \|\xi\| \leqq 1\}.$$

Let * be the ordinary adjoint operation, i.e.,

$$(\xi \mid A^*\eta) = (A\xi \mid \eta)$$

for all $\xi, \eta \in \mathscr{H}$. Then $\mathscr{L}(\mathscr{H})$ is a C^* algebra:

$$\begin{aligned}\|A^*A\| &= \sup \{\|A^*A\xi\| : \xi \in \mathscr{H}\|; \quad \|\xi\| \leqq 1\} \\ &= \sup \{|(\eta|A^*A\xi)| : \xi, \eta \in \mathscr{H}; \|\xi\| \leqq 1; \|\eta\| \leqq 1\} \\ &= \sup \{|(A\eta|A\xi)| : \xi, \eta \in \mathscr{H}; \|\xi\| \leqq 1; \|\eta\| \leqq 1\} = \|A\|^2.\end{aligned}$$

Evidently, any subalgebra of $\mathscr{L}(\mathscr{H})$ which is closed (and hence complete) in the operator norm, and which is self-adjoint (i.e., contains A^* if it contains A) is also a C^* algebra. It turns out that this example gives all C^* algebras, i.e., that every C^* algebra is isomorphic to a norm-closed self-adjoint algebra of bounded operators on a Hilbert space.

It is worth making a remark about terminology at this point: The term "C^* algebra" is sometimes (e.g., in Ruelle [1]) defined to mean a norm-closed self-adjoint algebra of operators on a Hilbert space, and the term "B^* algebra" is used for the more abstract object we have called C^* algebra. The distinction is not so very important since:

a) Every norm-closed self-adjoint algebra of operators on a Hilbert space is a C^* algebra.

b) Every C^* algebra is isomorphic to a norm-closed self-adjoint algebra of operators on a Hilbert space.

When we have occasion to distinguish between the two notions, we will refer to a norm-closed self-adjoint algebra of operators on a Hilbert space as a *concrete C^* algebra.*

The presence of an identity in an algebra is useful for many technical purposes. Therefore, if we have to deal with an algebra with no identity, it is useful to be able to imbed it in a larger algebra which has one. This we can do as follows: Let $\mathfrak{A}$ be an algebra; and let $\tilde{\mathfrak{A}}$ be the set of pairs (λ, A), with $\lambda \in \mathbb{C}$ and $A \in \mathfrak{A}$. Instead of writing pairs (λ, A), we will use the more suggestive notation $\lambda\mathbf{I} + A$. Define:

$$(\lambda\mathbf{I} + A) + (\mu\mathbf{I} + B) = (\lambda + \mu)\mathbf{I} + (A + B)$$

$$(\lambda\mathbf{I} + A)(\mu\mathbf{I} + B) = (\lambda\mu)\mathbf{I} + (\mu A + \lambda B + AB).$$

It is easy to see that, with this structure, $\tilde{\mathfrak{A}}$ is an algebra with identity $0\mathbf{I} + 0$, containing $\mathfrak{A}$ (the set of elements of the form $0\mathbf{I} + A$) as a twosided ideal. (Note that the construction works perfectly well even if $\mathfrak{A}$ has an identity, but that, in this case, the identity of $\mathfrak{A}$ is no longer an identity for $\tilde{\mathfrak{A}}$.) If $\mathfrak{A}$ has an involution, we may extend it to $\tilde{\mathfrak{A}}$ by defining $(\lambda\mathbf{I} + A)^* = \bar{\lambda}\mathbf{I} + A^*$. If $\mathfrak{A}$ is a Banach algebra, or a C^* algebra, we would like to extend the norm of $\mathfrak{A}$ to $\tilde{\mathfrak{A}}$ in such a way that $\tilde{\mathfrak{A}}$ becomes an object of the same sort. Defining

$$\|\lambda\mathbf{I} + A\| = |\lambda| + \|A\|$$

can easily be seen to make $\tilde{\mathfrak{A}}$ into a Banach algebra if $\mathfrak{A}$ is a Banach algebra, but need not make $\tilde{\mathfrak{A}}$ into a C^* algebra if $\mathfrak{A}$ is a C^* algebra. We will give a construction that does this, under the assumption that $\mathfrak{A}$ does not already have an identity. (There is another construction which works in the other case; but we do not need it. See Dixmier C^*A, 1.3.8.) The algebra $\tilde{\mathfrak{A}}$ acts by left multiplication on $\mathfrak{A}$ (since $\mathfrak{A}$ is a left ideal in $\tilde{\mathfrak{A}}$); we will define $\|\lambda\mathbf{I} + A\|$ to be the norm of this multiplication operator (i.e., $\|\lambda\mathbf{I} + A\| = \sup\{\|\lambda B + AB\| : B \in \mathfrak{A};\ \|B\| \leqq 1\}$.) It is immediate from the above description that this expression defines a semi-norm, and that

$$\|(\lambda\mathbf{I} + A)(\mu\mathbf{I} + B)\| \leqq \|(\lambda\mathbf{I} + A)\| \cdot \|(\mu\mathbf{I} + B)\|.$$

For any $A \in \mathfrak{A}$, $\|0\mathbf{I} + A\| \leqq \|A\|$, but also

$$\|(0\mathbf{I} + A)\| \geqq \frac{\|AA^*\|}{\|A^*\|} = \|A\|, \quad \text{so}$$

$\|0\mathbf{I} + A\| = \|A\|$. We have still to check

1) $\|\lambda\mathbf{I} + A\| = 0$ implies $\lambda = 0$ and $A = 0$.

2) $\|\bar{\lambda}\mathbf{I} + A^*\| = \|\lambda\mathbf{I} + A\|$.

3) $\|(\lambda\mathbf{I} + A)(\bar{\lambda}\mathbf{I} + A^*)\| = \|(\lambda\mathbf{I} + A)\|^2$.

4) $\tilde{\mathfrak{A}}$ is complete in the norm we defined

To prove 1: Since $\|0\mathbf{1} + A\| = \|A\|$, $\|0\mathbf{1} + A\| = 0$ only if $A \neq 0$. Hence, we have to show that $\|\lambda\mathbf{1} + A\| \neq 0$ if $\lambda \neq 0$. By homogeneity, it suffices to consider the case $\lambda = 1$. Thus, suppose $\|\mathbf{1} - I\| = 0$. Then $(\mathbf{1} - I)B = B - IB = 0$ for all $B \in \mathfrak{A}$. In other words, I is a left identity for $\mathfrak{A}$. But then, for any $B \in \mathfrak{A}$, $B \cdot (I)^* = ((I) \cdot B^*)^* = B$, so $(I)^*$ is a right identity for $\mathfrak{A}$. Now

$$I = I \cdot I^* = I^*,$$

so I is a two-sided identity for $\mathfrak{A}$, contradicting the assumption that $\mathfrak{A}$ has no identity. This proves 1.

To prove 2 and 3, consider:

$$\begin{aligned}\|\lambda\mathbf{1} + A\|^2 &= \sup\{\|\lambda B + AB\|^2 : \|B\| \leqq 1\} \\ &= \sup\{\|B^*(\lambda\mathbf{1} + A)^*(\lambda\mathbf{1} + A)B\| : \|B\| \leqq 1\} \\ &\leqq \|(\lambda\mathbf{1} + A)^*(\lambda\mathbf{1} + A)\|.\end{aligned}$$

Now in particular:

$$\|\lambda\mathbf{1} + A\|^2 \leqq \|(\lambda\mathbf{1} + A)^*(\lambda\mathbf{1} + A)\| \leqq \|(\lambda\mathbf{1} + A)^*\|\,\|\lambda\mathbf{1} + A\|,$$

so $\|\lambda\mathbf{1} + A\| \leqq \|(\lambda\mathbf{1} + A)^*\|$. The opposite inequality follows by replacing $\lambda\mathbf{1} + A$ by $(\lambda\mathbf{1} + A)^*$, so 2. is proved. Also

$$\|\lambda\mathbf{1} + A\|^2 \leqq \|(\lambda\mathbf{1} + A)^*(\lambda\mathbf{1} + A)\| \leqq \|(\lambda\mathbf{1} + A)^*\|\,\|\lambda\mathbf{1} + A\| = \|\lambda\mathbf{1} + A\|^2,$$

which proves 3. To prove that $\tilde{\mathfrak{A}}$ is complete, note first that $\mathfrak{A}$ is complete and hence closed in $\tilde{\mathfrak{A}}$. Since $\mathfrak{A}$ is the kernel of the linear functional $\lambda\mathbf{1} + A \mapsto \lambda$, this functional is continuous. Thus, if $\lambda_n\mathbf{1} + A_n$ is a Cauchy sequence in $\tilde{\mathfrak{A}}$, λ_n is a Cauchy sequence in $\mathbb{C}$, and hence A_n is a Cauchy sequence in $\mathfrak{A}$. Therefore $\lambda_n \to \lambda$ and $A_n \to A$, so $\lambda_n\mathbf{1} + A_n \to \lambda\mathbf{1} + A$.

The above construction provides a powerful tool for investigating C^* algebras without identity. Unfortunately, it does not solve all problems, and keeping track of what happens in algebras without identity introduces substantial complications into the theory of C^* algebras. By and large, these complications seem to be irrelevant for the application of C^* algebras to physics. We will therefore; in what follows, concentrate on algebras with identity, and make only occasional comments about the general case.

B Spectrum and Resolvent

Let $\mathfrak{A}$ be an algebra with identity. An element A of $\mathfrak{A}$ is said to be *invertible* if there exists an element A^{-1} of $\mathfrak{A}$ such that

$A A^{-1} = A^{-1} A = 1$

The element A^{-1} is uniquely determined if it exists; it is called the *inverse* of A. The *resolvent set* of $A(R(A))$ is the set of all complex numbers λ such that $\lambda\mathbf{1} - A$ is invertible; the inverse $(\lambda\mathbf{1} - A)^{-1}$ is called the *resolvent* of A. The *spectrum* of $A(\sigma(A))$ is the complement of the resolvent set. In general, the spectrum or the resolvent set may be quite arbitrary, but we will show that, in a Banach algebra, the spectrum of any element is a non-empty compact set.

PROPOSITION III. B. 1 *Let A be an element of a normed algebra $\mathfrak{A}$. Then* $\lim_{n\to\infty} \|A^n\|^{1/n}$ *exists and is equal to* $\inf_n \|A^n\|^{1/n}$.

Proof Choose any $m \geq 1$. For any n, write $n = j_n m + i_n$, with $0 \leq i_n < m$. Then

$$\|A^n\|^{1/n} \leq \|A^{j_n m}\|^{1/n} \|A^{i_n}\|^{1/n} \leq \|A^m\|^{j_n/n} \|A^{i_n}\|^{1/n}.$$

Since $\lim_{n\to\infty} j_n/n = 1/m$ and $\limsup_{n\to\infty} \|A^{i_n}\|^{1/n} = 1$ (unless $A^n = 0$ for some n, in which case the proposition is trivial), we get

$$\limsup_{n\to\infty} \|A^n\|^{1/n} \leq \|A^m\|^{1/m}.$$

This is true for all m, so

$$\limsup_{n\to\infty} \|A^n\|^{1/n} \leq \inf_n \|A^n\|^{1/n}.$$

But, trivially,

$$\inf_n \|A^n\|^{1/n} \leq \liminf_n \|A^n\|^{1/n},$$

so the proposition is proved.

The limit whose existence is proved in this proposition is called the *spectral radius* of A and is denoted by $\varrho(A)$. The reason for this terminology is the fact, to be proved shortly, that, if the algebra $\mathfrak{A}$ is complete, $\varrho(A) = \sup\{|\lambda| : \lambda \in \sigma(A)\}$.

PROPOSITION III. B. 2 *Let $\mathfrak{A}$ be a Banach algebra with identity, and let $A \in \mathfrak{A}$ be invertible. If $B \in \mathfrak{A}$ is such that $\varrho(BA^{-1}) < 1$* $\left(\text{*in particular, if* } \|B\| < \frac{1}{\|A^{-1}\|}\right)$, *then $A + B$ is invertible and*

$$(A + B)^{-1} = A^{-1} \sum_{n=0}^{\infty} (-1)^n (BA^{-1})^n;$$

the series converges absolutely in the norm.

Proof The fact that the series converges absolutely in the norm follows at once from the fact that the spectral radius of BA^{-1} is < 1; the fact that the sum of the series is an inverse for $A + B$ is a straightforward compu-

COROLLARY III. B. 3 *Let $\mathfrak{A}$ be a Banach algebra with identity; let $A \in \mathfrak{A}$. If λ is a complex number with $|\lambda| > \varrho(A)$, then λ is in the resolvent set of A, and*

$$(\lambda \mathbf{1} - A)^{-1} = \lambda^{-1} \sum_{n=0}^{\infty} \left(\frac{A}{\lambda}\right)^n.$$

As $\lambda \to \infty$, $\|(\lambda \mathbf{1} - A)^{-1}\| \to 0$.

COROLLARY III. B. 4 *Let $\mathfrak{A}$ be a Banach algebra with identity; let $A \in \mathfrak{A}$; and let λ be in the resolvent set of A. If $|\mu| < \dfrac{1}{\|(\lambda \mathbf{1} - A)^{-1}\|}$, then $\lambda + \mu$ is in the resolvent set of A and*

$$((\lambda + \mu)\mathbf{1} - A)^{-1} = (\lambda \mathbf{1} - A)^{-1} \sum_{n=0}^{\infty} [-\mu(\lambda \mathbf{1} - A)^{-1}]^n.$$

In particular, the resolvent set of A is open (so the spectrum of A is closed), and $\lambda \mapsto (\lambda \mathbf{1} - A)^{-1}$ is an analytic $\mathfrak{A}$-valued function on the resolvent set of A.

So far, we have not ruled out the possibility that the spectrum is empty.

PROPOSITION III. B. 5 *Let $\mathfrak{A}$ be a Banach algebra with identity, and let $A \in \mathfrak{A}$. Then the spectrum of A is non-empty.*

Proof Suppose not. Then, for any continuous linear functional ϕ on $\mathfrak{A}$ $\phi((\lambda \mathbf{1} - A)^{-1})$ is an entire function of λ going to zero at infinity. Hence, by Liouville's Theorem, $\phi((\lambda \mathbf{1} - A)^{-1}) \equiv 0$. This is true for all continuous linear functionals, so $(\lambda \mathbf{1} - A)^{-1} = 0$, which is impossible.

Remark The above proposition is valid without the requirement that $\mathfrak{A}$ is complete (but it is necessary that $\mathfrak{A}$ be normed). In the general case, use the above argument to show that the resolvent set of A in the completion $\bar{\mathfrak{A}}$ of $\mathfrak{A}$ cannot be all of $\mathbb{C}$, and then remark that the resolvent set of A as an element of $\mathfrak{A}$ is contained in the resolvent set of A as an element of $\bar{\mathfrak{A}}$.

THEOREM III. B. 6 (Gelfand) *Let $\mathfrak{A}$ be a Banach algebra with identity in which every non-zero element is invertible. Then $\mathfrak{A} = \{\lambda \mathbf{1} : \lambda \in \mathbb{C}\}$, i.e., $\mathfrak{A}$ is isomorphic to $\mathbb{C}$.*

Proof Let $A \in \mathfrak{A}$. Then $\sigma(A) \neq \phi$, so $A - \lambda \mathbf{1}$ is not invertible for some λ. This implies $A - \lambda \mathbf{1} = 0$, i.e., $A = \lambda \mathbf{1}$.

PROPOSITION III. B. 7 *Let A be a Banach algebra with identity, and let $A \in \mathfrak{A}$. Then $\varrho(A) = \sup\{|\lambda| : \lambda \in \sigma(A)\}$.*

Proof We know from Proposition III. B. 3 that $\varrho(A) \geqq \sup\{|\lambda| : \lambda \in \sigma(A)\}$. Hence, assume that $\varrho(A) > \sup\{|\lambda| : \lambda \in \sigma(A)\}$. Note that $\mu \mapsto (\mathbf{1} - \mu A)^{-1}$ is analytic on $|\mu| < \dfrac{1}{\sup\{|\lambda| : \lambda \in \sigma(A)\}}$. Choose r so that $\varrho(A) > r > \sup\{|\lambda| : \lambda \in \sigma(A)\}$. Then for any continuous linear functional φ on $\mathfrak{A}$

$\phi((\mathbf{I} - \mu A)^{-1})$ has a power series expansion in μ convergent for $|\mu| \leqq 1/r$. But we know that the expansion coefficients must be $\phi(A^n)$ (since $(\mathbf{I} - \mu A)^{-1} = \Sigma \mu^n A^n$ for $|\mu| < (1/\|A\|)$). Hence, $\phi(A^n)/r^n$ is bounded with respect to n for all continuous linear functionals ϕ. By the uniform boundedness principle, this implies that $\|(A/r)^n\|$ is bounded. This violates the assumption that $r < \varrho(A)$ and proves the proposition.

C Commutative Banach Algebras

In this section, we make a preliminary investigation of the structure of commutative Banach algebras. Since the proofs are somewhat technical, we will give a heuristic summary of the approach to be taken. The method of attack is to try to realize, in one sense or another, such algebras as algebras of complex-valued continuous functions. To see how to do this, we first suppose we have a function algebra $\mathfrak{A}$ and ask how we can recover the points of the space on which the functions are defined. One fact about the points of the space is clear: If x is a point, then evaluation at that point $f \leftrightarrow f(x)$ is a morphism from $\mathfrak{A}$ to the complex numbers. Such a morphism (or, at least, one which is not identically zero) we will call a character of $\mathfrak{A}$. Thus, given an abstract commutative algebra $\mathfrak{A}$, we will try to realize it as an algebra of functions on the set of its characters, suitably topologized. (Of course, if we start from an algebra of functions, there may be characters which do not come from points of the space on which the functions are defined. We may, however, regard all the characters as points of some "completion" of the original domain.)

The first technical problem which arises in this program is that of proving the existence of characters. There seems to be no very good direct way of constructing them in general. One therefore makes the remark that a character, as a morphism from $\mathfrak{A}$ to the complex numbers, is uniquely determined by its kernel, which is a maximal ideal in $\mathfrak{A}$. Thus, one wants to find maximal ideals. Now one can make an argument using Zorn's lemma to show that every non-invertible element of $\mathfrak{A}$ is in at least one maximal ideal (Proposition III. C. 2). Next, one has to argue that every maximal ideal is the kernel of a character. This argument has several parts: First one shows that a maximal ideal must be closed by showing that its closure must either be an ideal or all of $\mathfrak{A}$ and ruling out the latter possibility by showing that all elements sufficiently near the identity are invertible so the identity cannot be in the closure of any ideal. (Proposition III. C. 4.) Next one shows that the quotient of a Banach algebra by a closed ideal is again a Banach algebra. Finally, one argues that, in the quotient of an algebra by a maximal ideal, every non-zero element is invertible. (Proposition III. C. 3.) Combining these two remarks with the theorem that the only Banach algebra in which

every non-zero element is invertible in $\mathbb{C}$ itself one concludes that the quo-

tient of $\mathfrak{A}$ by any maximal ideal is isomorphic to $\mathbb{C}$, i.e., that every maximal ideal of $\mathfrak{A}$ is the kernel of a character. (Proposition III. C. 5.)

Now one topologizes the set of characters conveniently and associates with each element A of $\mathfrak{A}$ the function $\hat{A}$ on the set of characters of $\mathfrak{A}$ which, at the character χ, takes $\hat{A}(\chi) = \chi(A)$. This gives a morphism of $\mathfrak{A}$ into the algebra of continuous functions on the set of characters of $\mathfrak{A}$. This morphism is not very satisfactory in general: It need not be norm preserving (or even injective) and it need not take the involution on $\mathfrak{A}$ (if one exists) to complex conjugation. In the next section, we will see that the situation is much better if $\mathfrak{A}$ is a commutative C^* algebra.

We now turn to the technical details, starting with some remarks valid in general normed algebras: Let $\mathfrak{A}$ be a normed algebra, $\mathscr{C}$ a *closed* two-sided ideal in $\mathfrak{A}$. For $\dot{A} \in \mathfrak{A}/\mathscr{C}$, define $\|\dot{A}\| = \inf\{\|A\| : A \in \dot{A}\}$. It is not hard to verify that this defines a norm on $\mathfrak{A}/\mathscr{C}$ (If $\|\dot{A}\| = 0$, i.e., if A contains elements of arbitrarily small norm, then, since $\dot{A}$ is closed in $\mathfrak{A}$, $0 \in \dot{A}$, i.e. $\dot{A} = 0$ in $\mathfrak{A}/\mathscr{C}$) making $\mathfrak{A}/\mathscr{C}$ into a normed algebra. If $\mathfrak{A}$ is complete, so is $\mathfrak{A}/\mathscr{C}$. (This is a general fact about Banach spaces: Let $(\dot{A}_n)$ be a Cauchy sequence in $\mathfrak{A}/\mathscr{C}$. It suffices to show that a subsequence of $(\dot{A}_n)$ converges. Choose a subsequence $\dot{A}_{n_j}$ so that $\|\dot{A}_{n_j} - \dot{A}_m\| \leqq 1/2^j$ for $m \geqq n_j$. Then $\|\dot{A}_{n_j} - \dot{A}_{n_{j+1}}\| \leqq 1/2^j$. Choose representatives A_{n_j} of $\dot{A}_{n_j}$ such that $\|A_{n_j} - A_{n_{j+1}}\| \leqq 1/2^{j-1}$. Then $\sum_{j=1}^{\infty} \|A_{n_j} - A_{n_{j+1}}\| < \infty$, so (A_{n_j}) converges to, say, A; hence, $\dot{A}_{n_j}$ converges to $\dot{A}$).

If $\mathfrak{A}$ is an algebra with involution, and if $\mathscr{C}^* = \mathscr{C}$, then we can make $\mathfrak{A}/\mathscr{C}$ an algebra with involution by defining $(\dot{A} + \mathscr{C})^* = \dot{A}^* + \mathscr{C}$. The involution is compatible with the quotient norm on $\mathfrak{A}/\mathscr{C}$. (If $\mathfrak{A}$ is a C^* algebra, and if $\mathscr{C}$ is a closed, two-sided ideal, then $\mathscr{C}$ is automatically self-adjoint and $\mathfrak{A}/\mathscr{C}$ is a C^* algebra. We will not prove these facts; they are contained in Proposition I. 8.2 of Dixmier C^*A.)

Next, we investigate ideals in algebras:

PROPOSITION III. C.1 *Let $\mathfrak{A}$ be an algebra with identity and let $\mathscr{C}$ be a (left, right, two-sided) ideal in $\mathfrak{A}$. Then $\mathscr{C}$ is contained in a maximal (left, right, two-sided) ideal.*

Proof Order the set of ideals by inclusion. Let $(\mathscr{C}_\alpha)$ be an increasing family of ideals; then $\bigcup_\alpha \mathscr{C}_\alpha$ is an ideal containing all the $\mathscr{C}_\alpha$. $\left(\bigcup_\alpha \mathscr{C}_\alpha \text{ cannot be all of } \mathfrak{A}, \text{ since } \mathbf{I} \text{ cannot belong to any } \mathscr{C}_\alpha\right)$. The proposition follows by Zorn's Lemma.

PROPOSITION III. C. 2 *Let $\mathfrak{A}$ be a commutative algebra with identity, and let $A \in \mathfrak{A}$. Then A belongs to some maximal ideal of $\mathfrak{A}$ if and only if A*

is not invertible.

Proof If A is invertible, then it cannot belong to any ideal of $\mathfrak{A}$. If A is not invertible, then $\mathfrak{A} \cdot A$ is not all of $\mathfrak{A}$ (it cannot contain $\mathbf{I}$); hence, is an ideal of $\mathfrak{A}$; hence, is contained in a maximal ideal of $\mathfrak{A}$.

PROPOSITION III. C. 3 *Let $\mathfrak{A}$ be a commutative algebra with identity, $\mathscr{C}$ an ideol of $\mathfrak{A}$. Then $\mathscr{C}$ is maximal if and only if $\mathfrak{A}/\mathscr{C}$ is a field (i.e., if and only if every non-zero element of $\mathfrak{A}/\mathscr{C}$ is invertible).*

Proof Suppose $\mathfrak{A}/\mathscr{C}$ contains a non-invertible element different from 0. Then $\mathfrak{A}/\mathscr{C}$ contains an ideal different from $\{0\}$. The inverse image of this ideal under the canonical morphism $\mathfrak{A} \to \mathfrak{A}/\mathscr{C}$ is an ideal of $\mathfrak{A}$ properly containing $\mathscr{C}$. Thus, $\mathscr{C}$ is not maximal. Conversely, if $\mathscr{C}$ is not maximal, there is an ideal $\mathscr{I}$ in $\mathfrak{A}$ properly containing $\mathscr{C}$; the image of $\mathscr{I}$ in $\mathfrak{A}/\mathscr{C}$ is an ideal which is not equal to $\{0\}$; any element of this ideal must be non-invertible.

So far, we have been doing pure algebra. We now look specifically at Banach algebras.

PROPOSITION III. C. 4 *Let $\mathfrak{A}$ be a Banach algebra with identity, and let $\mathscr{C}$ be a (left, right, two-sided) ideal in $\mathfrak{A}$. Then $\overline{\mathscr{C}}$ (the closure of $\mathscr{C}$) is also (left, right, two sided) ideal in $\mathfrak{A}$. Every maximal (left, right, two-sided) ideal in $\mathscr{C}$ is closed.*

Proof We consider the case of $\mathscr{C}$ a left ideal. Let $A \in \overline{\mathscr{C}}$ and let $B \in \mathfrak{A}$; we want to show that $B \cdot A \in \overline{\mathscr{C}}$. Let A_n be a sequence in $\mathscr{C}$ converging to A; then $B \cdot A_n$ is a sequence in $\mathscr{C}$ converging to $B \cdot A$. Hence, $B \cdot A \in \overline{\mathscr{C}}$. Thus, all that remains to be shown is that $\overline{\mathscr{C}}$ is not all of $\mathfrak{A}$. By Proposition III. B. 2, if $\|A - \mathbf{I}\| < 1$, then A is invertible and hence cannot belong to $\mathscr{C}$. Thus $\{A : \|A - \mathbf{I}\| \geq 1\}$ is a closed set containing $\mathscr{C}$ and hence $\overline{\mathscr{C}}$, so $\mathbf{I} \notin \overline{\mathscr{C}}$. A maximal ideal is closed since, if it were not, its closure would be a strictly larger ideal.

PROPOSITION III. C. 5 *Let $\mathfrak{A}$ be a commutotive Banach algebra with identity. If $\mathscr{C}$ is a maximal ideal, then $\mathfrak{A}/\mathscr{C}$ is isomorphic to $\mathbb{C}$.*

Proof If $\mathscr{C}$ is a maximal ideal, then $\mathfrak{A}/\mathscr{C}$ is a Banach algebra in which every non-zero element is invertible; hence, by Theorem III. B. 6, is isomorphic to $\mathbb{C}$.

A *character* of a commutative algebra $\mathfrak{A}$ is a non-zero morphism of $\mathfrak{A}$ into the complex numbers. If $\mathfrak{A}$ has an identity, and if χ is a character of $\mathfrak{A}$, then $\chi(\mathbf{I}) = 1$. The kernel of a character of a commutative algebra with identity is a maximal ideal of the algebra. Conversely, by Proposition III. C. 5, every maximal ideal of a commutative Banach algebra with identity is the kernel of a character. Thus, characters are, in this case, in one-one

correspondence with maximal ideals

PROPOSITION III. C. 6 *Let $\mathfrak{A}$ be a commutative Banach algebra with identity, and let $A \in \mathfrak{A}$. There is a character χ of $\mathfrak{A}$ such that $\chi(A) = \lambda$ if and only if $\lambda \in \sigma(A)$. In particular, for any character χ of $\mathfrak{A}$,*

$$|\chi(A)| \leqq \varrho(A) \leqq \|A\|.$$

Proof There is a character χ such that $\chi(A) = \lambda$ if and only if $A - \lambda\mathbf{1}$ is sent to zero by some character, which is true if and only if $A - \lambda\mathbf{1}$ is not invertible (Proposition III. C. 2), i.e., if and only if $\lambda \in \sigma(A)$.

A character χ of $\mathfrak{A}$ is in particular a linear functional on $\mathfrak{A}$; the above proposition says that χ is bounded and has norm not greater than 1. (The norm of a character is in fact equal to 1, since $\chi(\mathbf{1}) = 1$.) Thus, the set of characters of $\mathfrak{A}$ is a subset of the unit ball of the dual $\mathfrak{A}^*$ of $\mathfrak{A}$. We claim that it is in fact a closed subset in the weak-* topology. Let χ_α be a net of characters converging in the weak topology to the linear functional χ; since $\chi_\alpha(\mathbf{1}) = 1$ for all α, $\chi(\mathbf{1}) = 1$; if $A, B \in \mathfrak{A}$, then

$$\chi(AB) = \lim_\alpha \chi_\alpha(AB) = \lim_\alpha \chi_\alpha(A) \cdot \lim_\alpha \chi_\alpha(B) = \chi(A) \cdot \chi(B),$$

so χ is a character. Hence we have:

PROPOSITION III. C. 7 *Let $\mathfrak{A}$ be a commutative Banoch algebra with identity. Then the separate set of characters of $\mathfrak{A}$ is compact in the weak-* topology on the dual of $\mathfrak{A}$.*

The set of characters of $\mathfrak{A}$, as a topological space, is called the *spectrum*, or *maximal ideal space*, of $\mathfrak{A}$; we will denote it by $S(\mathfrak{A})$. If $A \in \mathfrak{A}$, we may define a function $\hat{A}$ on $S(\mathfrak{A})$ by $\hat{A}(\chi) = \chi(A)$. We collect in the following proposition the elementary properties of the map $A \mapsto \hat{A}$, which is called the *Gelfand transform.*

PROPOSITION III. C. 8 *Let $\mathfrak{A}$ be a commutative Banach algebra with identity. The mapping $A \mapsto \hat{A}$ is a morphism of the algebra $\mathfrak{A}$ into the algebra of continuous complex-valued functions on $S(\mathfrak{A})$ with the pointwise operations. The range of $\hat{A}$ is precisely the spectrum of A; in particular, $\|\hat{A}\| = \sup_{\chi \in S(\mathfrak{A})} |\chi(A)| = \varrho(A) \leqq \|A\|$. The functions $\hat{A}$ separate points of $S(\mathfrak{A})$, i.e., given $\chi_1, \chi_2 \in S(\mathfrak{A})$, $\chi_1 \neq \chi_2$, we may find $A \in \mathfrak{A}$ such that $\hat{A}(\chi_1) \neq \hat{A}(\chi_2)$.*

Proof The fact that $\hat{A}$ is continuous for all $A \in \mathfrak{A}$ follows tautologically from the choice of the topology on $S(\mathfrak{A})$. The mapping $A \mapsto \hat{A}$ is certainly linear; also $\widehat{A \cdot B}(\chi) = \chi(AB) = \chi(A)\chi(B) = \hat{A}(\chi)\hat{B}(\chi)$, so $A \mapsto \hat{A}$ is a morphism. The fact that the range of $\hat{A}$ is the spectrum of A is Proposition III. C. 6. The statement that $\chi_1 \neq \chi_2$ means $\chi_1(A) \neq \chi_2(A)$ for some A i.e. $\hat{A}(\chi_1) \neq \hat{A}(\chi_2)$ for some A.

The Gelfand transform generalizes, in a certain sense, the Fourier transform. Let $l^1(\mathbf{I})$ denote the space of two-sided sequences of complex numbers $(\ldots, a_{-2}, a_{-1}, a_0, a_1, \ldots)$ such that $\sum_{n=-\infty}^{\infty} |a_n| < \infty$. If $A = (a_n)$, $B = (b_n)$, define $(A \cdot B)_n = \sum_j a_j \cdot b_{n-j}$, i.e., define products by convolution. Then it may be verified that $l^1(\mathbf{I})$ is a commutative Banach algebra with identity. It turns out that the spectrum of $l^1(\mathbf{I})$ may be identified with the complex numbers of absolute value 1 with the usual topology; the character identified with $e^{i\theta}$ sends $A = (a_n)$ to $\sum_n a_n e^{in\theta}$. Thus, the map $A \mapsto \hat{A}$ is just the classical Fourier transform. In this case, the range of the Gelfand transform is not all continuous functions but the set of continuous functions with absolutely convergent Fourier series, and the supremum norm of $\hat{A}$ is generally strictly smaller than the norm of A.

D Commutative C^* Algebras

At the end of the last section, we saw an example of a commutative Banach algebra with identity in which the Gelfand transform sends $\mathfrak{A}$ to a dense subalgebra of $C(S(\mathfrak{A}))$. The point of the present section is to show that for C^* algebras the situation is better. Specifically, we will prove:

THEOREM III. D. 1 *Let $\mathfrak{A}$ be a commutative C^* algebra with identity. Then the Gelfand transform is a norm-preserving isomorphism of $\mathfrak{A}$ onto the algebra of continuous functions on $S(\mathfrak{A})$.*

Before proving this theorem, we will establish some terminology, make a few remarks, and establish a subsidiary result. An element A in an algebra $\mathfrak{A}$ with involution is said to be *self-adjoint* if $A^* = A$ and *normal* if A commutes with A^*. For any A, A^*A is self-adjoint. Any element A of an algebra with involution may be written uniquely as $A = A_1 + iA_2$ with A_1 and A_2 self-adjoint; indeed $A = \frac{1}{2}(A + A^*) + i\left(\frac{A - A^*}{2i}\right)$. A is normal if and only if A_1 and A_2 commute. If A, B are self-adjoint elements of $\mathfrak{A}$, then $A \cdot B$ is self-adjoint if and only if A and B commute.

PROPOSITION III. D. 2 *Let $\mathfrak{A}$ be a C^* algebra, A a normal element of $\mathfrak{A}$. Then $\varrho(A) = \|A\|$.*

Proof

$$\|A^{2^n}\|^2 = \|A^{*2^n}A^{2^n}\| = \|(A^*A)^{2^n}\| = \|(A^*A)^{2^{n-1}}\|^2 = \cdots = \|A^*A\|^{2^n} = \|A\|^{2^{n+1}}.$$

Thus,

$\|\hat{A}\|^{1/2^n} = \|A\|$ so $\rho(A) = \lim_{n\to} \|A^{2^n}\|^{1/2^n} = \|A\|$

Proof of Theorem III. D. 1 Since $\mathfrak{A}$ is commutative, every element of $\mathfrak{A}$ is normal, so $\|\hat{A}\| = \rho(A) = \|A\|$ for all $A \in \mathfrak{A}$, i.e., the Gelfand transform is norm-preserving as a map from $\mathfrak{A}$ to $C(S(\mathfrak{A}))$. We will show that, if A is self-adjoint, $\hat{A}$ is real. This will imply that $\{\hat{A}: A^* = A\}$ is a norm-complete (hence, norm-closed) algebra of continuous real-valued functions on $S(\mathfrak{A})$ containing the constants and separating points. By the Stone-Weierstrass Theorem, then, every continuous real-valued function is the image under the Gelfand transform of a self-adjoint element of $\mathfrak{A}$, so the Gelfand transform is surjective and thus an isomorphism.

Let $A \in \mathfrak{A}$ be self-adjoint. Define $\exp(\pm iA) = \sum_{n=0}^{\infty} \frac{(\pm iA)^n}{n!}$. It is straightforward to verify that $[\exp(+iA)]^* = \exp(-iA)$ and that $\exp(+iA) \times \exp(-iA) = \mathbf{1}$. Hence, $\|\exp(\pm iA)\|^2 = \|\exp(\pm iA)\exp(\mp iA)\| = 1$. Thus; for any character χ of $\mathfrak{A}$,

$$1 \geq |\chi(\exp(\pm iA))| = \left|\chi\left(\sum_{n=0}^{\infty} 1/n!\,(\pm iA)^n\right)\right| = \left|\sum_{n=0}^{\infty} 1/n!\,(\pm i\chi(A))^n\right|$$

$$= |\exp(\pm i\chi(A))|,$$

so $\chi(A)$ is real.

Let us look briefly at what happens if $\mathfrak{A}$ does not have an identity. Then we can imbed $\mathfrak{A}$ in $\tilde{\mathfrak{A}}$. The Gelfand transform for $\tilde{\mathfrak{A}}$ then maps $\mathfrak{A}$ isometrically onto a subalgebra (in fact, a maximal ideal) of the algebra of continuous functions on $S(\tilde{\mathfrak{A}})$. Now every multiplicative linear functional χ on $\mathfrak{A}$ extends uniquely to a character $\tilde{\chi}$ on $\tilde{\mathfrak{A}}$ by $\tilde{\chi}(\lambda\mathbf{1} + A) = \lambda + \chi(A)$. Conversely, every character of $\tilde{\mathfrak{A}}$ restricts to a character of $\mathfrak{A}$ except for the character $\tilde{\chi}_\infty$ defined by $\tilde{\chi}_\infty(\lambda\mathbf{1} + A) = \lambda$ (which restricts to the zero functional). Thus, the set of characters of $\mathfrak{A}$, which we will again denote by $S(\mathfrak{A})$ and equip with the weak-* topology, may be identified with $S(\tilde{\mathfrak{A}}) \setminus \{\tilde{\chi}_\infty\}$ (which is a locally compact space since it is obtained by deleting one point from a compact space). Moreover, an element A of $\tilde{\mathfrak{A}}$ belongs to $\mathfrak{A}$ if and only if $\tilde{\chi}_\infty(A) = 0$, so the Gelfand transform for $\tilde{\mathfrak{A}}$ sends $\mathfrak{A}$ to the algebra of all continuous functions on $S(\tilde{\mathfrak{A}})$ vanishing at $\tilde{\chi}_\infty$, i.e., the Gelfand transform for $\mathfrak{A}$ sends $\mathfrak{A}$ isomorphically and isometrically onto the algebra of all continuous functions vanishing at infinity on the locally compact space $S(\mathfrak{A})$.

With the detailed information we have about commutative C^* algebras, we can easily analyze morphisms of such objects. Let $\mathfrak{A}$, $\mathfrak{B}$ be commutative C^* algebras with identity, and let $\varphi: \mathfrak{A} \to \mathfrak{B}$ be a morphism such that $\varphi(\mathbf{1}_{\mathfrak{A}}) = \mathbf{1}_{\mathfrak{B}}$. If χ is a character of $\mathfrak{B}$, then $\chi \circ \varphi$ is a character of $\mathfrak{A}$. Thus, we may define a mapping $\varphi^*: S(\mathfrak{B}) \to S(\mathfrak{A})$ by $\varphi^*(\chi) = \chi \circ \varphi$. It follows directly from the definition of the topologies on $S(\mathfrak{A})$ and $S(\mathfrak{B})$ that φ^* is continuous. Conversely, if φ^* is any continuous mapping of $S(\mathfrak{B})$ into $S(\mathfrak{A})$, then $\hat{A} \mapsto \hat{A} \circ \varphi^*$

defines a morphism from the continuous functions on $S(\mathfrak{B})$ into the contin-

uous functions on $S(\mathfrak{B})$, i.e. ,a morphism from $\mathfrak{A}$ to $\mathfrak{B}$. Thus, morphisms from $\mathfrak{A}$ to $\mathfrak{B}$ sending $\mathbf{1}_{\mathfrak{A}}$ to $\mathbf{1}_{\mathfrak{B}}$ are in one-one correspondence with continuous mappings of $S(\mathfrak{B})$ into $S(\mathfrak{A})$.

PROPOSITION III. D. 3 *Let $\mathfrak{A}$, $\mathfrak{B}$ be commutative C^* algebras with identity, φ a morphism of $\mathfrak{A}$ into $\mathfrak{B}$ sending $\mathbf{1}_{\mathfrak{A}}$ to $\mathbf{1}_{\mathfrak{B}}$, φ^* the associated mapping from $S(\mathfrak{B})$ to $S(\mathfrak{A})$. Then φ is injective if and only φ^* is surjective, and φ is surjective if and only if φ^* is injective. If φ is injective, then $\|\varphi(A)\| = \|A\|$ for all $A \in \mathfrak{A}$.*

Proof Since φ^* is continuous and $S(\mathfrak{B})$ is compact, $\varphi^*(S(\mathfrak{B}))$ is compact and hence closed in $S(\mathfrak{A})$. Now $\varphi(A) = 0$ if and only if $\chi(\varphi(A)) = 0$ for all $\chi \in S(\mathfrak{B})$, i.e., if and only if $\hat{A} = 0$ on $\varphi^*(S(B))$. This shows that φ is injective if and only if zero is the only continuous function on $S(\mathfrak{A})$ vanishing on $\varphi^*(S(\mathfrak{B}))$. Since $\varphi^*(S(\mathfrak{B}))$ is closed, this is equivalent to $\varphi^*(S(\mathfrak{B})) = S(\mathfrak{A})$. If φ is injective, then $\|\varphi(A)\| = \sup\{|\chi(\varphi(A))| : \chi \in S(\mathfrak{B})\} = \sup\{|\varphi^*(\chi)(A)| : \chi \in S(\mathfrak{B})\} = \sup\{|\chi(A)| : \chi \in S(\mathfrak{A})\} = \|A\|$.

The morphism φ is surjective if and only if, for all $B \in \mathfrak{B}$, there is an $A \in \mathfrak{A}$ such that $\hat{A} \circ \varphi^* = \hat{B}$. If φ^* is not injective this is clearly not possible since, if $\varphi^*(\chi_1) = \varphi^*(\chi_2)$, no $\hat{B}$ separating χ_1 from χ_2 can be so obtained. On the other hand, if φ^* is injective, then it is a continuous one-one mapping from the compact space $S(\mathfrak{B})$ to the compact set $\varphi^*(S(\mathfrak{B})) \subset S(\mathfrak{A})$. But such a mapping has a continuous inverse (it sends closed sets, i.e., compact sets, to sets which are compact; hence closed, so the inverse image under its inverse of a closed set is closed.) Thus, all we have to do is to extend the continuous function $\hat{B} \circ (\varphi^*)^{-1}$, defined on $\varphi^*(S(\mathfrak{B}))$, to a continuous function $\hat{A}$ on all of $S(\mathfrak{A})$. The Tietze Extension Theorem asserts the existence of such an extension, so the proposition is proved.

E The Spectral Theorem for Bounded Normal Operators on Hilbert Space

As an example of the power of the methods we have been discussing, we will show that they lead to a very quick proof of the spectral theorem for bounded normal operators. Let $\mathscr{H}$ be a Hilbert space, and let A be a bounded operator on $\mathscr{H}$ which is normal, i.e., which commutes with its adjoint. (In particular, A can be self-adjoint or unitary.) Let $\mathfrak{A}$ be the C^* algebra generated by A and $\mathbf{1}$, i.e., the norm closure of the set of all polynomials in A and A^*. $\mathfrak{A}$ is a commutative C^* algebra, and hence is isomorphic to the algebra of all continuous functions on a compact space. (It is worth remarking that in this case, $S(\mathfrak{A})$ may be identified with the spectrum of A, regarded as a subset of $\mathbb{C}$. The proof is as follows: Consider the mapping $S(\mathfrak{A}) \to \mathbb{C}$ defined by $\chi \mapsto \chi(A)$. This is continuous by the definition of the topology on $S(\mathfrak{A})$, and its image is exactly the spectrum of A. Since $\chi(A^*) = \overline{\chi(A)}$

For any character of $\mathfrak{A}$ and since a character is a multiplicative linear

functional, two characters which agree on A agree on all polynomials in A and A^* and, hence, by continuity, are equal on $\mathfrak{A}$. Thus, the mapping $\chi \leftrightarrow \chi(A)$ is continuous and one-one from $S(\mathfrak{A})$ to $\sigma(A)$. But we have already observed that a continuous one-one mapping between compact spaces is a homeomorphism, so $S(\mathfrak{A})$ is homeomorphic to $\sigma(A)$).

Now let us assume, temporarily, that $\mathfrak{A}$ has a cyclic vector, i.e., that there is a vector $\xi \in \mathscr{H}$, which we may take to have norm one, such that $\{B\xi: B \in \mathfrak{A}\}$ is dense in $\mathscr{H}$. We will define a linear functional μ on $C(S(\mathfrak{A}))$ by $\mu(\hat{B}) = (\xi \mid B\xi)$. We claim that μ is positive and hence defines a measure on $S(\mathfrak{A})$. To see this, suppose that $\hat{B} \geqq 0$ everywhere. Then $\sqrt{\hat{B}}$ ($\sqrt{\ }$ denotes the positive square root) is again a continuous function on $S(\mathfrak{A})$ and hence is the Gelfand transform of a self-adjoint element of $\mathfrak{A}$, which we will denote by $\sqrt{B}$, and which evidently has the property that $(\sqrt{B})^2 = B$. Then

$$\mu(\hat{B}) = (\xi \mid B\xi) = \left(\xi \,\middle|\, \sqrt{B}\cdot\sqrt{B}\,\xi\right) = \left(\sqrt{B}\,\xi \,\middle|\, \sqrt{B}\,\xi\right) \geqq 0.$$

This shows that μ defines a measure on $S(\mathfrak{A})$; we will also denote the measure by μ. We now claim that we can define a unitary mapping of $\mathscr{H}$ into $\mathscr{L}^2(\mu)$ by

$$B\xi \leftrightarrow \hat{B}.$$

It is not clear at the moment that this mapping is well defined, but we have, for $B, C \in \mathfrak{A}$,

$$(B\xi \mid C\xi) = (\xi \mid B^*C\xi) = \int d\mu\, \widehat{B^*C} = \int d\mu\, \bar{\hat{B}}\hat{C}.$$

so the proposed mapping preserves scalar products, so it is well defined and length-preserving; since it is defined on the dense subset $\mathfrak{A}\xi$ of $\mathscr{H}$ and has range the dense subset $C(S(\mathfrak{A})) \subset \mathscr{L}^2(\mu)$, it extends uniquely to a unitary operator U from $\mathscr{H}$ onto $\mathscr{L}^2(\mu)$. Now let $B \in \mathfrak{A}$ be arbitrary; we have

$$\hat{AB} = UAB\xi = UAU^{-1}UB\xi = UAU^{-1}\hat{B},$$

so UAU^{-1} is equal to the operator M_A of multiplication by A on the dense subset $C(S(\mathfrak{A}))$ of $\mathscr{L}^2(\mu)$; since both operators are bounded, they must agree on all of $\mathscr{L}^2(\mu)$. Stripping away the details of the construction, we see that we have proved the following: If $\mathfrak{A}$ has a cyclic vector, then A is unitarily equivalent to the operator of multiplication by a continuous function on $\mathscr{L}^2(\mu)$, where μ is a Borel measure on a compact space (Moreover, the space may be taken to be $\sigma(A)$ and the function to be the co-ordinate function z.)

We now must remove the requirement that $\mathfrak{A}$ have a cyclic vector. To do this, we use the fact, to be proved later, that if $\mathfrak{A}$ is any concrete C^* algebra containing the identity operator on a Hilbert space $\mathscr{H}$, then $\mathscr{H}$ may be decomposed into a direct sum of subspaces $\mathscr{H} = \bigoplus_{i \in I} \mathscr{H}_i$ such that each $\mathscr{H}_i$ is mapped into itself by each element of $\mathfrak{A}$ and such that the restriction of $\mathfrak{A}$ to each $\mathscr{H}_i$ has a cyclic vector (See Proposition III. F.2) Then we can

use the above argument to show that A, restricted to any $\mathscr{H}_i$, is unitarily equivalent to the operator of multiplication by $\hat{A}$ on $\mathscr{L}^2(\mu_i)$ for some measure μ_i on $S(\mathfrak{A})$. We may now construct a (possibly very large) space $\mathscr{M}$ by taking the disjoint union of one copy of $S(\mathfrak{A})$ for each $i \in I$ and defining a Borel measure μ on this union by requiring that its restriction to the i^{th} copy be just μ_i. Then $\mathscr{L}^2(\mu)$ may be identified with $\bigoplus_{i \in I} \mathscr{L}^2(\mu_i)$, and this identification makes A unitarily equivalent to the operator on $\mathscr{L}^2(\mu)$ of multiplication by the function whose restriction to each copy of $S(\mathfrak{A})$ is just $\hat{A}$. Thus, we can formulate the following theorem.

THEOREM III. E. 1 *Let A be a bounded normal operator on a Hilbert space. Then A is unitarily equivalent to the operator of multiplication by a bounded continuous function on $\mathscr{L}^2$ of a Borel measure on a locally compact space.*

The realization of A as a multiplication operator may be thought of as "diagonalizing" A. We may easily deduce from this statement the version of the spectral theorem expressing A in terms of a projection-valued measure: Let $A = UM_f U^{-1}$, where M_f is multiplication by f on $\mathscr{L}^2(\mu)$ and let S be a Borel subset of $\mathbb{C}$. Let $U^{-1}P(S)\,U$ be the operator of multiplication by the characteristic function of $\{x : f(x) \in S\}$ on $\mathscr{L}^2(\mu)$. It is easy to check that $S \mapsto P(S)$ is a projection-valued measure on $\mathbb{C}$ and that, for any $\xi \in \mathscr{H}$,

$$(\xi \mid M_f \xi) = \int_{\mathbb{C}} \lambda (\xi \mid dP(\lambda)\, \xi).$$

If A is self-adjoint, then f is real so the projection-valued measure P is concentrated on the real axis, and we get the usual representation of a bounded self-adjoint operator in terms of its spectral resolution.

In the classical versions of the spectral theorem, it is shown that, if A is self-adjoint and if B commutes with A, then B commutes with the spectral projections of A. We can prove this a follows: Let F be any closed set in the spectrum of A. Then there is a non-negative continuous function $\hat{C}(\lambda)$ on $\sigma(A)$ equal to one on F, and strictly less than one on the complement of F. $\hat{C}$ is the Gelfand transform of an element C of $\mathfrak{A}$. Since, by assumption, B commutes with A, it commutes with all polynomials in A, and hence with all elements of $\mathfrak{A}$; in particular, B commutes with C. By realizing A as a multiplication operator, and applying the monotone convergence theorem, we see that $C^n\xi$ converges to $P(F)\,\xi$ for all $\xi \in \mathscr{H}$. Thus

$$B\,P(F)\,\xi = \lim_{n \to \infty} BC^n\xi = \lim_{n \to \infty} C^n B\xi = P(F)\,B\xi$$

for all $\xi \in \mathscr{H}$, i.e., B commutes with $P(F)$ for all closed $F \subset \sigma(A)$. It is easy to see that the collection of subsets G of $\sigma(A)$ such that $P(G)$ commutes with B is a [illegible]

with B is a σ-field; since it contains closed sets by the above argument,

it contains all Borel sets, i.e., $P(E)$ commutes with B for all Borel subsets E of $\sigma(A)$.

One advantage of the derivation of the spectral theorem that we have given over the more classical ones is that it shows that any collection, finite or infinite, of commuting self-adjoint operators can be simultaneously diagonalized, i.e., realized as multiplication operators on an $\mathscr{L}^2$ space. The proof is the same; one has only to replace $\mathfrak{A}$ by the C^* algebra generated by the family of operators in question.

F Generalities on Representations

Let $\mathfrak{A}$ be an algebra with involution. A *representation* of $\mathfrak{A}$ on a Hilbert space $\mathscr{H}$ is a morphism of $\mathfrak{A}$ into the C^* algebra $\mathscr{L}(\mathscr{H})$ of bounded operators on $\mathscr{H}$. (Thus, we are defining "representation" to mean "representation by bounded operators". We will see later that, if $\mathfrak{A}$ is a *Banach* algebra with involution and identity and if π is any morphism of $\mathfrak{A}$ into the set of (possibly unbounded) operators on a dense domain in a Hilbert space with adjoint defined in the obvious way, then $\pi(A)$ is bounded for all $A \in \mathfrak{A}$ and, indeed,

$$\|\pi(A)\| \leqq \|A\|.$$

Two representations π and π' on Hilbert spaces $\mathscr{H}$ and $\mathscr{H}'$ are said to be *unitarily equivalent* ($\pi \cong \pi'$) if there is a unitary operator U from $\mathscr{H}$ to $\mathscr{H}'$ such that

$$U^{-1}\,\pi'(\cdot)\,U = \pi(\cdot).$$

Let $(\pi_i)_{i \in I}$ be an indexed set of representations of $\mathfrak{A}$ on Hilbert spaces $(\mathscr{H}_i)$. If, for each $A \in \mathfrak{A}$, $\sup_i \|\pi_i(A)\| < \infty$ (In particular, if $\mathfrak{A}$ is a Banach algebra with identity, by the remark made above), we can form the direct sum representation $\bigoplus_{i \in I} \pi_i$ on the direct sum Hilbert space $\bigoplus_{i \in I} \mathscr{H}_i$ as follows: $\bigoplus_{i \in I} \mathscr{H}_i$ is the set of indexed families (ξ_i) with $\xi_i \in \mathscr{H}_i$ such that $\sum_i \|\xi_i\|^2 < \infty$. For $A \in \mathfrak{A}$, the operator $\bigoplus_i \pi_i(A)$ is defined to take (ξ_i) to $(\pi_i(A)\,\xi_i)$.

If π is a representation of $\mathfrak{A}$ on $\mathscr{H}$, and if $\mathscr{H}^1$ is a linear subspace of $\mathscr{H}$, we say that $\mathscr{H}^1$ is an *invariant subspace* for π if $\pi(A)\,\mathscr{H}^1 \subset \mathscr{H}^1$ for all $A \in \mathfrak{A}$. If $\mathscr{H}^1$ is invariant, so is its orthogonal complement $\mathscr{H}^{1\perp}$, since if $\xi \in \mathscr{H}^1$, $\eta \in \mathscr{H}^{1\perp}$, and $A \in \mathfrak{A}$,

$$(\xi \mid \pi(A)\,\eta) = (\pi(A^*)\,\xi \mid \eta) = 0.$$

Now if $\mathscr{H}^1$ is *closed* and invariant, and if $P_{\mathscr{H}^1}$ is the projection onto $\mathscr{H}^1$, we have

$$\begin{aligned} P_{\mathscr{H}^1}\,\pi(A)\,\xi &= P_{\mathscr{H}^1}\,\pi(A)\,P_{\mathscr{H}^1}\xi + P_{\mathscr{H}^1}\,\pi(A)\,(\mathbf{1} - P_{\mathscr{H}^1})\,\xi \\ &= \pi(A)\,P_{\mathscr{H}^1}\xi \end{aligned}$$

since $\pi(A)\,P_{\mathscr{H}^1}\xi \in \mathscr{H}^1$ and $\pi(A)\,(\mathbf{1} - P_{\mathscr{H}^1})\,\xi \in \mathscr{H}^{1\perp}$,

Thus, $P_{\mathscr{H}^1}$ commutes with $\pi(A)$ for all $A \in \mathfrak{A}$. Conversely, if $P_{\mathscr{H}^1}$ commutes with $\pi(A)$ for all $A \in \mathfrak{A}$, then $\mathscr{H}^1$ is invariant. Thus: A closed subspace $\mathscr{H}^1$ of $\mathscr{H}$ is invariant for π if and only if $P_{\mathscr{H}^1}$ commutes with $\pi(A)$ for all $A \in \mathfrak{A}$.

There is a particularly trivial kind of representation for any algebra $\mathfrak{A}$, that in which $\pi(A) = 0$ for all $A \in \mathfrak{A}$. We want to split any representation into a part of this kind and a part in which such pathology is entirely absent.

PROPOSITION III. F. 1 *Let $\mathfrak{A}$ be an algebra with involution, π a representation of $\mathfrak{A}$ on a Hilbert space $\mathscr{H}$. Then π splits into the direct sum of two orthogonal invariant subspaces $\mathscr{H}_1$ and $\mathscr{H}_2$; $\mathscr{H}_1$ is the closed linear span of*

$$\{\pi(A)\,\xi : A \in \mathfrak{A};\ \xi \in \mathscr{H}\},$$

and

$$\mathscr{H}_2 = \{\xi \in \mathscr{H} : \pi(A)\,\xi = 0 \quad \text{for all} \quad A \in \mathfrak{A}\}.$$

If $\mathfrak{A}$ has a identity, then $\pi(\mathbf{1})$ is the projection onto $\mathscr{H}_1$.

Proof Let $\mathscr{H}_1$ and $\mathscr{H}_2$ be defined as in the statement of the proposition. We have to show that $\mathscr{H}_2 = \mathscr{H}_1^{\perp}$. First note that, for a fixed A, the orthogonal complement of the range of $\pi(A)$ is the null space of $\pi(A^*)$. This is true since η is orthogonal to the range of $\pi(A)$ if and only if $(\eta \mid \pi(A)\,\xi) = 0$ for all $\xi \in \mathscr{H}$, which is true if and only if $(\pi(A^*)\,\eta \mid \xi) = 0$ for all $\xi \in \mathscr{H}$, which is true if and only if $\pi(A^*)\,\eta = 0$. Thus, the orthogonal complement of $\mathscr{H}_1$ is the intersection over $A \in \mathfrak{A}$ of the null space of $\pi(A^*)$, and this is what we have called $\mathscr{H}_2$. It is clear that $\mathscr{H}_2$ is invariant, so $\mathscr{H}_1$ is also. If $\mathfrak{A}$ has an identity, then $\pi(\mathbf{1})$ is the identity on $\mathscr{H}_1$ and zero on $\mathscr{H}_2$, so $\pi(\mathbf{1})$ is the projection onto $\mathscr{H}_1$.

With the notation of the above proposition, we say that π is *non-degenerate* if $\mathscr{H}_2 = \{0\}$. A vector ξ is said to be a *cyclic vector* for π if $\{\pi(A)\,\xi : A \in \mathfrak{A}\}$ is dense in $\mathscr{H}$ and π is said to be a *cyclic representation* if it admits a cyclic vector. From the above proposition it is clear that a cyclic representation is non-degenerate. The converse is not true, but we have:

PROPOSITION III. F. 2 *Let $\mathfrak{A}$ be an algebra with involution. Then every non-degenerate representation of $\mathfrak{A}$ is a direct sum of cyclic representations.*
To prove this, we need the following:

PROPOSITION III. F. 3 *Let $\mathfrak{A}$ be an algebra with involution, π a non-degenerate representation of $\mathfrak{A}$ on $\mathscr{H}$, and $\xi \in \mathscr{H}$. Then $\overline{\pi(\mathfrak{A})\,\xi}$ is a closed invariant subspace of $\mathscr{H}$ containing ξ.*

Proof of III. F. 3 It is clear that $\overline{\pi(\mathfrak{A})\,\xi}$ is invariant; let P be the projection onto this subspace. Then P commutes with $\pi(A)$ for all $A \in \mathfrak{A}$, so $0 = (\mathbf{1} - P)\pi(A)\xi = \pi(A)\,(\mathbf{1} - P)\xi$, so $(\mathbf{1} - P)\xi = 0$ by the non-degeneracy of π, so $\xi \in \overline{\pi(\mathfrak{A})\xi}$

Proof of III. F. 2 We use Zorn's Lemma. Consider the collection of all sets $\{\mathscr{H}_i\}$ of pairwise orthogonal closed invariant non-zero subspaces of $\mathscr{H}$ such that the restriction of π to each $\mathscr{H}_i$ has a cyclic vector. Order this collection by inclusion. It is nearly obvious that the hypotheses of Zorn's Lemma are satisfied. Thus, there exists a maximal collection $\{\mathscr{H}_i\}_{i\in I}$. The maximality implies that there is no non-zero cyclic subspace orthogonal to all the $\mathscr{H}_i$. By III. F. 3, this means that $\mathscr{H} = \bigoplus_{i\in I} \mathscr{H}_i$. (Otherwise, there would exist $\xi \neq 0$ orthogonal to all the $\mathscr{H}_i$'s; then $\overline{\pi(\mathfrak{A})\,\xi}$ would be a cyclic subspace orthogonal to all the $\mathscr{H}_i$'s.)

If $\mathfrak{A}$ is an algebra with involution and π is a representation of $\mathfrak{A}$ on a Hilbert space $\mathscr{H}$, we say that π is *irreducible* if the only closed invariant subspaces for π are $\{0\}$ and $\mathscr{H}$. The term "topologically irreducible" is sometimes used for this concept, the term "irreducible" being defined to mean that the only invariant linear subspaces, closed or not, are $\{0\}$ and $\mathscr{H}$. We will not have occasion to use the stronger notion; anyway, it turns out that, for $\mathfrak{A}$ a C^* algebra, they are equivalent. (See Dixmier, C^*A, Corollaire 2.8.4, p. 45).

PROPOSITION III. F. 4 *Let $\mathfrak{A}$ be an algebra with involution, π a representation of $\mathfrak{A}$ on a Hilbert space $\mathscr{H}$. Then the following are equivalent.*

i) *π is irreducible.*

ii) *The only orthogonal projections on $\mathscr{H}$ commuting with $\pi(A)$ for all A are 0 and $\mathbf{1}$.*

iii) *The only bounded operators on $\mathscr{H}$ commuting with $\pi(A)$ for all A are scalar multiples of $\mathbf{1}$.*

If π is non-degenerate, these are all equivalent to

iv) *Every non-zero vector $\xi \in \mathscr{H}$ is a cyclic vector for π.*

Proof i) and ii) are equivalent, since we have already seen that a closed subspace of $\mathscr{H}$ is invariant for π if and only if its orthogonal projection commutes with $\pi(A)$ for all A. Clearly, iii) implies ii); we now show that ii) implies iii). Thus, let ii) hold, and let T be a bounded operator commuting with all $\pi(A)$'s. Then $[\pi(A), T^*] = [T, \pi(A^*)]^* = 0$ for all A, so T^* also commutes with all $\pi(A)$'s. Hence, $\dfrac{T+T^*}{2}$ and $\dfrac{T-T^*}{2i}$ also commute with all $\pi(A)$'s so if we can show that, if T is a self-adjoint operator commuting with all $\pi(A)$'s, then T is a scalar multiple of $\mathbf{1}$, we are through. But if T is self-adjoint and commutes with all $\pi(A)$'s, then the spectral projections of T must commute with all $\pi(A)$'s; hence, by ii), must all be zero or $\mathbf{1}$. This implies that T is a scalar multiple of $\mathbf{1}$.

Now let π be non-degenerate and irreducible, and let $\xi \in \mathscr{H}$, $\xi \neq 0$. By Proposition III. F. 3. $\overline{\pi(\mathfrak{A})\,\xi}$ is a non-zero closed invariant subspace

hence, must be all of $\mathscr{H}$, so ξ is a cyclic vector for π. Thus, i) implies iv). Conversely, suppose π is not irreducible, and let ξ be a non-zero element of a proper closed invariant subspace, $\mathscr{H}_1$. Then $\overline{\pi(\mathfrak{A})\,\xi} \subset \mathscr{H}_1 \neq \mathscr{H}$, so ξ cannot be a cyclic vector for π. Thus, iv) implies i) and the proof is complete.

G Positive Linear Functionals and the Gelfand-Segal Construction

We saw in the last section that every non-degenerate representation of an algebra $\mathfrak{A}$ with involution may be written as a direct sum of cyclic representations. We are now going to show how to describe a cyclic representation of $\mathfrak{A}$ by a linear functional on $\mathfrak{A}$ of a special kind. Thus, the study of representations is reduced, in a certain sense, to the study of the so-called positive linear functionals on $\mathfrak{A}$. The definition of positive linear functionals is motivated by the following remark: Let $\mathfrak{A}$ be an algebra with involution, π a representation of $\mathfrak{A}$ on a Hilbert space $\mathscr{H}$, ξ a vector in $\mathscr{H}$. Define a linear functional ϕ on $\mathfrak{A}$ by $\phi(A) = (\xi \mid \pi(A)\,\xi)$. Then, for any $A \in \mathfrak{A}$, $\phi(A^*A) = (\xi \mid \pi(A^*A)\,\xi) = (\pi(A)\,\xi \mid \pi(A)\xi) \geqq 0$. We are thus led to define: A *positive linear functional* on an algebra $\mathfrak{A}$ with involution is a linear functional ϕ on $\mathfrak{A}$ such that $\phi(A^*A) \geqq 0$ for all $A \in \mathfrak{A}$.

The study of positive linear functionals is largely based on the remark that, if ϕ is a positive linear functional on $\mathfrak{A}$, then we can define something which is almost an inner product on $\mathfrak{A}$ by

$$\langle A \mid B \rangle = \phi(A^*B).$$

We have therefore to review some elementary facts about spaces with inner products.

Let E be a vector space over the complex numbers. A *sesquilinear form* on E is a mapping $(\xi, \eta) \mapsto \langle \xi \mid \eta \rangle$ from $E \times E$ to $\mathbb{C}$, such that

$$\langle \alpha\xi + \beta\eta \mid \zeta \rangle = \bar{\alpha}\langle \xi \mid \zeta \rangle + \beta\langle \eta \mid \zeta \rangle,$$
$$\langle \zeta \mid \alpha\xi + \beta\eta \rangle = \alpha\langle \zeta \mid \xi \rangle + \beta\langle \zeta \mid \eta \rangle$$

for $\alpha, \beta \in \mathbb{C}$, $\xi, \eta, \zeta \in E$. A sesquilinear form is *positive semi-definite* if $\langle \xi \mid \xi \rangle \geqq 0$ for all $\xi \in E$, and *positive definite* if $\langle \xi \mid \xi \rangle > 0$ for $\xi \in E$, $\xi \neq 0$. A positive-definite sesquilinear form is also called an *inner product*. A vector space equipped with a positive semi-definite sesquilinear form is called a *pre-Hilbert space*; if the sesquilinear form is positive definite, we speak of a *strict pre-Hilbert space*. Elementary arguments show that a positive semi-definite sesquilinear form is *hermitian*, i.e.

$$\langle \xi \mid \eta \rangle = \overline{\langle \eta \mid \xi \rangle}$$

and satisfies the Schwarz inequality

$\langle \xi, \eta \rangle^2 \leq \langle \xi | \xi \rangle \langle \eta | \eta \rangle$.

If we define $\|\xi\| = \sqrt{\langle \xi \mid \xi \rangle}$ then $\|\cdot\|$ is a semi-norm on E; the semi-norm is a norm if E is a strict pre-Hilbert space. Any strict pre-Hilbert space may be completed, i.e., linearly imbedded as a dense subspace of a Hilbert space in an inner-product preserving way.

There is also a standard way of constructing from a pre-Hilbert space E, $\langle \cdot | \cdot \rangle$, a strict pre-Hilbert space: Let

$$I = \{\xi \in E : \langle \xi \mid \xi \rangle = 0\}.$$

Since $\|\cdot\|$ is a seminorm, I is a linear subspace. If $\xi - \xi'$ and $\eta - \eta'$ belong to I,

$$|\langle \xi \mid \eta \rangle - \langle \xi' \mid \eta' \rangle| \leq |\langle \xi - \xi' \mid \eta \rangle| + |\langle \xi' \mid \eta - \eta' \rangle| \leq \|\eta\| \cdot \|\xi - \xi'\| + \|\xi'\| \cdot \|\eta - \eta'\| = 0.$$

Thus, $\langle \xi \mid \eta \rangle$ only depends on $\xi + I$, $\eta + I$. In other words, the sesquilinear form $\langle \cdot \mid \cdot \rangle$ may be regarded as mapping $E/I \times E/I$ to $\mathbb{C}$; this mapping is easily seen to be a scalar product on E/I. We shall speak of E/I with the scalar product constructed in this way as the strict pre-Hilbert space *associated with* E, $\langle \cdot \mid \cdot \rangle$.

Returning to positive linear functionals, we have:

PROPOSITION III. G. 1 *Let $\mathfrak{A}$ be an algebra with involution, ϕ a positive linear functional on $\mathfrak{A}$. Then, for $\xi, \eta \in \mathfrak{A}$,*

$$\phi(\xi^* \eta) = \overline{\phi(\eta^* \xi)};$$

$$|\phi(\xi^* \eta)|^2 \leq \phi(\xi^* \xi) \cdot \phi(\eta^* \eta).$$

If $\mathfrak{A}$ has an identity, then

$$\phi(\xi^*) = \overline{\phi(\xi)};$$

$$|\phi(\xi)|^2 \leq \phi(\xi^* \xi) \cdot \phi(\mathbf{1}).$$

Proof Everything follows from the fact that $\langle \xi \mid \eta \rangle = \phi(\xi^* \eta)$ is a positive semi-definite sesquilinear form.

We now come to a less trivial result, that positive linear functionals on Banach algebras with identity and involution are automatically continuous.

PROPOSITION III. G. 2 *Let $\mathfrak{A}$ be a Banach algebra with identity and involution, and let ϕ be a positive linear functional on $\mathfrak{A}$. Then ϕ is continuous and $\|\phi\| = \phi(\mathbf{1})$.*

Proof Let $\xi \in \mathfrak{A}$, and assume $\|\xi\| < 1$. We will prove that

$$\phi(\xi^* \xi) \leq \phi(\mathbf{1}).$$

Then, by the Schwarz inequality

$$|\phi(\xi)|^2 \leq \phi(\xi^*\xi)\cdot\phi(\mathbf{I}) \leq \phi(\mathbf{I})^2$$

if $\|\xi\| < 1$; thus, ϕ is continuous and $\|\phi\| \leq \phi(\mathbf{I})$. The other inequality is immediate; $\|\phi\| \geq \phi(\mathbf{I})$ since $\|\mathbf{I}\| = 1$.

To show that $\phi(\xi^*\xi) \leq \phi(\mathbf{I})$, we prove that $\mathbf{I} - \xi^*\xi$ has a positive square-root. Consider the series

$$\eta = \mathbf{I} - \frac{1}{2}\cdot(\xi^*\xi) - \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{(\xi^*\xi)^2}{2!} - \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{3}{2}\cdot\frac{(\xi^*\xi)^3}{3!} - \cdots$$

(We have inserted $\xi^*\xi$ for z in the Taylor series for $\sqrt{1-z}$ about $z = 0$). The series converges since $\sqrt{1-z}$ is analytic for $|z| < 1$ and since $\|\xi^*\xi\| < 1$. Also, η is self-adjoint since it is a sum of self-adjoint terms. Exactly the same calculation as that required to show that

$$\left(1 - \frac{1}{2}z - \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{z^2}{2!} - \cdots\right)^2 = 1 - z$$

for z a complex number of modulus less than 1 shows that

$$\eta^2 = \mathbf{I} - \xi^*\xi.$$

Thus

$$0 \leq \phi(\eta^2) = \phi(\mathbf{I}) - \phi(\xi^*\xi), \quad \text{so} \quad \phi(\mathbf{I}) \geq \phi(\xi^*\xi).$$

Remark For C^* algebras, one can prove by a different argument that every positive linear functional is continuous without the assumption that $\mathfrak{A}$ has on identity. See Dixmier, C^*A, 2.1.8.

We define as *state* of a Banach algebra with involution to be a continuous positive linear functional ϕ of norm one. If $\mathfrak{A}$ has an identity, we may remove the requirement that ϕ be continuous, and replace the requirement $\|\phi\| = 1$ by $\phi(\mathbf{I}) = 1$. Notice a peculiar property of the norms of positive linear functionals: If $\mathfrak{X}$ is an arbitrary Banach space, and if ϕ, ψ are linear functionals on $\mathfrak{X}$, then $\|\phi + \psi\|$ is normally strictly less than $\|\phi\| + \|\psi\|$. For example, if $\mathscr{H}$ is a Hilbert space, then

$$\|\phi + \psi\| = \|\phi\| + \|\psi\|$$

only if ϕ and ψ are linearly dependent. However, if ϕ and ψ are positive linear functionals on a Banach algebra with identity and involution, then $\phi + \psi$ is also positive, and

$$\|\phi + \psi\| = (\phi + \psi)(\mathbf{I}) = \phi(\mathbf{I}) + \psi(\mathbf{I}) = \|\phi\| + \|\psi\|.$$

By this remark, the set of states of a Banach algebra with identity is convex.

Using the above proposition we can prove a result about the continuity of representations.

PROPOSITION III G. 3 *Let $\mathfrak{A}$ be a Banach algebra with involution, π a morphism of $\mathfrak{A}$ into the algebra of all linear operators on a pre-Hilbert space $\mathscr{H}$, such that, for $\xi, \eta \in \mathscr{H}$*

$$\langle \xi \mid \pi(A)\,\eta \rangle = \langle \pi(A^*)\,\xi \mid \eta \rangle.$$

Then $\|\pi(A)\,\xi\| \leqq \|A\|\,\|\xi\|$, for all $\xi \in \mathscr{H}$.

Proof If $\mathfrak{A}$ has no identity, extend π to $\tilde{\mathfrak{A}}$ by defining $\pi(\mathbf{I})$ to be the identity operator on $\mathscr{H}$. Thus, we may assume that $\mathfrak{A}$ has an identity. Now, for any $\xi \in \mathscr{H}$, define a linear functional ϕ_ξ on $\mathfrak{A}$ by $\phi_\xi(A) = \langle \xi \mid \pi(A)\,\xi \rangle$. ϕ_ξ is positive, so, for any $A \in \mathfrak{A}$,

$$\begin{aligned}\|\pi(A)\,\xi\|^2 &= \langle \xi \mid \pi(A^*A)\,\xi \rangle = \phi_\xi(A^*A) \leqq \|A\|^2\,(\phi_\xi(\mathbf{I})) \\ &= \|A\|^2\,\|\xi\|^2.\end{aligned}$$

We are now ready to construct a representation associated with a positive functional. The idea is as follows: We begin with a positive linear functional ϕ on an algebra $\mathfrak{A}$ with involution and identity. We make $\mathfrak{A}$ itself into a pre-Hilbert space by $\langle A \mid B \rangle = \phi(A^*B)$. Let $I = \{A \in \mathfrak{A} : \phi(A^*A) = 0\}$. For each $A \in \mathfrak{A}$, we define a linear mapping L_A of $\mathfrak{A}$ into itself by $L_A B = A \cdot B$.

Evidently, $L_A I_B = I_{AB}$ and

$$\langle L_A B \mid C \rangle = \langle B \mid L_{A^*} C \rangle.$$

We want to show that each I_A induces a linear mapping on the quotient space $\mathfrak{A}/I$. To do this, we have to show that

$$\phi(B^*B) = 0 \quad \text{implies} \quad \phi(B^*A^*AB) = 0.$$

To do this, we note that $\phi_B(\cdot) = \phi(B^*(\cdot)\,B)$ is a positive linear functional, and hence, by the Schwarz inequality, we have:

$$\begin{aligned}|\phi(B^*A^*AB)|^2 &= |\phi_B(A^*A)|^2 \leqq \phi_B(\mathbf{I}) \cdot \phi_B((A^*A)^2) \\ &= \phi(B^*B)\,\phi_B((A^*A)^2) = 0.\end{aligned}$$

Now let $\hat{\mathscr{H}}_\phi$ denote $\mathfrak{A}/I$, $\hat{\pi}_\phi(A)$ denote the linear operator induced on $\hat{\mathscr{H}}_\phi$ by L_A, and ξ_ϕ denote $\mathbf{I} + I \in \hat{\mathscr{H}}_\phi$. Then we have:

THEOREM III. G. 4 *Let ϕ be a positive linear functional on an algebra $\mathfrak{A}$ with involution and identity. Then there exist a strict pre-Hilbert space $\hat{\mathscr{H}}_\phi$, a morphism $\hat{\pi}_\phi$ from $\mathfrak{A}$ to the algebra of all linear mappings of $\hat{\mathscr{H}}_\phi$ into itself, and a vector $\xi_\phi \in \hat{\mathscr{H}}_\phi$ such that:*

i) $\phi(A) = (\xi_\phi \mid \hat{\pi}_\phi(A)\,\xi_\phi)$ *for all* $A \in \mathfrak{A}$.

ii) $(\hat{\pi}_\phi(A)\,\xi \mid \eta) = (\xi \mid \hat{\pi}_\phi(A^*)\,\eta)$ *for all* $A \in \mathfrak{A}$, $\xi, \eta \in \hat{\mathscr{H}}_\phi$.

iii) $\hat{\mathscr{H}}_\phi = \hat{\pi}_\phi(\mathfrak{A})\,\xi_\phi$.

Finally, these objects are unique up to unitary equivalence, i.e., if $\mathscr{H}'$, π' and ξ' also satisfy i), ii) *and* iii), *then there exists a mapping U of $\hat{\mathscr{H}}_\phi$ onto $\mathscr{H}'$, preserving inner products, such that*

$$U^{-1}\pi'(\cdot)\,U = \hat{\pi}_\phi(\cdot)$$

$$U\xi_\phi = \xi'.$$

Proof Conditions i), ii), iii) are straightforward verifications. To prove the uniqueness statement, show, again by a simple computation, that

$$\hat{\pi}_\phi(A)\,\xi_\phi \overset{U}{\mapsto} \pi'(A)\,\xi'$$

preserves inner products (and is therefore well-defined) and has the desired algebraic properties.

COROLLARY III. G. 5 *Let α be an automorphism of $\mathfrak{A}$ such that $\phi(\alpha A) = \phi(A)$ for all $A \in \mathfrak{A}$. Then there is a uniquely determined unitary operator $U_\phi(\alpha)$ on $\hat{\mathscr{H}}_\phi$ such that*

$$U_\phi(\alpha)\,\xi_\phi = \xi_\phi;\ U_\phi(\alpha)\,\hat{\pi}_\phi(A)\,U_{\phi(\alpha)}^{-1} = \hat{\pi}_\phi(\alpha A) \quad \text{for all} \quad A \in \mathfrak{A}.$$

Proof Apply the uniqueness part of the preceding theorem with $\hat{\mathscr{H}}_\phi = \mathscr{H}'$, $\xi_\phi = \xi'$, and $\pi_\phi \cdot \alpha = \pi'$.

We may regard $\hat{\mathscr{H}}_\phi$ as a dense linear subspace of its completion $\mathscr{H}_\phi$, and, if $\mathfrak{A}$ is a Banach algebra, Proposition III. G. 3, assures us that each $\hat{\pi}_\phi(A)$ may be extended to a bounded linear operator $\pi_\phi(A)$ on $\mathscr{H}_\phi$. It is worth proving the theorem in this form, allowing for representations by unbounded operators, since it gives the Wightman Reconstruction Theorem in field theory. Let $\mathscr{I}(\mathscr{S})$ be the set of all sequences $(f^{(n)})$, $n = 0, 1, 2, \ldots$ such that $f^{(0)}$ is a complex number, $f^{(n)}$ is a smooth, rapidly decreasing complex-valued function on $(\mathbb{R}^4)^n$ and $f^{(n)} = 0$ for all but finitely many n's. $\mathscr{I}(\mathscr{S})$ is a vector space in an obvious way, and we define

$$(f \cdot g)^{(n)}(x_1, \ldots, x_n) = \sum_{j=0}^{n} f^{(j)}(x_1, \ldots, x_j)\, g^{(n-j)}(x_{j+1}, \ldots, x_n)$$

$$(f^*)^{(n)}(x_1, \ldots, x_n) = \bar{f}(x_n, \ldots, x_1).$$

These operations make $\mathscr{I}(\mathscr{S})$ into an algebra with involution and identity (but not a Banach algebra). Representations π of $\mathscr{I}(\mathscr{S})$ with reasonable continuity properties may be thought of as specified by an operator-valued distribution $\phi(x)$ by

$$\pi((f)) = \sum_j \int \cdots \int dx_1 \ldots dx_j\, f^{(j)}(x_1, \ldots, x_n)\, \phi(x_1) \ldots \phi(x_n).$$

Thus, a scalar Wightman field defines a representation of $\mathscr{I}(\mathscr{S})$. Conversely, a set of Wightman functions $W_n(x_1, \ldots x_n)$ defines a positive linear criteria

tional on $\mathcal{J}(\mathcal{S})$ by

$$\phi(f) = \sum_j \int \cdots \int dx_1 \ldots dx_j f^{(j)}(x_1, \ldots, x_n) \mathscr{W}_n(x_1, \ldots, x_n),$$

so the above theorem enables us to reconstruct a Wightman field given a set of vacuum expectation values.

We now restate the above theorem specialized to the case in which $\mathfrak{A}$ is a Banach algebra with involution and identity, taking advantage of the fact that the $\pi_\phi(A)$'s are automatically continuous.

THEOREM III. G. 5 *Let $\mathfrak{A}$ be a Banach algebra with involution and identity, and let ϕ be a positive linear functional on $\mathfrak{A}$. Then there exist a Hilbert space $\mathscr{H}_\phi$, a representation π_ϕ of $\mathfrak{A}$ on $\mathscr{H}_\phi$, and a cyclic vector ξ_ϕ for π_ϕ such that*

$$\phi(A) = (\xi_\phi \mid \pi_\phi(A)\, \xi_\phi)$$

for all $A \in \mathfrak{A}$. If $\mathscr{H}', \pi', \xi'$ are another triple of objects satisfying these conditions, there exists a unique unitary operator U mapping $\mathscr{H}$ to $\mathscr{H}'$ such that

$$U\xi_\phi = \xi'; \quad U\pi_\phi(A)\, U^{-1} = \pi'(A) \quad \text{for all} \quad A \in \mathfrak{A}.$$

If α is an automorphism of $\mathfrak{A}$ such that $\phi(\alpha A) = \phi(A)$ for all $A \in \mathfrak{A}$, there is a unique unitary operator $U_\phi(\alpha)$ on $\mathscr{H}_\phi$ such that

$$U_\phi(\alpha)\, \xi_\phi = \xi_\phi; \; U_\phi(\alpha)\, \pi_\phi(A)\, U_\phi(\alpha)^{-1} = \pi_\phi(\alpha A).$$

We will speak of $(\mathscr{H}_\phi, \pi_\phi, \xi_\phi)$ as the *canonical cyclic representation of $\mathfrak{A}$ associated with ϕ*; the construction of this representation is called the *Gelfand-Segal construction.*

H Pure States and Irreducible Representations

We continue our investigation of the relation between positive linear functionals and representations by determining when the cyclic representation associated with a positive functional is irreducible. Suppose we start with a functional ϕ and construct the cyclic representation $(\mathscr{H}_\phi, \pi_\phi, \xi_\phi)$, and suppose that this representation is not irreducible. Let P be a nontrivial projection commuting with $\pi_\phi(A)$ for all $A \in \mathfrak{A}$. Consider the functional

$$\varrho(A) = (P\xi_\phi \mid \pi_\phi(A)\, P\xi_\phi).$$

This is certainly a positive, and

$$\begin{aligned}\phi(A) - \varrho(A) &= (\xi_\phi \mid \pi_\phi(A)\, \xi_\phi) - (P\xi_\phi \mid P\pi_\phi(A)\, \xi_\phi) = (\xi_\phi \mid \pi_\phi(A)\, \xi_\phi) \\ &\quad - (P\xi_\phi \mid \pi_\phi(A)\, \xi_\phi) - ((\mathbf{1} - P)\, \xi_\phi \mid \pi_\phi(A)\, \xi_\phi) \\ &= ((\mathbf{1} - P)\, \xi_\phi \mid \pi_\phi(A)\, (\mathbf{1} - P)\, \xi_\phi),\end{aligned}$$

so $\phi(A) - \varrho(A)$ is again a positive linear functional on $\mathfrak{A}$. Next, we claim that ϱ is not simply a numerical multiple of ϕ. Suppose the contrary is true, i.e., $\varrho(A) = (\lambda)\phi(A)$ for all $A \in \mathfrak{A}$. Using the fact that P commutes with $\pi_\phi(A)$ for all A we compute that

$$(\pi_\phi(B)\,\xi_\phi \mid P\pi_\phi(A)\,\xi_\phi) = \varrho(B^*A) = \lambda\phi(B^*A) = \lambda \cdot (\pi_\phi(B)\,\xi_\phi \mid \pi_\phi(A)\,\xi_\phi).$$

Since vectors of the form $\pi_\phi(B)\,\xi_\phi (B \in \mathfrak{A})$ are dense in $\mathscr{H}$, this equality implies that $P = \lambda\mathbf{I}$, which contradicts the assumption that P is a non-trivial projection.

We now make two definitions: If ϕ and ϱ are positive linear functionals on an algebra $\mathfrak{A}$ with involution and identity, we say that ϕ *majorizes* $\varrho (\phi \geq \varrho)$ if $\phi - \varrho$ is a positive linear functional. We say that a positive linear functional ϕ is *pure* if the only positive functionals majorized by ϕ are scalar multiples of ϕ. What the above argument shows is that if the representation π_ϕ is not irreducible, then ϕ is not pure.

Let us next try to prove the converse: Let ϱ be a positive linear functional majorized by ϕ. We note that

$$|\varrho(B^*A)|^2 \leq \varrho(B^*B)\,\varrho(A^*A) \leq \phi(B^*B)\,\phi(A^*A)$$
$$= \|\pi_\phi(B)\,\xi_\phi\|^2 \cdot \|\pi_\phi(A)\,\xi_\phi\|^2.$$

Thus, there exists a unique linear operator T such that

$$(\pi_\phi(B)\,\xi_\phi \mid T\pi_\phi(A)\,\xi_\phi) = \varrho(B^*A).$$

If ϱ is not a scalar multiple of ϕ, then T is not a scalar multiple of $\mathbf{I}$. Moreover, T is positive

$$((\pi_\phi(A)\,\xi_\phi \mid T\pi_\phi(A)\,\xi_\phi) = \varrho(A^*A) \geq 0),$$

and $\|T\| \leq 1$. For $A, B, C \in \mathfrak{A}$,

$$(\pi_\phi(B)\,\xi_\phi \mid T\pi_\phi(C)\,\pi_\phi(A)\,\xi_\phi) = \varrho(B^*CA)$$
$$= (\pi_\phi(C^*)\,\pi_\phi(B)\,\xi_\phi \mid T\pi_\phi(A)\,\xi_\phi),$$

so $[T, \pi_\phi(C)] = 0$. Thus, T commutes with $\pi_\phi(\mathfrak{A})$. But if any operator which is not a scalar multiple of $\mathbf{I}$ commutes with $\pi_\phi(\mathfrak{A})$, π_ϕ is not irreducible. We have thus shown that, if the functional ϕ is not pure, the representation π_ϕ is not irreducible. We have, in fact, proved the following proposition.

PROPOSITION III. H. 1 *Let $\mathfrak{A}$ be a Banach algebra with involution and identity, let ϕ be a positive linear functional on $\mathfrak{A}$, and let $(\mathscr{H}_\phi, \pi_\phi, \xi_\phi)$ be the associated cyclic representation. Then π_ϕ is irreducible if and only if ϕ is a pure positive functional. Moreover, there is a one-one correspondence between positive functionals on $\mathfrak{A}$ majorized by ϕ and positive operators on $\mathscr{H}_\phi$ of norm not greater than one commuting with $\pi_\phi(A)$ for all $A \in \mathfrak{A}$, the correspondence being defined to associate with the operator T the functional*

$$A \leftrightarrow (T\xi_\phi \mid \pi_\phi(A)\,\xi_\phi).$$

Suppose, now, we consider states of $\mathfrak{A}$ (i.e., positive linear functionals of norm one) and ask when a state ϕ is pure. We claim that this is the case if and only if ϕ is an extremal point of the set of states of $\mathfrak{A}$. Suppose first that ϕ is not extremal; then we can write

$$\phi = \tfrac{1}{2}\phi_1 + \tfrac{1}{2}\phi_2,$$

where ϕ_1 is a state different from ϕ. Then $\frac{1}{2}\phi_1$ is a positive functional on $\mathfrak{A}$ majorized by ϕ but not a scalar multiple of ϕ, so ϕ is not a pure positive functional. Conversely, suppose ϕ is not a pure positive functional, and let ϱ be a positive functional majorized by ϕ but not a scalar multiple of ϕ. Then:

$$\phi(\cdot) = \varrho(\cdot) + (\phi - \varrho)(\cdot) = \varrho(\mathbf{1}) \cdot \frac{\varrho(\cdot)}{\varrho(\mathbf{1})} + (\phi(\mathbf{1}) - \varrho(\mathbf{1})) \frac{(\phi - \varrho)(\cdot)}{\phi(\mathbf{1}) - \varrho(\mathbf{1})},$$

and since $\dfrac{\varrho(\cdot)}{\varrho(\mathbf{1})}$ and $\dfrac{(\phi - \varrho)(\cdot)}{(\phi - \varrho)(\mathbf{1})}$ are states of $\mathfrak{A}$, we have displayed ϕ as a non-trivial convex combination of states of $\mathfrak{A}$ and hence have shown that ϕ is not an extremal point of the set of states of $\mathfrak{A}$. We thus have:

COROLLARY III. II. 2 *Let $\mathfrak{A}$ be a Banach algebra with involution and identity, and let ϕ be a state of $\mathfrak{A}$. Then π_ϕ is irreducible if and only if ϕ is an extremal point of the set of states of $\mathfrak{A}$.*

Remark The set of states of $\mathfrak{A}$ is a weak-* closed subset of the unit ball of the dual of $\mathfrak{A}$; hence, is weak-* compact. If $\mathfrak{A}$ is separable, then the set of states of $\mathfrak{A}$ is metrizable in the weak-* topology. We may therefore use Choquet theory to decompose general states of $\mathfrak{A}$ into integrals of pure states. His decomposition is closely connected with the decomposition of the corresponding cyclic representation as a direct integral of irreducible representations. The decomposition of states into pure states is, however, usually non-unique. It may be shown that the set of states of a C^* algebra $\mathfrak{A}$ with identity is a simplex only if $\mathfrak{A}$ is commutative. (Dixmier, C^*A, Exercise 2.12.17, p. 57).

I Morphisms of C^* Algebras

In this section, we will prove some technical results about morphisms of C^* algebras. Specifically, we will show that any morphism of C^* algebras is norm-decreasing (in particular, continuous), and that any injective morphism is norm preserving. These two results imply that the norm on a C^* algebra is unique.

PROPOSITION III. I. 1 *Let $\mathfrak{A}$ be a Banach algebra with involution, $\mathfrak{B}$ a C^* algebra, and π a morphism from $\mathfrak{A}$ to $\mathfrak{B}$. Then*

$$\|\pi(A)\| \leq \|A\|$$

for all $A \in \mathfrak{A}$.

Proof We can assume that $\mathscr{B}$ has an identity (If not, adjoin one). If $\mathfrak{A}$ does not have an identity, or if it does have an identity but it is not sent by π to the identity of $\mathscr{B}$, adjoin an identity and extend π to send $\mathbf{I}_{\mathfrak{A}}$ to $\mathbf{I}_{\mathscr{B}}$. Thus, we can assume that $\mathfrak{A}$ and $\mathscr{B}$ both have identities and that π sends $\mathbf{I}_{\mathfrak{A}}$ to $\mathbf{I}_{\mathscr{B}}$. Let $A \in \mathfrak{A}$; we claim that the spectral radius of $\pi(A)$ is not greater than the spectral radius of A:

$$\varrho(\pi(A)) \leq \varrho(A).$$

Indeed, if $|\lambda| > \varrho(A)$, then $\lambda\mathbf{I}_{\mathfrak{A}} - A$ is invertible in $\mathfrak{A}$; since π sends identity to identity, $\pi((\lambda\mathbf{I}_{\mathfrak{A}} - A)^{-1})$ is an inverse for $\lambda\mathbf{I}_{\mathscr{B}} - \pi(A)$, so $\lambda \notin \sigma(\pi(A))$. Now for $A \in \mathfrak{A}$,

$$\|A\|^2 \geq \|(A^*A)\| \geq \varrho(A^*A) \geq \varrho(\pi(A^*A)).$$

But $\pi(A^*A)$ is a self-adjoint, and hence normal, element of the C^* algebra $\mathscr{B}$, so its norm is equal to its spectral radius, i.e.,

$$\varrho(\pi(A^*A)) = \|\pi(A^*A)\| = \|\pi(A)^*\,\pi(A)\| = \|\pi(A)\|^2.$$

Combining these relations gives

$$\|A\|^2 \geq \|\pi(A)\|^2.$$

PROPOSITION III. I. 2 *Let $\mathfrak{A}$, $\mathscr{B}$ be C^* algebras, and let π be an injective morphism of $\mathfrak{A}$ into $\mathscr{B}$. Then $\|\pi(A)\| = \|A\|$ for all $A \in \mathfrak{A}$.*

Proof We have already proved this result if $\mathfrak{A}$, $\mathscr{B}$ are both commutative C^* algebras with identity and if π sends identity to identity (Proposition III. D. 3) We will reduce the general statement to the commutative case by using relation $\|A^*A\| = \|A\|^2$. We first have to fix the identities. If $\mathfrak{A}$ has an identity, then $\pi(\mathbf{I}_{\mathfrak{A}})$ is an identity for $\overline{\pi(\mathfrak{A})}$; replacing $\mathscr{B}$ by $\overline{\pi(\mathfrak{A})}$ if necessary, we can assume that $\mathscr{B}$ has an identity and that π sends identity to identity. Suppose $\mathfrak{A}$ has no identity; then, since π is injective, $\mathbf{I}_{\mathscr{B}}$ (which it may be necessary to adjoin) cannot belong to $\pi(\mathfrak{A})$; then we can extend π to $\tilde{\mathfrak{A}}$ by $\pi(\mathbf{I}_{\tilde{\mathfrak{A}}}) = \mathbf{I}_{\mathscr{B}}$ without spoiling the injectiveness of π. Thus, we may assume that $\mathfrak{A}$ and $\mathscr{B}$ have identities and that π sends identity to identity.

Now let A be a self-adjoint element of $\mathfrak{A}$; then π defines an injective homomorphism from the commutative C^* sub-algebra of $\mathfrak{A}$ generated by A and $\mathbf{I}_{\mathfrak{A}}$ to the commutative C^* sub-algebra of $\mathscr{B}$ generated by $\pi(A)$ and $\mathbf{I}_{\mathscr{B}}$. Since we already know the result for commutative C algebras,

$$\|\pi(A)\| = \|A\|$$

for A self-adjoint. Now let A be as general element of $\mathfrak{A}$; then $\|A\|^2 = \|A^*A\| = \|\pi(A^*A)\| = \|\pi(A)\|^2$, so the result is proved.

We can now see that the norm on a C^* algebra is uniquely determined. The crudest form of this assertion is the remark that, if $\mathfrak{A}$ is an algebra with

involution and if $\|\cdot\|$, $\|\cdot\|'$ are two norms making $\mathfrak{A}$ into a C^*-algebra, then

the identity mapping is a morphism from $\mathfrak{A}$ with $\|\ \|$ to $\mathfrak{A}$ with $\|\ \|'$ and hence, by the first proposition, $\|A\|' \leq \|A\|$. Interchanging the roles of $\|\ \|$ and $\|\ \|'$ gives the opposite inequality and hence proves that $\|\ \| = \|\ \|'$. The second proposition enables us to prove a somewhat more subtle form of the uniqueness of the norm. Let $\mathfrak{A}$ be a C^* algebra with a norm $\|\ \|$, and let $\|\ \|'$ be another norm making $\mathfrak{A}$ into a normed algebra with involution and satisfying $\|A^*A\|' = (\|A\|')^2$, *but with respect to which* $\mathfrak{A}$ *is not necessarily complete.* Then the completion $\mathfrak{A}'$ of $\mathfrak{A}$ with respect to the prime norm is a C^* algebra; the identity mapping is an injective morphism from $\mathfrak{A}$ to $\mathfrak{A}'$, and is hence norm-preserving by the second proposition. Thus, $\|\ \|' = \|\ \|$. In other words, if $\mathfrak{A}$ is an algebra with involution, and if it admits one norm making it into a C^* algebra, then it admits no other norm with the right algebraic properties to make it a C^* algebra, i.e., the topology of $\mathfrak{A}$ is built into its algebraic properties.

There is another consequence of the above propositions which is sometimes useful. Let $\pi : \mathfrak{A} \to \mathfrak{B}$ be a morphism of C^* algebras. Then, since π is continuous (first proposition), $\ker(\pi)$ is closed in $\mathfrak{A}$, and, by general algebra, $\ker(\pi)$ is a self-adjoint two-sided ideal of $\mathfrak{A}$. We have already quoted the fact that the quotient of a C^* algebra by a closed, self-adjoint, two-sided ideal is a C^* algebra. π induces an injective morphism $\tilde{\pi} : \mathfrak{A}/\ker(\pi) \to \mathfrak{B}$. By the second proposition, $\tilde{\pi}$ is norm-preserving, so its image is complete and hence closed in $\mathfrak{B}$. But the image of $\tilde{\pi}$ is the same as the image of π, i.e., the image of any morphism of C^* algebras is closed.

J Positive Elements of C^* Algebras

In this section, we will show that every C^* algebra is isomorphic to a norm-closed self-adjoint algebra of operators on a Hilbert space. By Proposition III. I. 2, it suffices to show that every C^* algebra admits an injective (i.e., faithful) representation. The main ingredient of the proof is the fact that, if A is any element of a C^* algebra, then A^*A has spectrum contained in the positive real axis. We need first:

LEMMA III. J. 1 *Let* $\mathfrak{A}$ *be an algebra with identity,* $A, B \in \mathfrak{A}$. *If* $\lambda \neq 0$ *is in the resolvent set of* $A \cdot B$, *then* λ *is in the resolvent set of* $B \cdot A$ (*i.e., the spectrum of* $A \cdot B$ *is the same as the spectrum of* $B \cdot A$ *except possibly for* 0).

Proof We can assume $\lambda = 1$ by scaling, say, A. Thus, we have to show that $\mathbf{1} - BA$ is invertible if $\mathbf{1} - AB$ is. This we do by exhibiting an explicit formula:

$$(\mathbf{1} - BA)^{-1} = \mathbf{1} + B(\mathbf{1} - AB)^{-1}A.$$

It is trivial to verify that this formula is correct; the formula may be remembered by noting that, if $\|A\| < 1$, $\|B\| < 1$, then

$$\begin{aligned}(\mathbf{1} - BA)^{-1} &= \sum_{n=0}^{\infty} (BA)^n = \mathbf{1} + B(\mathbf{1} + AB + (AB)^2 + \ldots) \cdot A \\ &= \mathbf{1} + B(\mathbf{1} - AB)^{-1}A.\end{aligned}$$

Now if $\mathfrak{A}$ is a C^* algebra, we will say that an element A of $\mathfrak{A}$ is *positive* ($A \geqq 0$) if A is self-adjoint and if the spectrum of A is contained in $[0, \infty)$. The set of positive elements of $\mathfrak{A}$ will be denoted by $\mathfrak{A}_+$.

Remark Let A be a self-adjoint element of $\mathfrak{A}$, and let $\mathfrak{B}$ be the sub C^* algebra of $\mathfrak{A}$ generated by A and $\mathbf{I}$. We claim that A is positive as an element of $\mathfrak{A}$ if and only if it is positive as an element of $\mathfrak{B}$, and, more generally, that the spectrum of A as an element of $\mathfrak{A}$ is the same as its spectrum as an element of $\mathfrak{B}$. Certainly, the spectrum of A as an element of $\mathfrak{A}$ is contained in the spectrum of A as an element of $\mathfrak{B}$, since an inverse for $(\lambda \mathbf{I} - A)$ in $\mathfrak{B}$ is an inverse in $\mathfrak{A}$. In particular, the spectrum in $\mathfrak{A}$ is contained in the real axis. Let λ be a point of the spectrum of A as an element of $\mathfrak{B}$. Realizing $\mathfrak{B}$ as an algebra of continuous functions, we see that

$$\lim_{\varepsilon \to 0} \|((\lambda + i\varepsilon)\mathbf{I} - A)^{-1}\| = \infty.$$

But if λ were not in the spectrum of A as an element of $\mathfrak{A}$, we would have to have $\lim_{\varepsilon \to 0} ((\lambda + i\varepsilon)\mathbf{I} - A)^{-1} = (\lambda \mathbf{I} - A)^{-1}$; this is impossible by the above, so λ is in the spectrum of A as an element of $\mathfrak{A}$, i.e., the two spectra are identical.

PROPOSITION III. J. 2 *Let $\mathfrak{A}$ be a C^* algebra with identity. The set of positive elements of $\mathfrak{A}$ is a convex cone in the set of self-adjoint elements of $\mathfrak{A}$ with $\mathbf{I}$ as an interior point, and $\mathfrak{A}_+ \cap (-\mathfrak{A}_+) = \{0\}$.*

Proof It is clear that, if $\lambda > 0$ and $A \in \mathfrak{A}_+$ then $\lambda A \in \mathfrak{A}_+$. Thus, to prove that $\mathfrak{A}_+$ is a convex cone, it suffices to show that

$$\{A \in \mathfrak{A}_+ : \|A\| \leqq 1\} \text{ is convex.}$$

Now let A be self-adjoint. We can identify the C^* algebra generated by $\mathbf{I}$ and A as the algebra of continuous functions on a compact space. It is then clear that, if $\|A\| \leqq 1$, then A is positive if and only if $\|A - \mathbf{I}\| \leqq 1$. The set of such A's is clearly convex, so $\{A \in \mathfrak{A}_+ : \|A\| \leqq 1\}$ is convex. Furthermore $\mathbf{I}$ is an interior point of $\mathfrak{A}_+$ in the set of self-adjoint elements of $\mathfrak{A}$. Finally, if $A \in \mathfrak{A}_+ \cap (-\mathfrak{A}_+)$, then since $\sigma(A) = \{0\}$ and A is self-adjoint, $A = 0$.

The key technical result is now the following.

PROPOSITION III. J. 3 *Let $\mathfrak{A}$ be a C^* algebra with identity. An element of $\mathfrak{A}$ is positive if and only if it can be written as A^*A, with $A \in \mathfrak{A}$.*

Proof The "only if" is trivial; any positive element of $\mathfrak{A}$ has a positive square root. Thus, what we have to show is that $A^*A \in \mathfrak{A}_+$ for any $A \in \mathfrak{A}$. This is clear if A is self-adjoint. In the general case, by realizing the subalgebra generated by A^*A and $\mathbf{I}$ as an algebra of continuous functions, we see that we can write

$$A^*A = B - C, \quad \text{with} \quad B, C \in \mathfrak{A}_+, \; B \cdot C = C \cdot B = 0;$$

we want to show that $C = 0$. We have:

$$(AC)^* AC = C(B - C) C = -C^3 \in -\mathfrak{A}_+.$$

We can also write

$AC = S + iT$, S and T self-adjoint, so

$(AC)^* (AC) = (S - iT)(S + iT) = S^2 + T^2 + i[S, T]$.

$(AC)(AC)^* = S^2 + T^2 - i[S, T]$, so

$(AC)(AC)^* = -(AC)^* (AC) + 2(S^2 + T^2) \in \mathfrak{A}_+$ since

$(AC)^* (AC) \in -\mathfrak{A}_+$; $S^2 \in \mathfrak{A}_+$, $T^2 \in \mathfrak{A}_+$, and $\mathfrak{A}_+$ is a convex cone.

Thus, the spectrum of $(AC)(AC)^*$ is contained in $[0, \infty)$, By Lemma III. J. 1., this implies that the spectrum of $(AC)^* (AC)$ is contained in $[0, \infty]$ i.e., that $(AC)^* (AC) \in \mathfrak{A}_+$. Since we already know that $(AC)^* (AC) = -C^3 \in -\mathfrak{A}_+$; we have $-C^3 \in \mathfrak{A}_+ \cap -\mathfrak{A}_+$, so $C^3 = 0$, so $C = 0$.

Once we have this result, we know that a linear functional on $\mathfrak{A}$ is positive if and only if it is positive on $\mathfrak{A}_+$. We can therefore prove:

PROPOSITION III. J. 4 *Let $\mathfrak{A}$ be a C^* algebra with identity, $\mathfrak{B}$ a sub-C^* algebra containing $\mathbf{1}$, ϕ a positive linear functional on $\mathfrak{B}$. Then ϕ may be extended to a positive linear functional $\tilde{\phi}$ on $\mathfrak{A}$, and we have $\|\tilde{\phi}\| = \|\phi\|$. If ϕ is a pure state of $\mathfrak{B}$, then $\tilde{\phi}$ may be taken to be a pure state of $\mathfrak{A}$.*

Proof Consider the set $\mathfrak{A}_h$ of self-adjoint elements of $\mathfrak{A}$ as a real vector space; $\mathfrak{A}_+$ is a convex cone in $\mathfrak{A}_h$ with $\mathbf{1}$ as an interior point. The extension theorem for positive functionals (Sec. I. C.) tells us that the real-linear functional ϕ on $\mathfrak{B} \cap \mathfrak{A}_h$ positive on $\mathfrak{B} \cap \mathfrak{A}_+$ may be extended to a real-linear functional $\tilde{\phi}$ on $\mathfrak{A}_h$ positive on $\mathfrak{A}_+$. Extend $\tilde{\phi}$ to a complex-linear functional on $\mathfrak{A}$ by $\tilde{\phi}(A + iB) = \tilde{\phi}(A) + i\tilde{\phi}(B)$. Then $\tilde{\phi}$ is positive and extends ϕ. Also $\|\tilde{\phi}\| = \tilde{\phi}(\mathbf{1}) = \phi(\mathbf{1}) = \|\phi\|$; the middle equality holds because $\mathbf{1} \in \mathfrak{B}$. It remains to prove the last assertion. Let ϕ be a pure state of $\mathfrak{B}$ (i.e., an extremal point of the set of states of $\mathfrak{B}$), and let $\tilde{\Phi}$ be the set of all positive linear functionals on $\mathfrak{A}$ extending ϕ. Then $\tilde{\Phi}$ is a non-empty convex subset of the unit ball of the dual of $\mathfrak{A}$; mereover, $\tilde{\Phi}$ is weak-* closed and therefore weak-* compact. By the Krein-Milman Theorem, $\tilde{\Phi}$ has at least one extremal point $\tilde{\phi}$. We claim that $\tilde{\phi}$ must be a pure state of $\mathfrak{A}$. To see this, let $\tilde{\phi} = \frac{1}{2}(\tilde{\phi}_1 + \tilde{\phi}_2)$, where $\tilde{\phi}_1$ and $\tilde{\phi}_2$ are states of $\mathfrak{A}$. Then $\phi_1 = \tilde{\phi}_1 \mid \mathfrak{B}$ and $\phi_2 = \tilde{\phi}_2 \mid \mathfrak{B}$ are states of $\mathfrak{B}$, and $\frac{1}{2}(\phi_1 + \phi_2) = \phi$. But by assumption ϕ is a pure state of $\mathfrak{B}$, so $\phi_1 = \phi_2 = \phi$. Hence, $\tilde{\phi}_1$ and $\tilde{\phi}_2$ are both extensions of ϕ, i.e., $\tilde{\phi}_1$ and $\tilde{\phi}_2$ belong to $\tilde{\Phi}$. Since $\tilde{\phi}$ is an extremal point of $\tilde{\Phi}$, and since $\frac{1}{2}(\tilde{\phi}_1 + \tilde{\phi}_2) = \tilde{\phi}$, we have $\tilde{\phi}_1 = \tilde{\phi}_2 = \tilde{\phi}$, so $\tilde{\phi}$ is an extremal point of the set of states of $\mathfrak{A}$, i.e., a pure state of $\mathfrak{A}$.

From the above extension theorem, we get:

PROPOSITION III. J. 5 *Let $\mathfrak{A}$ be a C^* algebra with identity, and let $A \in \mathfrak{A}$, $A \neq 0$. Then there is a state ϕ of $\mathfrak{A}$ such that $\phi(A^*A) > 0$.*

Proof Let $\mathfrak{B}$ be the sub-C^* algebra of $\mathfrak{A}$ generated by $\mathbf{I}$ and A^*A. Since $A^*A \neq 0$; there is a character χ of $\mathfrak{B}$ such that $\chi(A^*A) > 0$. Then χ is a state of $\mathfrak{B}$; we may therefore extend χ to a state ϕ of $\mathfrak{A}$.

Now, finally:

THEOREM III. J. 6 *Let $\mathfrak{A}$ be a C^* algebra. Then $\mathfrak{A}$ is isomorphic to a norm-closed self-adjoint algebra of operators on a Hilbert space.*

Proof By adjoining an identity of necessary, we can assume that $\mathfrak{A}$ has an identity. Since an injective homomorphism of C^* algebras is norm-preserving, we have only to produce a representation π of $\mathfrak{A}$ such that, if $A \neq 0$, $\pi(A) \neq 0$. For each state ϕ of $\mathfrak{A}$, construct the associated cyclic representation $(\mathscr{H}'_\phi, \pi_\phi, \xi_\phi)$, and let $\pi = \bigoplus_\phi \pi_\phi$. We claim that π is injective. Let $A \in \mathfrak{A}$, $A \neq 0$; then by the preceding proposition, $\phi(A^*A) > 0$ for some state ϕ. This implies that $\|\pi_\phi(A)\,\xi_\phi\|^2 = (\xi_\phi \mid \pi_\phi(A^*A)\,\xi_\phi) = \phi(A^*A) > 0$, so $\pi_\phi(A) \neq 0$, so $\pi(A) \neq 0$.

IV Von Neumann Algebras

A Introduction and Preliminaries

We have seen that C^* algebras may be regarded as algebras with involution which are isomorphic to norm closed algebras of bounded operators on Hilbert space. We now want to investigate von Neumann algebras, i.e., self-adjoint algebras of operators on Hilbert space which are closed in the weak operator topology. Unlike C^* algebras, which have many important properties which can be investigated abstractly, i.e., without realizing the algebras concretely on Hilbert space, von Neumann algebras are very closely tied to the Hilbert space on which they act, and their study is based largely on Hilbert-space techniques.

Let $\mathscr{H}$ be a Hilbert space; we are going to define various topologies on the set $\mathscr{L}(\mathscr{H})$ of all bounded operators on $\mathscr{H}$.

1. The *strong operator topology* on $\mathscr{L}(\mathscr{H})$ is defined by requiring that a net A_α converges to A if and only if, for all $\xi \in \mathscr{H}$, $A_\alpha \xi \to A\xi$ in the Hilbert space $\mathscr{H}$. Alternatively, we say that a set $G \subset \mathscr{L}(\mathscr{H})$ is open in the strong operator topology if, for all $A \in G$, there exists a finite set $\xi_1, \ldots, \xi_n$ of elements of $\mathscr{H}$ and an $\varepsilon > 0$ such that G contains $\{B \in \mathscr{L}(\mathscr{H}): \|B\xi_i - A\xi_i\| \leq \varepsilon$ for $1 \leq i \leq n\}$. Note that this condition imposes no constraints at all on what B does on the orthogonal complement of the subspace generated by $\xi_1, \ldots, \xi_n$, so no non empty strongly open set in $\mathscr{L}(\mathscr{H})$ is bounded in the norm.

2. The *weak operator topology* is defined by requiring that a net A_α converges to A if and only if, for all $\xi, \eta \in \mathcal{H}$, $(\eta \mid A_\alpha \xi) \to (\eta \mid A\xi)$. As before, we can also describe explicitly the open sets: A set $G \subset \mathcal{L}(\mathcal{H})$ is open for the weak operator topology if, for every $A \in G$ there exist $\xi_1, \xi_2, \eta_1, \ldots, \eta_n \in \mathcal{H}$ and $\varepsilon > 0$ such that $|(\eta_i \mid B\xi_i) - (\eta_i \mid A\xi_i)| \leq \varepsilon$ for $1 \leq i \leq n$ implies $B \in G$.

It is clear that a net which converges in the norm, or *uniform*, topology on $\mathcal{L}(\mathcal{H})$ converges to the same limit in the strong operator topology, and that a net which converges in the strong operator topology converges to the same limit in the weak operator topology. In other words, the uniform topology is stronger (finer; has more open sets) than the strong operator topology, which in turn is stronger than the weak operator topology.

We can now make the essential definition: A *von Neumann algebra* is a self-adjoint algebra of operators on a Hilbert space, which contains the identity operator and is closed in the weak operator topology. We will see shortly that it is equivalent to require that it be closed in the strong operator topology. Requiring that the algebra be strongly closed is, however, much more restrictive than requiring that it be norm closed. For example: Consider the algebra of all continuous functions on $[0, 1]$, regarded as multiplication operators on $\mathcal{L}^2$ of Lebesgue measure. This is a norm-closed algebra of operators, but it is not closed in the weak operator topology. Its weak closure is the algebra of all bounded Borel functions modulo functions which are zero almost everywhere, these functions again being regarded as defining multiplication operators on $\mathcal{L}^2$.

There are two other topologies on $\mathcal{L}(\mathcal{H})$ whose usefulness is less immediately evident, but which turn out to be very important for technical purposes:

3. The *ultrastrong operator topology* is defined by requiring that a net A_α converges to A if and only if, whenever (ξ_i) is a sequence of elements of $\mathcal{H}$ such that

$$\sum_i \|\xi_i\|^2 < \infty, \quad \sum_i \|A_\alpha \xi_i - A\xi_i\|^2 \to 0.$$

Roughly speaking, to approximate an operator A in the strong operator topology, one must be able to approximate it on any finite set of vectors simultaneously while to approximate it in the ultrastrong topology one must be able to approximate it on a countable set of vectors simultaneously but the approximation on most of the vectors need not be very good.

4. The *ultraweak operator topology* is similarly defined by requiring that a net A_α converges to A if and only if, whenever (ξ_i) and (η_i) are two sequences of vectors in $\mathcal{H}$, such that

$$\sum_i \|\xi_i\| \cdot \|\eta_i\| < \infty, \quad \sum_i (\eta_i \mid A_\alpha \xi_i) \to \sum_i (\eta_i \mid A\xi_i).$$

It is nearly obvious that the ultrastrong operator topology is weaker than the uniform topology, but stronger than the strong operator topology,

and also stronger than the ultraweak operator topology, and that the ultraweak operator topology is stronger (!) than the weak operator topology. Note, however, that a *bounded* net A_α which converges to A in the strong operator topology also converges in the ultrastrong operator topology: Let (ξ_i) be any sequence of vectors such that $\sum_i \|\xi_i\|^2 < \infty$, and assume $\|A_\alpha\| \leq M$ for all α. We want to show

$$\lim_\alpha \sum_i \|A_\alpha \xi_i - A\xi_i\|^2 = 0,$$

given that

$$\lim_\alpha \|A_\alpha \xi_i - A\xi_i\| = 0 \quad \text{for each } i.$$

Now

$$\limsup_\alpha \sum_i \|A_\alpha \xi_i - A\xi_i\|^2 \leq \limsup_\alpha \sum_{i=1}^{n} \|A_\alpha \xi_i - A\xi_i\|^2 + \sum_{i=n+1}^{\infty} 4M^2 \|\xi_i\|^2 = \sum_{i=n+1}^{\infty} 4M^2 \|\xi_i\|^2.$$

This is true for all n, so

$$\limsup_\alpha \sum_i \|A_\alpha \xi_i - A\xi_i\|^2 = 0.$$

We may re-express this remark by saying that the strong and ultrastrong topologies on $\mathscr{L}(\mathscr{H})$ agree on bounded sets. A similar argument shows that the weak and ultraweak topologies agree on bounded sets.

There is another way of looking at the ultrastrong and ultraweak topologies which is sometimes useful: We form the direct sum $\bigoplus_{i=1}^{\infty} \mathscr{H}$ of countably many copies of $\mathscr{H}$, and we represent $\mathscr{L}(\mathscr{H})$ on $\bigoplus_{i=1}^{\infty} \mathscr{H}$ by the direct sum of countably many copies of the identity representation. More concretely, given $A \in \mathscr{L}(\mathscr{H})$, we define an operator $\tilde{A}$ on $\bigoplus_{i=1}^{\infty} \mathscr{H}$ by $\tilde{A}(\xi_i) = (A\xi_i)$. It is nearly trivial to verify that a net A_α converges ultrastrongly (ultraweakly) to A if and only if $\tilde{A}_\alpha$ converges strongly (weakly) to $\tilde{A}$. It is also sometimes useful to identify $\bigoplus_{i=1}^{\infty} \mathscr{H}$ as $\mathscr{H} \otimes l^2$, where l^2 is the space of all sequences $(a_n)_{n=1,2,\ldots}$ with $\sum_n |a_n|^2 < \infty$. The identification may be carried out by mapping (ξ_i) to $\sum_{i=1} \xi_i \otimes l_i$ where $l_i \in l^2$ is the sequence with zeros everywhere except in the i^{th} place where there is a one. With this identification $\tilde{A}$ corresponds to the operator $A \otimes \mathbf{1}$.

The ultrastrong, ultraweak, strong, and weak operator topologies all have some unpleasant features with respect to the algebraic operations. Consider, for example, the mapping $A \mapsto A^*$. This is continuous in the weak operator topology: A net A_α converges to A if and only if $(\eta \mid A_\alpha \xi) \to (\eta \mid A\xi)$ for all $\xi, \eta \in \mathscr{H}$; taking complex conjugates gives $(\xi \mid A_\alpha^* \eta) \to (\xi \mid A^* \eta)$, which

implies that A_n^* converges weakly to A^*. A similar argument shows that $A \mapsto A^*$ is ultraweakly continuous. On the other hand, this mapping is not continuous with respect to the strong or ultrastrong operator topologies: Assume for simplicity that $\mathcal{H}$ is separable and has a complete orthonormal set (ϕ_i). Define an operator S on $\mathcal{H}$ by $S\phi_1 = 0$; $S\phi_{i+1} = \phi_i$ for $i = 1, 2, 3 \ldots$ For any

$$\xi = \sum_i \lambda_i \phi_i \in \mathcal{H},$$

$$S^n\xi = \sum_{i=n+1}^{\infty} \lambda_i \phi_{i-n}, \quad \text{so} \quad \|S^n\xi\|^2 = \sum_{i=n+1}^{\infty} |\lambda_i|^2$$

which goes to zero as n goes to infinity. Thus S^n converges strongly to 0 as n goes to infinity. On the other hand, it is easy to check that S^* is given by $S^*\phi_i = \phi_{i+1}$, $i = 1, 2, 3 \ldots$ and hence that

$$\left\|(S^*)^n \left(\sum_i \lambda_i \phi_i\right)\right\|^2 = \left\|\sum_i \lambda_i \phi_{i+n}\right\|^2 = \sum_i |\lambda_i|^2,$$

so $(S^*)^n$ does not converge strongly to zero as n goes to infinity. Since the sequence (S^n) is bounded, and since the strong and ultrastrong topologies agree on bounded sets, this same example shows that $A \mapsto A^*$ is not continuous in the ultrastrong topology.

Furthermore, the mapping $(A, B) \mapsto A \cdot B$ is not continuous in any one of the four topologies we are considering (i.e., multiplication is not jointly continuous). This is easy to see for the weak topology: The sequence S^n constructed in the preceding paragraph converges strongly to zero and hence also converges weakly to zero; since taking adjoints is continuous in the weak operator topology, $(S^*)^n$ also converges weakly to zero. But $S^n(S^*)^n = I$ which certainly does not converge weakly to zero. The argument for the strong operator topology is a bit more subtle. We note first that, if a *sequence* A_n converges strongly, then $\|A_n\xi\|$ is bounded in n for each ξ, and hence, by the uniform boundedness principle, $\|A_n\|$ is bounded in n. Now let A_n be sequence converging strongly to A, and let B_n be a sequence converging strongly to B; we will show that A_nB_n converges strongly to $A \cdot B$. Thus, let $\xi \in \mathcal{H}$; then

$$\begin{aligned}\|A_nB_n\xi - AB\xi\| &\leq \|A_nB_n\xi - A_nB\xi\| + \|A_nB\xi - AB\xi\| \\ &\leq \|A_n\| \cdot \|(B_n - B)\xi\| + \|(A_n - A)(B\xi)\| \to 0.\end{aligned}$$

Nevertheless, the product is not jointly continuous in the strong topology. (And this shows, among other things, that one can make mistakes by arguing about sequences rather than general nets.) Indeed, if W is any strong neighborhood of zero in $\mathcal{L}(\mathcal{H})$, we will show that there exist $A, B \in W$ such that $A \cdot B = I$. By the definition of the strong topology, W contains a set of the form

$$\{C \in \mathcal{L}(\mathcal{H}) : \|C\zeta_i\| \leq \varepsilon \quad \text{for} \quad 1 \leq i \leq j\}.$$

Here, $\zeta_1, \ldots, \zeta_j$ are elements of $\mathscr{H}$ and $\varepsilon > 0$. Let

$$M = \sup_i \frac{\|\zeta_i\|}{\varepsilon}$$

We again assume for simplicity that $\mathscr{H}$ is separable, and we construct the operator S as above. Since (S^n) converges strongly to zero, we may choose n so that

$$\|S^n\zeta_i\| \leqq \frac{\varepsilon}{M} \quad \text{for} \quad i = 1, 2, \ldots, j.$$

Let $A = M \cdot S^n$; $B = \frac{1}{M}(S^*)^n$. Then

$$\|A\zeta_i\| = M \cdot \|S^n\zeta_i\| \leqq \frac{M \cdot \varepsilon}{M} = \varepsilon \quad \text{for} \quad i = 1, 2, \ldots, j,$$

so $A \in W$, and

$$\|B\zeta_i\| = \frac{1}{M}\|(S^*)^n\zeta_i\| = \frac{1}{M}\|\zeta_i\| \leqq \varepsilon \quad \text{for} \quad i = 1, 2, \ldots, j$$

(by the choice of M), so $B \in W$. But $A \cdot B = S^n(S^*)^n = \mathbf{1}$ as asserted. A similar argument shows that the product is not jointly continuous in the ultrastrong topology.

The pathologies of the operator topologies are somewhat mitigated by the following remarks:

1. The product is continuous in each variable separately in any one of the operator topologies. We will prove this for the weak operator topology: Let A_α be a net converging weakly to A, and let $B \in \mathscr{L}(\mathscr{H})$. Then, for any

$$\xi, \eta \in \mathscr{H}, (\eta \mid A_\alpha B\xi) \to (\eta \mid AB\xi),$$

so $A_\alpha B$ converges weakly to AB, and

$$(\eta \mid BA_\alpha\xi) = (B^*\eta \mid A_\alpha\xi) \to (B^*\eta \mid A\xi) = (\eta \mid BA\xi),$$

so BA_α converges weakly to $B \cdot A$.

2. The mapping $(A, B) \mapsto A \cdot B$ is jointly strongly continuous on bounded sets in $\mathscr{L}(\mathscr{H})$. Let A_α, B_α be bounded nets converging strongly to A, B respectively. Then, for any $\xi \in \mathscr{H}$

$$\begin{aligned}\|A_\alpha B_\alpha\xi - AB\xi\| &\leqq \|A_\alpha B_\alpha\xi - A_\alpha B\xi\| + \|A_\alpha B\xi - AB\xi\| \\ &\leqq \|A_\alpha\| \, \|B_\alpha\xi - B\xi\| + \|A_\alpha(B\xi) - A(B\xi)\| \to 0\end{aligned}$$

so $A_\alpha \cdot B_\beta$ converges strongly to $A \cdot B$. (Incidentally, we needed only the boundedness of A_α to make the above argument work). Since the strong and ultrastrong topologies agree on bounded sets, multiplication is also jointly continuous on bounded sets. We have, however, already given an example which shows that, even on bounded sets, multiplication is not jointly continuous in the weak operator topology.

We summarize the state of affairs in a table:

Topology	$A \mapsto A^*$	Multiplication
Uniform	Continous	Jointly continuous
Ultrastrong	Not continous	Separately continuous
strong		Not jointly continuous
		Jointly continuous on bounded sets
Ultraweak	Continuous	Separately continuous
weak		Not jointly continuous even on bounded sets

It should be remarked that the negative statement made above are correct only if the Hilbert space on which the operators are acting is infinite-dimensional. If the Hilbert space is finite-dimensional, all the operator topologies coincide. We also add a remark about terminology. One commonly says "weak topology", or "strong topology", rather than "weak operator topology" or "strong operator topology". This language is somewhat unfortunate, since, for example, the term "weak topology" ought to be applied to the weak topology of $\mathscr{L}(\mathscr{H})$ as a Banach space (which is not at all the same as the weak operator topology). Nevertheless, it is frequently convenient to use expressions like "weakly continuous" instead of saying "continuous in the weak operator topology". We will use the imprecise terminology freely, except in places where it seems to lead to serious ambiguities.

We will review here a few ideas from operator theory which will be needed later. Let $\mathscr{H}$ be a Hilbert space, $\mathscr{H}_1$ and $\mathscr{H}_2$ closed subspaces of $\mathscr{H}$. An operator U on $\mathscr{H}$ is said to be a *partial isometry with initial subspace* $\mathscr{H}_1$ *and terminal subspace* $\mathscr{H}_2$ if:

i) $U\xi = 0$ if ξ is orthogonal to $\mathscr{H}_1$.

ii) U, restricted to $\mathscr{H}_1$, is an isometry from $\mathscr{H}_1$ onto $\mathscr{H}_2$ (i.e., U is unitary from $\mathscr{H}_1$ to $\mathscr{H}_2$). If U is a partial isometry with initial subspace $\mathscr{H}_1$ and terminal subspace $\mathscr{H}_2$, then it may be verified that U^* is a partial isometry with initial subspace $\mathscr{H}_2$ and terminal subspace $\mathscr{H}_1$, and that

$$U^*U = P_{\mathscr{H}_1}; \quad UU^* = P_{\mathscr{H}_2}.$$

Conversely, if U is any operator such that $U^*U = P_{\mathscr{H}_1}$, then U is a partial isometry with initial subspace $\mathscr{H}_1$ (and terminal subspace $U\mathscr{H}_1$).

Now let $A \in \mathscr{L}(\mathscr{H})$; then A^*A is a positive self-adjoint operator, and therefore has a unique positive square-root, which we will denote by $|A|$. For any $\xi \in \mathscr{H}$,

$$\|A\xi\|^2 = (A\xi \mid A\xi) = (\xi \mid A^*A\xi) = (\xi \mid |A|^2 \xi) = \||A|\xi\|^2.$$

Thus, there is a unique unitary operator U_A from the closure of the range of $|A|$ to the closure of the range of A such that

$$U_A|A| = A.$$

We extend U_A to a partial isometry by making it zero on the orthogonal complement of the range of $|A|$. Note, incidentally, that the orthogonal complement of the range of $|A|$ is just the null space of $|A|$, which is the same as the null-space of A. Thus, any bounded operator A may be written uniquely as

$$A = U_A|A|,$$

where $|A|$ is a positive operator vanishing on the null-space of A and U_A is a partial isometry with initial subspace the orthogonal complement of the null-space of A and terminal subspace the closure of the range of A. This way of writing A is called the *polar decomposition* of A. From the uniqueness statement, it follows that, if W is any unitary operator such that $WAW^{-1} = A$, then $W|A|W^{-1} = |A|$ and $WU_AW^{-1} = U_A$, (since $A = (WU_AW^{-1}) \times (W|A|W^{-1})$ is another polar decomposition of A). Thus, any unitary operator commuting with A also commutes with $|A|$ and U_A. It is not necessary for the contruction that A map a Hilbert space into itself. If A is a bounded operator from $\mathscr{H}$ to $\mathscr{H}'$, then A can be written uniquely as $A = U_A|A|$; where $|A|$ is a positive operator on $\mathscr{H}$ and U_A is a partial isometry with initial subspace the orthogonal complement of the null-space of A (in $\mathscr{H}$) and terminal subspace the closure of the range of A (in $\mathscr{H}'$).

We also need some facts about increasing nets of operators: A net A_α of self-adjoint operators is *increasing* if $A_\alpha \geqq A_\beta$ whenever $\alpha \geqq \beta$. If $\|A_\alpha\|$ is bounded, then for each $\xi \in \mathscr{H}$, $(\xi \mid A_\alpha\xi)$ is a bounded increasing net of real numbers and hence converges. The polarization identity gives:

$$(\eta \mid A_\alpha\xi) = \tfrac{1}{4}[(\xi + \eta \mid A_\alpha(\xi + \eta)) - (\xi - \eta \mid A_\alpha(\xi - \eta))]$$

so $(\eta \mid A_\alpha\xi)$ converges for all $\xi, \eta \in \mathscr{H}$. The limit is a bounded bilinear form; hence, defines a bounded operator A, and we have shown that A_α converges weakly to A. It may also be shown that A_α converges strongly to A: Let $\xi \in \mathscr{H}$, and consider $\|(A - A_\alpha)\,\xi\|^2 = (\xi \mid (A - A_\alpha)^2\,\xi)$.
The sesquilnear form $\langle \xi \mid \eta \rangle = (\xi \mid (A - A_\alpha)\,\eta)$ is positive semi-definite, so the Schwarz inequality holds. Inserting $(A - A_\alpha)\,\xi$ for η, we get

$$(\xi \mid (A - A_\alpha)^2\,\xi) = \langle \xi \mid (A - A_\alpha)\,\xi \rangle \leqq \sqrt{\langle \xi \mid \xi \rangle \langle (A - A_\alpha)\,\xi \mid (A - A_\alpha)\xi \rangle}$$
$$= \sqrt{(\xi \mid (A - A_\alpha)\,\xi)}\,\sqrt{(\xi \mid (A - A)^3\,\xi)}.$$

The first term goes to zero and the second term remains bounded, so $\lim \|(A - A_\alpha)\,\xi\| = 0$, i.e., A_α converges strongly to A.

In particular, let $(P_i)_{i \in I}$ be a family of mutually orthogonal projections, and consider the net of finite partial sums of this family. This is an increasing net of projections, and hence converges strongly to a limit, which may easily be seen to be the projection onto the subspace generated by the union of the ranges of the P_i. We will write this limiting projection as $\sum_{i \in I} P_i$.

More generally, if $(P_i)_{i \in I}$ is any family of projections, not necessarily mutually orthogonal, we consider the net, labelled by the finite subsets

$(i_1, \ldots, i_n)$ of I, defined by $P_{(i_1, \ldots, i_n)} = P_{i_1} \vee \ldots \vee P_{i_n}$, i.e. the projection onto the subspace generated by the union of the ranges of $P_{i_1}, \ldots, P_{i_n}$. This is again an increasing net which converges strongly to the projection onto the subspace generated by the union of the ranges of all the P_i's; we denote this limiting projection by $\bigvee_{i \in I} P_i$.

B Linear Functionals on Operator Algebras

Before we can begin the study of von Neumann algebras proper, we need a few more technicalities having to do with linear functionals on operator algebras. Since the various operator topologies we have introduced are all weaker than the uniform topology, we do not expect, in general, that all norm-continuous linear functionals will be, say, strongly continuous. We want to see what can be said about special properties of functionals which are continuous in one or another of the operator topologies. The basic fact is that such functionals can always be written as sums (possibly infinite) of functionals of the form $A \mapsto (\xi \mid A\eta)$ with $\xi, \eta \in \mathscr{H}$.

PROPOSITION IV. B. 1 *Let (ξ_i), (η_i) be two sequences of elements of $\mathscr{H}$ such that $\sum_i \|\xi_i\| \cdot \|\eta_i\| < \infty$. Then the mapping $A \mapsto \sum_i (\eta_i \mid A\xi_i)$ is ultraweakly, and hence ultrastrongly, continuous. Conversely, if $\mathfrak{A}$ is a linear subspace of $\mathscr{L}(\mathscr{H})$, and if ϕ is an ultrastrongly continuous linear functional on $\mathfrak{A}$, then there exist two sequences (ξ_i), (η_i) as above such that*

$$\phi(A) = \sum_i (\eta_i \mid A\xi_i).$$

In particular, ϕ is ultraweakly continuous, i.e., a linear functional is ultraweakly continuous if and only if it is ultrastrongly continuous. The above statements remain true with "ultrastrong" replaced by "strong", "ultraweak" replaced by "weak", and infinite sequences replaced by finite ones.

Proof The first assertion is immediate: By the definition of the ultraweak topology, if a net A_α converges ultraweakly to A, then

$$\sum_i (\eta_i \mid A_\alpha \xi_i) \to \sum_i (\eta_i \mid A\xi_i),$$

i.e., the linear functional $\sum_i (\eta_i \mid A\xi_i)$ is ultraweakly continuous.

Now let ϕ be an ultrastrongly continuous functional. By the definition of the ultrastrong topology, there exists a sequence (ξ_i) in $\mathscr{H}$ such that $\sum_i \|\xi_i\|^2 < \infty$ and such that $\sum_i \|A\xi_i\|^2 < \varepsilon$ implies $|\phi(A)| < 1$. Regard (ξ_i) as an element of $\bigoplus_{i=1}^{\infty} \mathscr{H}$, and consider the linear subspace $\{(A\xi_i) : A \in \mathfrak{A}\}$ of $\bigoplus_{i=1}^{\infty} \mathscr{H}$. On this subspace, the correspondence $(A\xi_i) \mapsto \phi(A)$ is well-defined and continuous, since $\sum_i \|A\xi_i\|^2 < \varepsilon$ implies $\phi(A) < 1$.

Extending by continuity, we get a continuous linear functional on the closed linear subspace $\overline{\{(A\xi_i) : A \in \mathfrak{A}\}}$ of $\bigoplus_{i=1}^{\infty} \mathscr{H}$. By elementary Hilbert space theory, a continuous linear functional on a Hilbert space is always given by taking the scalar product with an element of the space, so there exists an element of $\bigoplus_{i=1}^{\infty} \mathscr{H}$, i.e., a sequence (η_i) with $\sum_i \|\eta_i\|^2 < \infty$, such that $\phi(A) = \sum_i (\eta_i \mid A\xi_i)$ for all $A \in \mathfrak{A}$. This, then, proves the assertions about the ultrastrong-ultraweak topologies; the proofs for the strong-weak topologies are similar.

Remark It follows from the above proposition that the ultraweak topology could have been defined as follows: A net A_α in $\mathscr{L}(\mathscr{H})$ converges to A in the ultraweak topology if and only if $\phi(A_\alpha)$ converges to $\phi(A)$ for all ultrastrongly continuous linear functionals ϕ. In other words, the ultraweak topology is the weakened topology associated with the ultrastrong topology. Similarly, the weak operator topology is the weakened topology associated with the strong operator topology. It is a general fact about topological vector spaces that a locally convex topology and its associated weakened topology have the same closed *convex* sets (although the initial topology may have many non-convex closed sets which are not closed in the weakened topology). We will prove this result for the case at hand; the proof is valid in general.

PROPOSITION IV. B. 2 *If K is a convex set in $\mathscr{L}(\mathscr{H})$, the strong operator closure of K is the same as its weak operator closure, and the ultrastrong closure of K is the same as its ultraweak closure.*

Proof We will prove only that the strong operator closure of K is the same as the weak operator closure. The weak operator closure of K is a strongly closed set containing K, hence, contains the strong operator closure of K. If we show that the strong operator closure of K is weakly closed, we get the opposite inclusion and thus finish the proof. Replacing K by its strong operator closure, which is again convex, we assume that K is closed in the strong operator topology but not in the weak operator topology, and attempt to derive a contradiction. Let A be in the weak operator closure of K but not in K. Since K is convex, and closed in the strong operator topology, we may apply the Hahn-Banach Theorem to show that Amay be separated from K by a strongly continuous linear functional, i.e. there is a strongly continuous linear functional ϕ such that

$$\mathrm{Re}\,\{\phi(A)\} > \sup\{\mathrm{Re}\{\phi(B)\} : B \in K\}.$$

But since ϕ is strongly continuous, it is also weakly continuous, and the above inequality contradicts the assumption that there is a net A_α in K converging weakly to A

Remark There do exist convex sets which are ultrastrongly closed but not strongly closed, e.g., the kernel of a linear functional which is ultrastrongly continuous but not strongly continuous.

We next want to see what happens when the functional ϕ is assumed to be positive:

PROPOSITION IV B. 3 *Let $\mathfrak{A}$ be a self-adjoint subalgebra of $\mathcal{L}(\mathcal{H})$, containing the identity operator, and let ϕ be an ultrastrongly continuous positive linear functional on $\mathfrak{A}$. Then there exists a sequence (ζ_i) of vectors of $\mathcal{H}$ such that $\sum_i \|\zeta_i\|^2 < \infty$ and such that*

$$\phi(A) = \sum_i (\zeta_i \mid A\zeta_i)$$

for all $A \in \mathfrak{A}$. If ϕ is strongly continuous, the sequence of ζ_i's may be taken to be finite.

Proof We will consider only the ultrastrong topology; the proof of the assertion for the strong operator topology is best obtained by repeating the argument with appropriate modifications. Furthermore, we may replace $\mathfrak{A}$ by its norm closure, i.e., we may assume that $\mathfrak{A}$ is a C^* algebra.

We know two things already:

1. Since ϕ is ultrastrongly continuous, we may write $\phi(A) = (\eta \mid \tilde{A}\xi)$, where $\xi, \eta \in \bigoplus_{i=1}^{\infty} \mathcal{H}$. (We have used the fact that, if $\xi = (\xi_i)$, $\eta = (\eta_i)$, $(\eta \mid \tilde{A}\xi) = \sum_i (\eta_i \mid A\xi_i)$).

2. $\phi(A) \geq 0$ for $A \geq 0$.

We want to combine these to show that $\phi(A) = (\zeta \mid \tilde{A}\zeta)$ for some $\zeta \in \bigoplus_{i=1}^{\infty} \mathcal{H}$. One might hope to prove this simply by being careful in the choice of η, but it turns out to be easier to use a trick. Consider the positive linear functional $\psi(A) = ((\xi + \eta) \mid \tilde{A}(\xi + \eta))$. Let A be a positive operator in $\mathfrak{A}$; then

$$\psi(A) = (\xi \mid \tilde{A}\xi) + (\eta \mid \tilde{A}\eta) + 2(\eta \mid \tilde{A}\xi) \geq 2\phi(A).$$

Hence, the positive functional ψ majorizes ϕ. Consider the cyclic subspace of $\xi + \eta$ in $\bigoplus_{i=1}^{\infty} \mathcal{H}$, i.e., the closure of $\{\tilde{A}(\xi + \eta) : A \in \mathfrak{A}\}$, and the representation of $\mathfrak{A}$ defined by restricting each $\tilde{A}$ to this cyclic subspace. By the uniqueness of the Gelfand-Segal construction, this representation is unitarily equivalent to the canonical cyclic representation associated with the positive linear functional ϕ. Since ψ majorizes ϕ, we know by Proposition III.H.1 that there is a positive operator T on the cyclic subspace, commuting with $\tilde{A}$ for each $A \in \mathfrak{A}$, such that

$$\phi(A) = (T(\xi + \eta) \mid \tilde{A}(\xi + \eta)) = (\sqrt{T}(\xi + \eta) \mid \tilde{A}\sqrt{T}(\xi + \eta)).$$

Letting $\zeta = (\zeta_i) = \sqrt{T}(\xi + \eta)$, we have the desired result.

Now let ϕ be a ultrastrongly continuous positive linear functional on a self-adjoint sub-algebra $\mathfrak{A}$ of $\mathscr{L}(\mathscr{H})$ containing the identity operator. We know that we can write:

$$\phi(A) = \sum (\zeta_i \mid A\zeta_i), \quad \text{where} \quad \sum_i \|\zeta_i\|^2 = \phi(\mathbf{I}) < \infty.$$

Define a linear operator ϱ on $\mathscr{H}$ by

$$\varrho\xi = \sum_i \zeta_i(\zeta_i \mid \xi).$$

Then ϱ is positive as $(\xi \mid \varrho\xi) = \sum_i |(\zeta_i \mid \xi)|^2$. Also, if $(\xi_1, \ldots, \xi_n)$ is any finite orthonormal set in $\mathscr{H}$,

$$\sum_{j=1}^{n} (\xi_j \mid \varrho\xi_j) = \sum_{j=1}^{n} \sum_i |(\zeta_i \mid \xi_j)|^2 \leqq \sum_i \|\zeta_i\|^2 = \phi(\mathbf{I}),$$

The bound on $\sum_{j=1}^{n} (\xi_j \mid \varrho\xi_j)$ which is independent of n implies that the operator ϱ is of trace class.

By the properties of (ζ_i), we have for any complete orthonormal set (ξ_α) and any $A \in \mathfrak{A}$,

$$\phi(A) = \sum_i (\zeta_i \mid A\zeta_i) = \sum_{\alpha,i} (\xi_\alpha \mid A\zeta_i)(\zeta_i \mid \xi_\alpha)$$
$$= \sum_\alpha (\xi_\alpha \mid A\varrho\xi_\alpha) = Tr(A\varrho).$$

i.e., $\phi(A) = \mathrm{Tr}\,(A\varrho)$. Thus we have:

PROPOSITION IV.B.4 *Let $\mathfrak{A}$ be a self-adjoint subalgebra of $\mathscr{L}(\mathscr{H})$ containing the identity operator, and let ϕ be a positive ultrastrongly continuous linear functional on $\mathfrak{A}$. Then there exists a positive linear operator ϱ of trace class such that*

$$\phi(A) = Tr(A\varrho)$$

for all $A \in \mathfrak{A}$. Conversely, if ϱ is a positive linear operator of trace class, $A \mapsto \mathrm{Tr}\,(A\varrho)$ is an ultrastrongly continuous positive linear functional on $\mathscr{L}(\mathscr{H})$.

It remains only to prove the converse statement. By the spectral theorem, we may write $\varrho\xi = \sum \zeta_i(\zeta_i \mid \xi)$, where the ζ_i are eigenvectors of ϱ and the corresponding eigenvalue is $\|\zeta_i\|^2$.

By the finiteness of the trace of ϱ, $\sum_i \|\zeta_i\|^2 < \infty$. We have

$$Tr\,(\varrho A) = \sum_\alpha (\xi_\alpha \mid \varrho A\xi_\alpha) = \sum_{\alpha,i} (\xi_\alpha \mid \zeta_i)(\zeta_i \mid A\xi_\alpha)$$
$$= \sum_i (\zeta_i \mid A\zeta_i),$$

(Here (ξ_α) is any complete orthonormal set in $\mathscr{H}$).

Thus $A \mapsto \mathrm{Tr}\,(\varrho A)$ is a positive ultraweakly continuous linear functional on $\mathscr{L}(\mathscr{H})$. Note incidentally that the ζ_i are mutually orthogonal, so

Proposition IV.B.3 can be refined by adding the requirement that the ζ_i be mutually orthogonal.

We will refer to ρ as a *density matrix* determining ϕ. It is not, in general, uniquely determined by ϕ (but is uniquely determined if $\mathfrak{A} = \mathscr{L}(\mathscr{H})$).

Note A positive linear functional ϕ on a von Neumann algebra $\mathfrak{A}$ is said to be *normal* if, whenever P_α is an increasing net of projections in $\mathfrak{A}$ (this means $P_\alpha \geqq P_\beta$ when $\alpha \geqq \beta$) converging to P (i.e., P is the projection onto the closure of $\bigcup_\alpha P_\alpha \mathscr{H}$), we have

$$\phi(P) = \lim_\alpha \phi(P_\alpha).$$

Since P_α converges to P in the ultrastrong topology, any ultrastrongly continuous positive linear functional is normal. It turns out that the converse is true: A positive linear functional on a von Neumann algebra is normal if and only if it is ultrastrongly continuous. (We will not give the proof that normality implies ultrastrong continuity; it is sketched in exercice 9, p. 65 of Dixmier AvN.). The usefulness of this result lies in the fact that it is frequently easier to verify that a functional is normal than that it is ultrastrongly continuous.

To summarize: A positive linear functional on a von Neumann algebra is normal if and only if it is ultrastrongly continuous, which is true if and only if it is given by a density matrix.

C The von Neumann and Kaplansky Density Theorems

We have defined a von Neumann algebra to be a weakly closed self-adjoint algebra of bounded operators on a Hilbert space containing the identity operator. In this section, we give a more algebraic characterization of von Neumann algebras.

Let M be any subset of $\mathscr{L}(\mathscr{H})$; we define the *commutant* of M, written M', to be the set of elements of $\mathscr{L}(\mathscr{H})$ commuting with every element of M.

PROPOSITION IV.C.1 *For any subset M of $\mathscr{L}(\mathscr{H})$, $(M \cup M^*)'$ is a von Neumann algebra.*

Proof It is clear that $(M \cup M^*)'$ is an algebra of operators, and that the identity operator belongs to $(M \cup M^*)'$. If $A \in (M \cup M^*)'$, then

$$[A^*, B] = ([B^*, A])^* = 0 \quad \text{for all} \quad B \in (M \cup M^*), \quad \text{so} \quad (M \cup M^*)'$$

is self-adjoint. Finally, $(M \cup M^*)'$ is weakly closed. If A_α is a net in $(M \cup M^*)'$ converging weakly to A, and if $B \in M \cup M^*$, then

$$[A, B] = AB - BA = \lim_\alpha (A_\alpha B - BA_\alpha) = \lim_\alpha 0 = 0,$$

by the separate continuity of multiplication, so $A \in (M \cup M^*)'$.

PROPOSITION IV.C.2 *If M, N are any subsets of $\mathscr{L}(\mathscr{H})$, and if $M \subset N$, then $M' \supset N'$, and $M \subset M''$.*

Proof If $M \subset N$ and $A \in N'$, then A commutes with every element of N; hence, with every element of M, so $A \in M'$. If $A \in M$, and $B \in M'$, then A commutes with B. Hence, A commutes with every element of M', i.e., $A \in M''$.

We now come to the crucial:

THEOREM IV.C.3 (*von Neumann Density Theorem, Bicommutant Theorem*). *Let $\mathfrak{A}$ be a self-adjoint subalgebra of $\mathscr{L}(\mathscr{H})$ containing the identity operator. Then the ultrastrong, strong, ultraweak, and weak closures of $\mathfrak{A}$ are all the same, and are equal to $\mathfrak{A}''$.*

Thus, a von Neumann algebra could alternatively have been defined as a self-adjoint subalgebra of $\mathscr{L}(\mathscr{H})$ equal to its bicommutant.

Proof We know that $\mathfrak{A}''$ is a von Neumann algebra, and that $\mathfrak{A}'' \supset \mathfrak{A}$. If we can show that $\mathfrak{A}$ is ultrastrongly dense in $\mathfrak{A}''$, then, since $\mathfrak{A}''$ is ultrastrongly closed, it is equal to the ultrastrong closure of $\mathfrak{A}$. But $\mathfrak{A}''$ is also strongly, weakly, and ultraweakly closed, so the ultrastrong closure of $\mathfrak{A}$ coincides with the strong, weak and ultraweak closure, and with $\mathfrak{A}''$. The problem, then, is to prove that $\mathfrak{A}$ is ultrastrongly dense in $\mathfrak{A}''$, i.e., that, given any $B \in \mathfrak{A}''$, any sequence (ξ_i) in $\mathscr{H}$ such that $\sum_i \|\xi_i\|^2 < \infty$, and any $\varepsilon > 0$, there exists $A \in \mathfrak{A}$ such that $\sum_i \|B\xi_i - A\xi_i\|^2 < \varepsilon$. We now use the trick of thinking of (ξ_i) as an element ξ of $\bigoplus_{i=1}^{\infty} \mathscr{H}$ and note that what we have to prove is precisely that $\tilde{B}\xi$ is in the closed linear span of $\{\tilde{A}\xi : A \in \mathfrak{A}\}$. Let P denote the projection from $\bigoplus_{i=1}^{\infty} \mathscr{H}$ onto this closed linear span; then what we have to show is that $P\tilde{B}\xi = \tilde{B}\xi$. Regard $A \mapsto \tilde{A}$ as a representation of $\mathfrak{A}$; then the projection P onto the cyclic subspace $\overline{\tilde{\mathfrak{A}}\xi}$ commutes with $\tilde{A}$ for every $A \in \tilde{\mathfrak{A}}$ (since the cyclic subspace is invariant). Thus, $P \in \tilde{\mathfrak{A}}'$. Let us admit, without proof for the moment, that $\tilde{B} \in \tilde{\mathfrak{A}}''$. Then

$$P\tilde{B}\xi = \tilde{B}P\xi \quad \text{(since } P \in \tilde{\mathfrak{A}}' \text{ and } \tilde{B} \in \tilde{\mathfrak{A}}'')$$

$$= \tilde{B}\xi, \quad \text{(since } \xi \text{ belongs to the cyclic subspace } \overline{\tilde{\mathfrak{A}}\xi}),$$

so we are through, except for the verification that $\tilde{B} \in \tilde{\mathfrak{A}}''$. To carry out this verification, we note that any operator C on $\bigoplus_{i=1}^{\infty} \mathscr{H}_i$ is defined by a matrix (C_{ij}) of bounded operators on $\mathscr{H}$, so that

$$(C\xi)_i = \sum_j C_{ij}\xi_j.$$

(Warning: The conditions on the matrix (C_{ij}) required to make it the matrix of a bounded operator are very complicated; we do not need to investigate them since we are starting from a bounded operator C). Now $C \in \tilde{\mathfrak{A}}'$ if and only if $\sum_j AC_{ij}\xi_j = \sum_j C_{ij}A\xi_j$ for all i and all (ξ_i), i.e. if and only if $C_{ij} \in \mathfrak{A}'$ for all i, j. Reversing the above calculation shows that, if $B \in \mathfrak{A}''$, and if $C_{ij} \in \mathfrak{A}'$, for all i, j, then $\tilde{B}$ commutes with C i.e., that $\tilde{B} \in \tilde{\mathfrak{A}}''$.

We can now read out some consequences. Let A be any element of a von Neumann algebra $\mathfrak{A}$. We may write $A = A_1 + iA_2$, with A_1 and A_2 self-adjoint. In the proof of the spectral theorem we showed that the spectral projections of any self-adjoint operator are strong limits of polynomials in that operator. In particular, the spectral projections of A_1 and A_2 belong to $\mathfrak{A}$. But on the other hand, any self-adjoint operator can be approximated arbitrarily well in norm by finite linear combinations of its spectral projections. Conclusion: Let $\mathfrak{A}$ be a von Neumann algebra; then the set of finite linear combinations of projections in $\mathfrak{A}$ is dense in $\mathfrak{A}$. Thus, an operator A' is in $\mathfrak{A}'$ if and only if it commutes with all the projections in $\mathfrak{A}$. If P is a projection, then $2P - \mathbf{1}$ is a unitary operator, so we can similarly conclude: An operator A' is in $\mathfrak{A}'$ if and only if it commutes with all the unitary operators in $\mathfrak{A}$. So far, we have not used the bicommutant theorem. We use it by remarking that, since $\mathfrak{A}$ is a von Neumann algebra, $\mathfrak{A} = \mathfrak{A}''$, so we may interchange the roles of $\mathfrak{A}$ and $\mathfrak{A}'$, to get:

PROPOSITION IV.C.4 *Let $\mathfrak{A}$ be a von Neumann algebra; then if $B \in \mathscr{L}(\mathscr{H})$ commutes with all unitary operators commuting with all operators in $\mathfrak{A}$, then $B \in \mathfrak{A}$. In particular, if $A \in \mathfrak{A}$, and if we write the polar decomposition*

$$A = U_A|A|,$$

then $|A|$ and U_A belong to $\mathfrak{A}$.

It is worth knowing what happens when we take closures of algebras without identity. Thus, let $\mathfrak{A}$ be a self-adjoint subalgebra of $\mathscr{L}(\mathscr{H})$. Regarding $\mathfrak{A}$ as a representation of itself, we may use the general decomposition of a representation of an algebra with identity into a non-degenerate part and a trivial part (Proposition III.F.1) to decompose $\mathscr{H} = \mathscr{H}_1 \oplus \mathscr{H}_2$, where

$$\mathscr{H}_2 = \{\xi \in \mathscr{H} : A\xi = 0 \quad \text{for all} \quad A \in \mathfrak{A}\}$$

and $\mathscr{H}_1$ = closed linear span of the union of the ranges of the operators in $\mathfrak{A}$. Let P be the projection on $\mathscr{H}_1$; then, for all $A \in \mathfrak{A}$, $PA = AP = A$, so $P \in \mathfrak{A}'$. It then turns out that the strong, weak, ultrastrong, and ultraweak closures of $\mathfrak{A}$ all coincide; the closure contains P, and is obtained by restricting all operators in $\mathfrak{A}$ to $\mathscr{H}_1$, taking the bicommutant (in $\mathscr{L}(\mathscr{H}_1)$), and then extending by zero. Thus, an arbitrary weakly (or strongly, etc.) closed self-adjoint subalgebra of $\mathscr{L}(\mathscr{H})$ splits into the direct sum of a von Neumann algebra and the zero algebra.

We sketch the proof of the above assertions. It is clear that the closure of $\mathfrak{A}$ in any of the topologies is obtained by restricting to $\mathscr{H}_1$, taking the closure in the corresponding topology on $\mathscr{L}(\mathscr{H}_1)$, then extending by zero. Since P restricted to $\mathscr{H}_1$ is the identity, we have only to prove that P is in the ultrastrong closure of $\mathfrak{A}$ and then to apply the bicommutant theorem to the ultrastrong closure of $\mathfrak{A}$ restricted $\mathscr{H}_1$.

To prove that P is in the ultrastrong closure of $\mathfrak{A}$, we note that P is the supremum of the projections onto the ranges of the self-adjoint elements of $\mathfrak{A}$. First, let A be a single self-adjoint operator in $\mathfrak{A}$. It follows easily from the spectral theorem and the dominated convergence theorem that, if $P_n(z)$ is a sequence of polynomials which is uniformly bounded on the spectrum of A and which converges pointwise to the function equal to 1 for $z \neq 0$ and to zero for $z = 0$, then the sequence of operators $P_n(A)$ is uniformly bounded and converges strongly (hence, ultrastrongly) to the projection onto the range of A. It is is clear from the Weierstrass Approximation Theorem that such a sequence of polynomials can be found and that, moreover, they can all be taken to have constant term zero. Then $P_n(A) \in \mathfrak{A}$ for all n, so the projection onto the range of A is in the ultrastrong closure of $\mathfrak{A}$. Now suppose A_1, A_2 are two self-adjoint elements of $\mathfrak{A}$ and let R_1, R_2 be the projections on to their ranges. Then $R_1 \vee R_2$ is the projection onto the range of $R_1 + R_2$ which, by the above argument, is a strong limit of a bounded sequence of operators which are obtained as polynomials in $R_1 + R_2$ with zero constant term. Thus, $R_1 \vee R_2$ is in the ultrastrong closure of $\mathfrak{A}$, and it is clear how to extend the argument to show that, if $A_1, \ldots, A_n$ is any finite set of elements of $\mathfrak{A}$, then the projection onto the closed linear span of union of the ranges of A_i is in the ultrastrong closure of $\mathfrak{A}$. Thus, P is a limit of an increasing net of projections in the ultrastrong closure of $\mathfrak{A}$ and therefore belongs to the ultrastrong closure of $\mathfrak{A}$.

We next prove a general principle asserting that, to approximate a self-adjoint operator, we can take the approximants to be self-adjoint.

PROPOSITION IV.C.5 *Let E be a vector subspace of $\mathscr{L}(\mathscr{H})$ which contains A^* if it contains A, and let B be a self-adjoint operator in the strong closure of E. Then there is a net of self-adjoint operators in E converging strongly to B.*

Proof Let A_α be a net in E converging strongly to B. Since taking adjoints is continuous in the *weak* operator topology, A_α^* converges weakly to $B^* = B$. Hence, $\frac{1}{2}(A_\alpha + A_\alpha^*)$ converges weakly to B. In other words, B is in the weak closure of the set of self-adjoint elements of E, and we want to show that it is in the strong closure. We do this by invoking the result that, for a *convex* set in $\mathscr{L}(\mathscr{H})$, the strong closure is the same as the weak closure (Proposition IV.B.2).

If $\mathfrak{A}$ is a self-adjoint algebra of operators containing $\mathbf{I}$, the von Neumann Density Theorem asserts that, given any $A \in \mathfrak{A}''$, there exists a net A_α in $\mathfrak{A}$ converging strongly to A. It does not, however, rule out the possibility that,

to approximate a given element A of $\mathfrak{A}''$, it might be necessary to use an *unbounded* net A_α. This is unfortunate since, if the net A_α is unbounded and converges to A, we cannot be sure that, for example, A_α^2 converges strongly to A^2. We will now prove a refinement of the von Neumann Density Theorem which eliminates all such problems:

THEOREM IV.C.5 (*Kaplansky Density Theorem*) *Let $\mathfrak{A}$ be a self-adjoint algebra of operators on a Hilbert space $\mathcal{H}$, and let $\overline{\mathfrak{A}}$ denote the strong closure of $\mathfrak{A}$. For any element A of $\overline{\mathfrak{A}}$, there exists a net A_α in $\mathfrak{A}$ such that:*

i) $\|A_\alpha\| \leqq \|A\|$ for all α.

ii) A_α converges strongly to A.

iii) A_α^* converges strongly to A^*.

If A is self-adjoint, A_α may be taken to be self-adjoint.

Proof We will first assume that A is self-adjoint. By scaling, we can assume that $\|A\| = 1$. Also, since any self-adjoint element of the norm closure of $\mathfrak{A}$ may be approximated arbitrarily well in norm by self-adjoint elements of $\mathfrak{A}$ of smaller norm, we may assume that $\mathfrak{A}$ is norm-closed, i.e., that $\mathfrak{A}$ is a C^* algebra (but we are not assuming that $\mathfrak{A}$ has an identity).

Consider the function

$$f(t) = \frac{2t}{1 + t^2} = \frac{2}{\frac{1}{t} + t}.$$

Its absolute value is nowhere greater than one; it sends zero to zero; and it maps the unit interval monotonically onto itself. Hence there exists a continuous function g on the unit interval such that

$$f(g(t)) = t \quad \text{for} \quad |t| \leqq 1.$$

Realizing the C^* algebra generated by A as an algebra of functions, we may construct $B = g(A)$ which is self-adjoint and which satisfies:

$$A = 2B(1 + B^2)^{-1}.$$

Since B is in the C^* algebra generated by A, B is in $\overline{\mathfrak{A}}$, so there is a net B_α of self-adjoint elements of $\mathfrak{A}$ converging strongly to B. The plan is to show that the net $A_\alpha = 2B_\alpha(1 + B_\alpha^2)^{-1}$ of self-adjoint elements of $\mathfrak{A}$ of norm not greater than one converges strongly to A. Thus, we look at:

$$\begin{aligned}
A - A_\alpha &= 2B(1 + B^2)^{-1} - 2B_\alpha(1 + B_\alpha^2)^{-1} \\
&= 2(1 + B_\alpha^2)^{-1}\,[(1 + B_\alpha^2)\,B - B_\alpha(1 + B^2)]\,(1 + B^2)^{-1} \\
&= 2(1 + B_\alpha^2)^{-1}\,[B - B_\alpha + B_\alpha^2 B - B_\alpha B^2]\,(1 + B^2)^{-1} \\
&= 2(1 + B_\alpha^2)^{-1}\,[(B - B_\alpha) + B_\alpha(B_\alpha - B)\cdot B]\,(1 + B^2)^{-1} \\
&= 2(1 + B_\alpha^2)^{-1}\,(B - B_\alpha)\,(1 + B^2)^{-1} \\
&\quad + 2B_\alpha(1 + B_\alpha^2)^{-1}\,(B_\alpha - B)\,B(1 + B^2)^{-1}
\end{aligned}$$

Now $(B_\alpha - B)(1 + B^2)^{-1}$ converges strongly to zero, and $\|(1 + B_\alpha^2)^{-1}\| \leqq 1$, $\|2B_\alpha(1 + B_\alpha^2)^{-1}\| \leqq 1$, so $A - A_\alpha$ converges strongly to zero, as asserted.

That, then, completes the proof if A is self-adjoint. To deal with the general case, we use a trick: Let $\mathfrak{A}_2((\overline{\mathfrak{A}})_2)$ be the self-adjoint algebra of operators on $\mathscr{H} \oplus \mathscr{H}$ given by matrices of the form $\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$, with $A_{11}, A_{12}, A_{21}, A_{22}$ in $\mathfrak{A}(\overline{\mathfrak{A}})$. It is easy to see that $(\overline{\mathfrak{A}_2}) = (\overline{\mathfrak{A}})_2$. Let A be an element of $\overline{\mathfrak{A}}$; then $\begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix}$ is a self-adjoint element of $\overline{\mathfrak{A}_2}$ with norm equal to $\|A\|$. Hence, there exists a net $\begin{pmatrix} A_{11}^{(\alpha)} & A_{12}^{(\alpha)} \\ A_{21}^{(\alpha)} & A_{22}^{(\alpha)} \end{pmatrix}$ of self-adjoint elements of $\mathfrak{A}_2$, with norm not greater than $\|A\|$ converging strongly to $\begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix}$. But then $\|A_{12}^{(\alpha)}\| \leqq \|A\|$ for all α, $(A_{12}^{(\alpha)})^* = A_{21}^{(\alpha)}$, $A_{12}^{(\alpha)}$ converges strongly to A, and $A_{21}^{(\alpha)} = (A_{12}^{(\alpha)})^*$ converges strongly to A^*, proving all the assertions of the theorem.

It is worth remarking that, if $\mathscr{H}$ is separable, we may replace "net" in the above theorem by "sequence". We will prove something slightly stronger: we will show that bounded sets in $\mathscr{L}(\mathscr{H})$ are metrizable in the strong topology if $\mathscr{H}$ is separable.

PROPOSITION IV.C.6 *Let $\mathscr{H}$ be a separable Hilbert space, $\mathscr{B}$ a bounded set in $\mathscr{L}(\mathscr{H})$. Then the topology induced on $\mathscr{B}$ by the strong topology on $\mathscr{L}(\mathscr{H})$ is metrizable. More precisely, there exists a norm $|||\cdot|||$ on $\mathscr{L}(\mathscr{H})$ such that a bounded net A_α converges strongly to A if and only if $\lim_\alpha |||A_\alpha - A||| = 0$.*

Proof Let (ξ_n) be a countable dense set in $\mathscr{H}$, such that no ξ_n is zero, and define

$$|||A||| = \sum_n \frac{1}{2^n} \cdot \frac{\|A\xi_n\|}{\|\xi_n\|}.$$

$|||\cdot|||$ is clearly a norm on $\mathscr{L}(\mathscr{H})$, and $|||A||| \leqq \|A\|$. We claim that, if A_α is a bounded net in $\mathscr{L}(\mathscr{H})$, then A_α converges strongly to A if and only if $|||A_\alpha - A||| \to 0$. (Warning: It is definitely *not* true that a general net A_α converges strongly to A if $|||A_\alpha - A||| \to 0$; the strong operator topology is not given by a single norm.) By passing to the net $A_\alpha - A$, we can assume $A = 0$. Now let A_α be a net converging strongly to zero, and assume $\|A_\alpha\| \leqq M$ for all α. Then

$$\limsup_\alpha |||A_\alpha||| \leqq \limsup_\alpha \left(\sum_{n=1}^{N} \frac{1}{2^n} \cdot \frac{\|A_\alpha \xi_n\|}{\|\xi_n\|} \right) + \limsup_\alpha \sum_{n=N+1}^{\infty} \frac{1}{2^n} \frac{\|A_\alpha \xi_n\|}{\|\xi_n\|}.$$

The first term on the right is zero since $\|A_\alpha \xi_n\| \to 0$ for all n; the second term on the right is $\leqq \dfrac{M}{2^N}$ since $\dfrac{\|A_\alpha \xi_n\|}{\|\xi_n\|} \leqq M$ for all n, α.

Thus,

$$\limsup_{\alpha} |||A_\alpha||| \leq \frac{M}{2^N} \quad \text{for all } N, \text{ so } \limsup_{\alpha} |||A_\alpha||| = 0.$$

Conversely, let A_α be a net satisfying $\|A_\alpha\| \leq M$ for all α, and suppose $|||A_\alpha||| \to 0$. We will show that $A_\alpha \to 0$ strongly. Thus, let $\xi \in \mathscr{H}$ and let $\varepsilon > 0$. Choose n such that

$$\|\xi_n - \xi\| < \varepsilon,$$

and choose α_0 such that

$$|||A_\alpha||| \leq \frac{\varepsilon}{2^n \cdot \|\xi_n\|} \text{ if } \alpha \geq \alpha_0.$$

Then, for

$$\alpha \geq \alpha_0, \quad \frac{1}{2^n} \cdot \frac{\|A_\alpha \xi_n\|}{\|\xi_n\|} \leq |||A_\alpha||| \leq \frac{\varepsilon}{2^n \cdot \|\xi_n\|},$$

so

$$\|A_\alpha \xi_n\| \leq \varepsilon.$$

Hence,

$$\|A_\alpha \xi\| \leq \|A_\alpha \xi - A_\alpha \xi_n\| + \|A_\alpha \xi_n\| \leq M \cdot \|\xi - \xi_n\| + \varepsilon = (M + 1)\varepsilon.$$

Since $(M + 1)\varepsilon$ may be made as small as desired, we have:

$$\lim_{\alpha} \|A_\alpha \xi\| = 0,$$

i.e., A_α converges strongly to zero.

As a corollary, we have:

COROLLARY IV. C.7 *Let $\mathfrak{A}$ be a self-adjoint algebra of operators on a separable Hilbert space, and let A be in the strong closure of $\mathfrak{A}$. Then there exists a sequence A_n in $\mathfrak{A}$ such that*

1) $\|A_n\| \leq \|A\|$ *for all* n.

2) A_n *converges strongly to* A.

3) A_n^* *converges strongly to* A^*.

Proof By the Kaplansky Density Theorem, we can find for each n an operator A_n in $\mathfrak{A}$ such that $\|A_n\| \leq \|A\|$ and such that

$$|||A_n - A||| \leq \frac{1}{n}; \qquad |||A_n^* - A^*||| \leq \frac{1}{n}.$$

Then A_n converges strongly to A and A_n^* converges strongly to A^*, by the preceding proposition.

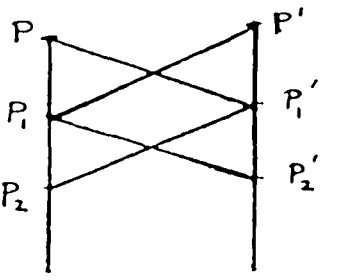
D Comparison of Projections in a von Neumann Algebra

We will begin with some motivation: Let $\mathfrak{A}$ be an algebra with involution, π a representation of $\mathfrak{A}$ on a Hilbert space $\mathscr{H}$, and let $\mathscr{B}$ be the von Neumann algebra $\pi(\mathfrak{A})'$. We will study the multiplicity of the representation π. To

begin with the most elementary of considerations, we note that a projection P is in $\mathcal{B}$ if and only if the subspace $P\mathcal{H}$ is invariant for π, i.e., that the closed invariant subspaces for π are parametrized by the projections in $\mathcal{B}$. Let P and P' be two projections in $\mathcal{B}$. The representation $\pi(\cdot) \mid P\mathcal{H}$ is unitarily equivalent to $\pi(\cdot) \mid P'\mathcal{H}$ if and only if there exists a unitary operator W from $P\mathcal{H}$ to $P'\mathcal{H}$, such that $\pi(\cdot)\, W = W\pi(\cdot)$ on $P\mathcal{H}$. Extending W to be zero on $(\mathbf{1} - P)\,\mathcal{H}$, we get a partial isometry with initial subspace $P\mathcal{H}$ and terminal subspace $P'\mathcal{H}$, and this partial isometry commutes with $\pi(A)$ for all A, i.e., belongs to $\mathcal{B}$. To summarize: The representation $\pi(\cdot) \mid P\mathcal{H}$ is unitarily equivalent to $\pi(\cdot) \mid P'\mathcal{H}$ if and only if there is a partial isometry W in $\mathcal{B}$ with initial subspace $P\mathcal{H}$ and terminal subspace $P'\mathcal{H}$. We thus make the following definition: Let $\mathcal{B}$ be a von Neumann algebra, P and P' two projections in $\mathcal{B}$. We say that P is *equivalent* to $P'(P \simeq P')$ if there is a partial isometry W in $\mathcal{B}$ with initial subspace $P\mathcal{H}$ and terminal subspace $P'\mathcal{H}$, i.e., if there is an operator W in $\mathcal{B}$ such that $W^*W = P$; $WW^* = P'$. (If $\mathcal{B} = \mathcal{L}(\mathcal{H})$, two projections are equivalent if and only if their ranges have the same dimension.) We extend this definition slightly by defining $P \prec P'$ or $P' \succ P$ to mean that P is equivalent to a subprojection of P' (i.e., to a projection in $\mathcal{B}$ whose range is contained in the range of P'). It is important to distinguish between $P' \geqq P$ (P is a subprojection of P') and $P' \succ P$ (P is equivalent to a subprojection of P'). The relation $P' \prec P$ clearly defines a pre-order on the set of equivalence classes of projections; the following proposition asserts that it actually defines an order:

PROPOSITION IV. D.1 *Let $\mathcal{B}$ be a Neumann algebra, P and P' two projections in $\mathcal{B}$ such that $P \succ P'$ and $P' \succ P$. Then $P \simeq P'$.*

Proof We have $P \simeq P_1' \leqq P'$ and $P' \simeq P_1 \leqq P$. Composing the partial isometry W taking P to P_1' with the partial isometry W' taking P' to P_1 gives partial isometry taking P to a subprojection P_2 of P_1, i.e., $P \simeq P_2$; Similarly, we get $P' \simeq P_2'$. The partial isometry taking P' to P_1 takes $P' - P_1'$ to $P_1 - P_2$ (i.e., the range of $W'(P' - P_1')$ is $P_1 - P_2$) so $P' - P_1' \simeq P_1 - P_2$, and similarly, $P - P_1 \simeq P_1' - P_2'$. We may represent the situation schematically by a diagram:



Lines sloping downward to the right indicate the action of the partial isometry W taking P to P_1'; lines sloping downward to the left indicate the action of the partial isometry W' taking P' to P_1, and opposite vertical ends of parallelograms represent equivalent projections. We may continue the diagram downward by defining P_3 as the range of $W'P_2'$, P_3' as the range of WP_2, etc. We get thus two decreasing sequences of projections:

$$P \geqq P_1 \geqq P_2 \geqq \cdots$$

$$P' \geqq P_1' \geqq P_2' \geqq \cdots$$

such that the range of $WP_i = P_{i+1}'$ and the range of $W'P_i' = P_{i+1}$. Then $P_i - P_{i+1} \simeq P_{i+1}' - P_{i+1}' - P_{i+2} - P_{i+2} \simeq P_i' - P_{i+1}'$ for $i = 0, 1, 2, \ldots$ (We let $P_0 = P$; $P_0' = P'$). Furthermore, let

$$P_\infty = \bigcap_i P_i \quad \text{and} \quad P_\infty' = \bigcap_i P_i' = \bigcap_i P_{i+1}'.$$

Since the range of WP_i is P_{i+1}', the range of $W \cdot P_\infty$ is exactly P_∞', i.e., P_∞ is equivalent to P_∞'. We may now write:

$$\begin{array}{ccccccccccc} P & = & P_\infty & + & (P & - & P_1) & + & (P_1 - P_2) & + & (P_2 - P_3) + (P_3 - P_4) + \cdots \\ & & \simeq & & & & \simeq & & \simeq & & \simeq \quad \simeq \\ P' & = & P_\infty' & + & (P_1' & - & P_2') & + & (P' - P_1') & + & (P_3' - P_4') + (P_2' - P_3') + \cdots \end{array}$$

A simple argument (which we omit) shows that P is equivalent to P'.

E Disjointness of Projections and Central Projections

We have indicated that our investigation of the structure of von Neumann algebras will be motivated by regarding the von Neumann algebra $\mathscr{B}$ in question as the commutant of a representation π of some algebra $\mathfrak{A}$, and translating questions about subrepresentations of π into questions about the von Neumann algebra $\mathscr{B}$. Let us elaborate on this approach a little by asking what the situation is if $\mathfrak{A}$ has the property that every representation may be written as a direct sum of irreducible representations. (This will be true, for example, if $\mathfrak{A}$ is finite-dimensional). In this case, every projection P in $\mathscr{B}$ can be written as a sum of mutually orthogonal projections onto irreducible subspaces, and two projections P and P' are equivalent if and only if each irreducible representation of $\mathfrak{A}$ appears the same number of times in $\pi(\cdot) \mid P\mathscr{H}$ as in $\pi(\cdot) \mid P'\mathscr{H}$. Thus, the structure of the set of equivalence classes of projections is entirely transparent. If we drop the requirement that every representation of $\mathfrak{A}$ decompose into a direct sum of irreducible representations, the situation becomes much more intricate. For example, the algebra of continuous functions on [0, 1] represent[illegible]

multiplication operators on $\mathscr{L}^2$ of Lebesgue measure on [0, 1], has no irreducible subspaces. Much more peculiar things can also happen, e.g., there exist C^* algebras $\mathfrak{A}$ and representations π of $\mathfrak{A}$ which are not irreducible, but which have the property that any two non-zero subrepresentations of $\mathfrak{A}$ are unitarily equivalent. (In fact, this is what occurs in quantum statistical mechanics for the cyclic representation of the algebra of quasilocal observables associated with an equilibrium state representing a pure phase at non-zero temperature.)

We will approach the general situation in the following way: We look for properties of subrepresentations of π which have a straightforward interpretation if π can be decomposed into a direct sum of irreducible representations, but which can be formulated without mentioning irreducible representations. We then translate these properties into properties of projections in the von Neumann algebra $\mathscr{B} = \pi(\mathfrak{A})'$, and make them into definitions which no longer depend on regarding $\mathscr{B}$ as the commutant of a representation. Of course, there is no way to see at the outset which properties of subrepresentations of π can *fruitfully* be translated into definitions in general von Neumann algebras....

We have already seen one example of this procedure: Unitary equivalence of subrepresentations of π translates into the definition of equivalence of projections in a von Neumann algebra. As a second example: If every representation of $\mathfrak{A}$ decomposes into a direct sum of irreducible representations, we will say that two subrepresentations of π are disjoint if no irreducible representation appearing in one of them also appears in the other. This we translate into: Two projections P and P' in a von Neumann algebra $\mathscr{B}$ are *disjoint* if no non-zero subprojection of P is equivalent to a subprojection of P'. Conversely, we say that P *covers* Q ($P \succ\succ Q$ or $Q \prec\prec P$) if no non-zero subprojection of Q is disjoint from P (Intuitively, if every irreducible contained in $\pi(\cdot) \mid Q\mathscr{H}$ is contained in $\pi(\cdot) \mid P\mathscr{H}$) and that P is *quasi-equivalent* to Q ($P \sim Q$) if P covers Q and Q covers P. One can thus think of quasi-equivalent projections as corresponding to subrepresentations which are the same except for multiplicity.

The key to analyzing the notion of disjointness is contained in the following proposition, which is really just a variant of Schur's Lemma.

PROPOSITION IV. E.1 *Let P and Q be projections in a von Neumann algebra $\mathscr{B}$. In order that P be disjoint from Q it is necessary and sufficient that $PBQ = 0$ for all $B \in \mathscr{B}$. In particular, if P is disjoint from Q, then $PQ = 0$, i.e., P is orthogonal to Q.*

Proof Assume first that P is not disjoint from Q, and let W be a non-zero partial isometry in $\mathscr{B}$ with initial domain contained in $Q\mathscr{H}$ and terminal domain contained in $P\mathscr{H}$. Then $W = PWQ$, so $PWQ \neq 0$, i.e., W is an element of $\mathscr{B}$ such that $PWQ \neq 0$. Next, assume that, for some $B \in \mathscr{B}$, $PBQ \neq 0$. Then $PBQ \in \mathscr{B}$ and we may use the polar decomposition to write

$PBQ = W \cdot |PBQ|$ where W is a partial isometry with initial subspace the orthogonal complement of the null-space of PBQ (which is contained in $Q\mathcal{H}$ since $PBQ = 0$ on $(\mathbf{1} - Q)\mathcal{H}$) and terminal subspace the closure of the range of PBQ (which is contained in $P\mathcal{H}$). Moreover, W is in $\mathcal{A}$ since it is the partial isometric part of PBQ which is in $\mathcal{A}$. Hence, Q is not disjoint from P (since the initial subspace of W is equivalent to the terminal subspace of W).

In any algebra the *center* of $\mathcal{A}$, denoted by $\mathfrak{Z}(\mathcal{A})$, is the set of elements of $\mathcal{A}$ commuting with all elements of $\mathcal{A}$. If $\mathcal{A}$ is a von Neumann algebra, we may equivalently define $\mathfrak{Z}(\mathcal{A}) = \mathcal{A} \cap \mathcal{A}'$. Thus $\mathfrak{Z}(\mathcal{A})$ is a commutative von Neumann algebra contained in $\mathcal{A}$.

COROLLARY IV. E.2 *Let $\mathcal{A}$ be a Neumann algebra, P a projection in $\mathcal{A}$. Then P is disjoint from $\mathbf{1} - P$ if and only if $P \in \mathfrak{Z}(\mathcal{A})$.*

Proof Suppose first that $P \in \mathfrak{Z}(\mathcal{A})$. Then, for all $B \in \mathcal{A}$, $PB(\mathbf{1} - P) = B \cdot P \cdot (\mathbf{1} - P) = 0$, so P is disjoint from $\mathbf{1} - P$. Conversely, if P is disjoint from $\mathbf{1} - P$, then $O = P \cdot B \cdot (\mathbf{1} - P) = PB - PBP$ i.e., $PB = PBP$ for all $B \in \mathcal{A}$. Taking adjoints gives

$$B^*P = PB^*P = PB^*$$

for all $B \in \mathcal{A}$, so $B^*P = PB^*$ for all $B \in \mathcal{A}$, so $P \in \mathfrak{Z}(\mathcal{A})$.

COROLLARY IV. E.3 *Let $\mathcal{A}$ be a von Neumann algebra, P a projection in $\mathcal{A}$. Let $(Q_i)_{i \in I}$ be a family of projections in $\mathcal{A}$ each of which is disjoint from P. Then $\bigvee_i Q_i$ is disjoint from P.*

Proof Since $Q_i BP = 0$ for all $B \in \mathcal{A}$, $\left(\bigvee_i Q_i\right) BP = 0$ for all $B \in \mathcal{A}$, so $\bigvee_i Q_i$ is disjoint from P.

PROPOSITION IV. E.4 *Let P be a projection in a von Neumann algebra $\mathcal{A}$. Then there exists a unique projection $Q \in \mathcal{A}$, $Q \geqq P$, such that P is quasi-equivalent to Q and disjoint from $\mathbf{1} - Q$. Furthermore, $Q \in \mathfrak{Z}(\mathcal{A})$.*

Proof Intuitively, we may think of Q as the projection onto the subspace generated by all irreducible subspaces equivalent to an irreducible subspace contained in P, and $\mathbf{1} - Q$ as the projection onto the subspace generated by all irreducible subspaces which are not equivalent to an irreducible subspace contained in P. The latter description gives the key to constructing Q; we define $\mathbf{1} - Q$ to be the supremum of all projections in $\mathcal{A}$ disjoint from P. Then $\mathbf{1} - Q$ is disjoint from P, and thus in particular orthogonal to P, by the preceding proposition. By the construction of Q, no non-zero subprojection of Q can be disjoint from P (otherwise, this projection would be in $\mathbf{1} - Q$), so P covers Q. On the other hand $Q \geqq P$, so Q covers P and thus Q is quasi-equivalent to P. This proves the existence of Q. Moreover, Q is

disjoint from $\mathbf{1} - Q$, *so* $Q \in \mathcal{B}(\mathcal{A})$. To prove the uniqueness, let Q' be another projection such that P is quasi-equivalent to Q' and disjoint from $\mathbf{1} - Q'$. Then Q' must also belong to $\mathcal{Z}(\mathcal{A})$, so $Q'(\mathbf{1} - Q)$ is a projection. Now $Q'(\mathbf{1} - Q)$ is covered by P (since it is a subprojection of Q') and disjoint from P (since it is a subprojection of $\mathbf{1} - Q$.) The only possibility, then, is that $Q'(\mathbf{1} - Q) = 0$, i.e., $Q' = Q'Q$. Interchanging the roles of Q and Q' we see that $Q = QQ'$, so $Q = Q'$ and uniqueness is proved.

The projection Q constructed in the preceding proposition is called the *central support* of P. There are some other ways to characterize it:

PROPOSITION IV. E.5 *Let P be a projection in the von Neumann algebra $\mathcal{A}$, and let Q be its central support. Then if Q' is any projection in $\mathcal{Z}(\mathcal{A})$ containing P, $Q' \geq Q$ (i.e., Q is the smallest projection in $\mathcal{Z}(\mathcal{A})$ containing P). Q is also equal to the projection onto the closed linear span of the set of vectors obtained by applying elements of $\mathcal{A}$ to elements of $P\mathcal{H}$ (we denote this closed linear span by $[\mathcal{A}P\mathcal{H}]$).*

Proof Let Q' be a projection in $\mathcal{Z}(\mathcal{A})$ which contains P. Then QQ' is a central projection containing P and contained in Q; hence it is quasi-equivalent to Q (we have $Q \succ\succ QQ' \geq P \succ\succ Q$). On the other hand $Q - QQ' \leq \mathbf{1} - QQ'$ and is therefore disjoint from QQ'. Thus we have $Q - QQ'$ is covered by QQ' and on the other hand $Q - QQ'$ is disjoint from QQ', so we must have $Q - QQ' = 0$, i. e. $Q = QQ'$, i.e. $Q' \geq Q$.

Now let P' denote the projection onto $[\mathcal{A}P\mathcal{H}]$. Since $[\mathcal{A}P\mathcal{H}]$ is invariant for $\mathcal{A}$, $P' \in \mathcal{A}'$. If Q is the central support of P, if $B \in \mathcal{A}$, and if $\xi \in \mathcal{H}$, then $QBP\xi = BQP\xi$ (since $Q \in \mathcal{Z}(\mathcal{A})$) $= BP\xi$ (since $Q \geq P$); hence $Q \geq P'$. To prove that $Q = P'$, we have thus only to show that $P' \in \mathcal{Z}(\mathcal{A})$, and we already know that $P' \in \mathcal{A}'$. Thus, what we have to show is that $[\mathcal{A}P\mathcal{H}]$ is invariant for $\mathcal{A}'$ (since this will imply $P' \in (\mathcal{A}')' = \mathcal{A}$.) But if $B' \in \mathcal{A}'$, $B \in \mathcal{A}$, and $\xi \in P\mathcal{H}$, $B'B\xi = BB'\xi \in [\mathcal{A}P\mathcal{H}]$ since $B'P\mathcal{H} \subset P\mathcal{H}$. Thus, B' maps $[\mathcal{A}P\mathcal{H}]$ into itself, so $P' \in \mathcal{A}$, so $P' = Q$.

COROLLARY IV. E.6 *Let P be a projection in a von Neumann algebra $\mathcal{A}$, and let Q be the central support of P. Then, for $B' \in \mathcal{A}'$, $B'|P\mathcal{H} = 0$ implies $B'|Q\mathcal{H} = 0$; i.e., the mapping $B'|Q\mathcal{H} \to B'|P\mathcal{H}$ defines an isomorphism of the algebra $B'|Q\mathcal{H}$ to the algebra $B'|P\mathcal{H}$.*

Proof If $B'|P\mathcal{H} = 0$, then, for any $B \in \mathcal{A}$, $B'BP\mathcal{H} = BB'P\mathcal{H} = 0$, so $B'|Q\mathcal{H} = 0$.

We can now analyze more or less completely the structure of the set of quasi-equivalence classes of projections.

1) Every projection is quasi-equivalent to a central projection, its central support.

2) If Q, Q' are central projections, then Q covers Q' if and only if $Q \geqq Q'$ (and hence Q is quasi-equivalent to Q' if and only if $Q = Q'$.)

Proof Suppose Q covers Q'; then since $(\mathbf{1} - Q) \cdot Q'$ is contained in Q' and disjoint from Q, it must be zero, so $Q' = QQ'$, i.e., $Q \geqq Q'$. Conversely, if $Q \geqq Q'$, the Q surely covers Q'.

3) Combining 1) and 2) we see that every quasi-equivalence class of projections contains exactly one central projection, i.e., the quasi-equivalence classes of projections are parametrized by the central projections and the relation $\succ\succ$ just translates into the relation $\geqq$.

4) If P_1, P_2 are projections in $\mathcal{B}$, then P_1 covers P_2 if and only if, for $B' \in \mathcal{B}'$, $B' \mid P_1\mathcal{H} = 0$ implies $B' \mid P_2\mathcal{H} = 0$, and P_1 is quasi-equivalent to P_2 if and only if the mapping $B' \mid P_1\mathcal{H} \leftrightarrow B' \mid P_2\mathcal{H}$ defines an isomosphism from $\mathcal{B}' P_1 \mid \mathcal{H}$ to $\mathcal{B}' P_2 \mid \mathcal{H}$.

From remark 4), it is a short step, but not a trivial one, to the following characterization of quasi-equivalence: Let $\mathfrak{A}$ be an algebra, π a representation of $\mathfrak{A}$, $\mathcal{B} = \pi(\mathfrak{A})$, P, P' two projections in $\mathcal{B}$. Then P is quasi-equivalent to P' if and only if there exists an isomorphism ϕ from the von Neumann algebra on $P\mathcal{H}$ generated by $\pi(\mathfrak{A}) \mid P\mathcal{H}$ to the von Neumann algebra on $P'\mathcal{H}$ generated by $\pi(\mathfrak{A}) \mid P'\mathcal{H}$, such that $\phi(\pi(A) \mid P\mathcal{H}) = \pi(A) \mid P'\mathcal{H}$ for all $A \in \mathfrak{A}$. To prove the "only if" statement from remark 4), we have only to show that the von Neumann algebra generated by $\pi(\mathfrak{A}) \mid P\mathcal{H}$ is just $\mathcal{B}' \mid P\mathcal{H}$, and this follows easily from the bicommutant theorem. To prove the "if" statement, we have to show that if ϕ is an isomorphism from $\mathcal{B}' \mid P\mathcal{H}$ to $\mathcal{B}' \mid P'\mathcal{H}$ extending the mapping $\pi(A)|P\mathcal{H} \leftrightarrow \pi(A)| P'\mathcal{H}$, then ϕ must be given the same formula on all of $\mathcal{B}' \mid P\mathcal{H}$. This follows easily from the fact that an isomorphism of von Neumann algebras is ultraweakly continuous, which we will not prove (See Dixmier, Av N., p. 57, Corollaire 1.)

There are some remarks to be made about the significance of the center $\mathcal{Z}(\mathcal{B})$ of a von Neumann algebra $\mathcal{B}$. First let P be a projection in $\mathcal{B}'$. Then, since $P\mathcal{H}$ and $(\mathbf{1} - P)\mathcal{H}$ are both invariant for $\mathcal{B}$, we may decompose $\mathcal{H} = P\mathcal{H} \oplus (\mathbf{1} - P)\mathcal{H}$ and get an analogous decomposition for every element of $\mathcal{B}$: $B = \begin{pmatrix} B_1 & 0 \\ 0 & B_1 \end{pmatrix}$. We can now ask what will be special about this decomposition if P is in $\mathcal{Z}(\mathcal{B})$, i.e., if it is in $\mathcal{B}$ as well as in $\mathcal{B}'$. It is not hard to see that P is in $\mathcal{Z}(\mathcal{B})$ if and only if the entries B_1 and B_2 in matrices $\begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix}$ representing elements of $\mathcal{B}$ are independent. By "independent" we mean that, given any two elements B', B'' in $\mathcal{B}$, we can find an element B of $\mathcal{B}$ such that

$$B \mid P\mathcal{H} = B' \mid P\mathcal{H}$$

$$B|(\mathbf{1} - P)\mathcal{H} = B''|(\mathbf{1} - P)\mathcal{H}.$$

In other words, if P is in the center of $\mathcal{B}$, then $\mathcal{B}$ is isomorphic to $\mathcal{B} \mid P\mathcal{H} \oplus \mathcal{B}(\mathbf{1} - P)\mathcal{H}$. Thus, we can study the von Neumann algebras $\mathcal{B} \mid P\mathcal{H}$ and $\mathcal{B} \mid (\mathbf{1} - P)\mathcal{H}$ separately, then reconstruct $\mathcal{B}$ from them. The algebras $\mathcal{B} \mid P\mathcal{H}$ and $\mathcal{B} \mid (\mathbf{1} - P)$ are smaller than $\mathcal{B}$, and ought to be less complicated. In particular, if it happens that P and $\mathbf{1} - P$ are the only non-trivial projections in $\mathcal{Z}(\mathcal{B})$, then it is easy to see that the centers of $\mathcal{B} \mid P\mathcal{H}$ and $\mathcal{B} \mid (\mathbf{1} - P)\mathcal{H}$ are both trivial, i.e., consist only of the scalars. A von Neumann algebra $\mathcal{B}$ such that $\mathcal{Z}(\mathcal{B}) = \{\lambda\mathbf{1}\}$ is called a *factor*, so what we have just shown is that, if $\mathcal{B}$ is a von Neumann algebra whose center is generated by a single non-trivial projection (i.e., whose center is two-dimensional), then $\mathcal{B}$ may be decomposed into a direct sum of two factors. Evidently, this argument extends trivially to any von Neumann algebra with finite-dimensional center (if $\mathcal{Z}(\mathcal{B})$ is n dimensional, $\mathcal{B}$ may be written as a direct sum of n factors) and even to a Neumann algebra whose center is generated by a countable set of mutually orthogonal projections. The center of $\mathcal{B}$ need not have this property, however, and it is therefore necessary to pass from direct sums to direct integrals. Very crudely speaking, the general situation (subject to certain technical restrictions) is the following: Starting from a Hilbert space $\mathcal{H}$ and a von Neumann algebra $\mathcal{B}$ on $\mathcal{H}$, one can find a measure space (X, μ), and, for each $x \in X$, a Hilbert space $\mathcal{H}_x$ and a factor $\mathcal{B}_x$ on $\mathcal{H}_x$ such that:

1) The Hilbert space $\mathcal{H}$ may be realized as the space of "all" functions $x \mapsto \xi_x \in \mathcal{H}_x$ such that $\int d\mu(x) \| \xi_x \|^2 < \infty$. The word "all" is in quotation marks since we really only want to consider only functions which are in some sense measurable; the way in which "measurable" should be defined is not at all obvious since the Hilbert spaces $\mathcal{H}_x$ for different x are to be thought of as different spaces.

2) In this realization, the elements of $\mathcal{B}$ are in one-one correspondence with functions $x \mapsto B_x \in \mathcal{B}_x$, where these functions are subjected to the condition $\sup_x \| B_x \| < \infty$ and to some measurability condition; the correspondence is defined by $(B\xi)_x = B_x \xi_x$.

The above theory, known as reduction theory, in some sense reduces the theory of general von Neumann algebras to the theory of factors. It has however two disadvantages: It is valid only under some restrictions on the size of the von Neumann algebra (these restrictions are not too serious for physical applications since they are always satisfied if the space on which $\mathcal{B}$ acts is separable), and applying it frequently leads to unpleasant problems of measurability. For many purposes, therefore, it is both simpler and clearer to make arguments directly about $\mathcal{B}$ itself rather than invoking the decomposition of $\mathcal{B}$ into factors. Nevertheless, thinking of a general von Neumann algebra as some kind of generalized direct sum of factors is frequently an invaluable heuristic device.

F The Comparability Theorem

We now want to get some idea of what the collection of equivalence classes of projections in $\mathscr{B}$ looks like as a partially ordered set. For orientation, we go back to the pictures of $\mathscr{B}$ as the commutant of a representation π of a C^* algebra $\mathfrak{A}$ and we assume, for heuristic purposes, that every representation of $\mathfrak{A}$ may be written as a direct sum of irreducible representations. Let I denote the set of equivalence classes of irreducible representations of $\mathfrak{A}$. For each projection $P \in \mathscr{B} = \pi(\mathfrak{A})'$, we associate an indexed set $(n_i)_{i \in I}$ of cardinal numbers, n_i being the number of times the representation i appears in a direct sum decomposition of $\pi(\cdot) \mid P\mathscr{H}$. If P' is another projection, and (n_i') is its set of multiplicities, then $P' \succ P$ if and only if $n_i' \geq n_i$ for each i. Hence, the relation $\succ$ does not linearly order the set of projections in $\mathscr{B}$, unless π is a direct sum of copies of a single irreducible representation. Moreover, in the simplified case we are considering, π is a direct sum of copies of a single irreducible representation if and only if no two non-zero projections in $\mathscr{B}$ are disjoint, i.e., if and only if the only central projections in $\mathscr{B}$ are 0 and $\mathbf{1}$, i.e. (since $\mathscr{Z}(\mathscr{B})$ is generated by its projections) if and only if $\mathscr{B}$ is a factor. We are thus led to the general conjecture that, if $\mathscr{B}$ is a factor, the set of projections in $\mathscr{B}$ is linearly ordered by $\succ$. In this conjecture of course, all mention of irreducible representations has disappeared. To prove the conjecture, we will prove the following slightly stronger result, which is also easy to interpret in the simplified case considered above:

THEOREM IV. F.1 *Let $\mathscr{B}$ be a von Neumann algebra, P and Q two projections in $\mathscr{B}$. There exists a central projection R such that $RP \succ RQ$; $(\mathbf{1} - R)\, Q \succ (\mathbf{1} - R)\, P$. In particular, if $\mathscr{B}$ is a factor (so either $R = \mathbf{1}$ or $\mathbf{1} - R = \mathbf{1}$), we have either $P \succ Q$ or $Q \succ P$. In other words, any two projections in a factor are comparable.*

Proof The idea of the proof is simple: We first find projections $S \leqq P$, $T \leq Q$ which are equivalent and as large as possible; we then see what can be done with what is left over. To carry out the first step, we form the set whose elements are sets $\{(S_i, T_i)\}$ of pairs of non-zero projections in $\mathscr{B}$ such that

1) $S_i \leqq P$, $T_i \leqq Q$ for all i.

2) $S_i \simeq T_i$ for all i.

3) If $i \neq j$, S_i is orthogonal to S_j and T_i is orthogonal to T_j.

Since the elements of this set are sets (of pairs of projections), we can order the set by inclusion: $\{(S_i, T_i)\} \leqq \{(S_k', T_k')\}$ means that each (S_i, T_i) is equal to (S_k', T_k') for some k. It follows at once from the definitions that the union of a linearly ordered family of such sets of pairs of projections is again such a set of pairs of projections; hence by Zorn's Lemma, there exists a maximal

such set $\{(S_t, T_t)\}$. Let $S = \sum_t S_t$; $T = \sum_t T_t$. Then $S \leqq P$, $T \leqq Q$, and $S \simeq T$. Furthermore, $P - S$ is disjoint from $Q - T$; (If this were not true, there would exist a non-zero projection $S_0 \leqq P - S$ which is equivalent to $T_0 \leqq Q - T$; then $\{(S_t, T_t)\} \cup \{(S_0, T_0)\}$ properly contains $\{(S_t, T_t)\}_{t \neq 0}$ and therefore violates maximality.) In a factor, no two non-zero projections can be disjoint, and this implies that either $P - S = 0$ (i.e., $P = S \simeq T \leqq Q$, so $P \prec Q$) or $T - Q = 0$ (i.e., $Q = T \simeq S \leqq P$, so $P \succ Q$), so if $\mathfrak{B}$ is a factor, either $P \prec Q$ or $Q \prec P$. To deal with the general case, let R be the central support of $P - S$. Then $R(P - S) = P - S$, and $R(Q - T) = 0$ since $Q - T$ is disjoint from $P - S$. Thus:

$$RP = R(P - S) + RS = P - S + RS,$$

while

$$RQ = R(Q - T) + RT = RT.$$

But since $R \in \mathfrak{Z}(\mathfrak{B})$ and $S \simeq T$, $RS \simeq RT$ (Proof. Let W be a partial isometry in $\mathfrak{B}$ with initial subspace $S\mathcal{H}$ and terminal subspace $T\mathcal{H}$. Then RWR is a partial isometry in $\mathfrak{B}$ with initial subspace $RS\mathcal{H}$ and terminal subspace $RT\mathcal{H}$). Thus, $RQ \prec RP$. Similarly,

$$(\mathbf{1} - R)P = (\mathbf{1} - R)(P - S) + (\mathbf{1} - R)S = (\mathbf{1} - R) \cdot S,$$

$$(\mathbf{1} - R)Q = (\mathbf{1} - R)(Q - T) + (\mathbf{1} - R)T = Q - T + (\mathbf{1} - R)T,$$

so

$$(\mathbf{1} - R)P \prec (\mathbf{1} - R)Q.$$

G Finite and Infinite Projections

Let π be a representation of an algebra $\mathfrak{A}$ with involution, and assume that every representation of $\mathfrak{A}$ can be written as a sum of irreducibles. We will say that π is finite if it contains at most finitely many copies of each irreducible representation of $\mathfrak{A}$; otherwise, we say π is infinite. It is easy to see that a representation is infinite if and only if it is unitarily equivalent to a proper subrepresentation of itself. Translating, in the usual way, statements about representations into statements about projections in von Neumann algebras, we say that a projection P in a von Neumann algebra $\mathfrak{B}$ is *infinite* if there exists a projection Q in $\mathfrak{B}$; $Q \lneq P$, such that $Q \simeq P$, and we say that a projection P is finite if it is not infinite. Note that, if P is an infinite projection and if $P' \geqq P$, then P' is also infinite (If Q is a proper subprojection of P equivalent to P, then $P' = (P' - P) + P \simeq (P' - P) + Q$.) We say that the von Neumann algebra $\mathfrak{B}$ is *finite* (*infinite*) if the projection $\mathbf{1}$ is finite (infinite). The algebra $\mathfrak{L}(\mathcal{H})$ is finite if and only if $\mathcal{H}$ is finite dimensional.

For some purposes, it is useful to have a strengthening of the notion of an infinite projection. In the simple case where $\mathfrak{A}$ is an algebra all representations

of which can be written as direct sums of irreducible representations, we have defined a representation π of $\mathfrak{A}$ to be infinite if it contains infinitely many copies of at least one irreducible representation. In the same context, we will say that a representation is properly infinite if every irreducible representation which appears once in it appears there infinitely often. We thus make the following definition: a projection P in a von Neumann algebra is *properly infinite* if, for every central projection R of $\mathfrak{B}$, either $RP = 0$ or RP is infinite.

There is another distinction, related to finiteness, which has no analogue in the irreducible representations example: We will say that a projection P in a von Neumann algebra $\mathfrak{B}$ is *purely infinite* if every non-zero subprojection of P is infinite and *semi-finite* if every non-zero subprojection of P contains a finite non-zero subprojection. We will say that a von Neumann algebra is *properly infinite* (*purely infinite, semifinite*) if the projection $\mathbf{1}$ is properly infinite (purely infinite, semifinite). It is not at all evident at this point that purely infinite von Neumann algebras exist.

We next show that every von Neumann algebra can be split uniquely into a finite part, a properly infinite but semifinite part, and a purely infinite part.

PROPOSITION IV. G.1 *Let $\mathfrak{B}$ be a von Neumann algebra. There exist three central projection R_f, $R_{s.f.}$ and $R_{p.i.}$ such that $\mathbf{1} = R_f + R_{s.f.} + R_{p.i.}$ and such that R_f is finite, $R_{s.f.}$ is semi-finite, and $R_{p.i.}$ is purely infinite. Furthermore, these projections are uniquely determined.*

Proof If $(R_t)_{t\in I}$ is a family of mutually orthogonal finite central projections, $\sum_{t\in I} R_t$ is finite. (*Proof* If $S \precsim T \leqq \sum_t R_t$, and if $S \simeq T$, then $SR_t \precsim TR_t$ for some t, so TR_t is an infinite projection contained in R_t, contradicting the assumed finiteness of R_t.) Hence, Zorn's Lemma gives the existence of a maximal finite central projection R_f. Also, if $(R_t)_{t\in I}$ is a mutually orthogonal family of purely infinite central projections, then $\sum_{t\in I} R_t$ is purely infinite. (If $Q \leqq \sum_{t\in I} R_t$, then $QR_t \neq 0$ for some t, so QR_t is a non-zero subprojection of the purely infinite projection R_t, so QR_t is infinite, but $QR_t \leqq Q$, so Q is infinite). Again by Zorn's Lemma, there exists a maximal purely infinite central projection $R_{p.i.}$, and clearly $R_{p.i.}R_{f.}$, being both finite and purely infinite, must be zero, so $R_{p.i.} \leqq \mathbf{1} - R_{f.}$. Define $R_{s.f.} = \mathbf{1} - R_{f.} - R_{p.i.}$. Then $R_{s.f.}$ is properly infinite (It is a central subprojection of the properly infinite central projection $\mathbf{1} - R_{f.}$), but it contains no purely infinite central projection. Now if Q is any purely infinite projection, the central support of Q is also purely infinite. (Any subprojection contained in the central support of Q has a non-zero sub-

projection equivalent to a subprojection of Q; hence, must be infinite.) Thus, $R_{s.f.}$ contains no purely infinite subprojection, i.e., is semifinite. The uniqueness is proved by remarking that $R_{f.}$ must contain every finite central projection and that $R_{p.i.}$ must contain every purely infinite central projection.

LEMMA IV. G.2 *A projection P in a von Neumann algebra $\mathscr{A}$ is infinite if and only if there exists a mutually orthogonal sequence of non-zero projections (Q_n) such that $Q_n \leqq P$ for all n and $Q_n \simeq Q_m$ for all n, m.*

Proof If projections (Q_n) exist, then $\sum_{n=1}^{\infty} Q_n \simeq \sum_{n=2}^{\infty} Q_n$, so $\sum_{n=1}^{\infty} Q_n$ is infinite, so $P \geqq \sum_{n=1}^{\infty} Q_n$ is infinite.

Now let P be infinite, and let W be a partial isometry in $\mathscr{A}$ with initial subspace $P\mathscr{H}$ and terminal subspace $P_1\mathscr{H}$, where $P_1 \leqq P$. Define a decreasing sequence P_i of projections by $P_{i+1}\mathscr{H} = WP_i\mathscr{H}$. Then $W(P_i - P_{i+1})$ is a partial isometry with initial subspace $(P_i - P_{i+1})\mathscr{H}$ and terminal subspace $(P_{i+1} - P_{i+2})\mathscr{H}$; hence $Q_1 = P - P_1 \simeq Q_2 = P_1 - P_2 \simeq Q_3 = P_2 - P_3 \ldots$

We now come to what seems to be the most delicate point on the theory of comparison of projections: The proof that the supremum of a finite number of finite projections is finite. The key to the proof is contained in the following lemma:

LEMMA IV. G.3 *Let (P_1, P_2) and (Q_1, Q_2) be two pairs of projections in the von Neumann algebra $\mathscr{A}$. Assume that P_1 is orthogonal to P_2, that Q_1 is orthogonal to Q_2, and that $P_1 + P_2 = Q_1 + Q_2$.*
Then there is a central projection R in $\mathscr{A}$ such that

$$RQ_1 \prec RP_1, (\mathbf{1} - R)Q_2 \prec (\mathbf{1} - R)P_2.$$

(If $\mathscr{A}$ is a factor, so $R = 0$ or $\mathbf{1}$, this implies that either $Q_1 \prec P_1$ or $Q_2 \prec P_2$.)

Proof We define projections P_1', P_2', Q_1', Q_2', by:

$$P_1 = P_1 \wedge Q_1 + P_1 \wedge Q_2 + P_1',$$
$$P_2 = P_2 \wedge Q_1 + P_2 \wedge Q_2 + P_2',$$
$$Q_1 = Q_1 \wedge P_1 + Q_1 \wedge P_2 + Q_1',$$
$$Q_2 = Q_2 \wedge P_1 + Q_2 \wedge P_2 + Q_2'.$$

We will show that P_1', P_2', Q_1', and Q_2' are all equivalent. Let us first show how this implies the lemma. Assume first that $\mathscr{A}$ is a factor; then either $P_1 \wedge Q_2 \prec P_2 \wedge Q_1$ or $P_2 \wedge Q_1 \prec P_1 \wedge Q_2$. If the first relation holds, then

$$P_1 = P_1 \wedge Q_1 + P_1 \wedge Q_2 + P_1' \prec P_1 \wedge Q_1 + P_2 \wedge Q_1 + Q_1' = Q_1;$$

if the second relation holds, a similar argument shows $Q_2 \prec P_2$. If $\mathfrak{B}$ is not a factor, then $P_1 \wedge Q_2$ and $P_2 \wedge Q_1$ may not be comparable, but Theorem IV. F.1 shows that there is a central projection R such that

$$R(P_2 \wedge Q_1) \prec R(P_1 \wedge Q_2), (\mathbf{1} - R)(P_2 \wedge Q_1) \succ (\mathbf{1} - R)(P_1 \wedge Q_2),$$

and then a straightforward modification of the above argument completes the proof. It remains to prove the equivalence of P_1', P_2', Q_1', and Q_2'. Consider first the operator $P_1'Q_1'$. We will show that its null space is exactly the null space of Q_1'; since its range is clearly contained in $P_1'\mathcal{H}$, this will imply that its polar decomposition gives a partial isometry in $\mathfrak{B}$ with initial subspace $Q_1'\mathcal{H}$ and terminal subspace contained in $P_1'\mathcal{H}$, i.e., it will show that $Q_1' \prec P_1'$. Now if the null space of $P_1'Q_1'$ is not contained in $(\mathbf{1} - Q_1')\mathcal{H}$, it must contain a non-zero vector in $Q_1'\mathcal{H}$. Thus, suppose $\xi \in Q_1'\mathcal{H}$; and suppose that $P_1'Q_1'\xi = P_1'\xi = 0$. Now the projection $Q_1' \leqq Q_1$; hence, $Q_2\xi = 0$, so $(Q_2 \wedge P_1)\xi = 0$. Similarly, since Q_1' is orthogonal to $P_1 \wedge Q_1$, $(P_1 \wedge Q_1)\xi = 0$. Since $P_1 = P_1 \wedge Q_1 + P_2 \wedge Q_2 + P_1'$, $P_1\xi = 0$. Since $Q_1' \leqq Q_1 + Q_2 = P_1 + P_2$, we must have $(P_1 + P_2)\xi = \xi$, i.e., $P_2\xi = \xi$, so $\xi \in P_2\mathcal{H} \wedge Q_1\mathcal{H}$. But Q_1' is orthogonal to $P_2 \wedge Q_1$, so $\xi = 0$. This then, shows $Q_1' \prec P_1'$. Interchanging the roles of P_1 and P_2 gives $Q_1' \prec P_2'$; interchanging the roles of Q_1 and Q_2 gives $Q_2' \prec P_1'$, $Q_2' \prec P_2'$; interchanging the roles of P's and the Q's gives $P_1' \simeq P_2' \simeq Q_1' \simeq Q_2'$ as desired.

THEOREM IV. G.4 *Let $P_1, \ldots, P_n$ be a finite set of finite projections in the von Neumann algebra $\mathfrak{B}$. Then $\bigvee_{i=1}^{n} P_i$ is finite.*

Proof By induction, it suffices to consider $n = 2$. Next, we reduce to the case P_1 and P_2 orthogonal. Let $P_2' = P_1 \vee P_2 - P_1$. Then P_2P_2' is injective from $P_2'\mathcal{H}$ into $P_2\mathcal{H}$, so its partial isometric part has initial subspace $P_2'\mathcal{H}$ and terminal subspace contained in $P_2\mathcal{H}$, so $P_2 \succ P_2'$. Since P_2 is finite, P_2' is finite, and $P_1 \vee P_2' = P_1 + P_2' = P_1 \vee P_2$, so we may replace P_2 by P_2', i.e., we may assume P_2 is orthogonal to P_1. Thus, let P_1, P_2 be mutually orthogonal finite projections in $\mathfrak{B}$, and suppose $P_1 + P_2$ is not finite. Then, by Lemma IV. G.2 there exists a mutually orthogonal sequence $S_1, S_2, \ldots$ of pairwise equivalent non-zero projections all contained in $P_1 + P_2$. Let $Q_1 = S_1 + S_3 \ldots$; $Q_2 = P_1 + P_2 - Q_1 \geqq S_2 + S_4 + \cdots$ Evidently, Q_1 and Q_2 are mutually orthogonal, and $Q_1 + Q_2 = P_1 + P_2$. By Lemma IV. G.3, there is a central projection R such that

$$RQ_1 \prec RP_1 \leqq P_1$$

$$(\mathbf{1} - R)Q_2 \prec (\mathbf{1} - R)P_2 \leqq P_2.$$

Since P_1 is finite and $RQ_1 = RS_1 + RS_2 + \cdots$ is the sum of infinitely many mutually orthogonal, mutually equivalent projections, we must have

$RS_i = 0$ for all i. But then

$$P_2 \geqq (\mathbf{1} - R)P_2 \succ (\mathbf{1} - R)Q_2 \geqq (\mathbf{1} - R)S_2 + (\mathbf{1} - R)S_4 + \cdots$$
$$= S_2 + S_4 + \cdots$$

contradicting the finiteness of P_2 and completing the proof.

II Tensor Product Decompositions

We have already seen how to decompose a von Neumann algebra into smaller pieces by using central projections. In this section, we will discuss another technique for breaking up von Neumann algebras into smaller pieces. The starting point here is a family $(P_i)_{i \in I}$ of projections in $\mathcal{B}$; (the index set may be finite of infinite); we assume that these projections are mutually orthogonal, pairwise equivalent, and add up to the identity operator. Let P be some fixed projection equivalent to each P_i. We will show that the Hilbert space $\mathcal{H}$ on which $\mathcal{B}$ acts may be decomposed as a tensor product $\mathcal{H} \cong \tilde{\mathcal{H}} \otimes \mathcal{H}_1 : \mathcal{H}_1$, where $\tilde{\mathcal{H}}$ has Hilbert space dimension equal to the cardinality of the index set I, $\mathcal{H}_1 \cong P\mathcal{H}$, and $\mathcal{B}$ goes over into the von Neumann algebra generated by $\mathcal{L}(\tilde{\mathcal{H}}) \otimes \mathbf{1}$ and $\mathbf{1} \otimes P\mathcal{B}P$: Here $P\mathcal{B}P$ denotes the collection of all operators on $P\mathcal{H}$ obtained by restricting to $P\mathcal{H}$ operators $B \in \mathcal{B}$ satisfying $B = PBP$, i.e., operators in $\mathcal{B}$ which are zero on $(\mathbf{1} - P)\mathcal{H}$ and whose range is contained in $P\mathcal{H}$. It is easy to verify from this description that PBP is a weakly closed self-adjoint algebra of operators on $P\mathcal{H}$, containing the identity operator, i.e., that $P\mathcal{B}P$ is a von Neumann algebra on $P\mathcal{H}$. This decomposition is particularly interesting when P is a *minimal projection* in $\mathcal{B}$, i.e., when there is no non-zero projection of $\mathcal{B}$ contained in P except P itself. In this case, the von Neumann algebra $P\mathcal{B}P$ contains no no-trivial projections; hence, consists only of the scalars, so $\mathcal{B} \cong \mathcal{L}(\tilde{\mathcal{H}}) \otimes \mathbf{1}$, i.e., $\mathcal{B}$ is isomorphic to $\mathcal{L}(\tilde{\mathcal{H}})$. Let us expand a bit on this point before proceeding to the proof of the existence of a tensor product decomposition. Let $\mathcal{B}$ be a factor and suppose $\mathcal{B}$ contains a minimal projection P. (We say that a factor which contains a minimal projection is of *type I*.) By Zorn's Lemma, there exists a maximal collection $(P_i)_{i \in I}$ of pairwise orthogonal projections, such that each P_i is equivalent to P. We claim that $\mathbf{1} = \sum_I P_i$. To see this, let $P' = \mathbf{1} - \sum_I P_i$. Then either $P' \succ P$ or $P \succ P'$, $P' \not\simeq P$. The first alternative is ruled out by the maximality of the family $(P_i)_{i \in I}$, and the second alternative implies $P' = 0$ by the minimality of P. Combining the above remarks, we get:

PROPOSITION IV. H.1 *Let $\mathcal{B}$ be a type I factor on a Hilbert space $\mathcal{H}$. Then $\mathcal{H}$ may be decomposed as a tensor product $\mathcal{H} \cong \tilde{\mathcal{H}} \otimes \mathcal{H}_1$, in such a way that $\mathcal{B}$ corresponds to the algebra of operators of the form $A \otimes \mathbf{1}$, $A \in \mathcal{L}(\tilde{\mathcal{H}})$.*

We now proceed to the construction of the tensor product decomposition. We first state the result precisely:

PROPOSITION IV. H.2 *Let $\mathfrak{A}$ be a von Neumann algebra on a Hilbert space $\mathcal{H}$. Let P be a projection in $\mathfrak{A}$, and suppose that we can decompose $\mathbf{1} = \sum_{i \in I} P_i$, where the P_i are mutually orthogonal projections in $\mathfrak{A}$ each of which is equivalent to P. Then there exists a Hilbert space $\hat{\mathcal{H}}$, with Hilbert-space dimension equal to the cardinality of I, and a unitary operator W from $\mathcal{H}$ to $\hat{\mathcal{H}} \otimes (P\mathcal{H})$, such that $W\mathfrak{A}W^{-1}$ is the von Neumann algebra generated by $\mathcal{L}(\hat{\mathcal{H}}) \otimes \mathbf{1}$ and $\mathbf{1} \otimes P\mathfrak{A}P$. (In other words, $\mathfrak{A}$ is unitarily equivalent to the tensor product of $\mathcal{L}(\hat{\mathcal{H}})$ and $P\mathfrak{A}P$.)*

Proof We will construct the tensor product decomposition in two steps. First, we note that the decomposition $\mathbf{1} = \sum_i P_i$ gives a decomposition $\mathcal{H} = \bigoplus_i P_i\mathcal{H}$. The subspace $P_i\mathcal{H}$ are all unitarily equivalent to $P\mathcal{H}$, and the unitary operator may be taken to be in $\mathfrak{A}$. More precisely, we can choose for each i a partial isometry $U_i \in \mathfrak{A}$ with initial subspace $P_i\mathcal{H}$ and terminal subspace $P\mathcal{H}$. Using this family of partial isometries, we can identify each $P_i\mathcal{H}$ with $P\mathcal{H}$ and thus transform the decomposition $\bigoplus_i P_i\mathcal{H}$ into a realization $\mathcal{H}' \cong \bigoplus_{i \in I} P\mathcal{H}$. To be completely explicit, we can define a unitary operator $U: \mathcal{H} \to \bigoplus_{i \in I} P\mathcal{H}$ by $U\xi = (U_i\xi)$ (Note that, since U_i has initial subspace $P_i\mathcal{H}$, $\|U_i\xi\| = \|P_i\xi\|$, so $\sum_i \|U_i\xi\|^2 = \sum_i \|P_i\xi\|^2 = \|\xi\|^2$.)

The second step is to transform the direct sum representation $\mathcal{H} \cong \bigoplus_{i \in I} P\mathcal{H}$ into a tensor product representation. This is a standard construction; we let $\hat{\mathcal{H}}$ be a Hilbert space with an orthonormal basis (ϕ_i) labelled by I, and we define a unitary operator $V: \bigoplus_{i \in I} P\mathcal{H} \to \hat{\mathcal{H}} \otimes P\mathcal{H}$ by $V: (\xi_i) \mapsto \sum_i \phi_i \otimes \xi_i$. Now, of course, we define $W = V \cdot U$.

The next step is to compute what happens to various operators on $\mathcal{H}$ when they are transformed by W. Tracing through the definitions, it is easy to see that

$$WU_i^*U_jW^{-1} = \hat{U}_{ij} \otimes \mathbf{1},$$

where $\hat{U}_{ij}$ is the operator on $\hat{\mathcal{H}}$ defined by

$$\hat{U}_{ij}\phi_j = \phi_i; \qquad \hat{U}_{ij}\phi_k = 0 \quad \text{if} \quad k \neq j.$$

It is also easily verified that the only operators on $\hat{\mathcal{H}} \otimes P\mathcal{H}$ commuting with $\hat{U}_{ij} \otimes \mathbf{1}$ for all i, j are those of the form

$$\mathbf{1} \otimes A, \quad \text{with} \quad A \in \mathcal{L}(P\mathcal{H}).$$

Next, suppose $B' \in \mathscr{B}'$. Then it may be seen that

$$WB'W^{-1} = \mathbf{1} \otimes (B' \mid P\mathscr{H}).$$

Finally, if $B \in P\mathscr{B}P$, then $\mathbf{1} \otimes B = W\tilde{B}W^{-1}$, where

$$\tilde{B} = \sum_i U_i^* B U_i \in \mathscr{B}.$$

Now using the fact that the commutant of $P\mathscr{B}P$ is exactly $\mathscr{B}' \mid P\mathscr{H}$, we see that the operators on $\tilde{\mathscr{H}} \otimes P\mathscr{H}$ commuting with $\tilde{U}_{ij} \otimes \mathbf{1}$ for all i, j and with $\mathbf{1} \otimes B$ for all $B \in P\mathscr{B}P$ are precisely the operators in $W\mathscr{B}'W^{-1}$. Thus:

1) For any $A \in \mathscr{L}(\tilde{\mathscr{H}})$, $A \otimes \mathbf{1} \in (W\mathscr{B}'W^{-1})' = W\mathscr{B}W^{-1}$.

2) The bicommutant of $(\mathscr{L}(\tilde{\mathscr{H}}) \otimes \mathbf{1}) \cup (\mathbf{1} \otimes P\mathscr{B}P)$ is the commutant of $W\mathscr{B}'W^{-1}$, i.e., is $W\mathscr{B}W^{-1}$, so $W\mathscr{B}W^{-1}$ is the von Neumann algebra generated by $\mathscr{L}(\tilde{\mathscr{H}}) \otimes \mathbf{1}$ and $\mathbf{1} \otimes P\mathscr{B}P$.

I Classification of Factors

In the preceding section, we defined a factor to be of type I if it contains a minimal projection. We then showed that, if $\mathscr{B}$ is such a factor, there is a tensor product decomposition of the Hilbert space $\mathscr{H}$ as $\tilde{\mathscr{H}} \otimes \mathscr{H}_1$, such that $\mathscr{B} = \mathscr{L}(\tilde{\mathscr{H}}) \otimes \mathbf{1}$. In particular, $\mathscr{B}$ is isomorphic to $\mathscr{L}(\tilde{\mathscr{H}})$. Conversely, if a von Neumann algebra $\mathscr{B}$ is isomorphic to $\mathscr{L}(\tilde{\mathscr{H}})$ for some Hilbert space $\tilde{\mathscr{H}}$, then $\mathscr{B}$ is a factor and contains a minimal projection (any one-dimensional projection is minimal in $\mathscr{L}(\tilde{\mathscr{H}})$); hence, is a factor of type I. Therefore, a von Neumann algebra is a factor of type I if and only if it is isomorphic to the algebra of all bounded operators on some Hilbert space. If that Hilbert space is finite-dimensional and of dimension n, we say that the algebra is of *type* I_n; if the Hilbert space is infinite-dimensional, we say that the algebra is of *type* I_∞.

At the opposite extreme from factors of type I are factors of type III; a factor $\mathscr{B}$ is said to be of *type III* if every non-zero projection in it is infinite. Finally, a factor $\mathscr{B}$ is said to be of type II of it is neither of type I nor of type III, i.e., if it contains no minimal projection but does contain a non-zero finite projection. A finite factor of type II is called a factor of *type* II_1; an infinite factor of type II is called a factor of type II_∞.

The decomposition theory of the preceding section provides a convenient way of passing from factors of type II_∞ to factors of type II_1. Let $\mathscr{B}$ be a factor of type II_∞, and let P be a finite projection in $\mathscr{B}$. We will prove shortly (Proposition IV. I.1) that there exists a mutually orthogonal family (P_i) of projections in $\mathscr{B}$ equivalent to P, such that $\mathbf{1} = \sum_i P_i$. Thus, we may write

$\mathcal{H} \cong \hat{\mathcal{H}} \otimes P\mathcal{H}$, and $\mathcal{B}$ is the von Neumann algebra generated by $\mathcal{L}(\hat{\mathcal{H}}) \otimes \mathbf{1}$ and $\mathbf{1} \otimes P\mathcal{B}P$. It is nearly immediate that $P\mathcal{B}P$ is a factor (if $A \in \mathcal{Z}(P\mathcal{B}P)$, then $\mathbf{1} \otimes A \in \mathcal{Z}(\mathcal{B})$, so $\mathbf{1} \otimes A$ is a scalar, so A is a scalar) and that $P\mathcal{B}P$ is finite (If W is a partial isometry in $P\mathcal{B}P$ with initial domain $P\mathcal{H}$ and terminal domain strictly contained in $P\mathcal{H}$, then since W is the restriction to $P\mathcal{H}$ of a partial isometry in $\mathcal{B}$ which is zero on $(\mathbf{1} - P)\mathcal{H}$, P is equivalent (in $\mathcal{B}$) to a strictly smaller projection, and this violates the assumed finiteness of P). Furthermore, $P\mathcal{B}P$ certainly cannot contain minimal projections (since it can be regarded as a subalgebra of $\mathcal{B}$), so $P\mathcal{B}P$ is a factor of type II_1. Thus, a type II_∞ factor can be decomposed into a tensor product (in a sense which should be clear from the above) of a factor of type II_1, and a factor of type I_∞. Conversely, starting with a factor of type II_1, one can form a factor of type II_∞ by taking the tensor product with a factor of type I_∞.

In the above discussion, we make use of the fact that the identity in a type II_∞ factor can be written as a sum of infinitely many mutually orthogonal, pairwise equivalent projections. The following proposition generalizes this assertion:

PROPOSITION IV, I.1 *Let $\mathcal{B}$ be a factor, P and Q projections in $\mathcal{B}$ with P finite and non-zero and Q infinite. Then there exists a family $(P_i)_{i \in I}$ of mutually orthogonal projections in $\mathcal{B}$ such that $Q = \sum_i P_i$ and such that $P_i \simeq P$ for all i.*

To prove this proposition, we first prove the following lemma:

LEMMA IV. I.2 *Let $\mathcal{B}$ be a factor, Q a projection in $\mathcal{B}$, and $(S_i')_{i \in I}$ an infinite mutually orthogonal family of pairwise equivalent non-zero projections contained in Q. Then we can write $Q = \sum_{j \in J} S_j$ where each S_j is equivalent to each S_i'.*

Proof of Lemma IV. I.2 By using Zorn's Lemma, we may extend $(S_i')_{i \in I}$ to a *maximal* mutually orthogonal family $(S_j')_{j \in J}$ of pairwise equivalent non-zero projections contained in Q (i.e., we make $(S_j')_{j \in J}$ by adjoining to $(S_i')_{i \in I}$ as many projections orthogonal to each other and to each S_i', but also equivalent to each S_i', as possible.) Let $Q' = Q - \sum_{j \in J} S_j'$. If j_0 is some element of J, then, since $\mathcal{B}$ is a factor, either $Q' \prec S_{j_0}'$ or $Q' \succ S_{j_0}'$. If the second relation were true, we could adjoin to $(S_j')_{j \in J}$ a projection $S'' \leqq Q'$ which is equivalent to S_{j_0}' and thus contradict the maximality of the family $(S_j')_{j \in J}$. Thus, we must have $Q' \prec S_{j_0}'$. Now:

$$\sum_{j \in J} S_j' = S_{j_0}' + \sum_{j \neq j_0} S_j' \succ Q' + \sum_{j \neq j_0} S_j'$$

(we have replaced S_{j_0}' by the smaller projection Q')

$$\simeq Q' + \sum_j S_j'$$

(since $J \setminus \{j_0\}$ and J have the same number of elements)

$$= Q,$$

so $\sum_{j \in J} S'_j \succ Q$, but, on the other hand,

$$\sum_{j \in J} S'_j \leqq Q, \quad \text{so} \quad \sum_{j \in J} S'_j \prec Q.$$

Thus, $\sum_{j \in J} S'_j \simeq Q_1$ and this means precisely that Q can be written as a sum of a mutually orthogonal family of subprojections labelled by J each of which is equivalent to each S'_j.

Proof of Proposition IV. I.1 Let $(P'_i)_{i \in I}$ be a maximal mutually orthogonal family of subprojections of Q equivalent to P, and let $Q' = Q - \sum_i P'_i$. The maximality of $(P'_i)_{i \in I}$ implies, as in the above lemma, that $Q' \prec P$. In particular, Q' is finite. If the family I were finite, then we could write $Q = P_{i_1} + P_{i_2} + \cdots + P_{i_n} + Q'$, so Q, as a finite sum of finite projections, would be finite. Since we have assumed Q to be infinite, this cannot be the case, so $(P'_i)_{i \in I}$ is an infinite family of subprojections of Q equivalent to P, and the proposition follows from the lemma.

We can also use Lemma IV. I.2 to prove

PROPOSITION IV. I.3 *Let $\mathfrak{A}$ be a factor on a separable Hilbert space. Then any two infinite projections in $\mathfrak{A}$ are equivalent.*

Proof Let P and Q be two infinite projections in $\mathfrak{A}$. Then there are sequences $S'_n(T'_n)$ of mutually orthogonal pairwise equivalent non-zero subprojections of $P(Q)$. Since $\mathfrak{A}$ is a factor, either $S'_n \prec T'_m$ or $S'_n \succ T'_m$ for all m, n. Assume for definiteness that the first alternative holds. Then P and Q each contain an infinite mutually orthogonal family of projections equivalent to, say, S'_1. Thus, by Lemma IV, I.2, we can write

$$P = \sum_{i \in I} P_i; \qquad Q = \sum_{j \in J} Q_j,$$

where each P_i and each Q_j is equivalent to S'_1, and where the index sets I and J are both infinite. Now, on a separable Hilbert space, any infinite mutually orthogonal family of non-zero projections must be countable, so the index sets I and J can both be taken to be $\{1, 2, 3, \ldots\}$. Thus, we have $P = \sum_{n=1}^{\infty} P_n; Q = \sum_{n=1}^{\infty} Q_n$, and $P_n \simeq Q_n$ for all n, so $P \simeq Q$.

COROLLARY IV. I.4 *Let $\mathfrak{A}$ be a factor of type III on a separable Hilbert space. Then any two non-zero projections in $\mathfrak{A}$ are equivalent.*

J Dimension Theory for Projections in Type II Factors

Let $\mathscr{H}$ be a separable Hilbert space, and consider the algebra $\mathscr{L}(\mathscr{H})$. For any projection $P \in \mathscr{L}(\mathscr{H})$, we define the dimension of P, dim (P), to be equal to the dimension of the range of P if this dimension is finite and to be ∞ if the dimension is infinite. Then the dimension function has two important properties:

a) $P \simeq Q$ if and only if dim (P) = dim (Q).

b) If P is orthogonal to Q, dim $(P + Q)$ = dim (P) + dim (Q).

It is easy to see that these two properties determine the dimension function uniquely up to multiplication by an overall constant factor. Since the equivalence of projections is preserved under isomorphisms of von Neumann algebras ($P \simeq Q$ in $\mathscr{B}$ if and only if there exists $W \in \mathscr{B}$ such that $W^*W = P$; $WW^* = Q$), and since every type I factor on a separable Hilbert space is isomorphic to $\mathscr{L}(\tilde{\mathscr{H}})$, where $\tilde{\mathscr{H}}$ is separable (but may be finite dimensional), we see that every type I factor $\mathscr{B}$ admits a dimension function satisfying a) and b) which may be normalized to take on the values $\{0, 1, 2, \ldots, n\}$ if $\mathscr{B}$ is of type I_n or $\{0, 1, 2, \ldots, \infty\}$, if $\mathscr{B}$ is of type I_∞. The point of the present section is to prove that, if $\mathscr{B}$ is a type II factor on a separable Hilbert space, $\mathscr{B}$ again admits a dimension function (a positive real-valued function on the set of projections of $\mathscr{B}$ satisfying a) and b)), that this dimension function is again uniquely determined up to multiplication by an overall constant but that this time the range of values is a closed interval $[0, a]$, where a is a finite number if $\mathscr{B}$ is of type II_1 and is $+\infty$ if $\mathscr{B}$ is of type II_∞. Thus, from the present point of view, the difference between a factor of type II and a factor of type I is that the dimension function for a factor of type II takes on a continuous range of values, while the dimension function for a factor of type I takes on a discrete set of values. (A factor of type III on a separable Hilbert space also admits a dimension function, but it is not very interesting. Since any two non-zero projections in such a factor are equivalent (Corollary IV. I.4), condition a) implies that a dimension function must take on exactly one value other than zero, and condition b) forces this value to be $+\infty$.) It should be remarked that the condition that $\mathscr{H}$ be separable is not essential to the discussion; it can be eliminated by allowing the dimension function to take on values which are infinite cardinal numbers and thus to distinguish between different infinite projections.

Before constructing the dimension function on a factor of type II, we prove some results about the comparison of projections in general von Neumann algebras. These results enable us to do simple "arithmetic operations" on equivalence classes of finite projections. By this, we mean the following: Suppose we have two equivalence classes of projections $[P]$ and $[Q]$. (We will use $[P]$ to denote the set of all projections equivalent to P.) If we can find $P \in [P]$ and $Q \in [Q]$ such that P is orthogonal to Q, we

define $[P] + [Q]$ to be the equivalence class of the projection $P + Q$. This definition is unambiguous since, if $P_1 \in [P]$, $Q_1 \in [Q]$ and if P_1 is orthogonal to Q_1, then $P_1 + Q_1 \simeq P + Q$. (Note, however, that $[P] + [Q]$ may not be defined; this is the case for $[P] = [Q] = [\mathbf{1}]$ in a finite factor.) The next thing we want to show is that, if $[Q]$ is a *finite* equivalence class and $[P] \prec [Q]$, then $[Q] - [P]$ is well-defined. Thus, we want to prove:

PROPOSITION IV. J.1 *Let $\mathscr{B}$ be a von Neumann algebra, Q_1 and Q_2 equivalent finite projections in $\mathscr{B}$, P_1 and P_2 equivalent projections in $\mathscr{B}$ such that $P_1 \leqq Q_1$; $P_2 \leqq Q_2$. Then $Q_1 - P_1 \simeq Q_2 - P_2$.*

Proof Let us first assume that $\mathscr{B}$ is a factor. Then, by Theorem IV. F.1, either $Q_1 - P_1 \succ Q_2 - P_2$ or $Q_2 - P_2 \succ Q_1 - P_1$. Assume for definiteness that the first relation holds but that $Q_1 - P_1$ is not equivalent to $Q_2 - P_2$. Then $Q_2 - P_2$ is equivalent to a projection S which is strictly smaller than $Q_1 - P_1$. But then $Q_1 \simeq Q_2 = (Q_2 - P_2) + P_2 \simeq S + P_1 \lneq Q_1$ and this violates the assumed finiteness of Q_1. If $\mathscr{B}$ is not a factor, then, again by Theorem IV. F.1, there is a central projection R such that

$$R(Q_1 - P_1) \succ R(Q_2 - P_2); \quad (\mathbf{1} - R)(Q_2 - P_2) \succ (\mathbf{1} - R)(Q_1 - P_1).$$

Straightforward modification of the above argument then yields:

$$R(Q_1 - P_1) \simeq R(Q_2 - P_2); \quad (\mathbf{1} - R)(Q_2 - P_2) \simeq (\mathbf{1} - R)(Q_1 - P_1),$$

and hence $Q_1 - P_1 \simeq Q_2 - P_2$.

Secondly, we want to prove that, for a finite equivalence class $[P]$ in $\mathscr{B}$,

(½) $[P]$ is uniquely defined if it makes sense at all. Specifically, we want to prove:

PROPOSITION IV. J.2 *Let P_1, P_2, Q_1, Q_2 be finite projections in a von Neumann algebra $\mathscr{B}$. Assume that P_1 and P_2 are mutually orthogonal and equivalent; that Q_1 and Q_2 are mutually orthogonal and equivalent; and that $P_1 + P_2$ is equivalent to $Q_1 + Q_2$. Then $P_1 \simeq Q_1$.*

Proof Assume first that $\mathscr{B}$ is a factor. Then we have either $P_1 \prec Q_1$ or $Q_1 \prec P_1$. Assume for definiteness $P_1 \prec Q_1$. Then, if P_1 is not equivalent to Q_1, we have $P_1 \simeq Q_1' \lneq Q_1$. Since $P_1 \simeq P_2$ and $Q_1 \simeq Q_2$, we also have $P_2 \simeq Q_2' \lneq Q_2$. Thus, $Q_1 + Q_2 \simeq P_1 + P_2 \simeq Q_1' + Q_2' \lneq Q_1 \lneq Q_2$. But by assumption Q_1 and Q_2 are both finite, so $Q_1 + Q_2$ is finite, so $Q_1 + Q_2$ cannot be equivalent to a proper subprojection of itself; hence, $P_1 \simeq Q_1$. Now, if $\mathscr{B}$ is not a factor, there is a central projection R such that $RQ_1 \prec RP_1$; $(\mathbf{1} - R)P_1 \prec (\mathbf{1} - R)Q_1$ (Theorem IV. F.1). The above argument can easily be reworked to show that $RQ_1 \simeq RP_1$; $(\mathbf{1} - R)Q_1 \simeq (\mathbf{1} - R)P_1$, so $Q_1 \simeq P_1$.

We now restrict ourselves to factors of type II, and show that dividing a finite equivalence class by two always makes sense.

PROPOSITION IV. J.3 *Let P be a projection in a factor $\mathscr{B}$ of type II. Then we can write $P = P_1 + P_2$, where P_1 and P_2 are mutually orthogonal and equivalent.*

Proof Consider the set of all collections of pairs of non-zero subprojections of P $\{(Q_i, Q_i')\}$ in $\mathscr{B}$, where

1) $Q_i \simeq Q_i'$ for all i.

2) Q_i is orthogonal to Q_j for $i \neq j$.
Q_i' is orthogonal to Q_j' for $i \neq j$.

3) Q_i is orthogonal to Q_j' for all i, j.

By Zorn's Lemma, there is a maximal such family $\{(Q_i, Q_i')\}$. Let $P_1 = \sum_i Q_i$; $P_2 = \sum_i Q_i'$. Then P_1 and P_2 are equivalent mutually orthogonal subprojections of P. If we can show $P_1 + P_2 = P$, we are through. Suppose not. Then $P - P_1 - P_2$ is a non-zero projection in $\mathscr{B}$. Since $\mathscr{B}$ is of type II, $P - P_1 - P_2$ cannot be a minimal projection, so we can write $P - P_1 - P_2 = Q_0 + \dot{Q}_0$, where neither Q_0 nor $\dot{Q}_0$ is zero. Either $Q_0 \prec \dot{Q}_0$ or $\dot{Q}_0 \prec Q_0$, and we can assume the former is the case. Then if Q_0' is a subprojection of $\dot{Q}_0$ equivalent to Q_0, $\{(Q_i, Q_i')\}$ u, $\{Q_0, Q_0'\}$ satisfies 1), 2), 3), and hence violates the maximality of $\{(Q_i, Q_i')\}$. Thus, $P - P_1 - P_2 = 0$, so we are through.

By induction, then, if $[P]$ is a finite equivalence class of projections in a type II factor, there is a uniquely defined equivalence class $(\frac{1}{2})^n [P]$ for each n, and hence a uniquely defined equivalence class $\frac{m}{2^n} [P]$ for any integer m between 0 and 2^n. If $\mathscr{B}$ is of type II_∞, then we may write $\mathbf{1}$ as a sum of a sequence of mutually orthogonal projections P_i such that each P_i is equivalent to P (Proposition IV. I.1). Thus, there is a uniquely defined equivalence class $\frac{m}{2^n} [P]$ for any integer m. Specifically, the statement $Q \in \frac{m}{2^n} [P]$ means that we can write $Q = \sum_{i=1}^{m} Q_i$; $P = \sum_{j=1}^{2^n} P_j$ with $Q_i \simeq P_j$ for all i, j. We define separately $0[P] = [0]$ for all P. It is important to recognize that the statement $Q \in \frac{m}{2^n} [P]$ has nothing to do with any assertion relating the projection Q to the linear operator obtained by multiplying the projection P by the real number $\frac{m}{2^n}$. We can now define the dimension function on projections in $\mathscr{B}$.

a) If $\mathscr{B}$ is of type II_1, and if P is a projection in $\mathscr{B}$, we define

$$\dim (P) = \sup \left\{ \frac{m}{2^n} : \frac{m}{2^n} [\mathbf{1}] \prec [P] \right\}.$$

Then $\dim (P)$ is a real number in $[0, 1]$.

b) *If $\mathscr{M}$ is of type II_∞, we choose an arbitrary finite non-zero projection $P_0 \in \mathscr{M}$, and we define for an arbitrary projection $P \in \mathscr{M}$*

$$\dim(P) = \sup\left\{\frac{m}{2^n} : \frac{m}{2^n}[P_0] \prec [P]\right\}.$$

This time, $\dim(P)$ is a real number or $+\infty$; it is $+\infty$ if and only if P is infinite (Proof: If P is infinite, then by Proposition IV. I.1, P may be written as a sum of infinitely many mutually orthogonal projections each of which is equivalent to P_0. Thus, $[P] \succ m[P_0]$ for all integers m, so $\dim(P) = +\infty$. Conversely, if P is finite, then a maximal mutually orthogonal family of subprojections of P equivaleot to P_0 must be finite, so $P = P_{i_1} + \cdots + P_{i_m} + P'$ where $P' \succ P_0$. Thus,

$$[P] \prec (m+1)[P_0], \quad \text{so} \quad \dim(P) \leqq m + 1 < \infty).$$

The choice of P_0 in this construction amounts to a choice of a normalization for the dimension function. Another choice of P_0 would have given a dimension function differing at most by multiplication by a constant. Similarly, in the construction of a dimension function for a type II_1 factor, there is an arbitrariness in the normalization, but we have made the natural convention of defining the dimension of $\mathbf{1}$ to be one. In order to have a unified notation for the two cases, we will sometimes write P_0 for $\mathbf{1}$ in the II_1 case.

PROPOSITION IV. J.2 *Let $\mathscr{M}$ be a type II factor, P a finite projection in $\mathscr{M}$. Then $[P] \succ \frac{m}{2^n}[P_0]$ if and only if $\dim(P) \geqq \frac{m}{2^n}$. In particular, $\dim\left(\frac{m}{2^n}[P_0]\right) = \frac{m}{2^n}$.*

Proof From the definition of the dimension function, it follows that $\dim(P) \geqq \frac{m}{2^n}$ if $[P] \succ \frac{m}{2^n}[P_0]$. Conversely, suppose it is not true that $[P] \succ \frac{m}{2^n}[P_0]$. Then, if $\frac{m'}{2^{n'}} \geqq \frac{m}{2^n}$, it is not true that $[P] \succ \frac{m'}{2^{n'}}[P_0]$, so $\dim(P) \leqq \frac{m}{2^n}$. To prove the final assertion, note that $\frac{m}{2^n} \leqq \dim\left(\frac{m}{2^n}[P_0]\right)$ but that, if $\frac{m'}{2^{n'}} > \frac{m}{2^n}$, then it is not true that $\frac{m'}{2^{n'}}[P_0] \prec \frac{m}{2^n}[P_0]$ (by the finiteness of $\frac{m'}{2^{n'}}[P_0]$), so $\frac{m'}{2^{n'}} \geqq \dim\left(\frac{m}{2^n}[P_0]\right)$. Thus, we have also $\frac{m}{2^n} \geqq \dim\left(\frac{m}{2^n}[P_0]\right)$, so equality holds.

PROPOSITION IV. J.3 *Let P and Q be mutually orthogonal projections in a type II factor $\mathfrak{A}$. Then* $\dim(P + Q) = \dim(P) + \dim(Q)$.

Proof If either $\dim(P)$ or $\dim(Q)$ is infinite, then $P + Q$ is infinite and $\dim(P + Q) = \infty = \dim(P) + \dim(Q)$. Thus, we have only to consider the case where $\dim(P)$ and $\dim(Q)$ are both finite. For each integer n, there exist integers m, m' such that

$$\frac{m}{2^n}[P_0] \prec [P] \prec \frac{m+1}{2^n}[P_0]; \quad \frac{m'}{2^n}[P_0] \prec [Q] \prec \frac{m'+1}{2^n}[P_0].$$

In other words, we can write

$$P = \hat{P}_n + P_n', \quad \text{with} \quad \hat{P}_n \in \frac{m}{2^n}[P_0] \quad \text{and} \quad [P_n'] \prec \left(\frac{1}{2^n}\right)[P_0],$$

$$Q = \hat{Q}_n + Q_n', \quad \text{with} \quad \hat{Q}_n \in \frac{m'}{2^n}[P_0] \quad \text{and} \quad [Q_n'] \prec \left(\frac{1}{2^n}\right)[P_0].$$

Then, since

$$P + Q = \hat{P}_n + \hat{Q}_n + P_n' + Q_n',$$

since

$$\hat{P}_n + \hat{Q}_n \in \frac{m+m'}{2^n}[P_0]$$

and since

$$[P_n' + Q_n'] \prec (\tfrac{1}{2}^{n-1})[P_0],$$

we have:

$$\frac{m+m'}{2^n} \leqq \dim(P+Q) \leqq \frac{m+m'+2}{2^n}.$$

As $n \to \infty$, $\frac{m}{2^n} \to \dim(P)$; $\frac{m'}{2^n} \to \dim(Q)$, so we get

$$\dim(P) + \dim(Q) = \dim(P + Q).$$

PROPOSITION IV. J.4 *Let P be a projection in a type II factor $\mathfrak{A}$, and suppose* $\dim(P) = 0$. *Then* $P = 0$.

Proof By induction, we can find a sequence $P_1, P_2, \ldots$ of mutually orthogonal subprojections of P_0 with

$$P_1 \in (\tfrac{1}{2})[P_0]; \qquad P_2 \in (\tfrac{1}{2})[P_0], \ldots, P_i \in (\tfrac{1}{2}^i)[P_0] \ldots$$

Since $\dim(P) = 0$, we cannot have $[P] \succ (\frac{1}{2}^i)[P_0]$ for any i, and, since any two equivalence classes of projections in a factor are comparable, we must have $[P] \prec (\frac{1}{2}^i)[P_0]$ for each i. Thus, for each i, we can find a subprojection S_i of P_i which is equivalent to P. The finite projection P_0 therefore contains a mutually orthogonal sequence of projections all of which are equivalent to P. This is possible only if $P = 0$. (Lemma IV. G.2).

PROPOSITION IV. J.5 *Let $\mathcal{A}$ be a type II factor on a separable Hilbert space. Then two projections P and Q in $\mathcal{A}$ are equivalent if and only if* dim (P) = dim (Q).

Proof The definition of dim (P) depends only on the equivalence class of P, so two equivalent projections have the same dimension. Conversely, suppose dim (P) = dim (Q). If the dimensions are infinite, then P and Q are both infinite, and hence are equivalent (Proposition IV. I.3). (This is the only use we make of the assumption that the Hilbert space on which $\mathcal{A}$ acts is separable.) We can therefore assume that the dimensions are finite. Since $\mathcal{A}$ is a factor, we have either $P \succ Q$ or $Q \succ P$, and we can assume that the first alternative holds. Thus, there is a projection Q' equivalent to Q with $Q' \leq P$. Now dim (P) = dim (Q') + dim $(P - Q')$ by Proposition IV. J.3, so dim $(P - Q')$ = 0. Hence, by Proposition IV. J.4, $P - Q' = 0$, so P is equivalent to Q.

Combining the above results, we get:

THEOREM IV. J.6 *Let $\mathcal{A}$ be a type II factor on a separable Hilbert space. Then there exists a function $P \mapsto$ dim (P) on the set of projections in $\mathcal{A}$, with values in* $[0, \infty]$, *such that:*

a) $P \simeq Q$ *if and only if* dim (P) = dim (Q).

b) *If and Q are orthogonal,* dim $(P + Q)$ = dim (P) + dim (Q).

Moreover, the dimension function is uniquely specified up to multiplication by an overall constant factor by conditions a) and b).

Proof We have already proved the existence statement. To prove the uniqueness statement, let dim′ be another function satisfying a) and b). We first claim that P is finite if and only if dim′ (P) is finite. Suppose P is any projection; then we can write $P = P_1 + P_2$ where P_1 is equivalent to P_2 (Proposition IV. 5.3). Thus, dim′ (P) = dim′ (P_1) + dim′ (P_2) = 2 dim′ (P_1). Now if P is finite, P is not equivalent to P_1, so dim′ (P) $\neq$ 2 dim′ (P_1), so dim′ $(P_1) < \infty$, so dim′ $(P) < \infty$. Conversely, if P is infinite, then P_1 must also be infinite (otherwise P would be the sum of two finite projections). Thus, $P \simeq P_1$ (Two infinite projections in a factor on a separable Hilbert space are equivalent.) This implies dim′ (P_1) = dim′ (P) = 2 dim′ (P_1). This implies dim′ $(P_1) = \infty$, so dim′ $(P) = \infty$.

Now consider the projection P_0 which we have chosen to have dimension one. Since P_0 is finite, dim′ $(P_0) \neq \infty$, and since $P_0 \neq 0$, dim′ $(P_0) \neq 0$. Hence, by multiplying dim′ by a suitably chosen constant, we can arrange dim′ $(P_0) = 1$. Having done this, we no longer have a normalization free, so we must prove dim′ (P) = dim (P) for all finite P. We can of course regard dim′ as a function on the set of equivalence classes of projections. Repeated

application of properties a) and b) of the function dim' gives:

$$\dim\left(\frac{m}{2^n}[P_0]\right) = \frac{m}{2^n} = \dim'\left(\frac{m}{2^n}[P_0]\right).$$

Now if P is any finite projection in $\mathcal{M}$, and if n is any positive integer, we can find a non-negative integer m such that

$$\frac{m}{2^n}[P_0] \prec [P] \prec \frac{m+1}{2^n}[P_0].$$

Condition b) implies that dim' is an increasing function on equivalence classes of projections, so

$$\frac{m}{2^n} = \dim'\left(\frac{m}{2^n}[P_0]\right) \leqq \dim'(P) \leqq \dim'\left(\frac{m+1}{2^n}[P_0]\right) = \frac{m+1}{2^n}.$$

Since, as $n \to \infty$, $\frac{m}{2^n}$ approaches $\dim(P)$, we get

$$\dim'(P) = \lim_{n\to\infty} \frac{m}{2^n} = \dim(P).$$

We have not yet shown that the dimension function takes on all values in $[0, 1]$ if $\mathcal{M}$ is of type II_1 and all values in $[0, \infty]$ if $\mathcal{M}$ is of type II_∞. This fact is easily obtained from the following proposition, which is of interest in its own right since it gives a continuity property of the dimension function.

PROPOSITION IV. J.7 *Let (P_n) be a mutually orthogonal sequence of projections in a type II factor. Then*

$$\dim\left(\sum_{n=1}^{\infty} P_n\right) = \sum_{n=1}^{\infty} \dim(P_n).$$

(In other words, the dimension function is countably additive.)

Proof $\sum_{n=1}^{\infty} P_n \geqq \sum_{n=1}^{N} P_n$ for all N, so

$$\dim\left(\sum_{n=1}^{\infty} P_n\right) \geqq \dim\left(\sum_{n=1}^{N} P_n\right) = \sum_{n=1}^{N} \dim(P_n)$$

for all N, so

$$\dim\left(\sum_{n=1}^{\infty} P_n\right) \geqq \sum_{n=1}^{\infty} \dim(P_n).$$

Let us assume that

$$\sum_{n=1}^{\infty} \dim(P_n) < \dim\left(\sum_{n=1}^{\infty} P_n\right)$$

and derive a contradiction. Choose Q so that

$$\sum_{n=1}^{\infty} \dim(P_n) < \dim(Q) < \dim\left(\sum_{n=1}^{\infty} P_n\right).$$

We may take, for example, $Q \in r[P_0]$ with r an appropriately chosen dyadic rational). We now define inductively a mutually orthogonal sequence $(Q_1, Q_2, \ldots)$ of subprojections of Q such that $Q_n \simeq P_n$ for each n. This we do as follows: Since $\dim(Q) > \dim(P_1)$, we can find $Q_1 \leqq Q$ with $Q_1 \simeq P_1$. Then

$$\dim(Q) = \dim(Q_1) + \dim(Q - Q_1) = \dim(P_1) + \dim(Q - Q_1)$$

But $\dim(Q) > \dim(P_1) + \dim(P_2)$, so

$$\dim(Q - Q_1) > \dim(P_2), \quad \text{so there exists} \quad Q_2 \leqq Q - Q_1 \quad \text{with} \quad Q_2 \simeq P_2.$$

Continuing in this way we get the desired sequence (Q_n). Then, on the one hand,

$$\sum_n Q_n \leqq Q, \quad \text{so} \quad \dim\left(\sum_n Q_n\right) \leqq \dim(Q);$$

on the other hand,

$$\sum_n Q_n \simeq \sum_n P_n, \quad \text{so} \quad \dim\left(\sum_n Q_n\right) = \dim\left(\sum_n P_n\right) > \dim(Q).$$

We have therefore obtained a contradiction and hence proved the proposition.

COROLLARY IV. J.8 *Let Q_n be a decreasing sequence of finite projections in a type II factor, and let Q_∞ be the infimum of this sequence. Then* $\dim(Q_\infty) = \lim_{n \to \infty} \dim(Q_n)$.

Proof We can write $Q_1 = Q_\infty + (Q_1 - Q_2) + (Q_2 - Q_3) + \cdots$ Thus,

$$\begin{aligned}\dim(Q_1) &= \dim(Q_\infty) + \dim(Q_1 - Q_2) + \dim(Q_2 - Q_3) + \cdots \\ &= \lim_{n \to \infty} [\dim(Q_\infty) + \dim(Q_1) - \dim Q_2 + \dim(Q_2) - \cdots \dim(Q_n)] \\ &= \lim_{n \to \infty} [\dim(Q_\infty) + \dim(Q_1) - \dim(Q_n)]\end{aligned}$$

so $\dim(Q_\infty) = \lim_{n \to \infty} \dim(Q_n)$ as asserted.

Bibliographic Note

In this section I will attempt to provide some indication of when to look for more information about the subjects treated in these notes. The listing is incomplete, unsystematic, and subjective; the works mentioned are some of those I have found useful in my own study.

Chapter I There are numerous excellent textbooks on general topology and measure theory; I have referred to Kelley [1] on topology and Halmos [1] on measure theory. The summaries of these subjects in the books of

Loomis [1] and Naimark [1] are also useful in providing a unencumbered exposition of the essential points. For more complete discussions of the theory of topological vector spaces, see Kelley and Namioka [1] or Köthe [1].

Chapter II The version of Choquet theory presented here is largely based on the review article of Choquet and Mayer [1], which is comprehensive but rather condensed. A more leisurely presentation of the subject may be found in the book of Phelps [1].

Chapter III The most useful single source is the treatise of Dixmier (Dixmier [2], referred to in the text as C^*A). In particular, the first two sections of this work give in sixty pages an excellent and complete exposition of the general theory of C^* algebras. The books of Loomis [1], Naimark [1], and Rickart [1] are also useful.

Chapter IV Here again, the standard reference is the treatise of Dixmier (Dixmier [1], referred to in the text as AvN). A more traditional exposition is given in Schwartz [1], and an excellent presentation of the elementary theory of von Neumann algebras, on about the same level as the present set of notes, may be found in Chapter VII of Naimark [1]. The exposition given in these notes of the comparison theory of projections and its relation to multiplicity theory of representations is based largely on Chapter I of Mackey [1]. For a brief survey of the theory of von Neumann algebras, see Kadison [1].

Bibliography

Choquet, G. and P. A. Mayer, [1] Existence et unicité des représentations intégrales dans les convexes compacts quelconques, *Ann. Inst. Fourier* (Grenoble), 13, 139–154 (1963).

Dixmier, J., [1] *Les algebres d'operateurs dans l'espace hilbertien (Algebres de von Neumann)*, Paris, Gauthier-Villars (1957).

[2] *Les C^* algebres et leurs représentations*, Paris, Gauthier-Villars (1964).

Halmos, P. R., [1] *Measure Theory*, New York, van Nostrand (1950).

Kadison, R. V., [1] Theory of Operators, Part II. Operator Algebras, *B.A.M.S.*, 64, No. 3, part 2, 61–85 (1958).

Kelley, J. L., [1] *General Topology*, New York, van Nostrand (1955).

Kelley, J. L. and I. Namioka, [1] *Linear Topological Spaces*, New York, van Nostrand (1963).

Köthe, G., [1] *Topologische Lineare Räume I*, Berlin/Göttingen/Heidelberg, Springer (1960).

Loomis, L. H., [1] *An Introduction to Abstract Harmonic Analysis*, New York, van Nostrand (1953).

Mackey, G. W., [1] *The Theory of Group Representations*, Mimeographed notes from lectures delivered at the University of Chicago (1955).

Naimark, M. A., [1] *Normed Rings*, trans. L. F. Boron, Groningen, P. Noordhoff (1959).

Phelps, R. R., [1] *Lectures on Choquet's Theorem*, New York, van Nostrand (1966).

Rickart, C., [1] *General Theory of Banach Algebras*, New York, van Nostrand (1960).

Ruelle, D., [1] *Statistical Mechanics: Rigorous Results*, New York, W. A. Benjamin (1969).

Schwartz, J. T., [1] *W-* Algebras*, New York, Gordon and Breach (1967).

Proceedings of the International Congress of Mathematicians
Vancouver, 1974

Time Evolution of Infinite Classical Systems

Oscar E. Lanford III

I will discuss in this article some recent progress in the problem of proving existence and uniqueness of solutions to Newton's equations of motion for infinite systems of classical particles interacting by two-body forces which go to zero reasonably quickly as the particle separation goes to infinity. For technical simplicity, I will assume that the interparticle potential Φ has a Lipschitz continuous derivative and finite range, but the results I will describe have extensions which require neither finite range nor the absence of singularities in the potential.

To establish notation: We consider systems of infinitely many particles with positions (q_i) and momenta (p_i), moving in $\boldsymbol{R}^\nu$. The equations of motion are

$$\frac{dq_i}{dt} = \frac{p_i}{m}, \qquad \frac{dp_i}{dt} = F_i = \sum_{j \neq i} F(q_i - q_j) \tag{1}$$

where m is the particle mass and $F = -\operatorname{grad} \Phi$ is the interparticle force. We assume that there are infinitely many particles, but that, initially at least, they are distributed so that there are only finitely many in each bounded region of space. Because of the infinite number of particles, these equations cannot be treated by the usual elementary techniques, and it is indeed not hard to imagine that some solutions may develop "singularities" in which, for example, infinitely many particles rush into a bounded region of space. What is needed is an existence result which assures us that such singularities are at least improbable.

The result to be described assumes, in addition to the regularity mentioned above, that the interparticle potential Φ has good thermodynamic properties. More specifically, we assume that Φ is *superstable* in the sense of Ruelle [**6**]. It is then possible to single out a class of probability measures on the phase space for the infinite system—the so-called Gibbs states—which represent thermodynamic

equilibrium for the interaction in question. (See [1], [4], or [6].) What we show is that there exists a set of solutions to the equations of motion forming a set of probability one for each Gibbs state. We are, however, not able to describe in any very explicit way the set of phase points which lie on such solution curves, nor are we able to prove existence of solutions for initial phase points representing situations which are globally not in thermodynamic equilibrium. In this respect, the results described here are much weaker than previous work on one-dimensional systems [2] which proved existence and uniqueness of solutions for all initial phase points satisfying some reasonable regularity conditions.

Mathematically, the main novelty in the argument we will give is that it exploits the formal fact that Gibbs states ought to be invariant under the flow we are trying to construct. This leads to an a priori estimate which is shown to hold almost everywhere with respect to each Gibbs state. The idea is that Gibbs states are concentrated on very well-behaved phase points, and the invariance ought to imply good behavior at all times. The following result illustrates the argument:

PROPOSITION 1. *Let μ be a Gibbs state, and assume that the equations of motion can be solved almost everywhere to give a flow T^t leaving μ invariant. Then, for almost every phase point $\boldsymbol{x} = (q_i, p_i)$, there exists a constant M such that*

$$|q_i(t) - q_i| \leqq M \log_+(q_i) \quad \textit{for all } i \textit{ and } |t| \leqq 1. \tag{2}$$

(Here $\log_+(q)$ denotes $\log(|q|)$ if $|q| \geqq e$ and 1 otherwise.)

To prove this result, we define a function B on the infinite system phase space by

$$B(\boldsymbol{x}) = \sup_i \left\{ \frac{|p_i/m|}{\log_+(q_i)} \right\}. \tag{3}$$

To say that $B(\boldsymbol{x})$ is finite says that velocity fluctuations grow at most like the logarithm of the distance from the origin. A simple argument, using the Maxwellian (i.e., Gaussian) character of the momentum distribution, shows that B is integrable with respect to μ. Now define

$$\bar{B}(\boldsymbol{x}) = \int_{-1}^{1} dt\, B(T^t \boldsymbol{x}).$$

By Fubini's theorem and the assumed invariance of μ under T^t,

$$\int \bar{B}\, d\mu = \int_{-1}^{1} dt \int B \circ T^t\, d\mu = 2 \int B\, d\mu < \infty.$$

Hence, $\bar{B}$ is finite almost everywhere. We now claim that, where $\bar{B}(\boldsymbol{x})$ is finite, there exists a constant M (depending only on $\bar{B}(\boldsymbol{x})$) such that (2) holds. In fact, we have, for each i,

$$\left| \int_0^t dt_1 \frac{|dq_i(t_1)/dt|}{\log_+(q_i(t_1))} \right| \leqq \int_{-1}^{1} dt_1 \sup_i \left\{ \frac{|p_i(t_1)/m|}{\log_+(q_i(t_1))} \right\} = \bar{B}(\boldsymbol{x}).$$

It is now a matter of elementary calculus to show that, for any number b, there exists an $M(b)$ such that

$$\int_0^t dt_1 \frac{|dq_i(t_1)/dt|}{\log_+(q_i(t_1))} \leqq b \quad \text{implies} \quad |q_i(t) - q_i| \leqq M(b) \log_+(q_i).$$

This proves the proposition. Of course, the proposition is of little direct use, since it assumes what was to be proved, the existence of solutions to the equations. Its usefulness derives from the fact that we can find approximate solutions to the equations of motion which leave μ invariant and to which we can apply the above argument. One way to do this is as follows: For each positive integer s, let Λ_s denote the ball of radius s centered about the origin, and let $T^t_{(s)}$ denote the solution flow for the following dynamics:

(a) particles initially outside Λ_s are frozen where they are (i.e., both positions *and momenta* remain fixed);

(b) particles initially inside Λ_s move under their mutual interaction, with constant external forces exerted by the particles outside and with elastic reflection at the boundary of Λ_s.

The definition of Gibbs state readily implies that every Gibbs state is invariant under $T^t_{(s)}$ for all s. We are going to construct solutions to the equations of motion as limits, as $s \to \infty$, of $T^t_{(s)}$.

To do this, we introduce functions

$$\bar{B}_{(s)}(\boldsymbol{x}) = \frac{1}{\pi} \int_{-\infty}^{\infty} dt \, \frac{B(T^t_{(s)}\boldsymbol{x})}{1 + t^2}$$

on the infinite system phase space. As before

$$\int \bar{B}_{(s)} \, d\mu = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{dt}{1 + t^2} \int d\mu(\boldsymbol{x}) \, B(T^t_{(s)}\boldsymbol{x}) = \int B \, d\mu \tag{4}$$

for all s, so $\bar{B}_{(s)} < \infty$ almost everywhere. Also, if $\bar{B}_{(s)}(\boldsymbol{x}) \leqq b < \infty$, then, for any i and $|t| \leqq \tau$,

$$\left| \int_0^t dt_1 \frac{dq_i^{(s)}(t_1)/dt}{\log_+(q_i^{(s)}(t_1))} \right| \leqq \left| \int_0^t dt_1 \, B(T^{t_1}_{(s)}\boldsymbol{x}) \right|$$

$$\leqq (1 + \tau^2)\pi \frac{1}{\pi} \int_{-\tau}^{\tau} dt_1 \frac{B(T^{t_1}_{(s)}\boldsymbol{x})}{1 + t_1^2} \leqq (1 + \tau^2) \cdot \pi \cdot b,$$

and hence, in the notation of Proposition 1,

$$|q_i^{(s)}(t) - q_i| \leqq M((1 + \tau^2) \cdot \pi \cdot b) \log_+(q_i)$$

for all i and all t with $|t| \leqq \tau$. This inequality is a kind of localization condition which says that particles stay relatively near their initial positions. We would like to have a bound like this which is uniform in s, for almost all $\boldsymbol{x}$. To get this bound, let

$$\bar{B}_\infty(\boldsymbol{x}) = \liminf_{s \to \infty} \bar{B}_{(s)}(\boldsymbol{x}).$$

By Fatou's lemma, $\bar{B}_\infty$ is integrable and hence is finite almost everywhere. We will show that, if $\bar{B}_\infty(\boldsymbol{x})$ is finite, then there is a solution of the equations of motion with initial data $\boldsymbol{x}$. If $\bar{B}_\infty(\boldsymbol{x})$ is finite, there is a real number b and a sequence (s_n) increasing to infinity such that $\bar{B}_{(s_n)}(\boldsymbol{x}) \leqq b$ for all n. By the argument just given, this

means that, for any positive τ,

$$|q_i^{(s_n)}(t) - q_i| \leqq M((1 + \tau^2)\cdot\pi\cdot b) \log_+(q_i) \tag{5}$$

for all i and all t with $|t| \leqq \tau$.

It is now easy to finish the proof. The bound (5) together with the finite range of the interaction places a bound on the number of particles which can interact with the ith particle for times between $-\tau$ and τ. Since the potential has no singularities, the force which can be exerted by any single particle is bounded so this gives a bound on $|dp_i^{(s_n)}(t)/dt|$ for any fixed i which is uniform in t between $-\tau$ and τ and uniform in n provided s_n is large enough so the inequality (5) prevents the ith particle from colliding with the wall between time $-\tau$ and τ. The Arzelà-Ascoli theorem implies that a subsequence of the $p_i^{(s_n)}(t)$ converges uniformly for t between $-\tau$ and τ. But i and τ are arbitrary, so a diagonal procedure gives a subsequence along which each $p_i^{(s_n)}(t)$ converges uniformly on every bounded interval of times. Let us denote the limits by $(p_i(t))$. If we define

$$q_i(t) = q_i + \int_0^t dt_1\, p_i(t_1)/m \tag{6}$$

then a straightforward passage to the limit in the corresponding equation for the $p_i^{(s_n)}(t)$ gives

$$p_i(t) = p_i + \int_0^t dt_1 \sum_{j\neq i} F(q_i(t_1) - q_j(t_1)) \tag{7}$$

and equations (6) and (7) are simply the integral form of Newton's equations of motion. We have thus proved

THEOREM 2. *Let Φ be a finite-range superstable potential with Lipschitz continuous derivative, and let μ be a Gibbs state for Φ. Let $\bar{B}_\infty$ be defined as above. Then*

(i) $\int \bar{B}_\infty\, d\mu < \infty$, *and in particular $\bar{B}_\infty$ is finite almost everywhere.*

(ii) *If $\bar{B}_\infty(\boldsymbol{x})$ is finite, there exists a solution $\boldsymbol{x}(t) = (q_i(t), p_i(t))$ of the equations of motion with $\boldsymbol{x}(0) = \boldsymbol{x}$ which satisfies the localization condition*

$$\sup_{|t|\leqq\tau} \sup_i \frac{|q_i(t) - q_i|}{\log_+(q_i)} < \infty \tag{8}$$

for all finite τ.

An existence theorem like this one, without a corresponding uniqueness result, is of very little use. Furthermore, examples can be found in which solutions are nonunique, at least for systems of infinitely many hard spheres. Fortunately, the localization condition (8), together with some mild restrictions on the initial phase point $\boldsymbol{x}$ (which hold almost everywhere with respect to each Gibbs state), suffices to determine the solution uniquely. The proof of uniqueness is straightforward: The equations are rewritten as integral equations:

$$q_i(t) = q_i + t(p_i/m) + \int_0^t dt_1 \int_0^{t_1} dt_2\, F_i(t_2),$$

$$F_i(t) = \sum_{j \neq i} F(q_i(t) - q_j(t));$$

it is assumed that these equations have two solutions satisfying the same initial condition; the equations are subtracted and the localization condition (8) is inserted to obtain an integral inequality which is iterated to show that the two solutions must have been identical.

Once uniqueness has been established, the solution mappings give a flow T^t on the infinite system phase which is defined almost everywhere with respect to each Gibbs state. It may further be shown that T^t leaves each Gibbs state invariant, and that $T^t_{(s)}$ converges in measure to T^t as s approaches infinity.

The results described above are discussed in more detail in [3]; a slightly different proof is given in [5]. A manuscript giving the extension to long-range and singular interactions is in preparation. It should be mentioned that there is another approach to the problem of infinite system dynamics, due to Sinaĭ, in which it is shown that almost all initial phase points admit solutions in which, over any bounded interval of time, the particles break up into finite noninteracting clusters. This is proved for arbitrary densities in one dimension [7] and for small densities in more than one dimension [8]; it is surely not true at high densities in more than one dimension.

References

1. R. L. Dobrušin, *Gibbsian random fields. General case*, Funkcional. Anal. i Priložen. **3** (1969), no. 1, 27–35. (Russian) MR **39** #1151.

2. O. E. Lanford III, *The classical mechanics of one-dimensional systems of infinitely many particles.* I. *An existence theorem*; II. *Kinetic theory*, Comm. Math. Phys. **9** (1968), 176–191; ibid. **11** (1968/69), 257–292. MR **38** #2409; **48** #3856.

3. ———, *Lectures on the time-evolution of large classical systems*, Proc. 1974 Battelle Rencontres on Dynamical Systems, Lecture Notes in Physics, Springer-Verlag, Berlin and New York (to appear).

4. O. E. Lanford III and D. Ruelle, *Observables at infinity and states with short range correlations in statistical mechanics*, Comm. Math. Phys. **13** (1969), 194–215. MR **41** #1343.

5. C. Marchioro, A. Pellegrinotti and E. Presutti, *Existence of time evolution for ν-dimensional statistical mechanics*, Comm. Math. Phys. (to appear).

6. D. Ruelle, *Superstable interactions in classical statistical mechanics*, Comm. Math. Phys. **18** (1970), 127–159. MR **42** #1468.

7. Ja. G. Sinaĭ, *Construction of the dynamics in one-dimensional systems of statistical mechanics*, Teoret. Mat. Fiz. **11** (1972), 248–258 = J. Theoret. Math. Phys. **11** (1972).

8. ———, *The construction of cluster dynamics for dynamical systems of statistical mechanics*, Vestnik Moskov. Univ. Ser. I Mat. Meh. **29** (1974), 152–158.

UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720, U.S.A.

Ergodic properties of simple model system with collisions*

Sheldon Goldstein†

Belfer Graduate School of Science, Yeshiva University, New York, New York

Oscar E. Lanford III‡

Department of Mathematics, University of California, Berkeley, California

Joel L. Lebowitz

Belfer Graduate School of Science, Yeshiva University, New York, New York

(Received 14 March 1973)

We investigate the ergodic properties of the discrete time evolution of a particle in a two-dimensional torus with velocity in the unit square. The dynamics consists of free motion for a unit time interval followed by a baker's transformation of the velocity.

1. INTRODUCTION

We are interested in the ergodic properties of dilute gas systems. These may be thought of as Hamiltonian dynamical systems in which the particles move freely except during binary "collisions". In a collision the velocities of the colliding particles undergo a transformation with "good" mixing properties (cf. Sinai's study of the billiard problem[1]). To gain an understanding of such systems we have studied the following simple discrete time model: The system consists of a single particle with coordinate $\underline{r} = (x, y)$ in a two-dimensional torus with sides of length (L_x, L_y), and "velocity" $\underline{v} = (v_x, v_y)$, in the unit square $v_x \in [0, 1], v_y \in [0, 1]$. The phase space Γ is thus a direct product of the torus and the unit square. The transformation T which takes the system from a dynamical state $(\underline{r}, \underline{v})$ at "time" j to a new dynamical state $T(\underline{r}, \underline{v})$ at time $j + 1$ may be pictured as resulting from the particle moving freely during the unit time interval between j and $j+1$ and then undergoing a "collision" in which its velocity changes according to the baker's transformation, i.e.,

$$T(\underline{r}, \underline{v}) = (\underline{r} + \underline{v}, B\underline{v}), \tag{1.1}$$

with

$$B(v_x, v_y) = \begin{cases} (2v_x, \tfrac{1}{2}v_y), & 0 \le v_x < \tfrac{1}{2} \\ (2v_x - 1, \tfrac{1}{2}v_y + \tfrac{1}{2}), & \tfrac{1}{2} < v_x \le 1. \end{cases} \tag{1.2}$$

The normalized Lebesgue measure $d\mu = dxdydv_xdv_y/L_xL_y = d\underline{r}\,d\underline{v}/L_xL_y$ in Γ is left invariant by T. We call U_T the unitary transformation induced by T on $L^2(d\mu)$, $U_T\varphi = \varphi \circ T$. Our interest lies then in the ergodic properties of T and in the spectrum of U_T.

We note first that the transformation B on the velocities is, when taken by itself as a transformation of the unit square with measure $d\underline{v}$, well known to be isomorphic to a Bernoulli shift. It therefore has very good mixing properties. The isomorphism is obtained by setting

$$v_x = \sum_{j=1}^{\infty} 2^{-j}u_j, \qquad v_y = \sum_{j=1}^{\infty} 2^{-j}u_{1-j}, \tag{1.3}$$

with the u_j independent random variables taking the values 0 and 1 each with probability $\frac{1}{2}$. We then have

$$\begin{aligned} (B\underline{v})_x &= \sum_{j=1}^{\infty} 2^{-j}u_{j+1} = 2v_x - u_1, \\ (B\underline{v})_y &= \sum_{j=1}^{\infty} 2^{-j}u_{2-j} = \tfrac{1}{2}v_y + \tfrac{1}{2}u_1. \end{aligned} \tag{1.4}$$

2. ERGODIC PROPERTIES

The ergodic properties of our system which combines B with free motion turn out to depend on whether L_x^{-1} and L_y^{-1} satisfy the independence condition (I),

$$n_xL_x^{-1} + n_yL_y^{-1} \notin Z \text{ for } n_x \text{ and } n_y \text{ integers unless } n_x = n_y = 0. \tag{I}$$

Theorem 1: When (I) holds, the spectrum of U_T, on the complement of the one-dimensional subspace generated by the constants, is absolutely continuous with respect to Lebesgue measure and has infinite multiplicity.

It follows from Theorem 1 that when (I) holds the dynamical system (Γ, T, μ) is at least mixing. We do not know at present whether it is also a Bernoulli shift or at least a K system.

Theorem 2: When (I) does not hold the system (Γ, T, μ) is not ergodic.

The proof of Theorem 1 has two parts: a general characterization of unitary operators with Lebesgue spectrum and a set of estimates.

Lemma: Let U by a unitary operatory on a Hilbert space h, with spectral representation $U = \int_0^{2\pi} e^{i\theta}\underline{P}(d\theta)$. Assume that there exists a total set of vectors $\{\varphi_i\}$ such that $\sum_{n=1}^{\infty} |(U^n\varphi_i | \varphi_i)| < \infty$ for all i. (A set of vectors is said to be *total* if the finite linear span of this set of vectors is dense.) Then the spectral measure $\underline{P}(d\theta)$ is absolutely continuous with respect to Lebesgue measure, i.e., if E is a Borel set of Lebesgue measure 0, then $\underline{P}(E) = 0$.

Proof: We have

$(U^n\varphi_i | \varphi_i) = \int e^{in\theta}(\underline{P}(d\theta)\varphi_i | \varphi_i)$, i.e., the function $n \to (U^n\varphi_i | \varphi_i)$ is the Fourier transform of the measure $(\underline{P}(d\theta)\varphi_i | \varphi_i)$. On the other hand, $\sum_n |(U^n\varphi_i | \varphi_i)| < \infty$, so we can compute its inverse Fourier transform in the elementary way. By the uniqueness of the Fourier transform, we get:

$$(\underline{P}(d\theta)\varphi_i | \varphi_i) = \frac{d\theta}{2\pi} \cdot \sum_{n=-\infty}^{\infty} e^{-in\theta}(U^n\varphi_i | \varphi_i),$$

so the numerical measure $(\underline{P}(d\theta)\varphi_i | \varphi_i)$ is absolutely continuous with respect to Lebesgue measure. If E is a Borel setof Lebesgue measure 0,

$$\|\underline{P}(E)\varphi_i\|^2 = (\underline{P}(E)\varphi_i | \varphi_i) = 0, \quad \text{so } \underline{P}(E)\varphi_i = 0 \text{ for all } \varphi_i.$$

But the vectors $\{\varphi_i\}$ form a total set, so $\underline{P}(E) = 0$ as desired.

Now the estimates: Let $\chi(1) = 1$, $\chi(0) = -1$. For each finite subset X of Z, we define

$$\chi_X(\underline{v}) = \prod_{j\in X} \chi(u_j).$$

The χ_X form an orthonormal basis for $L^2(d\underline{v})$. Similarly, the functions $\exp(ik\cdot r)$; $\{k = (k_x, k_y), k_x = 2\pi n_x L_x, k_y = 2\pi n_y L_y, n_x$ and n_y integers$\}$, form an orthonormal basis for $L^2(d\underline{r})$. Thus, the functions $\varphi_{X,\underline{k}} = \exp(i\underline{k}\cdot\underline{r})\cdot\chi_X(\underline{v})$ form an orthonormal basis for $L^2(d\mu)$. We will prove that

$$\sum_{n=1}^{\infty} |(U_T^n \varphi_{X_1,\underline{k}_1} | \varphi_{X_2,\underline{k}_2})| < \infty \quad \text{unless } \underline{k}_1 = \underline{k}_2 = 0, \quad X_1 = X_2 = 0.$$

By straightforward computation,

$$U_T^n \varphi_{X_1,\underline{k}_1} = \varphi_{X_1+n}(\underline{v}) \exp(i\underline{k}\cdot\underline{r}) \times \exp[i\underline{k}\cdot(\underline{v} + B\underline{v} + \cdots + B^{n-1}\underline{v})].$$

Thus

$$\int d\underline{r}(U_T^n \varphi_{X_1,\underline{k}_1})\bar{\varphi}_{X_2,\underline{k}_2} = 0 \quad \text{unless } \underline{k}_1 = \underline{k}_2 (= \underline{k}),$$

so we assume $\underline{k}_1 = \underline{k}_2 = \underline{k}$. Also,

$$\int d\underline{v}(U_T^n \varphi_{X_1,0})\bar{\varphi}_{X_2,0} = 0 \quad \text{unless } X_2 = X_1 + n,$$

so the result is trivially true for $\underline{k} = 0$. We therefore assume $\underline{k} \neq 0$.

Now

$$(L_x L_y)^{-1}\int d\underline{r}d\underline{v}(U_T^n \varphi_{X_1,\underline{k}})\bar{\varphi}_{X_2,\underline{k}} = \int d\underline{v}\chi_{X_1}(B^n\underline{v})\chi_{X_2}(\underline{v}) \exp[i\underline{k}(\underline{v} + B\underline{v} + \cdots + B^{n-1}\underline{v})],$$

$$(B^j\underline{v})_x = \sum_{i=1}^{\infty} u_{j+i} 2^{-i},$$

$$\sum_{j=0}^{n-1} (B^j\underline{v})_x = \sum_{j=0}^{n-1} \sum_{i=1}^{\infty} u_{j+i} 2^{-i} = \sum_{l=1}^{\infty} u_l \sum_{i=1\vee(l-n+1)}^{l} 2^{-i} = \sum_{l=1}^{\infty} u_l \alpha_l^n$$

(where this equation defines α_l^n),

$$(B^j\underline{v})_y = \sum_{i=1}^{\infty} 2^{-i} u_{j+1-i},$$

$$\sum_{j=0}^{n-1} (B^j\underline{v})_y = \sum_{j=0}^{n-1} \sum_{i=1}^{\infty} 2^{-i} u_{j+1-i} = \sum_{l=-\infty}^{n-1} u_l \sum_{i=1\vee(-l+1)}^{n-l} 2^{-i} = \sum_{l=-\infty}^{\infty} u_l \beta_l^n.$$

Now let $l_2 = 1\vee \max\{X_2\}$, $l_1 = \inf\{X_1\} \wedge 0$.

Then

$$U_T^n \varphi_{X_1,\underline{k}}\cdot\varphi_{X_2,\underline{k}} = \prod_{l=l_2+1}^{n+l_1-1} \exp[i(\alpha_l^n k_x + \beta_l^n k_y)u_l] \times [fn \text{ of the } u_l\text{'s for } l \notin (l_2, n+l_1)].$$

By independence, the integral of the product on the right is the product of the integrals, and the unspecified function of the u_l's, $l \notin (l_2, n+l_1)$ is no greater than one in absolute value, so

$$(L_x L_y)^{-1} |\int d\underline{v}d\underline{r}\, U_T^n \varphi_{X_1,\underline{k}}\cdot\varphi_{X_2,\underline{k}}| \leq \prod_{l=l_2+1}^{n+l_1-1} |\tfrac{1}{2}[\exp(i\alpha_l^n k_x + \beta_l^n k_y) + 1]|.$$

For l's within the limits of the product, we have

$$\alpha_l^n = \sum_{i=1}^{l} 2^{-i} = 1 - 2^{-l},$$

$$\beta_l^n = \sum_{i=1}^{n-1} 2^{-i} = 1 - 2^{-(n-l)}.$$

Thus, for most of the terms in the product, $\alpha_l^n \approx \beta_l^n \approx 1$, and the number of terms is n — const for large n. In particular, if we put

$$\gamma = \tfrac{1}{2}|\exp[i(k_x + k_y)] + 1| < 1 \quad \text{(by our fundamental assumption)},$$

$$|(U_T^n \varphi_{X_1,\underline{k}} | \varphi_{X_2,\underline{k}})| < \gamma^{n/2} \quad \text{for all sufficiently large } n,$$

we have

$$\sum_{n=1}^{\infty} |(U_T^n \varphi_{X_1,\underline{k}} | \varphi_{X_2,\underline{k}})| < \infty$$

as desired.

The fact that the multiplicity is infinite is trivial. We have $L^2(d\underline{v}) \subset L^2(d\underline{r}d\underline{v})$, and we already know that the spectrum of U_T restricted to $L^2(d\underline{v})$ has infinite multiplicity.

To obtain a proof of Theorem 2, we note that ergodicity is equivalent to

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \int d\mu\, (U_T^n \varphi)\bar{\Psi} = (\int d\mu\, \varphi)(\int d\mu\, \bar{\Psi}), \qquad \varphi, \Psi \in L^2(d\mu).$$

For φ or Ψ orthogonal to the constants we must then have Cesaro convergence to zero when the system is ergodic. We prove that the system is nonergodic by finding φ or Ψ orthogonal to the constants such that the above integral converges (strictly) to a nonzero number.

Let n_x, n_y be such that $n_x/L_x + n_y/L_y \in Z$ and n_x and n_y are not both 0, and let $k_x = 2\pi n_x/L_x, k_y = 2\pi n_y/L_y$. We set $\varphi = \Psi = \varphi_{0,\underline{k}}$ and compute as before the relevant integrals:

$$I_n = \int d\mu\, (U_T^n \varphi_{0,\underline{k}})\bar{\varphi}_{0,\underline{k}} = \int d\underline{v} \exp\left[i\underline{k}\cdot\left(\sum_{j=0}^{n-1} B^j\underline{v}\right)\right] = \int d\underline{v} \prod_{l=\infty}^{\infty} \exp[i(k_x\alpha_l^n + k_y\beta_l^n)u_l] = \prod_{l=-\infty}^{\infty} \tfrac{1}{2}[1 + \exp(i\alpha_l^n k_x + \beta_l^n k_y)].$$

Here

$$\alpha_l^n = \sum_{i=1\vee(l-n+1)}^{l} 2^{-i} = 2^{-l} \sum_{m=0}^{(n-1)\wedge(l-1)} 2^m = 2^{-l}(2^{n\wedge l} - 1)$$

for $l > 0$ and vanishes for $l \leq 0$, and

$$\beta_l^n = \sum_{i=1\vee(-l+1)}^{n-l} 2^{-i} = 2^{l-1} \sum_{m=0\vee l}^{n-1} 2^{-m} = 2^{0\wedge l} - 2^{l-n}$$

for $l < n$ and vanishes for $l \geq n$.

We thus have found that

$$I_n = \prod_{l=-\infty}^{0} \tfrac{1}{2}\{1 + \exp[i(2^l - 2^{l-n})k_y]\}$$
$$\times \prod_{l=1}^{n-1} \tfrac{1}{2}(1 + \exp\{i[(1 - 2^{-l})k_x + (1 - 2^{-(n-l)})k_y]\})$$
$$\times \prod_{l-n}^{\infty} \tfrac{1}{2}\{1 + \exp[ik_x(2^{-(l-n)} - 2^{-l})]\}$$
$$= F_n^1(\underline{k})F_n^2(\underline{k})F_n^3(\underline{k})$$

with

$$F_n^1(\underline{k}) = F_n^1(k_y) = \prod_{m=0}^{\infty} \tfrac{1}{2}\{1 + \exp[ik_y(2^{-m} - 2^{-(m+n)})]\},$$

$$F_n^3(\underline{k}) = F_n^3(k_x) = F_n^1(k_x),$$

$$F_n^2(\underline{k}) = \prod_{l=1}^{n-1} \tfrac{1}{2}(1 + \exp\{i[(1 - 2^{-l})k_x + (1 - 2^{-(n-l)})k_y]\}).$$

Since $k_x + k_y \in 2\pi Z$, we have

$$F_n^2(\underline{k}) = \prod_{l=1}^{n-1} \tfrac{1}{2}\{1 + \exp[-i(k_x 2^{-l} + k_y 2^{-(n-l)})]\}.$$

We now assert that (for $k_x + k_y \in \pi Z$)

$$\lim_{n\to\infty} F_n^i(\underline{k}) = \alpha^i \neq 0, \qquad i = 1, 2, 3.$$

This is verified by observing that the log $F_n^i(\underline{k})$ coverage to a finite limit, thus completing the proof.

(If k_x and k_y are such that some of the terms at the beginning of the series which one obtains from the log $F_n^i(\underline{k})$ are singular, one easily removes the difficulty by an appropriate change in the functions φ and Ψ introduced at the beginning of the proof of Theorem 2. We also note that for the case where L_x/L_y is rational we can find explicitly a nonconstant function f which is left invariant by U_T. From the fact that $U_B(v_x + 2v_y) = 2v_x + v_y$ it follows that $f(x - y - v_x - 2v_y)$ is invariant if f is doubly periodic with periods L_x and L_y, so that we can construct an infinite family of orthonormal invariant functions $f_n : f_n = \exp\{(i2\pi n/L)(x - y - v_x - 2v_y)\}$ with $L_x/r = L_y/s = L$, r and s integers.)

---

*Supported in part by USAFOSR Grant No. 73-2430.

†National Science Foundation Fellow.

‡Alfred P. Sloan Foundation Fellow; also supported in part by NSF Grant GP-15735.

[1] Ya. Sinai, Russ. Math. Surv. 25, 137 (1970).

[2] Giovanni Gallavotti, *Modern Theory of the Billiard. An Introduction* (to appear).

Journal of Statistical Physics, Vol. 16, No. 6, 1977

Time Evolution of Infinite Anharmonic Systems

Oscar E. Lanford III,[1] Joel L. Lebowitz,[2,3] and Elliott H. Lieb[4]

Received December 28, 1976

We prove the existence of a time evolution for infinite anharmonic crystals for a large class of initial configurations. When there are strong forces tying particles to their equilibrium positions then the class of permissible initial conditions can be specified explicitly; otherwise it can only be shown to have full measure with respect to the appropriate Gibbs state. Uniqueness of the time evolution is also proven under suitable assumptions on the solutions of the equations of motion.

KEY WORDS: Existence of time evolution; uniqueness; strong restoring forces.

1. INTRODUCTION

The time evolution of classical (or quantum) Hamiltonian dynamical systems containing an infinite number of particles is of great interest in statistical mechanics, being an essential ingredient in the study of nonequilibrium phenomena in macroscopic systems. There are many difficulties, however, in dealing with the dynamics of infinite systems and the available results on the existence of the time evolution of such systems are not entirely satisfactory. It is only for one-dimensional classical systems[(1)] or harmonic crystals[(2)] that we have a strong evolution theorem, i.e., we can specify explicitly a class of initial conditions for which a time evolution exists. This set of initial conditions is furthermore appropriate for a large class of interactions between the particles and has full equilibrium measure for all these interactions. In contrast, all that has been proven so far for higher dimensions[(3–5)] is the

Supported in part by NSF Grants #MCS 75-05576-A01 (to O.E.L.), #MPS 75-20638 (to J.L.L.), and #MCS 75-21684 (to E.H.L.).

[1] University of California at Berkeley, Berkeley, California.

[2] Institut des Hautes Etudes Scientifiques, Bures-sur-Yvette, France.

[3] Permanent address: Belfer Graduate School of Science, Yeshiva University, New York, New York.

[4] Princeton University, Princeton, New Jersey.

existence of a time evolution for a given interaction on some unspecified set of initial configurations which has full measure with respect to the equilibrium state for that interaction. It is the purpose of this paper to prove the existence of a strong time evolution for a certain class (where condition A_3 of Section 2 holds) of anharmonic crystals[6] in arbitrary dimensions and a weaker time evolution for very general anharmonic systems (Section 4).

2. EXISTENCE OF TIME EVOLUTION

The setting is the lattice $\mathbb{Z}^\nu$. At each point $i \in \mathbb{Z}^\nu$ we have an oscillator with coordinate $q_i \in \mathbb{R}$ and momentum $p_i \in \mathbb{R}$. Really, we should take q_i and p_i in $\mathbb{R}^k$ for some k; with $k = \nu$ this would represent, physically, the fact that each point of $\mathbb{Z}^\nu$ is the equilibrium position of a particle. To avoid complicating the notation, we take $k = 1$, but our results obviously go through for general k. By $\mathbf{q}$ (resp. $\mathbf{p}$) we denote the collection of oscillator coordinates (resp. momenta).

The oscillator variables are regarded as functions of time t, $\{q_i(t), p_i(t)\}$, and are represented collectively by $\mathbf{q}(t)$ and $\mathbf{p}(t)$. They satisfy the following infinite set of coupled differential equations:

$$dq_i(t)/dt = p_i(t) \tag{1a}$$

$$dp_i(t)/dt = F_i = -\partial U_i(q_i(t))/\partial q_i + R_i(\mathbf{q}(t)) \tag{1b}$$

In Eq. (1b) we wrote F_i, the force acting on the ith particle, as a sum of two terms: a gradient of a "self-energy" term $U_i(q_i)$ and a force R_i, which we shall take (but need not have) to be the gradient of some interaction energy

$$R_i = -\sum_j \partial V_j(\mathbf{q})/\partial q_i \tag{1c}$$

Our basic assumption in this part is that the self-energy $U_i(q_i)$ is such a steeply increasing function of q_i that it "dominates" the motion of the particles when they are far from their equilibrium positions. We also assume that the interactions have a finite range D (this is convenient but not essential). Stated precisely, we assume:

A_1. $V_i(q)$ depends only on those q_j for which the Euclidean distance $|i - j| \leqslant D$.

A_2. Each $U_j(q_j)$ and $V_j(\mathbf{q})$ is a twice continuously differentiable function of its arguments.

A_3. $|q_j| \leqslant C_1 U_j(q_j) + C_2$, C_1 and C_2 nonnegative constants.

A_4. There exist nonnegative bounded constants $A_{ij} < C$, $A_{ij} = 0$ for $|i - j| > D$, such that

$$|p_i R_i(\mathbf{q})| \leqslant \sum_j A_{ij} \mathscr{L}_j \tag{2a}$$

where

$$\mathscr{L}_i(p_i, q_i) = \tfrac{1}{2}p_i^2 + U_i(q_i) + K \geqslant 0, \qquad K \text{ a constant} \tag{2b}$$

Example. Conditions A will be satisfied if $U_i(q_i)$ is a polynomial of degree $2n$ whose leading coefficient λ_i is strictly positive, $\lambda_i \geqslant \lambda > 0$, and $R_i(q)$ is a multinomial of degree at most n.

The nonnegative functions $\mathscr{L}_i$ will play an important role in establishing the existence and uniqueness of solutions to the equations of motion (1). They are similar to self-energy or Lyapunov functions.

The problem posed by Eqs. (1a)–(1c) is the following: Given suitable initial data $\mathbf{q}(0)$, $\mathbf{p}(0)$, find $\mathbf{q}(t)$, $\mathbf{p}(t)$ that agree with the initial data at $t = 0$ and satisfy (1a)–(1c). This problem is equivalent to another one: Find $\mathbf{q}(t)$ such that

$$q_i(t) = q_i(0) + p_i(0)t + \int_0^t (t - s)F_i(\mathbf{q}(s))\, ds \tag{3}$$

Any solution to (3) will satisfy the initial condition and will be differentiable. One can then define $p_i(t) = dq_i(t)/dt$ and (1a)–(1c) will be satisfied. Conversely, any solution to (1a)–(1c) satisfies (3).

Definition. We denote by B_r the real Banach space of sequences $\xi = \{\xi_j\}$, $j \in \mathbb{Z}^\nu$, such that the norm

$$\|\xi\|_r = \sup_{j \in \mathbb{Z}^\nu} \{[\exp(-|j|r)]|\xi_j|\} \tag{4}$$

is finite.

Lemma 1. Let $\mathbf{q}(t)$, $\mathbf{p}(t)$ be solutions of (1a)–(1c) defined for $0 \leqslant t \leqslant T$, with initial data $\mathbf{q}(0)$ such that $\mathscr{L}(0) = \{\mathscr{L}_i(0)\} \in B_r$, where we have written $\mathscr{L}_i(t)$ for $\mathscr{L}_i(p_i(t), q_i(t))$. Then there is a constant a, independent of the initial condition but depending on r, such that

$$\|\mathscr{L}(t)\|_r \leqslant [\exp(at)]\|\mathscr{L}(0)\|_r \tag{5}$$

Proof. Using the equations of motion (1a)–(1c), we have

$$d\mathscr{L}_i(t)/dt = p_i(t)R_i(\mathbf{q}(t)) \tag{6}$$

By conditions A_3 and A_4

$$(d/dt)\mathscr{L}_i(t) \leqslant |d\mathscr{L}_i(t)/dt| \leqslant \sum_j A_{ij}\mathscr{L}_j(t) \tag{7}$$

where the A_{ij} are constants, independent of t; $0 \leqslant A_{ij} \leqslant C$; and $A_{ij} = 0$ for $|i - j| > D$, the range of the potential. If $\mathbf{A}$ denotes the matrix with elements A_{ij}, then[(2)] $\Psi(t) = [\exp(\mathbf{A}t)]|\mathscr{L}(0)|$ is a solution of the equations

$$d\Psi_i/dt = \sum A_{ij}\Psi_j(t), \qquad \Psi_j(0) = |\mathscr{L}_j(0)| \tag{8}$$

Standard arguments show that $|\mathscr{L}_i(t)| \leqslant \Psi_i(t)$, so Eq. (5) follows from (7) with a equal to the r-norm of the bounded operator $\mathbf{A}$ on B_r.

Theorem 1. Let $\mathbf{q}(0)$, $\mathbf{p}(0)$ be such that $\mathscr{L}(0)$ (defined in Lemma 1) belongs to B_r. There then exists a $\in B_r$ solution of Eqs. (1a)–(1c) defined for all t.

Proof. We shall first consider the case of a finite system in a bounded region $\Lambda_\alpha \subset \mathbb{Z}^\nu$. Let $q_i^\alpha(t)$, $p_i^\alpha(t)$ be the solutions of the equations

$$\left.\begin{aligned} dq_i^\alpha/dt &= p_i^\alpha(t) \\ dp_i^\alpha/dt &= F_i(\mathbf{q}^\alpha(t)) \end{aligned}\right\} \quad \text{for} \quad i \in \Lambda_\alpha \tag{9a, 9b}$$

$$dq_i^\alpha/dt = dp_i^\alpha/dt = 0 \quad \text{for} \quad i \notin \Lambda_\alpha \tag{9c}$$

with the initial conditions $q_i^\alpha(0) = q_i(0)$, $p_i^\alpha(0) = p_i(0)$, i.e., Eqs. (9a)–(9c) are time evolution equations for $\mathbf{q}^\alpha(t)$, $\mathbf{p}^\alpha(t)$ with *all* the particles outside Λ_α "tied down" to their initial positions.[2,3] Solutions of (9a)–(9c) are prevented from going to infinity in finite time by Lemma 1; they therefore exist for all time. The time evolution mappings T_t^α generated by (9a)–(9c) leave invariant the energy in Λ_α,

$$H_\alpha(\mathbf{q}, \mathbf{p}) = \sum_{i \in \Lambda_\alpha} [\tfrac{1}{2}p_i^2 + U_i(q_i)] + \sum_j{}' V_j(\mathbf{q})$$

where $\sum'$ is the sum over all j such that $\operatorname{dist}(j, \Lambda_\alpha) \leqslant D$. The solutions of (9a)–(9c) will satisfy the equations

$$\left.\begin{aligned} q_i^\alpha(t) &= q_i^\alpha(0) + p_i^\alpha(0)t + \int_0^t (t - s)F_i(\mathbf{q}^\alpha(s))\, ds \\ p_i^\alpha(t) &= p_i^\alpha(0) + \int_0^t F_i(\mathbf{q}^\alpha(s))\, ds \end{aligned}\right\} \quad \text{for} \quad i \in \Lambda_\alpha \tag{10a, 10b}$$

$$q_i^\alpha(t) = q_i^\alpha(0), \qquad p_i^\alpha(t) = p_i^\alpha(0) \qquad \text{for} \quad i \notin \Lambda_\alpha \tag{10c}$$

Using now the bound (5) for the time evolution T_t^α, we have, by condition A_3, that $|F_i(\mathbf{q}^\alpha(t))| < K_i$ for $t \in [0, T]$ with $K_i < \infty$ independent of Λ_α. Hence by the Arzela–Ascoli theorem we can choose sequences $\Lambda_\alpha \to \mathbb{Z}^\nu$ such that $q_i^\alpha(t)$, $p_i^\alpha(t) \to q_i(t)$, $p_i(t)$ uniformly on $[0, T]$. This is true for all finite T, so the sequence can be further refined to get uniform convergence on *every* bounded interval. The $q_i(t)$ will satisfy Eq. (3), so the $(\mathbf{q}(t), \mathbf{p}(t))$ satisfy (1a)–(1c), the equations of motion for the infinite system, with the initial conditions $(\mathbf{q}(0), \mathbf{p}(0))$.

By our assumption, $\mathscr{L}(0) \in B_r$. Hence, by (5), we also have an estimate of the form

$$\mathscr{L}_j(t) = \tfrac{1}{2}p_i^2(t) + U_j(q_j(t)) + K < K' \exp(r|j|), \qquad |t| \leqslant T$$

for each T (but where K' grows with T). By A_3 this implies

$$|q_j(t)| < C' \exp(r|j|), \qquad |p_j(t)| < C'' \exp(\tfrac{1}{2}r|j|) \tag{11}$$

This gives us rather good control over the time evolution, e.g., if the initial values are bounded, $|q_i(0)| < C$ and $|p_i(0)| < C$, then $q_i(t)$ and $p_i(t)$ will also be bounded for all finite t.

3. UNIQUENESS OF TIME EVOLUTION

Having established the existence of solutions of Eqs. (1a)–(1c) for a large class of initial conditions, we now consider their uniqueness. As is generally the case, e.g., for harmonic systems[2] we can obtain uniqueness only if we impose some conditions on how the solution $\{q_j(t), p_j(t)\}$ grows with $|j|$.

Definition. For any family $\mathbf{B} = \{B_i\}$ of positive constants, define $\Delta(\mathbf{B}) = \{\mathbf{q}: |q_i| \leqslant B_i \text{ for all } i\}$ and define $\bar{B}_k = \sup\{B_i: |i| \leqslant k\}$, $k = 1, 2, \ldots$. We will say $\mathbf{B}$ is a *sequence of uniqueness* if (a) the following holds:

$$\limsup_{k\to\infty} \bar{B}_k^{1/k} < \infty \tag{12a}$$

and (b) there exists a constant c such that

$$\sup_{\mathbf{q}\in\Delta(\mathbf{B})} \sum_j |\partial F_i(\mathbf{q})/\partial q_j| \leqslant ci^2 \qquad \text{for all } i \tag{12b}$$

Theorem 2. Let $\mathbf{B}$ be a sequence of uniqueness. Then two solutions $\mathbf{q}^{(1)}(t)$ and $\mathbf{q}^{(2)}(t)$ of (3), both defined on $[0, T]$ and both taking values in $\Delta(\mathbf{B})$, are identical on $[0, T]$.

Proof. Assume the contrary. Then we can assume that there are arbitrarily small, positive t's for which $\mathbf{q}^{(1)}(t) \neq \mathbf{q}^{(2)}(t)$. We will show that this leads to a contradiction. Writing out (3) for $\{q^{(1)}(t)\}$ and $\{q^{(2)}(t)\}$ and subtracting the two gives

$$q_i^{(1)}(t) - q_i^{(2)}(t) = \int_0^t dt_1\,(t - t_1)[F_i(\mathbf{q}^{(1)}(t_1)) - F_i(\mathbf{q}^{(2)}(t_1))]$$

Let

$$\delta_n(t) = \sup\{|q_i^{(1)}(t) - q_i^{(2)}(t)|: \quad |i| \leqslant nD\}$$

where D is the range of the potential, as defined in A_1. We then get, using (12b),

$$\delta_n(t) \leqslant \left[\int_0^t dt_1\,(t - t_1)\delta_{n+1}(t_1)\right] cn^2$$

Iterating this k times, then using the bound $\delta_{n+k}(t) \leqslant 2\bar{B}_{(n+k)D}$, we obtain

$$\delta_n(t) \leqslant \left[\int_0^t dt\,(t-t_1)^{2k-1}\delta_{n+k}(t)\right] c^k[n(n+1)\cdots(n+k-1)]^2$$

$$\leqslant \frac{t^{2k}}{(2k)!}\,(2\bar{B}_{(n+k)D})c^k[n(n+1)\cdots(n+k-1)]^2$$

Thus, letting $k \to \infty$, we find that $\delta_n(t) = 0$ for

$$0 < t < \left\{\limsup_{k\to\infty}\left[\frac{2\bar{B}_{(n+k)D}c^k[n(n+1)\cdots(n+k-1)]^2}{(2k)!}\right]^{1/2k}\right\}^{-1}$$

$$= 2(\tfrac{1}{2}\sqrt{c}\limsup_{k\to\infty}\bar{B}_k^{D/2k})^{-1}$$

This is true for all n, so

$$q_i^{(1)}(t) = q_i^{(2)}(t)$$

for all i, provided

$$t < 4[c\limsup_{k\to\infty}(B_k^{D/k})]^{-1/2}$$

which proves the theorem.

Example. If (as in the example of Section 2) there exists a constant c_1 such that

$$|\partial F_i/\partial q_j| \leqslant c_1(\sup\{|q_j|:\ \ |i-j| \leqslant D\})^{2n-2} \tag{13}$$

then any sequence of the form $B_j = b|j|^{1/(n-1)}$ is a sequence of uniqueness if $n \geqslant 2$.

This means that we have uniqueness in the class of solutions such that

$$\sup_{|t|\leq\tau}\sup\{|q_j(t)|/(|j|^{1/(n-1)}+1)\} < \infty \qquad \text{for all } \tau$$

Arguments similar to those leading to Eq. (11) show that if $\mathscr{L}_j(0)$ grows no faster than $|j|^{1/(n-1)}$, then there does exist a solution in this class.

In the harmonic case ($n = 1$), condition (12b) is vacuous and any sequence (B_j) such that

$$\sup_k\{\bar{B}_k^{1/k}\} < \infty$$

is a sequence of uniqueness (compare Ref. 2).

4. WEAK TIME EVOLUTION FOR GENERAL INTERACTIONS

In this section we sketch a proof that, under very general assumptions, solutions to the equations of motion exist for almost all initial conditions

with respect to any Gibbs state. We do not assume here that conditions A_3 and A_4 hold. The proof is very simple and almost nothing needs to be assumed about the interaction, but it should be noted that very reasonable interactions —such as the one-dimensional harmonic chain with formal interaction energy $\frac{1}{2}\sum_i (q_{i+1} - q_i)^2$—do not have any Gibbs states at all.[2] About such interactions our theorem evidently says nothing. We refer the reader to recent work for an analysis of Gibbs states for the kind of system considered here.[7-9]

We will assume as before that our interaction is of Hamiltonian form with range D, i.e., we assume that A_1 and A_2 hold. In addition, we assume that:

B_1. For each finite subset Λ_α of $\mathbb{Z}^\nu$, the equations of motion (9a)–(9c) admit solutions for all time for all initial points.

B_2. For each Λ_α, each $\beta > 0$, and each specification of the q_i for $i \notin \Lambda_\alpha$, the measure

$$\exp[-\beta H_\alpha(\mathbf{q}, \mathbf{p})] \prod_{j \in \Lambda_\alpha} dq_j\, dp_j \tag{14}$$

with H_α given in (10) is finite (normalizable) on $(R \times R)^{\Lambda_\alpha}$.

Condition B_2 makes it possible to define Gibbs states by an obvious adaptation of the definitions used in other cases, but it does not imply the *existence* of nontrivial Gibbs states.

We note that (i) by conservation of energy and Liouville's theorem, any Gibbs state is invariant under T_t^α for all α, t; (ii) with respect to any Gibbs state, the p_i are independent, identically distributed, Gaussian random variables of mean zero.

Theorem 3. Let μ be a Gibbs state for the interaction under consideration. For μ-almost all initial points $\{q_i, p_i\}$, there exists a solution $\{q_i(t)\}$ of Eq. (3) defined for all t and satisfying

$$\sup_t \sup_i \frac{|q_i(t) - q_i|}{(1 + t^2)[\log_+(i)]^{1/2}} < \infty \tag{15}$$

where $\log_+(j) = \sup[\log|j|, 1]$

Proof. (The argument here is similar to that used in Ref. 3.) For any $x = \{q_i, p_i\}$ define

$$B(x) = \sup_i \frac{|p_i|}{[\log_+(q_i)]^{1/2}}$$

$$\bar{B}_\alpha(x) = \int_{-\infty}^{\infty} \frac{dt}{1 + t^2} B(T_t^\alpha x), \qquad \bar{B}_\infty(x) = \liminf_{\alpha \to \infty} \bar{B}_\alpha(x)$$

It follows from (ii) that $\int B\, d\mu < \infty$; hence, from (i) and Fubini's theorem, $\int \bar{B}_\alpha\, d\mu$ is finite and independent of α. By Fatou's lemma, $\int \bar{B}_\alpha\, d\mu < \infty$. We will show: If $\bar{B}_\infty(x) < \infty$, then there exists a solution to (3) satisfying (15).

To see this, note first that there must then exist a sequence $\alpha_n \to \infty$ and a constant C such that

$$\bar{B}_{\alpha_n}(x) \leqslant C \qquad \text{for all } n$$

Hence,

$$\begin{aligned} |q_i^{\alpha_n}(t) - q_i| &\leqslant \left| \int_0^t dt_1\, p_i^{\alpha_n}(t_1) \right| \leqslant (1 + t^2) \int_{-\infty}^{\infty} \frac{dt_1}{1 + t_1^2}\, |p_i^{\alpha_n}(t_1)| \\ &\leqslant (1 + t^2)[\log_+(i)]^{1/2} \bar{B}_{\alpha_n}(x) \\ &\leqslant (1 + t^2)[\log_+(i)]^{1/2} C \qquad \text{for all } i, n, t \end{aligned} \tag{16}$$

Since F_i depends only on a finite number of the q_j, and since each V_j is continuously differentiable, this bound implies a family of bounds of the form

$$|dp_i^{\alpha_n}(t)/dt| \leqslant K_i(|t|)$$

where each K_i is a nondecreasing function of t (which does not depend on n). The proof of the existence of solutions is now completed in the same way as in Theorem 1; (15) follows from (16) by passage to the limit.

There remains the question of uniqueness. Suppose that (13) holds with some $n > 1$. If $\{q_i, p_i\}$ satisfies

$$\sup_{i \in \mathbb{Z}^\nu} (|q_i|/|i|^{1/(n-1)}) < \infty \tag{17}$$

and if there exists a solution to (3) satisfying (15), then

$$\sup_{|t| \leq \tau} \sup_i [|q_i(t)|/|i|^{1/(n-1)}] \quad \text{is finite for all } \tau$$

Theorem 2 asserts that the solution is unique in this class. We would therefore like to know whether the condition (17) holds μ almost everywhere. A sufficient condition is given by the following:

Proposition. If there exists $\gamma > \nu(n-1)$ and C such that

$$\int |q_i|^\gamma\, d\mu < C \qquad \text{for all } i \tag{18}$$

then

$$\sup_{i \in \mathbb{Z}^\nu} (|q_i|/|i|^{1/(n-1)}) < \infty \qquad \mu \text{ almost everywhere}$$

Proof.

$$\mu\{|q_i| > |i|^{1/(n-1)}\} = \mu\{|q_i|^\gamma > |i|^{\gamma/(n-1)}\}$$
$$\leqslant \left(\int |q_i|^\gamma \, d\mu\right)/i^{\gamma/(n-1)}$$

Since $\gamma(n-1) > \nu$,

$$\sum_{i\in\mathbb{Z}^\nu} \mu\{|q_i| > |i|^{1/(n-1)}\} < \infty$$

so by the Borel–Cantelli lemma[10]

$$\lim_i \sup(|q_i|/|i|^{1/(n-1)}) \leqslant 1 \qquad \mu \text{ almost everywhere}$$

Collecting the above results, we have the following:

Theorem 4. Let the interactions satisfy conditions A_1, A_2, B_1, and B_2 and also (13) with $n > 1$. Let μ be a Gibbs state for this interaction such that, for some $\gamma > \nu(n-1)$, (18) is satisfied. Then for μ-almost all $\{q_i, p_i\}$ there exists a solution to (3) such that

$$\sup_{|t|\leq\tau} \sup_{i\in\mathbb{Z}^\nu} [|q_i(t)|/|i|^{1/(n-1)}] < \infty \qquad \text{for all } \tau$$

and this solution is unique.

REFERENCES

1. O. E. Lanford III, *Commun. Math. Phys.* **9**:169 (1969); **11**:257 (1969).
2. O. E. Lanford III and J. L. Lebowitz, in *Lecture Notes in Physics*, No. 38, Springer-Verlag (1975), p. 144; J. L. van Hemmen, Thesis, University of Groningen (1976).
3. O. E. Lanford III, in *Lecture Notes in Physics*, No. 38, Springer-Verlag (1975), p. 1.
4. Ya. G. Sinai, *Vestnik Markov. Univ. Ser. I, Math. Meh.* **1974**:152.
5. C. Marchioro, A. Pellegrinotti, and E. Presutti, *Commun. Math. Phys.* **40**:175 (1975).
6. N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, Holt, Rinehart and Winston (1976).
7. H. J. Brascamp, E. H. Lieb, and J. L. Lebowitz, The Statistical Mechanics of Anharmonic Lattices, in *Proceedings of the 40th Session of the International Statistics Institute*, Warsaw (1975).
8. D. Ruelle, *Commun. Math. Phys.* **50**:189 (1976).
9. J. L. Lebowitz and E. Presutti, *Commun. Math. Phys.* **50**:195 (1976).
10. L. Breiman, *Probability*, Addison-Wesley, Section 3.14.

Journal of Statistical Physics, Vol. 22, No. 2, 1980

Equilibrium Time Correlation Functions in the Low-Density Limit

H. van Beijeren,[1] O. E. Lanford III,[2] J. L. Lebowitz,[3] and H. Spohn[4]

Received June 5, 1979

We consider a system of hard spheres in thermal equilibrium. Using Lanford's result about the convergence of the solutions of the BBGKY hierarchy to the solutions of the Boltzmann hierarchy, we show that in the low-density limit (Boltzmann–Grad limit): (i) the total time correlation function is governed by the linearized Boltzmann equation (proved to be valid for short times), (ii) the self time correlation function, equivalently the distribution of a tagged particle in an equilibrium fluid, is governed by the Rayleigh–Boltzmann equation (proved to be valid for all times). In the latter case the fluid (not including the tagged particle) is to zeroth order in thermal equilibrium and to first order its distribution is governed by a combination of the Rayleigh–Boltzmann equation and the linearized Boltzmann equation (proved to be valid for short times).

KEY WORDS: Time correlation functions; low-density limit; linearized Boltzmann equation; Boltzmann–Grad limit.

1. INTRODUCTION

In order to motivate the limits studied in this paper we consider first a fluid of hard spheres of diameter one and unit mass at low densities $\rho_\epsilon = \epsilon\rho$, $\epsilon \to 0$. In many cases of physical interest one expects in this regime typical spatial variations of the fluid to be of the order of a mean free path, $\sim 1/\epsilon$, and typical time variations to be of the order of a mean free time, $\sim 1/\epsilon$. Therefore, in order to study the dynamics of the fluid on its proper time and space scale, it is convenient to rescale time and space as

$$t' = \epsilon t, \qquad q' = \epsilon q \tag{1.1}$$

Supported in part by NSF Grant PHY 78-22302.

[1] Fachber. Physik, TH Aachen, Aachen, West Germany.
[2] Department of Mathematics, University of California at Berkeley, Berkeley, California.
[3] Department of Mathematics, Rutgers University, New Brunswick, New Jersey.
[4] Fachber. Physik, Universität München, Munich, West Germany.

0022-4715/80/0200-0237$03.00/0

Here t and q are the dynamical variables appearing in the equations of motion and t' and q' are the rescaled variables. The velocities and the mass of the particles remain unscaled.

In the rescaled t', q' variables typical time and space variations are of order one. On this scale a particle has diameter ϵ and the number of particles per unit volume increases as ϵ^{-2} in three dimensions. The $\epsilon \to 0$ limit is called the Boltzmann–Grad limit, since Grad[1] first wrote down and discussed this limit as the appropriate one for the exact validity of the Boltzmann equation. Subsequently, Lanford indeed proved[2,3] that, at least for short times, the nonlinear Boltzmann equation becomes exact in the Boltzmann–Grad limit for a rather general class of initial conditions on the n-particle correlation functions. The purpose of this paper is to study in the same limit equilibrium time correlation functions.

The self time correlation function can be regarded as describing the dynamics of a test particle in the fluid; e.g., imagine particle one painted red. Therefore this correlation function is governed, in the low-density limit, by the Rayleigh–Boltzmann equation, which is obtained from the nonlinear Boltzmann equation by replacing in the quadratic collision term the distribution function that is integrated over by the Maxwellian equilibrium distribution. The total time correlation function describes the time-dependent fluctuations of the fluid in thermal equilibrium. It is therefore governed, in the low-density limit, by the linearized Boltzmann equation which is obtained by linearizing the collision term at the Maxwellian.[4–6]

Our results are quite analogous to the fluctuation results obtained for the Vlasov equation by Braun and Hepp[7] (and for its quantum counterparts, the mean field models as studied by Hepp and Lieb[8]). They are only less complete in the sense that we can prove convergence only for short times and that, instead of proving a central limit theorem, we can show only convergence of the covariance.

2. THE LOW-DENSITY LIMIT. LANFORD'S THEOREM

We describe Lanford's result[2] about the convergence of the solutions of the BBGKY hierarchy to the solutions of the Boltzmann hierarchy. Since we will use an iteration argument later, we state the theorem as in King's thesis.[9]

We consider a system of hard spheres of diameter ϵ and unit mass inside a bounded region Λ with smooth boundary $\partial\Lambda$. The spheres (particles) are elastically reflected among themselves and at the boundary $\partial\Lambda$. Let the state of the system be specified by the absolutely continuous probabilities of finding exactly n particles at $dx_1 \cdots dx_n$

$$\left\{ f_n(x_1, \ldots, x_n) \frac{1}{n!} dx_1 \cdots dx_n | n \geqslant 0 \right\}$$

Here $x_i = (q_i, p_i) \in \Lambda \times R^3$ stands for the position of the center and the momentum of the ith particle. Then the distribution functions $\{\rho_n^\epsilon | n \geqslant 0\}$ corresponding to this state are defined by

$$\rho_n^\epsilon(x_1,\ldots, x_n) = \sum_{m=0}^{\infty} \frac{1}{m!} \int_{(\Lambda \times R^3)^m} dy_1 \cdots dy_m f_{n+m}(x_1,\ldots, x_n, y_1,\ldots, y_m) \tag{2.1}$$

The time evolution of a state of the hard-sphere system is studied by means of the time evolution of the corresponding distribution functions. A straightforward computation, which is, however, nontrivial to justify rigorously,[10,11] leads to the following evolution equation:

$$\begin{aligned} \frac{\partial}{\partial t} \rho_n^\epsilon(x_1,\ldots, x_n, t) &= H_n^\epsilon \rho_n^\epsilon(x_1,\ldots, x_n, t) \\ &\quad + \epsilon^2 \sum_{j=1}^{n} \int_{R^3} dp_{n+1} \int_{S^2} d\omega\, \omega \cdot (p_{n+1} - p_j) \rho_{n+1}^\epsilon(x_1,\ldots, x_n, q_j + \epsilon\omega, p_{n+1}, t) \end{aligned} \tag{2.2}$$

Here ω is a unit vector in R^3 and $d\omega$ is the surface measure of the unit sphere S^2 in three dimensions. H_n^ϵ describes the evolution of n hard spheres of diameter ϵ inside Λ. Equation (2.2) is the *BBGKY hierarchy* for hard spheres. The solutions of the BBGKY hierarchy are denoted by

$$\rho_n^\epsilon(x_1,\ldots, x_n, t) = (V_t^\epsilon \rho^\epsilon)_n(x_1,\ldots, x_n) \tag{2.3}$$

for the initial vector of distribution functions $\rho^\epsilon = (\rho_1^\epsilon, \rho_2^\epsilon,\ldots)$.

Remark. The phase space of n hard spheres in Λ is

$$\begin{aligned} \mathscr{X}(n, \epsilon) = \{(q_1, p_1,\ldots, q_n, p_n) \in (\Lambda \times R^3)^n | \; & |q_i - q_j| \geqslant \epsilon \text{ for } i \neq j, \\ & \operatorname{dist}(q_i, \partial\Lambda) \geqslant \epsilon/2\} \end{aligned}$$

In this space, boundary points of $\mathscr{X}(n, \epsilon)$ corresponding to a collision with the wall $\partial\Lambda$ and to a collision between two spheres are identified. E.g., if $q_j = q_i + \epsilon\omega$, $i \neq j$, and with incoming momenta p_i, p_j going over to p_i', p_j' in a collision, then $(q_1, p_1,\ldots, q_i, p_i,\ldots, q_i + \epsilon\omega, p_j,\ldots, q_n, p_n)$ is identified with $(q_1, p_1,\ldots, q_i, p_i',\ldots, q_i + \epsilon\omega, p_j',\ldots, q_n, p_n)$. There remains a set of "bad" points in $\partial\mathscr{X}(n, \epsilon)$ corresponding to triple and grazing collisions. In the interior of $\mathscr{X}(n, \epsilon)$ the time evolution is defined by free motion with infinitesimal generator $-\sum_{j=1}^n p_j \partial/\partial q_j$. This prescription extends smoothly through the points of $\partial\mathscr{X}(n, \epsilon)$ corresponding to pair collisions and to collisions with the wall $\partial\Lambda$. Points lying on trajectories leading to the bad points of $\mathscr{X}(n, \epsilon)$ form a set of Lebesgue measure zero. On this set the time evolution remains

undefined. (Cf. the thesis of Alexander[12] for a detailed treatment of the time evolution of hard spheres.)

At this stage we can formally lift the restriction that Λ has to be a bounded region. So Λ may be, for example, a slab or the whole three-dimensional space. It is also clear that specular reflection at $\partial\Lambda$ is only one choice out of many possible boundary conditions: we could consider, for example, a stochastic boundary condition at $\partial\Lambda$ corresponding to a wall with a certain temperature. All these boundary conditions would be included in the definition of H_n^ϵ.

We want to study the low-density limit of the solutions of the BBGKY hierarchy. The low-density (Boltzmann–Grad) limit is obtained by letting the fraction of volume occupied by the particles $\sim \rho\epsilon^3$, with ρ the average density, go to zero while keeping the mean free path of the hard spheres, $\sim 1/\epsilon^2\rho$, constant. This requires that, as $\epsilon \to 0$, the density is increased as ϵ^{-2}. Therefore for each hard-sphere diameter ϵ one chooses an initial state with distribution functions ρ_n^ϵ such that $\rho_n^\epsilon \sim \epsilon^{-2n}$. With this in mind we define the *rescaled distribution functions*

$$r_n^\epsilon(x_1,\ldots, x_n) = \epsilon^{2n}\rho_n^\epsilon(x_1,\ldots, x_n) \tag{2.4}$$

Then (2.2) reads

$$\frac{d}{dt} r_n^\epsilon(t) = H_n^\epsilon r_n^\epsilon(t) + C_{n,n+1}^\varepsilon r_{n+1}^\varepsilon(t) \tag{2.5}$$

where the collision term in that equation is abbreviated as $C_{n,n+1}^\varepsilon$. Regarding the sequence $\{r_n^\epsilon | n \geqslant 0\}$ as the vector r^ϵ, one can write (2.5) compactly as

$$\frac{d}{dt} r^\epsilon(t) = H^\epsilon r^\epsilon(t) + C^\epsilon r^\epsilon(t) \tag{2.6}$$

where H^ϵ is a diagonal matrix with entries H_n^ϵ, and C^ϵ is a matrix with entries $C_{n,n+1}^\varepsilon$ and zero otherwise.

Let us now consider H^ϵ as the unperturbed part of the operator $H^\epsilon + C^\epsilon$ and C^ϵ as the perturbation. The time-dependent (Dyson) perturbation series for the solution of (2.6) then reads

$$r^\epsilon(t) = \sum_{m=0}^{\infty} \int_{0\leqslant t_1\leqslant\cdots\leqslant t_m\leqslant t} dt_m \cdots dt_1\, S^\epsilon(t - t_m)C^\epsilon \cdots C^\epsilon S^\epsilon(t_1)r^\epsilon \tag{2.7}$$

where r^ϵ stands for $r^\epsilon(0)$, and where $(S^\epsilon(t)r^\epsilon)_n = ([\exp(H^\epsilon t)]r^\epsilon)_n = [\exp(H_n^\epsilon t)]r_n^\epsilon$ gives the evolution of n hard spheres of diameter ϵ inside Λ, always including the specular reflection at $\partial\Lambda$. Solutions of the BBGKY hierarchy are always understood in the sense of (2.7). Of course, one has to say in what sense (2.7) converges.

For $t \geqslant 0$ the time evolution of $r_n^{\epsilon}(t)$ is determined by backward streaming. Therefore it seems natural to replace, for a collision, the phase point $(x_1,..., q_j, p_j,..., q_j + \epsilon\omega, p_{n+1})$ with outgoing momenta by the phase point $(x_1,..., q_j, p_j',..., q_j + \epsilon\omega, p_{n+1}')$ with incoming momenta. (As explained before, these are just two different representations of the same phase point.) This leads to

$$\begin{aligned}
&\frac{\partial}{\partial t} r_n^{\epsilon}(x_1,..., x_n, t) \\
&\quad = H_n^{\epsilon} r_n^{\epsilon}(x_1,..., x_n, t) \\
&\qquad + \sum_{j=1}^{n} \int_{+} dp_{n+1}\, d\omega\, \omega\cdot(p_j - p_{n+1}) \\
&\qquad \times \{r_{n+1}^{\epsilon}(x_1,..., q_j, p_j',..., q_j - \epsilon\omega, p_{n+1}', t) \\
&\qquad - r_{n+1}^{\epsilon}(x_1,..., q_j, p_j,..., q_j + \epsilon\omega, p_{n+1}, t)\}
\end{aligned} \tag{2.8}$$

where $\int_+$ indicates that the integration over ω is restricted to the upper hemisphere $\omega\cdot(p_j - p_{n+1}) \geqslant 0$. Formally, the limiting form of (2.8), which the limiting distribution functions $r(t) = \lim_{\epsilon\to 0} r^{\epsilon}(t)$ might satisfy, for $t \geqslant 0$, is

$$\begin{aligned}
&\frac{\partial}{\partial t} r_n(x_1,..., x_n, t) \\
&\quad = -\sum_{j=1}^{n} p_j \frac{\partial}{\partial q_j} r_n(x_1,..., x_n, t) \\
&\qquad + \sum_{j=1}^{n} \int_{+} dp_{n+1}\, d\omega\, \omega\cdot(p_j - p_{n+1}) \\
&\qquad \times \{r_{n+1}(x_1,..., q_j, p_j',..., q_j, p_{n+1}', t) \\
&\qquad - r_{n+1}(x_1,..., q_j, p_j,..., q_j, p_{n+1}, t)\}
\end{aligned} \tag{2.9}$$

(Implicitly, the free motion $-\sum_{j=1}^{n} p_j\, \partial/\partial q_j$ includes the specular reflection at $\partial\Lambda$.)

For $t \leqslant 0$ the time evolution of $r_n^{\epsilon}(t)$ is determined by forward streaming. In that case, for a collision, the phase point $(x_1,..., q_j, p_j,..., q_j + \epsilon\omega, p_{n+1})$ with incoming momenta should be replaced by the phase point $(x_1,..., q_j, p_j',..., q_j + \epsilon\omega, p_{n+1}')$ with outgoing momenta. The formal limit of the resulting equation is then again (2.9) but with the sign of the collision term reversed.

Equation (2.9) for $t \geqslant 0$ and Eq. (2.9) with the sign of the collision term reversed for $t \leqslant 0$ is called the *Boltzmann hierarchy*, which can be written in the form

$$\frac{d}{dt} r_n(t) = H_n r_n(t) + C_{n,n+1} r_{n+1}(t)$$

or compactly

$$\frac{d}{dt} r(t) = Hr(t) + Cr(t) \tag{2.10}$$

Letting $(S(t)r)_n = (e^{Ht}r)_n = e^{H_n t} r_n$ denote the free motion of n particles inside Λ, the time-dependent perturbation series for the Boltzmann hierarchy reads

$$r(t) = \sum_{m=0}^{\infty} \int_{0 \leqslant t_1 \leqslant \cdots \leqslant t_m \leqslant t} dt_m \cdots dt_1 \, S(t - t_m) C \cdots C S(t_1) r \tag{2.11}$$

To prove that $r^\epsilon(t)$ defined by (2.7) indeed converges to $r(t)$ defined by (2.11) as $\epsilon \to 0$, we need two conditions.

First, the initial distributions r^ϵ have to be uniformly bounded in ϵ. This guarantees the uniform convergence of the perturbation series (2.7) for some interval $|t| \leqslant t_0$. If h_β denotes the normalized Maxwellian at inverse temperature β, then a suitable choice for this bound is as follows:

(C1) There exist a pair (z, β) such that

$$r_n^\epsilon(x_1, \ldots, x_n) \leqslant M z^n \prod_{j=1}^{n} h_\beta(p_j) \tag{2.12}$$

for all $\epsilon < \epsilon_0$ with a positive constant M independent of ϵ.

Second, r_n^ϵ has to converge to r_n in such a way that the series (2.7) converges term by term to the series (2.11). For the initial phase point $x^{(n)} = (x_1, \ldots, x_n) \in (\Lambda \times R^3)^n$ let $q_j(t, x^{(n)})$, $j = 1, \ldots, n$, be the position of the jth point particle at time t under the free motion. Then

$$\Gamma_n(t) = \{x^{(n)} = (x_1, \ldots, x_n) \in (\Lambda \times R^3)^n | q_i(s, x^{(n)}) \neq q_j(s, x^{(n)}) \text{ for } i \neq j = 1, \ldots, n \text{ and } -t \leqslant s \leqslant 0 \text{ if } t \geqslant 0, 0 \leqslant s \leqslant -t \text{ if } t \leqslant 0\}$$

In words, $\Gamma_n(t)$ is the restriction of the n-particle phase space to the set of phase points that under free backward streaming over a time t, if t is positive (or free forward streaming over a time $|t|$, if t is negative) do not lead to a collision between any pair of particles, regarded as point particles. By this restriction only a set of Lebesgue measure zero is excluded from $(\Lambda \times R^3)^n$.

Note that: (i) $\Gamma_n(t)$ depends only on the free motion, (ii) $\Gamma_n(t) \subset \Gamma_n(t')$ for $t' = \alpha t$, $\alpha \geqslant 1$, (iii) $\Gamma_n(t) \neq \Gamma_n(-t)$, and (iv) $x^{(n)} \in \Gamma_n(t)$ is equivalent to $\bar{x}^{(n)} \in \Gamma_n(-t)$, where $\bar{x}^{(n)}$ is the phase point obtained from $x^{(n)}$ under the reversal $p_j \mapsto -p_j$. In particular $\Gamma_n(t)$ is not invariant under reversal of velocities.

The suitable choice of convergence is then as follows:

(C2) There exists a continuous function r_n on $(\Lambda \times R^3)^n$ such that

$$\lim_{\epsilon \to 0} \epsilon^{2n} \rho_n^\epsilon = \lim_{\epsilon \to 0} r_n^\epsilon = r_n \tag{2.13}$$

uniformly on all compact sets of $\Gamma_n(s)$ for some $s \geqslant 0$.

Theorem (Lanford). Let $\{\rho_n^\epsilon | n \geqslant 0\}$ be a sequence of initial distribution functions of a fluid of hard spheres of diameter ϵ inside a region Λ and let the sequence $\{r_n^\epsilon | n \geqslant 0\}$ of rescaled distribution functions satisfy (C1) and (C2). Let $r_n^\epsilon(t)$ be the solution of the BBGKY hierarchy with initial conditions r_n^ϵ, and let $r_n(t)$ be the solution of the Boltzmann hierarchy with initial conditions r_n.

Then there exists a $t_0(z, \beta) > 0$ such that for $0 \leqslant t \leqslant t_0(z, \beta)$ the series (2.7) and (2.11) converge and such that $r_n^\epsilon(t)$ satisfies a bound of the form (C1) with $z' > z$ and $\beta' < \beta$. Furthermore,

$$\lim_{\epsilon \to 0} r_n^\epsilon(t) = r_n(t) \tag{2.14}$$

uniformly on compact sets of $\Gamma_n(s + t)$.

For $-t_0(z, \beta) \leqslant t \leqslant 0$, (2.14) holds provided that $s \leqslant 0$ and that in the Boltzmann hierarchy the collision term $C_{n,n+1}$ is replaced by $-C_{n,n+1}$.

Remark. It is the conditions for the validity of the limit (2.14) that make the irreversible nature of the Boltzmann hierarchy consistent with the reversibility of the BBGKY hierarchy (cf. Appendix A).

We now describe three interesting properties of the Boltzmann hierarchy. The first one is the well-known "propagation of chaos."

Property 1. If the initial conditions of the Boltzmann hierarchy factorize,

$$r_n(x_1, \ldots, x_n) = \prod_{j=1}^{n} f(x_j) \tag{2.15}$$

then the solutions with this initial condition stay factorized,

$$r_n(x_1, \ldots, x_n, t) = \prod_{j=1}^{n} f(x_j, t). \tag{2.16}$$

$f(x, t)$ is the solution of the *Boltzmann equation*

$$\frac{\partial}{\partial t} f(q, p, t) = -p \frac{\partial}{\partial q} f(q, p, t) + \int_+ dp_1 \, d\omega \, \omega \cdot (p - p_1) \times \{f(q, p', t) f(q, p_1', t) - f(q, p, t) f(q, p_1, t)\} \tag{2.17}$$

with initial condition $f(q, p)$.

The second property comes from considering one of the fluid particles as a test particle, e.g., imagine particle one painted red.

Property 2. If the initial conditions of the Boltzmann hierarchy are of the form

$$r_n(x_1, \ldots, x_n) = f(x_1) \prod_{j=1}^{n} \{z h_\beta(p_j)\} \tag{2.18}$$

corresponding to an initial test particle distribution of the form $f(x_1)zh_\beta(x_1)$, then its solutions are

$$r_n(x_1,\ldots, x_n, t) = f(x_1, t)\prod_{j=1}^{n}\{zh_\beta(x_j)\} \tag{2.19}$$

$f(x, t)$ is the solution of the *Rayleigh–Boltzmann equation*

$$\frac{\partial}{\partial t}f(q, p, t) = -p\frac{\partial}{\partial q}f(q, p, t) + z\int_+ dp_1\, d\omega\, \omega\cdot(p - p_1)h_\beta(p_1)$$
$$\times\{f(q, p', t) - f(q, p, t)\} = (Af(t))(q, p) \tag{2.20}$$

with initial condition $f(q, p)$. Equation (2.20) is also known as Lorentz–Boltzmann equation or linear Boltzmann equation.

Finally, we have the following property:

Property 3. If the initial conditions of the Boltzmann hierarchy are of the form

$$r_n(x_1,\ldots, x_n) = \left[\sum_{j=1}^{n} f(x_j)\right]\prod_{j=1}^{n}\{zh_\beta(x_j)\} \tag{2.21}$$

then its solutions are

$$r_n(x_1,\ldots, x_n, t) = \left[\sum_{j=1}^{n} f(x_j, t)\right]\prod_{j=1}^{n}\{zh_\beta(x_j)\} \tag{2.22}$$

$f(x, t)$ is the solution of the *linearized Boltzmann equation*

$$\frac{\partial}{\partial t}f(q, p, t) = -p\frac{\partial}{\partial q}f(q, p, t) + z\int_+ dp_1\, d\omega\, \omega\cdot(p - p_1)h_\beta(p_1)$$
$$\times\{f(q, p_1', t) + f(q, p', t) - f(q, p_1, t) - f(q, p, t)\} = (Lf(t))(q, p) \tag{2.23}$$

with initial condition $f(q, p)$.

Property 3 is proved by inserting the Ansatz (2.22) in the Boltzmann hierarchy and then by using repeatedly the fact that the collision operator acting on the Maxwellian h_β vanishes.

Properties 2 and 3 remain valid for $\prod_{j=1}^{n}\{zh(x_j)\}$ replaced by $\prod_{j=1}^{n} g(x_j)$, i.e., when the fluid is not in thermal equilibrium. In that case the analogs of A and L are time-dependent through the fluid distribution evolving according to the Boltzmann equation.

It should be understood that properties 1–3 are subject to the conditions of the theorem; in particular, the initial conditions have to satisfy the bound (C1) and the results are valid only up to $t_0(z, \beta)$. However, in contrast to the nonlinear Boltzmann equation, existence and uniqueness of the solutions of

the Rayleigh–Boltzmann equation and the linearized Boltzmann equation in suitable spaces of functions have been proved for all times.[13] In particular, $\{e^{At}|t \geqslant 0\}$ and $\{e^{Lt}|t \geqslant 0\}$ are contraction semigroups on the Hilbert space $\mathscr{H} = L^2(\Lambda \times R^3, h_\beta(p)\, dq\, dp)$.[14]

3. EQUILIBRIUM TIME CORRELATION FUNCTIONS

We consider the fluid of hard spheres of diameter ϵ to be in thermal equilibrium with fugacity z_ϵ and inverse temperature β; grand canonical ensemble.

The Boltzmann–Grad limit corresponds to letting $\epsilon \to 0$ while *increasing* the fugacity as $z_\epsilon = \epsilon^{-2}z$. Since the equilibrium distribution functions have the form

$$\rho^{\varepsilon}_{eq,n}(x_1,\ldots, x_n) = \prod_{j=1}^{n} \{z_\epsilon h_\beta(p_j)\} G_n(q_1,\ldots, q_n, z_\epsilon \epsilon^3) \tag{3.1}$$

with $G_n \to 1$ as $z_\epsilon \epsilon^3 \to 0$ for all $q_1,\ldots, q_n$ in which no two positions coincide, it is clear that as $\epsilon \to 0$ the system will resemble an ideal gas at infinite density. In particular, the rescaled distribution functions converge to

$$\lim_{\varepsilon \to 0} \epsilon^{2n} \rho^{\varepsilon}_{eq,n}(x_1,\ldots, x_n) = \prod_{j=1}^{n} \{z h_\beta(p_j)\} \tag{3.2}$$

uniformly on compact sets of $\Gamma_n(0)$. [As mentioned in the introduction, this limit can be viewed alternatively by considering a spatial scale on which the diameter of a sphere equals one while $q_j' = \epsilon^{-1} q_j$. On this scale the fugacity decreases as ϵz and as $\epsilon \to 0$ the fluid reaches an ideal gas at zero density. To discuss time-dependent properties on this scale we would have to let $t' = \epsilon^{-1} t$.]

We now want to study the *self* and the *total* equilibrium time correlation functions in the low-density limit.

3.1. The Self Correlation Function

We consider a bounded region Λ, but will later drop this restriction. Let $f, g\colon \Lambda \times R^3 \to R$ be bounded and continuous functions of compact support. On $(\Lambda \times R^3)^n$ let us consider the functions $f_j(x_1,\ldots, x_n) = f(x_j)$, $g_j(x_1,\ldots, x_n) = g(x_j)$, $j \leqslant n$. Then the time-dependent self correlation $C(g, f; t)$ is defined as the grand canonical average of $\sum_{j=1}^{n} g(x_j) f_j(x_1,\ldots, x_n, t)$,

$$C(g, f; t) = \left\langle \sum_{j=1}^{n} g_j f_j(t) \right\rangle \tag{3.3}$$

where $f_j(t)$ is f_j time-evolved under the dynamics of n hard spheres in Λ.

We transform (3.3) into a somewhat more manageable form. Let $e^{\epsilon}(q_1,\ldots,q_n)$ be the characteristic function, which is zero whenever $|q_i - q_j| \leqslant \epsilon$, $i \neq j$, or $d(q_i, \partial\Lambda) \leqslant \epsilon/2$, $i, j = 1,\ldots,n$, and which is one otherwise and let $S_n^{\epsilon}(t)$ denote, as before, the time evolution of n hard spheres of diameter ϵ inside Λ. Then, using the symmetry of the equilibrium distributions, we find

$$C(g,f;t) = \int dx_1 f(x_1) \sum_{m=1}^{\infty} \frac{1}{(m-1)!} \int dx_2 \cdots dx_m \times (g_1 \circ S_m^{\epsilon})(-t)(x_1,\ldots,x_m) e_m^{\epsilon}(q_1,\ldots,q_m) \prod_{j=1}^{m} \{z_{\epsilon} h_{\beta}(p_j)\} Z^{-1} \tag{3.4}$$

where Z is the grand canonical partition function. Defining the signed initial distribution functions

$$(\rho_{s,g}^{\varepsilon})_n(x_1,\ldots,x_n) = g(x_1)\rho_{\mathrm{eq},n}^{\varepsilon}(x_1,\ldots,x_n) \tag{3.5}$$

one can rewrite (3.4) as

$$C(g,f;t) = \int dx_1 f(x_1)(V_t^{\epsilon}\rho_{s,g}^{\varepsilon})_1(x_1) \tag{3.6}$$

where we have used the notation (2.3).

One may interpret the quantity $(V_t^{\epsilon}\rho_{s,g}^{\varepsilon})_1(x_1)$ as the test particle distribution function at time t resulting from an initial distribution $g(x_1)\rho_{\mathrm{eq},1}^{\varepsilon}(x_1)$ (cf. Section 4). Strictly speaking, this interpretation is allowed only if $g \geqslant 0$ and if $\int dx_1\, g(x_1)\rho_{\mathrm{eq},1}^{\varepsilon}(x_1) = 1$.

Remark. In (3.6), V_t^{ϵ} depends on the bounded region Λ. To obtain the result for an unbounded Λ, we choose a sequence of bounded regions Λ_m such that $\Lambda_m \to \Lambda$. The infinite-volume limit

$$\lim_{m\to\infty} V_t^{\epsilon}(\Lambda_m) = V_t^{\epsilon}(\Lambda)$$

can then be taken in the perturbation series (2.7) *before* taking the low-density limit $\epsilon \to 0$. The infinite-volume limit causes no difficulty, since all estimates in the proof of the theorem are uniform in Λ. With this prescription in mind, we drop the restriction of Λ being bounded.

Remark. For the computation of transport coefficients one has to consider such quantities as the velocity autocorrelation function $\langle p(t)\cdot p\rangle$ in the infinite-volume limit. In that case one has to show first the existence of

$$\lim_{\Lambda\to R^3} (1/|\Lambda|)C(g_{\Lambda}, f_{\Lambda}; t) = \langle p(t)\cdot p\rangle$$

with $f_{\Lambda}(q,p) = g_{\Lambda}(q,p) = \chi_{\Lambda}(q)p$, where χ_{Λ} is the characteristic function of the bounded region Λ. We have not studied this limit. The subsequent low-

density limit follows by the argument used in the proof of Theorem 3.4. In the low-density limit $\langle p(t) \cdot p \rangle$ is governed by the spatially homogeneous Rayleigh–Boltzmann equation.

To obtain the low-density limit of (3.6), Lanford's theorem has to be applied to $V_t^{\epsilon} \rho_{s,g}^{\epsilon}$. Therefore one has to check the conditions (C1) and (C2). Condition (C2) follows from (3.2) and (C1) from the following result:

Lemma 3.1. If $\sup_x |g(x)| < \infty$, then $\epsilon^{2n}(V_t^{\epsilon} \rho_{s,g}^{\epsilon})_n$ satisfies the bound (C1) for all t.

Proof. For a bounded region Λ we clearly have, by the invariance of $\rho_{\text{eq}}^{\epsilon}$,

$$(V_t^{\epsilon} \rho_{s,g}^{\epsilon})_n(x_1, \ldots, x_n) \leqslant \sup_x |g(x)| \prod_{j=1}^{n} \{h_\beta(p_j)\} \bar{\rho}_{\text{eq},n}^{\epsilon}(q_1, \ldots, q_n) \tag{3.7}$$

Here $\bar{\rho}_{\text{eq},n}^{\epsilon}$ are the spatial parts of the equilibrium distribution functions at fugacity z_ϵ, for which it is known[15] that

$$\bar{\rho}_{\text{eq},n}^{\epsilon}(q_1, \ldots, q_n) \leqslant (z_\epsilon)^n \tag{3.8}$$

independent of Λ. ∎

The fact that $C(g, f; t) \sim \epsilon^{-2}$ can also be seen directly. Consider as a typical example the case where f and g are of the form $\chi_\Delta \phi$ with $\Delta \subset \Lambda$, and ϕ some function of the momentum. Then $C(g, f; t) \sim \langle N \rangle / |\Lambda|$, with $\langle N \rangle$ the average number of particles in Λ, because the probability of finding a given particle initially within Δ is proportional to $1/|\Lambda|$ and the average number of particles contributing to this correlation is $\langle N \rangle$. In the limit as $\epsilon \to 0$, $\langle N \rangle / |\Lambda| \sim z_\epsilon$; hence $C(g, f; t) \sim \epsilon^{-2}$.

Theorem 3.2. Let $f, g \in \mathscr{H} = L^2(\Lambda \times R^3, h_\beta(p)\, dq\, dp)$. Then, for $t \geqslant 0$

$$\lim_{\epsilon \to 0} (z_\epsilon)^{-1} \left\langle \sum_i g_i f_i(t) \right\rangle_{z_\epsilon, \beta} = \int dx\, h_\beta(p) f(x) (e^{At} g)(x) \tag{3.9}$$

Proof. Let f, g be continuous and of compact support. By Lanford's theorem, Property 2, (3.2), and Lemma 3.1,

$$\lim_{\epsilon \to 0} \epsilon^{2n} (V_t^{\epsilon} \rho_{s,g}^{\epsilon})_n(x_1, \ldots, x_n) = (e^{At} g)(x_1) \prod_{j=1}^{n} \{z h_\beta(p_j)\} \tag{3.10}$$

uniformly on compact sets of $\Gamma_n(t)$ for $0 \leqslant t \leqslant t_0(z, \beta)$. At $t = t_0 = t_0(z, \beta)$ the uniform bound (C1) is still valid by Lemma 3.1. Therefore, using (3.10), Lanford's theorem can be applied again to conclude that

$$\lim_{\epsilon \to 0} \epsilon^{2n} (V_t^{\epsilon} (V_{t_0}^{\epsilon} \rho_{s,g}^{\epsilon}))_n(x_1, \ldots, x_n) = (e^{A(t+t_0)} g)(x_1) \prod_{j=1}^{n} \{z h_\beta(p_j)\} \tag{3.11}$$

uniformly on compact sets of $\Gamma_n(t_0(z,\beta)+t)$ for $0 \leqslant t \leqslant t_0(z,\beta)$. Iterating, the result is valid for all times. In particular

$$\lim_{\epsilon\to 0} \epsilon^2 (V_t^\epsilon \rho^\epsilon_{s,g})_1(x_1) = (e^{At}g)(x_1) z h_\beta(p_1) \tag{3.12}$$

uniformly on compact sets for all $t \geqslant 0$, which proves (3.9) for continuous f, g of compact support.

To extend (3.9) to all of $\mathscr{H}$ we use Schwarz's inequality and the invariance of the grand canonical equilibrium probability densities $\{f^\varepsilon_{\mathrm{eq},n} | n \geqslant 0\}$ to show

$$\begin{aligned}
&\left| \left\langle \sum_i g_i f_i(t) \right\rangle_{z_\epsilon,\beta} \right| \\
&\quad \leqslant \sum_{n=0}^{\infty} \sum_{i=1}^{n} \left| \int dx_1 \cdots dx_n f^\varepsilon_{\mathrm{eq},n}(x_1,\ldots,x_n) g(x_i) f_i(x_1,\ldots,x_n,t) \right| \\
&\quad \leqslant \sum_{n=0}^{\infty} n \left[\int dx_1 \cdots dx_n f^\varepsilon_{\mathrm{eq},n}(x_1,\ldots,x_n) g(x_1)^2 \right]^{1/2} \\
&\qquad \times \left[\int dx_1 \cdots dx_n f^\varepsilon_{\mathrm{eq},n}(x_1,\ldots,x_n) f(x_1)^2 \right]^{1/2} \\
&\quad \leqslant \left[\int dx_1 \, \rho^\varepsilon_{\mathrm{eq},1}(x_1) g(x_1)^2 \right]^{1/2} \left[\int dx_1 \, \rho^\epsilon_{\mathrm{eq},1}(x_1) f(x_1)^2 \right]^{1/2}
\end{aligned} \tag{3.13}$$

where we used (2.1) in the last step. Therefore $(z_\epsilon)^{-1} \langle \sum_i g_i f_i(t) \rangle_{z_\epsilon,\beta}$ is a bounded bilinear form on $\mathscr{H}$ and, since continuous functions of compact support are dense in $\mathscr{H}$, (3.9) extends by continuity. ■

3.2. The Total Correlation Function

We proceed with the total equilibrium time correlation functions. Let us define the sum $\sum g$ of one-particle functions g, which are assumed to be continuous and of compact support, as

$$\left(\sum g\right)(x_1,\ldots,x_n) = \sum_{j=1}^{n} g(x_j) \tag{3.14}$$

We define the total correlation functions of f and g as the grand canonical equilibrium average of

$$\left(\sum g\right)(x_1,\ldots,x_n) \left[\left(\sum f\right) \circ S_n^\epsilon(t) \right](x_1,\ldots,x_n) \tag{3.15}$$

In condensed form we write this average as

$$\left\langle \sum g \left(\sum f\right)(t) \right\rangle_{z_\epsilon,\beta} \tag{3.16}$$

It is not difficult to see that in the low-density limit

$$\lim_{\varepsilon\to 0}(z_\epsilon)^{-2}\left\langle \sum g\left(\sum f\right)(t)\right\rangle_{z_\epsilon,\beta} = \int dx_1\, h_\beta(p_1)g(x_1)\int dx_1\, h_\beta(p_1)f(x_1) \tag{3.17}$$

Therefore, a nontrivial result is only obtained upon subtracting out this limit, and the quantity to be considered is

$$\begin{aligned}(z_\epsilon)^{-1}&\left(\left\langle \sum g\left(\sum f\right)(t)\right\rangle_{z_\epsilon,\beta} - \left\langle \sum g\right\rangle_{z_\epsilon,\beta}\left\langle \sum f\right\rangle_{z_\epsilon,\beta}\right)\\ &= (z_\epsilon)^{-1}\left\langle (\delta g)(-t)\left(\sum f\right)\right\rangle_{z_\epsilon,\beta}; \qquad \delta g = \sum g - \left\langle \sum g\right\rangle_{z_\epsilon,\beta}\end{aligned} \tag{3.18}$$

where we have used the time invariance of the equilibrium measure.

We may think of (3.18) as giving the expectation value of $\sum f$ at time t when we start with a signed initial distribution obtained by multiplying the equilibrium density by δg. Equivalently, if at $t = 0$ the distribution functions $\rho_g{}^\epsilon$ are given by

$$\begin{aligned}\rho^\varepsilon_{g,n}(x_1,\ldots,x_n) = &\left[\sum_{j=1}^{n} g(x_j)\right]\rho^\varepsilon_{\mathrm{eq},n}(x_1,\ldots,x_n)\\ &+\int dx_{n+1}\, g(x_{n+1})\{\rho^\varepsilon_{\mathrm{eq},n+1}(x_1,\ldots,x_{n+1})\\ &-\rho^\varepsilon_{\mathrm{eq},n}(x_1,\ldots,x_n)\rho^\varepsilon_{\mathrm{eq},1}(x_{n+1})\}\end{aligned} \tag{3.19}$$

then

$$\left\langle (\delta g)(-t)\left(\sum f\right)\right\rangle_{z_\epsilon,\beta} = \int dx_1 f(x_1)(V_t^\epsilon \rho_g{}^\epsilon)_1(x_1) \tag{3.20}$$

To apply Lanford's theorem, the conditions (C1) and (C2) have to be verified for $\rho_g{}^\epsilon$. By (3.2) and (3.8) the first term in (3.19) clearly causes no problem. The second term is estimated by

$$\begin{aligned}&\left|\int dx_{n+1}\, g(x_{n+1})\{\rho^\varepsilon_{\mathrm{eq},n+1}(x_1,\ldots,x_{n+1}) - \rho^\varepsilon_{\mathrm{eq},n}(x_1,\ldots,x_n)\rho^\varepsilon_{\mathrm{eq},1}(x_{n+1})\}\right|\\ &\times \leqslant \sup_x|g(x)|\prod_{j=1}^{n}\{h_\beta(p_j)\}\\ &\qquad\times\int_\Lambda dq\,|\bar\rho^\varepsilon_{\mathrm{eq},n+1}(q,q_1,\ldots,q_n) - \bar\rho^\varepsilon_{\mathrm{eq},1}(q)\bar\rho^\varepsilon_{\mathrm{eq},n}(q_1,\ldots,q_n)|\end{aligned} \tag{3.21}$$

Then the uniform bound (C1) follows from the next result:

Lemma 3.3. Let $z' > ez$. Then there exists a constant $c > 0$ and

$\epsilon(z') > 0$ such that

$$\sup_{q_1,\ldots,q_n\in\Lambda} \epsilon^{2n} \int dq\, |\bar{\rho}^{\epsilon}_{\mathrm{eq},n+1}(q, q_1,\ldots, q_n) - \bar{\rho}^{\epsilon}_{\mathrm{eq},1}(q)\bar{\rho}^{\epsilon}_{\mathrm{eq},n}(q_1,\ldots, q_n)| \leqslant \epsilon c(z')^n \tag{3.22}$$

for all $\epsilon \leqslant \epsilon(z')$ independent of Λ.

Proof. Cf. Appendix B.

It is now easy to prove the following result:

Theorem 3.4. Let $f, g \in \mathscr{H}$. Then for $0 \leqslant t \leqslant t_0(ez, \beta)$

$$\lim_{\epsilon\to 0}(z_\epsilon)^{-1}\left[\left\langle\left(\sum g\right)\left(\sum f\right)(t)\right\rangle_{z_\epsilon,\beta} - \left\langle\sum g\right\rangle_{z_\epsilon,\beta}\left\langle\sum f\right\rangle_{z_\epsilon,\beta}\right] = \int dx\, h_\beta(p) f(x)(e^{Lt}g)(x) \tag{3.23}$$

Proof. Let f, g be continuous and of compact support. By Lemma 3.3, $\epsilon^{2n}\rho^{\epsilon}_{g,n}$ satisfies the uniform bound (C1) for the pair (ez, β). By (3.2), (3.19), and Lemma 3.3

$$\lim_{\epsilon\to 0} \epsilon^{2n}\rho^{\epsilon}_{g,n}(x_1,\ldots, x_n) = \left[\sum_{j=1}^{n} g(x_j)\right]\prod_{j=1}^{n}\{zh_\beta(p_j)\} \tag{3.24}$$

uniformly on compact sets of $\Gamma_n(0)$. Therefore by Lanford's theorem and by Property 3 [Eqs. (2.21) and (2.22)]

$$\lim_{\epsilon\to 0} \epsilon^{2n}(V_t^{\epsilon}\rho_g^{\epsilon})_n(x_1,\ldots, x_n) = \sum_{j=1}^{n}(e^{Lt}g)(x_j)\prod_{j=1}^{n}\{zh_\beta(x_j)\} \tag{3.25}$$

uniformly on compact sets of $\Gamma_n(t)$ for $0 \leqslant t \leqslant t_0(ez, \beta)$. Hence it follows from (3.20) that the left-hand side of (3.23) converges to $\int dx\, h_\beta(x)f(x)(e^{Lt})(x)$.

To extend (3.23) to all of $\mathscr{H}$, we use again Schwarz' inequality

$$\left|\left\langle\left(\sum g - \left\langle\sum g\right\rangle_{z_\epsilon,\beta}\right)\left[\left(\sum f\right)(t) - \left\langle\sum f\right\rangle_{z_\epsilon,\beta}\right]\right\rangle_{z_\epsilon,\beta}\right| \leqslant \left\langle\left(\sum g - \left\langle\sum g\right\rangle_{z_\epsilon,\beta}\right)^2\right\rangle^{1/2}_{z_\epsilon,\beta}\left\langle\left(\sum f - \left\langle\sum f\right\rangle_{z_\epsilon,\beta}\right)^2\right\rangle^{1/2}_{z_\epsilon,\beta} \tag{3.26}$$

Therefore for ϵ small enough

$$(z_\epsilon)^{-1}\left\langle\left(\sum g - \left\langle\sum g\right\rangle_{z_\epsilon,\beta}\right)\left[\left(\sum f\right)(t) - \left\langle\sum f\right\rangle_{z_\epsilon,\beta}\right]\right\rangle_{z_\epsilon,\beta}$$

is a bounded bilinear form on $\mathscr{H}$ and (3.23) follows by continuity.

Remark. Although $e^{Lt}g$ is known to exist for all $t \geqslant 0$, we have been unable to extend Theorem 3.4 beyond $t_0(ez, \beta)$.

Remark. The result of Theorem 3.4 can be viewed in a somewhat different way, which we feel to be rather instructive. Let us define the random variables

$$X_f^{\epsilon} = \sum f \tag{3.27}$$

on the phase space equipped with the equilibrium measure at fugacity z_ϵ and inverse temperature β. For the particular choice $f = \chi_\Delta$, X_f^{ϵ} is the number of particles in the region $\Delta \subset \Lambda \times R^3$. A straightforward equilibrium estimate shows that

$$\begin{aligned} \lim_{\epsilon \to 0} \langle \epsilon^2 X_f^{\epsilon} \rangle_{z_\epsilon, \beta} &= \int dx\, z h_\beta(p) f(x) \\ \lim_{\epsilon \to 0} \langle (\epsilon^2 X_f^{\epsilon})^2 \rangle_{z_\epsilon, \beta} - \langle \epsilon^2 X_f^{\epsilon} \rangle_{z_\epsilon, \beta}^2 &= 0 \end{aligned} \tag{3.28}$$

for all $f \in \mathscr{H}$. This means that the distribution of $\epsilon^2 X_f$ converges to a δ-function as $\epsilon \to 0$. In particular, the relative number of particles in Δ has a sharp value in this limit.

Let us now consider the fluctuations of X_f around its average value, i.e., the *fluctuation observables*

$$\xi_f^{\epsilon} = \epsilon(X_f^{\epsilon} - \langle X_f^{\epsilon} \rangle_{z_\epsilon, \beta}) \tag{3.29}$$

and also their time evolution $\xi_f^{\epsilon}(t)$. One expects and can prove[16] that $\xi_f^{\epsilon}(t)$ has a Gaussian distribution as $\epsilon \to 0$ with mean zero and variance $z \int dx\, h_\beta(p) f(x)^2$. In other words, the central limit theorem holds for the sequence of random variables $\xi_f^{\epsilon}(t)$. But one also expects that $\{\xi_f^{\epsilon}(t) | t \in R, f \in \mathscr{H}\}$ become *jointly* Gaussian. Now Theorem 3.4 tells us that at least their covariance exists in the limit $\epsilon \to 0$ for short times, i.e.,

$$\lim_{\epsilon \to 0} \langle \xi_f^{\epsilon}(t) \xi_g^{\epsilon}(s) \rangle_{z_\epsilon, \beta} = \int dx\, z h_\beta(p) f(x) (e^{L(t-s)} g)(x) \tag{3.30}$$

for $t \geqslant s$, $t - s \leqslant t_0(ez, \beta)$. So we conjecture that $\{\xi_f^{\epsilon}(t) | t \in R, f \in \mathscr{H}\}$ converges as $\epsilon \to 0$ to a Gaussian stochastic process indexed by $\mathscr{H}$ with mean zero and covariance (3.30).

4. A TAGGED PARTICLE IN AN EQUILIBRIUM HARD-SPHERE FLUID

As is well known, the self time correlation function can equally well be interpreted as describing the distribution of a tagged particle in a fluid. Thinking of this tagged particle as an external probe of the fluid, it is then of interest to consider also the response of the fluid to the perturbation caused by

the test particle. This in turn is related to the total time correlation function. But there are some new insights to be gained by looking at the problem from this point of view.

We consider a tagged particle in a fluid of hard spheres of diameter ϵ and mass one. The tagged particle is assumed to have the same properties. (We could allow the tagged particle to have a different mass and diameter. However, it is necessary that its diameter also decrease in proportion to ϵ.) The fluid plus tagged particle is enclosed in the region Λ. Initially, the fluid is assumed to be in thermal equilibrium at fugacity $z_\epsilon = \epsilon^{-2}z$ and inverse temperature β *conditioned* on the tagged particle being located at q_1 while the tagged particle has the distribution $f(x_1)\,dx_1$. Here $x_1 = (q_1, p_1)$ stands for the position and momentum of the tagged particle and $x_i = (q_i, p_i)$, $i \geqslant 2$, stands for the position and momentum of the $(i-1)$th fluid particle. Therefore the initial probability density of the joint system is proportional to

$$\left\{\frac{1}{n!} f(x_1)[\rho^\varepsilon_{\mathrm{eq},1}(x_1)]^{-1} e^\varepsilon_{n+1}(q_1,\ldots,q_{n+1}) \prod_{j=1}^{n+1} \{z_\epsilon h_\beta(p_j)\} \middle| n \geqslant 0\right\} \tag{4.1}$$

with $f(x_1) \geqslant 0$ and $\int dx_1 f(x_1) = 1$; $e^\varepsilon_n = 1$ if $|q_i - q_j| \geqslant \varepsilon$, zero otherwise.

We want to study the time evolution of the distribution functions of this system for $f(x_1) \leqslant c h_\beta(x_1)$. A straightforward computation shows that at $t = 0$ these distribution functions are

$$[\rho^\varepsilon_{\mathrm{eq},1}(x_1)]^{-1} f(x_1) \rho^\varepsilon_{\mathrm{eq},n+1}(x_1,\ldots,x_{n+1}) = (\rho^\varepsilon_{s,g})_{n+1}(x_1,\ldots,x_{n+1}) \tag{4.2}$$

with $g(x_1) = [\rho^\varepsilon_{\mathrm{eq},1}(x_1)]^{-1} f(x_1)$ in the notation (3.5). As before, $\rho^\varepsilon_{\mathrm{eq},n}$ are the unconditioned equilibrium distribution functions at fugacity z_ϵ and inverse temperature β. Since

$$\lim_{\epsilon \to 0} \epsilon^{-2}[\rho^\varepsilon_{\mathrm{eq},1}(x_1)]^{-1} = [z_\epsilon h_\beta(p_1)]^{-1} \tag{4.3}$$

we conclude from Property 2, (3.2), Lemma 3.1, and the iteration argument used in the proof of Theorem 3.2 that the following result holds:

Theorem 4.1. Let $(\rho^\varepsilon_{s,g})_{n+1}(x_1,\ldots,x_{n+1}, t)$ denote the time-evolved distribution functions of the fluid plus tagged particle system with initial distribution functions given by (4.2). Then for all $t \geqslant 0$

$$\lim_{\epsilon \to 0} \epsilon^{2n}(\rho^\varepsilon_{s,g})_{n+1}(x_1,\ldots,x_{n+1}, t) = (e^{At} f z^{-1} h_\beta^{-1})(x_1) \prod_{j=1}^{n+1} z h_\beta(p_j) \tag{4.4}$$

uniformly on compact sets of $\Gamma_{n+1}(t)$.

Integrating $(\rho^\varepsilon_{s,g})_{n+1}(x_1,\ldots,x_{n+1}, t)$ over x_1 yields the fluid distribution functions $(\rho^\varepsilon_{fl,g})_n(x_2,\ldots,x_{n+1}, t)$, which give the expectation of finding n fluid

particles at $x_2,\ldots, x_{n+1}$. From Theorem 4.1, by integrating over x_1, one obtains that in the low-density limit

$$\lim_{\epsilon\to 0} \epsilon^{2n}(\rho^{\epsilon}_{fl,g})_n(x_1,\ldots, x_n, t) = \prod_{j=1}^{n} \{zh_\beta(p_j)\} \tag{4.5}$$

In the limit the fluid is completely undisturbed by the presence of the tagged particle. This is of course to be expected, since in this limit the tagged particle will interact (directly or indirectly) during any fixed time interval only with a vanishing fraction of all particles in any fixed region. Consider now, however, the next order *correction*, i.e., the limiting behavior of

$$(\delta\rho^{\epsilon}_{fl,g})_n(x_1,\ldots, x_n, t) = \epsilon^{2n-2}\{(\rho^{\epsilon}_{fl,g})_n(x_1,\ldots, x_n, t) - \rho^{\epsilon}_{\mathrm{eq},n}(x_1,\ldots, x_n)\} \tag{4.6}$$

Theorem 4.2. For $0 \leqslant t \leqslant t_0(ez, \beta)$

$$\lim_{\epsilon\to 0}(\delta\rho^{\epsilon}_{fl,g})_n(x_1,\ldots, x_n, t) = \left\{\sum_{j=1}^{n} ([e^{Lt} - e^{At}]fz^{-1}h_\beta^{-1})(x_j)\right\}\prod_{j=1}^{n}\{zh_\beta(p_j)\} \tag{4.7}$$

uniformly on compact sets of $\Gamma_n(t)$.

Proof. With $g(x) = [\rho^{\epsilon}_{\mathrm{eq},1}(x)]^{-1}f(x)$ and the notations (3.5), (3.19), and (4.6) one checks the identity

$$\rho^{\epsilon}_{g,n}(x_1,\ldots, x_n) = \sum_{j=1}^{n} (\rho^{\epsilon}_{s,g})_n(x_j, x_1,\ldots, x_{j-1},\ldots, x_n) + (\delta\rho^{\epsilon}_{fl,g})_n(x_1,\ldots, x_n) \tag{4.8}$$

The assertion now follows from Theorem 4.1 and the proof of Theorem 3.4. ■

APPENDIX A

We wish to illustrate here by means of an example how the Lanford theorem, Eq. (2.14), can manage to get the irreversible Boltzmann hierarchy from the reversible BBGKY hierarchy.

For the sake of clarity let us introduce some notation. We denote the velocity reversal operator by R,

$$(R\rho)_n(q_1, p_1,\ldots, q_n, p_n) = \rho_n(q_1, -p_1,\ldots, q_n, -p_n) \tag{A1}$$

As before, V_t^{ϵ} denotes the solution operator of the BBGKY hierarchy and V_t^0 denotes the solution operator of the Boltzmann hierarchy. [We remind the reader that for $t \leqslant 0$, $V_t^0 r$ is defined as the solution of (2.9) with the sign of the collision term reversed.]

Let us now consider a situation in which the box Λ is divided into two parts Λ_1 and Λ_2 and the initial state, at $t = 0$, corresponds to a canonical

equilibrium state of N particles of diameter ϵ all in Λ_1. (We can imagine that there was an impenetrable barrier between Λ_1 and Λ_2 which was removed at $t = 0$.) It is clear that, since the initial state is invariant to reversal of velocities, its distribution functions $\rho^\epsilon = (\rho_1{}^\epsilon, \rho_2{}^\epsilon, \ldots)$ satisfy the equality

$$V_t{}^\epsilon \rho^\epsilon = R V^\varepsilon_{-t} \rho^\epsilon \tag{A2}$$

Furthermore,

$$V_t{}^\epsilon (R V_t{}^\epsilon \rho^\epsilon) = \rho^\epsilon \tag{A3}$$

while

$$V_t{}^\epsilon (V_t{}^\epsilon \rho^\epsilon) = V^\varepsilon_{2t} \rho^\epsilon \tag{A4}$$

This means that if at time t we reverse all velocities, then the system, after another time interval t, will return to its initial state in which all the particles are in Λ_1.

Consider now the sequence of initial states with distribution functions ρ^ϵ in which as $\epsilon \to 0$ the number of particles inside Λ_1 increases with fixed $N\epsilon^2 = z$. Then

$$\lim_{\epsilon \to 0} \epsilon^{2n} \rho_n{}^\epsilon(x_1, \ldots, x_n) = \lim_{\epsilon \to 0} r_n{}^\epsilon(x_1, \ldots, x_n) = r_n(x_1, \ldots, x_n)$$

$$= \prod_{j=1}^{n} \{\chi_{\Lambda_1}(q_j) z h_\beta(p_j)\} \tag{A5}$$

on $\Gamma_n(0)$, where χ_{Λ_1} is the characteristic function of the set Λ_1, and, since (C1) and (C2) are satisfied, by Lanford's theorem

$$\lim_{\epsilon \to 0} \epsilon^{2n} (V_t{}^\epsilon \rho^\epsilon)_n(x_1, \ldots, x_n) = (V_t{}^0 r)_n(x_1, \ldots, x_n) = \prod_{j=1}^{n} \{f(x_j, t)\} \tag{A6}$$

on $\Gamma_n(t)$ for $|t| < t_0(z, \beta)$, where $f(x, t)$ is the solution of the Boltzmann equation with initial conditions $f(q, p) = \chi_{\Lambda_1}(q) z h_\beta(p)$.

Let us now reverse the velocities at time t, $0 \leqslant |t| \leqslant t_0/2$, and let us consider $R V_t{}^\epsilon \rho^\epsilon$ as the new initial state. Clearly

$$V_t{}^0 (R V_t{}^0 r) \neq r \tag{A7}$$

in contrast to (A3), so the limiting r do not have the time reversibility of the r^ϵ. Indeed, the Boltzmann H-function decreases up to t, remains unchanged by R, and continues to decrease as $R V_t{}^0 r$ is evolved for a time interval t.

At first sight, Lanford's theorem seems to assert that

$$\lim_{\varepsilon \to 0} \epsilon^{2n} (V_t{}^\epsilon (R V_t{}^\epsilon \rho^\epsilon))_n = (V_t{}^0 (R V_t{}^0 r))_n \neq r_n$$

However, there is no such contradiction. The answer lies in the fact that while

$$\lim_{\varepsilon \to 0} \epsilon^{2n} (V_t{}^\epsilon \rho^\epsilon)_n = (V_t{}^0 r)_n \quad \text{on } \Gamma_n(t) \tag{A8}$$

it is also true that

$$\lim_{\epsilon \to 0} \epsilon^{2n}(RV_t^{\epsilon}\rho^{\epsilon})_n = (RV_t^{0}r)_n = (V_t^{0}r)_n \quad \text{on } \Gamma_n(-t) \neq \Gamma_n(t)$$

Therefore, continuing in the same time direction as before the reversal of velocities, $RV_t^{\epsilon}\rho^{\epsilon}$ no longer satisfies the second condition (C2) of Lanford's theorem. The theorem asserts nothing about the convergence of $\epsilon^{2n}(V_t^{\epsilon}(RV_t^{\epsilon}\rho^{\epsilon}))_n$ as $\epsilon \to 0$. [Of course, by (A3), we can say something about this limit. The point is that we cannot conclude from Lanford's theorem that the limit is $(V_t^{0}(RV_t^{0}r))_n$, since condition (C2) is violated.] For the theorem still to be applicable at time t one has only the two choices to consider, either $V_t^{\epsilon}(V_t^{\epsilon}\rho^{\epsilon})$ or $V_{-t}^{\epsilon}(RV_t^{\epsilon}\rho^{\epsilon})$. In both cases the system evolves further toward equilibrium.

The irreversible Boltzmann hierarchy is consistent with the reversible BBGKY hierarchy, since the approximation by the Boltzmann hierarchy is valid only for a *particular class* of initial states. The condition (C2) excludes highly correlated initial states such as the one just constructed by reversal of velocities.

APPENDIX B. PROOF OF LEMMA 3.3

Relation (3.22) is transformed to a spatial scale on which a sphere has diameter one. Then

$$\begin{aligned}
&\sup_{q_1,\cdots,q_n \in \Lambda} \epsilon^{2n} \int_{\Lambda} dq \, |\bar{\rho}_{\mathrm{eq},n+1}^{\epsilon}(q, q_1, \ldots, q_n) - \bar{\rho}_{\mathrm{eq},1}^{\epsilon}(q)\bar{\rho}_{\mathrm{eq},n}^{\epsilon}(q_1, \ldots, q_n)| \\
&\quad = \sup_{q_1,\cdots,q_n \in \epsilon\varepsilon^{-1}\Lambda} \epsilon \int_{\epsilon^{-1}\Lambda} dq \, \epsilon^{-(n+1)} \\
&\quad \times |\rho_{n+1}^{\epsilon z}(q, q_1, \ldots, q_n; \epsilon^{-1}\Lambda) - \rho_1^{\epsilon z}(q; \epsilon^{-1}\Lambda)\rho_n^{\epsilon z}(q_1, \ldots, q_n; \epsilon^{-1}\Lambda)|
\end{aligned} \tag{B1}$$

Here $\rho_n^{\epsilon z}(q_1, \ldots, q_n; \epsilon^{-1}\Lambda)$ represents the spatial part of the grand canonical equilibrium distribution functions of hard spheres of diameter one inside the region $\epsilon^{-1}\Lambda$ at fugacity $\epsilon^3 z_{\epsilon} = \epsilon z$. Lemma 3.3 follows now from the result:

Lemma B1. There exists an $\epsilon_0 > 0$ such that for all $\epsilon < \epsilon_0$

$$\begin{aligned}
&\sup_{q_1,\cdots,q_n \in \epsilon\varepsilon^{-1}\Lambda} \epsilon^{-(n+1)} \int_{\epsilon^{-1}\Lambda} dq \, |\rho_{n+1}^{\epsilon z}(q, q_1, \ldots, q_n; \epsilon^{-1}\Lambda) \\
&\quad - \rho_1^{\epsilon z}(q; \epsilon^{-1}\Lambda)\rho_n^{\epsilon z}(q_1, \ldots, q_n; \epsilon^{-1}\Lambda)| \leqslant M(z')^n
\end{aligned} \tag{B2}$$

where M is a constant and $z' > ez$.

Proof. We consider the particles at $q_1, \ldots, q_n$ as providing an external field and denote the equilibrium distribution functions of this system by

$\rho^z(\cdot\,|q_1,\ldots,q_n;\,\Lambda)$. Then, expanding in z,

$$\rho_1{}^z(q|q_1,\ldots,q_n;\,\Lambda) - \rho_1{}^z(q;\,\Lambda) = \sum_{j=0}^{\infty} c_j(q|q_1,\ldots,q_n)z^{j+1} \tag{B3}$$

In terms of the zero-field Ursell functions U_{j+1} the expansion coefficients are

$$c_j(q|q_1,\ldots,q_n) = \frac{1}{j!}\int dq_1' \cdots dq_j'\, U_{j+1}(q, q_1',\ldots,q_j')\Bigg\{\prod_{k=1}^{n}\Bigg[h(q-q_k) \times \prod_{i=1}^{j} h(q_i'-q_k)\Bigg] - 1\Bigg\} \tag{B4}$$

where, we let h be the overlap function, $h(q) = 0$ for $|q| \leqslant 1$ and $h(q) = 1$ otherwise. The second factor is negative and, according to Ref. 15, Chapter 4, (5.14), for a positive pair potential $(-1)^{j+1}U_j \geqslant 0$. Therefore

$$(-1)^{j+1}c_j(q|q_1,\ldots,q_n) \geqslant 0 \tag{B5}$$

and for $z > 0$

$$\begin{aligned}|\rho_1{}^z(q|q_1,\ldots,q_n;\,\Lambda) - \rho_1{}^z(q;\,\Lambda)| &\leqslant \sum_{j=0}^{\infty} |c_j(q|q_1,\ldots,q_n)|z^{j+1} \\ &= \sum_{j=0}^{\infty} c_j(q|q_1,\ldots,q_n)(-1)^{j+1}z^{j+1} \\ &= \rho_1^{-z}(q|q_1,\ldots,q_n;\,\Lambda) - \rho_1^{-z}(q;\,\Lambda)\end{aligned} \tag{B6}$$

For small enough ϵ, $\rho_n^{\epsilon z}(q_1,\ldots,q_n;\,\epsilon^{-1}\Lambda) \neq 0$ and therefore we have for $z > 0$

$$\begin{aligned}&\int_{\epsilon^{-1}\Lambda} dq\, |\rho_{n+1}^{\epsilon z}(q, q_1,\ldots,q_n;\,\epsilon^{-1}\Lambda) - \rho_1^{\epsilon z}(q;\,\epsilon^{-1}\Lambda)\rho_n^{\epsilon z}(q_1,\ldots,q_n;\,\epsilon^{-1}\Lambda)| \\ &\leqslant \rho_n^{\epsilon z}(q_1,\ldots,q_n;\,\epsilon^{-1}\Lambda)\int_{\epsilon^{-1}\Lambda} dq\, |\rho_1^{\epsilon z}(q|q_1,\ldots,q_n;\,\epsilon^{-1}\Lambda) - \rho_1^{\epsilon z}(q;\,\epsilon^{-1}\Lambda)| \\ &\leqslant \rho_n^{\epsilon z}(q_1,\ldots,q_n;\,\epsilon^{-1}\Lambda)\int_{\epsilon^{-1}\Lambda} dq\, [\rho_1^{-\epsilon z}(q|q_1,\ldots,q_n;\,\epsilon^{-1}\Lambda) - \rho_1^{-\epsilon z}(q;\,\epsilon^{-1}\Lambda)] \\ &= \rho_n^{\epsilon z}(q_1,\ldots,q_n;\,\epsilon^{-1}\Lambda)[\rho_n^{-\epsilon z}(q_1,\ldots,q_n;\,\epsilon^{-1}\Lambda)]^{-1} \\ &\quad \times \int_{\epsilon^{-1}\Lambda} dq\{\rho_{n+1}^{-\epsilon z}(q, q_1,\ldots,q_n;\,\epsilon^{-1}\Lambda) - \rho_1^{-\epsilon z}(q;\,\epsilon^{-1}\Lambda)\rho_n^{-\epsilon z}(q_1,\ldots,q_n;\,\epsilon^{-1}\Lambda)\} \\ &= \{\rho_n^{\epsilon z}(q_1,\ldots,q_n;\,\epsilon^{-1}\Lambda)[\rho_n^{-\epsilon z}(q_1,\ldots,q_n;\,\epsilon^{-1}\Lambda)]^{-1}\} \\ &\quad \times \{-n\rho_n^{-\epsilon z}(q_1,\ldots,q_n;\,\epsilon^{-1}\Lambda) + z\frac{d}{dz}\rho_n^{-\epsilon z}(q_1,\ldots,q_n;\,\epsilon^{-1}\Lambda)\}\end{aligned} \tag{B7}$$

(B7) is estimated using the Mayer expansion. For small enough ϵ the first factor is uniformly bounded. The second factor is bounded by

$$\left| -n\rho_n^{-\epsilon z}(q_1,\ldots,q_n;\epsilon^{-1}\Lambda) + z\frac{d}{dz}\rho_n^{-\epsilon z}(q_1,\ldots,q_n;\epsilon^{-1}\Lambda) \right|$$

$$\leqslant \sum_{m=0}^{\infty} |b_{n,m}(q_1,\ldots,q_n;\epsilon^{-1}\Lambda)|\,|-n+n+m|\,|\epsilon z|^{n+m}$$

$$\leqslant \sum_{m=1}^{\infty} n(n+m)^{m-1}m\left(\frac{4\pi}{3}\right)^m \frac{1}{m!}|\epsilon z|^{n+m} < M\frac{1}{1-4\pi|\epsilon z|/3}ne^n|\epsilon z|^{n+1} \tag{B8}$$

where we have used the uniform bound on the coefficients $b_{n,m}(q_1,\ldots,q_n;\epsilon^{-1}\Lambda)$ [cf. Ref. 15, Chapter 4, (4.30)]. Relation (B8) together with (B7) proves the lemma. ■

Remark. Lemma B1 holds for any positive pair potential V with $\int dq\,(1-e^{-\beta V(q)}) < \infty$.

REFERENCES

1. H. Grad, Principles of the Kinetic Theory of Gases, in *Handbuch der Physik*, S. Flügge, ed. (Springer, Berlin, 1958), Vol. 12.
2. O. E. Lanford, Time Evolution of Large Classical Systems, in *Dynamical Systems, Theory and Applications*, J. Moser, ed. (Lecture Notes in Physics No. 38, Springer, Berlin, 1975), p. 1.
3. O. E. Lanford, On the Derivation of the Boltzmann Equation. *Soc. Math. de France. Astérisque* **40**:117 (1976).
4. J. L. Lebowitz and J. Percus, *Phys. Rev.* **155**:122 (1967).
5. J. L. Lebowitz, J. Percus, and J. Sykes, *Phys. Rev.* **188**:487 (1969).
6. P. Resibois and J. L. Lebowitz, *J. Stat. Phys.* **12**:483 (1975).
7. W. Braun and K. Hepp, *Comm. Math. Phys.* **56**:101 (1977).
8. K. Hepp and E. H. Lieb, *Helv. Phys. Acta* **46**, 573 (1973).
9. F. King, Ph.D. Thesis, Department of Mathematics, University of California at Berkeley (1975).
10. C. Cercignani, *Transport Theory and Statistical Physics* **2**:211 (1972).
11. O. E. Lanford, Notes of the Lectures at the Troisième Cycle, Lausanne (1978), unpublished.
12. R. K. Alexander, Ph.D. Thesis, Department of Mathematics, University of California at Berkeley (1975).
13. C. Cercignani, *The Boltzmann Equation* (Elsevier, New York, 1976).
14. M. Klaus, *Helv. Phys. Acta* **48**:99 (1975).
15. D. Ruelle, *Statistical Mechanics: Rigorous Results* (Benjamin, Reading, Mass., 1969).
16. R. L. Dobrushin and B. Tirozzi, *Comm. Math. Phys.* **54**:173 (1977).

Physica 7D (1983) 124–125
North-Holland Publishing Company

PERIOD DOUBLING IN ONE AND SEVERAL DIMENSIONS

Oscar E. LANFORD III*†
Institute for Mathematics and its Applications, University of Minnesota, Minneapolis MN 55455, USA

Feigenbaum cascade—infinite sequences of successive period doublings–form a route from periodic to aperiodic behavior of dynamical systems. These sequences of bifurcations exhibit some striking universal features. The simplest of these features to formulate concerns the rate of accumulation of the bifurcations: If μ_n denotes the parameter value at which the nth doubling occurs, then, asymptotically,

$$\mu_n = \mu_\infty - c(4.6692\ldots)^{-n} + \text{"higher order terms"}.$$

The rate 4.6692... appears to be universal, i.e., it shows up in many apparently unrelated systems such as
–one-dimensional non-invertible mappings, such as the one-parameter family $x \to 1 - \mu x^2$ on $[-1, 1]$, where $0 < \mu < 2$;
–dissipative (volume-decreasing) invertible mappings such as the Hénon system (see below);
–dissipative differential equations, such as the Lorenz system and the five-component truncation of the two-dimensional Navier–Stokes equations studied by Franceschini et al.

The main point we want to make here is that, despite their apparent diversity, these are really all instances of *precisely the same* mathematical phenomenon, and can be understood relatively easily once one has understood period doubling for one-dimensional mappings. (There is, on the other hand, another period doubling process, occurring for area-preserving mappings of the plane, which, although analogous to dissipative period doubling, seems to be an independent mathematical phenomenon. See Eckmann, Koch, and Wittwer [1] and the references cited therein.)

To undergo dissipative period doubling, a family of mappings–or, more generally a restriction of some iterate of the mappings–must have a characteristic behavior illustrated by the Hénon family

$$(x, y) \to (1 - \mu x^2 + cy, -cx),$$

with c small. (We are restricting ourselves here, for definiteness, to the orientation-preserving case, and have called the parameters μ and $-c^2$ instead of the more traditional a and b. We think of the bifurcations as occurring as we vary μ with c held fixed.) These mappings can be visualized as acting by:
1) contracting vertically:

$$(x, y) \to (x, cy);$$

2) bending and stretching by an x-dependent vertical shift:

$$\to (x, 1 - \mu x^2 + cy);$$

3) rotating a quarter-turn clockwise:

$$\to (1 - \mu x^2 + cy, -x);$$

4) contracting again vertically:

$$\to (1 - \mu x^2 + cy, -cx).$$

The general feature we want to emphasize is that the mapping contracts its multi-dimensional domain to an almost one-dimensional one, then folds that approximately one-dimensional set back into the original domain. Furthermore, the region of strong

*Current address: IHES, 91440 Bures-sur-Yvette, France..
†Work supported in part by NSF grant MCS81-07086AO1

0167-2789/83/0000–0000/$03.00

folding–a vertical strip about the x-axis in the above example–is mapped away from itself and into a region of gentle folding. Finally, it is characteristic of mappings undergoing period doubling that the strong-folding region, although mapped away from itself by a first application of the mapping, is sent back into itself by a second.

The key to understanding repeated period doubling is the introduction of a *renormalization* or *doubling* operator $\mathcal{T}$ which carries a mapping F to one obtained by
–composing F with itself;
–restricting to an appropriate subdomain;
–making a change of coordinates to magnify the subdomain up to the original domain.

Roughly speaking, applying $\mathcal{T}$ divides the periods of all cycles by two but preserves their stability properties.

The idea now is to apply the renormalization group program to $\mathcal{T}$. To account for the observed universality, what one needs to show is that $\mathcal{T}$ has a fixed point and that, in the neighborhood of the fixed point, $\mathcal{T}$ is expanding in one direction and contracting in all others (i.e., that the linearization of $\mathcal{T}$ at the fixed point has a single simple eigenvalue with modulus greater than one and that the remainder of its spectrum is strictly inside the unit circle.) These facts have been established for one-dimensional mappings [2]; the proof rests on complicated numerical estimates verified (rigorously) by computer. Up to now, no one has succeeded in giving a conceptual proof.

By contrast, the theory of multi-dimensional period doubling can be reduced to the one-dimensional theory by a relatively simple conceptual argument. The argument goes roughly as follows: The space of one-dimensional mappings may be imbedded in the space of multi-dimensional mappings by associating with the one-dimensional mapping f the multi-dimensional mapping

$$F_0\colon (x, y)\to(f(x), 0).$$

(Here, y may have any number of components.) Such as F_0 is of course not invertible, but an arbitrarily small perturbation on F_0 can give an invertible mapping; the Hénon mapping with c small is an example. We can think of the space of F_0's as a surface M_0 in the space of all F's. What is now done is to construct a multi-dimensional doubling operator which
1) maps M_0 into itself;
2) agrees with the ordinary one-dimensional doubling operator on M_0;
3) is contractive in the directions transverse to M_0, i.e., when applied to an F near but not on M_0, gives a new mapping which is still closer to M_0.

In order to get 3) to hold, it is necessary to choose the change of variables in the construction of the doubling operator with some care.

A multi-dimensional doubling operator satisfying 1)–3) has as a fixed point the mapping

$$(x, y)\to(g(x), 0),$$

where g is the fixed point for the one-dimensional operator. Contractivity in directions transverse to M_0 guarantees that allowing the operator to act on mappings which are not strictly one-dimensional does not introduce any new expanding directions.

An analysis similar to the one described above was first given by Collet, Eckmann, and Koch [3]. In precisely this form, it is unpublished work of the author.

References

[1] J.-P. Eckmann, H. Koch and P. Wittwer, A computer-assisted proof of universality for area-preserving maps, Université de Genève preprint UGVA-DPT 1981/04-345, to appear in Memoirs A.M.S.

[2] O.E. Lanford, A computer-assisted proof of the Feigenbaum conjectures, Bull. A.M.S. (New Series) 6 (1982) 427–434.

[3] P. Collet, J.-P. Eckmann, and H. Koch, Period-doubling bifurcations for families of maps on $\mathbb{R}^n$, J. Stat. Phys. 25 (1981) 1–14.

Commun. Math. Phys. 76, 211–254 (1980)

Communications in Mathematical Physics

Universal Properties of Maps on an Interval

P. Collet[1], J.-P. Eckmann[2], and O. E. Lanford III[3]

[1] Harvard University, Cambridge, MA 02138, USA
[2] Département de Physique Théorique, Université de Genève, CH-1211 Genève 4, Switzerland
[3] University of California, Berkeley, CA 94720, USA

Abstract. We consider iterates of maps of an interval to itself and their stable periodic orbits. When these maps depend on a parameter, one can observe period doubling bifurcations as the parameter is varied. We investigate rigorously those aspects of these bifurcations which are universal, i.e. independent of the choice of a particular one-parameter family. We point out that this universality extends to many other situations such as certain chaotic regimes. We describe the ergodic properties of the maps for which the parameter value equals the limit of the bifurcation points.

1. Introduction

Continuous mappings of intervals into themselves display some remarkable properties when regarded as discrete dynamical systems. (For a survey, see May [9] or Collet and Eckmann [14].) One much-studied example is the one-parameter family

$$x \to 1 - \mu x^2 \tag{1.1}$$

which maps $[-1, 1]$ into itself for $0 \leqq \mu \leqq 2$. In this and similar examples, what is interesting is not so much the behavior of any particular mapping; rather, it is the way this behavior changes with μ.

The example (1.1), and the more general one-parameter families $\mu \to \psi_\mu$ we will study, have a simplifying qualitative feature: Each ψ_μ has a unique (differentiable) maximum – at $x=0$ in the example – below which it is increasing and above which it is decreasing. We will consider mappings ψ which satisfy

P1) ψ is a continuously differentiable mapping of $[-1, 1]$ into itself.

P2) $\psi(0)=1$; ψ is strictly increasing on $[-1, 0]$ and strictly decreasing on $[0, 1]$.

P3) $\psi(-x)=\psi(x)$.

The space of all such mappings will be denoted by $\mathscr{P}$. (The condition that the maximum of ψ occurs at zero and that ψ sends zero to one can frequently be arranged, if necessary, by making an affine change of variables.) We have included

0010-3616/80/0076/0211/$08.80

condition P3) mostly for convenience; it simplifies matters and is satisfied by the ψ's we are able to analyze in detail.

One important property of such a transformation is having – or not having – an attracting periodic orbit. (Existence of periodic orbits which are not attracting is much less important, directly at least, in accounting for the behavior of typical orbits.) The fact that $[-1, 1]$ is ordered and connected gives rise to powerful and general methods for proving the existence of periodic orbits – see, for example, Stefan [13] – but these methods do not help very much with the existence of attracting periodic orbits. Note, however, that if 0 is periodic for $\psi \in \mathscr{P}$, then, since $\psi'(0)=0$, its orbit is necessarily attracting. We will say that ψ is *superstable of period p* if 0 is periodic of (minimal) period p for ψ. If ψ_0 is superstable of period p, then any $\psi \in \mathscr{P}$ which is near enough to ψ_0 in the C^1 topology will also have an attracting periodic orbit of period p. Thus for example if

$$\mu \to \psi_\mu$$

is a one-parameter family of elements of $\mathscr{P}$ with ψ_{μ_0} superstable of period p, there is an open interval about μ_0 in the parameter space such that each corresponding ψ_μ has an attracting periodic orbit of period p.

The existence of superstable ψ's can sometimes be proved by simple topological arguments. For example, with our normalization, ψ is superstable of period 2 if and only if $\psi(1)=0$. If we now consider a continuous one-parameter family ψ_μ defined on some interval of μ's, and if $\psi_\mu(1)$ is sometimes positive and sometimes negative, then there must be at least one μ for which ψ_μ is superstable of period 2. We give in Sect. 3 an elaboration of this simple argument which shows that, if $\psi_\mu(1)$ is near 1 for μ near the left end of the parameter interval and near -1 near the right end, there exists a sequence

$$\mu_1 < \mu_2 < \mu_3 < \ldots$$

such that ψ_{μ_j} is superstable of period 2^j. (See Guckenheimer [4] for an alternative approach to the existence of the μ_j's.) It is clear that, if we allow arbitrary (non-monotone) reparametrizations, we cannot hope to prove the existence of unique μ_j's. Moreover, similar topological considerations guarantee that such a parametrized family has, for each large j, many values of μ where ψ_μ is superstable of period 2^j. Nevertheless, in examples like

$$x \to 1 - \mu x^2$$

the first superstable values of μ appear to occur with periods

$$2, 4, 8, 16, \ldots$$

in that order. We will denote the corresponding values of μ by μ_j and $\lim_{j\to\infty} \mu_j$ by μ_∞.

By investigating numerically a number of one-parameter families, Feigenbaum [3] discovered a striking universality property: For large j, $\mu_\infty - \mu_j$ is asymptotic to

$$\text{const} \times \delta^{-j},$$

where $\delta = 4.66920\ldots$ *is apparently the same whatever one-parameter family is considered.* (Note that, encouragingly, this property of the μ_j's is not changed by making a differentiable change of parameter with derivative which does not vanish at μ_∞.)

Having discovered the universality of δ experimentally, Feigenbaum went on to propose an explanation for it which was inspired by the renormalization group approach to critical phenomena in statistical mechanics. The principal result of this paper is to show that Feigenbaum's explanation is correct, at least in a certain limiting regime to be explained below. We will next sketch our version of Feigenbaum's theory, ignoring numerous technical details which will need to be made precise later.

Consider a mapping $\psi \in \mathscr{P}$ and define

$$a = a(\psi) = -\psi(1); \quad b = b(\psi) = \psi(a).$$

Assume

$$0 < a < b\,(< 1)$$

and assume also that

$$\psi(b) = \psi^2(a) < a.$$

ψ then maps

$$[-a, a] \quad \text{onto} \quad [b, 1] \quad \text{and} \quad [b, 1] \quad \text{onto} \quad [-a, \psi(b)] \subset [-a, a],$$

i.e. it exchanges the two non-intersecting intervals $[-a, a]$ and $[b, 1]$. Hence $\psi \circ \psi$ maps $[-a, a]$ into itself, and $-\psi \circ \psi$ is again unimodal on $[-a, a]$ (see Fig. 1). If we reverse orientation and scale up by a factor of $\frac{1}{a}$, i.e. if we make the linear change of variables

$$x_{\text{old}} = -a x_{\text{new}}$$

then $\psi \circ \psi$ on $[-a, a]$ is transformed to

$$-\frac{1}{a} \psi \circ \psi(-ax) \equiv (\mathscr{T}\psi)(x)$$

on $[-1, 1]$. It is easy to verify that, with our hypotheses, $\mathscr{T}\psi$ again has properties P1)–P3) [but the condition $a(\mathscr{T}\psi) > 0$ or $a(\mathscr{T}\psi) < b(\mathscr{T}\psi)$ may fail]. We will refer to the transformation $\mathscr{T}$ as the *doubling transformation.* The doubling transformation is essentially just composition of ψ with itself, but combined with restriction of $\psi \circ \psi$ to a subdomain of the original domain and then a scaling (and reversal of orientation) chosen to preserve the "normalization" $\psi(0) = 1$. This combined operation, in contrast with composition alone, does not give rise to a more complicated-looking transformation. The utility of $\mathscr{T}$ in studying superstable ψ's lies largely in the remark that, provided ψ satisfies the conditions given above for $\mathscr{T}\psi$ to be defined, ψ is superstable of period p if and only if $\mathscr{T}\psi$ is superstable of period $p/2$ (and, in particular, p must be even).

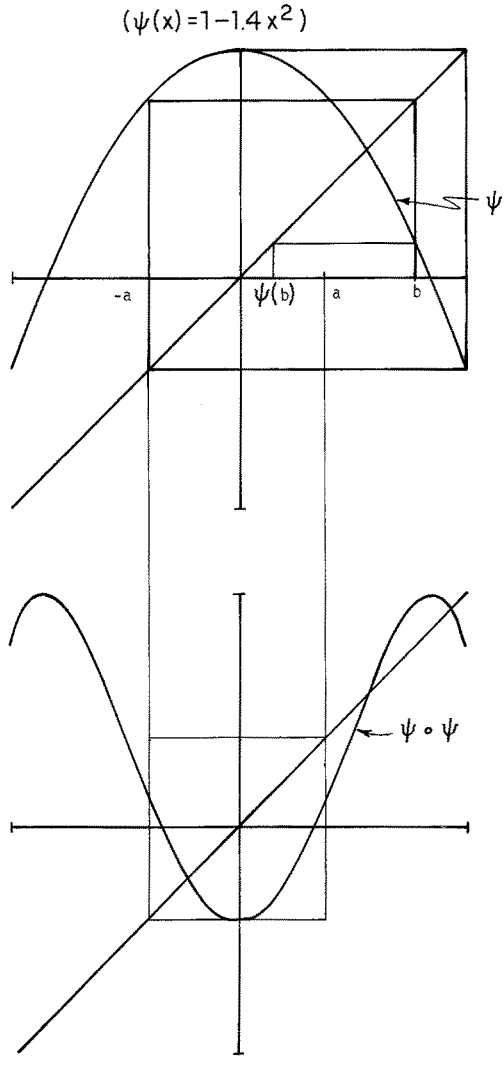


Fig. 1

We now, following Feigenbaum, propose some geometrical hypotheses about how $\mathscr{T}$ acts in the space $\mathscr{P}$ of transformations and show how these hypotheses account for the universality of δ. The picture is as follows:

a) $\mathscr{T}$ has a fixed point ϕ.

b) The derivative of $\mathscr{T}$ at the fixed point ϕ has a simple eigenvalue which is larger than one (and which will turn out to be δ); the remainder of its spectrum is contained in the open unit disk. $\mathscr{T}$ thus has a one-dimensional unstable manifold W_u and a codimension-one stable manifold W_s at ϕ.

c) The unstable manifold W_u intersects transversally the codimension-one surface Σ_1,

$$\Sigma_1 = \{\psi : \psi(1) = 0\}.$$

(Note that Σ_1 is exactly the set of ψ's which are superstable of period 2.) See Fig. 2.

Using this picture, we can account for the universality of δ as follows: Form successive inverse images $\Sigma_2, \Sigma_3, \ldots$ of Σ_1 under $\mathscr{T}$:

$$\Sigma_j = \mathscr{T}^{-(j-1)}\Sigma_1 .$$

Note that if $\psi \in \Sigma_j$ then $\mathscr{T}^{(j-1)}\psi \in \Sigma_1$, so $\mathscr{T}^{j-1}\psi$ is superstable of period 2, so ψ is superstable of period 2^j. The successive Σ_j's come closer and closer to W_s; in fact, a straightforward argument (which we will give in detail later) shows that the separation between Σ_j and W_s decreases exponentially like δ^{-j} for large j, where δ is the large eigenvalue of the derivative of $\mathscr{T}$ at ϕ.

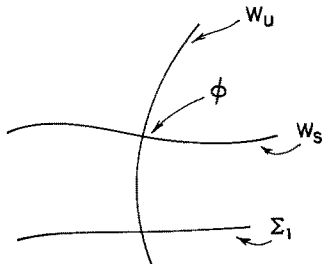


Fig. 2

Now consider a one-parameter family $\mu \to \psi_\mu$ of transformations and regard it as a curve in $\mathscr{P}$. Suppose this curve crosses the stable manifold W_s at $\mu = \mu_\infty$ with non-zero transverse velocity. It is then clear that, at least for large j, there will be a unique μ_j near μ_∞ such that $\psi_{\mu_j} \in \Sigma_j$ (which implies that ψ_{μ_j} is superstable of period 2^j) and that

$$\lim_{j \to \infty} \delta^j(\mu_\infty - \mu_j)$$

exists and is non-zero (see Fig. 3).

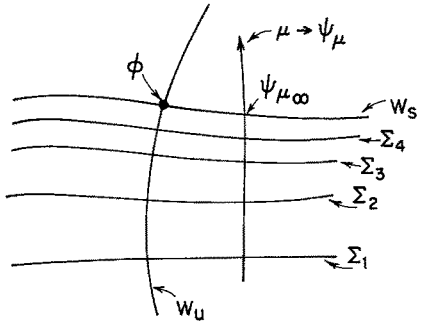


Fig. 3

Thus, Feigenbaum's hypotheses not only account for the universal rate at which μ_j approaches μ_∞; they also provide in principle an independent prescription for computing δ. They have other consequences as well; we will mention here just two of them:

1. For all j, $\mathscr{T}^{j-1}\psi_{\mu_j}$ is superstable of period 2. Because $\mathscr{T}$ contracts in the W_s direction, the $\mathscr{T}^{j-1}\psi_{\mu_j}$ converge as $j \to \infty$ to the point of intersection of Σ_1 with W_s,

which we will denote by ϕ_0^*. Thus, ϕ_0^* is also universal; for any one-parameter family as above, if we form

$$\psi_{\mu_j}^{2^{j-1}}$$

and scale properly, we get something near to ϕ_0^* for large j. Similarly, $\mathscr{T}^j \psi_{\mu_\infty}$ converges to ϕ.

2. Let $\tilde{\Sigma}_1$ denote the surface

$$\{\psi : \psi^3(1) = -\psi(1)\}$$

(i.e. the set of ψ's such that $-\psi(1)$ is a fixed point for ψ^2). Misiurewicz [11] has shown that there is an open set of ψ's on $\tilde{\Sigma}_1$ which admit an absolutely continuous invariant measure (and hence which have typical orbits which are not periodic). We will see that $\tilde{\Sigma}_1$ intersects W_u transversally, with point of intersection inside this open set. (The intersection point will lie *above* W_s in Figs. 2 and 3.) Again form successive inverse images of $\tilde{\Sigma}_1$ under $\mathscr{T}$

$$\tilde{\Sigma}_j = \mathscr{T}^{-(j-1)} \tilde{\Sigma}_1 ;$$

these surfaces converge to W_s, again exponentially with rate δ, from the side opposite to that of the Σ_j's. Again, for each large j, there will be a unique $\tilde{\mu}_j$ near μ_∞ with

$$\psi_{\tilde{\mu}_j} \in \tilde{\Sigma}_j ,$$

and the $\tilde{\mu}_j$'s converge to μ_∞ in the usual way:

$$\lim_{j \to \infty} (\tilde{\mu}_j - \mu_\infty) \cdot \delta^j$$

exists and is non-zero. For large enough j, $\mathscr{T}^{j-1} \psi_{\tilde{\mu}_j}$ will be near to the point of intersection of $\tilde{\Sigma}_1$ with W_u and hence will admit an absolutely continuous invariant measure. From this it is easy to show that $\psi_{\tilde{\mu}_j}$ itself admits an absolutely continuous invariant measure and hence also has orbits which are typically non-periodic.

Thus μ_∞ is the limit of values of μ for which ψ_μ is chaotic. We warn the reader, however, that not all ψ_μ's for μ just above μ_∞ are chaotic; for example, there is a sequence $\hat{\mu}_j$, again converging to μ_∞ from above, with the same exponential rate, such that

$$\hat{\psi}_{\hat{\mu}_j} \quad \text{is superstable with period} \quad 3 \cdot 2^j .$$

As indicated earlier, we are going to prove that Feigenbaum's hypotheses are correct in certain cases. As Feigenbaum has noted, the universality of δ is somewhat relative – its value depends on the function space in which the ψ's are assumed to lie. We will consider functions ψ of the form

$$\psi(x) = f(|x|^{1+\varepsilon}) .$$

where the function f is smooth. Except for a result on the uniqueness of the fixed point, we will in fact have to assume that f is analytic in a complex neighborhood of $[0, 1]$. We would of course like to deal with the case $\varepsilon = 1$, but the argument we are going to give is a perturbative analysis valid only for sufficiently small positive values of ε. Our results could be expressed in terms of convergent series expansions

in ε and various iterated logarithms of ε which are analogues of the ε-expansions occurring in the renormalization-group approach to the theory of critical phenomena. Work in progress, using quite different techniques, indicates that at least partial results can be obtained for $\varepsilon=1$ [8], see also note added in proof.

As an indication that there is some simple behavior at $\varepsilon=0$ about which we could hope to carry out a perturbative analysis, consider the family of functions

$$\psi_a(x)=1-(1+a)|x|,$$

for small positive a. A straightforward calculation shows that

$$(\mathscr{T}\psi_a)(x)=1-(1+a)^2|x|,$$

i.e.

$$\mathscr{T}\psi_a=\psi_{2a+a^2}.$$

Thus, the curve $a\to\psi_a$ is invariant under the action of $\mathscr{T}$, and the end point at $a=0$, although not in the domain of definition of $\mathscr{T}$, is a sort of virtual fixed point. If we consider instead

$$\psi(x)=1-(1+a)|x|^{1+\varepsilon}$$

we no longer get such a simple closed-form expression for $\mathscr{T}\psi$ but we do get

$$(\mathscr{T}\psi)(x)=1-(1+a)^2a^{\varepsilon}|x|^{1+\varepsilon}+O(\varepsilon).$$

This suggests that, for small ε, there might be a fixed point near ψ, where a is to be determined approximately by

$$1+a\simeq a^{\varepsilon}(1+a)^2,$$

i.e.

$$a^{\varepsilon}\simeq 1-a,$$

i.e.

$$\varepsilon\log a\simeq -a \quad \text{or} \quad \varepsilon\simeq\frac{a}{(-\log a)}.$$

Observe that this, if correct, implies that $\varepsilon\ll a$ and hence suggests that it should be possible to get a fixed point by adding to $1-(1+a)|x|^{1+\varepsilon}$ a correction which is small relative to a.

2. Statement of Results

We are going to consider functions $\psi\in\mathscr{P}$ of the form

$$\psi(x)=f(|x|^{1+\varepsilon}),$$

where f is the restriction to $[0,1]$ of a function analytic in some domain in the complex plane. Through most of our analysis, the domain Ω of analyticity will not affect our results very much – although it presumably affects how small we have to take ε in order to make our estimates work. There is one point in Sect. 7 where it is necessary to impose some conditions on Ω; these conditions are met if we take Ω to be an open disk with center 1/2 and radius not too much larger than 1/2. Aside

from this one argument, we can work with Ω any bounded connected open set in $\mathbb{C}$ containing $[0,1]$. We will mostly think of Ω as chosen once and for all and so will frequently suppress it from our notation. We will write $\mathfrak{H}(\Omega)$, or simply $\mathfrak{H}$, for the real Banach space of functions bounded and analytic on Ω, and real on $\Omega\cap\mathbb{R}$, equipped with the supremum norm. For $\varepsilon>0$, we denote by $\mathscr{P}_\varepsilon(\subset\mathscr{P})$ the set of functions ψ on $[-1,1]$ of the form

$$\psi(x)=f(|x|^{1+\varepsilon}),$$

with $f\in\mathfrak{H}$ and satisfying

$$f(0)=1;\quad \frac{df}{dt}<0 \text{ on } [0,1];\quad f(1)>-1.$$

We can identify $\mathscr{P}_\varepsilon$ in an obvious way with an open subset of the Banach space

$$\{g\in\mathfrak{H}(\Omega): g(0)=0\}.$$

Theorem 2.1. *For ε sufficiently small, $\mathscr{T}$ has a fixed point ϕ in $\mathscr{P}_\varepsilon$. If we write*

$$\phi_\varepsilon(x)=f_\varepsilon(|x|^{1+\varepsilon})$$

then $f_\varepsilon(t)$ extends to a function jointly analytic in (ε,t) for

$$\varepsilon\in\{z\in\mathbb{C}\setminus[-\infty,0]:|z|<\varepsilon_0\}$$

and $t\in\Omega$. We denote $-\phi_\varepsilon(1)$ by λ_ε; then

$$\begin{aligned}\lambda_\varepsilon&=-\varepsilon\log\varepsilon+O(\varepsilon)\\ f_\varepsilon(t)&=1-(1+\lambda_\varepsilon)t+O(\varepsilon^2\log\varepsilon).\end{aligned}\tag{2.1}$$

ϕ_ε is an isolated fixed point for $\mathscr{T}$ in $\mathscr{P}_\varepsilon$; it has negative Schwarzian derivative (see Singer [12]), i.e.

$$\frac{\phi_\varepsilon'''}{\phi_\varepsilon'}-\frac{3}{2}\left(\frac{\phi_\varepsilon''}{\phi_\varepsilon'}\right)^2<0.$$

Theorem 2.2. *For ε sufficiently small, ϕ_ε is an isolated fixed point for $\mathscr{T}$ in the space of functions*

$$\psi(x)=f(|x|^{1+\varepsilon})$$

with f twice continuously differentiable on $[0,1]$.

Note that, if $\hat\varepsilon>\varepsilon$, then $\phi_{\hat\varepsilon}(x)=f_{\hat\varepsilon}(|x|^{1+\hat\varepsilon})$ can also be written as $g(|x|^{1+\varepsilon})$ with g *continuously differentiable on* $[0,1]$. Thus, $\mathscr{T}$ has at least a one-parameter family of fixed points of the form $g(|x|^{1+\varepsilon})$ with g only once continuously differentiable.

Notational Convention: From now on the symbols ϕ_ε and λ_ε are permanently reserved to denote the above objects. We will frequently suppress the subscript ε.

Theorem 2.3. *The transformation $\mathscr{T}$ is infinitely differentiable in a neighborhood of ϕ_ε in $\mathscr{P}_\varepsilon$. The derivative of $\mathscr{T}$ at ϕ_ε has one simple eigenvalue $\delta_\varepsilon>1$ which approaches*

2 as ε approaches zero. The diameter of the smallest disk centered at zero containing the rest of its spectrum goes to zero with ε.

Theorem 2.4. *$\mathcal{T}$ has a smooth stable manifold, W_s, of codimension one and a smooth unstable manifold, W_u, of dimension one, at ϕ_ε. For each $a\in[-1,1]$ there is a unique point ϕ_a^* on W_u with*

$$\phi_a^*(1)=-a.$$

W_u crosses the surfaces Σ_1 and $\tilde{\Sigma}_1$ (defined in Sect. 1) transversally. Each ϕ_a^* has negative Schwarzian derivative.

Theorem 2.5. *Let $\mu\to\psi_\mu$ be a continuously differentiable parametrized curve in $\mathcal{P}_\varepsilon$ which crosses the stable manifold W_s with non-zero transverse velocity at $\mu=\mu_\infty$. There exist sequences μ_j and $\tilde{\mu}_j$ converging to μ_∞ from opposite sides such that*

$$\lim_{j\to\infty}\delta^j(\mu_\infty-\mu_j) \quad \text{and} \quad \lim_{j\to\infty}\delta^j(\mu_\infty-\tilde{\mu}_j)$$

are both finite and non-zero, and such that ψ_{μ_j} is superstable of period 2^j and $\psi_{\tilde{\mu}_j}$ admits an absolutely continuous invariant measure for each sufficiently large j. Moreover, the ratio of $\lim_{j\to\infty}\delta^j(\mu_\infty-\mu_j)$ to $\lim_{j\to\infty}\delta^j(\mu_\infty-\tilde{\mu}_j)$ is also universal, i.e. does not depend on the particular parametrized family under consideration.

Remark. One instance of such a parametrized family is

$$\psi_\mu(x)=\psi(\mu\cdot x)$$

for a fixed function ψ sufficiently near to ϕ_ε. We can then in particular take

$$\psi_\mu(x)=1-\mu|x|^{1+\varepsilon}.$$

[Actually, the first statement is not quite true. For if $\mu>1$, then $x\to\psi(\mu\cdot x)$ need not be in $\mathcal{P}_\varepsilon(\Omega)$, but it is in $\mathcal{P}_\varepsilon(\mu^{-(1+\varepsilon)}\Omega)$.]

Theorem 2.6. *If $\psi\in W_s$, then ψ has an invariant Cantor set J.*

1) *There is a decreasing chain of closed subsets of $[-1,1]$*

$$J^{(0)}\supset J^{(1)}\supset J^{(2)}\supset\dots$$

each of which contains 0, and each of which is mapped onto itself by ψ.

2) *Each $J^{(i)}$ is a disjoint union of 2^i closed intervals. $J^{(i+1)}$ is constructed by deleting an open subinterval from the middle of each of the intervals making up $J^{(i)}$.*

3) *ψ maps each of the intervals making up $J^{(i)}$ onto another one; the induced action on the set of intervals is a cyclic permutation of order 2^i.*

We let J denote $\bigcap_i J^{(i)}$. ψ maps J onto itself in a one-one fashion. Every orbit in J is dense in J. If, besides being on W_s, ψ has negative Schwarzian derivative – for which it suffices that it be near ϕ_ε – then we have:

4) *For each $k=1,2,\dots$ ψ has exactly one periodic orbit of period 2^{k-1}. This periodic orbit is repelling and does not belong to $J^{(k)}$; ψ has no periodic orbits other than these.*

5) *Every orbit of* ψ *either*

a) *lands after a finite number of steps exactly on one of the periodic orbits enumerated in* 4) *or*

or

b) *converges to the Cantor set* J *in the sense that, for each* k, *it is eventually contained in* $J^{(k)}$.

There are only countably many orbits of type a).

Theorem 2.7. *Again assume that* $\psi \in W_s$, *and let* $J^{(i)}$, J *be as in Theorem* 2.6. *Let* ν *denote the probability measure with support* J *which for each* i *assigns equal weight to each of the* 2^i *intervals making up* $J^{(i)}$.

1) ν *is invariant under the action of* ψ; *it is the only invariant probability measure on* J.

2) *The abstract dynamical system* (ν, ψ) *is ergodic but not weak mixing.*

3) *If* x *is any point of* $[-1,1]$ *whose orbit converges to* J, *and if* f *is any continuous function on* $[-1,1]$, *then*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(\psi^n(x)) = \int f d\nu .$$

In particular, if ψ is close enough to ϕ_ε so that Theorem 2.6 holds, then this equality holds for all but countably many x's. Similar results were obtained by Misiurewicz [10]. The analysis leading to the Cantor set also gives an attractive picture of how the bifurcation at μ_∞ looks. This is described in detail at the end of Sect. 8.

The proofs will be organized as follows:

In Sect. 3 we develop the elementary theory of the doubling transformation $\mathscr{T}$, and prove for a fairly general class of one-parameter families $\{\psi_\mu\}$ in $\mathscr{P}$ the existence of an increasing sequence μ_j of parameter values such that ψ_{μ_j} is superstable of period 2^j.

Section 4 gives the proofs of Theorem 2.1 – except for the estimate (2.1) on the precise form of the fixed point, which is deferred to Sect. 7 – and Theorem 2.3. Theorem 2.2 is proved in Sect. 5.

Section 5 gives the precise definitions of global stable and unstable manifolds that we use and proves a general theorem, sketched in the introduction, permitting us to deduce Theorem 2.5 immediately from Theorem 2.4. Theorem 2.4 is proved in Sect. 7, and Theorems 2.6 and 2.7 in Sect. 8.

3. The Doubling Transformation

In this section we develop the elementary theory of the doubling transformation $\mathscr{T}$. Let $\psi \in \mathscr{P}$, and define

$$a = a(\psi) = -\psi(1) .$$

If $a > 0$ we also define

$$b = b(\psi) = \psi(a) .$$

The domain of $\mathscr{T}$, $\mathscr{D}(\mathscr{T})$, is the set of all $\psi\in\mathscr{P}$ such that:

$$1)\ a>0$$
$$2)\ b>a$$
$$3)\ \psi(\psi(a))\leqq a,$$

and for $\psi\in\mathscr{D}(\mathscr{T})$ we define

$$(\mathscr{T}\psi)(x)=-\frac{1}{a}\psi\circ\psi(ax).$$

Remark. Although $\mathscr{D}(\mathscr{T})$ is defined by three conditions, the boundary of $\mathscr{D}(\mathscr{T})$ consists in fact of two surfaces:

$$a=0$$
$$\psi(\psi(a))=a.$$

This comes about because, in moving from $\mathscr{D}(\mathscr{T})$ to a region where 2) fails we must pass through a point where $b=a$, i.e. $\psi(a)=a$, and this implies $\psi(\psi(a))=a$. Normally, we would expect conditions 2) and 3) to fail simultaneously, but it is easy to find situations in which 3) fails and 2) does not.

Proposition 3.1. *Let $\psi\in\mathscr{P}$ satisfy $a(\psi)=0$, and let (ψ_n) be a sequence in $\mathscr{P}$ converging to ψ in the C^1 topology and with $a(\psi_n)>0$ for all n. Then*

1) *$\psi_n\in\mathscr{D}(\mathscr{T})$ for sufficiently large n.*
2) *$(\mathscr{T}\psi_n)(1)\to 1$ as $n\to\infty$.*

In other words: Every point on the surface $\{a(\psi)=0\}$ is part of the boundary of $\mathscr{D}(\mathscr{T})$, and $\mathscr{T}$ sends ψ's near this surface to functions near the constant function 1.

Proof. It is easy to see that, if $\psi^2(a)<0$, then $\psi\in\mathscr{D}(\mathscr{T})$. We will show:

$$\frac{\psi_n^2(a_n)-(-a_n)}{a_n}\to 0\quad (a_n=a(\psi_n))$$

which implies both 1) [since it implies $\psi_n^2(a_n)<0$ eventually] and 2) [since $\mathscr{T}\psi_n(1)=-\psi_n^2(a_n)/a_n$].

In view of the facts that

$$a_n\to 0;\quad \psi_n\to\psi \text{ in } C^1;\quad \psi'(0)=0,$$

we get

$$\frac{\psi_n(a_n)-\psi_n(0)}{a_n}\to 0.$$

On the other hand,

$$|\psi_n'(x)|\leqq M \text{ uniformly in } x,n$$

so

$$\left|\frac{\psi_n^2(a_n)-(-a_n)}{a_n}\right|=\left|\frac{\psi_n(\psi_n(a_n))-\psi_n(\psi_n(0))}{a_n}\right|\leqq M\left|\frac{\psi_n(a_n)-\psi_n(0)}{a_n}\right|\to 0$$

as claimed.

It is on the other hand clear that if ψ satisfies $\psi\circ\psi(a)=a$ and if $\psi_n\in\mathscr{D}(\mathscr{T})$; $\psi_n\to\psi$, then

$$\mathscr{T}\psi_n(1)=-\frac{1}{a_n}\psi_n\circ\psi_n(a_n)\to -1\,.$$

Consider now a mapping $\mu\to\psi_\mu$ from an interval $(\mu_0,\tilde{\mu}_0)$ into $\mathscr{P}$, i.e., a one-parameter family of elements of $\mathscr{P}$. We assume the mapping to be continuous in the C^1 topology. We will say that such a one-parameter family is *full* if

$$\psi_\mu(1)\to 1 \quad \text{as} \quad \mu\to\mu_0$$

and

$$\psi_\mu(1)\to -1 \quad \text{as} \quad \mu\to\tilde{\mu}_0$$

(e.g. $\psi_\mu(x)=1-\mu x^2$; $\mu_0=0$; $\tilde{\mu}_0=2$).

We have already remarked that for any such one-parameter family there must be at least one μ such that $\psi_\mu(1)=0$, i.e. such that ψ_μ is superstable. There may be many such μ's; in any case, we denote by μ_1 the *largest* such. Proposition 1 shows that, for μ slightly larger than μ_1, $\psi_\mu\in\mathscr{D}(\mathscr{T})$. We will denote by $\tilde{\mu}_1$ the *smallest* $\mu>\mu_1$ such that $\psi_\mu\notin\mathscr{D}(\mathscr{T})$. [Since $b(\psi_\mu)=\psi_\mu^2(1)\to -1$ as $\mu\to\tilde{\mu}_0$, whereas $a(\psi_\mu)=-\psi_\mu(1)\to 1$, condition 2) in the definition of $\mathscr{D}(\mathscr{T})$ must fail before μ reaches $\tilde{\mu}_0$.] By our earlier remarks, $\psi^2_{\tilde{\mu}_1}(a_{\tilde{\mu}_1})=a_{\tilde{\mu}_1}$.

Proposition 3.2. *If $\mu\to\psi_\mu$ is a full one-parameter family, then*

$$\mu\to\mathscr{T}\psi_\mu, \qquad \mu_1<\mu<\tilde{\mu}_1$$

is also a full one-parameter family.

Proof. By Proposition 1, $\mathscr{T}\psi_\mu(1)\to 1$ as $\mu\downarrow\mu_1$, and by the remark following Proposition 1,

$$\mathscr{T}\psi_\mu(1)\to -1 \quad \text{as} \quad \mu\uparrow\tilde{\mu}_1\,.$$

By induction, then, there exist two sequences

$$\mu_0<\mu_1<\mu_2<\ldots<\tilde{\mu}_2<\tilde{\mu}_1<\tilde{\mu}_0$$

such that, for $\mu_j<\mu<\tilde{\mu}_j$, $\psi_\mu\in\mathscr{D}(\mathscr{T}^j)$ and

$$\mu\to\mathscr{T}^j\psi_\mu, \qquad \mu_j<\mu<\tilde{\mu}_j$$

is a full one-parameter family. In particular these sequences are constructed in such a way that

$$\mathscr{T}^{j-1}\psi_{\mu_j}(1)=0$$

i.e. $\mathscr{T}^{j-1}\psi_{\mu_j}$ is superstable of period 2, i.e. ψ_{μ_j} is superstable of period 2^j.

4. Existence and Elementary Properties of the Fixed Point

If $\psi\in\mathscr{P}_\varepsilon$ is in $\mathscr{D}(\mathscr{T})$ as defined in Sect. 3, and if $\psi(x)=f(|x|^{1+\varepsilon})$, then

$$(\mathscr{T}\psi)(x)=F(|x|^{1+\varepsilon}) \qquad x\in[-1,1]\,,$$

where

$$F(t) = -\frac{1}{a} f(|(f(a^{1+\varepsilon}t))|^{1+\varepsilon}); \quad a = -f(1) > 0.$$

The conditions for the definition of $\mathscr{T}\psi$ imply that

$$f(a^{1+\varepsilon}t) > 0 \quad \text{for} \quad 0 \leqq t \leqq 1$$

and we can therefore drop the absolute value sign. If we define $\mathscr{D}_\varepsilon$ to be the set of such functions ψ satisfying in addition

$$a^{1+\varepsilon}\bar{\Omega} \subset \Omega; \quad f(a^{1+\varepsilon}\Omega) \cap (-\infty, 0) = \emptyset;$$
$$\overline{[f(a^{1+\varepsilon}\Omega)]^{1+\varepsilon}} \subset \Omega$$

then, if $\psi \in \mathscr{D}_\varepsilon$, $\mathscr{T}\psi$ is again in $\mathscr{P}_\varepsilon$. We are going to prove that $\mathscr{T}$ has a fixed point in $\mathscr{D}_\varepsilon$ for each sufficiently small positive ε.

It is convenient to introduce a new variable α related to ε by

$$\varepsilon = \frac{-\alpha}{1+\log(\alpha)}.$$

Note that for each small positive α there corresponds exactly one small positive ε and vice versa. Any $\psi \in \mathscr{P}_\varepsilon$ can be written uniquely as

$$\psi(x) = f(|x|^{1+\varepsilon}); \quad f(t) = 1 - t + \alpha t(g(t) - 1),$$

with $g \in \mathfrak{H}(\Omega)$.

Working with g rather than ψ is simply a (linear) change of variables in function space. If $\psi \in \mathscr{D}_\varepsilon$, we will write the g corresponding to $\mathscr{T}\psi$ as $T_\varepsilon g$. The domain of $\mathscr{T}$ is bounded on one side by the surface

$$\psi(1) = 0$$

which corresponds to

$$g(1) = 1.$$

We are going to show that, for small ε, T_ε is defined and well behaved on the open unit ball in $\mathfrak{H}$ and has a fixed point near zero.

To formulate our results concisely, we need some special terminology. If $\mathscr{X}$ is a normed space and ϱ a positive number, we write $\mathscr{X}_\varrho$ for the open ball in $\mathscr{X}$ with center 0 and radius ϱ. A mapping defined on $\mathscr{X}_1$ will be said to be *nearly bounded* if it is bounded on each $\mathscr{X}_\varrho$ with $\varrho < 1$. Similarly, functions will be said to converge *nearly uniformly* if they converge uniformly on each $\mathscr{X}_\varrho$ with $\varrho < 1$.

Proposition 4.1. *For $\varepsilon > 0$ sufficiently small, T_ε is defined on $\mathfrak{H}_1(\Omega)$. The mapping*

$$(\varepsilon, g) \to T_\varepsilon(g)$$

is jointly infinitely differentiable. For fixed ε, derivatives of all orders of T_ε with respect to g are nearly bounded on $\mathfrak{H}_1$. We can decompose T_ε as

$$T_\varepsilon(g) = T_0 g + r_\varepsilon(g),$$

where T_0 is a rank-one linear operator with range the constant functions:

$$(T_0 g)(t) = g(0) + g(1) + g'(1)$$

and r_ε and its g-derivatives of all orders converge almost uniformly to zero with ε.

We emphasize that: Although T_ε is highly non-linear, its zeroth order part T_0 is not only linear but very simple – dividing it by two gives a projection onto the constant functions. Its simplicity makes possible a detailed analysis of the behavior of T_ε for small ε.

Before proving the proposition we note its principal corollary.

Corollary 4.2. 1. *For each sufficiently small $\varepsilon > 0$, there is exactly one solution $g_\varepsilon^{(0)}$ for the fixed point problem*

$$g = T_\varepsilon(g)$$

in $\mathfrak{H}_{1/2}$. (Here, $\frac{1}{2}$ may be replaced by any number less than one.)

$$\varepsilon \to g_\varepsilon^{(0)}$$

is infinitely differentiable and $g_\varepsilon^{(0)}$ approaches zero with ε.

2. *$DT_\varepsilon(g_\varepsilon^{(0)})$ varies continuously with ε and approaches T_0 in operator norm as ε approaches zero.*

3. *Let $0 < \varrho < 1$. For sufficiently small ε, the only part of the spectrum of $DT_\varepsilon(g_\varepsilon^{(0)})$ at a distance greater than ϱ from 0 is a simple positive eigenvalue δ_ε which approaches two as ε approaches zero. The corresponding eigenspace converges to the space of constant functions.*

To prove 1., we write the fixed point problem as

$$g = T_0 g + r_\varepsilon(g)$$

or equivalently as

$$(I - T_0)g = r_\varepsilon(g).$$

Since $T_0^2 = 2T_0$, we have $(I - T_0)^2 = I$ and so the above equation is equivalent to

$$g = (I - T_0) r_\varepsilon(g).$$

Since r_ε and Dr_ε converge to zero nearly uniformly with ε,

$$g \to (I - T_0) r_\varepsilon(g)$$

is a contraction on $\mathfrak{H}_{1/2}$ for ε sufficiently small. The existence and uniqueness of $g_\varepsilon^{(0)}$ follows from the contraction mapping principle. The smoothness of the dependence of $g_\varepsilon^{(0)}$ on ε follows from the implicit function theorem in Banach space. (See, for example, Dieudonné [2].) That $g_\varepsilon^{(0)}$ approaches 0 with ε follows immediately from the nearly-uniform convergence of r_ε to zero.

Part 2 follows from the joint continuity of $DT_\varepsilon(g)$ in g, ε and the continuity of $g_\varepsilon^{(0)}$ in ε. Part 3 follows from 2 by standard perturbation theory (Kato [7]) and the fact that the spectrum of T_0 reduces to $\{0, 2\}$ with 2 a simple eigenvalue whose associated eigenspace is the constant functions.

The proof of the proposition is a relatively straightforward computation supported by some general theorems. We will give the computation first; then

sketch the justification that the remainder terms do indeed have the properties claimed.

We first do a computation whose result we will need to use again later. We have already seen that if

$$\psi(x)=f(|x|^{1+\varepsilon}); \qquad a=-\psi(1)$$

then the transformation $\psi\to\mathscr{T}\psi$ translates to

$$f\to-\frac{1}{a}f((f(a^{1+\varepsilon}t))^{1+\varepsilon}).$$

We will next write

$$f(t)=1-th(t)$$

and determine how h transforms. We will need the following notation: If $t_0\in\Omega$, we define a bounded linear operator Δ_{t_0} on $\mathfrak{H}(\Omega)$ by

$$(\Delta_{t_0}f)(t)=\frac{f(t)-f(t_0)}{t-t_0} \qquad t\neq t_0$$
$$=f'(t_0) \qquad t=t_0.$$

Now define η_1, η_2 by

$$f(a^{1+\varepsilon}t)=1-at\eta_1 \quad [\text{so } \eta_1=a^{\varepsilon}h(a^{1+\varepsilon}t)]$$
$$(1-at\eta_1)^{1+\varepsilon}=1-at\eta_2.$$

We now claim: Under the action of $\mathscr{T}$, h transforms as

$$h\to\eta_2\times\{h(1-at\eta_2)+(\Delta_1h)(1-at\eta_2)\}. \tag{4.1}$$

To verify this, write

$$f((f(a^{1+\varepsilon}t))^{1+\varepsilon})=1-(1-at\eta_2)h(1-at\eta_2)$$
$$=1-h(1-at\eta_2)+at\eta_2h(1-at\eta_2).$$

Now use the following expression for the first $h(1-at\eta_2)$,

$$h(1-at\eta_2)=h(1)-at\eta_2(\Delta_1h)(1-at\eta_2)$$

and recall that $a=-f(1)=h(1)-1$, to get

$$-a+at\eta_2\{(\Delta_1h)(1-at\eta_2)+h(1-at\eta_2)\}.$$

Thus

$$-\frac{1}{a}f((f(a^{1+\varepsilon}t))^{1+\varepsilon})=1-t\eta_2\{(\Delta_1h)(1-at\eta_2)+h(1-at\eta_2)\}$$

from which the formula (4.1) for the action of $\mathscr{T}$ on h can be read off.

We must next insert the expression

$$h(t)=1-\alpha(g(t)-1)$$

and extract the principal terms for small ε. In so doing, we generate a large number of remainder terms, and it is convenient to have a systematic notation for the spaces in which the remainder terms lie. Let $\mathfrak{R}$ denote the space of all mappings

$$r:(\varepsilon, g)\to r(\varepsilon, g)$$

defined on a set of the form $(0,\varepsilon_0)\times\mathfrak{H}_1(\Omega)$ with values in $\mathfrak{H}(\Omega)$. Here ε_0 is a strictly positive number which may vary with r. These mappings are required to be jointly infinitely differentiable in ε, g, and derivatives of all orders with respect to g are required to be nearly bounded in $\mathfrak{H}_1(\Omega)$ for each ε and to converge to zero nearly uniformly with ε. We will use $\mathfrak{B}$ to denote the analogous space of functions which, together with their g derivatives, will merely be required to remain *bounded* as ε approaches zero and $\mathfrak{R}_0$, $\mathfrak{B}_0$ to denote the analogous spaces of functions taking values in $\mathbb{R}$ rather than $\mathfrak{H}$. Recall, also, that ε and α are related by

$$\varepsilon=\frac{-\alpha}{(1+\log\alpha)};$$

whenever α appears in one of our formulas it is to be regarded as a function of ε.

Since $a=h(1)-1$ and we are writing $h(t)=1-\alpha(g(t)-1)$ we have

$$a=\alpha(1-g(1))$$

and hence

$$a^\varepsilon=\alpha^\varepsilon(1-g(1))^\varepsilon.$$

Now $\alpha^\varepsilon=\exp[\varepsilon\log\alpha]=\exp\left(\frac{-\alpha\log\alpha}{1+\log\alpha}\right)=\exp[-\alpha-\varepsilon]$.

Thus

$$a^\varepsilon=e^{-\alpha}+\varepsilon b_1;\qquad b_1\in\mathfrak{B}_0.$$

Also,

$$g(a^{1+\varepsilon}t)=g(0)+a^{1+\varepsilon}t(\Delta_0 g)(a^{1+\varepsilon}t)=g(0)+\alpha b_2,\qquad b_2\in\mathfrak{B}.$$

Thus

$$\begin{aligned}\eta_1&=a^\varepsilon(1+\alpha(1-g(a^{1+\varepsilon}t)))\\&=(e^{-\alpha}+\varepsilon b_1)(1+\alpha-\alpha g(0)-\alpha^2 b_2)\\&=1-\alpha g(0)+\alpha r_1,\qquad r_1\in\mathfrak{R}.\end{aligned}$$

[We have used the fact that ε/α goes to zero with ε to replace εb_1 by $\alpha(\varepsilon/\alpha)b_1$ with $\varepsilon/\alpha b_1\in\mathfrak{R}_0$.] Now

$$1-at\eta_2=(1-at\eta_1)^{1+\varepsilon},$$

so

$$\eta_2=\eta_1+\varepsilon\alpha b_3,\qquad b_3\in\mathfrak{B},$$

so

$$\eta_2=1-\alpha g(0)+\alpha r_2,\qquad r_2\in\mathfrak{R}.$$

Also

$$h(1-at\eta_2)+(\Delta_1 h)(1-at\eta_2)=1+\alpha-\alpha g(1-at\eta_2)-\alpha(\Delta_1 g)(1-at\eta_2)$$
$$=1+\alpha(1-g(1)-g'(1)+r_3),\quad r_3\in\mathfrak{R}.$$

Thus

$$\eta_2\{h(1-at\eta_2)+(\Delta_1 h)(1-at\eta_2)\}=(1-\alpha g(0)+\alpha r_2)(1+\alpha(1-g(1)-g'(1)+r_3))$$
$$=1-\alpha(g(0)+g(1)+g'(1)-1+r_4),\quad r_4\in\mathfrak{R}$$

and hence by inspection

$$T_\varepsilon g=g(0)+g(1)+g'(1)+r_4,$$

as desired.

Remark. These computations show that $\mathscr{T}$ admits a fixed point $\phi_\varepsilon\in\mathscr{P}_\varepsilon$ with

$$\phi_\varepsilon(x)=1-(1+\alpha(\varepsilon))|x|^{1+\varepsilon}+\text{correction},$$

where the correction vanishes more rapidly than α as ε goes to zero. They also show that there is no other fixed point in a ball about ϕ_ε whose radius is bounded below by const$\alpha(\varepsilon)$. If we write

$$\lambda_\varepsilon=-\phi_\varepsilon(1),$$

then

$$\lambda_\varepsilon=\alpha(\varepsilon)+O(\alpha),\quad\text{or}\quad\lambda_\varepsilon=-\varepsilon\log\varepsilon+o(\varepsilon\log\varepsilon).$$

More detailed computations to be done in Sect. 7 show that in fact

$$\lambda_\varepsilon=-\varepsilon\log\varepsilon+O(\varepsilon)$$

and that

$$\phi_\varepsilon(x)=1-(1+\lambda_\varepsilon)|x|^{1+\varepsilon}+O(-\varepsilon^2\log\varepsilon).$$

We turn now to the problem of justifying the above computations, i.e. of showing that the remainder terms do indeed have the asserted properties. The verifications are tedious and we will not do all of them, but we will work through one in full detail. We wrote, in the course of the computation,

$$a^{1+\varepsilon}t(\Delta_0 g)(a^{1+\varepsilon}t)=\alpha b_2\quad\text{with}\quad b_2\in\mathfrak{B}. \tag{4.2}$$

We now want to prove this. The proof is based on a number of principles which we list here:

a) The mapping $(\varepsilon,g)\to g$ is in $\mathfrak{B}$.

b) For any $g_0\in\mathfrak{H}$, the constant mapping $(\varepsilon,g)\to g_0$ is in $\mathfrak{B}$.

c) If $b_0\in\mathfrak{B}_0$, the mapping $(\varepsilon,g)\to$(the constant function with value $b_0(\varepsilon,g)$) is in $\mathfrak{B}$.

d) Let $\mathfrak{V}_1,\ldots,\mathfrak{V}_n$ be open sets in $\mathfrak{H}$ and let

$$\Phi:\mathfrak{V}_1\times\ldots\times\mathfrak{V}_n\to\mathfrak{H}$$

be bounded, infinitely differentiable, and have bounded derivatives. Further, let $b_1, \ldots, b_n \in \mathfrak{B}$, with the range of b_i contained in $\mathfrak{B}_i$ for each ε. Then

$$(\varepsilon, g) \to \Phi(b_1(\varepsilon, g), \ldots, b_n(\varepsilon, g))$$

is in $\mathfrak{B}$.

e) (Corollary of d.) If $b_1, b_2 \in \mathfrak{B}$, then

$$(\varepsilon, g) \to b_1(\varepsilon, g) \cdot b_2(\varepsilon, g)$$

(pointwise product of analytic functions) is in $\mathfrak{B}$. Similarly, if b_1, b_2 are in $\mathfrak{B}_0$, so is their product.

f) If Ω_1, Ω_2 are bounded open sets in $\mathbb{C}$, we write $\mathfrak{H}(\Omega_1, \Omega_2)$ for the set of all $g \in \mathfrak{H}(\Omega_1)$ with $\overline{g(\Omega_1)} \subset \Omega_2$. [Hence, if Ω_2 is the open unit disk, $\mathfrak{H}(\Omega, \Omega_2) = \mathfrak{H}_1(\Omega)$.] $\mathfrak{H}(\Omega_1, \Omega_2)$ is an open subset of $\mathfrak{H}(\Omega_1)$. Now let $\Omega_2' \subset \bar{\Omega}_2' \subset \Omega_2$. Then composition $(g_1, g_2) \to g_2 \circ g_1$ is a C^∞ function with bounded derivatives from $\mathfrak{H}(\Omega_1, \Omega_2') \times \mathfrak{H}(\Omega_2, \Omega_3)$ into $\mathfrak{H}(\Omega_1, \Omega_3)$.

We omit the proofs of these statements. Note that a)–e) remain true if we replace $\mathfrak{H}$ by C^2, but that f) depends upon the properties of analytic functions, and fails in C^2.

We are now ready to verify 2).

1. The mappings

$$(\varepsilon, g) \to 1 - g(1), \quad (1 - g(1))^\varepsilon, \quad \alpha, \quad e^{-\alpha}, \quad e^{-\varepsilon}$$

are all in $\mathfrak{B}_0$. This is readily proved by direct verification. Note that the g-derivatives of $(1-g(1))^\varepsilon$ have singularities as $g(1) \to 1$; it is for this reason (only) that we have to work with functions with derivatives which are nearly bounded – rather than bounded – on the unit ball of $\mathfrak{H}$.

We will from now on write, as a shorthand, statements like $(1-g(1))^\varepsilon \in \mathfrak{B}_0$ rather than the more logical

$$((\varepsilon, g) \to (1 - g(1))^\varepsilon) \in \mathfrak{B}_0 .$$

2. $a = \alpha(1-g(1)) \in \mathfrak{B}_0$; $a^\varepsilon = e^{-\alpha} e^{-\varepsilon}(1-g(1))^\varepsilon \in \mathfrak{B}_0$; $a^{1+\varepsilon} \in \mathfrak{B}_0$. [Use 1e).]

3. $a^{1+\varepsilon} t \in \mathfrak{B}$. [Use 2b), c), and e).] If ε_0 is small enough there exists a domain Ω_1 with $\bar{\Omega}_1 \subset \Omega$ such that $a^{1+\varepsilon} t \in \Omega_1$ for all $\varepsilon < \varepsilon_0$, $g \in \mathfrak{H}_1$, $t \in \Omega$.

4. $\Delta_0 g \in \mathfrak{B}$. [Use a), d).]

5. $\Delta_0 g(a^{1+\varepsilon} t) \in \mathfrak{B}$. [Use 3., 4., f), d).]

6. $\dfrac{1}{\alpha} a^{1+\varepsilon} \Delta_0 g(a^{1+\varepsilon} t) = (1 - g(1)) a^\varepsilon \Delta_0 g(a^{1+\varepsilon} t) \in \mathfrak{B}$. [Use 1., 2., 5., c), e).]

This completes the proof of (4.1).

The arguments given so far show that the fixed point $g_\varepsilon^{(0)}$ varies smoothly with ε. We next show that $g_\varepsilon^{(0)}(t)$ is jointly analytic in ε, t. The logarithm appearing in the relation between $\lambda_\varepsilon = -\phi_\varepsilon(1)$ and ε shows that there must be a singularity at $\varepsilon = 0$, and we want to clarify the structure of that singularity. For this purpose it turns out to be useful to consider a somewhat contrived generalized fixed point problem in which the relation between ε and α is partly relaxed. Recall that $\mathscr{T}$ takes the form

$$f \to -\frac{1}{a} f((f(a^{1+\varepsilon} t))^{1+\varepsilon})$$

with $a = -f(1) = \alpha(1-g(1))$. We modify this transformation as follows: In the innermost argument we replace a^ε by $e^{-\varepsilon-\alpha}(1-g(1))^\varepsilon$. We then replace ε wherever it appears by $\mu \cdot \alpha$ and we regard μ, α as independent parameters. This gives a two-parameter family of transformations whose action we can again express in terms of g related to f as above. Expressed in terms of g, we will denote the transformations by

$$g \to T_{\alpha,\mu}(g).$$

Our original transformations T_ε are recovered through

$$T_{\varepsilon(\alpha)} = T_{\alpha, -\frac{1}{1+\log\alpha}}.$$

The advantage of considering α, μ as independent variables is that $T_{\alpha,\mu}$ turns out to be, heuristically, jointly analytic in α, μ at $(0,0)$; the non-analyticity appears only in the relation between μ and α.

We now, temporarily, drop the requirement of reality for real values of the argument in the definition of the space $\mathfrak{H}$, and we consider the transformations $T_{\alpha,\mu}$ for arbitrary small *complex* values of α, μ. Exactly the same computations as were done in the proof of Proposition 4.1 go through, and we obtain

$$T_{\alpha,\mu}(g) = T_0 g + r_{\alpha,\mu}(g),$$

where $r_{\alpha,\mu}$ and its derivatives of all orders with respect to g converge nearly uniformly to zero as α, μ both go to zero. Hence, as before, the fixed point problem

$$g = T_{\alpha,\mu}(g)$$

can be rewritten as

$$g = (I - T_0) r_{\alpha,\mu}(g)$$

and the right hand side is contractive on $\mathfrak{H}_{1/2}$ for sufficiently small $|\alpha|, |\mu|$. Thus, there exist $\alpha_0, \mu_0 > 0$ such that for all α, μ with $|\alpha| < \alpha_0$, $|\mu| < \mu_0$, $T_{\alpha,\mu}$ has a unique fixed point $g^{(0)}_{\alpha,\mu}$ in $\mathfrak{H}_{1/2}$. Moreover, the computations which show that $r_{\alpha,\mu}$ is small for small α, μ also show that if $\alpha, \mu \to g_{\alpha,\mu}$ is a mapping from $\{(\alpha,\mu) : |\alpha| < \alpha_0, |\mu| < \mu_0\}$ into $\mathfrak{H}_{1/2}$ such that $g_{\alpha,\mu}(t)$ is jointly analytic in α, μ, t then $r_{\alpha,\mu}(g_{\alpha,\mu})(t)$ is again jointly analytic, and so the same is true for

$$(I - T_0) r_{\alpha,\mu}(g_{\alpha,\mu})(t).$$

From this it is easy to show – using the contraction mapping principle in an appropriate space of jointly analytic functions – that $g^{(0)}_{\alpha,\mu}(t)$ is jointly analytic in α, μ, t.

Since the parameter α was introduced artificially, it is desirable to express directly the dependence of $g^{(0)}$ on ε. To do this we need the following lemma.

Lemma 4.3. *There exists a function $r(z_1, z_2)$, analytic on a neighborhood of zero in $\mathbb{C}^2$ and with $r(0,0) = 1$ such that for small positive ε*

$$\alpha = -\varepsilon \log \varepsilon \, r\left(\frac{1}{\log\varepsilon}, \frac{\log(-\log\varepsilon)}{\log\varepsilon}\right)$$

is the unique small positive solution of the equation

$$\varepsilon = \frac{-\alpha}{1+\log\alpha}. \tag{4.3}$$

Proof. Write $\alpha = -\varepsilon\cdot\log\varepsilon\cdot r$ and insert in (4.3). The result is

$$1+\frac{1}{\log\varepsilon}+\frac{\log(-\log\varepsilon)}{\log\varepsilon}+\frac{\log r}{\log\varepsilon}=r.$$

For fixed ε in $(0,1)$ this equation has a unique solution r; for ε small, the solution is near 1. On the other hand, by the implicit function theorem, there is a uniquely determined function $r(z_1, z_2)$ defined and analytic on a neighborhood of zero in $\mathbb{C}^2$ and satisfying

$$1+z_1+z_2+z_1\log r=r;$$

and $r(0,0)=1$. This is thus our desired function r.

We can express the conclusion of the lemma more concisely by saying simply that $\alpha/\varepsilon\log\varepsilon$ is an analytic function of $1/\log\varepsilon$ and $\log(-\log\varepsilon)/\log\varepsilon$. It follows immediately that ε/α is also an analytic function of $1/\log\varepsilon$ and $\log(-\log\varepsilon)/\log\varepsilon$. Thus we get:

Proposition 4.4. *$g_\varepsilon^{(0)}(t)$ is an analytic function of $\varepsilon\log\varepsilon, \dfrac{1}{\log\varepsilon}, \dfrac{\log(-\log\varepsilon)}{\log\varepsilon}$, and t. In particular, $g_\varepsilon^{(0)}(t)$ is jointly analytic and bounded in ε, t for ε in a small disk about zero in $\mathbb{C}$ and off the negative real axis and t in Ω.*

5. Existence and Uniqueness in C^2

Write

$$f(t)=1-(1+\lambda)t+\lambda^2\ell(t). \tag{5.1}$$

We are looking for (ε,ℓ) such that

$$\lambda f(t)+f(f(\lambda^{1+\varepsilon}t)^{1+\varepsilon})=0. \tag{5.2}$$

We study the equation to be satisfied by the second derivative of f. If (5.2) holds, then

$$f'(t)=-\lambda^\varepsilon f'(f(\lambda^{1+\varepsilon}t)^{1+\varepsilon})f(\lambda^{1+\varepsilon}t)^\varepsilon(1+\varepsilon)f'(\lambda^{1+\varepsilon}t), \tag{5.3}$$

and

$$\begin{aligned} f''(t)=-\lambda^{1+2\varepsilon}\{&f''(f(\lambda^{1+\varepsilon}t)^{1+\varepsilon})f(\lambda^{1+\varepsilon}t)^{2\varepsilon}(1+\varepsilon)^2f'(\lambda^{1+\varepsilon}t)^2\\ &+f'(f(\lambda^{1+\varepsilon}t)^{1+\varepsilon})f(\lambda^{1+\varepsilon}t)^\varepsilon(1+\varepsilon)f''(\lambda^{1+\varepsilon}t)\\ &+f'(f(\lambda^{1+\varepsilon}t)^{1+\varepsilon})f(\lambda^{1+\varepsilon}t)^{\varepsilon-1}(1+\varepsilon)\varepsilon f'(\lambda^{1+\varepsilon}t)^2\}. \end{aligned} \tag{5.4}$$

Setting $t=1$ in (5.3) we get

$$\begin{aligned} 1+\lambda-\lambda^2\ell'(1)=&\lambda^\varepsilon\{(1+\lambda)-\lambda^2\ell'(f(\lambda^{1+\varepsilon})^{1+\varepsilon})\}\\ &\cdot\{1-(1+\lambda)\lambda^{1+\varepsilon}+\lambda^2\ell(\lambda^{1+\varepsilon})\}^\varepsilon(1+\varepsilon)\\ &\cdot\{1+\lambda-\lambda^2\ell(\lambda^{1+\varepsilon})\}, \end{aligned} \tag{5.5}$$

or

$$0=\lambda+\varepsilon+\varepsilon\log\lambda+\lambda^2\hat{N}_0(\ell,\varepsilon,\lambda), \tag{5.6}$$

i.e.

$$\varepsilon=-\frac{\lambda}{1+\log\lambda}+\lambda^2\hat{N}_1(\ell,\varepsilon,\lambda)\equiv N_\lambda(\ell,\varepsilon). \tag{5.7}$$

We also rewrite (5.4):

$$\ell''=\lambda K_\lambda(\ell,\varepsilon)\ell''+M_\lambda(\ell,\varepsilon), \tag{5.8}$$

where for $h\in C^0[0,1]$, and f defined as in (5.1),

$$\begin{aligned}(K_\lambda(\ell,\varepsilon)h)(t)=-\lambda^{2\varepsilon}\{&h(f(\lambda^{1+\varepsilon}t)^{1+\varepsilon})f(\lambda^{1+\varepsilon}t)^{2\varepsilon}(1+\varepsilon)^2f'(\lambda^{1+\varepsilon}t)\\&+h(\lambda^{1+\varepsilon}t)f'(f(\lambda^{1+\varepsilon}t)^{1+\varepsilon})f(\lambda^{1+\varepsilon}t)^{\varepsilon}(1+\varepsilon)\},\end{aligned} \tag{5.9}$$

and

$$M_\lambda(\ell,\varepsilon)(t)=-\lambda^{2\varepsilon}\frac{\varepsilon}{\lambda}f'(f(\lambda^{1+\varepsilon}t)^{1+\varepsilon})f(\lambda^{1+\varepsilon}t)^{\varepsilon-1}(1+\varepsilon)f'(\lambda^{1+\varepsilon}t)^2. \tag{5.10}$$

Finally, we define $\mathscr{L}$ by

$$(\mathscr{L}h)(t)=\int_0^t(t-\tau)h(\tau)d\tau-t\int_0^1(1-\tau)h(\tau)d\tau. \tag{5.11}$$

Instead of solving (5.2) directly, we rather discuss first the set of equations (for fixed, small $\lambda>0$)

$$\varepsilon=N_\lambda(\ell,\varepsilon), \tag{5.12a}$$

and

$$\ell=\mathscr{L}(I-\lambda K_\lambda(\ell,\varepsilon))^{-1}M_\lambda(\ell,\varepsilon). \tag{5.12b}$$

We claim that the solutions (ε,ℓ) of (5.12a, b) solve actually (5.2).

[*Proof*. (5.12a) implies (5.6) and (5.5), and hence (5.3) at $t=1$. Equation (5.12b) implies (5.4). Integrating, we find that (5.3) must hold up to an additive constant, but this constant is zero since we have already seen that (5.3) holds at $t=1$. Integrating again, we find that (5.12) holds up to an additive constant, which however must be zero, since, by (5.12b) and the definition of $\mathscr{L}$, $\ell(0)=\ell(1)=1$.]

We now discuss the existence and uniqueness of solutions of (5.12). [This also implies uniqueness of the solution of (5.2) since every solution of (5.2) solves (5.12).] The appropriate function space is

$$\mathfrak{E}=\{(\ell,\varepsilon),\ \ell\in C^2[0,1],\ \ell(0)=\ell(1)=0,\ \varepsilon\in\mathbb{C}\}$$

equipped with the norm

$$\|(\ell,\varepsilon)\|_{\mathfrak{E}}=\sup_{x\in[0,1]}|\ell''(x)|-|\varepsilon|\frac{\log\lambda+1}{2\lambda};\quad(\lambda\text{ is small}).$$

Let $\mathfrak{E}_1$ be the unit ball in $\mathfrak{E}$. $\mathfrak{E}$ is a Banach space. In fact $\ell = \mathscr{L}(\ell'')$, and $|(\mathscr{L}h)'(t)| + |(\mathscr{L}h)(t)| \leqq \text{const} \sup_{t\in[0,1]} |h(t)|$. (All spaces C^0, C^1 are on $[0,1]$.)

We claim that for sufficiently small $\lambda > 0$,

a) N_λ is a contraction[1] from $\mathfrak{E}_1$ to $\mathbb{C}$.

b) M_λ is a contraction from $\mathfrak{E}_1$ to C^1.

c) For $(\ell, \varepsilon) \in \mathfrak{E}_1$, $K_\lambda(\ell, \varepsilon)$ is a bounded linear map from C^0 to C^0 and from C^1 to C^1.

d) For $X_1, X_2 \in \mathfrak{E}_1$,

$$|K_\lambda(X_1)h - K_\lambda(X_2)h|_{C^0} \leqq \text{const}\, \|X_1 - X_2\|_{\mathfrak{E}} |h|_{C^1}.$$

e) $\mathscr{L}$ is bounded linear from C^0 to C^2, and $(\mathscr{L}h)(0) = (\mathscr{L}h)(1) = 0$.

Note that from c), d) above it follows that

$$\begin{aligned} &|(1-\lambda K_\lambda(X_1))^{-1}h - (1-\lambda K_\lambda(X_2))^{-1}h|_{C^0} \\ &\quad = \lambda |(1-\lambda K_\lambda(X_1))^{-1}(K_\lambda(X_1) - K_\lambda(X_2))(1-\lambda K_\lambda(X_2))^{-1}h|_{C^0} \\ &\quad = \lambda\, \text{const}\, \|X_1 - X_2\|_{\mathfrak{E}} |h|_{C^1}, \end{aligned}$$

and hence we see that (5.12) has a unique solution. It remains to verify the claims a)–e). This is tedious but straightforward. Each time we have to estimate

$$h(f_1(\lambda^{1+\varepsilon_1}t)^{1+\varepsilon_1}) - h(f_2(\lambda^{1+\varepsilon_2}t)^{1+\varepsilon_2}),$$

we use the formula

$$|h(u) - h(v)| \leqq |h|_{C^1} |u - v|. \tag{5.13}$$

This is responsible for the fact that the contractions in b) and d) lose a derivative.

We only comment on the otherwise trivial verifications:

a): Obvious from (5.5), (5.6), (5.7).

b): Obvious from (5.10). [Note that $f(\lambda^{1+\varepsilon}t) > \frac{1}{2}$ for small $\lambda \geqq 0$.]

c): Obvious from (5.9) since no derivatives of h occur.

d): Obvious from (5.9) using (5.13).

e): By construction, $(\mathscr{L}h)(0) = (\mathscr{L}h)(1) = 0$, cf. (5.11). On the other hand, a direct computation shows

$$|\mathscr{L}h|_{C^0} \leqq |h|_{C^0}; \quad |(\mathscr{L}h)'|_{C^0} \leqq |h|_{C^0}; \quad |(\mathscr{L}h)''|_{C^0} \leqq |h|_{C^0}.$$

This proves existence and uniqueness in C^2 and hence Theorem 2.2. The same proof works on any interval $[0, A]$ provided $\lambda > 0$ is sufficiently small, and $A > 1$.

6. Stable and Unstable Manifolds

In this section, we present a general argument deriving some analytic consequences from a geometric situation. The geometric situation is as follows:

We consider a twice continuously differentiable mapping $\mathscr{T}$ defined on an open set $\mathfrak{D}$ in a Banach space $\mathfrak{H}$ and taking values in $\mathfrak{H}$. We do not assume that $\mathscr{T}$

1 By contraction we mean a bounded, contractive map

maps $\mathfrak{D}$ into itself, but we do assume that it has a fixed point ϕ. We further assume that $D\mathscr{T}(\phi)$, the derivative of $\mathscr{T}$ at ϕ (which is a bounded linear operator on $\mathfrak{H}$) has spectrum which, except for a single simple eigenvalue $\delta > 1$, is entirely contained in the open unit circle. It then follows from invariant manifold theory that $\mathscr{T}$ admits a stable manifold W_s of codimension one and an unstable manifold W_u of dimension one. We will define these submanifolds precisely later; for present purposes, it is good enough to think of W_s as an invariant surface and W_u an invariant curve (the two of them crossing at the fixed point ϕ) with $\mathscr{T}$ acting in a purely contractive way on W_s and in a purely expansive way on W_u.

We also give ourselves two further objects:

A submanifold Π_1 of $\mathfrak{H}$ of codimension one which intersects W_u transversally at some point $\phi^* \neq \phi$.

A continuously differentiable parameterized curve $\mu \to \psi_\mu$ in $\mathfrak{H}$ which crosses the stable manifold W_s at $\mu = \mu_\infty$ with non-zero transverse velocity.

Although the notation in this section will be chosen to suggest the application of the results obtained here to the main topic of the paper, the reader should note that these results depend only on a few explicit assumptions about the objects under consideration. Symbols like $\mathscr{T}$, ϕ, W_u, etc. are used in a more general sense here than in the remainder of the paper. Moreover, Π_1 will be later on identified with Σ_1 (see Fig. 3).

From this set-up we want to conclude:

a) There exists a sequence μ_n (perhaps defined only for large n), converging to μ_∞, with $\mathscr{T}^{n-1}\psi_{\mu_n} \in \Pi_1$, and such that $\lim_{n\to\infty} \delta^n(\mu_n - \mu_\infty)$ exists and is non-zero.

b) The sequence $\mathscr{T}^{n-1}\psi_{\mu_n}$ converges to ϕ^*.

The significance of these conclusions has been discussed in the introduction. (We would like to be able to make the more precise assertion that μ_n is the unique value of μ near μ_∞ such that $\mathscr{T}^{n-1}\psi_\mu \in \Pi_1$. Whether this is true or not depends on relatively inaccessible global properties of $\mathscr{T}$; but we can say, informally, that μ_n is the unique such μ for which $\mathscr{T}\psi_\mu, \mathscr{T}^2\psi_\mu, \ldots, \mathscr{T}^{n-2}\psi_\mu$ all lie between W_s and Π_1.)

The first step in our analysis will be to define precisely what we mean by stable and unstable manifolds. This is not entirely routine since, in the application we have in mind, the transformation $\mathscr{T}$ is not invertible; in fact, it is not even locally one-one near ϕ.

If $\mathfrak{B}$ is a sufficiently small open ball in $\mathfrak{H}$ with center ϕ then

$$\{\psi \in \mathfrak{B} : \mathscr{T}^j\psi \in \mathfrak{B} \quad \text{for} \quad j = 1, 2, 3, \ldots\}$$

may be shown to be a smooth connected submanifold of $\mathfrak{B}$ of codimension one. We will call this set a *local stable manifold* for $\mathfrak{B}$ at ϕ and denote it by $W_s^{(0)}$. It passes through ϕ and is tangent there to the stable eigenspace for $D\mathscr{T}(\phi)$, i.e. the spectral subspace for the part of the spectrum which is inside the unit disk. The set $W_s^{(0)}$ is mapped into itself by $\mathscr{T}$, and the sequence of sets $\mathscr{T}^j W_s^{(0)}$ shrinks to $\{\phi\}$, i.e. is eventually contained in any neighborhood of ϕ. The proofs of these facts, as well as those to be cited in the next paragraph on local unstable manifolds, can be found in the monograph of Hirsch et al. [6].

If we define (again for $\mathfrak{B}$ a small open ball with center ϕ)

$$\mathfrak{B}_0 = \mathfrak{B}; \qquad \mathfrak{B}_{j+1} = (\mathscr{T}\mathfrak{B}_j) \cap \mathfrak{B},$$

then $\bigcap_j \mathfrak{B}_j$ is a smooth connected one-dimensional submanifold of $\mathfrak{B}$, passing through ϕ and tangent there to the eigenspace of $D\mathcal{T}(\phi)$ corresponding to the large eigenvalue δ. We call this set a *local unstable manifold* for $\mathcal{T}$ at ϕ and denote it by $W_u^{(0)}$.

We have:

$$\mathcal{T} W_u^{(0)} \supset W_u^{(0)}$$

and, for any $\psi \in W_u^{(0)}$ and any $j=1,2,3,\ldots$, there is a unique $\psi_j \in W_u^{(0)}$ such that

$$\mathcal{T}^j \psi_j = \psi;$$

moreover, the sequence (ψ_j) converges to ϕ.

The globalization of the stable and unstable manifolds is complicated by the non-invertibility of $\mathcal{T}$. We will simply define what we mean by *a* stable or unstable manifold without investigating the existence of a unique largest one. Thus we define:

A *stable manifold* for $\mathcal{T}$ is a smooth codimension-one submanifold W_s of the domain of $\mathcal{T}$ such that:

a) $\mathcal{T} W_s \subset W_s$.

b) If $\psi \in W_s$, then $\lim_{j\to\infty} \mathcal{T}^j \psi = \phi$. (Note that this implies that $\mathcal{T}^j \psi \in W_s^{(0)}$ for sufficiently large j.)

c) (Transversality.) For any ψ in W_s, the range of $D\mathcal{T}(\psi)$ is not contained in the tangent space to W_s at $\mathcal{T}\psi$.

An *unstable manifold* for $\mathcal{T}$ is a smooth one-dimensional submanifold W_u of $\mathfrak{H}$ (not necessarily contained in the domain of $\mathcal{T}$) such that

a) $\mathcal{T}(W_u \cap \mathfrak{D}(\mathcal{T})) \supset W_u$.

b) If $\psi \in W_u$, there is a sequence ψ_j converging to ϕ such that $\psi = \mathcal{T}^j \psi_j$. (This implies that $W_u \subset \bigcup_{j=1}^{\infty} \mathcal{T}^j W_u^{(0)}$.)

c) For any $\psi \in W_u \cap \mathfrak{D}(\mathcal{T})$, the tangential derivative of $\mathcal{T}$ along W_u at ψ does not vanish.

Since $W_s^{(0)}$ and $W_u^{(0)}$ are, respectively, stable and unstable manifolds, stable and unstable manifolds do exist.

We now need some special terminology. Let Π_j, $j=1,2,3,\ldots$ and W be submanifolds of $\mathfrak{H}$ of codimension one. We will say that the sequence Π_j *converges to* W *exponentially with rate* δ (δ a real number larger than one) if, for each $\psi \in W$ there is a diffeomorphism from $\mathcal{X}_1 \times (-1,1)$, $\mathcal{X}_1$ the open unit ball in some Banach space $\mathcal{X}$, onto a neighborhood $\mathfrak{B}$ of ψ (i.e. a set of local coordinates at ψ) such that

1. ψ is the image of $(0,0)$.

2. $W \cap \mathfrak{B}$ is the image of $\mathcal{X}_1 \times \{0\}$.

3. For each sufficiently large j, $\Pi_j \cap \mathfrak{B}$ is the image of the graph of a mapping $\hat{\Pi}_j : \mathcal{X}_1 \to (-1,1)$,

where

4. $\delta^j \hat{\Pi}_j$ converges in the C^1 topology on $\mathcal{X}_1$ to a nowhere vanishing limit.

Intuitively, this means that the separation between Π_j and W varies asymptotically (for large j) like δ^{-j} multiplied by a differentiable function of position on W.

The following proposition is nearly obvious:

Proposition 6.1. *Let Π_j converge exponentially to W with rate δ, and let $\mu \to \psi_\mu$ be a continuously differentiable parametrized curve in $\mathfrak{H}$ crossing W with non-zero transverse velocity at $\mu = \mu_\infty$. There is then a sequence $\mu_j \to \mu_\infty$ (defined for sufficiently large j) such that $\psi_{\mu_j} \in \Pi_j$; the quantity $\delta^j(\mu_\infty - \mu_j)$ converges as $j \to \infty$ to a finite non-zero limit.*

Returning to the principal objective of this section, we see that the proof of a) p. 233 now reduced to constructing appropriately localized preimages Π_j of Π_1 under $\mathcal{T}^{j-1}$ and showing that they converge exponentially to W_s with rate δ. The following theorem asserts that this is possible; it also asserts that b) holds.

Theorem 6.2. *Let $\mathcal{T}$, ϕ, W_s, W_u, δ, Π_1, and ϕ^* be as above. Then there exists a sequence (Π_j) of codimension-one submanifolds of $\mathfrak{H}$, converging exponentially to W_s with rate δ, such that*

$$\mathcal{T}^{j-1}\Pi_j \subset \Pi_1 .$$

Moreover, if $\psi \in W_s$ and if $\mathfrak{W}$ is a sufficiently small neighborhood of ψ in $\mathfrak{H}$, then

$$\mathcal{T}^{j-1}(\Pi_j \cap \mathfrak{W}) \to \{\phi^*\} \quad \text{as} \quad j \to \infty .$$

The first step in proving this theorem is to reduce it to a statement which is local at ϕ. More precisely, we claim that the theorem as stated is true if we can find an open neighborhood $\mathfrak{V}$ of ϕ such that it is true for W_s and W_u replaced by $W_s \cap \mathfrak{V}$ and $W_u \cap \mathfrak{V}$ respectively, with the added assumption that $\Pi_1 \subset \mathfrak{V}$. Proof of this claim is straight-forward, using (notably) the transversality conditions in the definition of stable and unstable manifolds. We will sketch one part of the argument, showing that there is no loss of generality in assuming that $\Pi_1 \subset \mathfrak{V}$.

As before, we let ϕ^* denote the (first) point where W_u intersects Π_1. Since $\phi^* \in W_u$, there is an integer k and a point $\phi_k^* \in W_u \cap \mathfrak{V}$ such that $\mathcal{T}^{k-1}\phi_k^* = \phi^*$. By our definition of W_u, the tangential derivative of $\mathcal{T}^{k-1}$ along W_u at ϕ_k^* does not vanish. From this (and the implicit function theorem) it follows that, for $\mathfrak{U}'$ a sufficiently small open ball about ϕ_k^*,

$$\Pi_1' = \{\psi \in \mathfrak{U}' : \mathcal{T}^{k-1}\psi \in \Pi_1\}$$

is a smooth codimension-one submanifold of $\mathfrak{V}$ intersecting W_u transversally at ϕ_k^*. The localized version of the theorem implies the existence of a sequence of surfaces Π_j' converging exponentially to $W_s \cap \mathfrak{V}$ with rate δ and with

$$\mathcal{T}^{j-1}\Pi_j' \subset \Pi_1' .$$

We can thus take

$$\Pi_{j+1} = \Pi_{j+1}' .$$

This establishes localizability in the expanding direction. A similarly straightforward argument, which we omit, establishes localizability in the contracting direction.

Thus, we have only to prove the localized version of the theorem. We do this by choosing special coordinates in which $\mathcal{T}$ and Π_1 take particularly simple forms. The result we need is the following:

If ϕ^* is close enough to ϕ, then there exists a C^1 diffoemorphism from a set of the form $\mathscr{X}_1 \times [-1,1]$, $\mathscr{X}_1$ the unit ball in some Banach space, onto a neighborhood $\mathfrak{V}$ of ϕ such that

ϕ is the image of $(0,0)$,

$W_s \cap \mathfrak{V}$ is the image of $\mathscr{X}_1 \times \{0\}$,

$W_u \cap \mathfrak{V}$ is the image of $\{0\} \times (-1,1)$,

$\Pi_1 \cap \mathfrak{V}$ is the image of $\mathscr{X}_1 \times \{1\}$.

If we regard $x \in \mathscr{X}_1$, $y \in [-1,1]$ as coordinates for their image in $\mathfrak{V}$, then in these coordinates $\mathscr{T}$ takes the form

$$\mathscr{T}: (x,y) \to (M(x,y), \delta y),$$

where

$$\|M(x_1,y) - M(x_2,y)\| \leqq \alpha \|x_1 - x_2\|,$$

with $\alpha < 1$, and $M(0,y) = 0$.

In terms of these coordinates we can take simply

$$\Pi_j = \text{image of } \mathscr{X}_1 \times \{\delta^{-(j-1)}\}$$

and this sequence of surfaces converges exponentially to $W_s \cap \mathfrak{V}$ with rate δ. Moreover, in view of the contractivity of $M(x,y)$ in x, the diameter of $\mathscr{T}^{j-1}\Pi_j$ goes to zero as $j \to \infty$. Since this set always contains ϕ^*, we have

$$\mathscr{T}^{j-1}\Pi_j \to \{\phi^*\} \quad \text{as} \quad j \to \infty.$$

The proof of the theorem is thus reduced to proving the existence of the indicated "normal coordinates" for $\mathscr{T}$ and Π_1. We will concentrate on showing the existence of coordinates in which $\mathscr{T}$ has the desired form, since this result may be of interest in other contexts; we will then at the end sketch how the argument can be modified to bring Π_1 into normal form as well. The proof of the following theorem is based on an analogous but more complicated result proved in Collet and Eckmann [1].

Theorem 6.3. *Let $\mathscr{T}$ be a twice continuously differentiable mapping from an open set $\mathfrak{D}$ in a Banach space into the Banach space. Let ϕ be a fixed point for $\mathscr{T}$. Assume that $D\mathscr{T}(\phi)$ has a single simple eigenvalue $\delta > 1$ and that the rest of its spectrum is in the interior of the unit disk. Then there exists a C^1 diffeomorphism of $\mathscr{X}_1 \times (-1,1)$ ($\mathscr{X}_1$ denoting the open unit ball in some Banach space) onto a neighborhood $\mathfrak{V}$ of ϕ – i.e. a set of local coordinates at ϕ – such that*

$(0,0)$ represents ϕ,

$\mathscr{X}_1 \times \{0\}$ represents $W_s \cap \mathfrak{V}$,

$\{0\} \times (-1,1)$ represents $W_u \cap \mathfrak{V}$,

$\mathscr{T}$ takes the form $(x,y) \to (M(x,y), \delta y)$,

where

$$M(0,y) = 0; \quad \|D_x M(x,y)\| \leqq \alpha < 1 \quad \text{for} \quad (x,y) \in \mathscr{X}_1 \times (-1,1).$$

We emphasize

1. In these coordinates the action of $\mathscr{T}$ in the y direction has been made *exactly* linear.

2. We do not require – and it is generally not possible – that $\mathcal{T}$ be linearized smoothly in the x direction as well.

The first step in the proof is to introduce coordinates which are approximately right. We do this in a sequence of steps:

– Make a translation to put ϕ at the origin.

– Write the Banach space as the direct sum of the one-dimensional eigenspace corresponding to the eigenvalue δ (y direction) and the complementary spectral subspace (x direction).

– Carry out an x-dependent translation in the y direction to bring the stable manifold to the surface $\{y=0\}$.

– Carry out a y-dependent translation in the x direction to bring the unstable manifold to the line segment $x=0$.

– Reparametrize the y coordinate in such a way that the action of $\mathcal{T}$ *on the unstable manifold* becomes exactly multiplication by δ. The possibility of doing this follows from a trivial case of the Sternberg Linearization Theorem. (See, for example, Hartmann [5].)

Thus we can assume that:

1. The domain of $\mathcal{T}$ is $\mathcal{X}_1 \times(-1,1)$ and $\mathcal{T}(x,y)=(M(x,y), \delta y+N(x,y))$ where M, N are continuously differentiable and $M(0,y)=0$; $N(x,0)=0$; $N(0,y)=0$.

In the course of the argument, we will need to assume that the nonlinear terms in $\mathcal{T}$ are small. This can be accomplished by *magnification*, i.e. by replacing (x,y) with new coordinates $x'=\lambda x$, $y'=\lambda y$ with λ large (and restricting the domain to the set $\|x'\| \leqq 1$, $|y'| \leqq 1$). This transformation leaves the linear terms unchanged but shrinks the nonlinear terms by a factor of at least $1/\lambda$. In this way (and possibly also renorming the space $\mathcal{X}$) we can arrange that, for some $\beta>0$,

2. $$\left.\begin{aligned} &\|D_x M(x,y)\| \leqq \alpha<1 \\ &1<\beta \leqq \frac{\delta y+N(x,y)}{y} \end{aligned}\right\} \quad \text{for} \quad \|x\|<1, \quad |y|<1.$$

It will also be convenient to assume that $\mathcal{T}(x,y)$ is defined and well behaved for *all* y. To extend it, we choose a smooth cut-off function $\varrho(y)$,

$$0 \leqq \varrho(y) \leqq 1,$$

with

$$\varrho(y)=1 \quad \text{for} \quad |y| \leqq \frac{1}{\beta}$$

$$\varrho(y)=0 \quad \text{for} \quad |y| \geqq 1$$

and modify $\mathcal{T}$ to the mapping

$$\begin{aligned} (x,y) &\to (\varrho(y)M(x,y), \delta y+\varrho(y)N(x,y)), && |y|<1 \\ &\to (0, \delta y), && |y| \geqq 1. \end{aligned}$$

The modified $\mathcal{T}$ agrees with the original $\mathcal{T}$ on all (x,y) with $\mathcal{T}(x,y) \in \mathcal{X}_1 \times[-1,1]$; it maps $\mathcal{X}_1 \times \mathbb{R}$ into itself and satisfies the inequalities (2.) for all y. We will work, from now on, with this modified $\mathcal{T}$.

To prove the theorem, it suffices to find a continuously differentiable function z such that

$$z \circ \mathscr{T} = \delta \cdot z, \qquad z(0, y) = y.$$

The inverse function theorem then assures us that (x, z) is a set of local coordinates at $(0,0)$ and $\mathscr{T}$ evidently has the desired form in these coordinates. To construct z, we let $y_n(x, y)$ denote the y coordinate of $\mathscr{T}^n(x, y)$, and we put

$$z_n(x, y) = \delta^{-n} y_n(x, y).$$

It is immediate that $z_n(x, y)$ is continuously differentiable and that $z_n(0, y) = 0$. Also,

$$z_n \circ \mathscr{T} = \delta \cdot z_{n+1}$$

so if

$$\lim_{n \to \infty} z_n \equiv z$$

exists, we have immediately

$$z \circ \mathscr{T} = \delta \cdot z$$

as desired. Evidently, $z_n(x, 0) = 0$, so the limit exists in this case. On the other hand, if $y \neq 0$, then, since

$$|y_n| \geqq \beta^n |y|,$$

$|y_n|$ is eventually larger than one. But because of the way $\mathscr{T}$ is cut off, $y_{n+1} = \delta y_n$ if $|y_n| \geqq 1$, and so $z_{n+1}(x, y) = z_n(x, y)$. Hence, for all x, y

$$z(x, y) = \lim_{n \to \infty} z_n(x, y)$$

exists, and the only problem is to prove that it is continuously differentiable. Note, incidentally, that if $y \neq 0$ and n is sufficiently large, $z = z_n$ on a neighborhood of (x, y) and so z is continuously differentiable on that neighborhood; the only place where differentiability could fail is for $y = 0$.

We will write

$$z_n(x, y) = y + r_n(x, y);$$

then a simple computation using

$$z_{n+1} = \frac{1}{\delta} z_n \circ \mathscr{T}$$

and the expression (1.) for $\mathscr{T}$ gives:

$$r_{n+1}(x, y) = \frac{1}{\delta} N(x, y) + \frac{1}{\delta} r_n \circ \mathscr{T}(x, y).$$

Defining a linear operator L by

$$Ls(x, y) = \frac{1}{\delta} s \circ \mathscr{T}(x, y)$$

we see that

$$r_{n+1} = \frac{1}{\delta} \sum_{j=0}^{n} L^j N,$$

so, if we can find a normed space containing N on which L is a contraction, we can write

$$z - y = \frac{1}{\delta} \sum_{j=0}^{\infty} L^j N.$$

We will use the norm

$$|||s||| = \max \left\{ \sup_{x,y} \frac{|s(x,y)|}{\|x\|}, \sup_{x,y} \frac{|D_y s(x,y)|}{\|x\|}, \sup_{x,y} \|D_x s(x,y)\| \right\}$$

on the space of those continuously differentiable functions vanishing for $x=0$ for which the norm is finite. (It is to guarantee that N belongs to this space that we have to assume that $\mathcal{T}$ is twice continuously differentiable. We could make do with less, but simple continuous differentiability does not seem to be enough.)

The proof that L is a contraction in this norm is straightforward; we will describe explicitly only the most sensitive of the estimates,

$$D_y(Ls)(x,y) = \frac{1}{\delta}(D_x s)(M(x,y), \delta y + N(x,y)) \cdot D_y M(x,y)$$

$$+ \frac{1}{\delta}(D_y s)(M(x,y), \delta y + N(x,y))(\delta + D_y N(x,y)),$$

$$\frac{1}{\delta} \frac{|D_x s(M(x,y), \delta y + N(x,y)) D_y M(x,y)|}{\|x\|} \leqq \frac{1}{\delta} |||s||| \sup_{x,y} \|D_x D_y M(x,y)\|.$$

Now $D_x D_y M(x,y)$ comes from the nonlinear terms in $\mathcal{T}$ which can be made small by magnification, so we can estimate this expression by an arbitrarily small multiple of $|||s|||$. From

$$M(0,y) = 0; \qquad \|D_x M(x,y)\| \leqq \alpha$$

it follows that

$$\|M(x,y)\| \leqq \alpha \cdot \|x\|$$

and thus

$$\left| \frac{1}{\delta} \frac{D_y s(M(x,y), \delta y + N(x,y))(\delta + D_y N(x,y))}{\|x\|} \right| \leqq \alpha |||s||| \left(1 + \frac{1}{\delta} \sup_{x,y} \|D_y N(x,y)\| \right).$$

Again, the last term can be made arbitrarily small by magnification. The final result is, then, that we have an estimate

$$\sup_{x,y} \left\{ \frac{|D_y Ls(x,y)|}{\|x\|} \right\} \leqq |||s||| \times (\alpha + \text{something small}).$$

The other two terms are estimated similarly.

Hence, L is a contraction so the series

$$\frac{1}{\delta}\sum_{j=0}^{\infty} L^j N$$

converges in this normed space; the sum is $z-y$ so z is continuously differentiable.

Remark. Similar estimates show that if $\alpha\cdot\delta^{r-1}<1$ then z is C^r (assuming that $\mathscr{T}$ is C^{r+1}).

On the other hand, no matter how smooth $\mathscr{T}$ is, it is usually not possible to find a C^r solution to $z\circ\mathscr{T}=\delta\cdot z$, $z(0,y)=y$ if one of $\frac{1}{\delta},\frac{1}{\delta^2},\ldots,\frac{1}{\delta^{r-2}}$ can be written as a product of points of the spectrum of $D_xM(0,0)$.

Finally, we have to show how to modify the above argument to bring a codimension-one surface crossing the unstable manifold transversally away from the fixed point simultaneously into the desired normal form, a flat horizontal surface. We start as before, choosing coordinates in a neighborhood of ϕ in which $\mathscr{T}$ takes the form (1.) and (2.) holds. Let $\hat{\Pi}_1$ intersect the unstable manifold inside this coordinate neighborhood; then a part of $\hat{\Pi}_1$ near the intersection point can be represented in our coordinate system as the graph of a function $x\to\hat{\Pi}_1(x)\in(-1,1)$ with $\hat{\Pi}_1$ defined in a neighborhood of 0. By magnifying in the x dimension we can assume that $\hat{\Pi}_1$ is defined and non-zero on all of $\mathscr{X}_1$.

Now define a new y coordinate by

$$y_{\text{new}}=\frac{y_{\text{old}}}{\hat{\Pi}_1(x)}.$$

In terms of the new coordinates, (1.) still holds; if (2.) doesn't then it can be made to hold by a further magnification in the x direction. We have thus arranged so that (1.) and (2.) hold, and, in addition, Π_1 corresponds to the surface $\{y=1\}$. We then proceed to cut off and extend N, M as indicated. Note that, because of the way we have done the cutting off, the inverse image under $\mathscr{T}$ of any set in $\mathscr{X}_1\times[-1,1]$ is exactly the same as it was before the modification of $\mathscr{T}$ but, on the other hand,

$$\mathscr{T}^n(x,1)=(0,\delta^n).$$

From this last equation, it follows that

$$z(x,1)=1$$

i.e. that Π_1 corresponds to the surface $\{z=1\}$.

It still has to be shown that z is continuously differentiable. To prove that, one needs to know that

$$\sup_{x,y}\|D_xD_yM(x,y)\|,\quad \sup_{x,y}\|D_yN(x,y)\|,\quad \sup_{x,y}\|D_xN(x,y)\|$$

are all small. This would normally be arranged by magnifying sufficiently *before* cutting off in the y direction. We note, however, that the magnification can just as well be done *after* cutting off, and that this does not spoil the flatness of Π_1 in the z coordinate.

7. The Global Unstable Manifold

The preceding section shows that the universality of the rate of period doubling can be understood if the unstable manifold for the fixed point ϕ_ε crosses the surface $\Sigma_1 = \{\psi : \psi(1) = 0\}$ transversally. In this section, we estimate the global structure of the unstable manifold. In essence, we show that it remains close to the line segment $\{\psi(x) = 1 - (1+a)|x|^{1+\varepsilon} : -1 < a < 1\}$ throughout the full length of this segment, provided that ε is small enough.

We need yet another realization of the action of $\mathscr{T}$ in convenient coordinates. Recall that we showed that if we write

$$\psi(x) = f(|x|^{1+\varepsilon}); \quad f(t) = 1 - th(t);$$

then in terms of h, $\mathscr{T}$ takes the form

$$h \to \eta_2 \{h(1 - at\eta_2) + \Delta_1 h(1 - at\eta_2)\}, \tag{7.1}$$

where the notation is as in Sect. 4. Deviating from the notation of that section, we will write

$$h(t) = 1 + a + (t-1)g(t) \tag{7.2}$$

and determine the form of $\mathscr{T}$ expressed in terms of the coordinates $a (\in (-1, 1))$ and $g (\in \mathfrak{H})$. Note that

$$1 + a = h(1)$$
$$g(t) = (\Delta_1 h)(t)$$

so it is easy, in principle, to read off this form from (7.1). We will need, however, to look with some care at the expression for η_2. Recall

$$\eta_1 = a^\varepsilon h(a^{1+\varepsilon} t)$$
$$(1 - at\eta_1)^{1+\varepsilon} = 1 - at\eta_2 .$$

A straightforward computation shows that

$$(1-z)^{1+\varepsilon} = 1 - z(1 + \varepsilon s(z, \varepsilon)),$$

where $s(z, \varepsilon)$ (which can easily be written explicitly) is analytic for all ε and all $z \notin [1, \infty)$. Hence

$$\eta_2 = \eta_1 (1 + \varepsilon s(-at\eta_1, \varepsilon)).$$

Also,

$$\begin{aligned} h(1 - at\eta_2) + \Delta_1 h(1 - at\eta_2) &= 1 + a - at\eta_2 g(1 - at\eta_2) + g(1 - at\eta_2) \\ &= 1 + a + (1 - at\eta_2) g(1 - at\eta_2). \end{aligned}$$

Thus the transformed h becomes

$$a^\varepsilon (1 + a + (a^{1+\varepsilon} t - 1) g(a^{1+\varepsilon} t))(1 + \varepsilon s(-at\eta_1, \varepsilon))(1 + a + (1 - at\eta_2) g(1 - at\eta_2)). \tag{7.3}$$

We will write $A(\varepsilon, a, g)$ and $G(\varepsilon, a, g)$ for the a, g components of the representation of (7.3) in the form (7.2). Thus, to get $A(\varepsilon, a, g)$ we have simply to insert $t = 1$ in (7.3)

and then subtract 1. Doing this, and grouping terms in a straightforward way, we find

$$A = a^{\varepsilon}(1+a)^2 - 1 + A_1 + \varepsilon A_2 ,$$

where $A_1 = 0$ for $g = 0$ and where A_1, A_2 are both "regular". The meaning of "regular" has to be specified with a bit of care. The formulas are full of factors of a^{ε} whose a-derivatives become infinite as a approaches zero. The idea is that these are the only singularities for small ε, a, g. One way to formulate this is indicated in the following proposition.

Proposition 7.1. *We can write*

$$\begin{aligned} A &= a^{\varepsilon}(1+a)^2 - 1 + A_1(\varepsilon, a, a^{\varepsilon}, g) + \varepsilon A_2(\varepsilon, a, a^{\varepsilon}, g) \\ G &= aG_1(\varepsilon, a, a^{\varepsilon}, g) + a\varepsilon G_2(\varepsilon, a, a^{\varepsilon}, g) \end{aligned} \tag{7.4}$$

(A_1 and A_2 take values in $\mathbb{R}$; G_1 and G_2 in $\mathfrak{H}(\Omega)$) where A_1, A_2, G_1, G_2 are all infinitely differentiable with bounded derivatives on

$$(0, \varepsilon_0) \times (0, a_0) \times (0, 1) \times \{g \in \mathfrak{H} : \|g\| < g_0\}$$

(for sufficiently small ε_0, a_0, g_0) and where A_1 and G_1 vanish identically for $g = 0$.

We have already indicated the proof for A. To get the formula for G, apply Δ_1 to (7.3) using the "product rule"

$$\Delta_1 g_1 g_2 = g_1 \Delta_1 g_2 + g_2(1) \Delta_1 g_1 ;$$

group the contributions from differencing in the first and third factors to form $a \cdot G_1$ and the contribution from the second factor to form $a \cdot \varepsilon \cdot G_2$. To extract the indicated explicit factors of a, make repeated use of the "chain rule"

$$(\Delta_{z_0}(g_1 \circ g_2))(z) = (\Delta_{g_2(z_0)} g_1)(g_2(z)) \Delta_{z_0} g_2(z)$$

and observe that every t in (7.3) is accompanied by a multiplicative factor of a.

Corollary 7.2. *There exist constants B_1, B_2 such that*

$$\|G\| \leqq a \cdot B_1 \|g\| + a\varepsilon B_2 \qquad (0 < a < a_0,\ 0 < \varepsilon < \varepsilon_0,\ \|g\| < g_0). \tag{7.5}$$

In particular, at the fixed point, $a = \lambda_{\varepsilon}$ and we write $g_{\varepsilon}^{(0)}$ for the corresponding g. Then

$$\|g_{\varepsilon}^{(0)}\| \leqq \lambda_{\varepsilon} B_1 \|g_{\varepsilon}^{(0)}\| + \lambda_{\varepsilon} \varepsilon B_2$$

from which it follows at once that

$$\|g_{\varepsilon}^{(0)}\| \leqq \frac{\varepsilon \lambda_{\varepsilon} B_2}{(1 - \lambda_{\varepsilon} B_1)} = O(\varepsilon \lambda_{\varepsilon})$$

for small ε. Since $\lambda_{\varepsilon} = O(-\varepsilon \log \varepsilon)$ for small ε, this estimate gives exactly the estimate

$$\phi_{\varepsilon}(x) = 1 - (1 + \lambda_{\varepsilon})|x|^{1+\varepsilon} + O(-\varepsilon^2 \log \varepsilon)$$

announced in Sect. 2.

Corollary 7.3. *If*

$$a(B_1 + B_2) \leqq 1,$$

and if

$$\|g\| \leqq \varepsilon,$$

then

$$\|G\| \leqq \varepsilon.$$

In other words, $\mathscr{T}$ can push a point out of the cylinder $\{(a, g): 0 < a < a_1, \|g\| \leqq \varepsilon\}$, [with a_1 the smaller of a_0, $1/(B_1 + B_2)$], only through the ends.

We have noted already that singularities appear when we differentiate a^ε for a near zero. However, since

$$\frac{d}{da} a^\varepsilon = \frac{\varepsilon}{a} a^\varepsilon$$

the singularities are not much in evidence in the first derivative for $a \geqq \varepsilon$. Since the fixed point occurs at $\lambda_\varepsilon \gg \varepsilon$, we don't have to worry about the singularities when looking at first derivatives near and beyond the fixed point. Thus:

Corollary 7.4. *There exist constants B_3, B_4, B_5 such that*

$$\left\| \frac{d}{da} G(a, g_a) \right\| \leqq B_3 \|g_a\| + B_4 \varepsilon + B_5 a \left\| \frac{dg}{da} \right\|$$

for any differentiable mapping $a \to g_a$ provided

$$\varepsilon < a < a_0; \qquad \|g_a\| \leqq g_0.$$

Proof. The B_3 term comes from the explicit aa^ε dependence in aG_1; the B_4 term from the explicit aa^ε dependence in aG_2; the B_5 term from the g dependence of the sum.

Corollary 7.5. *Consider a curve given as the graph of a function $a \to g_a$, defined in a subinterval of (ε, a_0). Assume*

$$\frac{dA(a, g_a)}{da} \geqq 1. \tag{7.6}$$

Then the image of this curve under the action of $\mathscr{T}$ is again representable as the graph of a function

$$\hat{a} \to \hat{g}_{\hat{a}}$$

and

$$\left| \frac{d\hat{g}}{d\hat{a}} A(\varepsilon, a, g) \right| \leqq B_3 \|g_a\| + B_4 \varepsilon + B_5 a \left\| \frac{dg_a}{da} \right\|. \tag{7.7}$$

We can apply this corollary in particular to a (possibly very small) local unstable manifold. Such a manifold can be represented as the graph of a mapping defined

on a small interval I about λ_ε; condition (7.6) is satisfied and we can assume

$$N = \sup_{a \in I} \|g_a\| \ll \varepsilon .$$

Put

$$M = \sup_{a \in I} \left\| \frac{dg}{da} \right\| .$$

Since the local unstable manifold is mapped onto itself by $\mathscr{T}$, we get

$$M \leqq B_3 N + B_4 \varepsilon + 2B_5 \lambda_\varepsilon M$$

or, for small ε,

$$M \leqq 2B_4 \varepsilon .$$

Consider now a manifold specified by

$$b \to g_b$$

defined on a subinterval of (ε, a), and satisfying

$$\|g_a\| \leqq \varepsilon ; \qquad \left\| \frac{dg_a}{da} \right\| \leqq 2(B_3 + B_4)\varepsilon . \tag{7.8}$$

From (7.4) it is easy to check that there exists a constant B_6 such that, under these hypotheses,

$$\frac{dA(a, g_a)}{da} \geqq 2 - B_6 \varepsilon$$

(and so $\geqq 1$ for small ε); then from (7.7) the image curve satisfies

$$\left\| \frac{d\hat{g}}{d\hat{a}} \right\| \leqq (B_3 + B_4)\varepsilon + B_5 a 2(B_3 + B_4)\varepsilon$$

so if

$$a < a_0 ; \qquad a(B_1 + B_2) < 1 ; \qquad 2aB_5 < 1$$

we get again

$$\left\| \frac{dg}{da} \right\| \leqq 2(B_3 + B_4)\varepsilon .$$

Summarizing and simplifying the notation, we get:

Proposition 7.6. *There exist constants $\varepsilon_1 > 0$; $a_1 > 0$, B_7 such that, if $0 < \varepsilon < \varepsilon_1$, and if we have a curve in $\mathscr{P}_\varepsilon$ specified by*

$$a \to g_a$$

defined on a subinterval of (ε, a_1) and satisfying

$$\|g_a\| \leqq \varepsilon ; \qquad \left\| \frac{dg_a}{da} \right\| \leqq B_7 \varepsilon \tag{7.9}$$

then

$$\frac{d}{da}A(\varepsilon, a, g_a) \geqq 1.5, \tag{7.10}$$

so the image curve also admits a representation as

$$a \to \hat{g}_a$$

and the bounds (7.9) *hold with* g *replaced by* $\hat{g}$. *A sufficiently small local unstable manifold satisfies these hypotheses.*

Now let $\varepsilon < \varepsilon_1$, a small local unstable manifold, and apply $\mathscr{T}$ to it repeatedly, throwing away at each step the part of the image curve with a outside of (ε, a_1). In view of (7.10), we obtain after a finite number of iterations a curve $a \to g_a^*$ defined on all of (ε, a_1) contained in the unstable manifold and satisfying the inequalities (7.9).

It remains to show that the unstable manifold in the form $a \to g_a^*$ can be extended to all $a \in (-1, 1)$. We will consider only the problem of extending it to values of $a \geqq a_1$; the extension to $a \in (-1, \varepsilon)$ is similar but easier.

If we write

$$f_a^*(t) = 1 - t(1 + a + (t-1)g_a^*), \qquad \phi_a^*(x) = f_a^*(|x|^{1+\varepsilon}),$$

it will suffice, by the general theory of Sect. 3, to show that g_a^* can be extended to a value of a such that

$$\phi_a^*(a) = a$$

(i.e. such that $(\phi_a^*)^4(0)$ is the fixed point of ϕ_a^* in $[0, 1]$); then one more iteration of $\mathscr{T}$ gives values of a running all the way to 1. By continuity, it will suffice to extend it to a value of a such that

$$\phi_a^*(a) < a. \tag{7.11}$$

As above, we will write ϕ^* in the form

$$\phi_a^*(x) = f_a^*(|x|^{1+\varepsilon}); \qquad f_a^*(t) = 1 - t[1 + a + (t-1)g(t)].$$

Ignoring g and terms of order ε, we get

$$\phi_a^*(a) \approx 1 - a(1 + a);$$

from this it is easy to see that (7.11) will hold provided $a(2 + a) > 1$, i.e. $a > \sqrt{2} - 1$ and provided $\|g\|$ and ε are small enough.

Next we will argue that we can take $a_0 > \sqrt{2} - 1$ in Proposition 7.1. To do this we need, for the first time, to be careful about our choice of the domain Ω. Examination of the proof of Proposition 7.1 shows that the only limitations on ε_0, a_0, and g_0 are that

1) $a^{1+\varepsilon}\bar{\Omega} \subset \Omega$

2) $f(a^{1+\varepsilon}\bar{\Omega}) \cap (-\infty, 0] = \emptyset$

3) $[f(a^{1+\varepsilon}\bar{\Omega})]^{1+\varepsilon} \subset \Omega$

for $0<a<a_0$; $0<\varepsilon<\varepsilon_0$; $\|g\|<g_0$. If these three conditions are satisfied with $g=0$ and $\varepsilon=0$, it follows by continuity (and compactness) that they will be satisfied for all sufficiently small ε and g. Thus, a_0 is limited by:

$$1') \quad a\bar{\Omega}\subset\Omega$$

$$2') \quad (1-a(1+a)\bar{\Omega})\subset\Omega\backslash(-\infty,0]$$

for $0<a<a_0$. It is easy to check that these conditions hold, for example, for $a_0=1/2>\sqrt{2}-1$ and for Ω the open disk with radius 3/4 and center 1/2.

We can thus assume that estimates (7.5), (7.7), and (7.10) hold for $\varepsilon<a<1/2$, and that the unstable manifold has been extended to a curve of the form $a\to g_a^*$ satisfying (7.9) and defined on (ε, a_1) where a_1 may be taken to be independent of ε. Because of (7.10), a number of iterations of $\mathscr{T}$ which is bounded uniformly in ε for small ε suffices to extend the unstable manifold to a curve of the above form defined on $(\varepsilon, 1/2)$. The bounds (7.9) are no longer necessarily propagated by application of $\mathscr{T}$, but each such application worsens them by only a finite amount [because of (7.5), (7.7)]. In this way (and also extending similarly in the direction $a\to 0$) we get:

Proposition 7.7. *Let Ω be the open disk of radius 3/4 and center 1/2. There exist constants B_8, B_9, and $\varepsilon_0>0$, such that for $0<\varepsilon<\varepsilon_0$, the unstable manifold for $\mathscr{T}$ in $\mathscr{P}_\varepsilon$ contains a curve*

$$a\to\phi_a^*: \phi_a^*(x)=1-(1+a)|x|^{1+\varepsilon}-|x|^{1+\varepsilon}(1-|x|^{1+\varepsilon})g_a^*(|x|^{1+\varepsilon})$$

defined on $a\in[0,\tilde{a}]$ and satisfying:

$$\|g_a^*\|\leqq B_8\varepsilon\,; \qquad \left\|\frac{dg_a^*}{da}\right\|\leqq B_9\varepsilon\,; \qquad \phi_{\tilde{a}}^*(\tilde{a})=\tilde{a}\,.$$

Moreover if $A(a)=-(\mathscr{T}\phi_a^*)(1)$, then $\dfrac{dA}{da}>1.5$ on $[0,\tilde{a}]$.

Remarks. The estimates developed in this section show that any ψ near ϕ_ε and not on the stable manifold will be driven out of $\mathfrak{D}(\mathscr{T})$ by a finite number of iterations of $\mathscr{T}$. This justifies the restriction in Sect. 6 to a non-recurrent fixed point. It also permits us to clarify the uniqueness of the μ_j's. Consider the cylinder in $\mathscr{P}_\varepsilon$ corresponding to

$$0<a<1\,; \qquad \|g\|\leqq g_1\,.$$

For fixed small ε, and sufficiently small g_1, the stable manifold cuts across this cylinder and thus divides it into two parts. We will refer to the part on the side of $a=0$ as "above" the unstable manifold and the other part as "below" it (see Fig. 4). The surfaces Σ_j further divide the part of the cylinder above the stable manifold into slabs. If ψ lies between Σ_j and Σ_{j+1}, then $\psi_j\equiv\mathscr{T}^j\psi$ is defined and $(\psi_j)(1)>0$. Thus, ψ_j maps all of $[-1,1]$ into the interval $[\psi_j(1),1]$, which does not contain 0, and hence $\psi_j^p(0)\neq 0$ for $p=1,2,\ldots$. But ψ_j differs from ψ^{2^j} only by a scale factor, so $\psi^{p2^j}(0)\neq 0$ for all p. This implies that $\psi^p(0)\neq 0$, so ψ is *not* superstable. Thus: The

only superstable ψ's lying in the part of the cylinder above W_s are those on the surfaces $\Sigma_j, j=1,2,\ldots$. If ψ_μ is a parametrized curve crossing W_s from above when $\mu=\mu_\infty$, at a point inside the cylinder, with non-zero vertical velocity, then for sufficiently large j, the μ_j are uniquely determined by the conditions

$$\psi_{\mu_j} \text{ is superstable of period } 2^j; \quad \mu_j<\mu_\infty; \quad \mu_\infty-\mu_j \text{ is small.}$$

On the other hand, it is not hard to see that, for large j, there are very many values of μ, larger than but near to μ_∞, where ψ_μ is superstable with period 2^j.

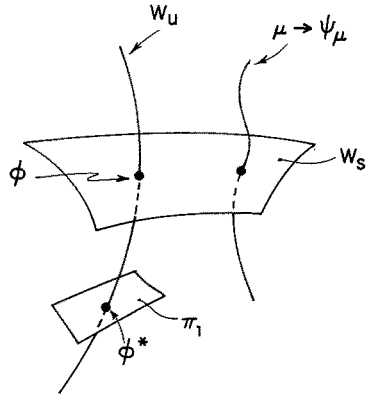


Fig. 4

In another direction: It is easy to verify (using the implicit function theorem) that, for small ε, g there is a uniquely determined $a=\hat{a}(g)$ such that ψ corresponding to (a,g) is superstable of period 3. Moreover $g\to\hat{a}(g)$ is smooth and defines a codimension-one surface crossing W_u transversally. Call this surface $\hat{\Sigma}_1$, and apply the theory of the preceding section to show that ots successive inverse images

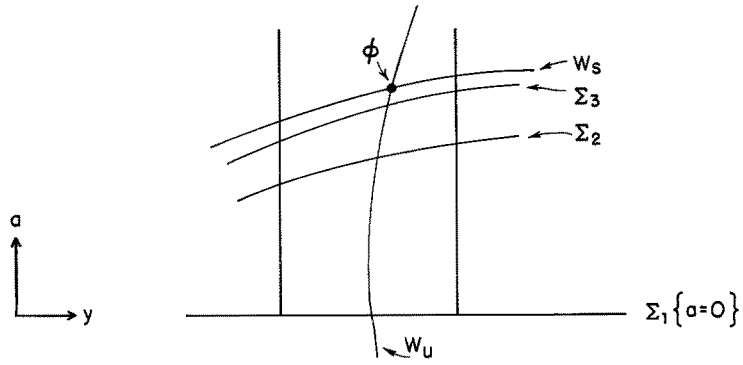


Fig. 5

$\hat{\Sigma}_2, \hat{\Sigma}_3, \ldots$ (under $\mathcal{T}$) converge to W_s exponentially with rate δ. If $\mu\to\psi_\mu$ is a parametrized curve as before, there exists a sequence $\hat{\mu}_j$, converging down to μ_∞, such that $\psi_{\hat{\mu}_j}\in\hat{\Sigma}_j$, i.e. $\psi_{\hat{\mu}_j}$ is superstable of period $3\cdot 2^{j-1}$. Moreover, just as before

$$(\hat{\mu}_j-\mu_\infty)\delta^j$$

converges to a finite non-zero limit. A warning, however: There is another sequence, say $\hat{\hat{\mu}}_j$, with $\psi_{\hat{\hat{\mu}}_j}$ superstable of period $3\cdot 2^{j-1}$ and $\hat{\mu}_{j-2}>\hat{\hat{\mu}}_j>\hat{\mu}_{j-1}$. In

fact, there are infinitely many more distinct, interleaved, sequences of periods $3 \cdot 2^{j-1}$.

Similarly, the equation

$$\psi(a) = a \quad [\text{i.e. } \psi^3(0) = \psi^4(0)]$$

defines a surface, say $\tilde{\Sigma}_1$, of codimension one crossing W_u transversally. A general result of Misiurewicz implies that any $\psi \in \tilde{\Sigma}_1$ sufficiently near to the unstable manifold admits an absolutely continuous invariant measure. It is easy to see, however, that if $\mathscr{T}\psi$ admits an absolutely continuous invariant measure, then so does ψ, so, applying the machinery described above, we see that, for each $\mu \to \psi_\mu$ as above, there is a sequence $\tilde{\mu}_j$ converging to μ_∞ such that each $\psi_{\tilde{\mu}_j}$ admits an absolutely continuous invariant measure and such that

$$\delta^j(\tilde{\mu}_j - \mu_\infty)$$

approaches a finite non-zero limit. Many other such examples could be considered, e.g., take for the initial surface the set of ψ's such that $\psi'(x_0) = -1$, where x_0 denotes the unique fixed point of ψ in $[0, 1]$. The μ_j's in this case will correspond to bifurcation points, where the orbit of period 2^{j-1} becomes unstable and the orbit of period 2^j appears.

8. Attracting Cantor Sets

Let $\psi \in \mathfrak{D}(\mathscr{T})$. As in Sect. 1, we write

$$a = -\psi(1); \qquad b = \psi(a);$$

and we will also write

$$c = \psi(b),$$

and we will assume $c \geqq 0$. [This would follow automatically if $\psi \in \mathfrak{D}(\mathscr{T}^2)$.] Since ψ maps $[-1, 1]$ into $[-a, 1]$, we may as well restrict its definition to the interval $[-a, 1]$. We have seen that ψ maps $[-a, a]$ onto $[b, 1]$ and it evidently maps $[b, 1]$ in a one-to-one fashion onto $[-a, c]$. Thus, the set

$$[-a, c] \cup [b, 1]$$

is mapped onto itself by ψ. Note that this invariant set is constructed out of the original interval $[-a, 1]$ by deleting an open subinterval (c, b) in the middle, i.e. as in the first step of constructing a Cantor set. Note also that $[-a, c]$ is mapped onto itself by $\psi \circ \psi$ and that $\mathscr{T}\psi$ is obtained from the restriction of $\psi \circ \psi$ to $[-a, c]$ by a linear change of variable $x \to -ax$. Observe, finally, also that if $K \subset [-a, c]$ is mapped onto itself by $\psi \circ \psi$, then

$$J = K \cup (\psi^{-1} K \cap [b, 1])$$

is mapped onto itself by $\mathscr{T}\psi$.

If $\mathscr{T}\psi$ is also in $\mathfrak{D}(\mathscr{T}^2)$, we can apply the same operation to $\mathscr{T}\psi$ and thus obtain an invariant set for ψ by deleting an open subinterval from the middle of each of $[-a, c]$ and $[b, 1]$. Continuing, if $\psi \in W_s$ and hence $\psi \in \mathfrak{D}(\mathscr{T}^j)$ for all j, we

can repeat this operation infinitely often and so obtain an invariant Cantor set for ψ. In this section, we will analyze the construction of this Cantor set in more detail. In particular, we determine the action of ψ on the Cantor set, show that orbits of ψ which converge to the Cantor set have simple statistical properties, and show that if ψ is near enough to the fixed point all but countably many orbits do indeed converge to the Cantor set.

We deal first with some combinatorial aspects of the construction of the Cantor set. For ψ as above we write

$$J_1^{(1)}=[b,1]\,; \quad J_2^{(1)}=[-a,c]\,; \quad J^{(1)}=J_1^{(1)}\cup J_2^{(1)}.$$

Observe that ψ maps $J_1^{(1)}$ onto $J_2^{(1)}$ and $J_2^{(1)}$ back onto $J_1^{(1)}$; also that the end-points of $J_1^{(1)}$ are $\psi(0)$ and $\psi^3(0)$ while those of $J_2^{(1)}$ are $\psi^2(0)$ and $\psi^4(0)$. The following proposition is proved in a straightforward way by induction, using the remarks already made.

Proposition 8.1. *Let $\psi\in\mathfrak{D}(\mathscr{T}^n)$ and assume $(\mathscr{T}^n\psi)(1)\leqq 0$. There then exists a decreasing sequence of closed sets*

$$J^{(1)}\supset J^{(2)}\supset J^{(3)}\supset \ldots \supset J^{(n)}$$

with the following properties:

1. *Each $J^{(i)}$ is mapped onto itself by ψ.*
2. *Each $J^{(i)}$ is a union of 2^i disjoint closed intervals which we can label $J_1^{(i)},\ldots,J_{2^i}^{(i)}$ in such a way that ψ maps $J_j^{(i)}$ onto $J_{j+1}^{(i)}$, where addition is understood modulo 2^i. The interval $J_{2^i}^{(i)}$ contains 0.*
3. *$J^{(i+1)}$ is constructed by removing an open subinterval from the middle of each of the intervals $J_j^{(i)}$. The resulting two intervals are labelled $J_j^{(i+1)}$ and $J_{j+2^i}^{(i+1)}$. The interval to be removed and the labelling are determined as follows: ψ^{2^i} restricted to $J_{2^i}^{(i)}$ differs from $\psi_i=\mathscr{T}^i\psi$ only by a scale factor. Remove from $J_{2^i}^{(i)}$ the interval corresponding, under this scaling, to $(c(\psi_i),b(\psi_i))$. Call the remaining subinterval which contains zero $J_{2^{i+1}}^{(i+1)}$ and the other $J_{2^i}^{(i+1)}$. (Note that ψ^{2^i} interchanges these intervals.) Then put $J_j^{(i+1)}=\psi^j[J_{2^{i+1}}^{(i+1)}]$, $j=1,2,\ldots,2^{i+1}-1$.*
4. *The end points of $J_j^{(i)}$ are $\psi^j(0)$ and $\psi^{j+2^i}(0)$.*

Corollary 8.2. *If $\psi\in W_s$ (so $\psi\in\mathfrak{D}(\mathscr{T}^n)$ for all n) then ψ admits an invariant Cantor set $J=\bigcap_i J^{(i)}$. This Cantor set is homeomorphic to $\{0,1\}^N$, with a correspondence such that $x\leftrightarrow(i_1,i_2,\ldots)\in J_j^{(i)}$ if and only if $j=i_1+2i_2+2^2i_3+\ldots+2^{j-1}i_j$.*

In this representation $\psi|_J$ takes the following form

$$\psi:(1,1,1,\ldots)\to(0,0,0,\ldots)$$

$$\psi:(1,1,\ldots,1,0,i_{n+1},\ldots)\to(0,0,0,\ldots,0,1,i_{n+1},\ldots).$$

The orbit of each element of J is dense in J; ψ is invertible on J. Heuristically, we can think of the sequence $(i_1,i_2,\ldots)$ as the binary representation of a (usually infinite) integer

$$k=\sum_{\ell=1}^{\infty} i_\ell 2^{\ell-1}.$$

In terms of k, the action of ψ is simply $k\to k+1$.

Next we continue to assume that $\psi \in W_s$ and we look at the ergodic theory of the action of ψ on J. Let ν be the unique probability measure on J defined by

$$\nu(J_j^{(i)}) = 2^{-i} \quad \text{for all} \quad i, j$$

(i.e. ν assigns equal weight to each of the intervals making up $J^{(i)}$.) Since

$$\psi^{-1}(J_j^{(i)}) \cap J^{(i)} = J_{j-1}^{(i)},$$

the uniqueness of ν implies that it is invariant under the action of ψ; on the other hand, this same equation shows that any ψ-invariant probability measure assigning measure zero to the complement of J must assign equal measure to each $J_j^{(i)}$ (i fixed but arbitrary, $j = 0, 1, 2, \ldots, 2^i - 1$) and hence must coincide with ν.

Proposition 8.3. *ν is the only ψ-invariant probability measure on J. If $x \in [-1, 1]$ has the property that $\psi^n(x)$ approaches J as $n \to \infty$ and if f is any continuous function on $[-1, 1]$ then*

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} f \circ \psi^n(x) = \int f d\nu.$$

The abstract dynamical system (ν, ψ) is ergodic but not weak mixing.

Proof. The first statement has already been proved. To prove the second, suppose it is not true. By standard compactness arguments, there then exists a sequence N_j going to ∞ with j such that

$$\bar{f} = \lim_{j \to \infty} \frac{1}{N_j} \sum_{n=0}^{N_j - 1} f \circ \psi^n(x)$$

exists for all continuous functions f on $[-1, 1]$ but such that $\bar{f} \neq \int f d\nu$ for some f. But $f \to \bar{f}$ is a positive linear functional on the space of continuous functions, taking the value 1 on the constant function 1, and vanishing if $f = 0$ on J [since $\psi^n(x) \to J$ by assumption]. Thus, there is a probability measure $\bar{\nu}$ on J such that

$$\bar{f} = \int f d\bar{\nu}.$$

A standard argument shows that $\overline{f \circ \psi} = \bar{f}$, so $\bar{\nu}$ is ψ-invariant, so $\bar{\nu} = \nu$, contradicting the fact that $\bar{f} \neq \int f d\nu$ for some f.

The ergodicity of (ν, ψ) follows at once from the fact that ν is the unique ψ-invariant probability measure on J. On the other hand, the set

$$J \cap J_0^{(1)}$$

is invariant under ψ^2 and has measure 1/2, so ψ^2 is not ergodic, so ψ is not weak mixing. One can also show that the spectrum is discrete.

We next show that, if ψ is sufficiently near to the fixed point (and on W_s), then all but countably many orbits of ψ converge to J. In general, if $\psi \in \mathscr{P}_\varepsilon$, it has a unique fixed point in $[0, 1]$, which we will denote by x_0.

Lemma 8.4. *For ε sufficiently small and ψ sufficiently close to the fixed point ϕ (but ψ not necessarily on W_s), the orbit of every $x \neq \pm x_0$ is eventually in $J^{(1)}$.*

Proof. Note first that, since $\psi(x_0) = x_0$ and $\psi(-x_0) = \psi(x_0) = x_0$, the orbits of x_0 and $-x_0$ are extremely simple. Note also that, for any $x \in [-1, 1]$, $\psi^n(x) \in [-a, 1]$

for all $n \geqq 1$ and that

$$[-a, 1] = J^{(1)} \cup \{x_0\} \cup (c, x_0) \cup (x_0, b).$$

Now ψ maps (x_0, b) onto (c, x_0) and (c, x_0) onto an interval containing (x_0, b). Any orbit must therefore either

land on x_0 after no more than one step

or

land in $J^{(1)}$ after no more than two steps (and remain there, by invariance)

or

land in (c, x_0) after no more than two steps.

(These are not mutually exclusive.)

We have to show that, in the third case, $\psi^n(x)$ is eventually in $J^{(1)}$. Now $\psi \circ \psi$ leaves x_0 fixed and maps (c, x_0) onto a larger interval contained in $(-a_0, x_0)$. The idea is that $\psi \circ \psi$ pushes points further and further from x_0 and we want to show that every orbit for $\psi \circ \psi$ starting in (c, x_0) eventually reaches $[-a_0, c]$. If it reaches $[c, a]$, then one more iteration of $\psi \circ \psi$ will put it in $[-a, c]$, so what have to show is that it is impossible that

$$(\psi \circ \psi)^n(x) \in (a, x_0) \quad \text{for all } n.$$

Since $\psi \circ \psi(x_0) = x_0$, it will suffice to prove

$$(\psi \circ \psi)'(x) > 1 \quad \text{on} \quad [a, x_0].$$

The proof of this last statement is straightforward, using the fact that since

$$\phi(x) = 1 - (1 + \lambda_\varepsilon)|x|^{1+\varepsilon} + O(\lambda_\varepsilon)$$

we can, by making ψ sufficiently close to ϕ, arrange that (for example)

$$\psi'(x) \leqq -(1 + \tfrac{2}{3}\lambda_\varepsilon)(1 + \varepsilon)|x|^\varepsilon \quad \text{on} \quad [0, 1]$$

and also that

$$a \geqq \lambda_\varepsilon / 2; \quad x_0 \geqq 1/3.$$

We will now iterate this arrangement to prove:

Proposition 8.4. *If $\psi \in W_s$ is near enough to ϕ, then*

1) *ψ has exactly one periodic orbit of each period $1, 2, 4, 8, \ldots$, and no periodic orbits of other periods. All these periodic orbits are repelling.*

2) *Every orbit of ψ which does not eventually fall exactly on one of the repelling periodic orbits converges to the invariant Cantor set J.*

Proof. The preceding lemma tells us that x_0 is a repelling fixed point for ψ and that every orbit which does not eventually fall exactly on x_0 is eventually in $J^{(1)}$. $J^{(1)}$ consists of the two pieces $J_0^{(1)}$ and $J_1^{(1)}$ which are exchanged by ψ, so to analyze orbits which are eventually in $J^{(1)}$ it suffices to analyze orbits of $\psi \circ \psi$ in $J_0^{(1)}$, i.e., of ψ in $[-1, 1]$. Since ψ is again in W_s and near ϕ, we can apply Lemma 3 to it. Thus, $\psi \circ \psi$ has a repelling fixed point in $J_0^{(1)}$, which corresponds to a repelling orbit of period 2 for ψ, and every orbit for $\psi \circ \psi$ in $J_0^{(1)}$ which does not eventually land on

the fixed point is eventually in $J_0^{(2)} \cup J_2^{(2)}$. Expressed in terms of ψ, every orbit which does not land eventually on either the fixed point x_0 or the orbit of period 2 just described is eventually in $J^{(2)}$.

Continuing in this way we find, for each n, a repelling periodic orbit of period 2^{n-1} and show that every orbit which does not fall exactly on one of the constructed orbits of periods $1, 2, 4, \ldots, 2^{n-1}$ is eventually in $J^{(n)}$.

This proves everything in the proposition except for the non-existence of periodic orbits other than the ones enumerated. From 2) any such orbit, if one exists, must be in J. But for any n, J can be broken into 2^n disjoint pieces

$$J \cap J_j^{(n)}, \quad j = 0, 1, 2, \ldots, 2^n - 1,$$

which are permuted cyclically by ψ. It follows that periodic orbits in J would have to have a period which is divisible by 2^n for all n. This is impossible, so there are no further periodic orbits.

Remarks. 1. We note that λ_ε appears as an asymptotic scaling parameter for the Cantor set J. Specifically, $J_0^{(n)} \cap J$ and $J_0^{(n+1)} \cap J$ asymptotically differ by a scale factor of λ_ε. This means the following: If we write

$$A_n = a(\psi) \ldots a(\mathscr{T}\psi) \ldots a(\mathscr{T}^{n-1}\psi)$$

then

$$J_0^{(n)} \cap J = A_n J(\mathscr{T}^n \psi).$$

As $n \to \infty$, $\mathscr{T}^n \psi \to \phi$, so, in a sense which is easy to make precise, $J(\mathscr{T}^n\psi) \to J(\phi)$. Thus,

$$A_n^{-1}(J_0^{(n)} \cap J) \quad \text{and} \quad A_{n+1}^{-1}(J_0^{(n+1)} \cap J)$$

look essentially the same for large n. In other words, $J_0^{(n+1)} \cap J$ looks almost the same as $J_0^{(n)} \cap J$ multiplied by a scale factor of $A_{n+1}/A_n = a(\mathscr{T}^n\psi)$. Again, since $\mathscr{T}^n\psi \to \phi$, this scale factor converges to $a(\phi) = \lambda_\varepsilon$. Observe, however, that this scaling is different for other pieces of the Cantor set. For example, successive terms in the decreasing sequence $J \cap J_1^{(1)} \supset J \cap J_1^{(2)} \supset J \cap J_1^{(3)} \supset \ldots$ differ asymptotically by a numerical factor of $\lambda_\varepsilon^{1+\varepsilon}$ rather than λ_ε, and the same is true for any of the sequences $J \cap J_j^{(n)}$ for fixed, non-zero j.

2. It is easy to see that, for ψ near ϕ, $J_0^{(2)}$ is longer than $J_1^{(2)}$. More generally for any n, $J_0^{(n)}$ is the longest of the 2^n intervals making up $J^{(n)}$ and $J_1^{(n)}$ is the shortest. The length of $J_0^{(n)}$ is $A_n(1 + a(\mathscr{T}^n\psi))$ which behaves asymptotically like const $\times \lambda_\varepsilon^n$. The conclusion that the longest interval in $J_0^{(n)}$ has length bounded by const $\times \lambda_\varepsilon^n$ holds even for all ψ on W_s since $\mathscr{T}^n\psi$ still converges to ϕ.

3. Proposition 4 remains true even for ψ not near ϕ provided that ψ is near the unstable manifold W_u, that $\psi \in \mathfrak{D}(\mathscr{T})$, and that $\mathscr{T}\psi(1) \leqq 0$. This leads to the following picture of the "bifurcation" which occurs on the stable manifold:

If $\psi \in \mathfrak{D}(\mathscr{T}^n)$, and if $(\mathscr{T}^n\psi)(1) \leqq 0$, then ψ admits a finite decreasing chain

$$J^{(1)} \supset J^{(2)} \supset \ldots \supset J^{(n)}$$

of invariant sets; $J^{(n)}$ is a sort of approximate Cantor set; it is a union of 2^n disjoint closed intervals permuted cyclically by ψ. If in addition ψ is not too far from ϕ, then the space between successive pairs of these intervals contains exactly one

periodic point of ψ. These periodic points have periods $1, 2, 4, \ldots, 2^{n-1}$; there is exactly one cycle of each of these periods, and they are all repelling. There are countably many orbits which fall onto one of these repelling orbits after finitely many steps; all others converge to $J^{(n)}$. If we collapse each of the intervals making up $J^{(n)}$ to a point, all such ψ's look the same – they have an attracting periodic orbit of period 2^n together with the simplest set of repelling periodic orbits between them required by simple considerations of connectedness. Each such ψ can thus be thought of as a sort of semi-direct product of the simplest possible ψ which is superstable of period 2^n with the transformation $\mathcal{T}^n\psi$ scaled down and made to act on $J_0^{(n)}$. These $\mathcal{T}^n\psi$'s can of course be very different – e.g., may on the one hand be superstable of period 2 or on the other hand admit an absolutely continuous invariant measure – but the differences act on a small spatial scale and will therefore not be very noticeable for large n. In the limit $n \to \infty$ the approximate Cantor set becomes a true Cantor set which remains attracting and which can crudely be thought of as a single attracting periodic orbit of period 2^∞; at the same time, the spatial scale of the difference between ψ's goes to zero and so the difference disappears entirely.

Even if ψ is not near enough to W_u for Lemma 4 to apply, it will still be true that $\mathcal{T}^n\psi$ is near enough for n larger than some n_0. Thus, although we cannot be sure that the gaps between the intervals in $J^{(n_0)}$ are free of extraneous recurrent behavior for ψ, each gap produced in passing from $J^{(n_0)}$ to $J^{(n)}$, $n > n_0$ will indeed contain exactly one repelling periodic point. Furthermore, any extraneous recurrent behavior has to be fairly tame. It is not hard to see that any orbit which never enters $J^{(n_0)}$ must be asymptotically periodic with period $1, 2, 4, \ldots$, or 2^{n_0}.

Note. a) In Lemma 8.4 and Proposition 8.5 the condition that ψ be near to ϕ can be replaced by the condition that ψ have negative Schwarzian derivative.

b) The computational proof of Lemma 8.4 can be replaced by a simple conceptual proof using the negativity of the Schwarzian derivative.

Acknowledgements. This work was made possible through the hospitality of the following institutions. EPF – Lausanne, IHES – Bures sur Yvette, Rockefeller University, Harvard University, ETH – Zurich, University of Geneva, and further financial support from 3e Cycle de Physique de la Suisse Romande (O.L.), Grant NSF-MCS78-06718 (O.L.), Grant NSF-PHY-77-18762 (P.C.), and Fonds National Suisse (P.C.).

References

1. Collet, P., Eckmann, J.-P.: A renormalization group analysis of the hierarchical model in statistical physics. Lecture Notes in Physics, Vol. 74. Berlin, Heidelberg, New York: Springer 1978
2. Dieudonné, J.: Foundations of modern analysis. New York, London: Academic Press 1969
3. Feigenbaum, M.: Quantitative universality for a class of nonlinear transformations. J. Stat. Phys. **19**, 25 (1978); **21**, 669 (1979)
4. Guckenheimer, J.: Bifurcations of dynamical systems. C.I.M.E. Lectures 1978
5. Hartmann, P.: Ordinary differential equations. New York, London: Wiley 1964
6. Hirsch, M.W., Pugh, C.C., Shub, M.: Invariant manifolds. Lecture Notes in Mathematics, Vol. 583. Berlin, Heidelberg, New York: Springer 1977
7. Kato, T.: Perturbation theory for linear operators. Berlin, Heidelberg, New York: Springer 1966
8. Lanford III, O.E.: To appear, and Lecture Notes in Physics, Vol. 116, Berlin, Heidelberg, New York: Springer 1980

9. May, R.M.: Simple mathematical models with very complicated dynamics. Nature **261**, 259–467 (1976)
10. Misiurewicz, M.: Structure of mappings of the interval with zero entropy. Preprint I.H.E.S. (1978)
11. Misiurewicz, M.: Absolutely continuous measures for certain maps on an interval. Preprint I.H.E.S./M/79/293 (1979)
12. Singer, D.: Stable orbits and bifurcations of maps of the interval. S.I.A.M. J. Appl. Math. **35**, 260–267 (1978)
13. Stefan, P.: A theorem of Sharkovskii on the existence of periodic orbits of continuous endomorphisms of the real line. Commun. Math. Phys. **54**, 237–248 (1977)
14. Collet, P., Eckmann, J.-P.: Iterated maps on the interval as dynamical systems. Progress in Physics. Birkhäuser Boston 1980 (to appear)

Communicated by D. Ruelle

Received February 28, 1980

Note added in proof. An alternate proof of existence of a fixed point for $\varepsilon = 1$ has been provided by M. Campanino, H. Epstein, D. Ruelle, (to appear). For the extension of the results to multidimensional dissipative maps, see P. Collet, J.-P. Eckmann, H. Koch, J. Stat. Phys. (to appear)

Ann. Rev. Fluid Mech. 1982. 14:347-364

THE STRANGE ATTRACTOR THEORY OF TURBULENCE

Oscar E. Lanford III

Department of Mathematics, University of California, Berkeley, California 94720

INTRODUCTION

It is a fact of experience almost too familiar to notice that dissipative physical systems subject to weak steady driving approach states of dynamic equilibrium that are independent of initial condition. As the strength of the driving is increased, these systems typically undergo a sequence of transitions—the details depending on the system—and arrive eventually at behavior that may be described as chaotic or turbulent. The turbulent motion is not entirely without regularity, but the regularity is statistical in character and appears only when long-term time averages are examined.

Ideally, the mechanisms producing the transition from steady to chaotic behavior, and the detailed nature of the motion in the chaotic regime, should be deducible directly from the equations of motion for the system in question, i.e. the Navier-Stokes or Boussinesq equations in the case of classical kinds of fluid systems. Direct attacks on these equations, however, meet with overwhelming difficulties. On the one hand, control over the analytic properties of the equations is not yet good enough either to prove or to disprove the existence of regular solutions for all times and arbitrary regular initial data. On the other hand, it seems quite hopeless to try to compute explicit analytic solutions with chaotic behavior, to say nothing of computing, from first principles, the statistical distribution describing the behavior of typical solutions. To circumvent the difficulties of a direct approach, a number of oblique lines of attack have been developed. One of these approaches, known as the *strange attractor* theory of turbulence, is the subject of this review.

This approach focuses on the time dependence of turbulent motion; the fundamental idea on which it is based is:

> Turbulent time dependence is not an exceptional feature of particular equations of motion but a property shared by a broad class of typical differential equations.

Adopting this point of view changes the perspective from one of studying particular—and intractable—equations to trying to answer the question:

> How does a typical solution of a typical differential equation behave over the long run?

A substantial body of deep mathematical theory is available to be applied to this question, and mathematical work in this area in recent years has been both invigorated and focused by interaction with the physical study of the chaotic behavior of dissipative systems.

The approach has at least two obvious drawbacks. One is that there is no guarantee that the Navier-Stokes equation will indeed turn out to be typical. This objection is not as serious as it might appear. The Navier-Stokes equation is after all only an approximation, albeit a very good one for most purposes. Even if it were to turn out to have nontypical properties, the very notion of "typical" means that most small perturbations on it would produce an equation with typical behavior. Furthermore, the discovery of a nontrivial exceptional qualitative property of the Navier-Stokes equation would be a great step towards understanding that equation, so the program can further our understanding even if its fundamental presupposition ultimately turns out to be wrong.

The second drawback is that, at best, the investigation of typical behavior can furnish only a list of alternatives. Which of the alternatives actually occurs for a particular equation can only be determined by a detailed study of that equation (or by performing an experiment, either a computer experiment on the equation or an actual experiment on the physical system it describes.)

Up to now, at least, this approach has not contributed very much to the solution of the traditional questions about turbulence or to the practical computation of critical parameter values, phenomenological parameters like effective turbulent viscosity, etc. Its successes have come more in suggesting new questions to be investigated experimentally than in explaining the results of prior experiments. Although the mathematical theory has developed some very powerful methods of analysis, it has generally not been possible to sum up the principal insights in a few concise theorems that can be applied without regard to the reasoning behind them. In short, it is a better source of tools than of recipes.

Terminology

We will be discussing differential equations. By a *state* for a differential equation, we mean a complete specification of initial condition; the space of all states will be called the *state space.* Thus, for a Hamiltonian system, the state space means the phase space rather than the configuration space. For an incompressible fluid system, a point of the state space is a velocity

field, defined on the physical region occupied by the fluid, with vanishing divergence and satisfying appropriate boundary conditions. We will use the term *orbit* to refer to a solution to the differential equation regarded as a curve in the state space, and call *solution flow* the motion on the state space that advances each point along its respective orbit. We will say that a stationary or periodic orbit is *stable* or *attracting* if all orbits starting sufficiently near to it converge to it; this property is frequently called *asymptotic stability in the sense of Lyapunov.*

The terms *turbulent, chaotic,* and *stochastic* (applied to describe time dependence of solutions to a differential equation) will be used interchangeably. Note, however, a slight and potentially question-begging difference in connotation; *stochastic,* as normally used, implies the existence of a well-defined average behavior.

Some General References

The idea that chaotic time dependence of turbulent fluid flows might be understood as a property of fairly general differential equations was first advanced in a way that attracted widespread attention in Ruelle & Takens (1971a,b). [A very suggestive example had been pointed out earlier by Lorenz (1963), but Ruelle & Takens were not aware of Lorenz's work.] The paper of McLaughlin & Martin (1975) was very influential in popularizing these ideas. Recent general surveys include Ruelle (1978a,b, 1980a,b), Lanford (1981), and Eckmann (1981).

CHAOTIC BEHAVIOR

It is often felt that there is something paradoxical about having solutions to a deterministic equation behave in a chaotic or stochastic fashion. There is, however, no real paradox; the solutions are, in fact, uniquely determined by the initial conditions, but the effects of small changes in the initial conditions are so amplified by the equations of motion that any *finite-precision* information about the initial conditions provides no *finite-precision* information about the state of the system at much later times. In other words:

> An important element in the explanation of the chaotic behavior of solutions of deterministic equations of motion is the sensitive dependence of solutions on initial conditions.

The use of probabilistic concepts in the analysis of deterministic motion should, in any case, be familiar from classical equilibrium statistical mechanics. In fact, sensitive dependence on initial condition, in the form of the notion of ergodicity, has long played a central role in one of the standard justifications for the foundations of that subject (see, for example,

Lebowitz & Penrose 1973). It needs to be noted, however, that there are substantial differences between Hamiltonian systems—to which the usual formalism of classical statistical mechanics applies—and the sort of dissipative systems under discussion here. The key distinction lies in the volume-preserving character of the solution flow for Hamiltonian systems (Liouville's Theorem) which has as a consequence the fact that the solution flow is *recurrent* (meaning, roughly, that almost all orbits come back arbitrarily near their initial points infinitely often). For dissipative systems, on the other hand, what usually happens is that most of the points of the instantaneous state space are *transient* in the sense that the orbits that start there eventually go to and stay in another part of the state space. A simple instance is provided by the stable dynamic equilibrium that is usually set up when a dissipative system is driven gently. In this situation, all orbits, no matter where in the state space they start, converge eventually to a single stationary solution corresponding to laminar motion. In a certain sense, the system has *no* effective degrees of freedom, although the state space may have large or even infinite dimension. It seems very likely that something similar happens for more strongly driven dissipative systems, even those whose motion is chaotic, viz., that

> There are one—or possibly a few—invariant sets of relatively low dimension in the state space to which almost all orbits converge.

These sets are what are called *attractors*. One of the great surprises in this subject is that attractors, except for the very simplest ones, are typically not smooth surfaces in the state space but rather more complicated kinds of sets.

Before taking up the notion of attractor in more detail, we need to elaborate on what is meant in practice by chaotic behavior of a physical system whose equation does not depend explicitly on time. Roughly, the idea is that the system behaves, over a long period of time, in a repetitive but not strictly periodic fashion. It should be emphasized that, in this article (and in the field it surveys), the analysis is focused on understanding temporal—not spatial—chaos. There is a tendency to identify chaotic motion with spatially complicated flow patterns. This identification is not entirely mistaken, since laminar flow generally has simple geometry, but it may be misleading. It is possible to have either

> a fluid flow with extremely complicated intrinsic spatial structure and no time dependence at all (in, for example, large-aspect-ratio convection)

or

> chaotic motion with relatively simple spatial structure (small-aspect-ratio convection).

It is not even true that *apparent* temporal complexity necessarily indicates chaotic behavior. Convective systems, for example, can undergo quite complicated periodic motion before they become aperiodic. When they do become stochastic, the motion can often be decomposed, at least roughly, into a small stochastic component superimposed on a much larger periodic component. It is often not easy to distinguish between such weakly stochastic motion and complicated but purely periodic motion simply by watching the system. The standard way to detect a stochastic component in the motion is to measure the *power spectrum* of some dynamical variable. Stripped of technicalities, this means the following: Measure, at equally spaced times t_i some numerical quantity such as one component of the velocity at a particular point of the fluid. Call the measured quantities $X_1, \ldots, X_N$. Subtract the mean and take the discrete Fourier transform, i.e. form

$$\tilde{X}(\omega) = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} (X_j - \overline{X})\, e^{i\omega j}, \tag{1}$$

where ω is of the form $2\pi k/N$ and where

$$\overline{X} = \frac{1}{N} \sum_{j=1}^{N} X_j . \tag{2}$$

Then see whether $|\tilde{X}(\omega)|^2$ is concentrated in a series of sharp peaks. The appearance of a "broad-band" component in $|\tilde{X}(\omega)|^2$ (beyond that due to the finite precision of the measurements) is generally taken as the operational definition of stochastic behavior of the system in question.

ATTRACTORS

One of the most fruitful ways of studying the mathematical structures underlying observed stochastic behavior of physical systems has been the careful study, mostly with the aid of computers, of simplified models, and one model that has been particularly informative is the *Lorenz system,* a set of three coupled differential equations:

$$\frac{dx}{dt} = -\sigma x + \sigma y ; \qquad \frac{dy}{dt} = rx - y - xz ;$$

$$\frac{dz}{dt} = -bz + xy , \tag{3}$$

where b, σ, and r are constants. It is hard to imagine a much simpler system that is neither linear nor two-dimensional, but the solutions to these equations nevertheless do very complicated things. E. N. Lorenz (1963) discovered numerically a striking mathematical structure which has

come to be known as the *Lorenz attractor* and which occurs for these equations with

$$b = 8/3\,; \qquad \sigma = 10\,; \qquad r = 28\,. \tag{4}$$

The exact parameter values are not crucial, but the behavior of typical solutions definitely does depend on the parameters and is quite different in other regions of parameter space. It should also be noted that, in spite of overwhelming numerical evidence, there is to my knowledge no complete proof that the structure about to be described actually does occur for these specific equations. It is not hard to see that it does occur for *some* equations.

The phenomenology is as follows: The equations admit three stationary solutions, one at the origin and the other two (which we will denote by $C_\pm$ and refer to as *centers*) at $x=y=\pm\sqrt{b(r-1)}$; $z=r-1$. All three stationary solutions are unstable. Orbits that start near the origin escape monotonically; those that start near the centers escape through growing oscillations. If a solution is computed starting from some more or less randomly chosen initial point, what is found without exception is that the orbit will, after an initial transient regime of variable length, settle down to a motion in which, most of the time, it can be thought of as performing oscillations about one of the centers. The oscillation grows in amplitude; when it reaches a critical size, the orbit abruptly makes a transition to oscillation about the other center. This oscillation again grows and the orbit eventually makes a transition back to oscillating about the first center, and so on. A representative orbit is shown in Figure 1.

The amplitude of oscillation immediately after transition varies from transition to transition, and it in turn determines the number of oscillations before the next transition. The sequence of numbers of oscillations between transitions appears random, and the power spectra of the coordinates are continuous (see Figure 2). Thus, the motion both appears chaotic and satisfies the standard operational test for chaotic behavior.

The mathematical object responsible for this behavior is sketched schematically in Figure 3. This sketch represents a family of orbits in the three-dimensional state space for the Lorenz system. It is not even approximately to scale; proportions have been distorted in the hope of making the mathematical structure more transparent. To a first approximation, the structure looks like two reasonably flat loops of ribbon, one lying above the other along a central band, and the two glued together at the bottom of that band. The motion flows around the loops, clockwise on the left and counter-clockwise on the right. Going once around the right-hand loop constitutes a single oscillation around C_+. Orbits beginning either above or below the ribbon are attracted quickly down to its immediate vicinity and then follow the flow on it. The double-loop structure is strictly invar-

iant under the solution flow; any point on it has an orbit that can be traced both forward and backward for all time without leaving it.

The central band is divided in half by orbits that flow essentially straight down to the stationary solution at the origin; these orbits are exceptions to the pattern of growing oscillations followed by transitions displayed by typical orbits. Orbits to the left of this boundary will make their next oscillation around C_-; those to the right will go next around C_+. The fact that oscillations around the centers are growing in amplitude means that, for example, a loop around C_+ brings the orbit back to the left of where it started out. A transition from oscillation around C_+ to oscillation around C_- occurs when an orbit making a loop around C_+ comes back to the left of the dividing boundary.

The central band divides in half laterally at the bottom of this boundary, and each half, after having made a loop around the appropriate center, has become wide enough to cover almost the entire top of the band. Thus,

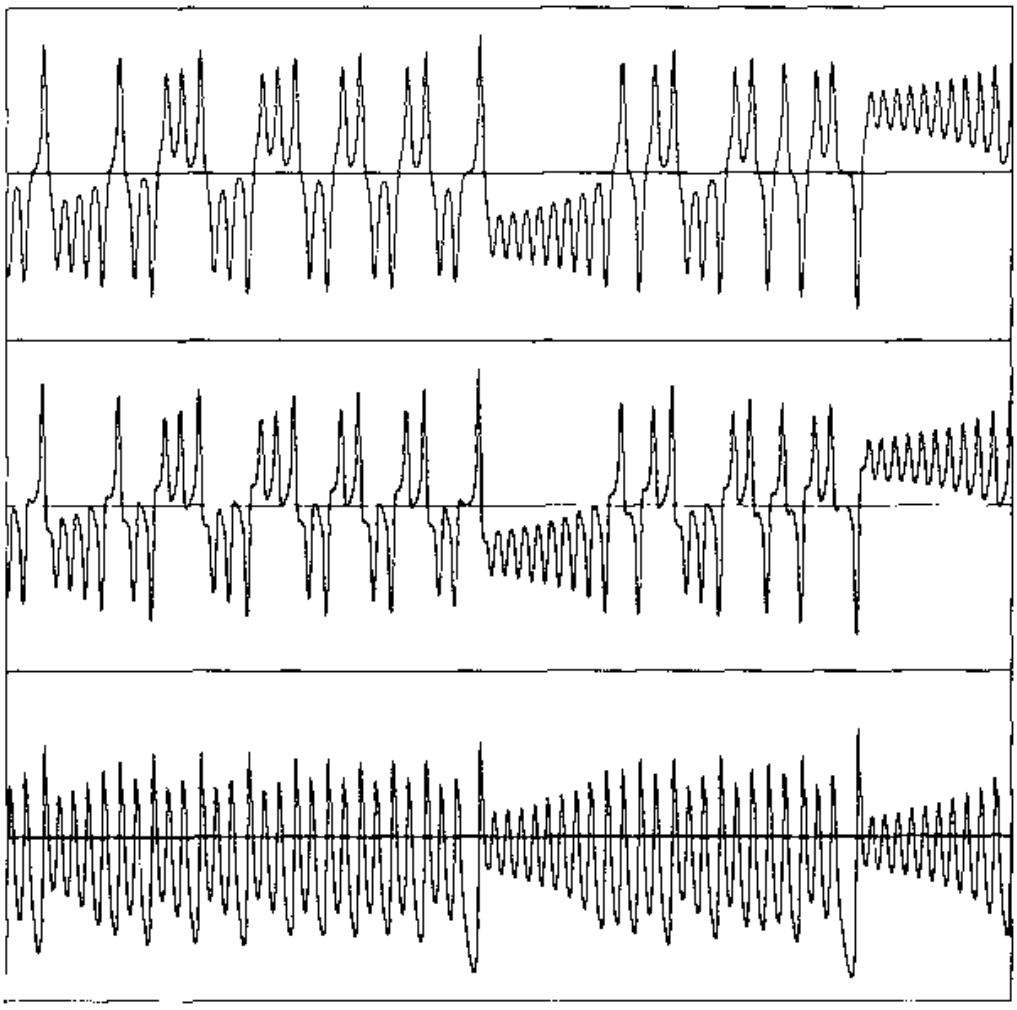


Figure 1 A representative orbit for the Lorenz system. From top to bottom: x, y, and z plotted against t.

orbits are pulled apart laterally as they flow around the loops, and this accounts for the observed sensitive dependence on initial conditions.

A typical orbit on this structure wanders over the surface, coming arbitrarily close to each point infinitely often. There are, however, a great many nontypical orbits. We have already mentioned the orbits in the middle of the central band which simply converge down toward the origin. There are also many periodic orbits, all unstable, as well as orbits with more subtle kinds of atypical behavior.

We next take a closer look at the ribbons and argue that they cannot be simple surfaces but must rather have infinitely many layers. Start at the top of the central band where there are two approximately parallel ribbons, one on top of the other. As we have drawn the picture, the upper ribbon is made up of orbits returning to the central band after a loop around C_+; the lower, around C_-. As the orbits flow down the central band, the two ribbons are drawn together. At the bottom, they form a two-sheeted surface which proceeds to split laterally in two with half going left around C_- and half right around C_+. Thus, the ribbon of orbits going around C_+ or C_- has at least two sheets, the upper one made up of orbits whose previous circuit was around C_+, the lower of orbits whose previous circuit was around C_-. These sheets are carried closer together by the flow but the separation remains nonzero, so the upper ribbon at the top of the central band actually has two layers rather than just one. The same is true for the lower ribbon, and therefore the structure at the bottom of

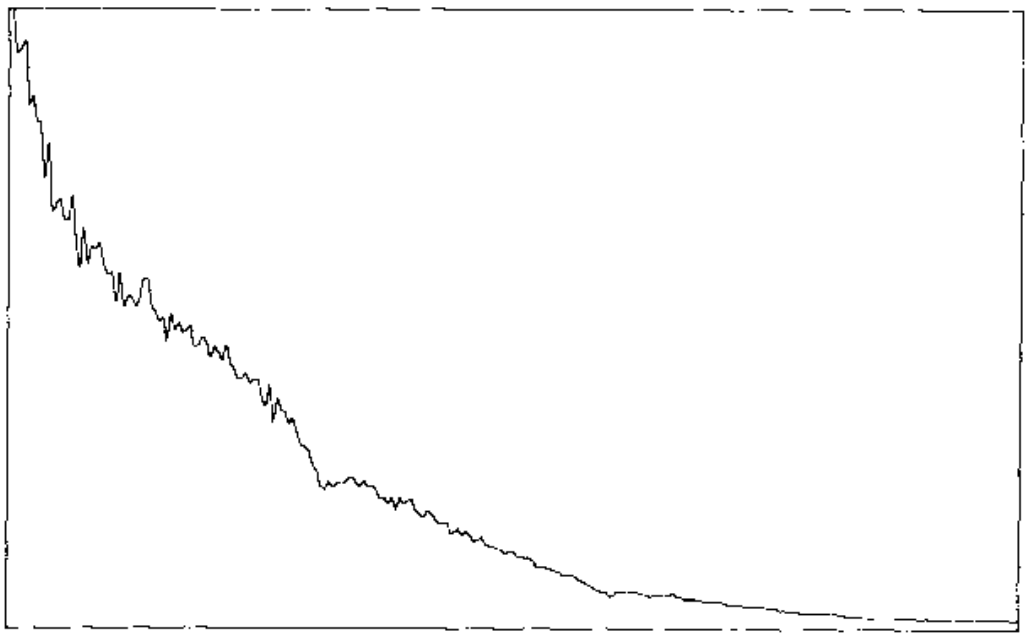


Figure 2 The power spectrum of the x coordinate of the Lorenz system (on a linear scale). The frequency ranges from 0 to 5; the oscillations apparent in Figure 1 have a frequency of about 1.5.

the central band actually has four layers rather than just two. Thus, the ribbons going around C_+ and C_- are actually four-sheeted, and so on. Continuing to argue in this way we see that all the ribbons must have infinitely many sheets.

This object is an instance of what has come to be called a *strange attractor.* The formulation of the definitive definition of the term *attractor* must await a more complete understanding of what the possible phenomena are. The general idea is clear, however: an attractor is a set that attracts nearby orbits (i.e. an orbit that starts near the attractor stays near it and converges to it as time goes to infinity). It should also be required that the set be closed and invariant under the solution flow (i.e. be made up of complete orbits defined for all time) and that the solution flow on the set be *recurrent*, i.e. that most orbits return infinitely often to the vicinity of their starting points. A situation which occurs frequently and which suffices to guarantee that the motion is adequately recurrent is that there is a single orbit in the attractor passing arbitrarily near to every point.

Stable stationary solution and limit cycles are elementary and trivial examples of attractors. It has turned out that the other kinds of attractors that occur most frequently are structures with infinitely many layers like the Lorenz attractor described above. Hence the epithet "strange." The general investigation of attractors has barely begun. Obtaining a complete classification looks, at present, extremely remote. On a less ambitious and more pragmatic level, there does not exist a convincing list of the "simplest" possibilities. Even worse, there are very simple examples—notably, the *Hénon* attractor (Hénon 1976, Hénon & Pomeau 1976, Curry 1979)—whose properties really aren't understood at all. There does exist, however,

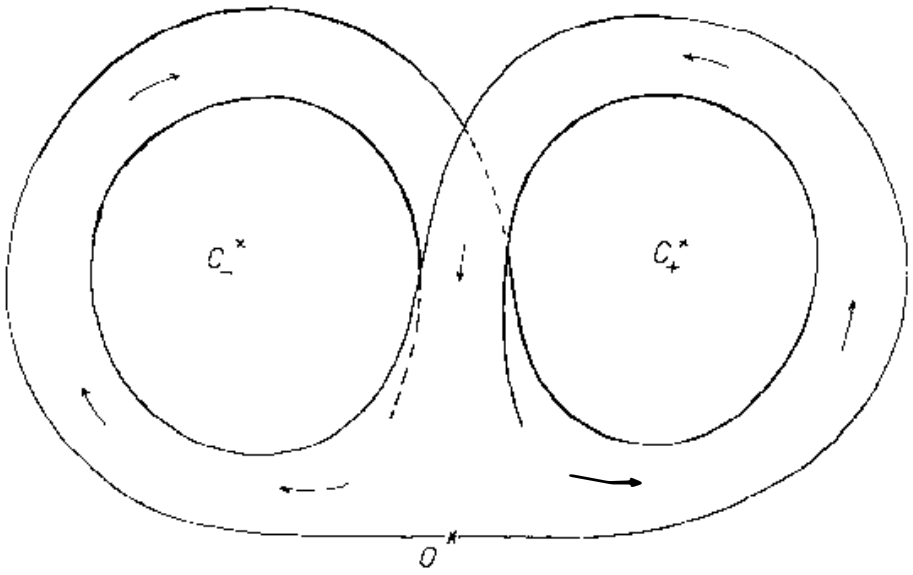


Figure 3 A schematic view of the Lorenz attractor.

one class of attractors for which there is a detailed and deep mathematical theory. These are the attractors satisfying a condition introduced by S. Smale (1967) and known as *Axiom A*. Roughly, Axiom A requires that orbits on or very near to the attractor have a strong and technically convenient form of sensitive dependence on initial condition, for both positive and negative time; it also requires that arbitrarily close to each point of the attractor there passes a periodic orbit. We will return in the final section of this review to one of the most important properties of attractors satisfying Axiom A.

For a more detailed discussion of the Lorenz attractor see Lorenz (1963), Guckenheimer (1976), Ruelle (1976a), Lanford (1978), Guckenheimer & Williams (1979), and Williams (1979). The literature on attractors satisfying Smale's Axiom A is extensive. For an introductory survey, see Bowen (1978).

SCENARIOS

In addition to asking how typical solutions of a typical differential equation behave, it is also possible to ask how this behavior varies as a parameter in the differential equation changes. Of particular interest is the study of the transition process from an equation whose solutions are asymptotically regular (i.e. stationary or periodic) to one with some kind of strange attractor. One can hope that there will turn out to be comparatively few such transition processes, each with distinctive characteristics permitting it to be identified without a detailed analysis of the underlying equations. J.-P. Eckmann has introduced the term *scenario* for such typical transition processes.

The Hopf bifurcation is an excellent classical example of a successful application of this approach. Let us review briefly how it works. The idea is to see what happens when, as a parameter is changed, a stationary solution to a differential equation loses stability. We consider, then, a differential equation depending on a parameter μ and a stationary solution for that equation which may also depend on μ. The particular situation we want to examine is that

> for $\mu < \mu_c$, all the eigenvalues of the linearization of the equation at the stationary solution have strictly negative parts,

but

> at $\mu = \mu_c$, a single complex-conjugate pair of simple nonreal eigenvalues crosses into the right half-plane with nonzero speed.

The Hopf Bifurcation Theorem says that the qualitative nature of the motion near the stationary solution is determined by the sign of a certain nonlinear function d of the first, second, and third derivatives of the dif-

ferential equation with respect to the state variables at the stationary solution for μ_c. (The explicit formula for d is extremely complicated but need not concern us here.) For $d>0$, what happens is that, for μ slightly larger than μ_c, the equation has an attracting periodic orbit—i.e. a stable nonlinear oscillation—in the neighborhood of the now unstable stationary solution. As μ decreases to μ_c, the oscillation collapses down to the stationary solution; its amplitude is asymptotically proportional to $\sqrt{\mu-\mu_c}$. For $d<0$, something equally specific (if less interesting) happens: for μ slightly less than μ_c, there is a nonattracting periodic orbit which, as μ increases to μ_c, collapses down to the stationary solution with amplitude asymptotically proportional to $\sqrt{\mu_c-\mu}$. One of the great strengths of the theorem is that it assures us that, under our assumptions about the eigenvalues of the linearization, these are the *only* two possibilities except in the degenerate case $d=0$. (For a detailed discussion of the Hopf bifurcation, and proofs of the results cited above, see Marsden & McCracken 1976.)

The most interesting applications of the Hopf Bifurcation Theorem to physical systems do *not* proceed by verifying that the equations of motion satisfy the hypotheses. What is done, rather, is to observe experimentally the way the system in question goes from a stationary to an oscillatory regime as the parameter passes through a critical value. If the amplitude of the oscillations grows like $\sqrt{\mu-\mu_c}$ as μ passes μ_c, it is fairly safe to conclude that the transition process is a Hopf bifurcation, the square-root behavior serving as an experimentally verifiable signature. In this way, it is possible to arrive at a fairly precise picture of the transition from steady to oscillatory behavior even in situations where it is impossible to compute the stationary solution accurately, either analytically or numerically.

The Hopf bifurcation is thus an extremely successful scenario for the comparatively elementary transition from steady to periodically oscillatory motion. A few scenarios for the more interesting transition from periodic to aperiodic motion have been analyzed; a useful practical summary of this area has been given recently by Eckmann (1981). We will concentrate here on just one scenario, known as the *Feigenbaum transition*, which has been identified unequivocally in numerical studies of a number of simple models and perhaps observed in convection experiments as well.

The transition process is not a single bifurcation but an infinite sequence of them; it may be described as follows: The starting point is an attracting periodic orbit which loses stability at a first critical parameter value μ_0. In losing stability, it produces a new attracting periodic orbit in much the same way as an attracting periodic orbit is produced in the Hopf bifurcation. The new orbit tracks the old one closely but goes around it twice before closing (see Figure 4).

Thus, the period of the stable oscillation for μ just above μ_0 is twice that for μ just below μ_0. We will therefore refer to this bifurcation as the *period-*

doubling bifurcation; it is also frequently called the *pitchfork bifurcation.* As the parameter is further increased, a second critical value μ_1 is reached at which the doubled orbit itself loses stability through a period-doubling bifurcation. Thus, for μ just above μ_1, the system has an attracting periodic orbit that follows four times around the fundamental orbit before closing, i.e. a stable oscillation with period about four times the base period. This orbit in turn undergoes a period doubling bifurcation (at $\mu=\mu_2$) producing an oscillation with period about eight times the base period, and in fact the cascade continues ad infinitum in a finite parameter interval. The parameter value μ_∞ at which the sequence of doublings accumulates represents the onset of chaotic behavior.

A qualification is necessary here. The fact that such infinite sequences of period doublings do occur and cannot be eliminated by small changes in the differential equation is well established, as are a number of striking features of the accumulation process. What is not well understood is the character of the motion for μ slightly above μ_∞. Computer experiments strongly suggest that for most parameter values just above μ_∞ there is a strange attractor or perhaps several strange attractors, similar in character to the Hénon attractor. Unfortunately, nothing like this has been proved as yet; it simply isn't known for certain that a strange attractor *ever* appears at the accumulation of period doublings for a differential equation depending on a parameter. One thing that is known is that, if such an attractor exists, it must be very unstable with respect to changes in μ; there are values of μ above but arbitrarily close to μ_∞ for which there is an attracting periodic orbit.

The accumulation of period doublings has some characteristic features, discovered by M. Feigenbaum (1978, 1979b, 1980), which are consider-

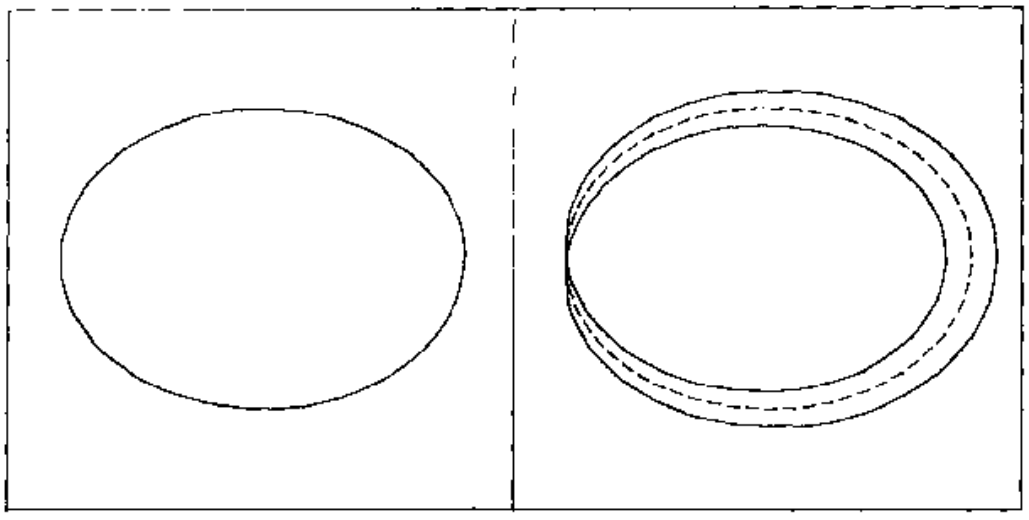


Figure 4 The period doubling bifurcation. On the left: the stable orbit for $\mu < \mu_0$. On the right: the solid curve represents the doubled stable orbit for $\mu_0 < \mu < \mu_1$; the dotted curve the now-unstable doubled orbit.

ably more distinctive than the square-root behavior of the amplitude in the Hopf bifurcation. The first, and easiest to state precisely, concerns the rate of convergence of the successive critical parameter values (μ_n) to their limit and says that, excluding degenerate cases, the convergence is geometric with a universal ratio:

$$\lim_{n\to\infty} \frac{\mu_n - \mu_{n-1}}{\mu_{n+1} - \mu_n} = 4.6692 \ldots . \tag{5}$$

This ratio has been observed in a number of model systems, and its origin and universality are well understood theoretically. Because of the relatively large value of the limiting ratio, all but the first few μ_n's can be expected to be very close together, and this makes the observation of successive period doublings and physical measurement of the ratio extremely difficult. Unless the parameter can be controlled extraordinarily well, what will be observed is one or two period doublings and then chaotic behavior. In fact, up to now, the limiting ratio has not been measured, even roughly, in any physical system.

Feigenbaum has also argued that the power spectrum for essentially any dynamical variable for μ near μ_∞ should display a universal ratio between the strengths of lines corresponding to periods of about $2^n\tau_0$ and those of nearby lines corresponding to periods of about $2^{n+1}\tau_0$. (Here, τ_0 denotes a base period). This part of Feigenbaum's analysis has not yet been put on a firm mathematical footing, but it does appear that something similar to what Feigenbaum predicts has been observed in convection experiments performed by Libchaber & Maurer (1980). This area is currently under very active investigation, and the situation should be clarified soon.

To close this section, we should point out that, although the notion of scenario is an appealing and powerful one, it does not apply to all transitions from regular to chaotic behavior. Discontinuous transitions, in which the chaotic component of the motion is large for parameter values even slightly above the critical value, can and do occur. The transition from a stable stationary solution to the Lorenz attractor is a mathematical example, and pipe flow appears to be a physical one. (These two situations are, however, not quite parallel. For the Lorenz system, there is a critical parameter value above which the stationary solution is no longer stable. The attractor and the stable stationary solution coexist for parameter values slightly below the critical one. For pipe flow, the stationary solution remains stable for arbitrarily large values of the Reynolds number, but becomes more and more sensitive to perturbations of small but finite amplitude).

For the theory of the Feigenbaum transition see, in addition to the works already cited, Feigenbaum (1979a), Collet et al. (1980, 1981), and Lanford (1980).

STATISTICAL THEORY

The investigation of strange attractors throws some light on the question of what should be meant, in a fundamental sense, by a statistical theory of a dissipative differential equation. The question needs to be turned around from the form usual in the classical statistical mechanics of Hamiltonian systems. Because of Liouville's Theorem, Hamiltonian systems—at least those with bounded and nonsingular energy surfaces—come equipped with natural time-independent probability distributions, the microcanonical ensembles of statistical mechanics. Such an ensemble is essentially just normalized area on an energy surface, and thus is a comparatively elementary and familiar construct; the investigation of whether there are other, radically different, stationary probability distributions can justifiably be dismissed as an empty mathematical exercise. Part of the reason why these ensembles are so natural is that the sets to which they assign probability zero conform to our intuitive notion of negligible sets of initial conditions. That is, if some particular behavior occurs only for a set of initial conditions of microcanonical probability zero, we can reasonably conclude that behavior will never be observed.

There is a general theorem about invariant probability distributions, the Birkhoff Pointwise Ergodic Theorem, which asserts that the long-time limit of the time average of any dynamical variable exists, except perhaps for a set of initial conditions of probability zero. (Here, "dynamical variable" simply means a function on the state space sufficiently well behaved for its integral to be defined and finite). Applied to Hamiltonian systems and the microcanonical ensemble, this theorem says that the time average of any dynamical variable will settle down if watched long enough. The existence of limiting time averages is an extremely familiar fact of experience, both for Hamiltonian and for dissipative systems, and it is commonly assumed that this is a general property of differential equations provided that their solutions remain in a bounded region of the state space. There is, however, no such general theorem for dissipative systems, and it is possible to find differential equations for which time averages do not exist. (An example is given in Ruelle 1980a).

Once the existence of time averages for Hamiltonian systems is established, attention turns to the problem of computing them. The most favorable situation is one where the time average does not depend on the initial condition, but is simply given by the ensemble average of the dynamical variable in question. It is not hard to show that this will be the case for all dynamical variables if and only if the system is *ergodic,* i.e. if and only if there is no way to decompose the energy surface into two parts, each of nonzero microcanonical probability, in a time-invariant way. Determining whether a particular Hamiltonian is ergodic is generally a

delicate mathematical problem, and a great deal of important and deep work has been done on such questions over the past fifty years.

The situation for dissipative systems looks entirely different. Liouville's Theorem does *not* hold; indeed, solution flows generally contract volumes in the state space. As already noted, it is not even automatic that limiting time averages exist. There are, on the other hand, general abstract theorems asserting the existence of time-invariant probability distributions, and even of ergodic probability distributions, provided that solution curves don't run off to infinity. These probabilities are not, as in the Hamiltonian case, spread out over the state space. Consider, for example, what happens in the vicinity of an attractor. Any invariant probability distribution must assign probability zero to the set of all orbits converging to the attractor but not actually lying in it. Thus, whatever invariant probability distribution we might choose to work with, it is no longer justified to interpret a set of initial states of probability zero as physically negligible. Moreover, on a typical complicated attractor like the Lorenz attractor, there are a great many invariant probability distributions and it is not at all apparent how to go about singling out the right one—analogous to the microcanonical ensemble for Hamiltonian systems—to represent the statistics of typical orbits.

Rather than focusing on the choice of an invariant probability distribution, it is more satisfactory to start from limiting time averages of dynamical variables. These are, in any case, the quantities of most direct interest. One might hope that, in good cases, limiting time averages would exist along all orbits and be independent of orbit. This is slightly too optimistic. In the first place, there is no reason why a solution flow should have only one attractor. If there are several, the limiting time average will depend on which attractor the orbit approaches. We will therefore concentrate on a single attractor and study time averages along orbits converging to it. The second qualification is less obvious. It turns out, for the nontrivial attractors that have been studied in detail, that there are always many orbits with exceptional long-term behavior. A simple example is that these attractors generally contain many (unstable) periodic orbits. The best one can hope for, then, is that time averages will be *essentially* independent of orbit, i.e. that among the orbits converging to our attractor there may be a small set of exceptional orbits but that time averages do exist and are independent of orbit as long as the orbit does not belong to the exceptional set. A sensible meaning to give to the word "small" in this case is that the set of exceptional orbits has zero volume in the space of instantaneous states; exactly as for Hamiltonian systems, it is reasonable to suppose that such a set will never be seen in an experiment. If this favorable situation obtains, we will say that the attractor is *ergodic*. It follows from standard theorems that there is then an invariant probability

distribution on the attractor—in physical terms, a stationary ensemble—such that the ensemble average reproduces the time average along nonexceptional orbits for all (continuous) dynamical variables.

Two questions now present themselves:

Are the attractors, which arise in practice, ergodic?

Can the ensemble reproducing the time average along nonexceptional orbits be described directly?

The only answer to the first question available at this time is that the few attractors whose structures are well understood—i.e. the Lorenz attractor and those satisfying Smale's Axiom A—turn out to be ergodic. For Axiom A attractors, this is the important Bowen-Ruelle Ergodic Theorem (Bowen & Ruelle 1975, Bowen 1975, Ruelle 1976b). For this same class a surprising answer to the second question is also available. Omitting numerous technicalities, this answer is roughly as follows: First take a hypersurface slicing through the attractor transversally in such a way that every orbit on the attractor crosses the hypersurface frequently. Introduce "coordinates" on the attractor by describing each point by giving the last place its orbit crossed the hypersurface, and the time since that crossing. Now cut up the intersection of the hypersurface with the attractor into a finite number of sufficiently small pieces, and describe an orbit on the attractor by saying which of these pieces it hits in which order. If the pieces are labeled with, say, 1,2,...,n, then this procedure associates with the orbit a two-sided infinite sequence of integers in the range from 1 to n. If the cutting-up is done with sufficient care, it can be arranged that

The set of sequences thus obtained from all orbits on the attractor can be described in a simple way; it is the set of all sequences in which certain pairs (i,j) never occur in succession;

The sequences are essentially in one-to-one correspondence with points on the intersection of the attractor with the hypersurface.

The ensemble that reproduces time averages along nonexceptional orbits can be transported to an ensemble on the space of sequences. This transported ensemble turns out to be the thermodynamic equilibrium ensemble for a one-dimensional array of copies of a system with a discrete set of n states, with some nearest-neighbor exclusions and otherwise interacting through a many-body potential, that decreases exponentially as the separation goes to infinity.

The "equilibrium ensemble" for an Axiom A attractor thus looks much more complicated than the microcanonical ensemble for a Hamiltonian system. To construct it, it is necessary both to have a great deal of detailed

information about the solution flow and to find the thermodynamic equilibrium ensemble for an infinite assembly of systems interacting in a nontrivial way. In simple cases at least, the description given might be used as a starting point for the construction of numerical approximations. The main interest of the Bowen-Ruelle Ergodic Theorem is, however, foundational. At least for very well behaved dissipative systems, it answers in a convincing and precise way the question of how, in principle, the equilibrium ensemble is to be defined.

ACKNOWLEDGMENTS

Preparation of this review was begun while the author was a visitor at the IHES in Bures-sur-Yvette, France. Financial support for that visit from the Volkswagen Foundation, and continuing financial support from the National Science Foundation (MCS78-06718), are gratefully acknowledged.

Literature Cited

Bowen, R. 1975. *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms, Lecture Notes in Mathematics 470*. Berlin/Heidelberg/New York: Springer. 108 pp.

Bowen, R. 1978. *On Axiom A Diffeomorphisms, CBMS Regional Conf. Ser. 35*. Providence: Am. Math. Soc. 45 pp.

Bowen, R., Ruelle, D. 1975. The ergodic theory of Axiom A flows. *Invent. Math.* 29: 181–202

Collet, P., Eckmann, J.-P., Koch, H. 1981. Period doubling bifurcations for families of maps on R^n, *J. Stat. Phys.* 25:1–14

Collet, P., Eckmann, J.-P., Lanford, O. E. 1980. Universal properties of maps on an interval. *Commun. Math. Phys.* 76:211–54

Curry, J. H. 1979. On the Hénon transformation. *Commun. Math. Phys.* 68:129–40

Eckmann, J.-P. 1981. Roads to turbulence in dissipative dynamical systems. *Rev. Mod. Phys.* In press

Feigenbaum, M. J. 1978. Quantitative universality for a class of nonlinear transformations. *J. Stat. Phys.* 19:25–52

Feigenbaum, M. J. 1979a. The universal metric properties of nonlinear transformations. *J. Stat. Phys.* 21:669–706

Feigenbaum, M. J. 1979b. The onset spectrum of turbulence. *Phys. Lett.* 74A:375–78

Feigenbaum, M. J. 1980. The transition to aperiodic behavior in turbulent systems. *Commun. Math. Phys.* 77:65–86

Guckenheimer, J. 1976. A strange, strange attractor. See Marsden & McCracken 1976, pp. 368–81

Guckenheimer, J. Williams, R. F. 1979. Structural stability of Lorenz attractors. *Publ. Math. IHES* 50:59–72

Hénon, M. 1976. A two-dimensional mapping with a strange attractor. *Commun. Math. Phys.* 50:69–77

Hénon, M., Pomeau, Y. 1976. Two strange attractors with a simple structure. In *Turbulence and Navier Stokes Equation*, ed. R. Temam, pp. 29–68. *Lecture Notes in Mathematics 565*. Berlin/Heidelberg/New York: Springer. 194 pp.

Lanford, O. E. 1978. Qualitative and statistical theory of dissipative systems. In *Statistical Mechanics: C.I.M.E. 1976*, pp. 26–98. Napoli: Liquori. 235 pp.

Lanford, O. E. 1980. Remarks on the accumulation of period doubling bifurcations. In *Mathematical Problems in Theoretical Physics*, pp. 340–42. *Lecture Notes in Physics 116*. Berlin/Heidelberg/New York: Springer. 412 pp.

Lanford, O. E. 1981. Strange attractors and turbulence. In *Hydrodynamic Instabilities and the Transition to Turbulence*, ed. H. L. Swinney, J. P. Gollub, pp. 7-26. *Topics in Applied Physics 45*. Berlin/Heidelberg/New York: Springer. 292 pp.

Lebowitz, J. L., Penrose, O. 1973. Modern ergodic theory. *Physics Today* 26(2):23–31

Libchaber, A., Maurer, J. 1980. Une expérience de Rayleigh-Bénard de géométrie réduite. *J. Phys., Colloques C3* 41:51–56

Lorenz, E. N. 1963. Deterministic nonperiodic flow. *J. Atmos. Sci.* 20:130–41

Marsden, J. E., McCracken, M. 1976. *The Hopf Bifurcation and its Applications. Applied Mathematical Sciences* 19. New York/Heidelberg/Berlin: Springer. 408 pp.

McLaughlin, J. B., Martin, P. C. 1975. Transition to turbulence in a statically stressed fluid system. *Phys. Rev. A* 12:186–203

Ruelle, D. 1976a. The Lorenz attractor and the problem of turbulence. See Hénon & Pomeau 1976, pp. 146–58

Ruelle, D. 1976b. A measure associated with Axiom A attractors. *Am. J. Math.* 98:619–54

Ruelle, D. 1978a. Dynamical systems with turbulent behavior. In *Mathematical Problems in Theoretical Physics*, pp. 341–60. *Lecture Notes in Physics 80.* Berlin/Heidelberg/New York: Springer. 438 pp.

Ruelle, D. 1978b. Sensitive dependence on initial condition and turbulent behavior of dynamical systems. *Ann. NY Acad. Sci.* 316:408–16

Ruelle, D. 1980a. Measures describing a turbulent flow. *Ann. NY Acad. Sci.* 357:1–9

Ruelle, D. 1980b. Les attracteurs étranges. *La Recherche* 108:132–44

Ruelle, D., Takens, F. 1971a. On the nature of turbulence. *Commun. Math. Phys.* 20:167–92

Ruelle, D., Takens, F. 1971b. Note concerning our paper "On the nature of turbulence." *Commun. Math. Phys.* 23:343–44

Smale, S. 1967. Differentiable dynamical systems. *Bull. Am. Math. Soc.* 73:747–817

Williams, R. F. 1979. The structure of Lorenz attractors. *Publ. Math. IHES* 50:73–99

BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 6, Number 3, May 1982

A COMPUTER-ASSISTED PROOF OF THE FEIGENBAUM CONJECTURES

BY OSCAR E. LANFORD III[1]

1. Introduction. Let M denote the space of continuously differentiable even mappings ψ of the interval $[-1, 1]$ into itself such that

M1. $\psi(0) = 1$,

M2. $x\psi'(x) < 0$ for $x \neq 0$.

M2 says that ψ is strictly increasing on $[-1, 0)$ and strictly decreasing on $(0, 1]$, so M is a space of mappings which are unimodal in a strict sense.

Condition M1 says that the unique critical point 0 is mapped to 1. We want to consider ψ's which map 1 slightly – but not too far – to the left of 0. It may then be possible to find nonoverlapping intervals I_0 about 0 and I_1 near 1 which are exchanged by ψ. Technically, we proceed as follows: Write a for $-\psi(1) = -\psi^2(0)$ and b for $\psi(a)$; we suppress from the notation the dependence of a and b on ψ. Define $\mathcal{D}(T)$ to be the set of all ψ's in M such that:

D1. $a > 0$,

D2. $b > a$,

D3. $\psi(b) \leqslant a$.

The two intervals $I_0 = [-a, a]$ and $I_1 = [b, 1]$ are then nonoverlapping and ψ maps I_0 into I_1 and vice versa. If $\psi \in \mathcal{D}(T)$, then $\psi \circ \psi|_{I_0}$ has a single critical point, which is a minimum. By making the change of variables $x \longrightarrow -ax$, we replace I_0 by $[-1, 1]$ and the minimum by a maximum, i.e., if we define

$$T\psi(x) = -\frac{1}{a}\,\psi \circ \psi(-ax) \quad \text{for } x \in [-1, 1]$$

then $T\psi$ is again in M. Thus, T defines a mapping of $\mathcal{D}(T)$ into M. (In general, $T\psi$ need not lie in $\mathcal{D}(T)$. If a is small, then $T\psi(1)$ will be approximately 1 so $T\psi$ will not satisfy D1. On the other hand, if $\psi(b)$ is near a, then $T\psi(1)$ will be near -1 from which it follows that $T\psi$ does not satisfy D2.)

M. Feigenbaum [6] has proposed an explanation for some universal features displayed by infinite sequences of period doubling bifurcations based on some conjectures about T. We will not review has argument here; a version with due regard for mathematical technicalities may be found in Collet and Eckmann [3],

Received by the editors October 27, 1981.

1980 *Mathematics Subject Classification.* Primary 58F14.

[1]The author gratefully acknowledges the financial support of the Stiftung Volkswagenwerk for a visit to the IHES during which this paper was written, and the continuing financial support of the National Science Foundation (Grant MCS 78-06718).

0273-0979/81/0000-0086/$03.00

Collet, Eckmann and Lanford [4], or in Lanford [8]. The purpose of this note is to announce a proof of essentially all of these conjectures and to indicate the kind of analysis used.

2. Statement of results.

THEOREM 1. *There exists a function g, analytic and even on* $\{z \in \mathbf{C}: |z| < \sqrt{8}\}$ *whose restriction to* $[-1, 1]$ *is a fixed point for* T. *The Schwarzian derivative of g is negative on* $[-1, 1]$.

Let Ω denote $\{z \in \mathbf{C}: |z^2 - 1| < 2.5\}$ and write

$\mathfrak{H}$ for the Banach space of even functions bounded and analytic on Ω, real on real points, equipped with the supremum norm.

$\mathfrak{H}_0$ for the subspace of $\mathfrak{H}$ consisting of those functions vanishing (to second order) at 0.

$\mathfrak{H}_1$ for $\mathfrak{H}_0 + 1$.

PROPOSITION 2. *There is an open neighborhood* V *of g in* $\mathfrak{H}_1$ *such that*
Every $\psi \in V$ *is in* $\mathcal{D}(T)$ *(i.e., its restriction to* $[-1, 1]$ *is).*
If $\psi \in V$, $T\psi \in \mathfrak{H}_1$.
T *is infinitely differentiable as a mapping from* V *into* $\mathfrak{H}_1$,
The derivative $DT(\psi)$ *is compact operator on* $\mathfrak{H}_0$ *for each* $\psi \in V$.

THEOREM 3. $DT(g)$ *is hyperbolic on* $\mathfrak{H}_0$ *with one-dimensional expanding subspace; the expanding eigenvalue* δ *is positive.*

In other words: The spectrum of $DT(g)$ does not intersect the unit circle, and the part of the spectrum outside the unit circle consists of a single simple positive eigenvalue δ.

It then follows from invariant manifold theory that T admits locally invariant local stable and local unstable manifolds, of codimension one and dimension one respectively. Because of the noninvertibility of T, we do not construct global stable and unstable manifolds; we will let W_s and W_u denote respectively some particular local stable and local unstable manifolds.

Let Σ_0 denote the bifurcation surface for the simple period-doubling bifurcation. By this we mean the following: Any ψ in M has exactly one fixed point x_0 in $[0, 1]$; Σ_0 then denotes $\{\psi \in M: \psi'(x_0) = -1; (\psi \circ \psi)'''(x_0) < 0\}$. As a one-parameter family of ψ's crosses Σ_0 (in the appropriate direction) the fixed point x_0 loses stability in favor of an attracting orbit of period 2.

THEOREM 4. *There is a positive integer j and an element* g_j^* *of* W_u *such that* $T^j g_j^* \in \Sigma_0$.

Except for the difficulties in defining a global unstable manifold, we could formulate this theorem by saying that the unstable manifold crosses Σ_0. We

would like to know more, viz., that the crossing is transversal. This – properly formulated – is almost certainly true, but we have not proved it.

Let $\psi_\mu^{(0)}(x)$ denote the quadratic mapping $1 - \mu x^2$.

THEOREM 5. *There is a positive integer j and a parameter value μ_∞ (between* 1.4011550 *and* 1.4011554*) such that $\psi_\mu^{(0)}$ is in $\mathcal{D}(T^j)$ for μ sufficiently near to μ_∞ and such that the curve $T^j\psi_\mu^{(0)}$ crosses W_s transversally at $\mu = \mu_\infty$.*

Except for technicalities, this says that $\psi_\mu^{(0)}$ crosses the stable manifold transversally at μ_∞.

3. Remarks on the method of proof. The heart of the proof is a set of complicated numerical estimates proved rigorously with the aid of a computer. To formulate these estimates, we have first to establish some notation. We will work, initially, not in $\mathfrak{H}_1$ but in a subspace equipped with a stronger norm. The idea is that we want to write ψ as

$$\psi(x) = 1 - x^2 h(x^2)$$

and to use the l^1 norm for the Taylor coefficients of h at 1. Formally, given an element (u, v) of $\mathbf{R} \oplus l^1$, we associate with it an element ψ of $\mathfrak{H}_1$ by

$$\psi(x) = 1 - x^2\left\{u/10 + \sum_{n=1}^{\infty} v_n\left(\frac{(x^2 - 1)}{2.5}\right)^n\right\}. \tag{3.1}$$

We denote the set of ψ's obtained in this way by A, and we equip A with the norm $|u| + \Sigma|v_n|$. Note that A contains any element of $\mathfrak{H}_1$ which is analytic on the closure of Ω. (Of course, $\mathbf{R} \oplus l^1$ could have been identified with l^1, but we have singled out the u component – and introduced the factor of 10 in the formula (3.1) for $\psi(x)$ – for convenience later on.) For the remainder of this section, the norm of an element of A will always mean the norm of l^1 type just introduced.

The first step is to choose an explicit polynomial ψ_0 which will turn out to be a good approximate fixed point. We will take ψ_0 to be the polynomial of degree 20 defined by the first ten terms of the series given in Table 1 below. It can be checked without difficulty that

For any $\psi \in \mathsf{A}$ with $\|\psi - \psi_0\| < .01$, $T\psi \in \mathsf{A}$

T is infinitely differentiable from $\{\|\psi - \psi_0\| < .01\}$ to A.

For any ψ in this ball, $DT(\psi)$ is a compact operator on A.

Identifying A with $\mathbf{R} \oplus l^1$, we can represent $DT(\psi)$ as a matrix

$$\begin{pmatrix} \alpha(\psi) & \beta(\psi) \\ \gamma(\psi) & \delta(\psi) \end{pmatrix}$$

with $\alpha \in \mathbf{R}$; $\beta \in (l^1)^*$; $\gamma \in l^1$; $\delta \in L(l^1, l^1)$. In this notation, we can formulate

ESTIMATE 1. If $\|\psi - \psi_0\| < .01$, then

$$|\alpha - 4.669| < .148; \quad \|\beta\| < .560; \quad \|\gamma\| < .756; \quad \|\delta\| < .719.$$

These bounds imply that the inequality

$$[\alpha(\psi) - 1][1 - \|\delta(\psi)\|] > \|\beta(\psi)\| \cdot \|\gamma(\psi)\| \tag{3.2}$$

holds uniformly on the ball of radius .01 about ψ_0. If T has a fixed point g in this ball, then hyperbolicity of $DT(g)$ acting on A follows readily from (3.2).

To prove the existence of a fixed point, we use a variant of Newton's method. Instead of studying

$$\psi \mapsto \psi - (DT(\psi) - \mathbb{1})^{-1}[T\psi - \psi],$$

we replace $(DT(\psi) - \mathbb{1})^{-1}$ by the approximation

$$J = \begin{pmatrix} \dfrac{1}{3.669} & 0 \\ 0 & -\mathbb{1} \end{pmatrix},$$

and we apply the contraction mapping principle to the mapping

$$\psi \mapsto \Phi(\psi) = \psi - J \cdot [T\psi - \psi]$$

which has the same fixed points as T.

A simple calculation using Estimate 1 shows that

$$\|D\Phi(\psi)\| < .9 \quad \text{for } \|\psi - \psi_0\| < .01.$$

It will then follow from the contraction mapping theorem that Φ has a fixed point in this ball provided that

$$\frac{\|\Phi(\psi_0) - \psi_0\|}{1 - .9} < .01.$$

For this we have

ESTIMATE 2.

$$\|\Phi(\psi_0) - \psi_0\| < 4 \times 10^{-6}.$$

Thus T has a fixed point in A, and DT at the fixed point, acting on A, has the hyperbolicity properties stated in Theorem 3. Domains of analyticity may be enlarged using the functional equation for g, and in this way we arrive at Theorems 1 and 3 as formulated.

Furthermore, Estimate 1 makes it possible to establish the existence of a system of expanding and contracting cones for T on $\{\psi: \|\psi - \psi_0\| < .01\}$, which in turn makes it possible to construct local stable and unstable manifolds which are not too small. This facilitates the proofs of Theorems 4 and 5.

The proofs of Estimates 1 and 2 are completely straightforward, if long. Consider, for example, Estimate 1. Since A is essentially l^1, we can think of $DT(\psi)$ as an infinite matrix. Norms of matrices acting on l^1 are easy to compute in terms of the matrix elements. Any matrix element can be expressed in terms of ψ. All but finitely many of these matrix elements are estimated analytically. For the remainder, strict upper and lower bounds are computed numerically from bounds on the Taylor coefficients for ψ. The arithmetic operations are performed in finite precision floating point arithmetic; the methods of interval arithmetic are used to control the effect of round-off error.

4. Supplementary remarks. 1. The results described here are descendants of (and improvements on) the results announced in [7]. Since that announcement, a completely different proof for the existence of g has been given by Campanino, Epstein, and Ruelle [1].

2, The approach to proving Theorem 1 outlined in the preceding section produces strict bounds on the difference between an approximate fixed point and the exact one. These estimates can be applied to higher precision calculations. Let

$$g^{(0)}(x) = 1 + \sum_{n=1}^{40} g_n^{(0)} x^{2n}$$

where the $g_n^{(0)}$ are given by Table 1.

We then have strict bounds

$$|g(z) - g^{(0)}(z)| \leqslant \begin{cases} 1.5 \times 10^{-23} & \text{for } |z|^2 \leqslant 1.5, \\ 5.5 \times 10^{-13} & \text{for } |z|^2 \leqslant 2, \\ 5 \times 10^{-7} & \text{for } |z|^2 \leqslant 6, \\ 1.7 \times 10^{-2} & \text{for } |z|^2 \leqslant 8. \end{cases}$$

These bounds are probably very conservative.

3. The domain Ω used in the statements of Proposition 2 and Theorem 3 was chosen for convenience. Many other domains, including arbitrarily small open neighborhoods of $[-1, 1]$, could have been used instead. The hyperbolicity statement of Theorem 3 is formally stronger for small domains than for larger ones. (For $\Omega_1 \subset \Omega_2$, any eigenfunction for $DT(g)$ on Ω_2 is also an eigenfunction on Ω_1). It can be shown, however, that any function analytic on a neighborhood of $[-1, 1]$ and satisfying there the formal functional equation for an eigenvector of $DT(g)$ is actually analytic and bounded on the domain Ω.

4. It follows easily from the functional equation for g that g is analytic on a neighborhood of the whole real axis. H. Epstein (private communication) has observed that a similar argument shows that it is analytic on a neighborhood

n	$g_n^{(0)}$
1	−1.52763 29970 36301 45403 58903 10240
2	0.10481 51947 87303 73321 67426 13801
3	0.02670 56705 25193 35403 26520 94944
4	−0.00352 74096 60908 70917 02341 90769
5	0.00008 16009 66547 53174 51721 90486
6	0.00002 52850 84233 96353 61762 62552
7	$-2.55631\ 71662\ 78493\ 84635\ 32541 \times 10^{-6}$
8	$-9.65127\ 15508\ 91203\ 21637\ 25768 \times 10^{-8}$
9	$2.81934\ 63974\ 50409\ 13707\ 56629 \times 10^{-8}$
10	$-2.77305\ 11607\ 99011\ 72437 \times 10^{-10}$
11	$-3.02842\ 70221\ 30566\ 32983 \times 10^{-10}$
12	$2.67058\ 92807\ 48075\ 55396 \times 10^{-11}$
13	$9.96229\ 16410\ 28482\ 31059 \times 10^{-13}$
14	$-3.62420\ 29829\ 04156\ 08455 \times 10^{-13}$
15	$2.17965\ 77448\ 27070\ 47701 \times 10^{-14}$
16	$1.52923\ 28994\ 80962\ 60560 \times 10^{-15}$
17	$-3.18472\ 87899\ 52775 \times 10^{-16}$
18	$1.13467\ 21062\ 11871 \times 10^{-17}$
19	$1.88167\ 60568\ 25439 \times 10^{-18}$
20	$-2.27561\ 25646\ 32121 \times 10^{-19}$
21	$-9.82244\ 76294\ 21762 \times 10^{-22}$
22	$2.06412\ 97560\ 04508 \times 10^{-21}$
23	$-1.24932\ 00592\ 43689 \times 10^{-22}$
24	$-1.07706\ 12046 \times 10^{-23}$
25	$1.87274\ 68082 \times 10^{-24}$
26	$-2.57770\ 82101 \times 10^{-26}$
27	$-1.55419\ 04560 \times 10^{-26}$
28	$1.28044\ 34650 \times 10^{-27}$
29	$5.58505\ 87986 \times 10^{-29}$
30	$-1.52783\ 46925 \times 10^{-29}$
31	$5.04174\ 26639 \times 10^{-31}$
32	$1.01653\ 68070 \times 10^{-31}$
33	-1.00690×10^{-32}
34	-5.24253×10^{-34}
35	1.72437×10^{-34}
36	-1.31439×10^{-35}
37	-1.85830×10^{-38}
38	8.05506×10^{-38}
39	-6.26717×10^{-39}
40	1.76882×10^{-40}

TABLE 1.

of the imaginary axis. On the other hand, it is essentially certain that g is not entire. Indeed, it appears – but has not been proved – that the singularities of g nearest to the origin occur at a set of 4 periodic points of period 2 for $z \mapsto g(-\lambda z)$, $\lambda = -g(1)$, located approximately at

$$z^2 = -3.8428 \pm i\ 9.8215.$$

5. Proposition 2 and Theorem 3 remain true if the requirement that ψ be even is dropped. In other words: No new expanding eigenvectors are introduced if we let $DT(g)$ act on functions which are not necessarily even (but which vanish *to second order* at 0).

6. Theorem 4 can be extended considerably. To formulate the extension, we need the theory of kneading sequences for unimodal mappings as developed, for example, in Chapter III.1 of Collet-Eckmann [3]. Let $\underline{K}$ be a finite kneading sequence. Except for the simple case $\underline{K} = RC$, there are associated with $\underline{K}$ three hypersurfaces in M:

The set of superstable ψ's with kneading sequence $\underline{K}$.

The saddle-node or period-doubling bifurcation surface where the attracting periodic orbit passing through the critical point on the preceding surface appears.

The period-doubling bifurcation surface where that periodic orbit becomes unstable.

It can be shown that, intuitively, the unstable manifold crosses these three surfaces for each $\underline{K}$; a precise version of this statement must be formulated with the same circumspection as Theorem 4. There is no reason to doubt that these crossings are all transverse.

A simple argument using the apparatus developed in [3] reduces the proof of Theorem 4 and the above extension to establishing the existence, on the local unstable manifold, of one point whose kneading sequence strictly precedes, and one whose kneading sequence strictly follows, that of g (in the combinatorial ordering for kneading sequences). The proof proceeds by finding with sufficient precision two points on the unstable manifold and computing initial segments of their kneading sequences.

7. Although done by computer, the computations involved in proving the results stated are just on the boundary of what it is feasible to verify by hand. I estimate that a carefully chosen minimal set of estimates sufficient to prove Theorems 1 and 3 could be carried out, with the aid only of a nonprogrammable calculator, in a few days.

ACKNOWLEDGEMENTS. It is a pleasure for me to thank:

P. Collet, J. P. Eckmann, H. Epstein, D. Ruelle, and S. Smale for helpful discussions and encouragement.

L. Michel for his assistance in making available the computing facilities needed to carry out this work.

Director N. Kuiper for his very gracious hospitality at the IHES.

The Stiftung Volkswagenwerk for financial support during my visit to the IHES.

The National Science Foundation for continuing financial support under Grant MCS 78-06718.

REFERENCES

1. M. Campanino, H. Epstein and D. Ruelle, *On Feigenbaum's functional equation,* (IHES preprint P/80/32 (1980)) Topology (to appear).

2. M. Campanino and H. Epstein, *On the existence of Feigenbaum's fixed point,* (IHES preprint P/80/35 (1980)) Comm. Math. Phys. (1981), 261–302.

3. P. Collet and J. P. Eckmann, *Iterated maps of the interval as dynamical systems,* Birkhäuser, Boston-Basel-Stuttgart, 1980.

4. P. Collet, J. P. Eckmann and O. E. Lanford, *Universal properties of maps on an interval,* Comm. Math. Phys. **76** (1980), 211–254.

5. M. Feigenbaum, *Quantitative universality for a class of non-linear transformations,* J. Statist. Phys. **19** (1978), 25–52.

6. ———, *The universal metric properties of non-linear transformations,* J. Statist. Phys. **21** (1979), 669–706.

7. O. E. Lanford, *Remarks on the accumulation of period-doubling bifurcations,* Mathematical Problems in Theoretical Physics, Lecture Notes in Physics, vol. 116, Springer-Verlag, Berlin and New York, 1980, pp. 340–342.

8. ———, *Smooth transformations of intervals,* Séminaire Bourbaki, 1980/81, No. 563, Lecture Notes in Math., vol. 901, Springer-Verlag, Berlin, Heidelberg and New York, 1981, pp. 36–54.

INSTITUT DES HAUTES ETUDES SCIENTIFIQUES, 35, ROUTE DE CHARTRES, 91440, BURES-SUR-YVETTE, FRANCE

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, BERKELEY, CALIFORNIA 94720 (Current address)

Journal of Statistical Physics, Vol. 34, Nos. 1/2, 1984

Functional Equations for Circle Homeomorphisms with Golden Ratio Rotation Number

Oscar E. Lanford III[1]

Received August 31, 1983

The investigation of a scaling limit for mappings of the circle to itself with golden ratio rotation number leads to a pair of functional equations with at least a formal resemblance to the functional equation using the accumulation of period-doubling bifurcations. We discuss the general theory of these functional equations, assuming that solutions exist.

KEY WORDS: Golden-ratio; rotation number; functional equation; renormalization group.

1. INTRODUCTION

The investigation of a "scaling limit" for iterates of mappings of the circle with rotation number equal to the golden ratio $(\sqrt{5}-1)/2$ and with a critical point leads to the functional equations

$$g(x) = \alpha g\big(g(\alpha^{-2}x)\big) \tag{A1}$$

$$g(x) = \alpha^2 g\big(\alpha^{-2} g(\alpha^{-1}x)\big) \tag{B}$$

(The analysis leading to these equations will be outlined in Section 2.) Here, α is a number and $g(x)$ a function defined on some interval; both α and $g(x)$ are to be determined. We are going to consider only solutions with (i) $\alpha < -1$, (ii) $g(x)$ a strictly decreasing function of x, and (iii) $g(0) = 1$. The condition $g(0) = 1$ is simply a normalization; if $\bar{g}(x)$ is a solution of

[1] IHES, 91440 Bures-sur-Yvette, France.

[2] Work supported in part by the National Science Foundation.

0022-4715/84/0100-0057$03.50/0

either equation with $\bar{g}(0) \neq 0$, then

$$g(x) = (\bar{g}(0))^{-1} \bar{g}(\bar{g}(0) \cdot x)$$

is also a solution (with the same α) satisfying $g(0) = 1$. We stress that the condition that g is strictly decreasing means

$$g(x_1) < g(x_2) \qquad \text{whenever} \quad x_1 > x_2$$

it does not rule out the existence of isolated points where g has vanishing derivative.

The analysis leading to the equations (A1) and (B) suggests that there should exist a function $g(x)$ which satisfies *both* of these equations. Furthermore, either of these equations, by itself, can be solved numerically, and the solution found appears, within computational error, to satisfy the other equation "automatically." These considerations lead to the suspicion that, under appropriate conditions, (A1) and (B) might be equivalent. A major step in the direction of proving this was made by Nauenberg,[4] who showed that an analytic solution of (A1) which satisfies in addition

$$g(\alpha^{-2}) = \alpha^{-2} \tag{A2}$$

must also satisfy (B). We will prove in Section 4 that, under mild differentiability and domain conditions, a solution of (B) also satisfies (A1). Note that condition (A2) follows immediately from (B) by putting $x = 0$ and using $g(0) = 1$. Thus, schematically, (A1) and (A2) together are equivalent to (B). *From now on, we will write* (A) *to denote the combination of conditions* (A1) *and* (A2).

It may be that (A2) is actually a consequence of (A1). It is not difficult to deduce from (A1) that $g^2(\alpha^{-2}) = \alpha^{-2}$, i.e., that α^{-2} is either a fixed point for g or a periodic point of period 2. The only solutions to (A1) (and the subsidiary conditions above) found so far do satisfy (A2), but, so far as I know, no general proof has been given.

Numerical computation of a solution to (A1) or (B) produces a function defined on a finite interval (or a bounded domain in the complex plane). In the analogous case of the functional equation for period-doubling (see, e.g., Collet and Eckmann[1]), the functional equation itself gives immediately an extension of the solution to the whole real axis. We investigate in Section 3 the problem of extending solutions of (B) defined on a finite interval to a larger set. What we show is that any solution defined on an interval of finite but reasonable size can be extended (uniquely) either to a solution defined on the whole real axis or to one defined on a finite interval which goes to infinity at both ends of that interval. The extended solution is as regular as the one we started from.

We also show that any solution arising as a scaling limit of a circle mapping with rotation number $(\sqrt{5} - 1)/2$ is extendable to the whole real axis.

Section 5 is also devoted to an extension argument. As Feigenbaum et al.[2] have remarked, equation (A)—in contrast to (B)—makes sense for functions defined only on the interval $[0, 1]$. We show that any solution of (A) on $[0, 1]$ can be extended to a function defined on a considerably larger interval and satisfying both (A) and (B). (The extension theorem for solutions of (B) can then be used to continue this solution still further.) The arguments used in this section are straightforward adaptations of ideas in Nauenberg[4]; the main interest of the argument we give is that it separates clearly the algebraic and analytic elements in Nauenberg's argument.

2. THE SCALING LIMIT FOR CIRCLE MAPPINGS

In this section we describe some fine structure in the behavior of certain mappings of the circle to itself. Almost nothing in this section is new; we will summarize, from a slightly different point of view, the results of Feigenbaum, Kadanoff, and Shenker.[2] An alternative, and more or less equivalent, development has been given by Ostlund, Rand, Sethna, and Siggia.[5]

Let $f(x)$ denote a continuous strictly increasing mapping of the real line $\mathbf{R}$ to itself satisfying

$$f(x + 1) = f(x) + 1 \tag{2.1}$$

By passage to quotients, such an f induces a one–one continuous mapping of the circle—represented as $\mathbf{R}/\mathbf{Z}$—onto itself, and any such mapping which is orientation preserving (increasing) is induced by an f which is unique up to an additive integer constant.

If f is as above, then

$$\lim_{n\to\infty} \frac{f^n(x) - x}{n}$$

exists for all x and is independent of x. This limit is called the *rotation number* of f; we will denote it by $\rho(f)$.

The mappings of the circle we want to discuss will be ones induced by f's with rotation number $(\sqrt{5} - 1)/2$, the *golden ratio*. For the remainder of this paper, the symbol σ will be reserved to denote $(\sqrt{5} - 1)/2$. We need some preliminaries on the relation between σ and the *Fibonacci sequence*, i.e., the sequence (Q_n) of integers satisfying

$$Q_{n+1} = Q_n + Q_{n-1}, \qquad Q_0 = 0, \qquad Q_1 = 1 \tag{2.2}$$

A well-known but relatively crude relation is

$$\lim_{n\to\infty} \frac{Q_{n-1}}{Q_n} = \sigma$$

This relation is sharpened considerably by the identity

$$Q_n \cdot \sigma - Q_{n-1} = (-1)^{n-1}\sigma^n \tag{2.3}$$

which can easily be proved by solving the recursion relation defining the Fibonacci sequence to get the explicit formula

$$Q_n = \left(\sigma^{-n} + (-1)^{n+1}\sigma^n\right)/(\sigma^{-1} + \sigma)$$

We give the identity (2.3) a "dynamical" interpretation as follows: Let $f(x)$ have rotation number σ. It follows at once from the definition of rotation number that

$$f_{(n)}(x) \equiv f^{Q_n}(x) - Q_{n-1}$$

has rotation number

$$Q_n \cdot \sigma - Q_{n-1} = (-1)^{n-1}\sigma^n$$

For n large, the rotation number of $f_{(n)}$ is small, so we might expect

$$f_{(n)}(x) \to x \qquad \text{as} \quad n \to \infty$$

We are going to look in detail at the limiting behavior of the $f_{(n)}$'s for functions f which are smooth but which, although strictly increasing, have zero as a critical point:

$$f'(0) = 0$$

and hence whose inverses are not smooth. A typical example is

$$f(x) = x + \omega_0 - \frac{1}{2\pi}\sin(2\pi x)$$

with the constant ω_0 adjusted to make the rotation number equal to σ. We want to concentrate particularly on the behavior near the critical point at zero, and we therefore magnify as follows: Let

$$\alpha^{(n)} = f_{(n)}(0)^{-1}$$

and

$$f_n(x) = \alpha^{(n-1)} f_{(n)}(x/\alpha^{(n-1)})$$

(We have chosen to rescale by $\alpha^{(n-1)}$ rather than $\alpha^{(n)}$ because this leads to slightly simpler formulas later on.) In the situation we want to look at, the $\alpha^{(n-1)}$ will tend to infinity, so the large-n behavior of f_n's on a fixed interval will reflect the behavior of $f_{(n)}$ and hence f^{Q_n} very near to zero.

As we have already remarked, $f_{(n)}$ has rotation number $(-1)^{n-1}\sigma^n$. This implies that

$$f_{(n)}(x) > x \qquad \text{everywhere for } n \text{ odd}$$
$$f_{(n)}(x) < x \qquad \text{everywhere for } n \text{ even}$$

and in particular that $\alpha^{(n)}$ has sign $(-1)^{n-1}$. We also note for use later that

$$f_n(x) < x \qquad \text{for all } n, x \tag{2.4}$$

Numerical experiments suggest that the following phenomenology is common for mappings f of the sort we are considering:

(i) The sequence of ratios $\alpha^{(n+1)}/\alpha^{(n)}$ approaches a limit $\alpha < -1$ (so that, roughly, $\alpha^{(n)} \approx \alpha^n$).

(ii) The sequence of functions f_n approaches a limit f^*.

(iii) The limiting ratio α and function f^* are *universal*, i.e., do not seem to depend on what f we start with in the class we are considering.

Note that the condition $f(x+1) = f(x)$ imposed on f will (presumably) *not* be satisfied by f^*. The functions f_n satisfy

$$f_n(x + \alpha^{(n-1)}) = f_n(x) + \alpha^{(n-1)}$$

and, since the $\alpha^{(n-1)}$ go to infinity, the condition becomes empty in the limit.

Independent of questions of universality, some striking consequences follow just from the convergence of the sequence f_n derived from a single f with rotation number σ. Suppose we have such an f. Write

$$\alpha_n = \alpha^{(n)}/\alpha^{(n-1)}$$

Then, first of all,

$$f_n(0) = \alpha^{(n-1)} f_n(0) = \alpha^{(n-1)}/\alpha^{(n)} = \alpha_n^{-1}$$

and so

$$\alpha^{-1} = \lim_{n\to\infty} f_n(0) = f^*(0)$$

From

$$Q_{n+1} = Q_n + Q_{n-1}$$
$$f_{(n)} = f^{Q_n} - Q_{n-1}$$
$$f(x+1) = f(x) + 1$$

it follows easily that

$$f_{(n+1)} = f_{(n)} f_{(n-1)} = f_{(n-1)} f_{(n)}$$

Rescaling by $\alpha^{(n)}$ and reorganizing in a straightforward way we get

$$f_{n+1}(x) = \alpha_n f_n\big(\alpha_{n-1} f_{n-1}\big(\alpha_n^{-1}\alpha_{n-1}^{-1} x\big)\big) \tag{2.5A}$$

from the first equality and

$$f_{n+1}(x) = \alpha_n \cdot \alpha_{n-1} f_{n-1}\left(\alpha_{n-1}^{-1} f_n\left(\alpha_n^{-1} x\right)\right) \tag{2.5B}$$

from the second. Taking limits as $n \to \infty$ we find

$$f^*(x) = \alpha f^*\left(\alpha f^*\left(\alpha^{-2} x\right)\right) \tag{2.6A}$$

$$f^*(x) = \alpha^2 f^*\left(\alpha^{-1} f^*\left(\alpha^{-2} x\right)\right) \tag{2.6B}$$

To summarize: *Using nothing but the assumption that appropriately magnified versions of* $f^{Q_n}(x) - Q_{n-1}$ *converge to a limit* f^*, *we have shown that* f^* *satisfies two functional equations.* Since solutions of these functional equations can be expected to be scarce, we get a simple intuitive explanation for "universality," i.e., the fact that many different f's produce the same f^*. Renormalization group ideas lead to an elaboration of this simple explanation for which we refer to Feigenbaum et al.,[2] MacKay,[3] or Ostlund et al.[5]

To get the functional equations (A1) and (B) as given in Section 1, we write

$$g(x) \equiv \alpha f^*(x)$$

and rewrite (2.6A) and (2.6B) accordingly. Since $\alpha^{-1} = f^*(0)$, the normalization condition $g(0) = 1$ is automatically satisfied. Since each f_n is increasing, f^* will be an increasing function so $g(x)$ will be a decreasing function. (Recall that α is negative and larger than one in magnitude.) We observed above that $f_n(x) < x$ for all n, x; hence,

$$f^*(x) \leqslant x \quad \text{and} \quad g(x) \geqslant \alpha x \tag{2.7}$$

at all points x where

$$f^*(x) = \lim_{n \to \infty} f_n(x)$$

The purpose of this paper is to investigate the properties of solutions of the equations (A1) and (B), assuming that the solutions exist. It may nevertheless be useful to sketch briefly the state of our knowledge about existence of analytic solutions. There is, first of all, a "trivial" solution

$$g(x) = 1 - x/\sigma, \qquad \alpha = -1/\sigma$$

This solution is what is obtained by taking the scaling limit described above for a smooth mapping with rotation number σ and derivative strictly positive everywhere. What we really want is a solution with $g'(0) = 0$. Monotonicity of g then implies $g''(0) = 0$, and it is natural to seek a solution with $g'''(0) \neq 0$. Assuming this, it is straightforward to show from either functional equation that $g^{(j)}(0) = 0$ for all j which are not multiples of 3, i.e., that g is an analytic function of x^3.

The usual way to solve (A1) or (B) numerically is to (i) rewrite the functional equation in terms of the variable x^3, (ii) truncate the functional equation to a finite number of conditions either by requiring only that it hold at a finite set of points or by requiring only that it hold to finite order at some point, (iii) use Newton's method to seek a polynomial satisfying the resulting finite set of nonlinear equations.

With a little care, Newton's method can be made to converge, and the result obtained always seems to be the same (i.e., independent of which equation is being solved and—within limits— of the truncation procedure used.)

These numerical results strongly suggest that there is function defined and analytic on a substantial interval of the real axis, satisfying (A) and (B), with

$$\alpha \approx -1.2885745\ldots$$

and with a second-order critical point at 0. It appears to be feasible to give a computer-assisted proof that this solution does in fact exist, and there are currently in progress at least two attempts to construct such a proof, one by D. Rand and B. Mestel, the other by R. de la Llave and the author.

3. EXTENSION OF SOLUTIONS OF (B)

We are going to prove in this section the existence of a maximal extension of a given solution of (B). We have first to specify carefully what we mean by a solution. We will, as always, only be concerned with *continuous*, *monotone-decreasing* solutions with $\alpha < -1$. Let $g(x)$ be a function defined on an interval I. For the sake of concreteness, we will write the formulas as if I is a closed interval $[a, b]$, but open and half-open intervals are also allowed. The domain I' of the function $\alpha^2 g(\alpha^{-2} g(\alpha^{-1} x))$, i.e., the set of numbers x such that

$$a \leqslant \alpha^{-1} x \leqslant b \quad \text{and} \quad a \leqslant \alpha^{-2} g(\alpha^{-1} x) \leqslant b$$

is either an interval or empty. For g to be a solution to (B), it is necessary at the very least that the following be true:

(1) $I \cap I'$ contains more than one point.

(2) $g(x)$ and $\alpha^2 g(\alpha^{-2} g(\alpha^{-1} x))$ agree on $I \cap I'$.

Condition (1) implies that 0 is in I, so the normalization condition $g(0) = 1$ makes sense. For our purposes, we add the following to these minimal conditions:

(3) The interval of definition $I = [a, b]$ is not too asymmetric:

$$|\alpha|^{-1} \leqslant \frac{|b|}{|a|} \leqslant |\alpha|$$

(4) $g(b) < 0$.

We will explain below our reasons for imposing these conditions. First, however, we note the following consequence:

Proposition 3.1. With the above assumptions, the equation

$$g(x) = \alpha^2 g\big(\alpha^{-2} g(\alpha^{-1} x)\big) \tag{B}$$

holds for all $x \in [a, b]$.

In other words, $[a, b]$ is a *self-defining interval* for (B).

Proof. We define

$$q(x) \equiv \alpha^{-2} g(\alpha^{-1} x)$$

so (B) reads

$$g(x) = \alpha^2 g(q(x))$$

Note that $q(x)$ is *increasing*. Also, let $I \cap I' = [a'', b'']$. What we want to show is that $a'' = a$, $b'' = b$.

From (3), $q(x)$ is defined on all of I, and from this it follows that either $b'' = b$ or $q(b'') = b$. If $b'' < b$, then

$$g(b'') = \alpha^2 g(q(b'')) = \alpha^2 g(b)$$

But $g(b) < 0$ by (4), and $\alpha^2 > 1$, so it follows that $g(b'') < g(b)$. Since $g(x)$ is decreasing and $b'' < b$, this is impossible, and so $b'' = b$. It can be shown in a similar way that $a'' = a$ [using, this time, the fact that $g(a) \geqslant g(0) = 1 > 0$]. ■

We can now explain why we impose (3). Suppose, for example, that $g(x)$ is defined on $[a, b]$ with $|a| > |\alpha|\,|b|$, (i.e., with $a < \alpha b$) and satisfies (B) on the interval $[a'', b'']$ where both sides are defined. The argument just given can readily be adapted to show that $a'' = \alpha b$ and that $q(a'') > a''$. Thus, the values taken on by $q(x)$ for $a'' \leqslant x \leqslant b''$ are all to the right of αb. Hence the values of $g(x)$ for $a \leqslant x < \alpha b$ do not in any way enter into Eq. (B), so $g(x)$ can be changed arbitrarily on that interval and remain a solution of (B). Condition (3) avoids this arbitrariness. Note that, in the case considered, we could simply replace a by αb; the restricted g would then satisfy (3).

We impose (4) simply because the functional equation has the form

$$g(x) = \alpha^2 g(q(x))$$

and thus offers no straightforward way to determine any negative values of g starting from positive values only.

Proposition 3.2. A solution of (B) satisfying conditions (1)–(4) above can be extended uniquely to a function defined and satisfying (B)

everywhere on an interval (a_∞, b_∞) with either

$$a_\infty = -\infty, \qquad b_\infty = \infty$$

or

$$\lim_{x \downarrow a_\infty} g(x) = +\infty, \qquad \lim_{x \uparrow b_\infty} g(x) = -\infty$$

Proof. From $g(0) = 1$, $g(b) < 0$, it follows that there is an x_0 in $[0, b]$ with

$$g(x_0) = 0$$

From

$$g(x) = \alpha^2 g(q(x))$$

and the fact that g is decreasing, it follows that

$$\begin{aligned} q(x) &< x \qquad \text{if} \quad x > x_0 \\ q(x) &> x \qquad \text{if} \quad x < x_0 \end{aligned} \tag{3.1}$$

We will use these inequalities repeatedly.

Consider now the functional equation written in the form

$$g(x) = \alpha^2 g(q(x)) \tag{3.2}$$

We will use the right-hand side of this equation to extend g and so want to show that its domain of definition, which we denote by $[a_1, b_1]$, properly contains the original domain $[a, b]$ of g. This is immediate: Since $q(x)$ is defined on the interval $[\alpha b, \alpha a]$ which properly contains $[a, b]$, and, by (3.1),

$$q(a) > a, \qquad q(b) < b$$

We thus use the functional equation to extend g to $[a_1, b_1]$; the extended g is still strictly decreasing and is as regular as the unextended g. The functional equation holds, by construction, everywhere on $[a_1, b_1]$, and it is easy to check that condition (3) is preserved, i.e., that

$$|\alpha|^{-1} \leqslant \frac{|a_1|}{|b_1|} \leqslant |\alpha|$$

Note that (i) if $b_1 \neq \alpha a$ then $q(b_1) = b$ and (ii) if $a_1 \neq \alpha b$ then $q(a_1) = a$.

The extension process can be repeated to generate a solution of (B) defined on an increasing sequence of intervals

$$[a, b] \subset [a_1, b_1] \subset [a_2, b_2] \subset \cdots$$

and hence defined on (a_∞, b_∞) where

$$a_\infty = \lim_{n \to \infty} a_n, \qquad b_\infty = \lim_{n \to \infty} b_n$$

It is clear from the construction that the extension is unique.

Since

$$|\alpha|^{-1} \leqslant \frac{|a_n|}{|b_n|} \leqslant |\alpha| \tag{3.3}$$

both a_∞ and b_∞ are finite if one of them is. To complete the proof of the proposition, we need only show that, if a_∞ and b_∞ are finite, then

$$\lim_{x \downarrow a_\infty} g(x) = +\infty, \qquad \lim_{x \uparrow b_\infty} g(x) = -\infty$$

We will give the proof only for the first of these relations; the second is completely analogous.

From (3.3),

$$\alpha b_\infty \leqslant a_\infty$$

and we start by considering the case

$$\alpha b_\infty < a_\infty$$

Then, for sufficiently large n,

$$\alpha b_{n-1} < a_\infty \leqslant a_n$$

But, as we noted above (in the case $n = 1$), if $a_n \neq \alpha b_{n-1}$, then $q(a_n) = a_{n-1}$. Combining this remark with the functional equation, we get that

$$g(a_n) = \alpha^2 g(q(a_n)) = \alpha^2 g(a_{n-1})$$

for all sufficiently large n. Since $g(a_n) > 0$ for all n, it follows that

$$\lim_{n \to \infty} g(a_n) = \lim_{x \downarrow a_\infty} g(x) = \infty$$

We now eliminate the hypothesis

$$\alpha b_\infty < a_\infty$$

by showing that the alternative

$$\alpha b_\infty = a_\infty$$

leads to a contradiction. Thus, suppose this latter equality holds. Then

$$\alpha a_\infty > b_\infty$$

and hence, by an argument analogous to the one just given,

$$\lim_{x \uparrow b_\infty} g(x) = -\infty$$

But, as $x \downarrow a_\infty$, $\alpha^{-1} x \uparrow b_\infty$, so $q(x) = \alpha^2 g(\alpha^{-1} x) \downarrow -\infty$. This, however, eventually contradicts the fact (3.1) that $q(x) > x$ for $x < x_0$, and thus completes the proof. ■

To close this section, we show that, for a solution $g(x)$ obtained as outlined in Section 2 as a scaling limit of a circle map with golden ratio rotation number, a_∞ and b_∞ are necessarily infinite.

Proposition 3.3. Let the hypotheses be as in Proposition 3.2, but suppose in addition that there is an increasing function $f(x)$ with $f(x+1) = f(x) + 1$ and with rotation number σ, such that (in the notation of Section 2) $f_n(x)$ converges to $\alpha^{-1}g(x)$ uniformly on compact sets in (a,b). Then $a_\infty = -\infty$, $b_\infty = +\infty$, and convergence is uniform on compact sets in $(-\infty, \infty)$.

Proof. If we can show that $f_n(x)$ converges to $\alpha^{-1}g(x)$ uniformly on compact sets in (a_∞, b_∞), it will follow from (2.7) that

$$g(x) \geqslant \alpha x$$

on (a_∞, b_∞). This bound, however, rules out the possibility that $g(x)$ goes to $-\infty$ for finite x; hence, it shows that b_∞ (and therefore also a_∞) must be infinite.

To prove uniform convergence on compact sets in (a_∞, b_∞), it suffices, by an obvious induction argument, to prove it on compact sets in (a_1, b_1).

Recall (2.5B):

$$f_{n+1}(x) = \alpha_n \cdot \alpha_{n-1} f_{n-1}\left(\alpha_{n-1}^{-1} f_n\left(\alpha_n^{-1}x\right)\right)$$

What we want to show is that the left-hand side of this identity converges uniformly on compact sets in (a_1, b_1) to $\alpha^{-1}g(x)$; we do this by proving convergence of the right-hand side to $\alpha g(q(x))$. Thus, let K be a compact set in (a_1, b_1). By the construction of the extension of g to (a_1, b_1), $\alpha^{-1}K$ and $q(K)$ are both contained in (a,b). Let K_1 be a compact set in (a,b) containing both $\alpha^{-1}K$ and $q(K)$ in its interior. For sufficiently large n, $\alpha_n^{-1}K$ is contained in K_1, so, from the uniform convergence of $f_n(x)$ to $\alpha^{-1}g(x)$ on K_1, it follows that $q_n(x) \equiv \alpha_{n-1}^{-1} f_n(\alpha_n^{-1}x)$ converges uniformly on K to $q(x)$. Similarly, $q_n(K) \subset K_1$ for all sufficiently large n, so uniform convergence of the right-hand side of (2.5) on K follows from the uniform convergence of f_{n-1} to $\alpha^{-1}g$ on K_1. ■

4. (B) IMPLIES (A)

In this section, $g(x)$ will denote a solution of (B) in the sense of Section 3, i.e., a strictly decreasing function with $g(0) = 1$, defined on an interval (a,b) and satisfying conditions (1)–(4) of Section 3. Invoking Proposition 3.2, we assume the g is maximally extended. As in Section 3, we let x_0 denote the unique point is (a,b) where $g(x_0) = 0$.

Proposition 4.1. If g is differentiable, if $g'(x_0) \neq 0$, and if the domain (a,b) of g contains 1, then

$$g(x) = \alpha g\left(g(\alpha^{-2}x)\right) \tag{A1}$$

everywhere on (a,b).

Remarks. 1. The assumption that $g'(x_0) \neq 0$ can be replaced by the assumption that, for some positive integer q, g is q times differentiable and $g^{(q)}(x_0) \neq 0$. (This extension seems, however, to be without practical interest.)

2. We actually need differentiability only at the two points x_0 and $\alpha^{-2}x_0$.

3. It is not very satisfactory to have to assume explicitly that g is defined at 1. (Note, however, that, if g is not defined at 1, the right-hand side of (A1) is not defined at 0.) This assumption can be replaced by the weaker one that the right and left sides of (A1) have at least one point of definition in common, but it has not, so far, been possible to eliminate it completely.

Proof. As in Section 3, we write

$$q(x) \equiv \alpha^{-2} g(\alpha^{-1} x)$$

We also write

$$r(x) \equiv g(\alpha^{-2} x)$$

The functional equations (A1) and (B) can now be rewritten, respectively, as

$$g(r(x)) = \alpha g(x) \tag{A1'}$$

$$g(q(x)) = \alpha^2 g(x) \tag{B'}$$

The obvious way for these equations to be consistent is to have

$$q(x) = r(r(x))$$

This identity is in fact a straightforward consequence of (B):

$$r(r(x)) = g\big(\alpha^{-2} g(\alpha^{-2} x)\big) = g\Big(\alpha^{-2} g\big(\alpha^{-1}(\alpha^{-1} x)\big)\Big)$$

$$= \alpha^{-2} g(\alpha^{-1} x) = q(x)$$

Note that, in this derivation, we have used (B) only at the point $\alpha^{-1}x$, so the identity holds for all x such that $\alpha^{-1}x \in (a, b)$, i.e., for all x in the domain of q.

We next argue that

$$r(x_0) = x_0$$

To see this, we note that r is a decreasing function on $[0, 1]$ with $r(0) = 1$ and so has exactly one fixed point in $[0, 1]$. This point is also fixed for $q = r^2$, and the only fixed point for q in (a, b) is x_0. [This argument is the only place where we need the assumption that g is defined at 1.]

Next: We differentiate (B) at x_0 and use $q(x_0) = x_0$ to get

$$g'(x_0) = \alpha^2 g'(x_0) q'(x_0)$$

and hence, since we are assuming $g'(x_0) \neq 0$,

$$q'(x_0) = \alpha^{-2}$$

Since $q = r^2$ and $r(x_0) = x_0$,

$$(r'(x_0))^2 = q'(x_0) = \alpha^{-2}$$

so, since r is decreasing and α negative

$$r'(x_0) = \alpha^{-1}$$

For any $x \in (a, b)$,

$$g(x) = \alpha^2 g(q(x)) = \cdots = \alpha^{2m} g(q^m(x))$$

Hence, as m goes to infinity, $g(q^m(x))$ goes to zero, so, since g is strictly decreasing,

$$q^m(x) \to x_0$$

We are first going to prove (A) assuming both $x \in (a, b)$ and $r(x) \in (a, b)$; then show that the second assumption is automatic. With the two assumptions, and using $q = r^2$, we get

$$\alpha g(r(x)) = \alpha^{2m+1} g(q^m(r(x))) = \alpha^{2m+1} g(r(q^m(x)))$$

We are going to show that

$$\lim_{m \to \infty} \alpha^{2m+1} g(r(q^m(x))) = g(x)$$

and hence that

$$\alpha g(r(x)) = g(x)$$

as desired.

Write x_m for $q^m(x)$. Then using $g(x_0) = 0$,

$$\alpha^{2m+1} g(r(x_m)) = \alpha \frac{g(r(x_m)) - g(x_0)}{r(x_m) - x_0} \cdot \frac{r(x_m) - x_0}{x_m - x_0} \cdot \alpha^{2m}(x_m - x_0) \quad (4.1)$$

On the other hand, by repeated application of (B),

$$g(x) = \alpha^{2m} g(x_m) = \frac{g(x_m) - g(x_0)}{x_m - x_0} \cdot \alpha^{2m}(x_m - x_0)$$

and so

$$\lim_{m \to \infty} \alpha^{2m}(x_m - x_0) = g(x)/g'(x_0)$$

Combining this with (4.1) and using $r'(x_0) = \alpha^{-1}$, we get

$$\lim_{m \to \infty} \alpha^{2m+1} g(r(x_m)) = \alpha g'(x_0) r'(x_0)\, g(x)/g'(x_0) = g(x)$$

This proves (A1) provided that both x and $r(x)$ are in the interval (a, b) on which g is defined. We now show that r maps this interval into

itself so the second condition is automatic. Suppose not. Then a and b must be finite and either a or b must be the image under r of some point in (a, b). For definiteness, assume that there is an $x_a \in (a, b)$ with $r(x_a) = a$. Since r is decreasing, $x_0 < x_a < b$. By what we have just shown,

$$g(x) = \alpha g(r(x)) \qquad \text{for} \quad x_0 < x < x_a$$

Let x approach x_a from below. Then $r(x)$ approaches a from above, so $g(r(x))$ approaches ∞, so $g(x) = \alpha g(r(x))$ approaches $-\infty$. But this contradicts the assumption that $x_a < b$. ■

5. (A) IMPLIES (B)

As noted in Section 1, Nauenberg[4] has shown that an analytic solution of

$$g(x) = \alpha g\big(g(\alpha^{-2}x)\big), \qquad g(\alpha^{-2}) = \alpha^{-2} \tag{A}$$

also satisfies (B). (His argument can actually be made to work assuming much less than analyticity.) We will prove here a complementary result, purely algebraic, showing that a solution of (A) on [0, 1] can be extended to a larger interval where it satisfies both (A) and (B).

For this section, our assumptions will be that g is a strictly decreasing function defined on [0, 1], with $g(0) = 1$, satisfying (A) on [0, 1]. As always, we assume $\alpha < -1$.

It follows from (A1) with $x = 0$ that

$$\alpha^{-1} = g(1)$$

hence, that $g(1) < 0$; hence, that g vanishes somewhere on [0, 1]. As previously, we denote this point by x_0. Note that the right-hand side of (A1) is defined for x up to $\alpha^2 x_0$, so we can take it as defining an extension of g to $[0, \alpha^2 x_0]$, and the extended function will still satisfy (A). We will from now on assume that this extension has been made.

At this point, it makes very little sense to ask whether g satisfies

$$g(x) = \alpha^2 g\big(\alpha^{-2} g(\alpha^{-1}x)\big) \tag{B}$$

the left-hand side is defined only for positive x and the right-hand side only for negative x. The only place where the two sides can be compared is at 0, and there (B) reduces to $g(\alpha^{-2}) = \alpha^{-2}$, i.e., to (A2). The left-hand side of (B) can, however, be taken as *defining* $g(x)$ for $\alpha x_0 \leqslant x < 0$. (B) then holds, by definition, on that interval. It is also possible to ask whether (i) (A1) holds for $\alpha x_0 \leqslant x < 0$ or (ii) (B) holds for $0 < x \leqslant \alpha^2 x_0$.

We are going to show, by straightforward verifications, that the answers to both questions are affirmative. Thus: Starting with a solution of

(A) defined only on [0, 1], we can extend, using the functional equations, to a function defined on $[\alpha x_0, \alpha^2 x_0]$ and satisfying both (A) and (B).

Let $\alpha x_0 \leqslant x < 0$; we want to show that

$$g\big(g(\alpha^{-2}x)\big) = \alpha^{-1}g(x)$$

The calculation is as follows:

$$\begin{aligned}
g\big(g(\alpha^{-2}x)\big) &= g\big(\alpha^2 g\big(\alpha^{-2}g(\alpha^{-3}x)\big)\big) && (1)\\
&= \alpha g\Big(g\big(\alpha^{-2}\alpha^2 g\big(\alpha^{-2}g(\alpha^{-3}x)\big)\big)\Big) && (2)\\
&= \alpha g\Big(g\big(g\big(\alpha^{-2}g(\alpha^{-3}x)\big)\big)\Big)\\
&= \alpha g\Big(\alpha^{-1}\big(\alpha g\big(g\big(\alpha^{-2}g(\alpha^{-3}x)\big)\big)\big)\Big)\\
&= \alpha g\Big(\alpha^{-1}\big(g\big(g(\alpha^{-3}x)\big)\big)\Big) && (3)\\
&= \alpha g\Big(\alpha^{-2}\big(\alpha g\big(g\big(\alpha^{-2}(\alpha^{-1}x)\big)\big)\big)\Big)\\
&= \alpha g\big(\alpha^{-2}g(\alpha^{-1}x)\big) && (4)\\
&= \alpha^{-1}g(x) && (5)
\end{aligned}$$

Only the numbered steps are substantive; we now have to justify each of them. Steps (1) and (5) are just application of the definition of the extended g for negative values of its argument; all that has to be checked is that the argument is in $[\alpha x_0, 0)$, and this is immediate in each case. The other three steps each use

$$\alpha g\big(g(\alpha^{-2}z)\big) = g(z) \qquad \text{(A1)}$$

and it has to be checked in each case that the corresponding z is in $[0, \alpha^2 x_0]$. For step (4), $z = \alpha^{-1}x$, and it is immediate that this is in the appropriate interval. For step (3), $z = g(\alpha^{-3}x)$. Since $\alpha x_0 \leqslant x < 0$,

$$0 < \alpha^{-3}x \leqslant \alpha^{-2}x_0 < x_0$$

so

$$g(\alpha^{-3}x) \in \big[\, g(\alpha^{-2}x_0),\, g(0)\big] \subset [0, 1] \subset [0, \alpha^2 x_0]$$

[To see that $1 < \alpha^2 x_0$, recall that

$$g(\alpha^{-2}) = \alpha^{-2} > 0 = g(x_0)$$

and thus $\alpha^{-2} < x_0$.]

For step (2),

$$z = g\big(\alpha^{-2}g(\alpha^{-3}x)\big)$$

From

$$g\big(g(\alpha^{-2}x_0)\big) = \alpha^{-1}g(x_0) = 0$$

it follows that

$$g(\alpha^{-2}x_0) = x_0$$

We have already verified that

$$g(\alpha^{-3}x) \in \big[g(\alpha^{-2}x_0), 1\big] = \big[x_0, 1\big]$$

Hence

$$g\big(\alpha^{-2}\big(g(\alpha^{-3}x)\big)\big) \in \big[g(\alpha^{-2}), g(\alpha^{-2}x_0)\big] = \big[\alpha^{-2}, x_0\big]$$

as desired.

The preceding calculation was used by Nauenberg in his proof that an analytic solution of (A) also satisfies (B).

We next want to verify that

$$g\big(\alpha^{-2}g(\alpha^{-1}x)\big) = \alpha^{-2}g(x) \qquad \text{for} \quad 0 \leqslant x \leqslant \alpha^2 x_0$$

The calculation is as follows:

$$\begin{aligned} g\big(\alpha^{-2}g(\alpha^{-1}x)\big) &= g\big(\alpha^{-2}\big(\alpha^2 g\big(\alpha^{-2}g(\alpha^{-2}x)\big)\big)\big) && (1)\\ &= \alpha^{-1}\alpha g\big(g\big(\alpha^{-2}\big(g(\alpha^{-1}x)\big)\big)\big) \\ &= \alpha^{-1}g\big(g(\alpha^{-2}x)\big) && (2)\\ &= \alpha^{-1}\alpha g\big(g(\alpha^{-2}x)\big) \\ &= \alpha^{-2}g(x) && (3) \end{aligned}$$

Step (1) is just the definition of $g(\alpha^{-1}x)$; step (3) is just (A1) applied at x. Step (2) is (A1) applied at $g(\alpha^{-2}x)$; we have to verify that this argument is in $[0, \alpha^2 x_0]$. But $0 \leqslant x \leqslant \alpha^2 x_0$ so $0 \leqslant \alpha^{-2}x \leqslant x_0$ so $g(\alpha^{-2}x) \in [g(x_0), g(0)] = [0, 1]$, as desired.

One thing the preceding analysis does *not* show is that, if g is analytic on a complex neighborhood of $[0, 1]$ (and hence defined for small negative x's), then our extension coincides with its analytic continuation. For this, we need the analytic part of Nauenberg's argument, which can be summarized as follows: Define

$$\tilde{g}(x) \equiv \alpha g\big(\alpha^{-2}g(\alpha^{-1}x)\big)$$

wherever the right-hand side is defined (and in particular on $[\alpha x_0, 0]$, where we used this function to extend g). The argument given above shows that

$$\tilde{g}(x) = \alpha g\big(\alpha^{-2}\tilde{g}(\alpha^{-1}x)\big) \qquad (5.1)$$

for small negative x and hence (by analytic continuation) on a complex neighborhood of zero. The problem is to show that $\tilde{g}$ and g agree near zero. This is done by showing that their Taylor series at zero coincide. There are two main steps:

1. Suppose $g(x) = 1 - ax^p + O(x^{p+1})$ with $a \neq 0$. Then $\tilde{g}(x)$ has the same form, with the same constant a. This follows from the definition of $\tilde{g}$ together with the formula

$$g'(\alpha^{-2}) = \alpha^p$$

which is easily deduced from (A).

2. For $j > p$, by differentiating (A1) and (5.1) repeatedly, putting $x = 0$, and rearranging, one gets formulas for $g^{(j)}(0)$ and $\tilde{g}^{(j)}(0)$ in terms of lower-order derivatives and derivatives of g at 1. From these expressions, it follows by induction on j that $\tilde{g}^{(j)}(0) = g^{(j)}(0)$ for $j = p + 1, p + 2, \ldots$. It also follows, incidentally, that $g^{(j)}(0) = 0$ unless j is a multiple of p, i.e., that g is an analytic function of x^p.

ACKNOWLEDGMENTS

This work was begun during the summer of 1982 at the University of California, Berkeley, with support from the National Science Foundation through Grant No. MCS8107068, and continued at the Institute for Mathematics and its Applications at the University of Minnesota in the fall of 1982. I am grateful to Professors H. Weinberger and G. Sell for their hospitality at the IMA. It is a pleasure to acknowledge also helpful discussions with P. Collet, J.-P. Eckmann, H. Epstein, and R. de la Llave.

REFERENCES

1. P. Collet and J.-P. Eckmann, *Iterated Maps on the Interval as Dynamical Systems* (Birkhäuser, Boston, 1980).
2. M. Feigenbaum, L. Kadanoff, and S. Shenker, Quasi-periodicity in dissipative systems: a renormalization group analysis, *Physica D* **5**:370–386 (1982).
3. R. S. MacKay, *Renormalization in Area Preserving Maps*, Ph.D. dissertation, Department of Astrophysical Sciences, Princeton University, October 1982.
4. M. Nauenberg, On fixed points for circle maps, *Phys. Lett. A B* **92**(7):319–320 (1982).
5. S. Ostlund, D. Rand, J. Sethna, and E. Siggia, Universal properties of the transition from quasi-periodicity to chaos in dissipative systems, NSF-ITP preprint 82-80 (July 1982).

Physica 124A (1984) 465-470
North-Holland, Amsterdam

COMPUTER-ASSISTED PROOFS IN ANALYSIS

Oscar E. LANFORD III

IHES, 91440 Bures-sur-Yvette, France. Work supported in part by NSF Grant MCS81-07086.

Computers have a number of uses in mathematics and mathematical physics. Two which are relatively familiar are heuristic numerical exploration and determining properties of discrete objects (e.g., testing integers for primality). I will describe here an example of a less familiar use of computers: the strict verification of estimates on continuously variable quantities (real numbers) for use in the proof of qualitative statements in analysis. This report is a case study rather than a general survey; it is based on my experience working on a concrete problem, the validity of Feigenbaum's renormalization group analysis of the accumulation of period-doubling bifurcations.[2] I am confident that the strict verification of estimates by computer will have other interesting applications, but it is too early to tell how much the particular methods described here will be useful in other situations.

An argument using a computer to prove a mathematical assertion can be thought of as divided into two stages. It is first necessary to derive a sufficient condition for the validity of the assertion which can be verified by a finite computation; then to carry out the verification. The first, analytic, stage is standard mathematics, although computer exploration is likely to be used to help choose a sufficient condition which is actually true. The second, computational, stage is in principle completely mechanical, but in practice considerable thought has to go into structuring the computation so as to make it as comprehensible as possible and hence to minimize the likelihood of error.

The concrete problem to be discussed, motivated by considerations in the analysis of infinite sequences of period-doubling bifurcations,[1] is as follows: We consider the operator

$$\mathcal{T}f(x) = \frac{1}{f^2(0)} \quad f \cdot f\,(f^2(0)\,x)$$

acting on an appropriate domain in the space of mappings f of [-1,1] into itself which are even, decreasing on [0,1] (and hence have a maximum at 0), and

Presented at the VIIth INTERNATIONAL CONGRESS ON MATHEMATICAL PHYSICS, Boulder, Colorado, 1983.

which map 0 to 1. We want to show that:

1. The operator $\mathfrak{T}$ has a fixed point g, analytic on a neighborhood of [-1,1], given approximately by

$$g(x) \simeq 1 - 1.5x^2 + .1\ x^4$$

2. The spectrum of the linearization $D\mathfrak{T}(g)$ of the operator $\mathfrak{T}$ at the fixed point g is strictly inside the unit disk except for a single simple positive eigenvalue larger than one.

To prove the existence of the fixed point we use a simplified version of Newton's method. Newton's method for solving

$$\mathfrak{T}(g)-g = 0$$

gives the iteration

$$g_{n+1} = g_n - (D\mathfrak{T}(g_n) - \mathbb{1})^{-1}\ (\mathfrak{T}(g_n) - g_n)$$

We use instead an iteration of the form

$$g_{n+1} = g_n - (\Gamma - \mathbb{1})^{-1}\ (\mathfrak{T}(g_n) - g_n) \equiv \Phi(g_n) \qquad (1)$$

where Γ is an operator (to be chosen) approximating $D\mathfrak{T}(f)$ reasonably well for f near g. It is easy to derive a sufficient condition for the convergence of the sequence (g_n) obtained from (1) with a given g_0 as follows: In order that Φ be contractive on a ball of radius ρ about g_0, it suffices that

$$\|D\Phi(f)\| = \|(\Gamma-\mathbb{1})^{-1}(D\mathfrak{T}(f)-\Gamma)\| \leq \kappa < 1 \quad \text{for } \|f-g_0\| \leq \rho \qquad (2)$$

and in order that Φ map this ball into itself it suffices that

$$\|(\Gamma - \mathbb{1})^{-1}\ (\mathfrak{T}(g_0) - g_0)\| \leq \rho(1 - \kappa) \qquad (3)$$

Thus, to prove the existence of a fixed point g, all we have to do is to find g_0, Γ, ρ, κ so that (2) and (3) hold. If, furthermore, (3) is sharpened to:

$$\|(\Gamma - e^{i\theta}\,\mathbb{1})^{-1}\ (D\mathfrak{T}(f) - \Gamma)\| \leq \kappa \qquad (4)$$

for all f with $\|f - g_0\| \leq \rho$ and all real θ, and if Γ has spectrum inside the unit disk except for a single simple expanding eigenvalue, then the same will

be true for $D\mathfrak{T}(g)$. We will from now on discuss only the proof of (2) and (3); (4) is proved in much the same way.

Next we have to choose a Banach space of functions f in which the inequalities (2) and (3) are to be proved. We will work with functions f analytic in

$$\{x: |x^2 - 1| < 2.5\}$$

(Many other choices of domain of analyticity would also work.) To define the norm we will use, we first write the general mapping f as

$$f(x) = 1 + x^2 h(x^2).$$

This form builds in the assumed evenness of f and the constraint $f(0) = 1$; working with h instead of f is simply a convenient change of coordinates in the space of mappings. We next expand

$$h(z) = \sum_{n=0}^{\infty} h_n \left(\frac{z-1}{2.5}\right)^n ;$$

define a norm by

$$\|h\| = \sum_{n=0}^{\infty} |h_n| ;$$

and work in the space of h's for which this norm is finite. (Finiteness of this norm says a little more than that h is analytic on the interior of $\overline{\Omega} = \{z: |z-1| \leq 2.5\}$ and continues on the boundary and a little less than that it is analytic on a neighborhood of $\overline{\Omega}$.) The reason for choosing this norm, instead of the more obvious supremum norm, is that it is well adapted to making accurate estimates of norms of linear operators: If T is a linear operator on the space of h's equipped with the above norm, then

$$\|T\| = \sup_{n \geq 0} \|Te_n\| ,$$

where the e_n are the "natural basis vectors" $\left(\frac{z-1}{2.5}\right)^n$.

Now comes a piece of good fortune: It turns out that we can use a very simple approximate derivative Γ, namely:

$$\Gamma e_0 = 4.669\, e_0$$

$$\Gamma e_n = 0 \quad \text{for } n > 0$$

We take an explicit approximate fixed point g_0 obtained as the result of solving he fixed-point problem numerically with good accuracy and we choose an explicit $\rho(\simeq .01)$ and $\kappa(\simeq .9)$. Once these choices are made the inequalities (2) and (3) which will imply the existence of the fixed point are completely explicit.

Before describing how to organize the verification of these inequalities, we need to compute the operator $D\mathfrak{T}(f)$. Heuristically, this is done by replacing f by $f + \delta f$ in the formula

$$\mathfrak{T}f(x) = \frac{1}{f(1)} \; f \cdot f\,(f(1) \cdot x)$$

and extracting the terms linear in δf. This produces a sum of four terms, one for each place f appears in the expression for $\mathfrak{T}f$. These expressions must then be rewritten in terms of h and δh (related to f, δf by

$$f(x) = 1 + x^2 h(x^2); \quad \delta f(x) = x^2 \delta h(x^2)\ )$$

To show what the final expressions look like, we write one of the four terms:

$$(D\mathfrak{T}^{(2)}(h)e_j)(z) = \frac{\lambda}{2.5}\;(1 + \lambda^2\, zh(z))^2\,(2 + \lambda^2\, zh(\lambda^2 z))$$
$$\times\; h(\lambda^2 z)\left[\frac{\lambda^2\, z\, h(\lambda^2 z)}{2.5}\right]^{j-1},$$

$$\lambda \equiv h(1) + 1$$

(This holds for $j \geq 1$; for $j = 0$, the left-hand side vanishes.) To verify (2), we estimate the location in function space of the right-hand side of (5) in tersm of information about the location of h (and similarly for the other three terms in the expression for $D\mathfrak{T}(h)$).

These estimates are completely straightforward in principle; in practice, they quickly become unreasonably complicated if not structured carefully. To structure the computation we use the notion of rectangle in function space. By this we mean a set $\mathbf{R}$, in the Banach space of h's, of the form

$$\{h = \sum_{j=0}^{\infty} h_j \left(\frac{z-1}{2.5}\right)^j : \ell_0 \leq h_0 \leq u_0, \ldots, \quad \ell_{n-1} \leq h_{n-1} \leq u_{n-1}, \quad \sum_{j=n}^{\infty} |h_j| \leq \varepsilon\},$$

determined by 2n+1 numbers

$$\ell_0 \leq u_0, \quad \ell_1 \leq u_1, \ldots, \quad \ell_{n-1} \leq u_{n-1}, \quad \varepsilon \geq 0$$

It is straightforward and reasonably simple to work out how to do elementary operations on these rectangles. For example: Given two rectangles $\mathcal{R}_1$ and $\mathcal{R}_2$, one constructs a rectangle $\mathcal{R}_3$ (" = $\mathcal{R}_1 \times \mathcal{R}_2$") such that, whenever $h_1 \in \mathcal{R}_1$ and $h_2 \in \mathcal{R}_2$, $h_1 \cdot h_2 \in \mathcal{R}_3$. Similarly for such other operations as addition, scalar multiplication, and composition.

In terms of these elementary operations, it is also not too difficult to give a prescription for finding, given a rectangle $\mathcal{R}_0$ and an integer j, another rectangle guaranteed to contain $D\mathcal{T}(h)e_j$ for any $h \in \mathcal{R}_0$. From this rectangle, it is easy to obtain a bound on

$$\|(\Gamma - \mathbb{1})^{-1}(D\mathcal{T}(h) - \Gamma)\, e_j\|$$

which holds for all $h \in \mathcal{R}_0$. Now recall that, to prove (2), we have only to show that

$$\|(\Gamma - \mathbb{1})^{-1}(D\mathcal{T}(h)-\Gamma)\, e_j\| \leq \kappa \quad \text{for all } j.$$

The above permits us to check this inequality for any given j. There are, however, infinitely many j's to be considered, so we are not yet reduced to a finite computation. Fortunately, only finitely many j's require estimates of the above detailed kind. To see why, consider the sample term written in (5) above, and note that it has the form

$$(D\mathcal{T}^{(2)}(h)e_j)(z) = u(z) \cdot (v(z))^{j-1}$$

with u(z), v(z) independent of j. Since

$$\|u \cdot v^{j-1}\| \leq \|u\|\, \|v\|^{j-1},$$

we can deal with all large j's simply by establishing a bound of the form

$$\|v\| \leq \sigma < 1 \qquad \text{for } \|h - h_0\| \leq \rho$$

Similar analyses can be done for the other three terms in the expression for $D\mathcal{T}(h)$.

There remains one more complication. Although it is possible to program a computer to do exact arithmetic (on rational numbers), this is usually impractical and arithmetic is instead normally done to some fixed finite precision. It is therefore necessary to control the effect of round-off error

if one is to make a strict verification of (2) and (3). There is a standard technique for the automatic estimation of round-off error, known as interval arithmetic. The idea of interval arithmetic is to represent numbers "with error bars" by specifying, instead of a single finite-precision approximation for a given number, the exact end-points of an interval guaranteed to contain the number in question. One can then construct computer procedures for "doing arithmetic operations on intervals." For example: Given two intervals $[\ell_1,u_1]$ and $[\ell_2, u_2]$, with ℓ_1, u_1, ℓ_2, u_2 all d-digit numbers, one finds another interval $[\ell_3, u_3]$, with ℓ_3, u_3 again d-digit numbers, guaranteed to contain all products $x_1 \cdot x_2$ with $x_1 \in [\ell_1, u_1]$ and $x_2 \in [\ell_2, u_2]$. (To be entirely explicit: The best possible lower bound ℓ_3 can be obtained by forming the four exact products $\ell_1 \cdot \ell_2$, $\ell_1 \cdot u_2$, $u_1 \cdot \ell_2$, $u_1 \cdot u_2$, each of which has no more than 2d digits; picking the smallest of them; and rounding down to the next smaller n-digit number. It is not really necessary, however, to find the best possible ℓ_3, and usually some shortcuts are taken, giving an ℓ_3 which is a correct lower bound but not necessarily the best possible one.)

We have now described all the elements needed to construct a computer program for verifying estimates (2) and (3). As indicated, this program can be organized into a number of reasonably simple pieces. At the lowest level is a set of computer procedures (subroutines) for doing the fundamental arithmetic operations on intervals. Built on these procedures is a higher-level set of procedures for doing elementary operations on rectangles in function space. The program for verifying (2) and (3) is constructed essentially by translating formulas like (5) into a sequence of invocations of these latter procedures.

REFERENCES

1) P Collet and J P Eckmann, Iterated Maps on the Interval as Dynamical Systems (Birkhäuser, 1980).

2) O E Lanford, A computer-assisted proof of the Feigenbaum conjectures. Bull. A.M.S. (New Series) 6 (1982) 427-434.

Commun. Math. Phys. 96, 521–538 (1984)

Communications in Mathematical Physics

A Shorter Proof of the Existence of the Feigenbaum Fixed Point

Oscar E. Lanford III

IHES, F-91440 Bures-sur-Yvette, France

Abstract. We use the Leray-Schauder Fixed Point Theorem to prove the existence of an analytic fixed point for the period doubling accumulation renormalization operator. Our argument does not, however, show that the linearization of the renormalization operator at this fixed point is hyperbolic.

1. Introduction

Two independent proofs (Campanino et al. [1, 2], Lanford [4]) have been given for the existence of an even analytic solution to the Feigenbaum-Cvitanović functional equation [3]

$$g(x) = -\frac{1}{\lambda} g(g(-\lambda x)), \quad g(0) = 1 \tag{1.1}$$

with

$$g''(0) < 0; \lambda \equiv -g(1) > 0 .$$

Both of these proofs rely on extensive computations. In this paper, we give yet another proof, based on the Leray-Schauder Fixed Point Theorem, which, if still fundamentally computational in nature, requires a substantially smaller amount of computation. It should be noted that the argument given here, like that of Campanino et al. and unlike the author's computer assisted proof, does *not* establish the spectral properties of the linearization of the renormalization operator at the fixed point g which are essential for the application of g to the analysis of period-doubling accumulation.

We will work in a space of even mappings f of $[-1, 1]$ to itself, satisfying the normalization condition

$$f(0) = 1 , \tag{1.2}$$

expressed as functions of x^2. Since we are working with x^2 as the independent variable, the renormalization operator has the form

$$\mathcal{T} f(x) = -\frac{1}{\lambda} f([f(\lambda^2 x)]^2), \quad \lambda \equiv -f(1) . \tag{1.3}$$

We will frequently write $\bar{f}$ for $\mathscr{T}f$. If we define μ by

$$f'(0) = -(1+\lambda+\mu),$$

and write

$$f_0(x) \equiv 1-(1+\lambda)x-\mu x(1-x),$$
$$f_1(x) \equiv f(x)-f_0(x),$$

then

$$f_1(0) = f_1(1) = f_1'(0) = 0,$$

so f_1 is uniquely determined by its third derivative. We will denote $f_1'''(x) = f'''(x)$ by $h(x)$. The triple (λ, μ, h) will serve as a set of coordinates for the space of mappings in which we work. Ultimately, h will be analytic on a complex neighborhood of $[0, 1]$, but for much of the argument we can take h to be simply a continuous function on $[0, 1]$. We will write $\bar{\lambda}(\lambda, \mu, h)$, $\bar{\mu}(\lambda, \mu, h)$, $\bar{h}(\lambda, \mu, h)$ for the coordinates of $\bar{f} = \mathscr{T}f$.

The correspondence which associates with any continuous function h on $[0, 1]$ the unique function f_1 such that

$$f_1'''(x) = h(x); \quad f_1(0) = f_1(1) = f_1'(0) = 0$$

is linear and can be written as an integral operator

$$f_1(x) = \int_0^1 K(x, y)\, h(y)\, dy \tag{1.4}$$

with kernel $K(x, y)$ which can easily be written explicitly (see Sect. 2). We will also use K to denote the operator, i.e., we will write $f_1 = Kh$ as a shorthand for (1.4). Thus, the f corresponding to the triple (λ, μ, h) can be written as

$$f(x) = 1-(1+\lambda+\mu)x+\mu x^2+Kh(x). \tag{1.5}$$

Because the renormalization operator $\mathscr{T}$ is expansive in one direction at the Feigenbaum fixed point, no small neighborhood of the fixed point can be invariant for $\mathscr{T}$. To get to a situation where we can apply the Leray-Schauder Fixed Point Theorem, we introduce an auxiliary operator with an equivalent fixed-point problem but which does admit small invariant neighborhoods of the fixed point we are looking for. This operator will act only on the h coordinate and will be constructed as follows: We first show that, for any h in an appropriate domain, the pair of equations

$$\bar{\lambda}(\lambda, \mu, h) = \lambda, \quad \bar{\mu}(\lambda, \mu, h) = \mu \tag{1.6}$$

has a unique solution $(\lambda^*, \mu^*) \approx (0.4, 0.1)$. The auxiliary operator is then defined to map h to $h^* \equiv \bar{h}(\lambda^*(h), \mu^*(h), h)$. Fixed points for this auxiliary operator correspond in an obvious way to fixed points for the renormalization operator itself.

To show that the auxiliary operator is well-defined and admits a fixed point, we are going to prove:

Lemma 1.1. *Let h be a continuous real-valued function on $[0,1]$ satisfying*

$$0 \leqq h(x) \leqq 0.32(1-0.36x) \quad \text{for} \quad 0 \leqq x \leqq 1. \tag{1.7}$$

Then there is a unique pair (λ^*, μ^*) with

$$0.396 \leqq \lambda^* \leqq 0.4031, \qquad 0.09 \leqq \mu^* \leqq 0.16$$

such that

$$\bar{\lambda}(\lambda^*, \mu^*, h) = \lambda^*, \qquad \bar{\mu}(\lambda^*, \mu^*, h) = \mu^*. \tag{1.8}$$

Furthermore, λ^* and μ^* vary continuously with h.

Lemma 1.2. *If h satisfies (1.7) and if λ^*, μ^* are as in Lemma 1.1 then $h^* \equiv \bar{h}(\lambda^*, \mu^*, h)$ also satisfies (1.7).*

For $\delta > 0$, let D_δ denote the set of complex numbers at distance less than δ from $[0,1]$.

Lemma 1.3. *If δ is sufficiently small, if h is analytic on D_δ, satisfies (1.7), and in addition satisfies*

$$|h(z)| \leqq 0.32(1-0.36|z|) \quad \text{on} \quad D_\delta, \tag{1.9}$$

then h^* is analytic on $D_{3/2\delta}$ and satisfies

$$|h^*(z)| \leqq 0.32(1-0.36|z|) \quad \text{on} \quad D_{3/2\delta}. \tag{1.10}$$

The correspondence $h \mapsto h^*$ is continuous from the space of functions analytic on D_δ satisfying (1.7) and (1.9) to the space of functions analytic on $D_{3/2\delta}$ satisfying (1.7) and (1.10), both spaces equipped with the topology of uniform convergence.

These three lemmas, and the Leray-Schauder Fixed Point Theorem, immediately imply:

Theorem. *There exists a function $g(x)$ defined and analytic on a neighborhood of $[-1,1]$, even, with $g(0)=1$, and satisfying*

$$g(x) = -\frac{1}{\lambda} g(g(-\lambda x)).$$

Furthermore, $\lambda (= -g(1)) \in [0.396, 0.4031]$,

$$-2(1+\lambda+0.16) \leqq g''(0) \leqq -2(1+\lambda+0.09),$$

and, writing $g(x) = G(x^2)$ (which is possible because g is even),

$$0 \leqq G'''(x) \leqq 0.32(1-0.36x) \quad \text{for} \quad 0 \leqq x \leqq 1.$$

Lemmas 1.1, 1.2, and 1.3 will be proved in Sects. 4–6, respectively. In Sect. 2, we establish some properties of the kernel $K(x, y)$ of (1.4), and in Sect. 3 we prove a number of estimates used repeatedly in later sections.

It will be a universal notational convention, for the remainder of this paper, that the symbols $f, \lambda, \mu, h, f_0, f_1$ are always assumed to be related as above. We also adopt

as standing assumptions that λ and μ denote real numbers satisfying

$$0.396 \leqq \lambda \leqq 0.4031; \quad 0.09 \leqq \mu \leqq 0.16, \tag{1.11}$$

and h a function defined and continuous (at least) on $[0, 1]$, *satisfying*

$$0 \leqq h(x) \leqq 0.32(1 - 0.36x). \tag{1.12}$$

These assumptions will be used very frequently, generally without explicit reference.

In the course of the proof, we need to make a considerable number of concrete numerical estimates. To take an example at random: At one point, we use the fact that, if $\lambda = 0.4031$ and $\mu = 0.16$, then $(1+\lambda)\lambda^2 + \mu\lambda^2(1-\lambda^2) < 0.2498$. Estimates like this were verified with the aid of a Hewlett-Packard HP-15C calculator. This calculator stores and manipulates numbers in a decimal floating point format with ten digit fraction and two digit exponent; we assumed only that it would perform correctly the operations of addition, subtraction, and multiplication on pairs of operands for which the result can be represented exactly in this format. In practice, this meant that intermediate results were rounded – up or down, depending on the sense of the inequality to be proved – to five digits before being multiplied together. Also, the results of divisions were verified by multiplying back after rounding. To take the above example:

$$\lambda^2 = 0.16248961 < 0.16249,$$

$$\lambda^2(1-\lambda^2) < 0.16249(1 - 0.16249) = 0.1360869999 < 0.13609,$$

$$\mu\lambda^2(1-\lambda^2) < 0.16 \times 0.13609 = 0.0217744 < 0.02178,$$

$$(1+\lambda)\lambda^2 < 1.4031 \times 0.16249 = 0.227989719 < 0.22799,$$

$$(1+\lambda)\lambda^2 + \mu\lambda^2(1-\lambda^2) < 0.22799 + 0.02178 = 0.24977 < 0.2498.$$

This approach to proving such numerical inequalities is no doubt more cautions than is really justified [1], but it does have the merit of relying as little as possible on the correctness of the calculator.

From a broader point of view, the question of what constitutes a satisfactory proof of an explicit numerical estimate like the one above provides an illuminating caricature of the issues involved in "computer assisted proofs" in general. It hardly seems reasonable to insist that the arithmetic operations be carried out by hand, but relying on results of individual arithmetic operations performed by an electronic calculator does not differ in a fundamental way from relying on results of more complicated sequences of operations performed by a larger computer.

---

1 Especially since Hewlett-Packard, in a departure from the standard practice of calculator manufacturers, has published an explicit and unambiguous statement on the accuracy the HP-15C is supposed to attain. For the four basic arithmetic operations, in the absence of underflow and overflow, it is asserted that the result returned differs from the exact result by no more than one-half unit in the last (i.e., tenth) place. This statement is labelled as a design objective which the designers believe that they can prove they have attained rather than as a guaranteed specification; it appears in the Appendix "Accuracy of Numerical Calculations", pp. 172–211 of *The HP-15C Advanced Functions Handbook*, part #00015-90011. I am indebted to W. Kahan for pointing this reference out to me

2. The Kernel K

Let

$$K(x,y) \equiv -x^2(1-y)^2/2 + \theta(x-y)(x-y)^2/2, \tag{2.1}$$

where $\theta(z)=0$ for $z<0$ and $\theta(z)=1$ for $z>0$. We will also write

$$K'(x,y) \equiv \frac{\partial K}{\partial x}(x,y) = -x(1-y^2) + \theta(x-y)(x-y), \tag{2.2}$$

$$K''(x,y) \equiv \frac{\partial K'}{\partial x}(x,y) = -(1-y^2) + \theta(x-y). \tag{2.3}$$

If h is a continuous function on $[0,1]$, we define

$$(Kh)(x) = \int_0^1 K(x,y)\, h(y)\, dy.$$

Proposition 2.1. *Kh is the unique three-times continuously differentiable function on $[0,1]$ such that*

$$Kh(0) = Kh(1) = Kh'(0) = 0, \qquad (Kh)'''(x) = h(x).$$

Furthermore:

$$(Kh)'(x) = \int_0^1 K'(x,y)\, h(y)\, dy, \tag{2.4}$$

$$(Kh)''(x) = \int_0^1 K''(x,y)\, h(y)\, dy. \tag{2.5}$$

Proof. It follows easily from standard results about differentiating under the integral sign that Kh is twice continuously differentiable and that (2.4) and (2.5) hold. From (2.5) and the formula for $K''(x,y)$,

$$(Kh)''(x) = \int_0^x h(y)\, dy - \int_0^1 (1-y)^2\, h(y)\, dy,$$

from which it follows that Kh is three times differentiable and that $(Kh)'''(x) = h(x)$. It is immediate from the definitions that $Kh(0) = Kh(1) = 0$, and from (2.4) that $(Kh)'(0) = 0$.

In the following proposition x and y denote general points of $[0,1]$, i.e., an assertion containing an unquantified x (or y) should be understood as holding for all x (or y) in $[0,1]$. Also, we write $K'_+(x,y)$ [respectively $K''_+(x,y)$, $K''_-(x,y)$] for the positive part of $K'(x,y)$ [respectively, the positive, negative parts of $K''(x,y)$], i.e., the larger of 0 and $K'(x,y)$ [respectively, $K''(x,y)$, $-K''(x,y)$].

Proposition 2.2. 1. $K(x, y) \leqq 0$.

2. $\int_0^1 K(x, y)\, dy = -x^2(1-x)/6$.

3. $\int_0^1 K(x, y)(1-r\cdot y)\, dy = -x^2(1-x)[(1-r/4)-(r/4)x]/6$ *for all real* r.

4. $K'(x, y) \leqq 0$ *for* $x \leqq 1/2$.

5. $K'(1, y) = y(1-y) \geqq 0$.

6. $\int_0^1 K'(1/2, y) = -1/24$.

7. $\int_0^1 K'(1, y)(1-r\cdot y)\, dy = 1/6 - r/12$ *for all real* r.

8. $\int_0^1 K'_+(x, y)\, dy \leqq 1/6$.

9. $\int_0^1 K''_+(x, y)\, dy = x^2 - x^3/3$.

10. $\int_0^1 K''_+(x, y)(1-r\cdot y)\, dy \leqq 2/3(1-5r/8)\cdot x$ *for* $0 \leqq r \leqq 4/7$.

11. $\int_0^1 K''_-(x, y)(1-r\cdot y)\, dy \leqq 1/3 - r/12$ *for* $0 \leqq r \leqq 1$.

Proof. 1. The assertion follows at once from the formula (2.1) for $K(x, y)$ if $x < y$. If $x > y$

$$\begin{aligned} K(x, y) &= -x^2(1-y)^2/2 + (x-y)^2/2 \\ &= 1/2[(x-y)+x(1-y)]\,[(x-y)-x(1-y)] \\ &= -1/2[(x-y)+x(1-y)]\, y(1-x) \leqq 0\,. \end{aligned}$$

2. This is a special case of 3..

3. By Proposition 2.1 (with $h = 1 - rx$), the left-hand side is the unique function vanishing to second order at 0 and to first order at 1 with third derivative equal to $1 - rx$; it is easy to see that the right-hand side has these properties.

4. It is immediate from the formula that $K'(x, y) \leqq 0$ for $x < y$. For $x > y$,

$$K'(x, y) = -x(1-y)^2 + (x-y) = [2x - 1 - xy]\, y\,,$$

which is manifestly negative if $x \leqq 1/2$.

5. Insert $x = 1$ in the preceding formula for $K'(x, y)$, which is valid when $x > y$.

6. Differentiate 2. and put $x = 1/2$.

7. Evaluate the integral explicitly, using 5. (or differentiate 3. and put $x = 1$).

8. By 4., the left-hand side vanishes for $x \leqq 1/2$, so we have only to consider $x > 1/2$. Also, by the proof of 4., $K'_+(x, y) = 0$ if either $y > x$ or $y > (2x-1)/x$. Since $(2x-1)/x \leqq x$, we can ignore the first condition: $K'_+(x, y) = [2x - 1 - xy]\, y$ for $y < (2x-1)/x$ and 0 otherwise. Hence:

$$\begin{aligned} \int_0^1 K'_+(x, y)\, dy &= \int_0^{(2x-1)/x} [(2x-1) - xy]\, y\, dy \\ &= (2x-1)\,[(2x-1)/x]^2 \int_0^1 (1-z)\, z\, dz = (2-1/x)^2\,(2x-1)/6 \leqq 1/6\,. \end{aligned}$$

9. From the definition of $K''(x, y)$,

$$\begin{aligned} K''_+(x, y) &= 0 && \text{for} \quad x < y \\ &= 1-(1-y)^2 = 2y-y^2 && \text{for} \quad x > y . \end{aligned}$$

Hence

$$\int_0^1 K''_+(x, y)\, dy = \int_0^x (2y-y^2)\, dy = x^2 - x^3/3 .$$

10. From the expression for $K''_+(x, y)$ obtained in the proof of 9.,

$$\int_0^1 K''_+(x, y)(1-ry)\, dy = x[x-(1+2r)x^2/3+rx^3/4] .$$

What we have to show, therefore, is that

$$[x-(1+2r)x^2/3+rx^3/4] \leqq 2/3(1-5r/8) \quad \text{for} \quad 0 \leqq r \leqq 4/7 .$$

Since both sides of this inequality are affine in r, it suffices to prove it for $r=0$ and $r=4/7$. For $r=0$, it reduces to $x-x^2/3 \leqq 2/3$, which is immediate and for $r=4/7$ to $x-5x^2/7+x^3/7 \leqq 3/7$, or $0 \geqq x^3-5x^2+7x-3=(x-1)^2(x-3)$, which is also immediate.

11. From the formula for $K''(x, y)$,

$$\begin{aligned} K''_-(x, y) &= (1-y)^2, && x < y \\ &= 0, && x > y . \end{aligned}$$

Thus

$$\begin{aligned} \int_0^1 K''_-(x, y)(1-ry)\, dy &= \int_x^1 (1-y)^2(1-ry)\, dy \\ &= (1-r)(1-x)^3/3 + r(1-x)^4/4 \leqq 1/3 - r/12 . \end{aligned}$$

3. Some Estimates

We collect in the following proposition a number of estimates which will be used repeatedly. In this section, as in the preceding one, x denotes a general point of $[0, 1]$.

Proposition 3.1. 1. $f''(x) > 0$.

2. $1.192 < -f'(x) \leqq 1+\lambda+\mu \leqq 1.5631$.
3. $0.7491 < f(\lambda^2) < 0.7692$.
4. $0.5611 < [f(\lambda^2)]^2 < 0.5917$.
5. $f(\lambda^2 x) \leqq 1-0.2308x$.
6. $\lambda^5(1+\lambda+\mu) f(\lambda^2 x) < 0.016638(1-0.2497x)$.
7. $\lambda^5(1+\lambda+\mu)[f(\lambda^2 x)]^2 < 0.016638(1-0.4370x)$.
8. $\lambda^5(1+\lambda+\mu)[f(\lambda^2 x)]^3 < 0.016638(1-0.5776x)$.
9. $[f(\lambda^2 x)]^2 \geqq 1-0.5082x$.
10. $-f'(x) < 1.4165$ *for* $x \geqq 0.5$.
11. $f''(\lambda^2 x) < 0.328$.

Proof. 1. Using Proposition 2.2.11,

$$f''(x) = 2\mu + (Kh)''(x) \geqq 2\mu - 0.32(1/3 - 0.36/12) > 0 .$$

2. By 1. and Proposition 2.2.7,

$$1 + \lambda + \mu = -f'(0) \geqq -f'(x) \geqq -f'(1) = 1 + \lambda - \mu - 0.32(1/6 - 0.36/12) > 1.192 .$$

3.

$$f(\lambda^2) = 1 - (1 + \lambda + \mu)\lambda^2 + \mu\lambda^4 + Kh(\lambda^2) .$$

Since, by Proposition 2.2.1, $Kh(\lambda^2) \leqq 0$, $f(\lambda^2) \leqq 1 - (1 + \lambda)\lambda^2 - \mu\lambda^2(1 - \lambda^2)$. The expression on the right is decreasing in μ and λ separately, so we get an upper bound by inserting $\lambda = 0.396$ and $\mu = 0.09$, which gives

$$f(\lambda^2) < 0.7692 .$$

To get a lower bound, we apply Proposition 2.2.3 to estimate

$$-Kh(\lambda^2) \leqq 1/6\lambda^4(1 - \lambda^2)[0.91 - 0.09\lambda^2] \cdot 0.32 .$$

It is easy to check (e.g., by taking logarithmic derivatives) that the expression on the right is increasing in λ, so we get an upper bound by inserting $\lambda = 0.4031$, so $-Kh(\lambda^2) < 0.00106$.

Thus

$$f(\lambda^2) \geqq 1 - (1 + \lambda)\lambda^2 - \mu\lambda^2(1 - \lambda^2) - 0.00106 > 0.7491 .$$

4. This follows from 3. by taking squares.

5. Since f is convex,

$$f(\lambda^2 x) \leqq (1 - x) f(0) + x f(\lambda^2) = 1 - (1 - f(\lambda^2))x ,$$

and, by 3., $1 - f(\lambda^2) > 1 - 0.7692 = 0.2308$.

6, 7, and 8. As in 5., $f(\lambda^2 x) \leqq 1 - (1 - f(\lambda^2))x$ and

$$1 - f(\lambda^2) = (1 + \lambda)\lambda^2 + \mu\lambda^2(1 - \lambda^2) - Kh(\lambda^2) \geqq (1 + \lambda)\lambda^2 + \mu\lambda^2(1 - \lambda^2) .$$

If we define

$$\varrho \equiv (1 + \lambda)\lambda^2 + \mu\lambda^2(1 - \lambda^2),$$

what we have just shown is that $f(\lambda^2 x) \leqq 1 - \varrho x$.

We now claim that, for $j = 1, 2, 3$,

$$\lambda^5(1 + \lambda + \mu)(1 - \varrho x)^j \tag{3.1}$$

is increasing in λ and μ separately. To show that (3.1) is increasing in λ, we take its logarithmic derivative; what we have to show is that

$$\frac{5}{\lambda} + \frac{1}{1 + \lambda + \mu} \geqq \frac{jx}{1 - \varrho x} \frac{\partial \varrho}{\partial \lambda} .$$

Since the expression on the right is increasing in j and x (and the expression on the left doesn't depend on these quantities), it suffices to consider $j = 3$ and $x = 1$, i.e., to

show that

$$\frac{5}{\lambda}+\frac{1}{1+\lambda+\mu}\geqq\frac{3}{1-\varrho}\frac{\partial\varrho}{\partial\lambda}. \tag{3.2}$$

Now

$$\frac{\partial\varrho}{\partial\lambda}=2\lambda+3\lambda^2+2\mu\lambda-4\mu\lambda^3<2\lambda+3\lambda^2+2\mu\lambda<1.5,$$

and $\varrho<0.2498$ so $1-\varrho>0.75$. Thus

$$\frac{3}{1-\varrho}\frac{\partial\varrho}{\partial\lambda}<\frac{3}{0.75}1.5=6,$$

while

$$\frac{5}{\lambda}>10,$$

so (3.2) is established.

Similarly, showing that (3.1) is increasing in μ reduces to showing

$$\frac{1}{1+\lambda+\mu}\geqq\frac{3}{1-\varrho}\frac{\partial\varrho}{\partial\mu}. \tag{3.3}$$

Since

$$\frac{\partial\varrho}{\partial\mu}=\lambda^2(1-\lambda^2)<0.1361,\quad \frac{3}{1-\varrho}\frac{\partial\varrho}{\partial\mu}\leqq 0.55,$$

while

$$\frac{1}{1+\lambda+\mu}>\frac{1}{1.6}=0.625,$$

which proves (3.3).

Thus, we can get an upper bound for $\lambda^5(1+\lambda+\mu)(1-\varrho)^j$ by inserting $\lambda=0.4031$, $\mu=0.16$. The corresponding value of ϱ is greater than 0.2497, and of $\lambda^5(1+\lambda+\mu)$ is less than 0.016638. Finally,

$$(1-0.2497x)^2\leqq 1-\varrho_2 x,\quad \varrho_2=1-(1-0.2497)^2>0.4370$$
$$(1-0.2497x)^3\leqq 1-\varrho_3 x,\quad \varrho_3=1-(1-0.2497)^3>0.5776.$$

9.

$$\frac{d^2}{dx^2}[f(\lambda^2 x)]^2=2\lambda^4\{[f'(\lambda^2 x)]^2+f(\lambda^2 x)\,f''(\lambda^2 x)\}>0.$$

Hence,

$$[f(\lambda^2 x)]^2\geqq[f(0)]^2+x\frac{d}{dx}[f(\lambda^2 x)]^2|_{x=0}$$
$$=1-2\lambda^2(1+\lambda+\mu)x\geqq 1-0.5082x.$$

10. By 1. and Proposition 2.2.6,

$$-f'(x) \leqq -f'(0.5) = 1 + \lambda - (Kh)'(0.5) \leqq 1.4031 + 0.32/24 < 1.4165 .$$

11. By Proposition 2.2.9,

$$f''(\lambda^2 x) = 2\mu + (Kh)''(\lambda^2 x) \leqq 2\mu + \lambda^4(1 - \lambda^2/3) \cdot 0.32 < 0.328 .$$

4. Solving for λ^* and μ^*

The objective of this section is to prove Lemma 1.1. The first step is to reduce the pair of simultaneous equations

$$\bar{\lambda}(\lambda^*, \mu^*, h) = \lambda^*, \qquad \bar{\mu}(\lambda^*, \mu^*, h) = \mu^* \tag{4.1}$$

to a single equation by solving for μ in terms of λ. To do this, we note that

$$(\mathscr{T}f)'(0) = -2\lambda f'(0) f'(1),$$

and hence

$$(\mathscr{T}f)'(0) = f'(0) \quad \text{if and only if} \quad f'(1) = -1/(2\lambda) .$$

Recalling that

$$f'(0) = -(1 + \lambda + \mu); \qquad (\mathscr{T}f)'(0) = -(1 + \bar{\lambda} + \bar{\mu}),$$

we see that, if $\bar{\lambda} = \lambda$, then $\bar{\mu} = \mu$ if and only if $f'(1) = -1/(2\lambda)$ i.e., if and only if

$$\mu = 1 + \lambda - 1/(2\lambda) - (Kh)'(1) . \tag{4.2}$$

Thus, to solve (4.1), we insert (4.2) into $\bar{\lambda}$ and solve the single equation $\bar{\lambda} = \lambda$; the solution is λ^* and

$$\mu^* = 1 + \lambda^* - 1/(2\lambda^*) - (Kh)'(1) .$$

For γ a real number, we let C_γ denote the intersection of the curve $\mu = 1 + \lambda - 1/(2\lambda) - \gamma$ with the rectangle $\{(\lambda, \mu) : 0.396 \leqq \lambda \leqq 0.4031, 0.09 \leqq \mu \leqq 0.16\}$. What we want to show is that, for $\gamma = (Kh)'(1)$, $\bar{\lambda}(\lambda, \mu, h) - \lambda$ vanishes exactly once on $C\gamma$. We break up the proof of this fact into a sequence of lemmas.

Lemma 4.1.

$$\frac{\partial \bar{\lambda}}{\partial \lambda}(\lambda, \mu, h) > 1; \qquad \frac{\partial \bar{\lambda}}{\partial \mu}(\lambda, \mu, h) > 0 .$$

Since $1 + \lambda - 1/(2\lambda) - \gamma$ is increasing in λ, this shows that $\bar{\lambda}(\lambda, \mu, h) - \lambda$ increases with λ along C_γ, and hence vanishes at most once along this arc.

Lemma 4.2.

$$0 \leqq (Kh)'(1) \leqq 0.0438 .$$

Lemma 4.3. *If* $0 \leqq \gamma \leqq 0.0438$, C_γ *intersects the boundary of the rectangle* $\{(\lambda, \mu) : 0.396 \leqq \lambda \leqq 0.4031, 0.09 \leqq \mu \leqq 0.16\}$ *twice; once (near the lower-left corner) at a point satisfying either* $\lambda = 0.396$, $\mu < 0.1334$ *or* $\lambda < 0.3962$, $\mu = 0.09$, *and once*

(near the upper-right corner) at a point satisfying either $\lambda = 0.4031$, $\mu > 0.1189$ *or* $\lambda > 0.4024$, $\mu = 0.16$.

Lemma 4.4.

$$\bar{\lambda}(0.396, 0.1334, h) < 0.396,$$

$$\bar{\lambda}(0.3962, 0.09, h) < 0.3962,$$

$$\bar{\lambda}(0.4031, 0.1189, h) > 0.4031,$$

$$\bar{\lambda}(0.4024, 0.16, h) > 0.4024.$$

Lemma 1.1 follows almost immediately from these four lemmas. By Lemmas 4.1 and 4.4, we see that $\bar{\lambda}(\lambda, \mu, h) < \lambda$ if either $\lambda = 0.396$, $\mu \leqq 0.1334$ or $\lambda \leqq 0.3862$, $\mu = 0.09$, and that $\bar{\lambda}(\lambda, \mu, h) > \lambda$ if either $\lambda = 0.4031$, $\mu > 0.1189$ or $\lambda > 0.4024$, $\mu = 0.16$. Hence, by Lemmas 4.2 and 4.3, $\bar{\lambda}(\lambda, \mu, h) - \lambda$ changes sign along $C_{(Kh)'(1)}$, and so vanishes at some point of this arc. We have already noted that Lemma 4.1 implies that it cannot vanish more than once. Furthermore, since the inequalities of Lemma 4.1 are strict, it follows from the Implicit Function Theorem that the unique solution λ^*, μ^* of $\bar{\lambda}(\lambda, \mu, h) = \lambda$ on $C_{(Kh)'(1)}$ varies continuously with h.

Proof of Lemma 4.1. We have

$$\bar{\lambda} = \frac{1}{\lambda} f([f(\lambda^2)]^2),$$

so

$$\frac{\partial \bar{\lambda}}{\partial \lambda} = -\frac{1}{\lambda^2} f([f(\lambda^2)]^2) + \frac{1}{\lambda} \frac{\partial f}{\partial \lambda}([f(\lambda^2)]^2) + \frac{2}{\lambda} f'([f(\lambda^2)]^2) f(\lambda^2) \left\{ \frac{\partial f}{\partial \lambda}(\lambda^2) + 2\lambda f'(\lambda^2) \right\}.$$

Since

$$f(x) = 1 - (1+\lambda)x - \mu x(1-x) + Kh(x), \quad \frac{\partial f}{\partial \lambda}(x) = -x, \tag{4.3}$$

so we get

$$\frac{\partial \bar{\lambda}}{\partial \lambda} = -\frac{\bar{\lambda}}{\lambda} - \frac{\lambda [f(\lambda^2)]^2}{\lambda} + 4f'([f(\lambda^2)]^2) f'(\lambda^2) f(\lambda^2) + 2\lambda(-f'([f(\lambda^2)]^2)) f(\lambda^2). \tag{4.4}$$

We first estimate

$$\frac{\bar{\lambda}}{\lambda} = \frac{1}{\lambda^2} f(a) = \frac{1}{\lambda^2}[1 - (1+\lambda)a - \mu a(1-a) + Kh(a)],$$

where we have written a for $[f(\lambda^2)]^2$. Since $Kh(a) \leqq 0$ (Proposition 2.2.1),

$$\frac{\bar{\lambda}}{\lambda} \leqq \frac{1}{\lambda^2}[1 - (1+\lambda)a - \mu a(1-a)].$$

The expression on the right is manifestly decreasing in λ and μ (for fixed a); it is easy to see that it is also decreasing in a. We can thus get an upper bound by replacing λ by 0.396, μ by 0.09, and a by 0.5611 (see Proposition 3.1.4); this gives

$$\frac{\bar{\lambda}}{\lambda} < 1.25 .$$

Next, from Proposition 3.1.4,

$$\frac{[f(\lambda^2)]^2}{\lambda} < \frac{0.5917}{0.396} < 1.5 .$$

To estimate the third term in (4.4) we use:

$$-f'(\lambda^2) \geqq -f'(0.5) = 1 + \lambda - (Kh)'(0.5) \geqq 1 + \lambda \geqq 1.396$$

(by Propositions 3.1.1 and 2.2.4), $-f'(a) > 1.192$ (by Proposition 3.1.2), $f(\lambda^2) > 0.7491$ (by Proposition 3.1.4). Hence

$$4 f(\lambda^2)\, f'(a)\, f'(\lambda^2) > 4.98 .$$

Since, finally, the last term in (4.4) is non-negative,

$$\frac{\partial \bar{\lambda}}{\partial \lambda} \geqq -1.25 - 1.5 + 4.98 = 2.23 .$$

Next

$$\frac{\partial \bar{\lambda}}{\partial \mu} = \frac{1}{\lambda} \frac{\partial f}{\partial \mu}(a) + \frac{2}{\lambda} f'(a)\, f(\lambda^2) \frac{\partial f}{\partial \mu}(\lambda^2) .$$

From (4.3)

$$\frac{\partial f}{\partial \mu}(x) = -x(1-x) ,$$

so

$$\frac{\partial \bar{\lambda}}{\partial \mu} = \frac{1}{\lambda} [2 f(\lambda^2)(-f'(a))\, \lambda^2 (1-\lambda^2) - a(1-a)] .$$

Now $-f'(a) = 1 + \lambda - \mu(2a-1) - (Kh)'(a)$. By Proposition 2.2.8 $(Kh)'(a) \leqq 0.32/6 < 0.0534$, and hence (using Proposition 3.1.4)

$$-f'(a) > 1 + 0.396 - 0.16(2 \times 0.5917 - 1) - 0.0534 > 1.313 .$$

Thus

$$\frac{\partial \bar{\lambda}}{\partial \mu} > \frac{1}{\lambda} [2 \times 0.7491 \times 1.313 \times \lambda^2 (1-\lambda^2) - 1/4] > \frac{1}{\lambda} [0.26 - 0.25] > 0 .$$

Proof of Lemma 4.2. By Propositions 2.2.5 and 2.2.7,

$$0 \leqq (Kh)'(1) \leqq 0.32(1 - 0.36/2)/6 < 0.0438 .$$

Proof of Lemma 4.3. If $0 \leqq \gamma \leqq 0.0438$, the graph of $\mu = 1 + \lambda + 1/(2\lambda) - \gamma$ crosses the vertical line $\lambda = 0.396$ at

$$\mu = 1 + 0.396 - 1/(2 \times 0.396) - \gamma < 0.1334$$

and the horizontal line $\mu = 0.09$ at a value of $\lambda < 0.3962$ [since $1 + 0.3962 - 1/(2 \times 0.3962) - \gamma > 0.09$]. One of these crossings belongs to the boundary of the rectangle $\{(\lambda, \mu) : 0.396 \leqq \lambda \leqq 0.4031,\ 0.09 \leqq \mu \leqq 0.16\}$.

Similarly, the graph crosses the vertical line $\lambda = 0.4031$ at a value of $\mu > 0.1189$ and the horizontal line $\mu = 0.16$ at a value of $\lambda > 0.4024$, and, again, one of these crossings belongs to the boundary of the rectangle.

Proof of Lemma 4.4. We use the notation f_0, f_1 as in Sect. 1; we also (as above) write a for $[f(\lambda^2)]^2$ and a_0 for $[f_0(\lambda^2)]^2$. With this notation

$$\bar{\lambda} = \frac{1}{\lambda} f(a) = \frac{1}{\lambda} f_0(a) + \frac{1}{\lambda} f_1(a) = \frac{1}{\lambda} f_0(a_0) + \frac{1}{\lambda} f_0'(\tilde{a})(a - a_0) + \frac{1}{\lambda} f_1(a) \tag{4.5}$$

for some $\tilde{a}$ between a_0 and a. Now $f_1(a) \leqq 0$ (Proposition 3.1.1) while

$$a - a_0 = [f_0(\lambda^2) + f_1(\lambda^2)]^2 - [f_0(\lambda^2)]^2 = f_1(\lambda^2)[f(\lambda^2) + f_0(\lambda^2)] \leqq 0,$$

and $f_0'(\tilde{a}) < 0$ so, rewriting (4.5) as

$$\bar{\lambda}(\lambda, \mu, h) = \bar{\lambda}(\lambda, \mu, 0) + \frac{1}{\lambda} f_1(a) + \frac{1}{\lambda} f_0'(\tilde{a})(a - a_0), \tag{4.6}$$

we see that the second term on the right is negative and the third positive. We will next bound each of these terms.

From Proposition 2.2.3,

$$-\frac{f_1(a)}{\lambda} \leqq \frac{a^2(1-a)[0.91 - 0.09a]}{6\lambda} \times 0.32. \tag{4.7}$$

From Proposition 3.1.4, $0.5611 < a < 0.5917$, and it is easy to check that the right-hand side of (4.7) is increasing in a in this range. We thus get an upper bound by substituting 0.5917 for a and 0.396 for λ in (4.7); this gives

$$-\frac{f_1(a)}{\lambda} < 0.01651.$$

To estimate the third term on the right of (4.6), we first remark that, by Proposition 3.1.4, both a and a_0 are between 0.5611 and 0.5917, so the same is true of $\tilde{a}$. Hence, by Proposition 3.1.10, $-f_0'(\tilde{a}) < 1.4165$. Similarly, by Proposition 3.1.3, both $f_0(\lambda^2)$ and $f(\lambda^2)$ are smaller than 0.7692, so $[f(\lambda^2) + f_0(\lambda^2)] < 2 \times 0.7692$. Finally, by Proposition 2.2.3,

$$-f_1(\lambda^2) \leqq \frac{\lambda^4(1-\lambda^2)}{6}[0.91 - 0.09\lambda^2] \times 0.32.$$

Combining these estimates, we get

$$\frac{f_0'(\tilde{a})(a - a_0)}{\lambda} < 0.00572.$$

We have thus established that

$$\bar{\lambda}(\lambda, \mu, 0) - 0.01651 < \bar{\lambda}(\lambda, \mu, h) < \bar{\lambda}(\lambda, \mu, 0) + 0.00572 .$$

Hence, to prove, for example, $\bar{\lambda}(0.4031, 0.1189, h) > 0.4031$, it suffices to prove $\bar{\lambda}(0.4031, 0.1189, 0) - 0.01651 > 0.4031$, and similarly for the other three statements of Lemma 4.4. The required estimates with $h = 0$ are established by straightforward explicit computation.

5. Bounding the Third Derivative: Real Points

The objective of this section is to prove Lemma 1.2. Thus, we can assume that λ and μ are the λ^* and μ^* of Lemma 1.1. For most of the argument, we will not need to use this fact but only our standing assumptions (1.11) and (1.12) about λ, μ, and h (and we will accordingly drop the *'s). At one point, however, it will be convenient to use the identity

$$f'(1) = -1/(2\lambda), \tag{5.1}$$

which was shown in Sect. 4 to be a consequence of $\lambda = \lambda^*$, $\mu = \mu^*$.

For this section, we introduce the notation

$$a(x) \equiv [f(\lambda^2 x)]^2 ; \tag{5.2}$$

note that this is not quite consistent with the use of the symbol a in Sect. 4. We also, as above, write $\bar{f}$ for $\mathscr{T} f$.

By differentiating the definition of $\bar{f}$ we get

$$\begin{aligned} \bar{h}(x) \equiv \bar{f}'''(x) = {} & 8\lambda^5 [f(\lambda^2 x)]^3 [-f'(\lambda^2 x)]^3 h(a(x)) \\ & + 2\lambda^5 f(\lambda^2 x) [-f'(a(x))] h(\lambda^2 x) \\ & + 12\lambda^5 f(\lambda^2 x) [-f'(\lambda^2 x)]^3 f''(a(x)) \\ & + 12\lambda^5 [f(\lambda^2 x)]^2 [-f'(\lambda^2 x)] f''(\lambda^2 x) f''(a(x)) \\ & - 6\lambda^5 [-f'(\lambda^2 x)] [-f'(a(x))] f''(\lambda^2 x) . \end{aligned} \tag{5.3}$$

Since $f(\lambda^2 y)$, $-f'(y)$, and $f''(y)$ are all non-negative on $[0, 1]$ (Proposition 3.1), all terms in the above expression for $\bar{h}$ are positive except the last. We will show that $\bar{h}(x) \geqq 0$ by showing that the sum of the third and last terms is already positive. To do this, we first note that, since $f''' \geqq 0$ (by assumption), $f''(a(x)) \geqq f''(\lambda^2 x)$, and, since $f'' \geqq 0$ (Proposition 3.1.1), $-f'(\lambda^2 x) \geqq -f'(a(x))$. Thus

$$\begin{aligned} & 12\lambda^5 f(\lambda^2 x) [-f'(\lambda^2 x)]^3 f''(a(x)) - 6\lambda^5 [-f'(\lambda^2 x)] [-f'(a(x))] f''(\lambda^2 x) \\ & \quad \geqq 6\lambda^5 [-f'(\lambda^2 x)]^2 f''(a(x)) \{2[-f'(\lambda^2 x)] f(\lambda^2 x) - 1\} . \end{aligned}$$

Again using $f'' \geqq 0$, $-f'(\lambda^2 x) \geqq -f'(0.5) \geqq 1 + \lambda$. [By Proposition 2.2.4, the contribution of h to $-f'(0.5)$ is positive.] Also, by Proposition 3.1.3, $f(\lambda^2 x) \geqq f(\lambda^2) > 0.7491$. Hence,

$$2[-f'(\lambda^2 x)] f(\lambda^2 x) - 1 > 2 \times 1.396 \times 0.7491 - 1 > 0 ,$$

which completes the proof that $\bar{h} \geqq 0$.

We now rework (5.3) by using

$$f''(a(x)) = 2\mu + (Kh)''(a(x))$$

in the third and fourth terms on the right and

$$f''(\lambda^2 x) = 2\mu + (Kh)''(\lambda^2 x)$$

in the fifth. Expanding and regrouping, we get

$$\begin{aligned}\bar{h}(x) = {} & 8\lambda^5[f(\lambda^2 x)]^3[-f'(\lambda^2 x)]^3 h(a(x)) + 2\lambda^5 f(\lambda^2 x)[-f'(a(x))]h(\lambda^2 x) \\ & + 12\mu\lambda^5\{2f(\lambda^2 x)[-f'(\lambda^2 x)]^3 - [-f'(\lambda^2 x)][-f'(a(x))]\} \\ & + 24\mu\lambda^5 f''(\lambda^2 x)[-f'(\lambda^2 x)][f(\lambda^2 x)]^2 \\ & + 12\lambda^5 f(\lambda^2 x)[-f'(\lambda^2 x)]\{[-f'(\lambda^2 x)]^2 + f(\lambda^2 x) f''(\lambda^2 x)\}(Kh)''(a(x)) \\ & - 6\lambda^5[-f'(\lambda^2 x)][-f'(a(x))](Kh)''(\lambda^2 x). \end{aligned} \tag{5.4}$$

We call the six terms on the right $T_1, \ldots, T_6$ respectively; we will now proceed to estimate them one at a time.

1.

$$T_1 = 8\lambda^5[f(\lambda^2 x)]^3[-f'(\lambda^2 x)]^3 h(a(x)).$$

We use $-f'(\lambda^2 x) \leqq 1+\lambda+\mu$. (Proposition 3.1.2), $\lambda^5(1+\lambda+\mu)[f(\lambda^2 x)]^3 < 0.016638(1-0.5776x)$ (Proposition 3.1.8), $a(x) \geqq 1-0.5082x$ (Proposition 3.1.9). Combining the last of these with the assumption

$$h(x) \leqq 0.32(1-0.36x),$$

we get

$$h(a(x)) \leqq 0.32 \times 0.64 \times \left(1 + \frac{0.36 \times 0.5082}{0.64} x\right).$$

Thus,

$$T_1 < 0.066608 \times (1-0.5776x) \times (1+0.2859x) \leqq 0.066608(1-0.2917x).$$

2.

$$T_2 = 2\lambda^5 f(\lambda^2 x)[-f'(a(x))]h(\lambda^2 x).$$

We use $-f'(a(x)) < 1.4165$ (Proposition 3.1.10), $f(\lambda^2 x) \leqq 1-0.2308x$ (Proposition 3.1.5), $h(\lambda^2 x) \leqq 0.32(1-0.0564x)$ [from $h(x) \leqq 0.32(1-0.36x)$], and

$$(1-0.2308x) \times (1-0.0564x) \leqq (1-cx)$$

where

$$c = 1-(1-0.2308)(1-0.0564) > 0.2741.$$

Thus,

$$T_2 < 0.00965(1-0.2741x).$$

3.

$$T_3 = 12\mu\lambda^5\{2f(\lambda^2 x)[-f'(\lambda^2 x)]^3 - [-f'(\lambda^2 x)][-f'(a(x))]\}.$$

Here, we will use (5.1), which, combined with Proposition 3.1.1, implies

$$-f'(a(x)) \geqq -f'(1) = 1/(2\lambda).$$

We also use $-f'(\lambda^2 x) \leqq 1+\lambda+\mu$ (Proposition 3.1.2), and $f(\lambda^2 x) \leqq 1-0.2308x$ (Proposition 3.1.5). Combining these estimates:

$$T_3 \leqq 12\mu\lambda^4(1+\lambda+\mu)\{2\lambda(1+\lambda+\mu)^2(1-0.2308x)-1/2\}.$$

The right-hand side is manifestly increasing in λ and μ, so we can get an upper bound by substituting $\lambda = 0.4031$ and $\mu = 0.16$. This gives

$$T_3 < 0.11653 - 0.03602x.$$

4.

$$T_4 = 24\mu\lambda^5 f''(\lambda^2 x)[-f'(\lambda^2 x)][f(\lambda^2 x)]^2.$$

We use $f''(\lambda^2 x) < 0.328$ (Proposition 3.1.11), $-f'(\lambda^2 x) \leqq 1+\lambda+\mu$ (Proposition 3.1.2), and

$$\lambda^5(1+\lambda+\mu)[f(\lambda^2 x)]^2 < 0.016638(1-0.4370x)$$

(Proposition 3.1.7). Thus,

$$T_4 < 0.02096(1-0.4307x).$$

5.

$$T_5 = 12\lambda^5 f(\lambda^2 x)[-f'(\lambda^2 x)]\{[-f'(\lambda^2 x)]^2 + f(\lambda^2 x) f''(\lambda^2 x)\}(Kh)''(a(x)).$$

We use

$$(Kh)''(a(x)) \leqq 0.32 \times \tfrac{2}{3}(1 - \tfrac{5}{8} \times 0.36)[f(\lambda^2 x)]^2$$

[Proposition 2.2.10 and the definition (5.2) of $a(x)$], $-f'(\lambda^2 x) \leqq (1+\lambda+\mu)$ (Proposition 3.1.2),

$$\lambda^5(1+\lambda+\mu)[f(\lambda^2 x)]^3 < 0.016638(1-0.5776x)$$

(Proposition 3.1.8), $f(\lambda^2 x) \leqq 1$ (for the occurrence inside braces), and $f''(\lambda^2 x) < 0.328$ (Proposition 3.1.11). Thus,

$$T_5 < 0.09149(1-0.5776x).$$

6.

$$T_6 = -6\lambda^5[-f'(\lambda^2 x)][-f'(a(x))](Kh)''(\lambda^2 x).$$

We use

$$-(Kh)'' \leqq 0.32 \times \left(\frac{1}{3} - \frac{0.36}{12}\right)$$

(Proposition 2.2.11),

$$-f'(a(x)) < 1.4165$$

(Proposition 3.1.10), and

$$-f'(\lambda^2 x) \leqq 1+\lambda+\mu$$

(Proposition 3.1.2). Thus,

$$T_6 < 0.01373\,.$$

Collecting the estimates established above, we get

$$\begin{aligned}\bar{h}(x) &< 0.06661(1-0.2917x)+0.00965(1-0.2714x)+0.11653-0.03602x \\ &\quad +0.02096(1-0.4370x)+0.09149(1-0.5776x)+0.01373 \\ &< 0.319-0.12x < 0.32(1-0.36x)\,.\end{aligned}$$

6. Bounding the Third Derivative: Complex Points

We will use the notation of the preceding section but will assume in addition that h is analytic on a complex neighborhood of $[0,1]$. For $x \in [0,1]$,

$$\left|\frac{d}{dx}a(x)\right| = |2\lambda^2 f(\lambda^2 x) f'(\lambda^2 x)| \leqq 2\lambda^2(1+\lambda+\mu) < 0.508\,.$$

Assume, now, that h is analytic on D_δ and satisfies

$$|h(x)| \leqq 0.32(1-0.36|x|) \tag{6.1}$$

there. Then, provided δ is small enough, $a(x)$ (which is analytic on D_{δ/λ^2}) satisfies

$$\left|\frac{d}{dx}a(x)\right| \leqq 2/3 \quad \text{on} \quad D_{3/2\delta} \subset D_{\delta/\lambda^2}\,, \tag{6.2}$$

and hence maps $D_{3/2\delta}$ into D_δ. Thus,

$$\bar{f}(x) = \frac{1}{\lambda} f(a(x))$$

is analytic on $D_{3/2\delta}$. We will from now on assume that δ is small enough so that (6.2) follows from (6.1). We want to show that, possibly by making δ smaller still, we can guarantee that

$$|\bar{h}(x)| \leqq 0.32(1-0.36|x|) \quad \text{on} \quad D_{3/2\delta}\,. \tag{6.3}$$

The calculation leading to (5.3) holds for complex x as well as for real x. The estimates of Sect. 5 show that the sum of the last three terms on the right of (5.3) (or, what is the same thing, $T_3+T_4+T_5+T_6$, in the notation of Sect. 5) is bounded by $0.24272-0.09802x$ for x in $[0,1]$. Since these three terms involve only f and its first two derivatives, whereas (6.1) gives a bound on the third derivative of f, we can, by taking δ small enough, guarantee that

$$|T_3+T_4+T_5+T_6| \leqq 0.24273-0.09802x \quad \text{for} \quad x \in D_{3/2\delta}\,.$$

(Note that we have added one to the last digit of the constant term on the right.) Bounding T_1 and T_2 requires a slightly different argument. We will consider only T_1 explicitly; T_2 is handled in essentially the same way.

By definition,

$$T_1 = 8\lambda^5[f(\lambda^2 x)]^3\,[-f'(\lambda^2 x)]^3\,h(a(x))\,.$$

The estimates of Sect. 5 show that

$$8\lambda^5[f(\lambda^2 x)]^3\,[-f'(\lambda^2 x)]^3 < 0.32522(1-0.5776x) \quad \text{on} \quad [0,1]\,.$$

Again using the bound (6.1) on f''', we see that, if δ is small enough,

$$|8\lambda^5[f(\lambda^2 x)]^3\,[-f'(\lambda^2 x)]^3| < 0.32523(1-0.5776|x|) \quad \text{on} \quad D_{3/2\delta}\,.$$

Also, $a(x) \geqq 1-0.5082x$ on $[0,1]$ (Proposition 3.1.9), and hence, again if δ is small enough, $|a(x)| \geqq 0.9999-0.5083|x|$ on $D_{3/2\delta}$. Hence,

$$|T_1| \leqq 0.066616(1-0.2916|x|) \quad \text{on} \quad D_{3/2\delta}\,.$$

Similarly, we can ensure that

$$|T_2| \leqq 0.009651(1-0.2740|x|) \quad \text{on} \quad D_{3/2\delta}\,.$$

Combining, we get

$$|\bar{h}(x)| < 0.319-0.12|x| \leqq 0.32(1-0.36|x|) \quad \text{on} \quad D_{3/2\delta}\,,$$

as desired.

Continuity of the mapping $h \mapsto h^*$ follows immediately from the formula (5.3) for $\bar{h}$ and the continuity of the dependence of λ^* and μ^* on h.

References

1. Campanino, M., Epstein, H.: On the existence of Feigenbaum's fixed point. Commun. Math. Phys. **79**, 261–302 (1981)
2. Campanino, M., Epstein, H., Ruelle, D.: On Feigenbaum's functional equation. Topology **21**, 125–129 (1982)
3. Feigenbaum, M.: Quantitative universality for a class of non-linear transformations. J. Statist. Phys. **19**, 25–52 (1978)
4. Lanford, O.E.: A computer-assisted proof of the Feigenbaum conjectures. Bull. Am. Math. Soc. (New Series) **6**, 427–434 (1982)

Communicated by A. Jaffe

Received August 3, 1984

Physica 14D (1985) 403–408
North-Holland, Amsterdam

A NUMERICAL STUDY OF THE LIKELIHOOD OF PHASE LOCKING

Oscar E. LANFORD III†*
Department of Mathematics, University of California, Berkeley CA 94720, USA

Received 31 May 1983
Revised 12 June 1984

We report the results of numerical computation of the size of the set of Ω's for which the homeomorphism of the circle

$$\tilde{f}_{K,\Omega}(x) = x + \Omega + K/2\pi \sin(2\pi x) \bmod (1)$$

is phase locked, for values of K ranging from 0.1 to 1. The results suggest that for moderate values of K phase locking is unlikely except for small periods, but that for $K = 1$ almost all Ω's give phase locking, with large periods accounting for a surprisingly large portion of the total.

1. Introduction

We are going to discuss *orientation-preserving homeomorphisms* of the circle, i.e., continuous one-to-one mappings of the circle to itself which are, intuitively, increasing. More precisely, we mean the following: Let f denote a continuous strictly increasing mapping of the real line to itself such that

$$f(x+1) = f(x) + 1.$$

Such a mapping defines by passing to quotients a continuous one-to-one mapping $\tilde{f}$ of the circle – identified with the real numbers modulo the integers – to itself. We can *define* an orientation-preserving homeomorphism of the circle to be a mapping obtained in this way. The mapping f is not quite uniquely determined by the induced mapping $\tilde{f}$; two f's induce the same $\tilde{f}$ if and only if they differ by an integer constant. We can use this arbitrariness to arrange, say, that $0 \leq f(0) < 1$; f will then be unique.

From the point of view of dynamics, orientation-preserving homeomorphisms of the circle can be grouped broadly into two classes, those which have at least one periodic cycle and those which do not (or, equivalently, those with respectively rational and irrational rotation numbers). If an orientation-preserving homeomorphism has a cycle of some period, then every orbit for that homeomorphism is asymptotic to a cycle of that period (so in particular every cycle must have the same period). On the other hand, a (sufficiently smooth) *diffeomorphism* (differentiable mapping of the above kind, with f' strictly positive everywhere), which has no periodic cycles, can be converted, by a continuous change of coordinates, into an irrational rotation. (The precise smoothness requirement is that f' be of bounded variation. This is a theorem of Denjoy, a proof may be found in Cornfeld et al. [2], §3.4.)

A mapping with at least one cycle is frequently said to be *phase locked*. The terminology arises from the application of theory of homeomorphisms of the circle to the analysis of the behavior of a system of two weakly coupled oscillators. In connection with this and other applications, it is of interest to have some idea of how "likely" it is that an "arbitrary" homeomorphism will be phase locked. This is a question in which qualitative theory gives contradictory indications. From the

†Work supported in part by National Science Foundation Grant MCS81-07086A01.
*Current address: IHES, F91440 Bures-sur-Yvette, France.

topological standpoint, phase locking is generic, i.e., absence of periodic cycles is exceptional (Arnold [1]). From a measure-theoretic standpoint, on the other hand, absence of periodic cycles is *not* exceptional. In the paper referred to just above, for example, Arnold proves a theorem to the effect that, in a one-parameter family of *analytic* mappings which is sufficiently close (in an analytic-function topology) to a (non-constant) family of pure rotations, "most" (in the sense of Lebesgue measure) parameter values correspond to mappings which are not phase-locked. (Arnold's results have been greatly extended by Herman [5, 6]. For useful reviews of this subject, see Rosenberg [8] and Deligne [3].)

The purpose of this note is to report some numerical results on a concrete example which may give some feeling for what the quantitative situation is.

2. Statement of results

The example we study is the two-parameter family of mappings

$$f_{K,\Omega}(x) = x + \Omega + \frac{K}{2\pi}\sin(2\pi x).$$

We may restrict Ω to lie in $[0, 1)$. For $K = 0$, $\tilde{f}_{K,\Omega}$ is simply a rotation by $2\pi\Omega$ radians. We may think of K as a "non-linearity parameter" governing the deviation of $\tilde{f}_{K,\Omega}$ from a pure rotation. In order that $f_{K,\Omega}$ be non-decreasing we have to require $|K| \leq 1$; $K = 1$ corresponds to the maxi-

Table I

n	$e_n(0.1)$	$e_n(0.3)$	$e_n(0.5)$	$e_n(0.7)$	$e_n(0.8)$	$e_n(0.9)$	$e_n(1.0)$
1	0.031831	0.095493	0.159155	0.222817	0.254648	0.286479	0.318310
2	0.000398	0.007109	0.019498	0.037534	0.048500	0.060670	0.073966
3	0.000070	0.001871	0.008473	0.022486	0.032874	0.045725	0.061119
4	4.23e-06	0.000333	0.002443	0.008728	0.014274	0.021842	0.031705
5	5.51e-07	0.000129	0.001564	0.007733	0.014354	0.024503	0.039089
6	3.01e-08	0.000020	0.000352	0.002080	0.004063	0.007211	0.011891
7	5.64e-09	0.000011	0.000342	0.003002	0.006968	0.014454	0.027330
8	5.04e-10	2.70e-06	0.000115	0.001193	0.002949	0.006502	0.013059
9	7.89e-11	1.18e-06	0.000077	0.001042	0.002858	0.006884	0.014893
10	1.08e-11	4.09e-07	0.000034	0.000474	0.001327	0.003310	0.007515
11	2.14e-12	2.21e-07	0.000029	0.000612	0.002089	0.006200	0.016207
12	3.88e-13	9.32e-08	0.000013	0.000218	0.000665	0.001872	0.004953
13	8.91e-14	5.44e-08	0.000011	0.000311	0.001233	0.004309	0.013219
14	2.10e-14	2.92e-08	6.44e-06	0.000142	0.000508	0.001712	0.005321
15	5.81e-15	1.79e-08	5.09e-06	0.000137	0.000532	0.001856	0.005810
16	1.94e-15	1.11e-08	3.54e-06	0.000088	0.000354	0.001415	0.005323
17	1.33e-15	7.36e-09	2.94e-06	0.000101	0.000495	0.002295	0.009513
18		4.95e-09	2.12e-06	0.000049	0.000172	0.000691	0.003097
19		3.48e-09	1.81e-06	0.000063	0.000330	0.001729	0.008293
20			1.40e-06	0.000037	0.000148	0.000669	0.003070
21			1.19e-06	0.000038	0.000175	0.000839	0.004077
22			9.73e-07	0.000026	0.000107	0.000556	0.003178
23			8.42e-07	0.000030	0.000162	0.001033	0.006532
24			7.03e-07	0.000018	0.000066	0.000320	0.002099
25			6.16e-07	0.000021	0.000110	0.000697	0.004707
26			5.26e-07	0.000015	0.000060	0.000352	0.002628
27			4.65e-07	0.000016	0.000074	0.000461	0.003624
28			4.04e-07	0.000012	0.000046	0.000273	0.002122
29			3.61e-07	0.000013	0.000067	0.000518	0.004869
30			3.18e-07	8.83e-06	0.000028	0.000121	0.001320

mum permissible non-linearity. This family has been extensively studied. The article of Arnold cited above gives a picture of the region in K, Ω-space in which phase-locking occurs, and Arnold's theorem says that the intersection of this region with a line of constant K has Lebesgue measure going to zero with K. The (very modest) contribution of this note is to offer some detailed numerical data confirming Arnold's picture.

Given K and a positive integer n, we let $E_n(K)$ denote the set of Ω's in [0,1) for which $\tilde{f}_{K,\Omega}$ has a cycle of period n. We will see shortly that E_n is a finite union of intervals, and we will denote by $e_n(K)$ the total length of these intervals, i.e., the fraction of [0,1) occupied by Ω's such that $\tilde{f}_{K,\Omega}$ is phase-locked with period n. Table I shows the results of a numerical computation of $e_n(K)$ for several values of K and n up to 30.

3. Interpretation

One inference which can be drawn from these results is that, for moderately non-linear mappings, phase locking – especially with large periods – does not seem to be very likely. For example: When $K = 0.5$, i.e., with nonlinearity of half-maximum strength, the sum of the e_n's for n up to 30 is only about 0.192, and the steady decrease of e_n with n makes it seem unlikely that adding on the contributions for n larger than 30 will increase the sum much. (This intuitive statement can be supported, somewhat, by quantitative estimates; see below.) Furthermore, most of the sum comes from small values of n; e.g., the sum for n from 4 to 30 is only about 0.005.

The situation at $K = 1$ (maximal non-linearity) is quite different. The sum of the e_n for n up to 30 is about 0.709 and the decrease of e_n with increasing n is both slow and irregular, making direct extrapolation difficult. A little more regularity can, however, be discerned as follows: For $n > 1$, E_n is the union of $\phi(n)$ intervals, where $\phi(n)$ (the Euler ϕ function) is the number of integers between 1 and $n-1$ prime relative to n. The ratio $e_n(1)/\phi(n)$

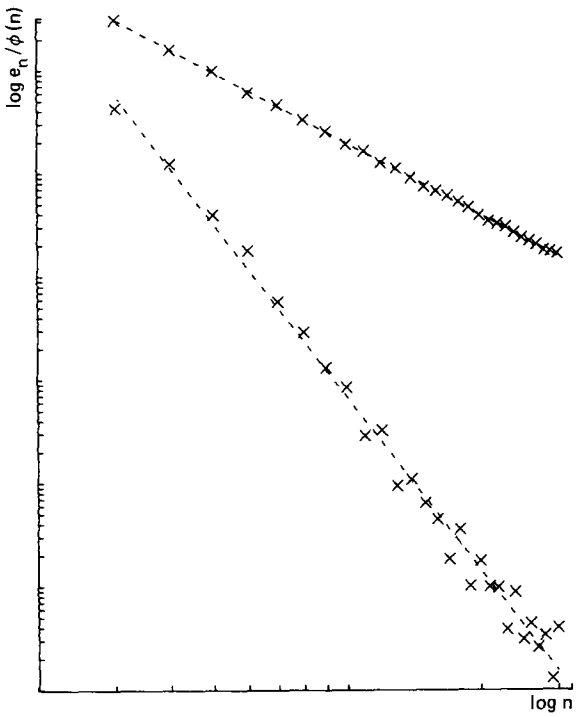


Fig. 1. Log–log plot of $e_n/\phi(n)$ vs. n for n from 1 to 30. The upper line represents $K = 1$, the lower line $K = 0.5$.

turns out to vary regularly with n for the values of n we have examined. The upper part of Fig. 1 shows a log–log plot of $e_n(1)/\phi(n)$ versus n for n from 3 to 30. The points lie remarkably close to a straight line, corresponding to a power-law dependence of $e_n(1)$ on n. (Why this should be true is a mystery.) A least-squares fit of a straight line to these points gives the empirical formula

$$e_n(1) \approx 0.377 \cdot \phi(n) \cdot n^{-2.288},$$

(shown as a dashed line in fig. 1) which has an error of no more than 6% for n from 3 to 30. Using this empirical formula to extrapolate $e_n(1)$ beyond $n = 30$, we can estimate the contributions of larger periods to the probability of phase locking. We have

$$\sum_{n=31}^{N} 0.377 \cdot \phi(n) \cdot n^{-2.288} = 0.2885 \cdots,$$

(The sum converges too slowly to be evaluated directly. To evaluate it, first show analytically that

$$\sum_{n=1}^{\infty} \phi(n) \cdot n^{-\gamma} = \zeta(\gamma - 1)/\zeta(\gamma),$$

where ζ is the Riemann zeta function; then subtract off the finite sum from $n = 1$ to 30.)

Since the e_n for n from 1 to 30 add up to about 0.709, it appears very plausible that

$$\sum_{n=1}^{\infty} e_n(1) = 1,$$

i.e., that the mappings $\tilde{f}_{1,\Omega}$ are phase-locked for almost all Ω (in the sense of Lebesgue measure). It also appears, however, that large periods are quite common, e.g., that Ω's which give phase locking with period greater than 100 comprise about 20% of the unit interval and those with period greater than 1000 about 8%.

Remarks:

1) D. Ruelle has shown me results of quite a different numerical investigation also tending to the conclusion that phase locking occurs for almost all Ω at $K = 1$. For a discussion of other related work, see section 5.

2) The empirical fit of a power of n to $e_n/\phi(n)$ works less well for $K = 0.5$ than for $K = 1$. The lower part of Fig. 1 is a log–log plot of $e_n(0.5)/\phi(n)$ against n, and the empirical formula obtained by fitting a straight line to the points shown is

$$e_n(0.5) \approx 2.47 \cdot \phi(n) \cdot n^{-5.55},$$

which is off by as much as a factor of 2.5. Furthermore, disturbingly, the fit is least good at $n = 30$. Nevertheless, for what it is worth, the empirical formula gives an estimate of only 2.3×10^{-6} for the sum of the $e_n(0.5)$ from n from 31 to infinity.

3) It is natural to ask whether the excellent fit of e_n (1) by a power of n is a special property of the family of mappings considered here or is more general, and whether the exponent $2.288\cdots$ is "universal." To check on this, preliminary computations were done on two other families. In both cases, the fit by a power law is not at all good. If, nevertheless, one insists on making a power law fit, the general trend is not inconsistent with an exponent of about 2.3.

4. Method of computation

The computation of the e_n rests on the following simple fact: In order that $\tilde{f}_{K,\Omega} \in [0,1)$, have a cycle of period $n > 1$, it is necessary and sufficient that there exist an x_0 in $[0,1)$ and an integer j between 1 and $n-1$, prime relative to n, such that

$$f^n_{K,\Omega}(x_0) = x_0 + j.$$

(In this case, the rotation number of $\tilde{f}_{K,\Omega}$ is j/n.) We thus get, immediately, the following criterion: Let

$$\phi^+(n, K, \Omega) = \max_{0 \le x < 1} \{ f^n_{K,\Omega}(x) - x \},$$

$$\phi^-(n, K, \Omega) = \min_{0 \le x < 1} \{ f^n_{K,\Omega}(x) - x \}.$$

Then $\tilde{f}_{K,\Omega}$ has a cycle of period n if and only if there is an integer j between 1 and $n-1$, prime relative to n, such that

$$\phi^-(n, K, \Omega) \le j \le \phi^+(n, K, \Omega).$$

A simple induction argument shows that $\phi^{\pm}$ are strictly increasing in Ω and indeed that the rate of increase is at least one, i.e.,

If $\Omega_1 > \Omega_2$ then

$$\phi^{\pm}(n, K, \Omega_1) \ge \phi^{\pm}(n, K, \Omega_2) + (\Omega_1 - \Omega_2).$$

Thus, for each j, there are uniquely determined $\Omega^{\pm}(n, j, K)$ such that

$$\phi^{\pm}(n, K, \Omega^{\pm}) = j,$$

and

$$E_n = \bigcup_j E_{n,j}, \quad E_{n,j} = [\Omega^+(n, j, K), \Omega^-(n, j, K)],$$

where the union is taken over j's between 1 and $n-1$ prime relative to n.

The preceding analysis does not quite apply to fixed points ($n = 1$), but it is easy to adapt the arguments given to show that $f_{K,\Omega}$ has a fixed point if and only if $0 \leq \Omega \leq K/2\pi$ or $1 - K/2\pi \leq \Omega < 1$ and hence that

$$e_1(K) = K/\pi.$$

The e_n's for $n > 1$ were computed by solving the equations

$$\phi^{\pm}(n, K, \Omega^{\pm}) = j$$

for the $\Omega^{\pm}$. These are non-linear equations whose left-hand side has to be computed numerically; they were solved using the FORTRAN subroutine ZEROIN (see Forsythe et al. [4], chap. 7), based on R. Brent's rootfinding algorithm. All calculations were done in double precision PDP-11 floating point arithmetic, i.e., with a precision of 56 bits or about 16 decimal digits. Because the $\phi^{\pm}$ increase with Ω with at least unit rate, the $\Omega^{\pm}$ can be computed as accurately as the $\phi^{\pm}$ can. Computing the $\phi^{\pm}$ is a global extremization problem. The approach used was to compute $f^n_{K,\Omega}(x) - x$ at each of 51 uniformly-spaced points running from 0 to 1; to find all places on this grid where the function is larger (smaller) than at both neighboring grid points; to locate with high accuracy a local maximum (minimum) somewhere in the two grid intervals around such a local grid extremum; and to take the largest (smallest) function value found in this way as $\phi^+(\phi^-)$. The local-extremum calculation was done using the one-dimensional minimum-finding routine FMIN (again, see Forsythe et al. [4], chap. 8) which reliably finds *some* local minimum in a specified interval with accuracy limited only by round-off error. It is possible, however, because of the finite grid spacing, that the global search will miss the true extremum. As a check on whether this is happening, some extrema were recomputed using a grid spacing of 1/200 instead of 1/50; in no case did this more refined search give a significantly different result. Nevertheless, the possibility that the true extremum is sometimes missed, and hence that the computed e_n's are seriously in error, cannot be completely ruled out. If there are no such gross errors, the e_n should be in error by no more than something like 10^{-14}.

5. Addendum

After completing this manuscript I learned of the work of Jensen, Bak and Bohr [7] who also study the fraction of the Ω interval corresponding to phase locking at $K = 1$. Jensen et al. analyze their data in a different way from what we have done; they examine the dependence on r of the fraction of the parameter interval not covered by phase-locking intervals of length at least r (rather than the dependence of n of the fraction of the parameter interval covered by phase-locking intervals of period n.) They find a power law dependence on r and present strong evidence that this power law is universal (in contrast to the power law dependence on n which we find in our particular example but which seems to fit other examples less well.) It is thus fair to say that their work provides stronger evidence than ours that phase-locking occurs except for a set of parameters of measure zero; it furthermore indicates that this is the general situation for critical circle mappings.

On the other hand, precisely because Jensen et al. don't analyze their numerical results in terms of the periods of the cycles considered, their work does not exhibit the fact that it is necessary to consider rather large periods in order to account for as much as, say, 90% of the parameter interval. It also seems to me to be worthy of note that the situation is quantitatively radically different even for K's as large as 0.5.

References

[1] V.I. Arnold, Small denominators I, Amer. Math. Soc. Translations, ser. 2, vol. 46 (1965) 213–284.

[2] I.F. Cornfeld, S.V. Fomin amd Ya.G. Sinai, Ergodic Theory (Springer, New York, 1982).

[3] P. Deligne, Les difféomorphismes du cercle [d'après M.R.

Herman], Lecture Notes in Mathematics 567 (1977) 99–121.
[4] G.E. Forsythe, M.A. Malcolm and C.B. Moler, Computer Methods for Mathematical Computations (Prentice-Hall, Englewood Cliffs, 1977).
[5] M.R. Herman, Mesure de Lebesgue et nombre de rotation, Lecture Notes in Mathematics 597 (1977) 271–293.
[6] M.R. Herman, Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations, Publ. Math. IHES 49 (1979) 5–233.
[7] M. Høgh Jensen, P. Bak and T. Bohr, Complete devil's staircase, fractal dimension, and universality of mode-locking structure in the circle map, Phys. Rev. Lett. 50 (1983) 1637–1639.
[8] H. Rosenberg, Les difféomorphismes du cercle [d'aprés M.R. Herman], Lecture Notes in Mathematics 567 (1977) 81–98.

TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 348, Number 4, April 1996

TRANSFER OPERATORS ACTING ON ZYGMUND FUNCTIONS

VIVIANE BALADI, YUNPING JIANG, AND OSCAR E. LANFORD III

ABSTRACT. We obtain a formula for the essential spectral radius ρ_{ess} of transfer-type operators associated with families of $C^{1+\delta}$ diffeomorphisms of the line and Zygmund, or Hölder, weights acting on Banach spaces of Zygmund (respectively Hölder) functions. In the uniformly contracting case the essential spectral radius is strictly smaller than the spectral radius when the weights are positive.

1. INTRODUCTION

During the last decade, a generalised theory of Fredholm determinants has been obtained using tools from statistical mechanics, often in a dynamical setting. Typically, one considers

- a transformation f, with finitely or countably many inverse branches, of a metric space M to itself,
- a weight $g : M \to \mathbb{C}$;

and one defines the associated *transfer operator*

$$\mathcal{L}\varphi(z) = \sum_{f(w)=z} g(w)\varphi(w)$$

acting on a Banach space of functions $\varphi : M \to \mathbb{C}$. Transfer operators are useful in the study of "interesting" invariant measures for f. They sometimes arise in a surprising fashion: It has been proved that the period-doubling renormalization spectrum is exactly the spectrum of a suitably defined transfer operator (see e.g. Jiang-Morita-Sullivan [6]). Transfer operators are usually bounded but non-compact; however, it has been possible in many cases to compute an upper bound, or even an exact value for the *essential spectral radius* ρ_{ess} of $\mathcal{L}$. This is the first step towards a generalised Fredholm theory. The second step is to introduce a *generalised Fredholm determinant,* which is often closely connected to weighted *dynamical zeta functions* (see Section 5). One then shows under suitable assumptions that the determinant is an analytic function in a subset of the complex plane, or that the zeta function is meromorphic in some domain, where its zeroes (respectively poles) describe exactly the spectrum of $\mathcal{L}$ outside of a disc of radius $r \geq \rho_{\text{ess}}$.

Received by the editors March 30, 1995.
1991 *Mathematics Subject Classification.* Primary 47A10, 47B38, 58F03, 26A16.
Y. Jiang is partially supported by an NSF grant (contract DMS-9400974), and PSC-CUNY awards.

This program has been successfully carried out in an Axiom A framework with various degrees of smoothness (Hölder, analytic, differentiable: see Parry-Pollicott [12]; and Rugh [18] for more recent developments), for families of contractions on finite dimensional manifolds and $C^{k+\alpha}$ smoothness, $0 \leq k \leq \omega$, $0 \leq \alpha \leq 1$ (Ruelle [15, 16], Fried [4]). In dimension one, one may consider test functions of bounded variation (see Ruelle [17] and references therein, Baladi-Ruelle [1]), and under Markov-type assumptions also C^k Banach spaces (Collet-Isola [2]).

One Banach space which had not yet been investigated in this context is the space $Z(I)$ of Zygmund functions on an interval or circle I (see Section 2 for definitions). The space $Z(I)$, which has been much used in dynamical systems in recent years, notably in Sullivan's analysis of renormalisation (Sullivan [19]), is interesting not only because $\Lambda^1 \subsetneq Z \subsetneq \Lambda^\alpha$ for all $0 < \alpha < 1$, where Λ^α denotes the space of α-Hölder functions ($\Lambda^1 = \mathrm{Lip(I)}$) but also because it arises in the study of quasiconformal mappings and Teichmüller theory, as we explain now.

Let I denote the circle $\mathbb{R}/\mathbb{Z}$, and choose three points $p_1 < p_2 < p_3$ in I. A homeomorphism h of I fixing p_i for $i = 1$, 2, and 3 is called quasisymmetric if

$$||h||_{qs} = \sup_{x \in I; x+t, x-t \in I} \frac{|h(x+t) - h(x)|}{|h(x) - h(x-t)|} < \infty \, .$$

Let T be the set of all orientation preserving quasisymmetric homeomorphisms of I which fix p_i for $i = 1$, 2, 3, endowed with the distance $d(h_1, h_2) = \log ||h_1 \circ h_2^{-1}||_{qs}$. The set T with distance d is a model for universal Teichmüller space (see Lehto [9]). For a fixed quasisymmetric homeomorphism h_0 in T, the right composition $R_{h_0}(h) = h \circ h_0$ acting on T is a continuous map, and sends a neighborhood of the identity to a neighborhood of h_0. This makes T into a homogeneous space. It is also known that T is a complex manifold (see Gardiner [5], Lehto [9]). Thus T has a tangent space at the identity, which is also the tangent space at any point h_0. This tangent space is a Banach space of continuous vector fields $\phi(x)d/dx$ defined on I, and, when factored by the two-dimensional subspace of affine functions, can be identified with $Z(I)$, the Zygmund function space (Reimann [14]). Therefore a transfer operator $\mathcal{L}$ acting on $Z(I)$ can be viewed as acting on the tangent space of universal Teichmüller space. It is hoped that the knowledge of the spectral properties of such operators may be applied to the study of Teichmüller theory. An especially interesting case is when $\mathcal{L}$ is the tangent map $\mathcal{DR}$ to some nonlinear operator $\mathcal{R}$ acting on universal Teichmüller space.

In this paper, we carry out the first step towards a generalised Fredholm determinant theory on Zygmund spaces: We obtain an exact formula (Theorem 1) for the essential spectral radius of transfer operators $\mathcal{L}$ acting on $Z(I)$, or Λ^α for $0 < \alpha \leq 1$ (the Λ^α case was treated by Lanford [8]), and under additional assumptions a strict inequality between the essential spectral and the spectral radii. Section 2 contains definitions and results on the essential spectral radius. To obtain the essential spectral radius, we prove an upper bound in Section 3, and a lower bound in Section 4 (our method to get the lower bound differs from the one used by Pollicott [13] and Collet-Isola [2], but is similar to the one applied by Keller [7]). Section 5 contains results on the spectral radius and two conjectures on the second part of the program mentioned above.

V.B. and Y.J. are grateful to J. Dodziuk and F. Gardiner for very useful remarks. O.E.L. thanks A. Davies for suggesting the approach used in the proof of the lower

bound. Y.J. is grateful to FIM/ETH Zürich for its kind hospitality and support during an invitation which made the present work possible.

2. Definitions and statement of results

Throughout, I denotes a compact interval (minor modifications yield results for $I = \mathbb{R}/\mathbb{Z}$), and $C > 0$ a generic constant (in particular we admit identities such as $C = 2C$).

Zygmund functions. The Zygmund space Z on I (Zygmund [20]) is the complex vector space of continuous (or, equivalently, locally bounded) functions $\varphi : I \to \mathbb{C}$ such that

$$Z(\varphi) = \sup_{\substack{x\in I\\ t>0: x\pm t\in I}} |Z(\varphi, x, t)| < \infty,$$

where $Z(\varphi, x, t) = (\varphi(x + t) + \varphi(x - t) - 2\varphi(x))/t$. The vector space Z becomes a Banach space when endowed with the norm $\|\varphi\|_Z = \max(\sup_I |\varphi|, Z(\varphi))$.

For $0 < \alpha \le 1$, let Λ^α denote the space of α-Hölder functions, i.e. functions $\varphi : I \to \mathbb{C}$ satisfying

$$|\varphi|_\alpha = \sup_{x\neq y\in I} \frac{|\varphi(x) - \varphi(y)|}{|x - y|^\alpha} < \infty\,.$$

In particular, Λ^1 is the space of Lipschitz functions. Each Λ^α is a Banach space for the norm $\|\varphi\|_\alpha = \max(\sup_I |\varphi|, |\varphi|_\alpha)$; $Z \subsetneq \Lambda^\alpha$ for $0 < \alpha < 1$; and $\Lambda^1 \subsetneq Z$. (For a proof of the second assertion, see e.g. de Melo-van Strien [10, p. 293]; for an example showing that $\Lambda^1 \neq Z$, see the remark following the proof of Lemma 1.) We shall also consider the Banach space $\mathcal{B}$ of bounded functions on I endowed with the supremum norm.

Note that the norms $\|\varphi\|_{Z,\alpha} = \max(\sup_I |\varphi|, Z(\varphi), |\varphi|_\alpha)$ for $0 < \alpha < 1$ on Z are all equivalent with the norm $\|\cdot\|_Z$. (Indeed, for each $0 \le \alpha < 1$ the space Z is a Banach space for the norm $\|\cdot\|_{Z,\alpha}$; the open mapping theorem may then be applied to the identity maps $(Z, \|\cdot\|_{Z,\alpha}) \to (Z, \|\cdot\|_Z)$.) In other words, for each $0 \le \alpha < 1$, there is a constant $K = K(\alpha)$ such that

$$|\varphi|_\alpha \le K(\alpha)\,(\sup|\varphi| + Z(\varphi))\,, \quad \forall \varphi \in Z\,.$$

The following key lemma may be proved by direct computation:

Zygmund derivation of a product. *For all φ, ψ in $Z(I)$, $x \in I$, and $t > 0$,*

$$\begin{aligned} Z(\varphi\psi, x, t) = {} & \varphi(x)Z(\psi, x, t) + \psi(x)Z(\varphi, x, t) \\ & + t\cdot\Delta_+(\varphi, x, t)\Delta_+(\psi, x, t) + t\cdot\Delta_-(\varphi, x, t)\Delta_-(\psi, x, t)\,, \end{aligned} \tag{2.1}$$

where $\Delta_+(\upsilon, x, t) = (\upsilon(x + t) - \upsilon(x))/t$ and $\Delta_-(\upsilon, x, t) = (\upsilon(x) - \upsilon(x - t))/t$.

The following result is also useful (the constant $1/2$ is not optimal):

Skewed Zygmund bound. *For all $\varphi \in Z$, $x, y \in I$, $0 < t < 1$,*

$$\big|\big((1 - t)\varphi(x) + t\varphi(y)\big) - \varphi((1 - t)x + ty)\big| \le \frac{1}{2}Z(\varphi)|x - y|.$$

Proof of the skewed Zygmund bound. Fix x and y. There is nothing to prove if $Z(\varphi) = 0$. Otherwise, by subtracting off an affine function, making an affine change of variables, and multiplying by a constant, we can reduce to the case $x = 0$, $y = 1$, $\varphi(0) = \varphi(1) = 0$, $Z(\varphi) = 4$. We then have

$$\left|\frac{1}{2}\varphi(u) + \frac{1}{2}\varphi(v) - \varphi(\frac{1}{2}u + \frac{1}{2}v)\right| \le |u - v|,$$

and we want to prove that $|\varphi(t)| \le 2$ for $0 \le t \le 1$. By continuity, it is enough to prove the desired bound for t a dyadic rational. We will construct recursively an increasing sequence of bounds γ_n such that $|\varphi(t)| \le \gamma_n$ for t of the form $\frac{j}{2^n}$.

We start with $\gamma_1 = 1$. For the induction step, it is evidently enough to consider

$$t = \frac{2j+1}{2^{n+1}} = \frac{1}{2}\frac{j}{2^n} + \frac{1}{2}\frac{j+1}{2^n}.$$

By the induction hypothesis, $\varphi(\frac{j}{2^n}) \le \gamma_n$ and $\varphi(\frac{j+1}{2^n}) \le \gamma_n$; by the Zygmund condition

$$\left|\varphi(t) - \big(\frac{1}{2}\varphi(\frac{j}{2^n}) + \frac{1}{2}\varphi(\frac{j+1}{2^n})\big)\right| \le \frac{1}{2^n}.$$

Hence, the bound holds inductively if we set $\gamma_{n+1} = \gamma_n + \frac{1}{2^n}$, and, since $\lim_{n\to\infty} \gamma_n = 2$, the assertion follows. □

The transfer operator. The basic data entering into the definition of the transfer operator are a dynamical system and a weight. Let $\mathcal{I}$ be a finite or countable set and $0 \le \delta < 1$. The *dynamical system* here is a family of $C^{1+\delta}$ diffeomorphisms, $f_i : I \to J_i$, for $i \in \mathcal{I}$, where the intervals $J_i \subset I$ have disjoint interiors. We assume further that $\sup_i \|f_i'\|_\delta < \infty$, in particular $\lambda := 1/\sup_{i,x} |f_i'(x)| > 0$.

The *weight* is a family of functions $g_i : I \to \mathbb{C}$, $i \in \mathcal{I}$. Such a family g_i is called *summably bounded* if $\sup{}^{\Sigma}|g| = \sum_i \sup |g_i| < \infty$.

A summably bounded family is called *summably* Λ^α if $|g|_\alpha^\Sigma = \sum_i |g_i|_\alpha < \infty$ for some $0 < \alpha \le 1$; it is called *summably Zygmund* if $Z(g)^\Sigma = \sum_i Z(g_i) < \infty$.

Define formally the transfer operator $\mathcal{L}$ associated with the families f_i and g_i, and acting on functions $\varphi : I \to \mathbb{C}$, by

$$\mathcal{L}\varphi(x) = \sum_{i\in\mathcal{I}} g_i(x)\varphi(f_i(x))\,. \tag{2.2}$$

A typical example is when the f_i are the finitely many *inverse branches* of a piecewise expanding, piecewise surjective interval map f, or the finitely many inverse branches of a one-dimensional hyperbolic repeller, and $g_i = |f_i'|$.

The following lemma is a "warm-up":

Lemma 1. *The linear operator $\mathcal{L}$ is bounded when acting on $\mathcal{B}$ (respectively Λ^α, for any $0 < \alpha \le 1$) if the family g_i is summably bounded (respectively summably Λ^α) and $\delta \ge 0$; the operator $\mathcal{L}$ is bounded when acting on Z if the family is summably Zygmund and $\delta > 0$.*

Proof of Lemma 1. It follows immediately from the definitions that

$$\sup_I |\mathcal{L}\varphi| \le \sup_I |\varphi| \sum_{i\in\mathcal{I}} \sup_I |g_i| \le \sup{}^{\Sigma}|g| \sup_I |\varphi|\,.$$

To bound the α-Hölder seminorm, we use $|x-y| \geq \lambda |f_i x - f_i y|$ for all i and get

$$
\begin{aligned}
|\mathcal{L}\varphi|_\alpha &= \sup_{x,y\in I} \frac{|\sum_i g_i(x)\varphi(f_i x) - g_i(y)\varphi(f_i y)|}{|x-y|^\alpha} \\
&\leq \sup_{x,y\in I} \frac{\sum_i |g_i(x)(\varphi(f_i x) - \varphi(f_i y))| + |\varphi(f_i y)(g_i(x) - g_i(y))|}{|x-y|^\alpha} \\
&\leq \sup{}^{\Sigma}|g| \frac{|\varphi|_\alpha}{\lambda^\alpha} + |g|_\alpha^{\Sigma} \sup_I |\varphi| \,.
\end{aligned} \tag{2.3}
$$

For the Zygmund bound, we first note that for each $x \in I$ and $t > 0$ with $x \pm t \in I$, the Zygmund derivation formula yields for any $0 < \alpha < 1$:

$$
\begin{aligned}
\big|Z(\mathcal{L}\varphi, x, t)\big| &= \left|\sum_{i\in\mathcal{I}} Z(g_i \cdot (\varphi \circ f_i), x, t)\right| \\
&\leq \sup{}^{\Sigma}|g| \sup_i \big|Z(\varphi \circ f_i, x, t)\big| + Z^{\Sigma}(g) \sup|\varphi| + \frac{2}{\lambda^\alpha} |g|_{1-\alpha}^{\Sigma} |\varphi|_\alpha \,.
\end{aligned} \tag{2.4}
$$

Defining $0 < |t_i| \leq t/\lambda$ for each $i \in \mathcal{I}$ by $f_i(x+t) = f_i(x) + t_i$, we observe that, since $\delta > 0$, there is a constant $C > 0$ such that for all i, and all $x \in I$, $t > 0$ with $x \pm t \in I$,

$$
\begin{aligned}
|f_i(x-t) - (f_i(x) - t_i)| &= |(f_i(x+t) - f_i(x)) - (f_i(x) - f_i(x-t))| \\
&= |f_i'(x+u)t - f_i'(x-v)t| \leq |f_i'|_\delta 2^\delta t^{1+\delta} \leq C t^{1+\delta} \,,
\end{aligned} \tag{2.5}
$$

where we used $0 \leq u + v \leq 2t$ and $\sup_i |f_i'|_\delta < \infty$. For each $i \in \mathcal{I}$, we decompose

$$
Z(\varphi \circ f_i, x, t) = \frac{t_i}{t} Z(\varphi, f_i(x), t_i) - \frac{\varphi(f_i(x) - t_i) - \varphi(f_i(x-t))}{t} = \mathrm{I}_i + \mathrm{II}_i \,. \tag{2.6}
$$

Clearly,

$$
\sup_{i\in\mathcal{I}} |\mathrm{I}_i| \leq \frac{1}{\lambda} Z(\varphi) \,. \tag{2.7}
$$

Now, using (2.5), we get for all i with $\mathrm{II}_i \neq 0$:

$$
|\mathrm{II}_i| \leq C \frac{|\varphi(f_i(x) - t_i) - \varphi(f_i(x-t))|}{|f_i(x) - t_i - f_i(x-t)|^{1/(1+\delta)}} \leq C\, |\varphi|_{1/(1+\delta)} \,. \tag{2.8}
$$

To finish, put (2.4) and (2.6)–(2.8) together, observing that for any $(1+\delta)^{-1} \leq \alpha < 1$ there is a constant $K(\alpha)$ with $|\varphi|_{1/(1+\delta)} \leq |\varphi|_\alpha \leq K(\alpha)\|\varphi\|_Z$, and $|g|_\alpha^{\Sigma} \leq K(\alpha) Z^{\Sigma}(g)$. □

Remark. We would like to point out that the transfer operator $\mathcal{L}$ acting on Z may be unbounded if $\delta = 0$ (even for constant weights). Indeed, it is well known that there exist Zygmund functions φ and C^1 diffeomorphisms f such that $\varphi \circ f$ is not Zygmund. For example, let $I = [-\epsilon, \epsilon]$ be a small neighbourhood of 0, let $\varphi(x) = x \log|x|$ on I, and let $f : I \to f(I) \subset I$ be a C^1 diffeomorphism with $f(0) = 0$, $f'(x) = 1$ for $x \leq 0$ and $f'(x) = 1 - 1/\sqrt{|\log(x)|}$ for $x > 0$ (in particular,

there is a constant $C > 0$ with $C < f'(x) \le 1$ on I). To check that $\varphi \circ f$ is not Zygmund, we first show, by straightforward computation, that

$$Z(\varphi \circ f, 0, t) = \left(\frac{f(t)}{t} - 1\right) \log(t) + \frac{f(t)}{t} \log \frac{f(t)}{t}, \quad \text{for } t > 0\,.$$

The second term on the right goes to zero as $t \to 0^+$; the first, on the other hand, is unbounded since

$$\left(\frac{f(t)}{t} - 1\right) \sqrt{|\log t|} = -\frac{\sqrt{|\log t|}}{t} \int_0^t \frac{ds}{\sqrt{|\log s|}} \to -1 \text{ when } t \to 0^+.$$

The essential spectral radius of the transfer operator. For each $n \ge 1$ and $i_\ell \in \mathcal{I}$, $1 \le \ell \le n$, introduce the maps $f_{\vec{i}}^{(n)} = f_{i_n} \circ \cdots \circ f_{i_1}$, and the weights $g_{\vec{i}}^{(n)}(x) = g_{i_n}(f_{i_{n-1}} \cdots f_{i_1}(x)) \cdots g_{i_2}(f_{i_1}(x)) \cdot g_{i_1}(x)$. Note that for all $n \ge 1$

$$\mathcal{L}^n \varphi(x) = \sum_{\vec{i} \in \mathcal{I}^n} g_{\vec{i}}^{(n)}(x) \varphi(f_{\vec{i}}^{(n)} x)\,.$$

Our main result is:

Theorem 1.

1. *Assume that the family g_i is summably Zygmund and that $\delta > 0$. The essential spectral radius $\rho_{\rm ess}(\mathcal{L})$ of the operator $\mathcal{L}$ acting on Z is equal to*

$$\rho_{\rm ess}(\mathcal{L}) = \lim_{n\to\infty} \left[\sup_{x\in I} \sum_{\vec{i}\in\mathcal{I}^n} |g_{\vec{i}}^{(n)}(x)|\, |f_{\vec{i}}^{(n)\prime}(x)|\right]^{1/n}$$

   (*in particular, the limit on the right exists*).

2. *If the family g_i is summably Λ^α for some $0 < \alpha \le 1$, the essential spectral radius $\rho_{\rm ess}(\mathcal{L})$ of the operator $\mathcal{L}$ acting on Λ^α is equal to*

$$\rho_{\rm ess}(\mathcal{L}) = \lim_{n\to\infty} \left[\sup_{x\in I} \sum_{\vec{i}\in\mathcal{I}^n} |g_{\vec{i}}^{(n)}(x)|\, |f_{\vec{i}}^{(n)\prime}(x)|^\alpha\right]^{1/n}.$$

The proof of Theorem 1 is based on the following result of Nussbaum [11], which holds for any bounded linear operator $\mathcal{L}$ on a Banach space:

$$\rho_{\rm ess}(\mathcal{L}) = \lim_{n\to\infty} \left(\inf\{\|\mathcal{L}^n - \mathcal{K}\| \mid \mathcal{K} \text{ compact }\}\right)^{1/n}.$$

Indeed, using the above equality and the expression of $\mathcal{L}^n$ as a sum over $\mathcal{I}^n$, the theorem will be an immediate consequence of the two following lemmas:

Lemma 2 (Upper bound). *There is a universal constant $C > 0$ so that, for any family f_i with $\delta > 0$ and summably Zygmund g_i,*

$$\inf\{\|\mathcal{L} - \mathcal{K}\|_Z \mid \mathcal{K} : Z \to Z \text{ compact }\} \le C \cdot \sup_{x\in I} \sum_{i\in\mathcal{I}} |g_i(x)|\, |f_i'(x)|\,;$$

and for any family f_i with $\delta \ge 0$ and summably Λ^α weights g_i

$$\inf\{\|\mathcal{L} - \mathcal{K}\|_\alpha \mid \mathcal{K} : \Lambda^\alpha \to \Lambda^\alpha \text{ compact }\} \le C \cdot \sup_{x\in I} \sum_{i\in\mathcal{I}} |g_i(x)|\, |f_i'(x)|^\alpha\,.$$

Lemma 3 (Lower bound). *For any family f_i with $\delta > 0$ and summably Zygmund g_i,*

$$\inf\{\|\mathcal{L} - \mathcal{K}\|_Z \mid \mathcal{K} : Z \to Z \ \ \textit{compact}\ \} \geq \sup_{x\in I} \sum_{i\in\mathcal{I}} |g_i(x)|\,|f_i'(x)|\,.$$

For any family f_i with $\delta \geq 0$ and summably Λ^α weights g_i,

$$\inf\{\|\mathcal{L} - \mathcal{K}\|_\alpha \mid \mathcal{K} : \Lambda^\alpha \to \Lambda^\alpha \ \textit{compact}\ \} \geq \sup_{x\in I} \sum_{i\in\mathcal{I}} |g_i(x)|\,|f_i'(x)|^\alpha\,.$$

The essential spectral radius of restrictions of linear operators. If the family g_i is summably Zygmund and $\delta > 0$, it follows from Theorem 1 that the essential spectral radius of $\mathcal{L}$ acting on $Z(I)$ is the limit of its essential spectral radii on Λ^α as $\alpha \to 1$. Moreover, if the family g_i is summably Lipschitz, $\mathcal{L}$ has the same essential spectral radius when acting on Λ^1 or Z. Although this is hardly surprising, we believe that part 1 of Theorem 1 cannot be easily deduced from part 2, i.e., that the Zygmund result cannot be deduced immediately from the Λ^α, $0 < \alpha \leq 1$, results: The essential spectrum of a bounded operator contains its residual spectrum, which can be very badly behaved under restriction (see e.g. Dowson [3]). In this respect, we recall the very well known example of the shift operator acting on the Hilbert space $\ell^2 = \{(x_k)_{k\in\mathbb{Z}} \mid x_k \in \mathbb{C}\,, \sum_k |x_k|^2 < \infty\}$ by $(T(\vec{x}))_j = x_{j-1}$, whose spectrum is the unit circle, but which has the property that the spectrum of its restriction to the closed invariant space of sequences $\{(x_k)_{k\in\mathbb{Z}} \in \ell^2 \mid x_k = 0\,, k \leq 0\}$ fills the whole unit *disc.*

It can happen that the essential spectral radius *decreases* when one lets $\mathcal{L}$ act on the bigger spaces Λ^α for $\alpha < 1$ instead of Z. A simple example can be constructed as follows: We take $I = [0, 1]$ and the index set $\mathcal{I}$ to have one member 1. We then take for f_1 an analytic diffeomorphism $I \to I$ satisfying $f_1'' < 0$, and having exactly two fixed points 0 and 1, with $f_1'(0) > 1$ and $f_1'(1) < 1$. If g_1 is analytic and satisfies $g_1(0) = 1$ and $0 \leq g_1(x) \leq 1$ for all $x \in I$, then Theorem 1 yields that the essential spectral radius of $\mathcal{L}$ acting on Z or Λ^1 is $f_1'(0) > 1$, but shrinks to $f_1'(0)^\alpha$ when $\mathcal{L}$ acts on Λ^α for $0 < \alpha < 1$. If $\sup |f_i'| \leq 1$ for all i, this shrinking phenomenon is of course not possible.

3. The upper bound

To prove the upper bound we consider an explicit sequence of compact projections. Assuming that $I = [0, 1]$ to fix ideas, define for integers $n \geq 1$

$$\tau_j^{(n)} = \frac{j}{n}, \qquad j = 0, \dots, n, \tag{3.1}$$

and let $P^{(n)}$ be the compact operator of piecewise affine interpolation at the $\tau_j^{(n)}$. (I.e., $P^{(n)}\varphi$ is the unique function which is affine on each interval $[\tau_{j-1}^{(n)}, \tau_j^{(n)}]$ and which agrees with φ at the points $\tau_j^{(n)}$.) We write $Q^{(n)} = 1 - P^{(n)}$, where 1 denotes the identity operator. For simplicity, we often drop the superscript (n). We will use the compact operators $\mathcal{K} = \mathcal{K}^{(n)} = \mathcal{L} - Q^{(n)}\mathcal{L}Q^{(n)}$.

For each fixed $n \geq 1$, it will be convenient to use the auxiliary seminorms

$$
\begin{aligned}
|\varphi|_\alpha^{(n)} &= \sup_{0\leq j\leq n-1} \sup_{\tau_j\leq x<y\leq\tau_{j+1}} \frac{|\varphi(x)-\varphi(y)|}{|x-y|^\alpha}\,, \text{ for } \varphi\in\Lambda^\alpha(I)\,,\ 0<\alpha\leq 1\,,\\
Z^{(n)}(\varphi) &= \sup_{0\leq j\leq n-2} \sup_{\substack{x\in I\\ t>0: x\pm t\in[\tau_j,\tau_{j+2}]}} |Z(\varphi,x,t)|\,, \text{ for } \varphi\in Z(I)\,.
\end{aligned}
\tag{3.2}
$$

Obviously, $|\varphi|_\alpha^{(n)} \leq |\varphi|_\alpha$ and $Z^{(n)}(\varphi) \leq Z(\varphi)$. We summarize properties of the operators $Q^{(n)}$ and the seminorms $|\cdot|_\alpha^{(n)}$ and $Z^{(n)}(\cdot)$:

Sublemma 4. *For any $n \geq 1$ and $0 < \alpha \leq 1$:*

1. $\sup|Q^{(n)}\varphi| \leq 2\sup|\varphi|$ *and* $\sup|Q^{(n)}\varphi| \leq (2n)^{-\alpha}|Q^{(n)}\varphi|_\alpha^{(n)}$, *for each* $\varphi\in\Lambda^\alpha$.
2. $|Q^{(n)}\varphi|_\alpha^{(n)} \leq 2|\varphi|_\alpha^{(n)}$ *and* $|Q^{(n)}\varphi|_\alpha \leq 2|Q^{(n)}\varphi|_\alpha^{(n)}$, *for each* $\varphi\in\Lambda^\alpha$.
3. $Z(Q^{(n)}\varphi) \leq 4Z^{(n)}(\varphi)$, *for each* $\varphi\in Z$.

Proof of Sublemma 4. Clearly, $\sup|P\varphi| \leq \sup|\varphi|$, which yields the first bound by the definition of Q. The other claim is immediate too, since $Q\varphi$ vanishes at the τ_j and any point x is within distance at most $1/(2n)$ of some τ_j.

To prove the first bound for the α-Hölder seminorm it suffices to control P. Consider a pair of points $x < y$ belonging to the same interval $[\tau_{j-1},\tau_j]$. Then $(P\varphi(y)-P\varphi(x))/(y-x) = (\varphi(\tau_j)-\varphi(\tau_{j-1}))/(\tau_j-\tau_{j-1})$. Therefore

$$
\frac{|P\varphi(y)-P\varphi(x)|}{|y-x|^\alpha} = \frac{|\varphi(\tau_j)-\varphi(\tau_{j-1})|}{|\tau_j-\tau_{j-1}|^\alpha}\cdot\left(\frac{|y-x|}{|\tau_j-\tau_{j-1}|}\right)^{1-\alpha} \leq |\varphi|_\alpha^{(n)}\,.
\tag{3.3}
$$

To prove the second bound, write $\psi = Q\varphi$ and consider $x < y$. If there is some j with $\tau_{j-1} \leq x < y \leq \tau_j$, then we have by definition $|\psi(y)-\psi(x)| \leq |\psi|_\alpha^{(n)}|y-x|^\alpha$. Otherwise, there are j and k such that $\tau_{j-1} \leq x < \tau_j \leq \tau_{k-1} < y \leq \tau_k$. Then, since $\psi(\tau_j) = 0 = \psi(\tau_{k-1})$, we have

$$
\begin{aligned}
|\psi(y)-\psi(x)| &\leq |\psi(y)-\psi(\tau_{k-1})| + |\psi(\tau_j)-\psi(x)|\\
&\leq |\psi|_\alpha^{(n)}\big(|y-\tau_{k-1}|^\alpha + |\tau_j-x|^\alpha\big)\\
&\leq 2|\psi|_\alpha^{(n)}|y-x|^\alpha\,.
\end{aligned}
$$

To prove the claim on the Zygmund seminorm, we first show that

$$
Z^{(n)}(P^{(n)}\varphi) \leq Z^{(n)}(\varphi)\,.
\tag{3.4}
$$

Since both $P^{(n)}$ and $Z^{(n)}$ can be built a pair of successive intervals at a time, it is enough to consider the case $n = 2$, in which case we can write simply Z rather than $Z^{(n)}$. By an affine change of variable, we can assume that the working interval is $[-1,1]$, and, by subtracting a linear function from φ, then multiplying by an overall constant, we can assume that $\varphi(-1) = \varphi(1) = 1$ and $\varphi(0) = 0$ or $\varphi(0) = 1$, i.e., $P^{(n)}\varphi(x) = |x|$ or $P^{(n)}\varphi(x) \equiv 0$. It suffices to consider the case $\varphi(0) = 0$. Then, on the one hand,

$$
Z(\varphi) \geq |\varphi(1)+\varphi(-1)-2\varphi(0)| = 2,
$$

and, on the other hand, for $t > 0$,

$$0 \le \frac{|x+t| - |x|}{t} - \frac{|x| - |x-t|}{t} \le 1 - (-1) = 2,$$

so that $Z(|\,.\,|) = 2$, proving (3.4).

We now show that $Z(Q^{(n)}\varphi) \le 4Z^{(n)}(\varphi)$. Recall that $Z(Q^{(n)}\varphi)$ is defined as the supremum of $|Z(Q^{(n)}\varphi, x, t)|$ over an appropriate set of pairs x, t; $Z^{(n)}(Q^{(n)}\varphi)$ as the supremum of the same quantity over the set of pairs such that $x \pm t$ lie in the union of some pair of successive subintervals. By (3.4), this latter supremum can be majorized by $Z^{(n)}(P^{(n)}\varphi) + Z^{(n)}(\varphi) \le 2Z^{(n)}(\varphi)$, so the asserted bound holds when $x \pm t$ lie in the union of a pair of successive intervals.

If, on the other hand, $x \pm t$ do not lie in the union of two successive subintervals, then $|t|$ must be $> 1/2n$. By the skewed Zygmund bound and the fact that $Q^{(n)}(\varphi)$ vanishes at the division points,

$$|Q^{(n)}(\varphi)(s)| \le \frac{1}{2} Z^{(n)}(\varphi) \frac{1}{n} \quad \text{for all relevant } s,$$

so we can estimate

$$|Z(Q^{(n)}\varphi, x, t)| \le 4 \cdot \frac{1}{2} Z^{(n)}(\varphi) \frac{1}{n} \cdot \frac{1}{|t|} \le 4Z^{(n)}(\varphi),$$

using $|t| \ge 1/(2n)$. Thus, the asserted bound also holds when $x \pm t$ do not lie in the union of two successive subintervals. □

For each fixed $n \ge 1$ and each $0 < \alpha \le 1$, we define

$$\beta_\alpha^{(n)} = \sup_{0 \le j \le n-1} \sup_{x,y \in [\tau_j^{(n)}, \tau_{j+1}^{(n)}]} \sum_{i \in \mathcal{I}} |g_i(x)|\, |f_i'(y)|^\alpha .$$

For $0 < \alpha \le 1$, and large enough n, $\beta_\alpha^{(n)}$ is arbitrarily close to

$$\sup_{x \in I} \sum_{i \in \mathcal{I}} |g_i(x)|\, |f_i'(x)|^\alpha .$$

The next sublemma shows the usefulness of the seminorms $|\cdot|_\alpha^{(n)}$, $Z^{(n)}(\cdot)$:

Sublemma 5. *If g_i is summably Λ^α for $0 < \alpha \le 1$, then for each $n \ge 1$ and $\varphi \in \Lambda^\alpha$*

$$|\mathcal{L}\varphi|_\alpha^{(n)} \le |g|_\alpha^\Sigma \sup|\varphi| + \beta_\alpha^{(n)} |\varphi|_\alpha .$$

If g_i is summably Zygmund and $\delta > 0$, there are constants $K > 0$ and $\epsilon > 0$, depending only on the families f_i and g_i, such that for any $n \ge 1$, and $\varphi \in Z$,

$$Z^{(2n)}(\mathcal{L}\varphi) \le K \sup|\varphi| + \big(\beta_1^{(n)} + \frac{K}{n^\epsilon}\big) Z(\varphi) .$$

Proof of Sublemma 5. We first prove the bound on the Λ^α seminorm by refining (2.3). Let $\varphi \in \Lambda^\alpha$ and $\tau_{j-1} \le x < y \le \tau_j$. Then there are points $z_i \in [x, y]$ with

$$\begin{aligned}|\mathcal{L}\varphi(y) - \mathcal{L}\varphi(x)| &\le \sum_{i\in\mathcal{I}} \Big(|g_i(y) - g_i(x)|\,|\varphi(f_i(y))| + |g_i(x)|\,|\varphi(f_i(y)) - \varphi(f_i(x))| \Big) \\ &\le |g|_\alpha^\Sigma \sup|\varphi|\,|x-y|^\alpha + \sum_{i\in\mathcal{I}} |g_i(x)|\,|\varphi|_\alpha |f_i(y) - f_i(x)|^\alpha \\ &= \left(|g|_\alpha^\Sigma \sup|\varphi| + \sum_{i\in\mathcal{I}} |g_i(x)|\,|f_i'(z_i)|^\alpha |\varphi|_\alpha \right) |x-y|^\alpha \\ &\le \left(|g|_\alpha^\Sigma \sup|\varphi| + \beta_\alpha^{(n)} |\varphi|_\alpha \right) |x-y|^\alpha ,\end{aligned}$$

as claimed.

To prove the Zygmund bound, we fix $0 < \alpha < 1$ and consider x, $x \pm t$ in some $[\tau_j^{(2n)}, \tau_{j+2}^{(2n)}]$. We first rewrite (2.4) more carefully:

(3.5)
$$\big|Z(\mathcal{L}\varphi, x, t)\big| \le \sum_{i\in\mathcal{I}} |g_i(x)|\,|Z(\varphi\circ f_i, x, t)| + Z^\Sigma(g) \sup|\varphi| + \frac{2}{\lambda^\alpha} \left(\frac{1}{n}\right)^\epsilon |g|_{1-\alpha+\epsilon}^\Sigma |\varphi|_\alpha ,$$

where $0 < \epsilon < \delta$ is such that $1 - \alpha + \epsilon < 1$, and we used $t < 1/n$. To bound the first term in the right-hand side of (3.5), we may use the decomposition (2.6) of $Z(\varphi\circ f_i, x, t)$ into $\mathrm{I}_i + \mathrm{II}_i$. Then, by definition of the t_i, there are points $z_i \in [x, x+t]$ so that

$$\mathrm{I}_i = f_i'(z_i) Z(\varphi, f_i(x), t_i) . \tag{3.6}$$

Using again $t < 1/n$, we may rewrite (2.8) as

$$|\mathrm{II}_i| \le C \left(\frac{1}{n}\right)^\epsilon |\varphi|_{\frac{1+\epsilon}{1+\delta}} . \tag{3.7}$$

Setting $\alpha = (1+\epsilon)/(1+\delta) < 1$, the bounds (3.5)–(3.7) yield a constant $C > 0$, depending only on the f_i, with

$$\begin{aligned}Z^{(n)}(\mathcal{L}\varphi) \le\ & Z^\Sigma(g) \sup|\varphi| + \beta_1^{(n)} Z(\varphi) \\ & + \left(\frac{1}{n}\right)^\epsilon \Big(\frac{2}{\lambda^{\frac{1+\epsilon}{1+\delta}}} |g|_{\delta\cdot\frac{1+\epsilon}{1+\delta}}^\Sigma + C \sup{}^\Sigma|g| \Big) |\varphi|_{\frac{1+\epsilon}{1+\delta}} ,\end{aligned}$$

To finish the proof, we proceed as in Lemma 1 to bound the Λ^α seminorms. □

Proof of Lemma 2. It suffices to show that there is a universal constant $C > 0$ so that for each $n \ge 1$

$$\limsup_{n\to\infty} \|Q^{(n)} \mathcal{L} Q^{(n)}\|_\alpha \le C \sup_{x\in I} \sum_{i\in\mathcal{I}} |g_i(x)|\,|f_i'(x)|^\alpha ,$$

when the g_i are summably Λ^α and $\delta \ge 0$, and

$$\limsup_{n\to\infty} \|Q^{(2n)} \mathcal{L} Q^{(2n)}\|_Z \le C \sup_{x\in I} \sum_{i\in\mathcal{I}} |g_i(x)|\,|f_i'(x)| ,$$

when the g_i are summably Zygmund and $\delta > 0$.

Applying Sublemma 4, we get for each $\varphi \in \Lambda^\alpha$, $n \geq 1$:

$$\begin{aligned}\sup|Q^{(n)}\mathcal{L}Q^{(n)}\varphi| &\leq C\sup|\mathcal{L}Q^{(n)}\varphi| \leq C\sum_{i\in\mathcal{I}}\sup|g_i|\,\sup|Q^{(n)}\varphi| \\ &\leq C\sup{}^{\Sigma}|g|\,\frac{|\varphi|_\alpha}{(2n)^\alpha}\,.\end{aligned}$$

Applying again Sublemma 4, and also Sublemma 5, we get for any $\varphi \in \Lambda^\alpha$, $n \geq 1$:

$$\begin{aligned}|Q^{(n)}\mathcal{L}Q^{(n)}\varphi|_\alpha &\leq C|\mathcal{L}Q^{(n)}\varphi|_\alpha^{(n)} \leq C\cdot\big(|g|_\alpha^\Sigma\sup|Q^{(n)}\varphi| + \beta_\alpha^{(n)}|Q^{(n)}\varphi|_\alpha\big) \\ &\leq C\cdot\big(|g|_\alpha^\Sigma\frac{|\varphi|_\alpha}{(2n)^\alpha} + \beta_\alpha^{(n)}|\varphi|_\alpha\big)\,.\end{aligned}$$

Finally, with Sublemmas 4 and 5, we obtain for each $\varphi \in Z(I)$, $0 < \alpha < 1$, and $n \geq 1$:

$$\begin{aligned}Z(Q^{(2n)}\mathcal{L}Q^{(2n)}\varphi) &\leq CZ^{(2n)}(\mathcal{L}Q^{(2n)}\varphi) \\ &\leq C\big(K\sup|Q^{(2n)}\varphi| + (\beta_1^{(n)} + \frac{K}{n^\epsilon})Z(Q^{(2n)}\varphi)\big) \\ &\leq C\cdot\big(\frac{K}{n^\alpha}|\varphi|_\alpha + (\beta_1^{(n)} + \frac{K}{n^\epsilon})Z(\varphi)\big)\,,\end{aligned}$$

where $C > 0$ is universal and $K > 0$, $\epsilon > 0$ depend on the f_i and g_i (but not on n). □

4. The lower bound

The idea for the argument yielding the lower bound on the Banach spaces Λ^α $(0 < \alpha \leq 1)$ is originally due to A. Davies (Lanford [8]). The Zygmund case can be treated similarly, as will be shown now.

Proof of Lemma 3. To prove the Zygmund claim, we introduce the continuous function

$$\beta_1(x) = \sum_{i\in\mathcal{I}}|g_i(x)|\,|f_i'(x)|\,.$$

Writing $\bar{\beta}_1 = \sup_{x\in I}\beta_1(x)$, the first assertion of Lemma 3 is that the infimum of $\|\mathcal{L} - \mathcal{K}\|_Z$ for $\mathcal{K}$ compact is not less than $\bar{\beta}_1$. Fix $\epsilon > 0$ small. We may assume that $\mathcal{I}$ is finite, since otherwise replacement of $\mathcal{I}$ by a large finite subset of $\mathcal{I}$ in the definition of $\beta_1(x)$ yields a supremum arbitrarily close to $\bar{\beta}_1$. The strategy is now to construct an infinite-dimensional subspace $\chi_\epsilon \subset Z(I)$ (with, in fact, $\chi_\epsilon \subset \Lambda^1$) such that $\|\mathcal{L}\varphi\|_Z \geq (\bar{\beta}_1 - \epsilon)\|\varphi\|_Z$ for each $\varphi \in \chi_\epsilon$.

Then, if $\mathcal{K}$ is a compact operator on $Z(I)$, there is a function $\varphi \in \chi_\epsilon$ with $\|\varphi\|_Z = 1$ and such that $\|\mathcal{K}\varphi\|_Z \leq \epsilon$, and hence such that $\|(\mathcal{L}-\mathcal{K})\varphi\|_Z \geq (\bar{\beta}_1 - 2\epsilon)$. Therefore the norm of $\mathcal{L} - \mathcal{K}$ cannot be less than $\bar{\beta}_1 - 3\epsilon$.

The construction of these subspaces goes as follows: We take a point x_∞ where $\beta_1(x_\infty) = \bar{\beta}_1$ and choose—with some care—a sequence $x_1, x_2, \ldots$ of distinct points in I converging to x_∞. We then construct a sequence of functions $\psi_1, \psi_2, \ldots$ in $\Lambda^1(I)$ such that

(P1) $\|a_1\psi_1 + \cdots + a_N\psi_N\|_Z = \max_j\{|a_j|\}$ for any $N \geq 1$ and complex numbers $a_1, \ldots, a_N$—in particular $\|\psi_j\|_Z = 1$ for every j;

(P2) $\limsup_{t\to 0} |Z(\mathcal{L}\psi_j, x_j, t)| = \beta_1(x_j) \to \bar{\beta}_1$ as $j \to \infty$;
(P3) $\mathcal{L}\psi_j$ vanishes on a neighbourhood of x_ℓ for all $j \neq \ell$.

From (P2) and (P3) we get $\|\mathcal{L}(a_1\psi_1 + \cdots + a_N\psi_N)\|_Z \geq \max_{1\leq j\leq N}\{\beta_1(x_j)|a_j|\}$, and hence, using (P1), we get for any φ in the linear span of ψ_k, $\psi_{k+1}, \cdots$

$$\|\mathcal{L}\varphi\|_Z \geq \inf_{j\geq k}\{\beta_1(x_j)\}\|\varphi\|_Z\,.$$

Thus, we can take χ_ϵ to be the closed linear span of the ψ_j's with $j \geq k$ for any sufficiently large $k = k(\epsilon)$. The problem is therefore reduced to constructing (x_j), (ψ_j) so that (P1), (P2) and (P3) hold.

We first specify how to choose the x_j's. For x_∞ as defined above, we choose inductively a sequence of x_j's converging to, but distinct from, x_∞, assuming furthermore that the $f_i(x_j)$, for $i \in \mathcal{I}$ and $j \geq 1$, are distinct from each other and from the $f_i(x_\infty)$. Suppose $x_1, \ldots, x_k$ have been chosen so that the $f_i(x_j)$ for $1 \leq j \leq k$ are distinct from each other and from the $f_i(x_\infty)$. We then choose a point x'_{k+1} near enough to x_∞ so that each $f_i(x'_{k+1})$ is nearer to its one or two neighbours in the set of $\{f_\ell(x_\infty)\}$ than *any* previous $f_\ell(x_j)$ but still not in this set. (We use here that no f_i can be locally constant.) Then, by moving x'_{k+1} a little, and using the fact that no two f_is coincide on any non-trivial interval, we find x_{k+1} so that the $f_i(x_{k+1})$ are distinct from each other, but so that the preceding "inequalities" still hold. Constructed in this way, the $f_i(x_j)$ for $i \in \mathcal{I}$ and $j \geq 1$ are all different and no $f_i(x_j)$ is an accumulation point of the others.

Now, let $\phi \in \Lambda^1(]-1,1[)$ be of Zygmund norm one, with compact support, and such that

$$\phi(t) = |t|/2 \quad \text{for small } t. \tag{4.1}$$

For any γ between 0 and 1, the rescaled function $\gamma\phi(t/\gamma)$ has the same properties, and by taking γ small we can make its support and supremum norm as small as we like. We simply construct ψ_j as a sum of functions $\psi_{i,j}$, for $i \in \mathcal{I}$, each of which is a rescaled ϕ translated to $f_i(x_j)$ (up to a complex phase), i.e., has the form

$$\psi_{i,j}(x) = \omega_{i,j}\cdot\gamma_{i,j}\cdot\phi((x - f_i(x_j))/\gamma_{i,j}),$$

where $|\omega_{i,j}| = 1$ will be chosen later, and $\gamma_{i,j} > 0$ is such that the support of $\psi_{i,j}$ is a subset of the interior of $f_i(I)$, and may be reduced further in Sublemma 6 below.

Now $\mathcal{L}\psi_j(x)$ is non-zero only if some $f_i(x)$ is in the support of some $\psi_{k,j}$. Since we can make the supports of the $\psi_{k,j}$ disjoint by making the $\gamma_{k,j}$ small enough, for any $\ell \neq j$ there is a neighbourhood of x_ℓ on which no $f_i(x)$ is in the support of any $\psi_{k,j}$. Thus, assertion (P3) holds.

We next check that by making the $\gamma_{i,j}$ sufficiently small we can guarantee that (P1) is satisfied. To carry out the verification it is convenient to relabel our objects: We label the pairs (i,j), $i \in \mathcal{I}$, $j \geq 1$, with a positive integer m, and we write $\xi_m = f_i(x_j)$. It suffices to prove the following sublemma:

Sublemma 6. *Let ξ_m, $m \geq 1$, be distinct points in I such that no ξ_m is an accumulation point of the others, and let γ_m be a sequence of positive numbers. For ϕ as defined in* (4.1)*, we set $\phi_m(x) = \omega_m\gamma_m\phi((x-\xi_m)/\gamma_m)$ for arbitrary $|\omega_m| = 1$. If the γ_m's are small enough, then, for any $N \geq 1$, and any $b_1, \ldots, b_N$,*

$$\|b_1\phi_1 + \cdots + b_N\phi_N\|_Z = b_{\max} = \max\{|b_m| \mid 1 \leq m \leq N\}\,. \tag{4.2}$$

Proof of Sublemma 6. We define $\eta = b_1\phi_1 + \cdots + b_N\phi_N$ and set $d_m = \inf\{|\xi_m - \xi_{m'}| \mid m \neq m'\} > 0$.

We claim that it suffices to take γ_m small enough so that

$$\phi_m(x) \text{ vanishes for } |x - \xi_m| \geq d_m/4 \text{ and } \sup|\phi_m| < d_m/8 \tag{4.3}$$

to get (4.2) to hold. Half of (4.2) is immediate: If we take $x = \xi_m$ and $t > 0$ very small, we have (since the supports of the ϕ_m are disjoint)

$$\eta(x+t) + \eta(x-t) - 2\eta(x) = \eta(\xi_m + t) + \eta(\xi_m - t) = \omega_m b_m \cdot t\,,$$

so that $\|\eta\|_Z \geq Z(\eta) \geq |b_m|$ for each m.

To prove the opposite inequality, we first observe that the disjointness of the supports and the fact that $\sup|\phi_m| \leq 1$ imply $\sup|\eta| \leq b_{\max}$. To prove the corresponding estimate for $Z(\eta)$ we consider general $x \in I$ and $t > 0$ with $x \pm t \in I$. We need to show that

$$|\eta(x+t) + \eta(x-t) - 2\eta(x)| \leq t b_{\max}\,.$$

Since $Z(\phi_m) = 1$ for all m, this is immediate unless $\{x, x+t, x-t\}$ intersects the supports of at least two *different* ϕ_m's. Assume thus that x is in the support of ϕ_{m_1}, $x - t$ in the support of ϕ_{m_0}, and $x + t$ in the support of ϕ_{m_2}, where the set $\{m_0, m_1, m_2\}$ is not a singleton. We leave to the reader the easier case where this set has only two elements, and suppose that $m_0 \neq m_1 \neq m_2$. By our assumption (4.3),

$$|x - t - \xi_{m_0}| \leq d_{m_0}/4 \leq |\xi_{m_0} - \xi_{m_1}|/4 \quad \text{and} \quad |x - \xi_{m_1}| \leq d_{m_1}/4 \leq |\xi_{m_1} - \xi_{m_0}|/4\,.$$

Therefore

$$t = |x - t - x| \geq |\xi_{m_0} - \xi_{m_1}|/2\,. \tag{4.4}$$

Similarly, we get $t \geq |\xi_{m_2} - \xi_{m_1}|/2$. On the other hand,

$$|\eta(x)| = |b_{k_1}||\phi_{m_1}(x)| \leq b_{\max} \sup|\phi_{m_1}| \leq b_{\max} d_{m_1}/8 \leq b_{\max}|\xi_{m_0} - \xi_{m_1}|/8\,.$$

Analogously $|\eta(x+t)| \leq b_{\max}|\xi_{m_2} - \xi_{m_1}|/8$, and $|\eta(x-t)| \leq b_{\max}|\xi_{m_0} - \xi_{m_1}|/8$. Finally, recalling (4.4),

$$|\eta(x \pm t) - \eta(x)| \leq |\eta(x \pm t)| + |\eta(x)| \leq \frac{b_{\max}}{4} \max(|\xi_{m_0} - \xi_{m_1}|, |\xi_{m_2} - \xi_{m_1}|) \leq b_{\max}\frac{t}{2}\,.$$

Since

$$|\eta(x+t) + \eta(x-t) - 2\eta(x)| \leq |\eta(x+t) - \eta(x)| + |\eta(x-t) - \eta(x)|\,,$$

this ends the proof of Sublemma 6 and therefore of assertion (P1). □

Going back to the notation with pairs $i \in \mathcal{I}$ and $j \geq 1$, it remains to choose the $\omega_{i,j}$ so that (P2) holds. To do this, we fix j and we decompose as before, using (2.1), (2.6), and the property of the support of $\psi_{i,j}$:

$$\begin{aligned} tZ(\mathcal{L}\psi_j, x_j, t) &= \sum_{i\in\mathcal{I}} t_i g_i(x_j) Z(\psi_j, f_i(x_j), t_i) - \sum_{i\in\mathcal{I}} g_i(x_j)(\psi_j(f_i(x_j)-t_i) - \psi_j(f_i(x_j-t)) \\ &\quad + \sum_{i\in\mathcal{I}} (g_i(x_j+t) - g_i(x_j))(\psi_j(f_i(x_j+t))) \\ &\quad + \sum_{i\in\mathcal{I}} (g_i(x_j) - g_i(x_j-t))(-\psi_j(f_i(x_j-t))) \\ &= Z_a - Z_b + Z_c + Z_d \end{aligned}$$

(we used the $t_i = t_{i,j}$ defined by $f_i(x_j + t) = f_i(x_j) + t_i$, and the fact that $\psi_j(f_i(x_j)) = 0$ for all i). Now, since each $\psi_j \in \Lambda^1$ (use e.g. $\#\mathcal{I} < \infty$), we have

$$|\psi_j(f_i(x_j \pm t))| \leq |\psi_j|_1 |f_i(x_j \pm t) - f_i(x_j)| \leq |\psi_j|_1 |f_i'(u)| t,$$

for some u, by construction. Therefore, using $|g|_\alpha^\Sigma < \infty$, for any $0 < \alpha < 1$, we get $Z_c + Z_d = o(t)$ when $t \to 0$. Since $\sup^\Sigma |g| < \infty$, we get, using again $\psi_j \in \Lambda^1$, that $Z_b = o(t)$ by applying (2.5) once more (recall that $\delta > 0$). By definition, $Z(\psi_j, f_i(x_j), u) = \omega_{i,j} + o(u)$ for $u \to 0$, uniformly in $i \in \mathcal{I}$ (using $\#\mathcal{I} < \infty$). Finally, $t_i/t = f_i'(x_j + u) = f_i'(x_j) + O(|u|^\delta)$ for some $|u| \leq t$. Therefore, if we choose the complex phases $\omega_{i,j}$ properly, we find:

$$tZ(\mathcal{L}\psi_j, x_j, t) = t \sum_{i\in\mathcal{I}} \omega_{i,j} g_i(x_j) f_i'(x_j) + o(t) = t \sum_{i\in\mathcal{I}} |g_i(x_j) f_i'(x_j)| + o(t)\,, \quad t \to 0\,,$$

which gives assertion (P2) above, and thus the first claim of Lemma 3.

A simple modification of the construction in the proof yields the second claim of Lemma 3: Instead of $\phi(t) = |t|/2$ for small t, we take $\phi(t) = |t|^\alpha$ (assuming that $\|\phi\|_\alpha = 1$) and we rescale by $\gamma^\alpha(\phi(t/\gamma))$, i.e., we have

$$\phi_m(x) = \omega_m(\gamma_m)^\alpha \phi((x - \xi_m)/\gamma_m),$$

where $|\omega_m| = 1$ and the points ξ_m are chosen exactly as above. The scalars γ_m are then chosen similarly as in Sublemma 6, condition (4.3) being naturally replaced by

$$\phi_m(x) \text{ vanishes for } |x - \xi_m| \geq d_m/4 \text{ and } \sup|\phi_m| < (d_m)^\alpha/4\,.$$

A slight variation on the above arguments (replacing the Zygmund seminorm by $|\cdot|_\alpha$, and using the decomposition (2.3) as a starting point) then yields

- (P1$_\alpha$) $\|a_1\psi_1 + \cdots + a_N\psi_N\|_\alpha = \max_j\{|a_j|\}$ for any N, $a_1, \ldots, a_N$. In particular, $\|\psi_j\|_\alpha = 1$ for every i;
- (P2$_\alpha$) $\limsup_{x\to x_j} |\mathcal{L}\psi_j(x) - \mathcal{L}\psi_j(x_j)|/|x_j - x|^\alpha = \beta_\alpha^{\mathcal{I}}(x_j) \to \bar\beta_\alpha^{\mathcal{I}}$ as $j \to \infty$ (we set $\beta_\alpha^{\mathcal{I}}(x) = \sum_{i\in\mathcal{I}} |g_i(x)|\,|f_i'(x)|^\alpha$ and $\bar\beta_\alpha^{\mathcal{I}} = \sup_{x\in I} \beta_\alpha^{\mathcal{I}}(x)$);
- (P3$_\alpha$) $\mathcal{L}\psi_j$ vanishes on a neighbourhood of x_ℓ for all $j \neq \ell$. □

5. The spectral radius and two conjectures

In this section, it is convenient to use the notation $\mathcal{L}_g$ instead of $\mathcal{L}$. We have the following result (the statements for Λ^α and $\mathcal{B}$ were obtained previously by Ruelle [15, 16]):

Theorem 2. *If the family g_i is summably bounded, then the spectral radius of $\mathcal{L}_{|g|}$ acting on $\mathcal{B}$ is equal to*

$$e^P := \lim_{n\to\infty} \left(\sup_{x\in I} \sum_{\vec{i}\in\mathcal{I}^n} |g_{\vec{i}}^{(n)}(x)| \right)^{1/n},$$

and the spectral radius of $\mathcal{L}_g$ on $\mathcal{B}$ is bounded above by e^P.

If the g_i are summably Zygmund and $\delta > 0$, respectively Λ^α and $\delta \geq 0$, the spectral radius of $\mathcal{L}_{|g|}$ acting on Z (respectively Λ^α) is equal to $\max(e^P, \rho_{\mathrm{ess}}(\mathcal{L}_{|g|}))$, *and the spectral radius of $\mathcal{L}_g$ acting on Z, respectively Λ^α, is bounded above by* $\max(e^P, \rho_{\mathrm{ess}}(\mathcal{L}_{|g|}))$.

Under the additional assumption that $\lambda > 1$, Theorem 2 together with Theorem 1 yields that $\rho_{\mathrm{ess}}(\mathcal{L}_{|g|})$ is strictly smaller than the spectral radius of $\mathcal{L}_{|g|}$ (except when both vanish) acting on Z (respectively Λ^α).

Proof. Since $e^{n(P-\epsilon)} \leq \sup \mathcal{L}_{|g|}^n \psi \leq e^{n(P+\epsilon)}$ for $\psi \equiv 1$, $\epsilon > 0$, and $n \geq n(\epsilon)$, the proof of Theorem 2 for the Banach space $\mathcal{B}$ is an immediate consequence of the spectral radius formula together with the easy inequality $\sup |\mathcal{L}_g^n \varphi| \leq \sup |\varphi| \sup \mathcal{L}_{|g|}^n \psi$, for all $\varphi \in \mathcal{B}$.

For the other Banach spaces, use the definition of the essential spectral radius.□

Maximal eigenfunctions, zeta functions, and two conjectures. In this subsection, we assume throughout that $\lambda > 1$.

Consider $\mathcal{L}_{|g|}$ acting on Λ^α: When our family f_i consists of the finitely many inverse branches of a (mixing) map $f : I \to I$, it is known that e^P is the only point in the spectrum of modulus e^P, that it is a simple eigenvalue, and that $\mathcal{L}_{|g|}$ admits a positive maximal eigenfunction φ (i.e., $\mathcal{L}_{|g|}\varphi = e^P\varphi$). Finally, P is the topological pressure of $\log|g|$ and f. For all these results, and a theory of equilibrium states, see Ruelle [15], where it was proven that the essential spectral radius of $\mathcal{L}_g$ acting on Λ^α is not bigger than e^P/λ^α, a result which follows from our Theorem 1, part 2. By Theorem 1, part 1, the essential spectral radius of $\mathcal{L}_{|g|}$ acting on Z is smaller than $e^P/\lambda < e^P$. Since each eigenfunction of $\mathcal{L}_{|g|}$ in Z is also an eigenfunction in Λ^α, the eigenvalue e^P is the unique point in the spectrum with modulus e^P, and it is simple with a positive eigenfunction $\varphi \in Z$. The case of countably many branches can be treated similarly.

In Ruelle [15, 16] and Fried [4], zeta functions associated with Λ^α (for $0 < \alpha \leq 1$) systems of finitely or countably many branches f_i and weights g_i were studied (in a slightly different setting—in particular the dimension was not limited to one). In our case, the zeta function is defined by

$$\zeta_g(z) = \exp \sum_{n\geq 1} \frac{z^n}{n} \sum_{\substack{\vec{i}\in\mathcal{I}^n \\ x\,:\, f_{\vec{i}}^{(n)}(x)=x}} \prod_{k=0}^{n-1} g_{i_{k+1}}(f_{i_k}\cdots f_{i_1})(x)\,.$$

The poles ω of $\zeta_g(z)$ in the disc of radius $\lambda^\alpha e^{-P}$ (where the function was shown to be meromorphic) were proved to be in bijection with the eigenvalues $\nu = 1/\omega$ of $\mathcal{L}_g$ acting on Λ^α of modulus $> e^P/\lambda^\alpha$.

For summably Λ^α weights g, we *conjecture* that $\zeta_g(z)$ is meromorphic in the disc of radius $\rho_{\rm ess}^{-1}(\mathcal{L}_g)$, where $\rho_{\rm ess}(\mathcal{L}_g)$ is given by part 2 of Theorem 1, and that its poles there are the inverses of the Λ^α-eigenvalues of $\mathcal{L}_g$ of modulus $> \rho_{\rm ess}(\mathcal{L}_g)$.

Also, if g is summably Zygmund and $\delta > 0$, we *conjecture* that $\zeta_g(z)$ is meromorphic in the disc of radius $\rho_{\rm ess}^{-1}(\mathcal{L}_g)$ for $\mathcal{L}_g$ acting on $Z(I)$, and that its poles there are in bijection $\omega = 1/\nu$ with the eigenvalues of $\mathcal{L}_g : Z(I) \to Z(I)$ of modulus $> \rho_{\rm ess}(\mathcal{L}_g)$.

The Λ^α conjectures do not immediately imply the Zygmund one, since Z is a strict subset of $\bigcap_{\alpha<1} \Lambda^\alpha$. The proof of these two conjectures would complete the second part of the program described in the introduction.

References

1. V. Baladi and D. Ruelle, *Sharp determinants*, IHES preprint (1994); to appear Invent. Math.
2. P. Collet and S. Isola, *On the essential spectrum of the transfer operator for expanding Markov maps*, Comm. Math. Phys. **139** (1991), 551–557. MR **92h:**58157
3. H.R. Dowson, *Spectral theory of linear operators*, Academic Press, London, 1978. MR **80c:**47022
4. D. Fried, *The flat-trace asymptotics of a uniform system of contractions*, Preprint (1993).
5. F. Gardiner, *Infinitesimal earthquaking and bending in universal Teichmüller space*, Preprint (1993).
6. Y. Jiang, T. Morita, and D. Sullivan, *Expanding direction of the period doubling operator*, Comm. Math. Phys. **144** (1992), 509–520. MR **93c:**58169
7. G. Keller, *On the rate of convergence to equilibrium in one-dimensional systems*, Comm. Math. Phys. **96** (1984), 181–193. MR **86k:**58071
8. O.E. Lanford III, *Essential norms of some operators on spaces of Hölder continuous functions*, Unpublished notes (1992).
9. O. Lehto, *Univalent functions and Teichmüller spaces*, Springer-Verlag, New York Berlin, 1987. MR **88f:**30073
10. W. de Melo and S. van Strien, *One-dimensional dynamics*, Springer-Verlag, Berlin, 1993. MR **95a:**58035
11. R.D. Nussbaum, *The radius of the essential spectrum*, Duke Math. J. **37** (1970), 473–478. MR **41:**9028
12. W. Parry and M. Pollicott, *Zeta functions and the periodic orbit structure of hyperbolic dynamics*, Société Mathématique de France (Astérisque, vol. 187-188), Paris, 1990. MR **92f:**58141
13. M. Pollicott, *Meromorphic extensions of generalized zeta functions*, Invent. Math. **85** (1986), 147–164. MR **87k:**58218
14. M. Reimann, *Ordinary differential equations and quasiconformal mappings*, Invent. Math. **33** (1976), 247–270. MR **53:**13556
15. D. Ruelle, *The thermodynamic formalism for expanding maps*, Comm. Math. Phys. **125** (1989), 239–262. MR **91a:**58149
16. D. Ruelle, *An extension of the theory of Fredholm determinants*, Inst. Hautes Etudes Sci. Publ. Math. **72** (1990), 175–193. MR **92b:**58187
17. D. Ruelle, *Dynamical zeta functions for piecewise monotone maps of the interval,* CRM Monograph Series, Vol. 4, Amer. Math. Soc., Providence, RI, 1994. CMP 94:12
18. H.H. Rugh, *Generalized Fredholm determinants and Selberg zeta functions for Axiom A dynamical systems*, Preprint (1994), to appear Ergodic Theory Dynamical Systems.

19. D. Sullivan, *Bounds, quadratic differentials and renormalization conjectures*, A.M.S. Centennial Publications, vol. 2, Mathematics into the Twenty-first Century (1988), 1992, pp. 417–466. MR **93k:**58194
20. A. Zygmund, *Smooth functions*, Duke Math. J. **12** (1945), 47–76. MR **7:**60b

ETH Zurich, CH-8092 Zurich, Switzerland (on leave from CNRS, UMR 128, ENS Lyon, France)
Current address: Mathématiques, Université de Genève, 1211 Geneva 24, Switzerland
E-mail address: baladi@sc2a.unige.ch

Department of Mathematics, Queens College, The City University of New York, Flushing, New York 11367-1597
E-mail address: yunqc@qcunix.acc.qc.edu

ETH Zurich, CH-8092 Zurich, Switzerland
E-mail address: lanford@math.ethz.ch

Statistical mechanical methods and continued fractions

By O. E. Lanford III and L. Ruedin

Mathematics Department, ETH-Zürich
ETH-Zentrum, 8092 Zürich, Switzerland

Abstract. For $a_1, \ldots, a_n$ a finite sequence of strictly positive integers, we denote by $q_n(a_1, \ldots, a_n)$ the denominator of the finite continued fraction $[a_1, \ldots, a_n]$ written as a quotient of two relatively prime integers. We show that the sequence of functions $\log q_n(a_1, \ldots, a_n)$, $n = 1, 2, \ldots$, have the formal properties of a Hamiltonian for a one-dimensional lattice system, to which the methods of statistical mechanics can be applied, and we investigate the properties of the system so defined.

1 Introduction.

We are going to discuss here a one-dimensional statistical-mechanical system constructed out of continued fractions. We will write continued fractions with the notation $[a_1, \ldots, a_n]$ instead of the typographically inconvenient

$$\cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots \cfrac{}{a_{n-1} + \cfrac{1}{a_n}}}}}$$

Formally, we can define this notation recursively:

$$[a_1] := \frac{1}{a_1}, \quad [a_1, a_2, \ldots, a_n] := \frac{1}{a_1 + \dfrac{1}{[a_2, \ldots, a_n]}}.$$

From this definition, it follows easily by induction that

$$[a_1, \ldots, a_k, \ldots a_n] = [a_1, \ldots, (a_k + [a_{k+1}, \ldots, a_n])] \quad \text{for any } k < n.$$

In general, the entries a_j can be elements of an arbitrary field (but it is then necessary to pay attention to the possibility of encountering a zero denominator.) In our application, the a_j will almost always be strictly positive integers; the only exception will be that it is occasionally convenient to let the last entry a_n be a real number ≥ 1. In these cases, there are no problems with zero denominators. Infinite continued fractions are defined as limits of finite ones: It is well known that, if all a_1, a_2, ... is a sequence of numbers all ≥ 1, then the sequence of finite ("truncated") partial fractions $[a_1, \ldots, a_n]$ converges as $n \to \infty$; we denote the limit by $[a_1, \ldots]$.

A finite continued fraction $[a_1, \ldots, a_n]$ with positive integer entries is a rational number between 0 and 1; we define $p_n(a_1, \ldots, a_n)$ and $q_n(a_1, \ldots, a_n)$ as the numerator and denominator of the reduced-form representation of this number:

$$[a_1, \ldots, a_n] =: \frac{p_n(a_1, \ldots, a_n)}{q_n(a_1, \ldots, a_n)},$$

with p_n, q_n relatively prime positive integers.

The starting point for the work reported here is the observation that *the sequence of functions*

$$H_n(a_1, \ldots, a_n) := \log q_n(a_1, \ldots, a_n)$$

can be taken to be the energy functions for a one-dimensional classical lattice system, with single-site state space $\mathbb{N}_+ := \{1, 2, \ldots\}$. In essence what this means is that this sequence of functions is *extensive* in the sense that

$$H_{n+m}(a_1, \ldots, a_{n+m}) \approx H_n(a_1, \ldots, a_n) + H_m(a_{n+1}, \ldots, a_{n+m}).$$

In the case at hand, the "$\approx$" sign in the above equation can be taken to mean that the difference in the two sides is bounded uniformly in n, m, and $a_1, \ldots, a_{n+m}$ (although a bit less would suffice for the construction of a statistical mechanical system.) In many respects, this Hamiltonian defines an extremely well-behaved statistical-mechanical system; notably, the interaction is exponentially decreasing. On the other hand, since the single-site state space is infinite, this system isn't quite covered by the standard theory, and does indeed turn out to display a few – inessential – pathologies.

The investigation of this system is the subject of the second author's Ph.D. dissertation (Ruedin, 1994). This dissertation contains results of two different kinds: On the one hand, extensions of many standard results to a a general framework adequate to cover the Hamiltonian $H_n = \log q_n$ as well as many others, and, on the other hand, proofs of a specific results for this Hamiltonian. We will concentrate here on results specific to this system, referring to Ruedin (1994) for the general theory.

This article is organized as follows: Section 2 reviews a few facts about continued fractions used in the remainder of the article, and the most basic properties of the specific Hamiltonian are established in Section 3. In Sections 4 and 5 respectively, we summarize the properties of the canonical

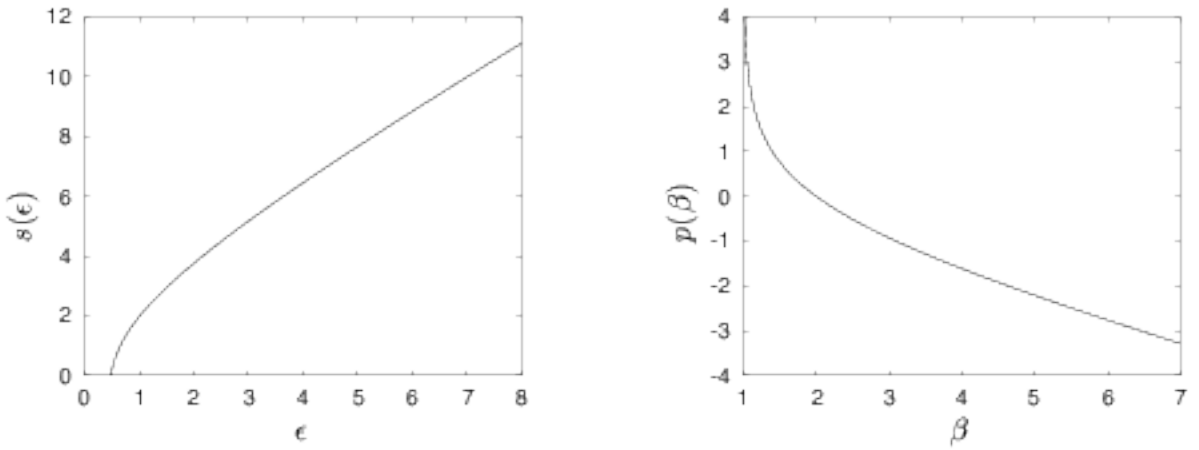


Figure 1: The thermodynamic functions

and microcanonical partition functions, and we introduce some ideas about "letting the size of the system fluctuate," which are natural for applications to continued fractions. The effect of letting the size of the system fluctuate is to fix the temperature at the particular value at which the pressure vanishes; in the present case this turns out to correspond to inverse temperature $\beta = 2$. By applying ideas about equivalence of ensembles, we show that, among rational numbers between 0 and 1 with reduced-form denominators not larger than some given large integer q, an overwhelming majority have continued fraction expansions $[a_1, \ldots, a_n]$ whose length n is approximately equal to $(\epsilon^*)^{-1} \log q$, for a constant ϵ^* (which we later show is equal to $\frac{\pi^2}{12 \log 2}$.)[1] In Section 6, we investigate the question of how thick the energy surface has to be made in order to get the microcanonical ensemble to function for our model, and in Section 7 we show that our system has no zero-point entropy, i.e., satisfies the third law of thermodynamics.

In Section 8, we introduce a second observable (in addition to H_n), the sequence

$$F_n(a_1, \ldots, a_n) = a_1 + \cdots + a_n,$$

and we investigate the joint distribution of H_n and F_n for large n. The quantity F_n has an interesting interpretation: it is the depth in the *Farey tree* enumeration of the rational numbers at which $[a_1, \ldots, a_n]$ occurs (see e.g. Kim and Ostlund (1989), §3). Ideas about equivalence of ensembles in statistical mechanics suggest that there should be a constant k_F such that most rational numbers $[a_1, \ldots, a_n]$ (n variable) with reduced form denominator $q_n \approx q$ have $F_n \approx k_F \log q$. One of the stimuli for this investigation was a considerable body of numerical evidence that this is *not* the case; in Section 8, we show that, in fact, typical values of $F_n / \log q_n$ go to ∞ as q_n does (i.e., loosely, that $k_F = \infty$.) In Section 9, we introduce an alternative representation for our system which is convenient for some kinds of computation, and we evaluate the constant ϵ^* referred to above. In Section 10, we state – without proof – the solution to the problem of maximizing the $q_n(a_1, \ldots, a_n)$ for fixed n and $a_1 + \cdots + a_n$, and we use this result to determine explicitly the "set of compatible values of H_n/n and F_n/n for large n," i.e., the set of points in the plane representable as limits of values of $(H_n/n, F_n/n)$ as $n \to \infty$.

For completeness, we show in Fig. 1 the microcanonical and canonical thermodynamic func-

[1] As we learned after this work was nearly completed, sharper results in this direction were proved more than twenty-five years ago by J. D. Dixon. See the discussion at the end of Section 5.

tions for our system. Aside from a few qualitative features – e.g., $p(\beta) \to \infty$ as $\beta \to 1$ – which will be explained in the course of the development, these graphs seem entirely unremarkable.

The work reported here certainly has some connection with classical ideas about "ergodic properties of the Gauss map," as presented, for example, in Cornfeld *et al.* (1982), §7.4. Exactly what the connection is remains something of a mystery for us; we do not see any strict mathematical implications in either direction. There is a more transparent connection with the work of D. Mayer (Mayer, 1990), who investigates an operator which turns out to be exactly the Ruelle transfer operator for our system and proves a number of striking results about its spectrum. Mayer, however, approaches the subject from a different point of view, and his results and ours seem to be more complementary than overlapping.

The first author thanks D. Ruelle for a number of helpful remarks in the course of this work and H. Epstein for many fruitful discussion.

2 Continued fractions.

It is a standard fact from the classical theory of continued fractions[2] that, if we write p_n/q_n as before for the reduced-form representation of the rational number $[a_1, \ldots, a_n]$, then the p_n and q_n both satisfy the same recursion relation, namely

$$p_n = a_n p_{n-1} + p_{n-2}; \quad q_n = a_n q_{n-1} + q_{n-2}.$$

Hence, given the a_n's, and given p_n (or q_n) for two successive values of n, we can determine all other p_n's (respectively q_n's) from the recursion relations. It is immediate that $p_1(a_1) = 1$ and $q_1(a_1) = a_1$. Although p_0 and q_0 are not defined by the above, it is easy to check that, if we set $p_0 = 0$ and $q_0 = 1$, the recursion relations give the correct p_2 and q_2 and hence all later ones as well. Thus, we can alternatively define the p_n's and q_n's by:

$$\begin{aligned} p_n &= a_n p_{n-1} + p_{n-2}; \quad p_0 = 0, \quad p_1 = 1, \\ q_n &= a_n q_{n-1} + q_{n-2}; \quad q_0 = 1, \quad q_1 = a_1. \end{aligned}$$

From these formulas it is clear that $q_n(a_1, \ldots, a_n)$ is a strictly increasing function of each of its arguments, and that $p_n(a_1, \ldots, a_n)$ is independent of a_1 but strictly increasing in all its other arguments.

The smallest value of $q_n(a_1, \ldots, a_n)$ is thus $q_n(1, \ldots 1)$, and these numbers are the *Fibonacci sequence*. To fix the normalization, we define the Fibonacci sequence F_n by:

$$F_n = F_{n-1} + F_{n-2} \quad \text{with } F_0 = F_1 = 1;$$

it then follows from the recurrences for p_n and q_n that

$$\begin{aligned} q_n(1, \ldots, 1) &= F_n, \quad \text{and also} \\ p_n(1, \ldots, 1) &= F_{n-1}. \end{aligned}$$

[2]Results quoted in this section without proof or explicit citation can be found in any of the standard classical texts, e.g., Hardy and Wright (1960), Chapter X

The Fibonacci numbers can be written explicitly via the *Binet formula*

$$F_n = \frac{\gamma^{n+1} + (-1)^n(1/\gamma)^{n+1}}{\sqrt{5}}, \quad \text{where } \gamma := \frac{\sqrt{5}+1}{2}, \text{ the } \textit{golden number.}$$

Since $\gamma > 1$, we get $F_n \approx \gamma^{n+1}/\sqrt{5}$ for large n and, in particular,

$$F_n > \frac{1}{2}\gamma^n \quad \text{for large enough } n.$$

We prove here a simple result which we will need later concerning exponential falloff of dependence of $[a_1, \ldots, a_n]$ on arguments "far to the right."

Proposition 2.1 *Let*

$$a_1, \ldots, a_n, a_{n+1}, \ldots a_m \quad \textit{and} \quad a_1, \ldots a_n, a'_{n+1}, \ldots a'_{m'}$$

be two sequences of strictly positive integers, both of length at least n, which agree through the nth place. Then

$$\left| \frac{[a_1, \ldots, a_m]}{[a_1, \ldots a'_{m'}]} - 1 \right| \leq \frac{1}{q_n p_n}.$$

The right-hand side of this inequality is majorized by $1/(F_n F_{n-1})$ and hence by $4\gamma^{-2n-1}$ for large enough n.

Proof. Let

$$x := \begin{cases} [a_{n+1}, \ldots a_m] & \text{for } m > n \\ 0 & \text{for } m = n \end{cases}$$

Then

$$[a_1, \ldots, a_n, \ldots a_m] = [a_1, \ldots, a_n + x],$$

and we can write $[a_1, \ldots, a'_{m'}]$ similarly. The proof of the recurrences for p_n and q_n shows that

$$[a_1, \ldots a_m] = \frac{p_n + x p_{n-1}}{q_n + x q_{n-1}},$$

where p_n and q_n denote $p_n(a_1, \ldots, a_n)$ and $q_n(a_1, \ldots, a_n)$ respectively. Hence

$$\begin{aligned} & \frac{[a_1, \ldots, a_m]}{[a_1, \ldots, a'_{m'}]} - 1 \\ = \quad & \frac{(p_n + x p_{n-1})(q_n + x' q_{n-1}) - (p_n + x' p_{n-1})(q_n + x q_{n-1})}{(q_n + x q_{n-1})(p_n + x' p_{n-1})} \\ = \quad & \frac{x p_{n-1} q_n + x' p_n q_{n-1} - x p_n q_{n-1} - x' p_{n-1} q_n}{(q_n + x q_{n-1})(p_n + x' p_{n-1})} \\ = \quad & \frac{(x' - x)(p_n q_{n-1} - p_{n-1} q_n)}{(q_n + x q_{n-1})(p_n + x' p_{n-1})} \end{aligned}$$

Now $|p_n q_{n-1} - p_{n-1} q_n| = 1$ and, since $0 \le x, x' \le 1$, $|x' - x| \le 1$, and $(q_n + x q_{n-1})(p_n + x' p_{n-1}) \ge q_n p_n$, so the desired estimate follows. □

It is a classical fact, and not difficult to prove, that $q_n(a_1, \ldots, a_n)$ is symmetric under reversal of its arguments, i.e., that

Proposition 2.2 $q_n(a_1, \ldots, a_n) = q_n(a_n, \ldots, a_1)$.

This equation follows easily from the *Euler bracket function* representation for q_n, also known as the *Euler-Minding formula*. See Roberts (1977), Ch. XIII, or Perron (1954), §3. None of our proofs actually depend on this fact; we mention it only to avoid having to justify some otherwise odd-looking choices for orders of arguments.

3 The statistical mechanical system.

We consider the sequence of functions

$$H_n(a_1, \ldots, a_n) = \log q_n(a_n, \ldots a_1) \quad \text{on } \mathbb{N}_+^n .$$

for $n = 1, 2, 3, \ldots$ (The reversal of the order of arguments on the right is inconsequential in view of Prop. 2.2.) The first thing to be seen is that this system of functions is "extensive," in the sense explained in the introduction. Once this has been established, we can interpret H_n as the "energy" of a lattice system with n sites occupied by identical molecules with countably infinite state space. We start by defining

$$\begin{aligned} h_n(a_1, \ldots, a_n) &= H_n(a_1, \ldots, a_n) - H_{n-1}(a_2, \ldots a_n) \\ &= \log \frac{q_n(a_n, \ldots, a_1)}{q_{n-1}(a_n, \ldots, a_2)} \end{aligned}$$

for $n > 1$ and

$$h_1(a_1) = \log(a_1)$$

Then

$$H_n(a_1, \ldots, a_n) = h_n(a_1, \ldots a_n) + h_{n-1}(a_2, \ldots, a_n) + \cdots + h_1(a_n).$$

We now have

Proposition 3.1

$$\frac{q_n(a_n, \ldots, a_1)}{q_{n-1}(a_n, \ldots, a_2)} = \frac{1}{[a_1, \ldots, a_n]},$$

(with $q_0 := 1$) and hence

$$h_n(a_1, \ldots, a_n) = -\log [a_1, \ldots, a_n].$$

Proof. By induction on n. The asserted formula is true for $n = 1$. The recursion for the q_j gives

$$q_n(a_n, \ldots, a_1) = a_1 q_{n-1}(a_n, \ldots a_2) + q_{n-2}(a_n, \ldots, a_3).$$

Dividing by q_{n-1} gives

$$\begin{aligned} \frac{q_n}{q_{n-1}} &= a_1 + \frac{q_{n-2}}{q_{n-1}} \\ &= a_1 + \frac{1}{[a_2, \ldots a_n]} \quad \text{by the induction hypothesis.} \\ &= \frac{1}{[a_1, \ldots a_n]} \quad \text{by the definition of continued fraction.} \end{aligned}$$

This proves the induction step and hence the formula. □

We now split the energy into an self-interaction part $H^{(0)}$ and a remainder $H^{(I)}$ by:

$$\begin{aligned} H_n^{(0)}(a_1, \ldots, a_n) &:= h_1(a_1) + h_1(a_2) + \cdots + h_1(a_n) \\ H_n^{(I)}(a_1, \ldots, a_n) &:= H_n(a_1, \ldots, a_n) - H_n^{(0)}(a_1, \ldots a_n) \\ &= h_n^{(\mathrm{I})}(a_1, \ldots a_n) + \cdots + h_2^{(\mathrm{I})}(a_{n-1}, a_n), \end{aligned}$$

where

$$\begin{aligned} h_j^{(\mathrm{I})}(a_1, \ldots a_j) &:= h_j(a_1, \ldots a_j) - h_1(a_1) \\ &= -\log(a_1 \cdot [a_1, \ldots, a_j]) \\ &= -\log(\frac{a_1}{a_1 + [a_2, \ldots, a_j]} \\ &= \log(1 + \frac{[a_2, \ldots, a_j]}{a_1}) \end{aligned}$$

for $j > 1$ and $h_1^{(\mathrm{I})}(a_1) = 0$.

Proposition 3.2 *We have*

$$0 \le h_n^{(\mathrm{I})}(a_1, \ldots, a_n) \le \log 2,$$

and there is a constant c such that, for all $n \ge 1$, and all pairs of sequences

$$a_1, \ldots, a_n, a_{n+1}, \ldots a_m \quad \textit{and} \quad a_1, \ldots a_n, a'_{n+1}, \ldots a'_{m'} \in \mathrm{N}_+,$$

both with length $\ge n$, and agreeing in the first n places, we have

$$|h_m^{(\mathrm{I})}(a_1, \ldots, a_m) - h_{m'}^{(\mathrm{I})}(a_1, \ldots, a'_{m'})| \le c\frac{1}{\gamma^{2n}}$$

Proof. The first estimate follows at once from the formula

$$h_m^{(\mathrm{I})}(a_1, \ldots, a_m) = \log(1 + \frac{[a_2, \ldots, a_m]}{a_1}).$$

The second assertion follows from Proposition 2.1 and the preceding formula, together with the observation that the derivative of the logarithm is ≤ 1 on the interval $[1, 2]$. □

Note that it follows that

$$H_n^{(0)} \leq H_n \leq H_n^{(0)} + n \log 2$$

for all n.

The proof that the sequence of functions H_n is extensive is now nearly immediate.

Proposition 3.3 *The difference*

$$H_{n+m}(a_1, \ldots, a_{n+m}) - H_n(a_1, \ldots, a_n) - H_m(a_{n+1}, \ldots, a_{n+m})$$

is bounded uniformly in n, m, a_1, ..., a_{n+m}.

Proof. Since $H_{n+m}^{(0)} = H_n^{(0)} + H_m^{(0)}$ – with the obvious arguments – we get

$$\begin{aligned}
&H_{n+m}(a_1, \ldots, a_{n+m}) - H_n(a_1, \ldots, a_n) - H_m(a_{n+1}, \ldots a_{n+m}) \\
&\quad = H_{n+m}^{(I)}(a_1, \ldots, a_{n+m}) - H_n^{(I)}(a_1, \ldots, a_n) - H_m^{(I)}(a_{n+1}, \ldots, a_{n+m}) \\
&\quad = (h_{n+m}^{(\mathrm{I})}(a_1, \ldots a_{m+n}) - h_n^{(\mathrm{I})}(a_1, \ldots, a_n)) \\
&\qquad + \cdots + (h_{m+2}^{(\mathrm{I})}(a_{n-1}, \ldots a_{n+m}) - h_2^{(\mathrm{I})}(a_{n-1}, a_n)) \\
&\qquad + h_{m+1}^{(\mathrm{I})}(a_n, \ldots a_{n+m}).
\end{aligned}$$

The modulus of the right-hand side is majorized by

$$\log 2 + \sum_{j=2}^{\infty} \frac{c}{\gamma^{2j}},$$

which is finite and independent of n, m, and the a_i. □

It is now also easy to give a potential from which the sequence of H_n's can be reconstructed: We put

$$\begin{aligned}
\Phi_{\{j\}}(a_j) &= \log a_j, \\
\Phi_{\{j,\ldots k\}}(a_j, \ldots, a_k) &= h_{k-j+1}^{(\mathrm{I})}(a_j, \ldots, a_k) - h_{k-j}^{(\mathrm{I})}(a_j, \ldots, a_{k-1}) \qquad \text{for } j < k,
\end{aligned}$$

and $\Phi_J \equiv 0$ for finite subsets J of $\mathbb{Z}$ other than intervals. It is then easy to check that

$$H_n(a_1, \ldots, a_n) = \sum_{J \subset \{1,\ldots,n\}} \Phi_J(a|_J),$$

and it follows from Proposition 3.2 that

$$\|\Phi_J\|_\infty = \mathcal{O}(\gamma^{-2\,\mathrm{diam}(J)}),$$

and hence, in particular, that the interaction is exponentially decreasing.

4 The canonical ensemble.

The first observation we need to make is that the canonical ensemble only makes sense for inverse temperature $\beta > 1$. This is already true for the finite system. The canonical partition function for a finite system with n (adjacent) lattice sites is

$$Z_n(\beta) = \sum_{a_1,\ldots a_n=1}^{\infty} \exp(-\beta H_n(a_1,\ldots,a_n)).$$

We denote – temporarily – the corresponding sum for $H_n^{(0)}$ by $Z_n^{(0)}$. From the inequalities

$$H_n^{(0)} \leq H_n \leq H_n^{(0)} + n \cdot \log 2,$$

it follows that

$$Z_n^{(0)}(\beta) \geq Z_n(\beta) \geq 2^{-n\beta} Z_n^{(0)}(\beta).$$

But

$$Z_n^{(0)}(\beta) = (\sum_{a=1}^{\infty} e^{-\beta \log a})^n = (\sum_{a=1}^{\infty} \frac{1}{a^\beta})^n = (\zeta(\beta))^n,$$

and $\zeta(\beta)$, the Riemann zeta function, goes to infinity as β decreases to 1. Hence, the same is true for the finite-system partition function for any n.

On the other hand, for $\beta > 1$, the finite-system partition function is finite for all n. Using the bound on $H_n - H_n^{(0)}$, it is easy to adapt the standard proof of the existence of the thermodynamic limit of the canonical partition function for lattice systems – which assumes that the system at each lattice site has only finitely many states, rather than countably many as in the case at hand – to show that

$$p(\beta) = \lim_{n\to\infty} \frac{1}{n} \log Z_n(\beta)$$

exists for any $\beta > 1$. The limiting function is convex on $(1, \infty)$, since this is true before passing to the limit. From the estimates proved above we get the bounds

$$\log \zeta(\beta) \geq p(\beta) \geq \log \zeta(\beta) - \beta \log 2.$$

It is a standard and simple fact that the zeta function has a simple pole with unit residue at 1; from this it follows that

$$p(\beta) = \log(\beta - 1) + \mathcal{O}(1) \quad \text{as } \beta \to 1.$$

Remark. The above lower bound can be improved – for β near its minimum value 1 – as follows: Since

$$h_j(a_j,\ldots,a_n) = \log(a_j + [a_{j+1},\ldots,a_{j_n}]) \leq \log(a_j + 1),$$

we get

$$Z_n(\beta) \geq (\sum_{a=2}^{\infty} \frac{1}{a^\beta})^n = (\zeta(\beta) - 1)^n,$$

and hence

$$p(\beta) \geq \log(\zeta(\beta) - 1) \approx \log \zeta(\beta) - \frac{1}{\zeta(\beta)}.$$

These estimates do not however tell us much about the behavior for $\beta \to \infty$; the upper bound goes to 0 and the lower bound is asymptotic to $-\beta \log 2$. We will get better information about this limiting regime later.

We will need here generalizations of a certain number of results which are standard for one-dimensional lattice systems with finite single-site state spaces to our model (which has $\mathbb{N}_+$ as single-site state space.) We referred above to one such result, the existence of the thermodynamic limit for the canonical partition function. The generalization for that particular result is easy, but we will require here generalizations of two other circles of ideas – Gibbs states and the transfer-operator formalism – which are not quite so straightforward. These extensions have been carried out in all detail in Ruedin (1994); we summarize the results here:

Proposition 4.1 *1. For each $\beta > 1$ there is a unique Gibbs state σ_β, (which is then necessarily translation-invariant,) and*

$$p(\beta) = s(\sigma_\beta) - \beta\, \overline{\epsilon}_\beta,$$

where $s(\sigma_\beta)$ is the Kolmogorov-Sinai entropy of σ_β and $\overline{\epsilon}_\beta$ the mean energy per lattice site of σ_β.

2. $p(\beta)$ is a real-analytic function of β on $(1, \infty)$ and is strictly convex in the strong sense that its second derivative is everywhere strictly positive.

5 The microcanonical entropy.

Once again we need an extension of some standard results to our slightly-nonstandard technical situation. The standard results can be found in Lanford (1973); an extension adequate to the present situation is given in Ruedin (1994). To formulate the result we need, we use the following notation: Let $-\infty \leq \epsilon_1 < \epsilon_2 \leq \infty$; then $\mathcal{V}_n(\epsilon_1, \epsilon_2)$ will denote the number of sequences $a_1, \ldots, a_n$ of length n with

$$\epsilon_1 < \frac{1}{n} H_n(a_1, \ldots, a_n) < \epsilon_2$$

Proposition 5.1 *There is a non-negative concave function $s(\epsilon)$, defined on an open interval $(\epsilon_{\min}, \epsilon_{\max})$, (where $\epsilon_{\min}$ may be $-\infty$ and $\epsilon_{\max}$ may be $+\infty$) such that*

1. $\lim_{n\to\infty} \frac{1}{n} \log \mathcal{V}_n(\epsilon_1, \epsilon_2) = \sup_{\epsilon_1 < \epsilon < \epsilon_2} s(\epsilon)$ *for all intervals (ϵ_1, ϵ_2) intersecting $(\epsilon_{\min}, \epsilon_{\max})$*
2. $\mathcal{V}_n(\epsilon_1, \epsilon_2) = 0$ *for all sufficiently large n for intervals (ϵ_1, ϵ_2) whose closures do not intersect the closure of $(\epsilon_{\min}, \epsilon_{\max})$.*

The formulation of the preceding proposition is a little dense, and it may be helpful to elaborate on it a bit. For purposes of this explanation, let us say that ϵ is an *asymptotically excluded value for* H_n/n if there exists a neighborhood of ϵ which is disjoint from the image of H_n/n for all sufficiently large n and an *asymptotically allowed value* otherwise. The set of asymptotically excluded values is manifestly open. The first non-trivial assertion of the proposition is that the complementary set of asymptotically allowed values is an interval; its interior is the interval $(\epsilon_{\min}, \epsilon_{\max})$ of the proposition. We will accordingly – if not quite precisely – refer to $(\epsilon_{\min}, \epsilon_{\max})$ as the *allowed interval*. The idea is then that, for any "sampling interval" (ϵ_1, ϵ_2), the number $\mathcal{V}_n(\epsilon_1, \epsilon_2)$ of configurations with $H_n/n \in (\epsilon_1, \epsilon_2)$ should be asymptotically – for large n – about $\exp(n s_{\epsilon_1,\epsilon_2})$, with a particular form for the dependence of the exponent $s_{\epsilon_1,\epsilon_2}$ on the sampling interval. Part 1. of the proposition says that this behavior does hold for sampling intervals which overlap the allowed interval, and part 2. says that a natural – and rather strong – variant holds for sampling intervals which stay away from the allowed interval. It turns out, however, that the asserted exponential behavior can fail – or at least is much more delicate to prove – if the sampling interval touches but does not overlap the allowed interval, i.e., if $\epsilon_2 = \epsilon_{\min}$ or $\epsilon_1 = \epsilon_{\max}$; the proposition says nothing in these cases. We remark that the proof of this statement depends not just on the "extensivity" of the sequence of functions H_n in the sense described above; it is also necessary that they "grow at infinity" in an appropriate way so as to guarantee, in particular, that $\mathcal{V}_n(\epsilon_1, \epsilon_2)$ is finite for all finite n, ϵ_1, and ϵ_2. An appropriate general formulation of the growth at infinity condition is given in Ruedin (1994); we note here only that, in the case at hand, adequate growth at infinity is guaranteed by the fact that $H_n \geq H_n^{[0]}$ and that $h_1(a)$ grows adequately fast as $a \to \infty$.

The above proposition is a general result, using only qualitative properties of H_n. In the case at hand, we can be more specific.

Proposition 5.2 *For $H_n(a_1, \ldots, a_n) = \log q_n(a_1, \ldots, a_n)$, we have:*

- $\epsilon_{\min} = \log \gamma$ *(with, as above,* $\gamma = \dfrac{\sqrt{5}+1}{2}$*),*
- $\epsilon_{\max} = \infty$,
- $s(\epsilon)$ *is strictly increasing on* $(\log \gamma, \infty)$, *and* $s(\epsilon) \to \infty$ *as* $\epsilon \to \infty$.

Proof. We have already noted that

$$q_n(a_1, \ldots, a_n) \geq q_n(1, \ldots, 1) = F_n \approx \text{const}\, \gamma^n.$$

Hence,

$$\mathcal{V}_n(-\infty, \epsilon_2) = 0 \quad \text{for large } n, \text{ if } \epsilon_2 < \log \gamma,$$

and

$$\mathcal{V}_n(-\infty, \epsilon_2) \geq 1 \quad \text{for large } n, \text{ if } \epsilon_2 > \log \gamma.$$

From these two assertions it follows that $\epsilon_{\min} = \log \gamma$.

We now claim that

$$\sup_{\epsilon < \epsilon_1} s(\epsilon) \to \infty \quad \text{as } \epsilon_1 \to \infty.$$

From this claim, it follows that $s(\epsilon)$ is not bounded; hence, since it is concave, that it it is strictly increasing on its whole interval of definition and goes to ∞ with ϵ, and these are the remaining assertions of the proposition

To prove the claim, we begin by considering the sequence $q_n(p, \ldots, p)$ for general $p \in \mathbb{N}_+$. By the recursion relation

$$q_n(p, \ldots, p) = p\, q_{n-1}(p, \ldots, p) + q_{n-2}(p, \ldots, p).$$

It follows from simple standard arguments – generalizing the derivation of the Binet formula for the Fibonacci numbers, see also §10 – that

$$q_n(p, \ldots, p) \approx \text{const} \cdot \gamma_p^n,$$

where γ_p is the positive root of the quadratic equation $t^2 - pt - 1 = 0$, i.e.,

$$\gamma_p = \frac{1}{2}(p + \sqrt{p^2 + 4}) (\approx p \quad \text{for } p \text{ large.})$$

Thus, if $\epsilon_1 > \log \gamma_p$ and n is large enough,

$$\frac{1}{n} \log q_n(p, \ldots, p) \leq \epsilon_1,$$

and hence the same inequality holds for $q_n(a_1, \ldots, a_n)$ provided that all the a_i are $\leq p$. Thus:

$$\mathcal{V}_n(-\infty, \epsilon_1) \geq p^n \quad \text{for } \epsilon_1 > \log \gamma_p \text{ and } n \text{ sufficiently large.}$$

Taking logarithms, dividing by n, and letting $n \to \infty$ gives

$$\sup_{\epsilon \leq \epsilon_1} s(\epsilon) \geq \log p \quad \text{for } \epsilon_1 > \log \gamma_p;$$

letting ϵ_1 decrease to $\log \gamma_p$ gives

$$\sup_{\epsilon \leq \log \gamma_p} s(\epsilon) \geq \log p.$$

By letting p go to ∞ we see that $s(\epsilon)$ is unbounded, as asserted, and this completes the proof of the proposition. We can now in fact say a little more about the behavior of $s(\epsilon)$ as $\epsilon \to \infty$. Now that we know that $s(\epsilon)$ is increasing, we can simplify the above lower bound to

$$s(\log \gamma_p) \geq \log p,$$

On the other hand, $\gamma_p / p \to 1$ as $p \to \infty$, so $s(\epsilon)$ in fact grows at least as fast as ϵ as $\epsilon \to \infty$. □

The next step is to argue that $p(\beta)$ is the Legendre transform of $s(\epsilon)$ and to deduce analyticity and strict concavity for $s(\epsilon)$ from analyticity and strict convexity for $p(\beta)$.

Proposition 5.3 *$s(\epsilon)$ is real-analytic, strictly increasing, and strictly convex on $(\epsilon_{\min}, \infty)$. The function $\beta(\epsilon) = -s'(\epsilon)$ maps $(\epsilon_{\min}, \infty)$ diffeomorphically onto $(1, \infty)$; its inverse is $\epsilon(\beta) = p'(\beta)$. For every β between 1 and ∞,*

$$p(\beta) = \sup_{\epsilon}(s(\epsilon) - \beta\epsilon);$$

the supremum is taken on at $\epsilon = \epsilon(\beta)$, and nowhere else.

Proof. The argument is standard, but we give it in detail anyway, since there are a few places where special features of the situation at hand have to be invoked to rule out pathologies. We begin from the fact that, since $s(\epsilon)$ is concave, it is differentiable except at most at a countable set of points and its derivative – where defined – is monotone decreasing. We define temporarily

$$\beta_{\min} := \lim_{\epsilon \to \infty} s'(\epsilon) \quad \text{and} \quad \beta_{\max} := \lim_{\epsilon \to \epsilon_{\min}^+} s'(\epsilon),$$

with the understanding that the limits are to be taken along the set where the derivative exists. Nothing said so far rules out the possibility that $\beta_{\min} = \beta_{\max}$. Nevertheless, the microcanonical analysis leads to

Proposition 5.4 *1. If $\beta_{\min} < \beta < \beta_{\max}$, then*

$$p(\beta) = \sup_{\epsilon}(s(\epsilon) - \beta\epsilon),$$

and the supremum is taken on.

2. If $\beta < \beta_{\min}$, then

$$\lim_{n \to \infty} \frac{1}{n} \log Z_n(\beta) = +\infty.$$

3. If $\beta > \beta_{\max}$, then

$$p(\beta) = s_0 - \beta\epsilon_{\min},$$

where s_0 denotes $\lim_{\epsilon \to \epsilon_{\min}^+} s(\epsilon)$.[3]

Again, we refer to Ruedin (1994) for the proof. The argument is essentially the standard one, but a little extra effort is needed to work around the fact that $\mathcal{V}_n(\epsilon, \infty)$ is infinite.

It follows from 1 and 2, together with what we know about $p(\beta)$, that $\beta_{\min} = 1$. If $\beta_{\max}$ were finite, $p(\beta)$ would have to be linear from $\beta_{\max}$ to ∞, and this violates the strict convexity of $P(\beta)$; hence, $\beta_{\max}$ must be infinite. Thus, the Legendre transformation formula

$$p(\beta) = \sup_{\epsilon}(s(\epsilon) - \beta\epsilon)$$

holds for $1 < \beta < \infty$, with the supremum taken on. Furthermore, except for at most a countable set of β's, the supremum is taken on at a single point. We denote this point by $\epsilon(\beta)$; if the supremum is taken on at more than one point – i.e., on an interval of non-zero length – then $\epsilon(\beta)$ is not defined.

For all relevant ϵ and β, we have

$$p(\beta) + \epsilon\beta \leq s(\epsilon),$$

with equality for $\epsilon = \epsilon(\beta)$. Out of this we can read the following: Let β_0 be such that $\epsilon(\beta_0)$ is defined; this excludes at most countably many values. Put $\epsilon_0 = \epsilon(\beta_0)$. Then $\beta \mapsto p(\beta) + \beta\epsilon_0$

[3] We are in fact going to show in §7 that $s_0 = 0$, but we don't need this fact for the moment.

takes on its maximum at $\beta = \beta_0$, which implies $p'(\beta_0) = -\epsilon_0$. In other words: $\epsilon(\beta) = -p'(\beta)$ whenever $\epsilon(\beta)$ is defined. But the only way $\epsilon(\beta)$ can fail to be defined is for the graph of $s(\epsilon)$ to contain a linear segment with slope β, and this implies that $\epsilon(\beta)$ has a jump discontinuity there. This, however, is ruled out by the fact that $p'(\beta)$ is real-analytic. The conclusion is that $\epsilon(\beta)$ is defined for all $\beta \in (1, \infty)$, and that $\epsilon(\beta) = -p'(\beta)$ for all these values of β.

Substituting into an earlier formula, we thus get the parametric representation

$$s(-p'(\beta)) = p(\beta) - \beta p'(\beta),$$

which, together with the analyticity and strict convexity of $p(\beta)$, ensures that $s(\epsilon)$ is real-analytic on the image of the mapping $\beta \mapsto -p'(\beta)$. By continuity, this image is an interval. Since $p(\beta) \to \infty$ as $\beta \to 1^+$, the same must be true of $-p'(\beta)$, i.e., the image interval must extend to ∞. On the other hand, the fact that $\beta_{\max} = \infty$ means that $s'(\epsilon)$ goes to ∞ as ϵ approaches $\epsilon_{\min}$ and hence implies that there exists a sequence ϵ_n converging to $\epsilon_{\min}$ such that $s'(\epsilon_n)$ exists for all n. Denoting $s'(\epsilon_n)$ by β_n, we get that $s(\epsilon) - \beta_n \epsilon$ takes on its supremum at ϵ_n, i.e., that $\epsilon_n = \epsilon(\beta_n) = -p'(\beta_n)$. Hence, the image interval also extends to $\epsilon_{\min}$, so the above formula represents $s(\epsilon)$ over its full range of definition. Thus, $s(\epsilon)$ is real-analytic where defined, and differentiation of the formula gives $s''(\epsilon) < 0$ everywhere.

We have just argued that $\beta \mapsto \epsilon(\beta) = -p'(\beta)$ sends $(1, \infty)$ diffeomorphically onto $(\log \gamma, \infty)$. We denote the inverse mapping by $\beta(\epsilon)$; a standard calculation shows that $\beta(\epsilon) = s'(\epsilon)$. We then have:

$$p(\beta) = s(\epsilon(\beta)) + \beta \cdot \epsilon(\beta) \qquad \text{for all } \beta \in (1, \infty).$$

This completes the proof of Prop. 5.3 □

Everything said so far has used only qualitative properties of $H_n = \log q_n$. We will now make a first contact with "number theory." The argument is the reverse of what we ultimately want to do – we will use some classical facts from number theory to prove something about $s(\epsilon)$. The argument is also illuminating as a simple example of how to compute more concrete quantities in terms of $s(\epsilon)$. The question we want to address is:

Question. *How does the total number $N(q)$ of sequences $a_1, \ldots, a_n$ with*

$$q_n(a_1, \ldots, a_n) < q$$

(n variable) behave as $q \to \infty$?

Determining $N(q)$ is almost the same as counting rational numbers between 0 and 1 with reduced-form denominator $< q$. There is in fact exactly a factor of 2 difference between the two question: A rational number has *exactly two* continued fraction representations:

- the "standard" one – given by the Euclidean algorithm – which has the form $[a_1, \ldots, a_n]$ with $a_n \geq 2$, and
- a second one $[a_1, \ldots, a_n - 1, 1]$

(e.g., $1/2 = [2] = [1,1]$.)Furthermore, the number of rational numbers between 0 and 1 with reduced-form denominator q is exactly $\varphi(q)$, the Euler φ-function. Thus, we have the exact formula

$$N(q) = 2\sum_{j=2}^{q-1} \varphi(j).$$

By classical number theory (e.g. Hardy and Wright, 1960, Theorem 330) this sum is asymptotically a constant multiple of q^2. We now proceed to compute the asymptotic behavior of $N(q)$ in terms of the function $s(\epsilon)$; comparing that answer with the one just obtained will tell us something about $s(\epsilon)$.

We start from the fact that the number of sequences *of length* n with $\log q_n < n\epsilon$ is, by definition, $\mathcal{V}_n(-\infty, \epsilon)$. Unraveling the notation: The number of sequences of length n with $q_n < q$ is $\mathcal{V}_n(-\infty, (\log q)/n)$. Thus, the total number of sequences – of arbitrary n – is

$$N(q) = \sum_{n=1}^{\infty} \mathcal{V}_n(-\infty, \frac{\log q}{n}).$$

Although we have written the sum as running to ∞, there are in fact only finitely many non-zero terms for given q: $q_n(a_1, \ldots, a_n) \geq F_n$, so

$$\mathcal{V}_n(-\infty, \frac{\log q}{n}) = 0 \quad \text{for } F_n > q,$$

i.e., for

$$\frac{1}{n}\log F_n > \frac{1}{n}\log q,$$

i.e., for

$$n \geq \log q \frac{n}{\log F_n} \approx \frac{\log q}{\log \gamma}.$$

In particular: The number of terms in the above sum is $\mathcal{O}(\log q)$ for large q. From this, we want to argue that *for our purposes, it is adequate to approximate the above sum by its largest term.* The justification for this assertion runs as follows: For given q, let $n(q)$ be such as to make

$$\mathcal{V}_{n(q)}(-\infty, \frac{\log q}{n(q)})$$

as large as possible, and put

$$y(q) := \frac{n(q)}{\log q}.$$

Then $y(q)$ is certainly not much larger than $1/\log\gamma$ for large q. We will argue later that $y(q)$ converges to a finite non-zero limit as $q \to \infty$; for the moment, we accept this without proof to see how the rest of the argument goes. The largest term in the sum is then

$$\mathcal{V}_{n(q)}(-\infty, \frac{1}{y(q)}) \sim \exp(n(q)s(1/y(q))) = \exp(t(q)\log q)$$

where $t(q)$ denotes $y(q)s(1/y(q))$. Taking logarithms and dividing by $\log q$ gives a sequence which has a chance of remaining of order unity as $q \to \infty$. We can now justify the claim that the largest term is an adequate approximation to the sum: We have

$$\frac{1}{\log q} \log \mathcal{V}_{n(q)}(-\infty, \frac{1}{y(q)}) \leq \frac{1}{\log q} \log\Big(\sum_{n=1}^{\infty} \mathcal{V}_n(-\infty, \frac{\log q}{n})\Big)$$
$$\leq \frac{1}{\log q} \log \mathcal{V}_{n(q)}(-\infty, \frac{1}{y(q)}) + \mathcal{O}(\frac{\log\log q}{\log q}),$$

so the sequence built from the sum and the one built from its largest term do have the same limiting behavior in the sense that their difference goes to zero.

We now make a heuristic argument, intended as motivation for a subsequent precise result: Assuming that $y(q)$ converges to a limit y^*, and assuming also the validity of an obvious exchange of limits, we would expect that

$$\lim_{q\to\infty} \frac{1}{\log q} \log N(q) = \lim_{q\to\infty} \frac{1}{\log q} \log \mathcal{V}_{n(q)}(-\infty, \frac{1}{y(q)}) = y^* s(1/y^*).$$

Furthermore, $n(q)$ was chosen to make the corresponding term in the sum as large as possible, and y^* is the limit of the $n(q)/\log q$'s, so it should be at least plausible that

$$y^* s(1/y^*) = \sup_y y\, s(1/y) = \sup_\epsilon s(\epsilon)/\epsilon.$$

On the other hand, we showed above that

$$N(q) \approx \text{const} \cdot q^2 \quad \text{so} \quad \lim_{q\to\infty} \frac{1}{\log q} \log N(q) = 2,$$

so, finally, we expect that

$$y^* s(1/y^*) = 2, \quad \text{i.e.,} \quad \sup_\epsilon \frac{s(\epsilon)}{\epsilon} = 2.$$

With this as introduction and motivation, we formulate the following result:

Proposition 5.5 *The function $p(\beta)$ has a unique zero, which we denote by β^*. The function $s(\epsilon)/\epsilon$ takes on its supremum at $\epsilon = \epsilon^* := \epsilon(\beta^*)$. This is the only place where the supremum is taken on, and the function is strictly increasing to the left of ϵ^* and strictly decreasing to the right. We have, furthermore,*

$$2 = \beta^* = \sup_\epsilon \frac{s(\epsilon)}{\epsilon} = \lim_{q\to\infty} \frac{1}{\log q} \log N(q),$$

where $N(q)$ denotes as above the number of sequences $a_1, \ldots, a_n$ (n variable) with

$$q_n(a_1, \ldots, a_n) < q.$$

Proof. The logic is:

- We investigate first the problem of maximizing $s(\epsilon)/\epsilon$. We show that on the one hand the supremum is taken on exactly at $\epsilon(\beta^*)$ and on the other hand the supremum is also *equal to* β^*.
- We then fill in the gaps in the earlier heuristic analysis to show that

$$\lim_{q\to\infty} \frac{1}{\log q} \log N(q) = \sup_{\epsilon} \frac{s(\epsilon)}{\epsilon}.$$

- Comparing with the formula for $N(q)$ in terms of the Euler φ-function, we find

$$\lim_{q\to\infty} \frac{1}{\log q} \log N(q) = 2,$$

  which completes the proof.

We will prove later – by quite different methods – the explicit formula

$$\epsilon^* = \frac{\pi^2}{12 \log 2}.$$

That $s(\epsilon)/\epsilon$ takes on its supremum at ϵ^*, and that the value of the supremum is β^*, can be motivated by putting the derivative of $s(\epsilon)/\epsilon$ equal to zero. For a proof, it is more convenient to proceed less directly.[4] From the fact that $p'(\beta) = -\epsilon(\beta) < -\log\gamma < 0$, it follows that $p(\beta) \to -\infty$ for $\beta \to \infty$. We know, on the other hand, that $p(\beta) \to +\infty$ for $\beta \to 1$. Hence, there is a β^* such that

$$p(\beta^*) = 0,$$

and since $p'(\beta) < 0$ everywhere, this β^* is unique. From the Legendre transform

$$0 = p(\beta^*) = \sup_{\epsilon}(s(\epsilon) - \beta^*\epsilon),$$

and the supremum is taken on exactly for $\epsilon = \epsilon^*$. In other words:

$$s(\epsilon) \le \beta^*\epsilon, \quad \text{with equality if and only if } \epsilon = \epsilon^*.$$

Since all relevant ϵ's are $> \epsilon_{\min} = \log\gamma > 0$, we can divide by ϵ to get

$$\frac{s(\epsilon)}{\epsilon} \le \beta^*, \quad \text{with equality if and only if } \epsilon = \epsilon^*,$$

which is the desired assertion about where the supremum is taken on and what its value is. To show that $s(\epsilon)/\epsilon$ is strictly decreasing with increasing separation from ϵ^*, we use the general fact that (strict) concavity of $s(\epsilon)$ on $(\epsilon_{\min}, \infty)$ implies (strict) concavity of

$$g(y) := ys(\frac{1}{y}) \quad \text{on } (0, 1/\epsilon_{\min}).$$

[4] The argument we are about to give is standard in the application of statistical mechanics to dynamical systems.

This is true without smoothness assumptions, but can be proved particularly easily in the smooth case by verifying that

$$g''(y) = \frac{s''(1/y)}{y^3}.$$

Since $s(\epsilon)/\epsilon$ takes on its supremum at an interior point of its interval of definition, the same is true for $g(y)$; since $g(y)$ is concave, it is strictly monotone decreasing with increasing distance from the place where it takes on its supremum, so the same is true for $s(\epsilon)/\epsilon$.

This completes our analysis of the behavior of $s(\epsilon)/\epsilon$; we turn now to the behavior of $N(q)$ for large q. We have already given an outline of the argument; what remains to be shown is

$$\frac{1}{\log q} \log \mathcal{V}_{n(q)}(-\infty, \frac{\log q}{n(q)}) \longrightarrow \frac{s(\epsilon^*)}{\epsilon^*},$$

(where, as before, $n(q)$ denotes a value of n maximizing $\mathcal{V}_n(-\infty, \log q/n)$.) The first step in proving this is:

Lemma 5.6 *Let $m(q)$ be a sequence of integers such that*

$$\frac{\log q}{m(q)} \longrightarrow \bar{\epsilon} \in (\epsilon_{\min}, \infty)$$

Then

$$\frac{1}{\log q} \log \mathcal{V}_{m(q)}(-\infty, \frac{\log q}{m(q)}) \longrightarrow \frac{s(\bar{\epsilon})}{\bar{\epsilon}}.$$

Proof. Let $\epsilon_1 < \bar{\epsilon}$. Then, for sufficiently large q, $\log q/m(q) > \epsilon_1$, so

$$\mathcal{V}_{m(q)}(-\infty, \frac{\log q}{m(q)}) \geq \mathcal{V}_{m(q)}(-\infty, \epsilon_1),$$

again for sufficiently large q. Taking logarithms and dividing by $\log q$ gives

$$\frac{1}{\log q} \log \mathcal{V}_{m(q)}(-\infty, \frac{\log q}{m(q)}) \geq \frac{m(q)}{\log q} \frac{1}{m(q)} \log \mathcal{V}_{m(q)}(-\infty, \epsilon_1) \longrightarrow \frac{s(\epsilon_1)}{\bar{\epsilon}}.$$

Hence,

$$\liminf_{q \to \infty} \frac{1}{\log q} \log \mathcal{V}_{m(q)}(-\infty, \frac{\log q}{m(q)}) \geq \frac{s(\epsilon_1)}{\bar{\epsilon}}.$$

This holds for all $\epsilon_1 < \bar{\epsilon}$, and $s(\epsilon)$ is continuous, so we can replace ϵ_1 on the right by $\bar{\epsilon}$. In exactly the same way – starting with $\epsilon_1 > \bar{\epsilon}$ – we show that

$$\limsup_{q \to \infty} \frac{1}{\log q} \log \mathcal{V}_{m(q)}(-\infty, \frac{\log q}{m(q)}) \leq \frac{s(\bar{\epsilon})}{\bar{\epsilon}},$$

and the lemma follows. It is clear that this argument also works, with the obvious modification in the formulation, if $m(q)$ is only defined for a subsequence of q's going to ∞. □

As a consequence of the lemma, we note that the sequence $(\log q)/n(q)$ – with $n(q)$ defined as above – cannot have any accumulation point in $(\epsilon_{\min}, \infty)$ other than ϵ^*. Otherwise, letting $\hat{n}(q)$ be a sequence such that $(\log q)/\hat{n}(q) \to \epsilon^*$, we would eventually encounter a q for which

$$\mathcal{V}_{\hat{n}(q)}(-\infty, \frac{\log q}{\hat{n}(q)}) > \mathcal{V}_{n(q)}(-\infty, \frac{\log q}{n(q)}),$$

contradicting the assumed maximizing property of $n(q)$.

By the same sort of argument as used in the proof of the preceding lemma, we see that if

$$\limsup_{q\to\infty} \frac{\log q}{m(q)} \leq \epsilon_{\min},$$

then

$$\limsup_{q\to\infty} \frac{1}{\log q} \log \mathcal{V}_{m(q)}(-\infty, \frac{\log q}{m(q)}) \leq \lim_{\epsilon\to\epsilon^+_{\min}} \frac{s(\epsilon)}{\epsilon} < \sup_{\epsilon} \frac{s(\epsilon)}{\epsilon},$$

Thus, it is also impossible that $(\log q)/n(q)$ have an accumulation point in $[0, \epsilon_{\min}]$, so the only remaining possible accumulation points for the sequence $\log q/n(q)$ are ϵ^* and $+\infty$. If we eliminate the second possibility, it will follow that $\log q/n(q) \to \epsilon^*$ and hence, applying again the lemma, that

$$\frac{1}{\log q} \log \mathcal{V}_{n(q)}(-\infty, \frac{\log q}{n(q)}) \longrightarrow \frac{s(\epsilon^*)}{\epsilon^*},$$

as asserted.

Lemma 5.7 *Let $m(q)$ be a sequence such that*

$$\frac{m(q)}{\log q} \longrightarrow 0.$$

Then

$$\limsup \frac{1}{\log q} \log \mathcal{V}_{m(q)}(-\infty, \frac{\log q}{m(q)}) \leq 1$$

Proof. Fix $\beta > 1$, and let B denote an upper bound for the $Z_m(\beta)^{1/m}$. For any m and any real number r, we get

$$\begin{aligned} B^m &\geq Z_m(\beta) = \sum_{a_1,\ldots,a_m} \exp(-\beta H_m(a_1, \ldots, a_m)) \\ &\geq \sum \{\exp(-\beta H_m(a_1, \ldots, a_n)) : H_m(a_1, \ldots, a_m) < rm\} \\ &\geq \exp(-\beta rm)\mathcal{V}_m(-\infty, r), \end{aligned}$$

i.e.,

$$\mathcal{V}_m(-\infty, r) \leq \exp(\beta rm) B^m.$$

Thus,

$$\frac{1}{\log q} \log \mathcal{V}_{m(q)}(-\infty, \frac{\log q}{m(q)}) \leq \frac{m(q)}{\log q} \log B + \beta.$$

The first term on the right drops out as $q \to \infty$, so we get

$$\limsup_{q\to\infty} \frac{1}{\log q} \log \mathcal{V}_{m(q)}(-\infty, \frac{\log q}{m(q)}) \leq \beta.$$

Since this holds for all $\beta > 1$, the lemma – and hence also the proposition – is proved.

We can expand on the above argument to establish a statistical relation between n and q. Let $\epsilon_{\min} < \epsilon_1 < \epsilon^*$ and let $N^{(<)}(q, \epsilon_1)$ denote the number of n-tuples $a_1, \ldots, a_n$ (n variable) with

$$q_n(a_1, \ldots, a_n) \leq q \quad \text{and} \quad \frac{\log q_n(a_1, \ldots, a_n)}{n} < \epsilon_1,$$

i.e., with

$$q_n(a_1, \ldots, a_n) \leq q \quad \text{and} \quad n > (\epsilon_1)^{-1} \log q_n(a_1, \ldots, a_n).$$

The proof of Proposition 5.5 can easily be generalized to show that

$$\lim_{q\to\infty} \frac{1}{\log q} \log N^{(<)}(q, \epsilon_1) = \sup_{\epsilon<\epsilon_1} \frac{s(\epsilon)}{\epsilon} < \frac{s(\epsilon^*)}{\epsilon^*}.$$

Hence, in particular,

$$\lim_{q\to\infty} \frac{N^{(<)}(q, \epsilon_1)}{N(q)} = 0.$$

The ratio $N^{(<)}(q, \epsilon_1)/N(q)$ is the fraction of sequences $a_1, \ldots, a_n$ with $q_n < q$ which satisfy the further condition that $n > (\epsilon_1)^{-1} \log q_n$. Thus, we can say that, for q large, the overwhelming majority of sequences with $q_n < q$ have $n \leq (\epsilon_1)^{-1} \log q_n$, and this holds *for all* $\epsilon_1 < \epsilon^*$.

In exactly the same way, we argue that, for all $\epsilon_2 > \epsilon^*$, the fraction of these sequences with $n < (\epsilon_2)^{-1} \log q_n$ also goes to zero as $q \to \infty$. Hence:

Proposition 5.8 *Let $\epsilon_1 < \epsilon^* < \epsilon_2$. Then for q sufficiently large, an overwhelming majority of configurations with $q_n < q$ satisfy*

$$(\epsilon_2)^{-1} \log q_n \leq n \leq (\epsilon_1)^{-1} \log q_n.$$

Loosely formulated: For q large, nearly all configurations with $q_n < q$ satisfy $n \approx (\epsilon^*)^{-1} \log q$. We have already observed that the number of configurations with $q_n < q$ is twice the number of rational numbers between 0 and 1 with reduced-form denominator $< q$; each rational number has exactly two continued-fraction representations. The two representations of such a number have lengths differing by one, which is unimportant at the resolution at which we are working. Thus, we can reformulate what we have shown to say: *For q large, nearly all rational numbers with reduced-form denominator $< q$ have continued fraction expansions of length $\approx (\epsilon^*)^{-1} \log q$.*

After we had obtained the result formulated in the preceding paragraph, we learned that a sharper assertion had been proved in Dixon (1970). (See also Knuth (1981), §4.5.3, for a readable survey of work in this direction.) Part of what Dixon proves can be formulated as follows: For any $\epsilon > 0$, and

for q large, the overwhelming majority of rational numbers r between 0 and 1 with reduced-form denominator $q(r) \leq q$, have continued fraction expansion with length $n(r)$ satisfying

$$|n(r) - \lambda \log q(r)| \leq (\log q(r))^{1/2+\epsilon},$$

where λ denotes $\frac{12 \log 2}{\pi^2}$. Loosely: For typical rational numbers r, with large $q(r)$, $n(r)$ differs from $\lambda \log q(r)$ by something not much larger that $(\log q(r))^{1/2}$, whereas our results show only that the difference is typically $o(\log q(r))$. Dixon also gives an estimate for the number of configurations which do *not* satisfy the asserted inequality. Although Dixon's proof is based on detailed estimates in the spirit of analytic number theory, rather than the general statistical mechanical ideas we have used, there are many points of resemblance between his argument and ours.

6 Full entropy.

We are now going to explore the possibility of improving, e.g., Proposition 5.8 by replacing the phrase "the overwhelming majority of configurations with $q_n < q$" by "the overwhelming majority of configurations with $q_n \approx q$." This is a version of a standard problem in statistical mechanics: How thick must the energy shell be to get the microcanonical ensemble to work properly? We can formulate the question somewhat more precisely as follows: Suppose we choose, for each q, a quantity $\delta(q)$ between zero and one, and we let $\widetilde{N}(q)$ denote the number of configurations with $q - \delta(q) \cdot q < q_n < q$. If we can show that

$$\lim_{q \to \infty} \frac{1}{\log q} \log \widetilde{N}(q) = \lim_{q \to \infty} \frac{1}{\log q} \log N(q),$$

i.e., if we can show that the set of configurations with $q - \delta(q) \cdot q < q_n < q$ has "the same entropy" as the larger set of all configurations with $q_n < q$, then the argument of the preceding section applies to show that the overwhelming majority of configurations with $q - \delta(q) \cdot q < q_n < q$ have $n \approx \epsilon^\star \log q$. The question thus becomes: How small can $\delta(q)$ be without excluding too many configurations? In particular: Is $\delta(q)$ small and constant allowed?

We will cast this question in slightly more general terms: We consider two sequences $\epsilon_n^{(1)}$ and $\epsilon_n^{(2)}$ with

- $\epsilon_n^{(1)} < \epsilon_n^{(2)}$ for all n, and
- $\epsilon_n^{(2)} \to \epsilon$, with $\epsilon_{\min} < \epsilon < \infty$,

and we ask for condition sufficing to guarantee

$$\lim_{n \to \infty} \frac{1}{n} \log \mathcal{V}_n(\epsilon_n^{(1)}, \epsilon_n^{(2)}) = s(\epsilon).$$

The following proposition gives such a condition:

Proposition 6.1 *Let the general setup be as described in the preceding paragraph, and assume that* $\epsilon_n^{(2)} - \epsilon_n^{(1)}$ *goes to zero, if at all, more slowly than exponentially in* n*, in the sense that*

$$\limsup_{n\to\infty} \frac{-\log(\epsilon_n^{(2)} - \epsilon_n^{(1)})}{n} \leq 0.$$

Then

$$\lim_{n\to\infty} \frac{1}{n} \log \mathcal{V}_n(\epsilon_n^{(1)}, \epsilon_n^{(2)}) = s(\epsilon).$$

Proof. We note first that it is always true that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathcal{V}_n(\epsilon_n^{(1)}, \epsilon_n^{(2)}) \leq \lim_{n\to\infty} \frac{1}{n} \log \mathcal{V}_n(-\infty, \epsilon_n^{(2)}) = s(\epsilon),$$

so we have only to prove

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathcal{V}_n(\epsilon_n^{(1)}, \epsilon_n^{(2)}) \geq s(\epsilon).$$

We can thus assume without loss of generality that $\epsilon_n^{(2)} - \epsilon_n^{(1)} \to 0$. Fix $\bar{\epsilon} < \epsilon$; we are going to argue that, for n sufficiently large, any configuration $a_1, \ldots, a_{n-1}$ of length $n-1$ with

$$\log q_{n-1}(a_1, \ldots, a_{n-1}) < (n-1)\bar{\epsilon}$$

can be extended, by proper choice of a_n, to a configuration of length n with

$$n\epsilon_n^{(1)} < \log q_n(a_1, \ldots, a_n) < n\epsilon_n^{(2)}.$$

This will imply

$$\mathcal{V}_n(\epsilon_n^{(1)}, \epsilon_n^{(2)}) \geq \mathcal{V}_{n-1}(-\infty, \bar{\epsilon}),$$

and hence

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathcal{V}_n(\epsilon_n^{(1)}, \epsilon_n^{(2)}) \geq \lim_{n\to\infty} \frac{1}{n} \log \mathcal{V}_{n-1}(-\infty, \bar{\epsilon}) = s(\bar{\epsilon}).$$

This holds for all $\bar{\epsilon} < \epsilon$, so

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathcal{V}_n(\epsilon_n^{(1)}, \epsilon_n^{(2)}) \geq s(\epsilon),$$

which is what we want to prove.

It remains only to prove the assertion about extension of configurations $a_1, \ldots, a_{n-1}$ with

$$\log q_{n-1}(a_1, \ldots, a_{n-1}) < n\,\bar{\epsilon}.$$

For any a_n we have

$$q_n(a_1, \ldots, a_n) = a_n q_{n-1}(a_1, \ldots, a_{n-1}) + q_{n-2}(a_1, \ldots, a_{n-2}).$$

Clearly, by taking a_n large enough, we can make $q_n > \exp(n\epsilon_n^{(1)})$. We choose the smallest a_n which accomplishes this and show that then $q_n < \exp(n\epsilon_n^{(2)})$, provided that n is large enough. Note first that

$$a_n = \frac{q_n}{q_{n-1}} - \frac{q_{n-2}}{q_{n-1}} \geq \exp(n(\epsilon_n^{(1)} - \bar{\epsilon})) - 1,$$

which is $\geq \exp(\alpha n)$ for all sufficiently large n, for an appropriately chosen $\alpha > 0$. Next,

$$\begin{aligned} \frac{q_n(a_1, \ldots, a_{n-1}, a_n)}{q_n(a_1, \ldots, a_{n-1}, a_n - 1)} &= 1 + \frac{1}{a_n - 1 + \frac{q_{n-2}}{q_{n-1}}} \\ &\leq 1 + \mathcal{O}(\exp(-\alpha n)). \end{aligned}$$

Hence,

$$\log q_n(a_1, \ldots, a_n) - \log q_n(a_1, \ldots, a_n - 1) = \mathcal{O}(\exp(-\alpha n));$$

since – by the choice of a_n –

$$\log q_n(a_1, \ldots, a_n - 1) \leq n\epsilon_n^{(1)},$$

and since, by assumption,

$$\epsilon_n^{(2)} - \epsilon_n^{(1)} \gg \exp(-n\alpha) \quad \text{for large } n,$$

it follows that

$$\log q_n(a_1, \ldots, a_n) < n\epsilon_n^{(2)} \quad \text{for } n \text{ large enough.}$$

This completes the proof of the proposition. □

It is easy to translate this result to apply to the statistical relation between q and n:

Proposition 6.2 *Let $\delta(q)$ be a sequence in $(0, 1]$ such that*

$$\limsup_{q \to \infty} \frac{\log(1/\delta(q))}{\log q} = 0,$$

(i.e., $\delta(q)$ goes to zero, if at all, less rapidly than any inverse power of q), and let $\widetilde{N}(q)$ denote the number of configurations $a_1, \ldots, a_n$ with

$$(1 - \delta(q))q < q_n(a_1, \ldots, a_n) < q. \qquad (*)$$

Then

$$\lim_{q \to \infty} \frac{1}{\log q} \log \widetilde{N}(q) = \sup_{\epsilon} \frac{s(\epsilon)}{\epsilon}. \qquad (\dagger)$$

Hence: For any $\eta > 0$, and for large q, the overwhelming majority of configurations satisfying $(*)$ have continued-fraction expansion of length between $(1 - \eta)(\epsilon^*)^{-1} \log q$ and $(1 + \eta)(\epsilon^*)^{-1} \log q$.

We omit the proof; it is a straightforward adaptation of the proofs of Propositions 5.5 and 5.8, using Proposition 6.1. We remark that the condition that $\delta(q)$ decrease less rapidly than any inverse power of q is also *necessary* for $(\dagger)$; this follows from the elementary upper bound

$$\widetilde{N}(q) \leq 2 \sum_{j=(1-\delta(q))q}^{q} j = \mathcal{O}(q^2\delta(q)).$$

7 The third law of thermodynamics.

We are going to show here that

Proposition 7.1

$$\lim_{\epsilon \to \epsilon_{\min}^{+}} s(\epsilon) = 0.$$

In other words: Our statistical mechanical system has no zero-point entropy, i.e., it satisfies the third law of thermodynamics. There are available general methods for proving the third law; see, for example Simon (1993) §III.9 or Schrader (1970). Although these methods could certainly be used here, we will instead give a simple "bare-hands" argument. In general terms, the argument goes as follows:

- We show that our system has, in a particularly clean sense, a unique ground state and a "mass gap."
- From this, we argue that the unique Gibbs state σ_β converges to the point mass at the ground state as $\beta \to \infty$. Intuitively, this means that the entropy of σ_β should go to zero, and we show that convergence takes place in a sufficiently strong sense that this expectation is realized.
- To finish the argument, we invoke a version of the "variational principle," saying that the entropy of σ_β is equal to the microcanonical entropy for $\epsilon = \bar{\epsilon}_\beta$, where $\bar{\epsilon}_\beta$ means the mean energy per lattice site in σ_β. (For this conclusion, we need only the "easy" half of the variational principle, i.e., the fact that Gibbs states maximize the free energy, and not the converse assertion that states maximizing the free energy are Gibbs states.)

The heart of the matter is the "mass gap." Recall that the Hamiltonian of an n-site finite system is $\log q_n(a_1, \ldots, a_n)$, and that q_n is strictly increasing in each a_i separately. Hence, the unique ground state of the n-site system is the configuration $(1, \ldots, 1)$. Something much stronger is true in our case: If we start from a non-ground state $(a_1, \ldots, a_n)$, pick any i for which $a_i \neq 1$, and replace the corresponding a_i by 1, keeping all the other a_j's fixed, then the energy decreases by at least a fixed nonzero amount independent of n, i, and the configuration.

Lemma 7.2 *There is a strictly positive number ϵ_g such that,*

$$\log q_n(a_1, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots a_n) \geq \log q_n(a_1, \ldots, a_{i-1}, 1, a_{i+1}, \ldots a_n) + \epsilon_g$$

for all n, all i between 1 and n, and all configurations $(a_1, \ldots, a_n)$ with $a_i \neq 1$.

Proof. Since $q_n(a_1, \ldots, a_i, \ldots, a_n)$ is nondecreasing in a_i, it suffices to consider $a_i = 2$. We will write q_j for $q_j(a_1, \ldots, a_{i-1}, 1, a_{i+1}, \ldots a_j)$ (with the obvious simplifications for $j \leq i+1$), and we put

$$\begin{aligned} d_j &:= q_j(a_1, \ldots, a_{i-1}, 2, a_{i+1}, \ldots, a_j) - q_j \quad \text{for } j \geq i \text{ and} \\ d_j &:= 0 \quad \text{for } j < i. \end{aligned}$$

Since

$$\begin{aligned} q_i(a_1,\ldots,a_{i-1},2) &= 2q_{i-1}+q_{i-2} \quad \text{and} \\ q_i(a_1,\ldots,a_{i-1},1) &= q_{i-1}+q_{i-2}, \end{aligned}$$

we get

$$d_i = q_{i-1}.$$

The d_{i+k} satisfy the recurrence

$$d_{i+k} = a_{i+k}d_{i+k-1} + d_{i+k-2},$$

i.e., the same recurrence as the $q_k(a_{i+1},\ldots,a_{i+k})$. In view of the initial condition

$$d_{i-1} = 0 \quad \text{and} \quad d_i = q_{i-1},$$

(which differs only by a factor of q_{i-1} for that for $q_k(a_{i+1},\ldots,a_{i+k})$), we see

$$d_{i+k} = q_{i-1}q_k(a_{i+1},\ldots,a_{i+k}).$$

Hence, setting $i+k=n$,

$$q_n(a_1,\ldots,2,\ldots,a_n) - q_n = q_{i-1}(a_1,\ldots,a_{i-1})q_{n-i}(a_{i+1},\ldots,a_n),$$

so we have only to show that

$$\frac{q_{i-1}(a_1,\ldots,a_{i-1})q_{n-i}(a_{i+1},\ldots,a_n)}{q_n(a_1,\ldots,a_{i-1},1,a_{i+1},\ldots a_n)} \qquad (*)$$

is bounded away from zero.

Now let $\tilde{q}_k$ and $\tilde{p}_k$ denote $q_k(a_{i+1},\ldots,a_{i+k})$ and $p_k(a_{i+1},\ldots,a_{i+k})$ respectively. The $\tilde{q}_k$ and $\tilde{p}_k$ satisfy the same recurrence as the q_{i+k}, but with initial conditions

$$\begin{aligned} \tilde{q}_0 &= 1 \qquad \tilde{q}_{-1} = 0 \\ \tilde{p}_0 &= 0 \qquad \tilde{p}_{-1} = 1. \end{aligned}$$

It follows that

$$q_{i+k} = q_i\tilde{q}_k + q_{i-1}\tilde{p}_k.$$

Setting $k=n-i$, we see that we can rewrite $(*)$ as

$$\begin{aligned} \frac{q_{i-1}\tilde{q}_{n-i}}{q_i\tilde{q}_{n-i}+q_{i-1}\tilde{p}_{n-i}} &= \frac{1}{q_i/q_{i-1}+\tilde{p}_{n-i}/\tilde{q}_{n-i}} \\ &= \frac{1}{1+q_{i-2}/q_{i-1}+\tilde{p}_{n-i}/\tilde{q}_{n-i}} \quad \text{since } q_i = q_{i-1}+q_{i-2} \\ &= \frac{1}{1+[a_{i-1},\ldots,a_1]+[a_{i+1},\ldots,a_n]} \\ &> \frac{1}{3}, \end{aligned}$$

so the assertion is proved □

The next step will be to show, using this lemma, that as $\beta \to \infty$ the Gibbs state converges to the point mass on the unique ground state configuration in a strong enough way to ensure that the entropy of the Gibbs state goes to 0. We first need to recall the definition of entropy in the present context. Let μ denote a translation-invariant probability measure on $\{1, 2, \ldots\}^{\mathbb{Z}}$, and let $\mu(a_0, \ldots, a_{n-1})$ denote the μ-probability of the configuration $(a_0, \ldots, a_{n-1})$ in the finite subset $\{0, \ldots, n-1\}$ of the index set ("lattice") $\mathbb{Z}$. We then define

$$S_n(\mu) = \sum_{a_0, \ldots, a_{n-1}} -\mu(a_0, \ldots, a_{n-1}) \log \mu(a_0, \ldots, a_{n-1}),$$

with, as usual, the convention $0 \log 0 = 0$. Then S_n is a subadditive function of n, so $\lim_{n\to\infty} \frac{1}{n} S_n(\mu)$ exists and is equal to $\inf_n \frac{1}{n} S_n(\mu)$; the common value is the entropy $s(\mu)$ of μ. We can now formulate:

Proposition 7.3 *Let σ_β denote the unique Gibbs state with inverse temperature β. Then $s(\sigma_\beta) \to 0$ when $\beta \to \infty$.*

Proof. We are going to show that, in the notation of the preceding paragraph, $S_1(\sigma_\beta)$ converges to 0 as $\beta \to \infty$; since

$$0 \leq s(\sigma_\beta) \leq \cdots \leq S_n(\sigma_\beta) \leq \cdots \leq S_1(\sigma_\beta),$$

the assertion follows. We will need some notation related to Gibbs states. For Λ any subset of $\mathbb{Z}$, we denote by X_Λ the set of configurations in Λ, i.e., of mappings $\Lambda \to \{1, 2, \ldots\}$. For Λ a *finite* subset of $\mathbb{Z}$, the interaction gives rise to a function Ψ_Λ, defined on $X_\Lambda \times X_{\Lambda^c}$ with the interpretation that $\Psi_\Lambda(a_\Lambda, a_{\Lambda^c})$ is the sum of the self-energy of the finite configuration a_Λ and its energy of interaction with the outside configuration a_{Λ^c}.[5] A Gibbs state with inverse temperature β is a probability measure on $X_{\mathbb{Z}}$ with the property that, for any finite Λ, the conditional probability of finding a_Λ inside Λ given that the configuration outside Λ is a_{Λ^c} is

$$\frac{\exp(-\beta \Psi_\Lambda(a_\Lambda, a_{\Lambda^c}))}{Z_\Lambda(\beta, a_{\Lambda^c})},$$

with

$$Z_\Lambda(\beta, a_{\Lambda^c}) = \sum_{a'_\Lambda \in X_\Lambda} \exp(-\beta \Psi_\Lambda(a'_\Lambda, a_{\Lambda^c})).$$

It follows from the considerations of §3 that

$$\Psi_\Lambda(a_\Lambda, a_{\Lambda^c}) - \sum_{i \in \Lambda} \log(a_i)$$

is bounded (for fixed Λ.)

We apply these considerations in the very simple case $\Lambda = \{0\}$. We define

$$V(a_0, \hat{a}) := \Psi_{\{0\}}(a_0, \hat{a}) - \Psi_{\{0\}}(1, \hat{a}),$$

where $\hat{a}$ denotes a general configuration of $\mathbb{Z} \setminus \{0\}$. Then

[5] Although these individual energies are not unambiguously defined, the sum is unambiguous, at least up to an additive constant.

– $V(1, \hat{a}) = 0$

– $V(a_0, \hat{a}) \geq \epsilon_g$ for $a_0 > 1$, by Lemma 7.2

– $V(a_0, \hat{a}) - \log a_0$ is bounded.

We put

$$Z_0(\beta, \hat{a}) := \sum_{a_0 \geq 1} \exp(-\beta V(a_0, \hat{a})) = 1 + \sum_{a_0 > 1} \exp(-\beta V(a_0, \hat{a})).$$

By the preceding remarks, $\exp(-\beta V(a_0, \hat{a}))$ converges to zero for any fixed $a_0 > 1$ and is furthermore $< a_0^{-\beta/2}$ (for example) for any sufficiently large a_0, all uniformly in $\hat{a}$. Hence, in particular, $Z_0(\beta, \hat{a}) \to 1$ as $\beta \to \infty$. From the definition of Gibbs state, the conditional probability of finding a_0 at the origin given the configuration $\hat{a}$ away from the origin is

$$\frac{\exp(-\beta V(a_0, \hat{a}))}{Z_0(\beta, \hat{a})},$$

which converges to 1 for $a_0 = 1$ and to 0 for $a_0 > 1$, and is bounded by $a_0^{-\beta/2}$ for all sufficiently large a_0, again uniformly in $\hat{a}$.

Now let $\sigma_\beta(a_0)$ denote the probability, with respect to the unique Gibbs state σ_β of having a_0 at the origin. Since this probability is a convex combination of the above conditional probabilities, it follows that

$$\sigma_\beta(1) \to 1, \quad \sigma_\beta(a_0) \to 0 \quad \text{for} \quad a_0 > 1, \quad \text{as} \quad \beta \to \infty,$$

and

$$\sigma_\beta(a_0) < a_0^{-\beta/2} \quad \text{for all sufficiently large } a_0.$$

Hence,

$$-\sigma_\beta(a_0) \log \sigma_\beta(a_0)$$

converges to zero with $\beta \to \infty$, for all a_0, and furthermore is $\leq a_0^{-\beta/4}$ for all sufficiently large a_0 (since $-t \log t \leq t^{1/2}$ for t positive and sufficiently small.) From this it follows that

$$S_1(\sigma_\beta) = \sum_{a_0 = 1}^{\infty} -\sigma_\beta(a_0) \log \sigma_\beta(a_0) \longrightarrow 0 \quad \text{with} \quad \beta \longrightarrow \infty.$$

□

It remains to relate $s(\sigma_\beta)$ to the microcanonical entropy. To avoid confusion, we will, for this section, denote the microcanonical entropy by $s_{\text{mc}}(\epsilon)$; s without subscript means the entropy of a translation-invariant measure, as defined above.

– from Proposition 4.1,

$$p(\beta) = s(\sigma_\beta) - \beta\, \bar{\epsilon}(\beta),$$

where $\bar{\epsilon}(\beta)$ denotes the mean energy per lattice site in the Gibbs state σ_β.

– from the theory of thermodynamic limits for partition functions and the Legendre transform, we have

$$p(\beta) = s_{\mathrm{mc}}(\epsilon(\beta)) - \beta\,\epsilon(\beta),$$

where $\epsilon(\beta)$ is defined as the unique ϵ for which $s_{\mathrm{mc}}(\epsilon) - \beta\epsilon$ takes on its supremum,

– and finally, from a standard argument using Propositions 4.1 and 5.3

$$\bar{\epsilon}(\beta) = \epsilon(\beta),$$

Putting all this together, we see that

$$s(\sigma_\beta) = s_{\mathrm{mc}}(\epsilon(\beta)).$$

As $\beta \to \infty$, one the one hand $s(\sigma_\beta) \to 0$ – by what was shown above – and on the other hand, $\epsilon(\beta)$ is continuous and strictly decreasing, and converges to ϵ_{min}. This completes the proof of Proposition 7.1

8 Joint distribution of $\log q$ and the Farey depth.

We consider here, in addition to $H_n = \log q_n(a_1, \ldots, a_n)$, the function

$$F_n(a_1, \ldots, a_n) := a_1 + \cdots + a_n$$

on $\mathbb{N}_+^n$. As noted in the introduction, F_n has a number-theoretic significance: It is the level in the Farey-tree representation of rational numbers at which $[a_1, \ldots, a_n]$ appears. We will accordingly refer to F_n as the *Farey depth*. The sequence of functions F_n is trivially extensive; if interpreted as an energy, it corresponds to a non-interacting system. We put the two quantities $\log q_n$ and F_n together and regard them as components of a single $\mathbb{R}^2$-valued extensive quantity

$$g_n(a_1, \ldots, a_n) := (\log q_n(a_1, \ldots, a_n), F_n(a_1, \ldots, a_n)).$$

The theory of the microcanonical entropy of such vector-valued extensive quantities is developed under technically favorable assumptions in Lanford (1973) and has been generalized to apply to the present situation in Ruedin (1994). To formulate the results, we will use the following notation: For J a subset of $\mathbb{R}^2$ and $n = 1, 2, \ldots$, $\mathcal{V}_{q,F}(n, J)$ will denote the number of configurations $(a_1, \ldots, a_n)$, of length n, with

$$\frac{g(a_1, \ldots, a_n)}{n} \in J.$$

The main results are as follows:

Proposition 8.1 *There exist:*

– *a (non-empty) convex open set $\mathcal{D}_{q,F}$ in $\mathbb{R}^2$ and*

– *a non-negative concave function* $s_{q,F}$ *on* $\mathcal{D}_{q,F}$

such that:

– *If* J *is an open convex subset of* $\mathbb{R}^2$ *with* $J \cap \mathcal{D}_{q,F} \neq \emptyset$, *then*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathcal{V}_{q,F}(n, J) = \sup_{x \in J \cap \mathcal{D}_{q,F}} s_{q,F}(x).$$

– *if* J *is an open convex set whose distance from* $\mathcal{D}_{q,F}$ *is strictly positive, then* $\mathcal{V}_{q,F}(n, J)$ *vanishes for all sufficiently large* n.

The intuitive meaning is: If $(\epsilon, f) \in \mathcal{D}_{q,F} \subset \mathbb{R}^2$, then there exist, for arbitrarily large n, configurations with – simultaneously – $\log q_n \approx n\,\epsilon$ and $F_n \approx n\,f$; the number of such configurations is furthermore $\approx \exp(n s_{q,F}(\epsilon, f))$. If, on the other hand, (ϵ, f) is outside the closure of $\mathcal{D}_{q,F}$, then $(n\,\epsilon, n\,f)$ is *excluded* as a value for $(\log q_n, F_n)$ for large n. As in the single-observable case, this proposition evades the potentially delicate question of the behavior of $\mathcal{V}_{q,F}(n, J)$ when J has distance zero from $\mathcal{D}_{q,F}$ but does not actually intersect it. Comparing the defining properties of $s_{q,F}$ with those of the single-observable s, we see that

$$s(\epsilon) = \sup_f s_{q,F}(\epsilon, f).$$

Furthermore, $s_{q,F}$ has the following interpretation: If I is any interval for which

$$\sup_{f \in I} s_{q,F}(\epsilon, f) < s(\epsilon),$$

then, for large n, among configurations of length n with $q_n \approx \exp(n\epsilon)$, only a vanishingly small fraction have $F_n/n \in I$. Somewhat less precisely: For large n, among configurations with $q_n \approx \exp(n\epsilon)$, the values of F_n/n are strongly concentrated around values of f where $s_{q,F}(\epsilon, f)$ is maximal.

The preceding proposition is a version of a result which holds with great generality. A first special feature of the particular situation we are considering is

Proposition 8.2 *Let* $(\epsilon_0, f_0) \in \mathcal{D}_{q,F}$, *and let* $f_1 > f_0$. *Then* $(\epsilon_0, f_1) \in \mathcal{D}_{q,F}$, *and* $s_{q,F}(\epsilon_0, f_1) \geq s_{q,F}(\epsilon_0, f_0)$.

Roughly: $s_{q,F}(\epsilon, f)$ is non-decreasing in f for fixed ϵ.

Proof. For purposes of this argument, we denote by $R_\delta(\epsilon, f)$ the open square of side-length 2δ centered at $(\epsilon, f) \in \mathbb{R}^2$. We are going to argue:

Claim. *For sufficiently large* n,

$$\mathcal{V}_{q,F}(n+1, R_{2\delta}(\epsilon_0, f_1)) \geq \mathcal{V}_{q,F}(n, R_\delta(\epsilon_0, f_0)).$$

Before proving the claim, we show how it implies the proposition. In the first place, it follows at once that (ϵ_0, f_1) must be in the *closure* of $\mathcal{D}_{q,F}$; otherwise, for small enough δ, $\mathcal{V}_{q,F}(n+1, R_{2\delta}(\epsilon_0, f_1))$ would have to vanish for large n whereas $\mathcal{V}_{q,F}(n, R_\delta(\epsilon_0, f_0))$ is non-zero. By applying the same argument, with ϵ_0 moved a little, we see that *a neighborhood of* (ϵ_0, f_1) *lies inside the closure of* $\mathcal{D}_{q,F}$. But $\mathcal{D}_{q,F}$ is a convex open set, so this implies that (ϵ_0, f_1) itself is in $\mathcal{D}_{q,F}$. It then follows that $s_{q,F}$ is continuous at (ϵ_0, f_1) (as it is at (ϵ_0, f_0)). Thus,

$$s_{q,F}(\epsilon_0, f_1) = \inf_{\delta>0} \lim_{n\to\infty} \frac{1}{n+1} \log \mathcal{V}_{q,F}(n+1, R_{2\delta}(\epsilon_0, f_1)),$$

but, on the other hand, applying the claim again,

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{n+1} \log \mathcal{V}_{q,F}(n+1, R_{2\delta}(\epsilon_0, f_1)) \\
&\qquad \geq \lim_{n\to\infty} \frac{1}{n+1} \log \mathcal{V}_{q,F}(n, R_\delta(\epsilon_0, f_0)) \\
&\qquad = \sup\{s_{q,F}(\epsilon, f) : (\epsilon, f) \in R_\delta(\epsilon_0, f_0)\} \\
&\qquad \geq s_{q,F}(\epsilon_0, f_0),
\end{aligned}$$

so the proposition follows from the claim.

Proof of Claim: Let $(a_1, \ldots, a_n)$ be a configuration with

$$\frac{1}{n}(\log q_n(a_1, \ldots, a_n), F_n(a_1, \ldots, a_n)) \in R_\delta(\epsilon_0, f_0).$$

We are going to make a configuration of length $n+1$ by adjoining a single large a_n and show that, for sufficiently large n, the augmented configuration always has

$$\frac{1}{n+1}(\log q_{n+1}(a_1, \ldots, a_{n+1}), F_n(a_1, \ldots a_{n+1})) \in R_{2\delta}(\epsilon_0, f_1);$$

this will establish the claim. We choose a_{n+1} to be the smallest integer with $a_1 + \ldots + a_n + a_{n+1} \geq (n+1)f_1$. Since $a_1 + \ldots + a_n \approx nf_0$, it is easy to find upper and lower bounds for a_{n+1} both of which go to infinity linearly with n (We need to assume here, as we may without loss of generality, that δ is chosen small enough so that $f_0 + \delta < f_1$.) Since

$$q_{n+1}(a_1, \ldots, a_{n+1}) = a_{n+1}q_n + q_{n-1},$$

we get

$$a_{n+1}q_n < q_{n+1} < (a_{n+1}+1)q_n,$$

and hence – in view of the growth rate of the a_n's –

$$\log q_{n+1} = \log q_n + \mathcal{O}(\log n).$$

Since

$$\frac{1}{n} \log q_n \in (\epsilon_0 - \delta, \epsilon_0 + \delta),$$

it follows that, for n sufficiently large,

$$\frac{1}{n+1}\log q_{n+1} \in (\epsilon_0 - 2\delta, \epsilon_0 + 2\delta),$$

which completes the argument. □

This gives us at least a rough picture of $\mathcal{D}_{q,F}$: We know from the outset that it lies to the right of the vertical line $\{\epsilon = \epsilon_{\min}\}$, since smaller values of ϵ are asymptotically excluded without any condition on F. It also extends arbitrarily far to the right, since $s(\epsilon)$ is defined for arbitrary large ϵ. In view of the preceding proposition, it is a union over ϵ of semi-infinite vertical lines:

$$\mathcal{D}_{q,F} = \{(\epsilon, f) : \epsilon > \epsilon_{\min}, f > f_{\min}(\epsilon)\}.$$

The function $f_{\min}(\epsilon)$ defined in the preceding formula is convex – since its epigraph $\mathcal{D}_{q,F}$ is – and hence continuous. It is not difficult to see that $f_{\min}(\epsilon)$ is monotone non-decreasing and not constant; hence, by convexity, that it goes to ∞ with ϵ. We will in fact determine $f_{\min}(\epsilon)$ explicitly in §10; somewhat surprisingly, it turns out to be piecewise linear.

We turn next to the canonical ensemble. We set

$$Z_n(\beta, \gamma) := \sum_{a_1,\dots,a_n} \exp(-\beta H_n(a_1, \dots, a_n) - \gamma F_n(a_1, \dots a_n)).$$

In view of

$$\begin{aligned} H_n &= \log a_1 + \cdots + \log a_n + \text{bounded}, \\ F_n &= a_1 + \cdots + a_n, \end{aligned}$$

it is easy to see that the sum converges for *all* β, positive and negative, for $\gamma > 0$; for $\beta > 1$ for $\gamma = 0$ – the case already studied – and not at all for $\gamma < 0$. We refer to Ruedin (1994) for the proof of the following result, which involves only straightforward generalizations of standard results:

Proposition 8.3 *1. For $\gamma > 0$,*

$$\lim_{n\to\infty} \frac{1}{n} \log Z_n(\beta, \gamma) =: p_{q,F}(\beta, \gamma)$$

exists, and is given by the Legendre transform of $s_{q,F}$:

$$p_{q,F}(\beta, \gamma) = \sup\{s_{q,F}(\epsilon, f) - \beta\epsilon - \gamma f : (\epsilon, f) \in \mathcal{D}_{q,F}.\}$$

2. $p_{q,F}(\beta, \gamma)$ is a real-analytic and strictly convex function of (β, γ) in $\{\gamma > 0\}$.

By "Legendre duality," $s_{q,F}$ is the inverse Legendre transform of $p_{q,F}$. To be precise:

For $(\epsilon, f) \in \mathcal{D}_{q,F}$,

$$s_{q,F}(\epsilon, f) = \inf_{\beta,\gamma} p_{q,F}(\beta, \gamma) + \beta\epsilon + \gamma f;$$

the infimum on the right-hand side is $-\infty$ *for* (ϵ, f) *outside the closure of* $\mathcal{D}_{q,F}$.

(There are a number of versions of "Legendre duality." The preceding assertions follow from Theorem I.6.4 of Simon (1993) and the following observation: We have taken the definition domain of $s_{q,F}$ to be open. If we extend $s_{q,F}$ to the closure of the domain by defining it at boundary points to be the lim sup of values at nearby interior points, then the subgraph of the extended function is convex and closed. Hence, except for a sign, the extended function and the closure of the original domain form a *Fenchel pair* in the sense of Simon (1993).)

In particular, if we define

$$\epsilon_{q,F}(\beta,\gamma) := -\frac{\partial p_{q,F}}{\partial \beta}(\beta,\gamma), \quad f_{q,F}(\beta,\gamma) := -\frac{\partial p_{q,F}}{\partial \gamma}(\beta,\gamma),$$

for $\gamma > 0$ and β arbitrary, then $p(\beta,\gamma) + \beta\epsilon_{q,F}(\beta_0,\gamma_0) + \gamma f_{q,F}(\beta_0,\gamma_0)$ has vanishing gradient – and therefore a minimum – at $\beta = \beta_0$, $\gamma = \gamma_0$. Hence, by Legendre duality, $(\epsilon_{q,F}(\beta_0,\gamma_0), f_{q,F}(\beta_0,\gamma_0))$ is in the closure of $\mathcal{D}_{q,F}$, and this holds for all (β_0,γ_0) with $\gamma_0 > 0$. By strict convexity of $p_{q,F}$, the mapping

$$(\beta,\gamma) \mapsto (\epsilon_{q,F}(\beta,\gamma), f_{q,F}(\beta,\gamma))$$

is open and injective. Its image must therefore lie in $\mathcal{D}_{q,F}$, not just in its closure, and we have

$$s_{q,F}(\epsilon_{q,F}(\beta,\gamma), f_{q,F}(\beta,\gamma)) = p(\beta,\gamma) + \beta\epsilon_{q,F}(\beta,\gamma) + \gamma f_{q,F}(\beta,\gamma).$$

The right-hand side is a real-analytic function of (β,γ), and the inverse of

$$(\beta,\gamma) \mapsto (\epsilon_{q,F}(\beta,\gamma), f_{q,F}(\beta,\gamma))$$

is real-analytic by the inverse function theorem, so $s_{q,F}$ is real-analytic and, by a straightforward computation, strictly concave *on the image* $\mathcal{D}^{(0)}_{q,F}$ *of the upper half plane* $\{\gamma > 0\}$ *under*

$$(\beta,\gamma) \mapsto (\epsilon_{q,F}(\beta,\gamma), f_{q,F}(\beta,\gamma)).$$

Our next task is to determine the image domain $\mathcal{D}^{(0)}_{q,F}$ of the "analytic" Legendre transform. We do this in a way which produces some extra information which we will need later.

Lemma 8.4 *For any* $\gamma > 0$ *and any* $\epsilon > \epsilon_{\min}$, *there is a unique* β *with*

$$\epsilon_{q,F}(\beta,\gamma) = \epsilon.$$

We will denote this β *by* $\hat{\beta}(\epsilon,\gamma)$.

- $\hat{\beta}$ *is a real-analytic function of* ϵ, γ.
- $f_{q,F}(\hat{\beta}(\epsilon,\gamma),\gamma)$ *is strictly decreasing in* γ.
- $\frac{\partial s_{q,F}}{\partial f}(\epsilon, f_{q,F}(\hat{\beta},\gamma)) = \gamma$.

– *For* $\epsilon > \epsilon_{\min}$, $\hat{\beta}(\epsilon,\gamma) \to \beta(\epsilon)$ *as* $\gamma \to 0^+$*; convergence is uniform on compact sets in* $(\epsilon_{\min},\infty)$.

– $f_{q,F}(\hat{\beta}(\epsilon,\gamma),\gamma) \longrightarrow \begin{cases} f_{\min}(\epsilon) & \gamma \to \infty \\ \bar{f}(\beta(\epsilon)) & \gamma \to 0^+, \beta(\epsilon) > 2 \\ \infty, & \gamma \to 0^+, \beta(\epsilon) \le 2. \end{cases}$

Here, $\bar{f}(\beta)$ *means the means value of* F_n/n *in the Gibbs state with inverse temperature* β *(and* $\gamma = 0$*.).*

Proof. For fixed γ, $\epsilon_{q,F}(\beta,\gamma)$ is a strictly decreasing real-analytic function of β. We are going to argue that it converges to $\epsilon_{\min}$ as $\beta \to \infty$ and ∞ as $\beta \to -\infty$; continuity then implies that it takes on every intermediate value exactly once, i.e., that $\hat{\beta}(\epsilon,\gamma)$ is defined. For this argument, we use the fact that $\epsilon_{q,F}(\beta,\gamma)$ is equal to the mean value, in the unique Gibbs state for (β,γ), of the function $-\log([a_0,a_1,\ldots])$.

– An easy extension of the arguments of §7 shows that, as $\beta \to \infty$ with fixed γ, the corresponding Gibbs state converges to the point mass at $(\ldots,1,1,\ldots)$, in a strong enough sense to allow us to conclude that $\epsilon_{q,F}(\beta,\gamma) \to -\log([1,1,\ldots]) = \epsilon_{\min}$.

– The difference between $-\log([a_0,\ldots])$ and $\log a_0$ is bounded, so it is enough to show that the mean value of $\log a_0$ goes to ∞ as $\beta \to -\infty$. By the arguments of §7, the Gibbs state assigns a probability to a_0 which can be written as

$$c_1(\beta)\exp(-\gamma a_0 - \beta(\log a_0 + c_2(\beta,a_0))$$

with $c_2(\beta,a_0)$ uniformly bounded in a_0, β. From this form, it is clear that, as $\beta \to -\infty$, the probability distribution becomes concentrated on large values of a_0 and hence that the mean value of $\log a_0$ goes to ∞.

Thus, the existence of $\hat{\beta}(\epsilon,\gamma)$ is established; real analyticity follows from the inverse function theorem (using the strict convexity of $p_{q,F}(\,.\,)$.) Strict positivity of the derivative of $f_{q,F}(\hat{\beta}(\epsilon,\gamma),\gamma)$ with respect to γ follows from the strict convexity of $p_{q,F}(\,.\,)$ by a straightforward computation. The formula

$$\frac{\partial s_{q,F}}{\partial f}(\epsilon_{q,F}(\beta,\gamma), f_{q,F}(\beta,\gamma)) = \gamma$$

holds for all (β,γ) by the elementary properties of the Legendre transform; inserting $\hat{\beta}$ for β and remembering how $\hat{\beta}$ was defined gives

$$\frac{\partial s_{q,F}}{\partial f}(\epsilon, f_{q,F}(\hat{\beta},\gamma) = \gamma.$$

As $\gamma \to 0$ with β fixed > 1, $\epsilon_{q,F}(\beta,\gamma) \to \epsilon(\beta)$ and the convergence is uniform for β is any compact set in $(1,\infty)$. Hence, for fixed $\epsilon > \epsilon_{\min}$ and any $\beta_1 < \beta(\epsilon) < \beta_2$,

$$\epsilon_{q,F}(\beta_1,\gamma) > \epsilon(\beta) > \epsilon_{q,F}(\beta_2,\gamma) \quad \text{for all sufficiently small } \gamma.$$

For γ small enough so that these inequalities hold, $\beta_1 < \hat{\beta}(\epsilon,\gamma) < \beta_2$; this shows that $\hat{\beta}(\epsilon,\gamma)$ converges for $\gamma \to 0^+$ to $\beta(\epsilon)$, and it is easy to see that the convergence is in fact uniform on compact subintervals of $(\epsilon_{\min},\infty)$. It is also easy to see that, for $\gamma \to 0$, $f(\beta,\gamma)$ converges to $\bar{f}(\beta)$ for $\beta > 2$ and to ∞ for $1 < \beta < 2$; the convergence is uniform on compact subsets of $(1,\infty)$. Hence, also for $\gamma \to 0$, $f_{q,F}(\hat{\beta}(\epsilon,\gamma),\gamma)$ converges to $\bar{f}(\beta(\epsilon))$ for $\epsilon < \epsilon^*$ and to ∞ for $\epsilon \geq \epsilon^*$, as asserted. □

Proposition 8.5 *$\mathcal{D}^{(0)}_{q,F} = \{(\epsilon,f) \in \mathcal{D}_{q,F} : \epsilon \geq \epsilon^*$ or $f < \bar{f}(\beta(\epsilon))\}$. For $\epsilon \geq \epsilon^*$, $f \mapsto s_{q,F}(\epsilon,f)$ is strictly increasing and real-analytic on $(f_{\min}(\epsilon),\infty)$; for $\epsilon_{\min} < \epsilon < \epsilon^*$, $f \mapsto s_{q,F}(\epsilon,f)$ is real-analytic and strictly increasing on $(f_{\min}(\epsilon),\bar{f}(\beta(\epsilon)))$, but constant – equal to $s(\epsilon)$ – for $f \geq \bar{f}(\beta(\epsilon))$.*

Proof. The image under the inverse Legendre transformation of the parametrized curve

$$\gamma \mapsto (\hat{\beta}(\epsilon_0,\gamma),\gamma)$$

is a vertical segment above ϵ_0 in the (ϵ,f) plane which evidently lies $\mathcal{D}^{(0)}_{q,F}$. As γ runs from 0 to ∞, the segment is traversed downward. The f-coordinate of the upper end of the segment is $\lim_{\gamma\to 0^+} f_{q,F}(\hat{\beta}(\epsilon_0,\gamma),\gamma)$, which, by the preceding lemma is ∞ for $\epsilon_0 \geq \epsilon^*$ and $\bar{f}(\beta(\epsilon))$ otherwise. We temporarily denote the lower end of the segment by (ϵ_0, f_∞). f_∞ is evidently $\geq f_{\min}(\epsilon_0)$; we want to show that equality actually holds. To see this, we note that $f \mapsto s_{q,F}(\epsilon_0,f)$ is concave and nondecreasing on $(f_{\min}(\epsilon_0),\infty)$. At $f = f_{q,F}(\hat{\beta}(\epsilon_0),\gamma)$, its derivative is equal to γ. Hence, as $f \to f_\infty^+$, the derivative goes to ∞. This is not compatible with concavity unless $f_\infty = f_{\min}$. Furthermore, in the case $\epsilon_0 < \epsilon^*$, the derivative approaches 0 as $f \to \bar{f}(\beta(\epsilon_0))$; concavity and monotonicity then imply that the function must be constant for $f \geq \bar{f}$. As a consequence: If $\epsilon_0 < \epsilon^*$ and $f \geq \bar{f}$, then (ϵ_0,f) cannot lie in the image $\mathcal{D}^{(0)}_{q,F}$ of the analytic Legendre transform, i.e., the set of points above ϵ_0 in $\mathcal{D}^{(0)}_{q,F}$ are exactly those with $f_{\min}(\epsilon_0) < f < \bar{f}(\beta(\epsilon_0))$. Together with the analyticity and strict monotonicity of the analytic Legendre transform onto $\mathcal{D}^{(0)}_{q,F}$, this proves all the assertions of the proposition. □

Our intuition about statistical-mechanical systems suggests that fixing the energy – in the absence of phase transitions – determines all other extensive quantities. In the present context, this suggests that fixing $\log q_n/n$ – within some appropriate thickened energy surface – ought to determine F_n/n statistically. We will argue here, on the basis of the above results, that this is *not* the case, if we allow the size of the system to fluctuate. What happens instead is that the typical values of F_n/n go to infinity as the size of the system goes to infinity.

We consider a q – which will tend to ∞ – and a fixed parameter y and denote by $p(q,y)$ the fraction of set of configurations – of whatever length – with $q_n < q$ which also satisfy $F_n < y \log q$. Since configurations with $q_n < q$ nearly all have $\log q_n \approx \log q$, this means roughly the set of configurations with $F_n/\log q_n < y$. The kinds of arguments used in the proof of Proposition 5.5 show that

$$\lim_{q\to\infty} \frac{1}{\log q} \log p(q,y) = \sup_\epsilon \frac{s_{q,F}(\epsilon,\epsilon y)}{\epsilon} - \sup_\epsilon \frac{s(\epsilon)}{\epsilon},$$

provided that y is large enough so that the line $f = y\epsilon$ intersects $\mathcal{D}_{q,F}$. We are going to argue that the right-hand side is < 0 for all values of y, i.e., the probability that $F_n/\log q_n$ is $< y$ – given that $q_n < q$ – becomes exponentially small at $q \to \infty$ for all y. The argument goes as follows: It is always true that

$$s_{q,F}(\epsilon, f) \leq s(\epsilon),$$

and it follows from Proposition 5.3 that $s'(\epsilon) \to 1$ for $\epsilon \to \infty$ and hence that $s(\epsilon)/\epsilon \to 1$ in the same limit. Hence, if

$$\sup_{\epsilon} \frac{s_{q,F}(\epsilon, y\epsilon)}{\epsilon}$$

is not taken on at a finite ϵ, then it is ≤ 1, whereas $\sup_\epsilon s(\epsilon)/\epsilon$ was shown earlier to be equal to 2. Thus, the assertion is proved if the supremum is not taken on. Suppose now that the supremum is taken on, at, say, ϵ_1. If $\epsilon_1 \neq \epsilon^*$, then

$$\frac{s_{q,F}(\epsilon_1, y\epsilon_1)}{\epsilon_1} \leq \frac{s(\epsilon_1)}{\epsilon_1} < \frac{s(\epsilon^*)}{\epsilon^*};$$

the last inequality is strict since $s(\epsilon)/\epsilon$ takes on its supremum only at ϵ^*. Thus, the assertion is also proved if the supremum is taken on at any ϵ other than ϵ^*. Finally, if the supremum is taken on at ϵ^*, then we have

$$\frac{s_{q,F}(\epsilon^*, y\epsilon^*)}{\epsilon^*} < \frac{s(\epsilon^*)}{\epsilon^*},$$

since $f \mapsto s_{q,F}(\epsilon^*, f)$ is strictly increasing on $(f_{\min}(\epsilon^*), \infty)$ with asymptotic value $s(\epsilon^*)$. Thus, the assertion is proved in this case too, so all cases are covered.

9 The continuum representation.

We describe here a neat and convenient representation for the Ruelle transfer operator for our system. We deviate here from Ruedin (1994), where the theory of the transfer operator is extended to a general class of systems including the one we are treating. This extension turns out to be technical and complicated. What we do here is to use the special features of our system to give a simple, if limited, treatment.

We start from the observation that the mapping

$$(a_0, a_1, \ldots) \longmapsto [a_0, a_1, \ldots]$$

sends the space of semi-infinite configurations

$$\Omega_+ := \{1, 2, \ldots\}^{\mathbb{N}}$$

bijectively onto the irrational numbers in $[0, 1]$. We will use this mapping to transport various objects from the configuration space to the unit interval, where they may be easier to work with. We start by looking at

– the Gibbs state of the semi-infinite system (index set $\{0, 1, \ldots\}$)

– the Gibbs state for the two-sided infinite system (index set $\mathbb{Z}$) projected onto the semi-infinite configuration space.

These are both probability measures on Ω_+; the second of them is shift-invariant and the first presumably not. However, as Ruelle observed, the Gibbs state for the semi-infinite system has the advantage of satisfying a relatively simple equation. This comes about as follows: We can construct the semi-infinite Gibbs state by

– Constructing the semi-infinite Gibbs state "with one fewer lattice site," i.e., on configurations labeled by $1, 2, 3, \ldots$ rather than $0, 1, 2, 3, \ldots$. Because of uniqueness, this is the same as the semi-infinite Gibbs state of Ω_+ shifted one place to the right.

– appending a new lattice site at the left-hand end,

– assigning weights proportional to $\exp(-A(a_0, a_1, \ldots))$ to the possible state a_0 at the new lattice site, and

– normalizing.

In other words, the assertion is that

$$e^{-A(a_0,a_1,\ldots)}\, da_0\, \sigma_+(da_1, da_2, \ldots) = \lambda \sigma_+(da_0, da_1, da_2, \ldots). \qquad (*)$$

Here,

– σ_+ denotes the semi-infinite Gibbs state

– $A(a_0, a_1, \ldots)$ denotes $-\beta \log[a_0, a_1, \ldots]$ (or $-\beta \log[a_0, \ldots] + \gamma a_0$, if we are talking about the two-observable situation.)

– the da_i's appearing inside σ_+ are "symbolic," but the da_0 on the left stands for counting measure on $\{1, 2, \ldots\}$.

– λ – or perhaps its reciprocal – is the normalizing factor.

The left-hand side of $(*)$ defines a linear operator $\mathcal{L}^*$ from measures on Ω_+ to measures on Ω_+; $(*)$ says that σ_+ is an eigenvector for $\mathcal{L}^*$ with eigenvalue λ. Iterating $(*)$ n times corresponds to adding n sites to the left. It is easy to see, using standard ideas from the theory of Gibbs states, that the semi-infinite Gibbs state σ_+ is the *unique* probability measure satisfying $(*)$, i.e., the unique probability measure which is an eigenvector of $\mathcal{L}^*$. The same set of considerations shows that the partition function $Z_n(\beta, \gamma)$ admits upper and lower bounds of the form $c\lambda^n$, c a strictly positive constant independent of n; hence, that

$$p(\beta, \gamma) = \log \lambda.$$

We now transport this whole picture to the unit interval. We will generally use the same notation for objects on the sequence space and the corresponding transported objects on the unit interval; for

example, σ_+ will also denote the measure on the unit interval obtained by transporting the semi-infinite Gibbs state. We recall that the left shift

$$(a_0, a_1, a_2, \ldots) \longmapsto (a_1, a_2, \ldots)$$

carries over to the Gauss map

$$t \longmapsto \text{fract}(\frac{1}{t}).$$

The construction on the right-hand side of $(*)$ translates into the following: Given a measure μ on $[0,1]$ (assigning measure zero to the rational numbers), we construct a new measure $\mathcal{L}^*\mu$ by specifying that the $\mathcal{L}^*\mu$-measure of any set contained in one of the intervals $(1/(a_0+1), 1/a_0)$ is the integral of $e^{\gamma a_0}/(a_0+t)^\beta$ over the preimage of the set in question under $t \mapsto 1/(a_0+t)$. Then σ_+ is characterized as the unique probability measure transformed into a multiple of itself by $\mathcal{L}^*$.

We now have:

Proposition 9.1 *For $\beta = 2$ and $\gamma = 0$,*

- *σ_+ is Lebesgue measure on $[0,1]$, and*
- *The transported projected two-sided infinite Gibbs state σ is the Gauss measure $\frac{1}{\log 2}\frac{dt}{1+t}$*
- *$\epsilon^* = \frac{\pi^2}{12 \log 2}$*

Proof. σ_+ is uniquely characterized by the fact that it is transformed into a multiple of itself by the operation of the preceding paragraph. To prove the first assertion, it is therefore enough to show that Lebesgue measure is unchanged by this operation. Concretely, it is enough to show that Lebesgue measure itself and the transform of Lebesgue measure assign the same measure to any interval J contained in some one of the intervals $(1/(a_0+1), 1/a_0), a_0 = 1, 2, \ldots$. In other words, we want to show that the length of J is the integral of the function $(a_0+t)^{-2}$ over the preimage of the interval under $t \mapsto 1/(a_0+t)$, and this follows at once from the fact that the absolute value of the derivative of $t \mapsto 1/(a_0+t)$ is $(a_0+t)^{-2}$. Thus, the transform of the one-sided Gibbs state is identified with Lebesgue measure, and the rescaling factor λ is shown to be one.

We now turn to the determination of the image of the projected Gibbs measure for the doubly infinite system under the mapping

$$(a_0, a_1, \ldots) \mapsto [a_0, a_1, \ldots].$$

We denote both the projected Gibbs state and the corresponding measure on $[0,1]$ by σ. From the general theory of Gibbs states, we know that

- σ_+ and σ are equivalent, i.e., have the same null sets.
- σ is ergodic (with respect to the left shift respectively the Gauss map)

From these facts it follows that σ *is the only invariant probability measure equivalent to* σ_+. For $\beta = 2$, in the unit interval representation, this means that σ is the only probability measure on the unit interval equivalent to Lebesgue measure and invariant under the Gauss map. But it is well known – and in any case follows from an easy computation – that the Gauss measure is invariant under the Gauss map; hence, the Gauss measure must coincide with σ. Since ϵ^* is the mean value of $-\log[a_0, a_1, \ldots]$ with respect to the projected Gibbs state, we conclude that

$$\epsilon^* = -\frac{1}{\log 2}\int_0^1 \frac{\log t\, dt}{1+t} = \frac{\pi^2}{12\log 2}.$$

It is easy to see that the operator $\mathcal{L}^*$ on measures described above is the adjoint of an operator on continuous functions given by

$$(\mathcal{L}f)(t) = \sum_{a_0=1}^{\infty} \frac{e^{-\gamma a_0}}{(a_0+t)^\beta} f(\frac{1}{a_0+t}).$$

This operator – with $\gamma = 0$ – has been studied extensively in Mayer (1990). It is easy to see from the preceding formula that this operator is compact when restricted to act on a Banach space of functions bounded and analytic on an appropriate domain. A relatively elementary version of the Perron-Frobenius theorem applies and says that the eigenvalue of largest modulus is positive and simple. As might be expected, it can be shown that this eigenvalue is exactly λ. Efficient numerical methods are available for the computation of this principal eigenvalue; this provides an effective method for the numerical computation of the thermodynamic functions of our system.

10 Determining $\mathcal{D}_{q,F}$.

We need the solution to the following elementary (finite!) optimization problem:

Given n *and* F, *find the maximum of* $q_n(a_1, \ldots, a_n)$ *over all configurations with* $a_1 + \cdots + a_n = F$.

To formulate the answer, we need to introduce some notation. We write

$$F = m\,n + r, \quad \text{with } 0 \le r < n;$$

in order that there be any configurations at all, it is necessary that $m \ge 1$, and this will always be assumed in what follows.

Proposition 10.1 *If* $r = 0$, *there is only one maximizing configuration – the one with* $a_i = m$ *for* $1 \le i \le n$. *For* $r > 0$, *the configurations*

$$a_1 = m+1,\ a_2 = \cdots = a_{n-r+1} = m,\ a_{n-r+2} = \cdots = a_n = m+1$$

is maximizing, as is its reversal, and there are no others. Thus, there is a unique maximizing configuration for $r = 0$ *and* $r = 2$, *and exactly two maximizing configurations for each other value of* r.

Although this fact must be known, we have seen no trace of it in the literature. The proof we have found is relatively straightforward but neither short nor particularly enlightening, so we will not give it here. We will nevertheless use the result to get a simple description of the right-hand boundary of the domain $\mathcal{D}_{q,p}$ of asymptotically allowed values for $((\log q_n)/n, F_n/n)$.

Corollary 10.2 *Let $n_j \to \infty$ and $F_j \to \infty$, with*

$$\frac{F_j}{n_j} \to p + \alpha, \quad \text{with } p = 1, 2, \dots \text{ and } 0 \le \alpha < 1,$$

and let $H_{j,\max}$ denote the maximum of $\log(q_{n_j})$ over all configurations of length n_j with $a_1 + \cdots + a_{n_j} = F_j$. Then

$$\frac{H_{j,\max}}{n_j} \to (1-\alpha)\log\gamma_p + \alpha \log\gamma_{p+1},$$

where

$$\gamma_p := \frac{1}{2}\left(p + \sqrt{p^2+4}\right).$$

Proof. Let M_p denote the 2×2 matrix

$$M_p := \begin{pmatrix} p & 1 \\ 1 & 0 \end{pmatrix}.$$

Then, if q_j satisfies the recurrence

$$q_{j+1} = p\, q_j + q_{j-1} \quad \text{for } j = n, \dots, n+m-1,$$

we get

$$\begin{pmatrix} q_{n+m} \\ q_{n+m-1} \end{pmatrix} = M_p^m \begin{pmatrix} q_n \\ q_{n-1} \end{pmatrix}.$$

A simple computation shows that the eigenvalues of $M(p)$ are γ_p (as defined above) and $-\gamma_p^{-1}$. We let Φ_p and Ψ_p denote eigenvectors of M_p and its transpose respectively with eigenvalue γ_p; we can take these vectors to have strictly positive entries and to be normalized so that their scalar product is unity. (There is no particular difficulty in writing explicit formulas ...) Then

$$M_p^n = \gamma_p^n \Phi_p \otimes \Psi_p + \mathcal{O}(\gamma_p^{-n}) \quad \text{for } n \to \infty.$$

The case $\alpha = 0$ requires a slightly special argument, and we treat first the contrary case $\alpha > 0$. Then, if we define r_j by

$$F_j = p n_j + r_j, \tag{$*$}$$

we get $r_j \to \infty$ and $n_j - r_j \to \infty$. By Proposition 10.1, and denoting by e_0 the 2-vector (1,0),

$$\begin{aligned} \exp(H_{j,\max}) &= q_n(p+1, \underbrace{p, \dots p}_{n_j - r_j}, \underbrace{p+1, \dots, p+1}_{r_j - 1}) \\ &= (e_0, M_{p+1}^{r_j-1} M_p^{n_j - r_j} M_{p+1} e_0) \\ &= \gamma_{p+1}^{r_j-1} \gamma_p^{n_j - r_j - 1} (e_0, \Phi_{p+1})(\Psi_{p+1}, \Phi_p)(\Psi_p, M_{p+1} e_0) + o(1) \end{aligned}$$

Since the coefficient $(e_0, \Phi_{p+1})(\Psi_{p+1}, \Phi_p)(\Psi_p, M_{p+1}e_0)$ is non-zero, it follows that

$$H_{j,\max} = r_j \log \gamma_{p+1} + (n_j - r_j) \log \gamma_p + \mathcal{O}(1),$$

so, since $r_j/n_j \to \alpha$,

$$\frac{1}{n_j} H_{j,\max} \to (1-\alpha) \log \gamma_p + \alpha \log \gamma_{p+1}$$

as asserted.

For $\alpha = 0$ we can still use $(*)$ to define r_j, but this time all we know is that $r_j/n_j \to 0$. By passing to subsequences, we can reduce to the cases

- $r_j \to \infty$, in which case the above argument works as it stands.
- $r_j \to -\infty$, in which case a straightforward modification of the above argument – replacing $p, p+1$ by $p-1, p$ – works.
- $r_j = r$ independent of j, in which case we write

  $$\exp(H_{j,\max}) = (e_0, M_{p+1} M_p^{n_j - r} M_{p+1}^{r-1} e_0)$$

  and argue as before.

□

It follows easily from the preceding corollary that the intersection of $\mathcal{D}_{q,P}$ with the horizontal line $f = p + \alpha$ is the interval $(e_{\min}, (1-\alpha) \log \gamma_p + \alpha \log \gamma_{p+1})$ In other words:

Proposition 10.3 *The right-hand boundary of $\mathcal{D}_{q,P}$ is the polygonal arc consisting of the segments joining $(\log \gamma_p, p)$ to $(\log \gamma_{p+1}, p+1)$, for $p = 1, 2, \ldots$.*

References

Cornfeld, I. P., Fomin, S. V., and Sinai, Y. G. (1982). *Ergodic Theory*. Springer-Verlag, New York/Heidelberg/Berlin.

Dixon, J. D. (1970). The number of steps in the euclidean algorithm. *J. Number Theory*, **2**, 414–422.

Hardy, G. H. and Wright, E. M. (1960). *An Introduction to the Theory of Numbers*. Clarendon Press, Oxford, UK, fourth edition.

Kim, S. and Ostlund, S. (1989). Universal scaling in circle maps. *Physica D*, **39**, 365–392.

Knuth, D. E. (1981). *Seminumerical Algorithms*, volume 2 of *The Art of Computer Programming*. Addison-Wesley, Reading, MA, USA, second edition.

Lanford, O. E., III (1973). Entropy and equilibrium states in classical statistical mechanics. In A. Lenard, editor, *Statistical Mechanics and Mathematical Problems, Battelle Seattle 1971 Rencontres*, volume 20 of *Springer Lecture Notes in Physics*, pages 1–113, Berlin/Heidelberg/New York. Springer-Verlag.

Mayer, D. (1990). On the thermodynamic formalism for the Gauss map. *Comm. Math. Phys.*, **130**, 311–333.

Perron, O. (1954). *Die Lehre von den Kettenbrücken*, volume 1. B. G. Teubner, Stuttgart, third edition.

Roberts, J. (1977). *Elementary Number Theory: A Problem Oriented Approach*. The MIT Press, Cambridge, MA.

Ruedin, L. (1994). *Statistical Mechanical Methods and Continued Fractions*. Ph.D. thesis, Mathematics Department, ETH-Zürich.

Schrader, R. (1970). Ground states in classical lattice systems with hard core. *Comm. Math. Phys.*, **16**, 247–264.

Simon, B. (1993). *The Statistical Mechanics of Lattice Gases*, volume 1. Princeton University Press, Princeton, NJ, USA.

Informal Remarks on the Orbit Structure of Discrete Approximations to Chaotic Maps

Oscar E. Lanford III

CONTENTS

We report the results of some computer experiments on the orbit structure of the discrete maps on a finite set which arise when an expanding map of the circle is iterated "naively" on the computer. We also comment on what mathematical questions ought to be answered in order to account for the reliability in practice of orbit following on the computer as an indicator of the ergodic properties of the underlying map.

1. INTRODUCTION

It is a fact of experience that computer simulations — of a relatively naive sort — are generally fairly reliable indicators of the properties of concrete dynamical systems. In the interest of brevity, let me explain what I mean by giving an explicit example, and leave it to the reader to think about generalizations. Consider the mapping

$$x \mapsto f(x) := 2x + 0.5x(1-x) \pmod 1, \quad \text{for } 0 \le x \le 1. \tag{1–1}$$

It is perhaps better to think of this mapping as acting on the unit interval with endpoints identified, i.e., on the circle. Note that $f'(x) \ge 1.5$ everywhere, so f is strictly expanding in a particularly clean and simple sense. As a consequence of expansivity, this mapping has about the best imaginable ergodic properties:

- It admits a unique invariant measure μ equivalent to Lebesgue measure.
- The abstract dynamical system (f, μ) is ergodic and in fact isomorphic to a Bernoulli shift.
- A central limit theorem holds, etc.

1058-6458/1998 $0.50 per page

(One respect in which our example is less than optimal is that, when regarded as acting on the circle, it has a discontinuous first derivative at the origin. This could be fixed by studying instead, say, $x \mapsto 2x + \alpha \sin 2\pi x \pmod 1$ with $|\alpha| < 1/(2\pi)$. We chose the above example because it is cheap to compute and because the effects of round-off error are relatively easy to analyze.)

One consequence of the ergodicity of f relative to μ is that, for Lebesgue almost all initial points x in the unit interval, the corresponding orbit $f^n(x)$ is asymptotically distributed over the unit interval according to μ, i.e.,

$$\lim_{m\to\infty} \frac{1}{m} \sum_{n=0}^{m-1} \varphi(f^n(x)) = \int \varphi(t)\, \mu(dt)$$

for all well-behaved functions φ. (This assertion is simply a clever reading of the pointwise ergodic theorem applied to (f, μ) together with the fact that μ-almost all and Lebesgue-almost all are the same.) Intuitively: If we choose an initial point at random, then we should be very surprised if the corresponding orbit fails to distribute itself as indicated.

This mathematically rigorous result leads one to expect something quite similar to happen when the mapping is iterated *on the computer*. More formally:

Choose an initial point which is not too special (the points 0 and 1, for example, are obviously not typical); compute numerically a few tens of millions of points on the orbit of the point in question; divide the working interval into a few hundred equal subintervals; count the number of points of the computed orbit lying in each of these intervals; and plot the resulting histogram. It would be surprising if this process failed to produce a graph looking very much like that of the density of the absolutely continuous invariant measure.

I want to begin by making the very simple and general point that, *reasonable as this expectation is, it is not so obvious how to* ***prove*** *that it is correct.* The reason is that, because of the expansivity of the mapping, the growth of round-off error normally means that the computed orbit will remain near the true orbit with the chosen initial condition only for a relatively small number of steps — typically, of the order of the number of bits of precision with which the calculation is done. It is true that the above mapping — like many "chaotic" mappings — satisfies a *shadowing theorem* which ensures that the computed orbit stays near to some true orbit over arbitrarily large numbers of steps. The flaw in this idea as an explanation of the behavior of computed orbits is that the shadowing theorem says that the computed orbit approximates *some* true orbit, but not necessarily that it approximates *a typical* one. In fact, the simple example $x \mapsto 2x \pmod 1$ shows that computed orbits do not always approximate typical exact orbits (and also makes clear that the expectation expressed in the preceding paragraph is not always fulfilled).

It appears to me that the precise formulation of these questions will require the setting-up of a limiting regime in which precision of calculation and number of steps of iteration both go to infinity, with relations between the rates. In fact: In a very general way, this problem reminds me quite a lot of the notoriously difficult one of deriving non-equilibrium statistical mechanics from atomic physics. As with statistical mechanics, there are two very different length scales, a macroscopic one on which the state space looks like a continuum and the mapping smooth, and a microscopic one on which the state space looks like a discrete set of points and the mapping has a certain amount of jaggedness. This suggests the discouraging possibility that this problem may be as hard of that of non-equilibrium statistical mechanics. As with statistical mechanics, the problem can probably be made *much* easier by the judicious introduction of a stochastic element in the microscopic evolution. As in statistical mechanics, I think this is cheating: For me, a satisfactory solution will have to take seriously the fact that computer iteration is perfectly deterministic.

2. EXPERIMENTAL RESULTS

But these are issues of general philosophy, and I don't want to discuss them in detail here. I mention them only as a preface to saying that what I do want to talk about is an experimental study of an *inappropriate* limiting regime — one in which the number of iterations is *too large* relative to the precision of calculation. The question addressed is the following: The exact mathematical problem concerns iterating a smooth mapping on an interval. The computer, working with fixed finite precision, is able to represent finitely many points in the interval in question. It is probably good, for purposes of orientation, to think of the case where the representable points are uniformly spaced in the interval. The true smooth mapping is then *approximated* by a discretized mapping, sending the finite set of representable points in the interval to itself. Describing the discretized mapping exactly is usually complicated, but it is *roughly* the mapping obtained by applying the exact smooth mapping to each of the discrete representable points and "rounding" the result to the nearest representable point. (The reason why this simple description is not quite realistic is that, in practice, intermediate quantities, and not just the final result, undergo rounding.) However the discretization is done, the upshot is that what is really iterated on the computer is *a mapping of a finite — albeit large — set to itself.* Every orbit of such a mapping is, trivially, eventually periodic, i.e., eventually lands exactly on a periodic cycle. The question addressed here is the orbit structure of such a discretized map:

- How many periodic cycles are there, and what are their periods?
- How large are their respective basins of attraction, i.e., for each periodic cycle, how many initial points give orbits which eventually land on the cycle in question?

I have done two kinds of experiments to explore this question:

- For relatively coarse discretizations — say about 10^7 representable points — determine the orbit structure completely, i.e., find all the periodic cycles and the exact sizes of their basins of attraction.
- For iteration using ordinary (IEEE-754) double-precision arithmetic — so that the working interval contains of the order of 10^{16} representable points — sample the orbit structure by choosing some number — 1000 in the case reported here — initial points at random and determining the cycles to which they converge.

For purposes of logical simplicity, it seemed mildly advantageous to look at evenly spaced discretizations. For the experiments with double precision, this was accomplished by shifting the working interval from $[0, 1]$ to $[1, 2]$, i.e., the mapping actually iterated was

$$x \mapsto 2x + 0.5(x-1)(2-x) \pmod 1$$

from $[1, 2]$ to itself.

Some representative results results are given on the next page (Table 1).

Many more examples could be given, but those given may serve to illustrate the intriguing character of the results: The outcome proves to be extremely sensitive to the details of the experiment, but the results all have a similar flavor: A relatively small number of cycles attract nearly all orbits, and the lengths of these significant cycles are much larger than one but much smaller than the number of representable points.

3. MODELLING

It seems clear that there are regularities here which ought to be understood. *I know of* ***no*** *ideas which contribute, in my judgment, to a* ***fundamental*** *understanding of these regularities.* There is, on the other hand, a very persuasive idea about how one might *model* them. The idea, which I first heard from D. Ruelle (see Section 7 and the note following it), runs as follows: Since the mapping

$N = 2^{22} = 4{,}194{,}304$ 13 cycles		
period	basin size	
3,864	2,523,929	60%
1,337	538,712	13%
718	513,839	12%
295	238,486	6%
130	203,587	5%
1,338	152,942	4%
297	12,359	0.3%
169	5,056	0.1%
97	3,012	$< 0.1\%$
17	2,346	$< 0.1\%$
6	21	$< 0.1\%$
1	8	$< 0.1\%$
7	7	$< 0.1\%$

$N = 2^{24} = 16{,}777{,}216$ 2 cycles		
period	basin size	
5,300	16,777,214	100%
1	2	$< 0.1\%$

$N = 2^{23} = 8{,}388{,}608$ 7 cycles		
period	basin size	
4,898	5,441,432	65%
1,746	2,946,734	35%
13	205	$< 0.1\%$
6	132	$< 0.1\%$
30	96	$< 0.1\%$
4	8	$< 0.1\%$
1	1	$< 0.1\%$

$N = 2^{24} - 1 = 16{,}777{,}215$ 10 cycles		
period	basin size	
3,081	7,502,907	45%
699	3,047,369	18%
3,469	2,905,844	17%
1,012	2,774,926	17%
563	290,733	2%
2,159	221,294	1%
138	21,610	0.1%
421	12,477	$< 0.1\%$
9	54	$< 0.1\%$
1	1	$< 0.1\%$

$N = 2^{25} = 33{,}554{,}432$ 8 cycles		
period	basin size	
4,094	32,114,650	96%
621	918,519	3%
283	516,985	2%
126	2,937	$< 0.1\%$
6	887	$< 0.1\%$
55	433	$< 0.1\%$
4	20	$< 0.1\%$
1	1	$< 0.1\%$

double precision (sampling) 1000 initial points 7 cycles found		
period	"basin size"	
27,627,856	517	52%
88,201,822	320	32%
4,206,988	147	15%
4,837,566	17	2%
802,279	8	1%
6,945,337	6	1%
2,808,977	1	0.1%

TABLE 1. Census of cycles for representative discretizations. N is the order of the discretization; thus the representable points are the numbers j/N, with $0 \le j < N$.

is "chaotic," it is reasonable to think of modeling computed orbits by simply choosing the successive points at random. This model only makes sense, however, until some point has been chosen twice; thereafter, the fundamentally deterministic character of the mapping can no longer be neglected and the future of the orbit is unambiguously determined. An elementary calculation shows that the probability that there is *no* repetition in a set of n points chosen independently from a population of N (with equal weights) is about

$$e^{-n^2/(2N)},$$

provided that N is large and n not too much larger than $\sqrt{N}$. Loosely: The number of steps before the first repetition is typically of the order of $\sqrt{N}$. This rule of thumb is roughly consistent with the experimental results cited above. (A moment's thought shows that the above computation of distribution of "first-repeat times" is equivalent to computing the first-repeat time for a randomly chosen map. Hence, one speaks of the "random-map model" for computing periods of cycles, etc.)

These remarks also permit me to be a little more specific about what I think the right limiting regime for theorem proving about the reliability of numerical experiments of the sort described earlier should be: One should look at a discretization to N points,

and at time averages over m steps, with m and N both large but with

$$\log N \ll m \ll \sqrt{N}.$$

The first $\ll$ allows the computed orbit to deviate macroscopically from the true one over most of its length; the second is in any case usually satisfied in practice and ought to mean that the times considered are short enough so that the effects of the strict finiteness of the space of states are not important. In fact: it might be prudent to replace the second $\ll$ by the stronger condition

$$\log m \ll \log N,$$

but it isn't so clear that this condition is satisfied in practice.

4. SPATIAL DISTRIBUTION OF CYCLES

As noted at the beginning, almost all the orbits of the mathematically exact mapping distribute themselves asymptotically over the working interval according to the unique invariant probability measure μ which is absolutely continuous with respect to Lebesgue measure. On the other hand, a very long computed orbit simply runs many times around whatever periodic cycle it lands on, i.e., has the same asymptotic distribution as that cycle. It is therefore interesting to know whether the periodic cycles are distributed according to the absolutely continuous invariant measure. This can of course only hold in an approximate sense, since a periodic cycle is only a finite set of points. Somewhat surprisingly, it does appear that, at least in the example we are considering, the periodic cycles do approximate the absolutely continuous invariant measure quite well. To be somewhat more concrete: I looked, from this point of view, only at the double-precision discretization. The working interval was partitioned into 200 equal subintervals and, for each of the seven cycles listed in Table 1, the points of the cycle in each subinterval were counted. The resulting histogram for the first cycle — the one which seems to attract a majority of the orbits — is shown in Figure 1. The same

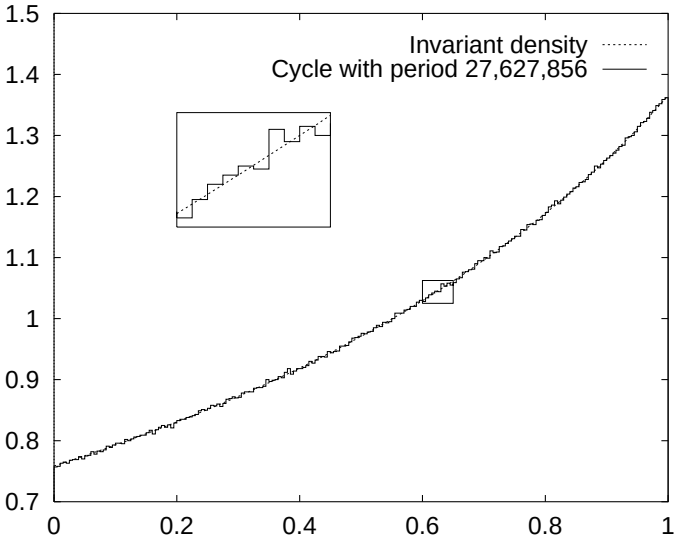


FIGURE 1. Histogram of the main cycle and density of the absolutely continuous invariant measure.

figure shows the density of the absolutely continuous invariant measure, but the agreement is so good that it is nearly impossible to distinguish the two plots (except when magnified, as in the inset box). The results for the other cycles (not shown) are similar, with the following systematic variation: Very simple ideas lead to the surmise that the occupation numbers n_i of the various intervals should show fluctuations of the order of $\sqrt{n_i}$. The n_i's are roughly proportional to the period of the cycle, and the periods of the cycles vary by about a factor of 100; thus, it is to be expected that the histograms for cycles with relatively small periods look significantly noisier than those for cycles with relatively large periods. This is indeed what happens.

To get a more quantitative measure of the agreement between the distribution of the cycles and the invariant measure, we compute for each cycle the χ^2-statistic,

$$\chi^2 = \sum_{i=1}^{200} \frac{(n_i - \bar{n}_i)^2}{\bar{n}_i}, \qquad (4\text{–}1)$$

where n_i is the number of points on the cycle lying in the ith subinterval, and $\bar{n}_i$ the "expected" number, i.e., the period of the orbit multiplied by the probability assigned to the ith subinterval by the

period	χ^2
27,627,856	170.480
88,201,822	229.728
4,206,988	122.258
4,837,566	264.920
802,279	184.270
6,945,337	215.295
2,808,977	197.632

TABLE 2. The χ^2-statistic as defined in (4–1) for each of the seven cycles found for the double-precision discretization.

absolutely continuous invariant probability measure. The results are as shown in Table 2. For comparison: For sample of p points chosen independently according to the invariant probability measure, χ^2 has probability about 0.05 of being smaller than 167 and also probability about 0.05 of being larger than 233.[1] The third and fourth orbits look a little suspect at first glance; the probability that χ^2 is < 123 (respectively > 264) is only about 5.4×10^{-6} (respectively 1.4×10^{-3}). On reflection, however, it is clear that these probabilities should be taken with a grain of salt; it would be more convincing to compare the fluctuations of the distributions of periodic cycles about the invariant measure with the typical sizes of fluctuations of long segments of true orbits around this same measure. The latter fluctuations are indeed Gaussian – i.e., the system satisfies a central limit theorem — but the covariance (presumably) does not have the simple form corresponding to an independent choice of p points (and furthermore does not seem to be easy to compute).

5. ANOTHER MAPPING

To show that the above is not the whole story, we present the results of one other experiment — a sampling study in double precision of a discretization of the mapping

$$x \mapsto 1 - 2x^2 \quad \text{on } [-1, 1]. \tag{5–1}$$

This mapping also has excellent ergodic properties but in a more subtle and unstable way than the previous example. (The precise discretization studied is obtained by first exploiting evenness to fold the interval $[-1, 0]$ onto $[0, 1]$, i.e. replacing (5–1) by

$$x \mapsto |1 - 2x^2| \quad \text{on } [0, 1]. \tag{5–2}$$

It is not difficult to see that the folded mapping has the same set of periods as the original one. The working interval is then shifted from $[0, 1]$ to $[1, 2]$ by translation, and the iteration of the translated folded mapping is programmed in a straightforward way.) Out of 1000 randomly chosen initial points,

- 890 — the overwhelming majority — converged to the fixed point corresponding to the fixed point -1 in the original representation (5–1);
- 108 converged to a cycle of period 3,490,273;
- the remaining 2 converged to a cycle of period 1,107,319.

Thus, in this case at least, the very long-term behavior of numerical orbits is, for a substantial fraction of initial points, in flagrant disagreement with the true behavior of typical orbits of the original smooth mapping.

6. SOME DETAILS

As should now be apparent, the orbit structure of a discretized map tends to depend sensitively on the details of the discretization. One consequence is that an attempt to reproduce the reported orbit structures is not likely to give the same results unless care is taken to use *exactly* the same discretization. For the relatively low-precision "artificial" discretization, it is not difficult to describe the discretization precisely. We denote by N the

[1] As is customary, the computation of these probabilities is done under the assumption that the central limit theorem gives a "sufficiently accurate" description of the fluctuations of the n_i about their mean values $\mu_i p$. This assumption should certainly hold very well in this situation, but it seems to be hard to estimate the error reliably.

number of points in the discretized working interval, so that the points themselves are the

$$x_j = \frac{j}{N}, \quad \text{with } 0 \le j \le N-1.$$

The discretized map we study can be described as the result of the following three-step procedure:

1. Apply the exact mapping (1–1) to x_j. The result lies between 0 and $2 - 1/N$.
2. If the result of step 1 is $\ge 1 - 1/2N$, subtract 1; otherwise, leave it unchanged. In either case, the result in now in $[-1/2N,\ 1-1/2N)$.
3. Round the result of step 2 to the nearest number of the form k/N. If it lies exactly halfway between two of these lattice points, choose the one with k even.

(We emphasize that this is an "implementation-independent"characterization of the the discretized map, not an algorithm adapted to computing it.) The prescription in step 2 to reduce to the interval $[-1/2N,\ 1-1/2N)$, rather than the more natural-seeming $[0, 1]$, serves to ensure that we round to a number of the form k/N with $k \le N-1$. Note that the plausible alternative of rounding *before* reducing to the interval $[0, 1)$ gives the same results for N even, but not necessarily for N odd, because of rounding-to-even in the case of a tie.

For the sampling experiments in double precision, it is not so easy to give a short precise specification of how the mapping is discretized. The way the experiment was actually done was to write reasonably straightforward C-language code for the mapping and pass it through a compiler. The C code was:

```
double f (double x) {
   double w;
   w = 2*(x-1)+0.5*(x-1.0)*(2.0-x);
   if (w < 1.0) w += 1.0;
   return(w);
}
```

The compiler used for the experiment reported here was the GNU C compiler, version 2.7.0, running on an Intel Pentium under the Linux operating system. "Unoptimized" compilation was requested. Because the Pentium's internal floating point registers provide 64 bit precision rather than the 52 bits of double precision numbers, the exact discretization algorithm depends on which intermediate quantities are kept in registers and which are rounded to double precision in order to be stored.[2] I have not attempted to sort out in all detail what was really happening in the experiment performed; the objective was simply to imitate how such a computation would be done in practice. It may be worth noting that the above code produces *different* results when compiled and run on a Sparc (again using the GNU C compiler).

7. A FEW BIBLIOGRAPHIC REMARKS

As noted earlier, I learned from D. Ruelle the idea that the "random iteration" described earlier might be a sensible way to model the structure of periodic cycles of chaotic maps (and hence that typical periods should be of the order of the square root of the number of accessible states). Ruelle proposed this idea to account for some anomalies in numerical experiments performed by Y. Levy, and the idea appears in [Levy 1982]. These ideas were developed more generally — and a technical emendation proposed to Levy's Ansatz — in [Grebogi et al. 1988]. Recently, a group at Deakin University and the University of Queensland has been studying discretized mappings systematically. One direction they have pursued is the quantitative study of how well the random-maps model predicts the orbit structure of discretizations. They have also studied ways of improving on the random-maps model to take into account relevant properties of the map being discretized, notably the presence of

[2] There is in fact yet another complication. In its default mode, the Intel floating point hardware performs floating point operations by first rounding to 64 bits, then rounding *that* to 52 when it is stored. Because of the round-to-even tie-breaking rule, this is not the same as rounding directly to 52 bits. Thus, even if all intermediate quantities are stored, the results will not always be the same as with pure 52-bit arithmetic.

critical points. See, for example, [Diamond et al. 1996]; the extensive bibliography of this article also provides a more thorough overview of prior work than can be given here.

NOTE ADDED

I am indebted to the anonymous referee for calling to my attention the very relevant work of T. Erber and his collaborators, in which, among many other things, the idea of modelling the orbit structure of a chaotic map by that of a random map appears prior to the aforementioned work of Levy. In [Erber et al. 1979], orbit structure of discretizations of $x \mapsto 2 - x^2$ and its dependence on the precision of the discretization are studied in very much the same spirit as the experiments reported here. (Curiously, the phenomenon of collapse of a significant fraction of the orbits onto the fixed point — at $x = -2$ in this way of writing the mapping — which was so prominent in our experiments turned up only rarely in those reported in this paper.) This paper and the related one [Erber et al. 1983] present a wealth of intriguing heuristic ideas bearing on the modelling of the orbit structure of these discretizations.

ACKNOWLEDGEMENTS

The programming on which this report is based was mostly done during an extended stay at the Institut des Hautes Études Scientifiques in Bures-sur-Yvette, France, in the Fall and Winter of 1993–94. I thank the IHES and especially Professor Marcel Berger, its Director at that time, for their hospitality. I am also grateful to Henri Epstein for many stimulating discussions and for encouragement.

REFERENCES

[Diamond et al. 1996] P. Diamond, A. Klemm, P. Kloeden, and A. Pokrovskii, "Basin of attraction of cycles of discretizations of dynamical systems with SRB invariant measures", *J. Statist. Phys.* **84** (1996), 713–733.

[Erber et al. 1979] T. Erber, P. Everett, and P. W. Johnson, "The simulation of random processes on digital computers with Čebyšev mixing transformations", *J. Comput. Phys.* **32** (1979), 168–211.

[Erber et al. 1983] T. Erber, T. M. Rynne, W. F. Darsow, and M. J. Frank, "The simulation of random processes on digital computers: unavoidable order", *J. Comput. Phys.* **49** (1983), 394–419.

[Grebogi et al. 1988] C. Grebogi, E. Ott, and J. A. Yorke, "Roundoff-induced periodicity and the correlation dimension of chaotic attractors", *Phys. Rev. A* **38** (1988), 3688–3692.

[Levy 1982] Y. E. Levy, "Some remarks about computer studies of dynamical systems", *Phys. Lett. A* **88** (1982), 1–3.

Oscar E. Lanford III, Mathematics Department, ETH-Zürich, 8092 Zürich, Switzerland (oscar.lanford@math.ethz.ch, http://www.math.ethz.ch/~lanford)

Received December 3, 1997; accepted in revised form June 22, 1998