Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

The CWI is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

CWI Syllabus 17

Mark Kac seminar on probability and physics Syllabus 1985-1987

edited by

F. den Hollander
H. Maassen

Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

ISBN 90 6196 350 8
NUGI-code: 811

Printed in the Netherlands

PREFACE

This syllabus is a bundle of reports of lectures delivered at the "Mark Kac seminar on probability and physics" during the academic years 1985-1987. This seminar is a monthly meeting in Amsterdam, held between probabilists and statistical physicists who discuss topics in their common field of interest. As such this booklet shows a cross-section of the activities in the Netherlands in this area of interaction.

Two reports deal with the sequences of lectures delivered by prof. G. Grimmett from Bristol on interacting-particle systems and on random networks, and by prof. J.T. Lewis from Dublin on large deviation techniques in statistical mechanics. The other reports are accounts of single lectures and treat the reader to a variety of specialized topics.

We would like to thank the speakers for their contribution to this syllabus.

Frank den Hollander (Delft)
Hans Maassen (Nijmegen)

November 1987

PROGRAM 1985/1986:

27-9 E. Hendriks (Amsterdam)

On a model of spatially homogeneous polymerization with a phase transition

E. Thomas (Groningen)

Positive forms on nuclear *-algebras and their integral representations

25-10 H. van Beijeren (Utrecht)

How does one characterize macroscopic diffusion?

J. Villain (Jülich)

Random fields in the Ising-model

29-11 H. Hilhorst (Leiden)

Some recent developments and problems in the theory of spin glasses

J. Naudts (Antwerpen)

Distributions with the Lee-Yang property

31-1 G. Grimmett (Bristol)

Interacting-particle systems and random networks I

E. Aarts (Eindhoven)

Statistical cooling: a general approach to combinatorial optimization problems

28-2 G. Grimmett (Bristol)

Interacting-particle systems and random networks II

R. Kuik (Groningen)

On a noncommutative approach to equilibrium states in the q-state Potts model

25-4 G. Grimmett (Bristol)
Interacting-particle systems and random networks III

C. Scheffer and M. Dekking (Delft)
Random fields on trees

30-5 E. Lieb (Princeton)
A model for crystallization: a variation on the Hubbard model

G. Grimmett (Bristol)
Interacting-particle systems and random networks IV

PROGRAM 1986/1987:

3-10 L. van Hemmen (Heidelberg)
The theory of large deviations in statistical mechanics:
I. Theoretical considerations

A. Gandolfi (Delft)
On the uniqueness of the infinite cluster in percolation

7-11 L. van Hemmen (Heidelberg)
The theory of large deviations in statistical mechanics:
II. Neural nets and spin glasses

A. Verbeure (Leuven)
Applications of dynamical semigroups

5-12 T. Dorlas (Groningen)
Renormalization and the continuum limit

W. Vervaat (Nijmegen)
Algebraic duality of Markov processes

6-2 M. van Lambalgen (Amsterdam)

Complexity, randomness and unpredictability

J. Lewis (Dublin)

The large deviation principle in statistical mechanics and its application to the Boson gas I

6-3 J. Lewis (Dublin)

The large deviation principle in statistical mechanics and its application to the Boson gas II

H. Dekker (Den Haag)

Dynamics of the dissipative two-state system

3-4 Ph. de Smedt (Leuven)

The quantum Langevin equation and its stationary state

J. Lewis (Dublin)

The large deviation principle in statistical mechanics and its application to the Boson gas III

8-5 A. Greven (Heidelberg)

Large systems with locally interacting components

J. Lewis (Dublin)

The large deviation principle in statistical mechanics and its application to the Boson gas IV

ON A MODEL OF SPATIALLY HOMOGENEOUS POLYMERIZATION WITH A PHASE TRANSITION

E. Hendriks, Koninklijke/Shell-Laboratorium, Amsterdam

This contribution is about a simple model describing condensation polymerization[1] or coagulation[2]. Consider a fixed volume containing M basic building blocks (monomers), which combine to produce clusters[3-6]. We assume spatial homogeneity, such that the state of the system can be described by the vector variable $x = (x_1, x_2, \ldots, x_M)$, where x_k is the number of clusters containing k basic units (k-mers). Conservation of mass implies $\sum_k kx_k = M$, for each possible state x. Under a "detailed balance" condition for the transition rates between the different states of the system[5], the equilibrium probability of the state x equals

$$(1) \qquad P^{eq}(x) = Z^{-1} \prod_{1 \le k \le M} (a_k M/q)^{x_k} / x_k! \ ,$$

where q is a control parameter (like temperature) and a_k is a degeneracy factor. The degeneracy factor may depend e.g. on the structure of the monomers and on the degree in which "steric hindrance" limits the construction of a k-mer out of its constituent monomers. Eq.(1) can be derived from the master equation[5] or from statistical mechanical considerations. The numbers a_k are assumed to show the following asymptotic behaviour for large k:

$$(2) \qquad a_k \simeq Ak^{-\beta-1} z_c^{-k} \quad (k \to \infty,\ 1 < \beta < 2)$$

and completely characterize the model. The so-called classical model of polymerization, introduced by Stockmayer[4], is the special case where $a_k k!$ just equals the number of ways a k-mer can be built out of its constituents ($\beta = 3/2$).

In the range $1 < \beta < 2$ assumed here, the model shows a phase transition in the thermodynamic limit $M \to \infty$, q fixed. There is a critical value q_c such that for $q > q_c$ a finite part of the total mass is in a single large cluster (the gel), which is not present for $q < q_c$. This can be shown by locating the most probable state (using Lagrange's multipliers method[4,5]) and investigating its behaviour for $M \to \infty$, or by studying the ensemble averaged size distribution

$$(3) \qquad \langle x_k \rangle = \sum_{\{x\}} P^{eq}(x)\, x_k \ ,$$

where the summation is over all possible states. From eqs. (1) and (3) and the mass conservation requirement, the following integral representation for $\langle x_k \rangle$ can be derived:

(4) $$\langle x_k \rangle = M a_k I_k / q I_0$$

$$I_k = \oint \frac{dz}{2\pi i} \exp\{ (M/q)F(z) + (k-M-1) \log z \} ,$$

where

(5) $$F(z) = \sum_{k \geq 1} a_k z^k$$

is the generating function of the a_k. The integration path in eq.(4) must be chosen such that it encircles the origin of the complex plane once, counterclockwise, and does not enclose singularities of F(z). From eq.(4) a recursion relation for $\langle x_k \rangle$ follows, whose numerically obtained solution is displayed in fig.1, for two values of M. The presence of the "bump" indicates the occurrence of a single large molecule: the gel. (Note that for $k > M/2$, $\langle x_k \rangle$ is simply the probability of occurrence of a cluster of size k.) The distribution is bimodal: it consists of a sol part and a gel part (the particular value of q used here exceeds q_c). The average number of molecules present in the gel phase equals 1.

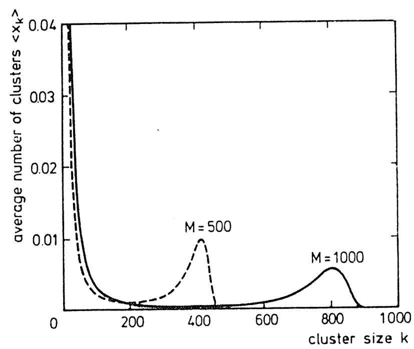


fig. 1

Through application of the method of stationary phase[7] to the integral in eq.(4), the following results are obtained[3,5]:

(1) The essential properties of the system are all embodied in the generating function $F(z)$ of the degeneracy factor a_k for k-mers. Of particular importance for the properties of the phase transition is the asymptotic behaviour of this factor for $k \to \infty$.

(2) The control parameter q plays a role analogous to that of temperature in statistical mechanics. It goes through a critical value $q_c = z_c F'(z_c)$ when the system passes from the sol to the gel phase (z_c is the radius of convergence of $F(z)$).

(3) In the sol phase ($q < q_c$), the ensemble averaged cluster size distribution decays exponentially for large values of k (except for an algebraic factor) and is determined by the numbers a_k and by the "fugacity" z_* that is the solution of the equation $zF'(z) = q$ closest to the origin:

(6) $$\lim_{M\to\infty} \langle x_k \rangle / M = a_k z_*^k / q.$$

At $q = q_c$, the exponential factor disappears ($z_* = z_c$) and the decay becomes purely algebraic. Critical exponents of the phase transition in this model are derived in ref.5.

(4) In the gel phase ($q > q_c$), the ensemble averaged size distribution has a clearly bimodal character. It is the superposition of a distribution of sol species, which is still given by eq.(6) (with $z_* = z_c$), and of a distribution of gel species, which is characterized by a maximum near $k = Mg$, where $g = 1 - q_c/q$ is the gel fraction of the system (for $M \to \infty$). Hence the gel consists of a single large cluster.

(5) For $M \to \infty$, the distribution of gel species is described by an asymptotic distribution depending only on the asymptotic behaviour of a_k. The width of this distribution is proportional to $M^{1/\beta}$, and the distribution is such that the scaled random variable $y = (q/AM)^{1/\beta}(k_g - Mg)$, with k_g the size of the largest cluster, has limiting distribution $G(y)$ which depends only on β and is found from:

(7) $$\lim_{\substack{M\to\infty \\ y \text{ fixed} \\ k = Mg + M^{1/\beta} y}} \langle x_k \rangle = (q/AM)^{1/\beta}\, G\big((q/A)^{1/\beta} y \big) .$$

(Recall that for k/M large, $\langle x_k \rangle = P^{eq}(k_g = k)$.) An explicit integral representation for the scaling function G(y) is derived in ref.3. It is essentially the Fourier transform of the function $h(s)=\exp\{\Gamma(-\beta)(-is)^{\beta}\}$. It is highly nonsymmetric and its second moment diverges (see fig. 2).

The above results are a generalization of similar results in ref.8, valid for the classical Stockmayer model.

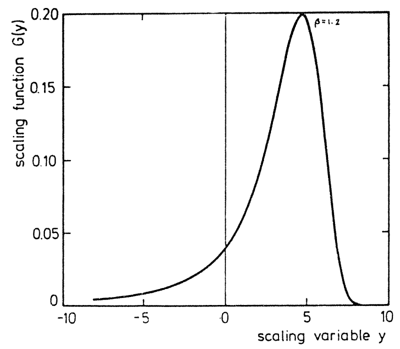


fig. 2

The equilibrium probability distribution in eq.(1) is the stationary solution of the master equation[5,9]. From this equation, by taking the limit $M\to\infty$, the Smoluchowski coagulation equation[6,9] can be derived (under certain conditions). A complete proof, valid for all choices of the transition rates in the master equation, has not yet been given. The coagulation equation is the subject of much recent research in the field of growth phenomena[2,10,11]. The occurrence of a (dynamical) phase transition and the values of corresponding critical exponents can be derived from this equation, but not the detailed shape of the gel distribution (which requires $M<\infty$).

References:

(1) P. J. Flory, "Principles of Polymer Chemistry", Cornell University Press, New York, 1953.

(2) R. L. Drake, in: "Topics in Current Aerosol Research (part 2)", G. Hidy and J. Brook (eds.), Pergamon Press, Oxford, 1972.

(3) E. M. Hendriks, J. Phys. A18 (1985) L579.

(4) W. H. Stockmayer, J. Chem. Phys. 11 (1934) 45.

(5) E. M. Hendriks, Z. Phys. B57 (1984) 307.
(6) J. L. Spouge, Macromolecules 16 (1983) 121.
(7) R. B. Dingle, "Asymptotic Expansions, their derivation and interpretation", Academic Press, London, 1973.
(8) E. Donoghue and J. H. Gibbs, J. Chem. Phys. 70 (1979) 2346.
E. Donoghue, J. Chem. Phys. 77 (1982) 4236.
(9) E. M. Hendriks, J. L. Spouge, M. Eibl and M. Schreckenberg, Z. Phys. B58 (1985) 219.
(10) A. Lushnikov, J. Coll. Interf. Sc. 65 (1978) 276, Izv. Acad. Sci. USSR Atmos. Oceanic Phys. 14 (1978) 738 (these two papers deal with the connection between Smoluchowski's equation and the master equation).
(11) R. M. Ziff, J. Stat. Phys. 23 (1980) 241.
E. M. Hendriks, M. E. Ernst and R. M. Ziff, J. Stat. Phys. 31 (1983) 363.
F. Leyvraz and H. R. Tschudi, J. Phys. A14 (1981) 3389, A15 (1982) 1951.
P. C. J. Van Dongen and M. H. Ernst, J. Stat. Phys. 37 (1984) 301.

POSITIVE FORMS ON NUCLEAR $\star$-ALGEBRAS AND THEIR INTEGRAL REPRESENTATIONS

E. Thomas, Mathematisch Instituut, RU Groningen

Both in Harmonic analysis and in Field theory one encounters nuclear $\star$-algebras, positive forms on them, and integral representations of these forms. We attempt a synthesis of these two cases, which up to now seem to have been developed almost entirely disjointly. In such an attempt the analogy with positive states on C$\star$-algebras (which are nuclear in the sense intended here only if they are finite dimensional) obviously imposes itself; for the present summary this aspect will be left aside however.

Let A be an algebra over C, equipped with an involution $\phi \to \phi^\star$ and a locally convex Hausdorff topology with respect to which it is a nuclear space in the sense of Grothendieck[5], such that the product is separately continuous, and the involution continuous; briefly: a nuclear $\star$-algebra.

The two most important examples for the present purpose are:

Harmonic analysis: Let G be a unimodular Lie group, and let $A = D(G) = C_c^\infty(G)$ be the space of test functions equipped with the usual inductive limit topology, with the convolution product

(1) $\quad \phi \cdot \psi(g) = \int \phi(x)\psi(x^{-1}g)dx$

and with the involution:

(2) $\quad \phi^\star(g) = \overline{\phi(g^{-1})}.$

(This is the global approach. In the infinitesimal approach A would be the universal enveloping algebra of a Lie algebra.)

Field Theory:[6,10] Here A is equal to the tensor algebra $S = C + \sum_{\nu=1}^{\infty} S(R^{4\nu})$ composed of sequences $(\phi_\nu)_{\nu\geq 0}$, having only a finite number of terms different from 0, such that ϕ_0 is a complex number, and ϕ_ν is, for $\nu \geq 1$, a function belonging to the Schwartz space of infinitely differentiable functions on $R^{4\nu}$ which, together with all their derivatives, are rapidly decreasing at infinity. The topology is the direct sum topology. The product is defined as follows:

(3) $$(\phi\cdot\psi)_n(x_1,\dots,x_n) = \sum_{\mu+\nu=n} \phi_n(x_1,\dots,x_\mu)\psi_\nu(x_{\mu+1},\dots,x_n)$$

and the involution is defined by the formula:

(4) $$(\phi^\star)_\nu(x_1,\dots,x_\nu) = \overline{\phi(x_\nu,\dots,x_1)}.$$

The term tensor-algebra is appropriate because $S(R^{4\nu})$ can be identified with the completed tensor product of ν copies of $S(R^4)$.

In both examples A is a nuclear $\mathcal{LF}$-space satisfying the further condition

(C) The map p from the inductive tensor product $A\overline{\otimes}A$ to A, defined by $p(\phi\otimes\psi) = \phi\cdot\psi$, is a surjection.

and having an identity 1 or an approximate identity $(\chi_n)_{n\in N}$ such that $\lim_n \chi_n\cdot\phi = \phi$ for all $\phi\in A$.

A positive functional on a $\star$-algebra A is a continuous linear form ω on A such that $\omega(\phi\cdot\phi^\star)\geq 0$ for all $\phi \in A$. The set A'_+ of all positive forms is a closed convex cone in the dual space A'.

A functional $\omega \in A'$ is called central (or abelian) if $\omega(\phi\cdot\psi) = \omega(\psi\cdot\phi)$ for all $\phi,\psi \in A$. The central elements of A'_+ obviously form a closed convex sub-cone. Other closed sub-cones of A'_+ may play an important role.

An element ω of a convex cone Γ is called extremal, briefly $\omega \in \mathrm{ext}(\Gamma)$, if every decomposition $\omega = \omega_1 + \omega_2$, with $\omega_i \in \Gamma$, is trivial, in as much as ω_i is proportional to ω, for $i = 1,2$.

Theorem 1. Let A be a nuclear $\star$-algebra which is an $\mathcal{LF}$-space satisfying the condition (C). Let Γ be a closed convex sub-cone of A'_+.
(A) Every element $\underline{\omega}$ in Γ has at least one integral representation by means of extreme elements: $\underline{\omega} = \int_{\mathrm{ext}(\Gamma)} \omega dm(\omega)$.
(B) The element $\underline{\omega} \in \Gamma$ has a unique integral representation by means of extremal elements if and only if the face $\Gamma(\underline{\omega})$ generated in Γ by $\underline{\omega}$ is a lattice with respect to its own order.

The proof of this theorem can be given by showing that Γ has bounded order intervals, and then applying the nuclear integral representation theorem[11,12].

In the case where A = D(G), positive forms are normally called positive definite distributions. The sub-cone of all central positive definite distributions plays a particularly important role in Harmonic analysis. It is known to be a lattice cone. Thus one has a theorem analogous to the theorem of Bochner-Schwartz in the case of R^n: Every central distribution of positive type has a unique integral representation

(5) $T = \int \chi_\pi dm(\pi)$

by means of extremal elements. In the case of type I Lie groups (e.g. semi-simple or nilpotent) these can, after appropriate normalisation, be identified with the characters of irreducible unitary representations. In the particular case where $T=\delta$ the measure m is the Plancherel measure, and (5) becomes the Plancherel decomposition, or "inversion formula":

(5') $\delta = \int \chi_\pi d\pi.$

Another type of sub-cone is the cone of H-invariant positive definite distributions, H being a closed subgroup of G. These are important in the harmonic analysis on the homogeneous space G/H. The cone of positive definite H-invariant distributions is a lattice cone if and only if (G,H) is a generalised Gelfand pair[(13)]. In that case the extreme generators are the spherical distributions.

In the case of the tensor algebra S, the cone of positive forms has been extensively studied, especially by H.J. Borchers[(1-3)] and J. Yngvason[(14)], who interpret Field theory from the point of view of positive functionals on the algebra S. One particularly considers the sub-cone of Wightman functionals. These are the positive functionals which are invariant under the natural action of the (connected) inhomogeneous Lorentz group, and satisfy the so-called spectrum and locality conditions.

Now consider the well known GNS construction, which to every $\omega \in A'_+$ associates a Hilbert space H_ω. According to L. Schwartz's theory of reproducing kernels[(9)], H_ω can be canonically embedded in the dual A'. In the case of Field theory this is not generally done (though it may turn out to be very useful). In the case of Harmonic analysis it is very natural. Thus the Hilbert space associated to δ is nothing but the space $L^2(G)$, which is of course regarded as a space of distributions. The equation (5') then turns into a direct integral decomposition of $L^2(G)$ into minimal translation-invariant Hilbert spaces $H_\omega \subset D'(G)$.

If $T \in A'$ it is convenient to define $\langle T,\phi\rangle = T(\phi^\star)$, and to define the product $\phi.T$ and the element $T^\star$ by the formulas

(6) $\quad (\phi\cdot T)(\psi) = T(\phi\cdot\psi), \qquad T^\star(\phi) = T(\phi^\star),$

just as one does in distribution theory. The space $H_\omega \subset A'$ then contains the set

$$D_\omega = \{\phi\cdot\omega\}_{\phi\in A}$$

as a dense subspace, the inner product (anti-linear on the right) being defined on D_ω by the formula:

(7) $\quad (\phi\cdot\omega|\psi\cdot\omega) = \langle\omega,\phi^\star\cdot\psi\rangle.$

On D_ω one can define the operator $\pi_\omega(\phi)$ by the assignment $T \to \phi\cdot T$. The map $\phi \to \pi_\omega(\phi)$ is just the usual $\star$-algebra representation associated to the GNS construction.

In the case of Field theory the GNS construction, and the representation of the algebra S, amounts to the reconstruction of a field in terms of vacuum expectation values: the operators $\pi_\omega(\phi)$, $\phi\in S$, are just the field operators. When in (7) $\psi=1=(1,0,0,\ldots)$ we in fact obtain:

(8) $\quad \omega(\phi) = (\pi(\phi)\Omega \mid \Omega)$

where Ω is the cyclic vector corresponding to the GNS construction (if we view the space H_ω as a Hilbert space of A', $\Omega=\omega$). In particular, if $\phi_1,\phi_2,\ldots,\phi_n$ belong to S, we have:

(9) $\quad \omega(\phi_1\otimes\phi_2\otimes\ldots\otimes\phi_n) = (\pi(\phi_1)\pi(\phi_2)\ldots\pi(\phi_n)\Omega \mid \Omega).$

The locality condition insures that the operators $\pi(\phi)$ and $\pi(\psi)$ commute when $\phi,\psi\in S$ have spacially separated supports. The invariance of ω under the natural action of the homogeneous Lorentz group gives rise to a representation of this group in H_ω in which the element Ω is fixed. The spectral condition on ω just expresses the assumption that the spectral measure corresponding to the representation of the translation subgroup is concentrated on the forward light cone. Those field theories in which there is a unique vacuum (i.e. for which the space of vectors fixed under the action of the inhomogeneous Lorentz group is one-dimensional) are certainly extremal in the cone of all Wightman functionals. Conversely, however, a field theory corresponding to an extremal Wightman functional does not necessarily have a unique vacuum[1].

In the case of Harmonic analysis the operators $\pi_\omega(\phi)$ are bounded in H_ω (smoothing convolution operators). In general, e.g. in the case of Field theory, they are not bounded. It can happen, however, that the representation π_ω is self-adjoint in the sense of Powers[7,8], in which case we will simply say that the element ω is self-adjoint.

We can now formulate an extension of the Bochner-Schwartz theorem to the general framework discussed in this lecture:

Theorem 2. Let A be a nuclear $\star$-algebra which is an $\mathcal{LF}$-space satisfying the condition (C), and having an approximate identity. Let Γ be the cone of central positive forms. Then every self-adjoint element $\underline{\omega}$ of Γ has a unique integral representation by means of extreme elements of Γ.

By theorem 1 one only has to prove that the face $\Gamma(\underline{\omega})$ is a lattice in its own order. The idea is to prove that it is isomorphic to the positive part of a commutative $C^\star$-algebra, which by Gelfand's theorem is a lattice. For $f\in D_{\underline{\omega}}$ we have $\pi_{\underline{\omega}}(\phi)f = \phi\cdot f$. Similarly let $\rho_{\underline{\omega}}(\phi)f = f\cdot\phi^\star$. Then $\pi_{\underline{\omega}}$ and $\rho_{\underline{\omega}}$ are representations of A. Also let $J(f) = f^\star$. Then J is an anti-unitary involution in the space $H_{\underline{\omega}}$ and $\rho_{\underline{\omega}} = J\pi_{\underline{\omega}}J$. Consider the weak commutant $\{\pi_{\underline{\omega}}(A)\}'$. This in the case of a self-adjoint representation is known to be a von Neumann algebra[8] (the proof of lemma 4.6 in ref. 8 does not appear to make use of the unit). Thus the weak commutant $\{\rho_{\underline{\omega}}(A)\}' = J\{\pi_{\underline{\omega}}(A)\}'J$ is also a von Neumann algebra. The intersection M of two commutants is a von Neumann algebra stable under conjugation by J. Now $\Gamma(\underline{\omega})$ is isomorphic to the positive part of M. In fact $\omega\in\Gamma(\underline{\omega})$ iff $\omega(\psi^\star\cdot\phi) = (T\ \phi\cdot\underline{\omega} \mid \psi\cdot\underline{\omega})$, with $T\in M^+$. It suffices therefore to prove that M is commutative. This can be done in the same way as in the bounded case, by showing that for $A\in M$, $A^\star = JAJ$ (cf. refs. 4 and 13).

Theorem 3. Let $\underline{\omega}$ be a positive self-adjoint Wightman functional on the algebra S. Then $\underline{\omega}$ has a unique integral representation by means of extreme Wightman functionals.

Here again one can show that the face $\Gamma(\underline{\omega})$ in the cone of all Wightman functionals is isomorphic to the positive part of a commutative von Neumann algebra, in the case $\{\pi_{\omega}(A)\}'$. The fact that this commutant is commutative is a consequence of the PCT symmetry, as explained in the work of H. J. Borchers[2]. A priori one would have to consider the subalgebra of $\{\pi_{\omega}(A)\}'$ composed of operators commuting with the action of the inhomogeneous Lorentz group. Actually, it is shown in ref. 2 that, thanks to the spectrum condition, every operator in the commutant commutes with the group action.

In this same article, and later in more detail in ref 1, the existence of the integral decomposition has been proved. The uniqueness statement, based on theorem 1, appears to be new.

References.

(1) H.J. Borchers and J. Yngvason, "On the Algebra of Field operators. The Weak commutant and integral Decompositions of states", Comm. Math. Phys. 42, (1975) 231-252.

(2) H.J. Borchers, "On the structure of the algebra of Field operators", Il Nuovo Cimento XXIV, no. 2, (1962) 214-236.

(3) H.J. Borchers, "Algebraic Aspects of Wightman Field theory, in Statistical Mechanics and Field theory", R.N. Sen and C. Weil (eds.), Halsted Press (1972) 31-79.

(4) O. Bratteli, D.W. Robinson, "Operator algebras and quantum statistical mechanics", Springer Verlag (1979).

(5) A. Grothendieck, "Produits tensoriels topologiques et espaces nucléaires", Memoir Amer. Math. Soc. 16 (1955).

(6) R. Jost, "The general Theory of Quantised Fields", Amer. Math. Soc. (1965).

(7) R.T. Powers, "Algebras of unbounded operators', Proc. of Symp. in Pure Mathematics 38,(1982) 389-406.

(8) R.T Powers, "Self-Adjoint algebras of unbounded operators", Comm. Math. Phys. 21, (1971) 85-124.

(9) L .Schwartz, "Sous espaces Hilbertiens et noyaux associés, Journal d'Analyse Mathématique de Jérusalem, (1964) 115-256.
(10) R.F. Streater and A.S. Wightman, "PCT, Spin and Statistics, and all that", Mathematical Physics Monograph Series, Benjamin (1978).
(11) E.G.F. Thomas, "Représentation intégrale dans les cônes convexes", C.R. Acad. Sci. Paris t. 286, (1978) 515-518.
(12) E.G.F. Thomas, "Integral representations in convex cones", Report University of Groningen (1977).
(13) E.G.F. Thomas, "The theorem of Bochner-Schwartz-Godement for Generalised Gelfand Pairs", Proceedings of the third Paderborn Conference on Functional Analysis, North Holland Publ. Co. (1983).
(14) J.Yngvason, "On the decomposition of Wightman functionals in the Euclidean Framework", Reports on Mathematical Physics 13, (1978)101-115.

HOW DOES ONE CHARACTERIZE MACROSCOPIC DIFFUSION?

H. van Beijeren, Instituut voor Theoretische Fysica, RU Utrecht

Under what conditions can a microscopic process be characterized macroscopically as a diffusion process?

It is not enough to require that the mean square displacement is asymptotically linear in time: any superposition of diffusion processes with different diffusion constants has this property. In such a case the Einstein or the Green-Kubo-formula for the diffusion constant just gives the mean of the diffusion constants associated with the separate processes.

A stronger requirement is that under the scaling transformation ($\tau=\epsilon t, \vec{\rho}=\epsilon^{\frac{1}{2}}\vec{r}$) and in the limit as $\epsilon \to \infty$ the microscopic process tends to Brownian motion in the variables τ and $\vec{\rho}$. This, however, does not suffice to ensure the existence of the diffusion constant according to the Einstein-formula. Such may be illustrated by a model in which a particle combines a simple random walk with occasional jumps over a large distance immediately followed by the same jump backwards. In such a situation the process does scale to Brownian motion, but the mean square displacement may diverge for all times so that the diffusion coefficient does not exist.

This leads to several "physical" consequences: although some properties, like e.g. the spectrum of light scattered from jumping particles, may be almost the same as for normal diffusion, other properties, like e.g. the stationary flow between two reservoirs, may be different in an essential way.

To characterize diffusion macroscopically one should require both that the process scales to Brownian motion and that the diffusion coefficient exists.

References

1. R.J. ZWANZIG, J. Chem. Phys. **40** (1964) 2527.
2. D. DÜRR, S. GOLDSTEIN and J. LEBOWITZ, J. Stat. Phys. **30** (1983) 519.
3. Y. GEFEN, A. AHARONY and S. ALEXANDER, Phys. Rev. Lett. **50** (1983) 77.

RANDOM FIELDS IN THE ISING MODEL

J. Villain, IFF der KFA, Jülich (BRD)
(Abstract made by the organizers)

The model discussed here consists of a d-dimensional lattice of Ising spins with Hamiltonian

$$H = -\sum_{\langle i,j\rangle} J s_i s_j - \sum_i H_i s_i ,$$

where s_i is the spin at site i, J is a positive constant (the coupling constant) and the H_i are i.i.d. random variables of mean zero and variance H^2 (the random field). The notation $\langle .,. \rangle$ denotes nearest neighbours. This "random field Ising model" has a wide range of applicability: it describes reasonably well e.g. solid-state systems with impurities, spin glasses, and ferromagnetic materials subjected to a random magnetic field. Most of the recent experimental interest has concentrated on the latter. The main characteristic of the model is that a competition takes place between the J-term, having a tendency to align the spins, and the random H-term, tending to destroy this long range order. We discuss several aspects of the model and some recent theoretical developments.

1. The lower critical dimension.

The lower critical dimension is the dimension d_{cl} below which long range order is destroyed by the field, no matter how weak the field is. Imry and Ma[1] have given the following qualitative argument why d_{cl} should be 2.

Consider a large domain of size R in which all the spins are up, surrounded by a region of downspins. Flipping all the spins inside this domain would yield a gain in energy which, for large R, behaves like

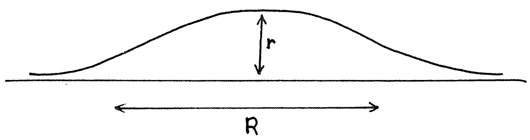
$$(1) \qquad \delta W \sim -gR^{d-1} + H\,(R^d)^{\frac{1}{2}}.$$

Here g is the surface tension, to which we shall return shortly. For $d<2$ the field term dominates, so that flipping the spins is not energetically favorable, and no long range order is established. For $d>2$, on the other hand, the surface term dominates (provided $g>0$), so that the domain is unstable.

2. Domain wall roughening.

The above argument, being heuristic, has been criticized, but more detailed considerations and calculations show that it is essentially correct[2-4]. The main point of the criticism is that, as a result of the erratic influence of the field, the domain walls may become quite rough, so that their surface area may in fact become much larger than is assumed in eq.(1) while the energy lost in abolishing them becomes much less. This may be taken into account by letting g depend on R.

Whether or not this roughening indeed takes place can be found out, again heuristically, by comparing the energy of a flat domain wall with that of one developing a bump:



For $r \ll R$ the difference goes like

(2) $\delta W \sim gR^{d-1}(r/R)^2 - H\,(R^{d-1}r)^{\frac{1}{2}}$.

(To understand the first term, put e.g. d=2. If $r \ll R$ then the length of the arc is $\int(dx^2+dz^2) \simeq \int(1+\frac{1}{2}(dz/dx)^2)dx \simeq R+\frac{1}{2}R(r/R)^2$.) Minimization of δW leads to

(3) $r \sim (H/g)^{2/3}R^{(5-d)/3}$.

For d>2 this is consistent with $r \ll R$ and hence no roughening takes place for d>2. Eq. (2) may be used to give an estimate of g(R), showing that for d>2 indeed g(R) remains strictly positive, while for d<2 it collapses and becomes negative (for large R). The case d=2 requires a more refined analysis.

3. Metastable states.

For d>2, large domains with very rough walls, though not stable, are fenced off from the ground state by an energy barrier of the order of $H(H/g)^{1/3}R^{(d+1)/3}$, obtained by substitution of eq.(3) into eq.(2). As their relaxation times become exceedingly large at low temperatures, these domains then will seem stable for all practical purposes. Experimental investigations and computer simulations meet with great difficulties because of the existence of such metastable states.

4. Breakdown of the $(d_{cu}\text{-}d)$-expansion.

The upper critical dimension is the dimension d_{cu} above which all mean-field results for the critical exponents are exact. For the random field Ising model $d_{cu}=6$ (i.e., 2 higher than the ordinary Ising model). In the standard theory of critical phenomena one probes the region below d_{cu} by an expansion in $d_{cu}\text{-}d$. However, for the random field Ising model this expansion turns out to yield incorrect results, which is believed to be due to the occurrence of metastable states. For instance, one of the outcomes of this expansion is that d_{cl} lies 2 above the value of the ordinary Ising model, i.e., $d_{cl}=3$ instead of 2.

5. Expansion in powers of $d\text{-}d_{cl}$.

Bray and Moore[5] have obtained results based on a $(d\text{-}d_{cl})$-expansion, on scaling assumptions and on renormalization methods. No field renormalization occurs, but the line tension is renormalized. The correlation length ξ is found to behave near T_c like

$$(4) \qquad \xi \sim |T-T_c|^{-3/2(d-2)}.$$

However, the exponent in eq.(4) is controversial. In the ordinary Ising model the exponent equals $1/(d-2)$.

References:

(1) Y. Imry and S. K. Ma, Phys. Rev. Lett. 35 (1975) 1399.

(2) J. Villain, J. Phys. Lett. 43 (1982) L551.

(3) G. Grinstein and S. K. Ma, Phys. Rev. Lett. 49 (1982) 685, Phys. Rev. B28 (1983) 2588.

(4) J. Z. Imbrie, Phys. Rev. Lett. 53 (1984) 1747.

(5) A. J. Bray and M. A. Moore (preprint 1985).

SOME RECENT DEVELOPMENTS AND PROBLEMS IN THE THEORY OF SPIN GLASSES

H. Hilhorst, Instituut-Lorentz voor Theoretische Natuurkunde, RU Leiden

One of the most important developments during the last five years in the theory of random lattice systems has been Parisi's work on the Ising spin glass. This work, and its ramifications, were the subject of this lecture.

In 1980 Parisi gave a "solution" of a mean field model of the Ising spin glass. In subsequent years this solution turned out to lead to a number of fascinating conclusions:

(1) below the critical temperature such a spin glass has infinitely many thermodynamic phases (as opposed to the usual ferromagnet which has only two: spinup or spin-down);

(2) these phases are not related via symmetry transformations but can only be discussed as statistical objects;

(3) the order parameter is a function.

In this lecture the derivation of these results was discussed in detail, and several mathematical gaps in the argument leading to Parisi's "solution" were pointed out. Finally, it was indicated how one is currently trying to find alternative routes to give the theory a better mathematical foundation.

References

1. S. Kirkpatrick and D. Sherrington, Phys. Rev. B **17** (1978) 4384.
2. J.R.L. de Almeida and D.J. Thouless, J. Phys. A **11** (1978) 983.
3. G. Parisi, J. Phys. A **13** (1980) 1101.
4. G. Parisi, Phys. Rev. Lett. **50** (1983) 1946.
5. K. Binder and A.P. Young, Rev. Mod. Phys. **58** (1986) 801.

DISTRIBUTIONS WITH THE LEE-YANG PROPERTY

J. Naudts, Departement Natuurkunde, Universitaire Instelling Antwerpen (Belgium)

In statistical mechanics phase transitions are described as non-analyticities in the free energy density f as a function of some model parameter. For example, in the 2-dimensional Ising-model at low enough temperature the first derivative $\delta f/\delta h$ with respect to the magnetic field h shows a discontinuity at h=0. Our understanding of the origin of such non-analyticities is partially based on the Lee-Yang theorem(1), which states that for a certain class of polynomials all zeros lie on the unit-circle in the complex plane. Since the publication of this theorem in 1952 many generalizations have appeared in the literature and new applications have been found. Of particular interest here is the work of Newman(2) and of Lieb and Sokal(3), in which a generalized Lee-Yang theorem is formulated in terms of probability distributions. Following ref. 2, one says that a probability distribution μ on R has the Lee-Yang property if the moment-generating function given by $f_\mu(z)=\int \exp(zx)d\mu(x)$ is analytic in the halfplane Re z >0 and has no zeros in this halfplane. The generalized Lee-Yang theorem can be formulated as a construction theorem, stating that if certain probability distributions have the Lee-Yang property then so do certain other probability distributions derived from these (see refs. 2 and 3).

In the present approach a further generalization is obtained by formulating the Lee-Yang theorem in terms of moment sequences. Given any sequence $\mu=\{\mu_n\}_{n\geq 0}$ of complex numbers the moment $M_\mu p$ of a polynomial $p(z)=\sum_{0\leq m\leq n} a_m z^m$ equals $\sum_{0\leq m\leq n} a_m \mu_m$ (convention $\mu_0=1$).

Definition 1. The support D_μ (in C) of a sequence $\mu=\{\mu_n\}_{n\geq 0}$ is the intersection of all closed sets D with the property that for each polynomial p there exists a point z in D such that $M_\mu p=p(z)$.

Proposition 2. The zeros of the polynomials

$$p_n^\mu(z) = \sum_{0\leq m\leq n} \binom{n}{m} (-z)^m \mu_{n-m}$$

are dense in the support of D_μ.

Proposition 3. Consider the probability distribution μ of the total magnetization of a ferromagnetic Ising-model and let $\mu=\{\mu_n\}_{n\geq 0}$ be the corresponding sequence of moments of μ. Then the support D_μ is a subset of the imaginary axis Re z = 0.

Let μ be a probability distribution for which the generating function f_μ is analytic in the halfplane Re z > 0 and in a neighborhood of the origin. If the support D_μ is a subset of the halfplane Re $z \leq 0$ then it follows from Hurwitz's theorem that f_μ does not vanish on the halfplane Re z > 0, which means that μ has the Lee-Yang property. Hence proposition 3 is a strong Lee-Yang theorem in the sense that the result of the original Lee-Yang theorem, when applied to the Ising model, can be obtained as a consequence of the present result.

An important tool in the derivation of the above results is a theorem of Grace[4]. A closed subset D of the complex plane is called a "circular region" if it is the inside or outside of a circle or if it is a halfplane. Grace's theorem can be formulated as follows.

Theorem 4. Given a circular region D and a sequence μ of complex numbers, then the following two statements are equivalent:

(a) for each polynomial p there exists a point z in D such that $M_\mu p=p(z)$;

(b) $D_\mu \subset D$.

This theorem is the essential ingredient in the proof of proposition 2 [5].

Let us now fix a finite subset Λ of the lattice Z^d and let

$$D = \{ z\in C^\Lambda : \text{Re } z_j \leq 0 \text{ for all } j\in\Lambda \}.$$

Let $G_{D^\sim}$ denote the set of all multi-index sequences $\mu=\{\mu_m\}_{m\in N^\Lambda}$ with the property that for each polynomial p on C^Λ there exists a point z in D such that $M_\mu p=p(z)$. In the proof of proposition 3 the following construction theorems are used (similarly as in Newman's proof of the Lee-Yang theorem):[6]

(i) If a sequence μ over N has support D_μ in the halfplane Re $z \leq 0$, then the multi-index sequence $\nu=\{\nu_m\}_{m\in N^\Lambda}$ with elements $\nu_m=\prod_{j\in\Lambda}\mu_{m_j}$ belongs to $G_{D^\sim}$.

(ii) If $\nu\in G_{D^\sim}$ and p is a polynomial which does not vanish on D, then the multi-index sequence $\rho=\{\rho_m\}_{m\in N^\Lambda}$ with elements $\rho_m=M_\nu(z\mapsto z^m p(z))/M_\nu p$ belongs to

$G_{D^\sim}$.

(iii) If $\rho \in G_{D^\sim}$, then the sequence σ over N with elements $\sigma_n = M_\rho(z \to (\sum_{j\in\Lambda} z_j)^n)$ has its support D_σ in the halfplane Re $z \le 0$.

References:

(1) T. D. Lee and C. N. Yang, Phys. Rev. 87, 410-419 (1952).

(2) Ch. M. Newman, Commun. pure appl. Math. 27, 143-159 (1974).

(3) E. H. Lieb and A. D. Sokal, Commun. Math. Phys. 80, 153-179 (1981).

(4) J. H. Grace, Proc. Cambridge Philos. Soc. 11, 352-357 (1902).

(5) J. Naudts, "Duality and the support of a sequence" (preprint 1986).

(6) J. Naudts, "Lee-Yang theory" (preprint 1987).

INTERACTING-PARTICLE SYSTEMS AND RANDOM NETWORKS

G. Grimmett, School of Mathematics, Bristol University (United Kingdom)

I. Interacting-particle systems.

Principal reference: Thomas Liggett, "Interacting particle systems", Grundlehren der mathematischen Wissenschaften 276, Springer, 1985.

Set S of sites (usually $S=Z^d$, $d\geq 1$).
Set W of local states, each site $x\in S$ being in some member of W at each time t (usually $W=\{0,1\}$).
State space $X = W^S \equiv \{ \eta : S \to W \}$, so that the composite state of S at any time may be described as a mapping $\eta : S \to W$.

We assume henceforth that $W=\{0,1\}$, the states 0/1 representing directions of spins or vacancy/occupancy of the site in question by a particle.
Topological considerations: X is compact, metrizable space, with product discrete topology.

Let $\{ \eta_t : t\geq 0 \}$ be a Markov process on X with infinitesimal changes given by

$$P(\eta_{t+\delta}(x) = 1-\eta(x) \mid \eta_t = \eta) = c(x,\eta)\ \delta + o(\delta),$$
$$P(\eta_{t+\delta}(x) = \eta(y),\ \eta_{t+\delta}(y) = \eta(x) \mid \eta_t = \eta) = c(x,y,\eta)\ \delta + o(\delta),$$

where c is the speed function of the process, and we assume that $c(x,y,\eta)=0$ unless one of $\eta(x)$ and $\eta(y)$ equals 0 and the other 1.

Case 1. $c(x,y,\eta)=0$ for all x and y: spin system.
Case 2. $c(x,\eta)=0$ for all x: particle system.

Examples.

Voter model: $c(x,\eta) = \sum_{y\in S} p(x,y)\ I_{\{ \eta(x) \neq \eta(y) \}}$,

where $p(x,y)\geq 0$, $\sum_y p(x,y) = 1$, $p(x,x)=0$.

Antivoter model: $c(x,\eta) = \sum_{y \in S} p(x,y)\, I_{\{\eta(x) = \eta(y)\}}$.

Contact process:
$$c(x,\eta) = \begin{cases} 1 & \text{if } \eta(x) = 1 \\ \lambda\, |\{ y \in A : \eta(x+y) = 1 \}| & \text{if } \eta(x) = 0 , \end{cases}$$
where A is a finite set.

Stochastic Ising model: $c(x,\eta) = \exp(-\beta \sum_{y \sim x} \chi(x)\chi(y))$,

where $\beta \geq 0$, $\chi(x) = 2\eta(x) - 1$, and $\sim$ is a relation on S.

Simple exclusion process:
$$c(x,y,\eta) = \begin{cases} p(x,y) & \text{if } \eta(x) = 1,\ \eta(y) = 0 \\ 0 & \text{otherwise,} \end{cases}$$
where $p(x,y) \geq 0$, $\sum_y p(x,y) = 1$, $p(x,x) = 0$.

Principal questions.

(1) For a given speed function, when does there exist a corresponding Markov process with uniquely determined transition probabilities?
(2) Identify the set I of invariant probability measures.
(3) Identify the set I_e of extremal invariant probability measures.
(4) When is it the case that $|I| = 1$?
(5) If $\nu \in I$, identify all measures μ such that $\mu S(t) \to \nu$ as $t \to \infty$, where S is the underlying semigroup.
(6) Under what condition is the process ergodic, in that $I = \{\nu\}$ and $\mu S(t) \to \nu$ as $t \to \infty$ for all μ?

We discuss the answers to these questions in general, beginning with an account of the Hille-Yosida theory of interacting particle systems, leading to Liggett's theorem that the process is uniquely determined if, in the case of a spin system for example,

1. $0 \leq c(x,\eta) \leq M$ for all x and η,
2. $c(x,\eta)$ is continuous in η for all x,
3. $\sup_{x \in S} \sum_{y \in S} \sup_{\eta} |c(x,\eta^y) - c(x,\eta)| < \infty$,

   where η^y is the composite state η except for a spin flip at y.

We discuss general properties of ergodicity.

I.A. Spin systems.

A spin system with speed function c is attractive if, whenever $\eta \le \xi$ (i.e. $\eta(x) \le \xi(x)$ for all x), then

$$c(x,\eta) \ge c(x,\xi) \quad \text{if } \eta(x) = \xi(x) = 1, \quad c(x,\eta) \le c(x,\xi) \quad \text{if } \eta(x) = \xi(x) = 0 .$$

We discuss the coupling of pairs of spin systems, particularly for attractive systems, and indicate the value of coupling two realizations of the same attractive process on a single probability space.

Theorem. If a spin system with speed function c is attractive, then there exists a coupled pair (η_t, ξ_t) of spin systems defined on the same probability space such that η_t and ξ_t both have speed function c and

$$P(\eta_t \le \xi_t \text{ for all t})$$

whenever $\eta_0 \le \xi_0$. □

Two processes η_t and ξ_t with state spaces X and Y are called dual with respect to the bounded measurable function H on XxY if

$$E^{\eta} H(\eta_t, \xi) = E^{\xi} H(\eta, \xi_t) \quad \text{for all } \eta \text{ and } \xi,$$

where $\eta = \eta_0$, $\xi = \xi_0$. We discuss duality for spin systems, with particular reference to the voter model, showing that η_t is dual to a system of finitely many coalescing random walks, when

$$H(\eta, A) = \prod_{x \in A} (1 - \eta(x))$$

for $\eta \in X$ and finite subsets A of S. As an application of duality we prove the following theorem for the voter model.

Theorem. Let U and V be continuous-time random walks on S with generator $\{p(x,y),\ x \ne y\}$, and let $\delta_\emptyset$ and δ_S be point masses for η on the composite states

$\eta \equiv 0$ and $\eta \equiv 1$, respectively. Then $I_e = \{ \delta_\emptyset, \delta_S \}$ if and only if

$$P(U_t = V_t \text{ for some } t) = 1 \text{ for all starting positions } U_0 \text{ and } V_0. \qquad \square$$

We discuss applications of this theorem to nearest-neighbour voter models on Z^d, $d \geq 1$. Finally, we discuss the biased voter model, and the shape problem thereof, and the antivoter model together with an application to the problem in computational complexity of determining whether or not a given finite graph has a k-coloring.

I.B. Exclusion processes.

Exclusion process with speed change:

$$c(x,y,\eta) = \eta(x)c(x,\eta)p(x,y) + \eta(y)c(y,\eta)p(y,x).$$

Simple exclusion:

$$c(x,y,\eta) = \begin{cases} p(x,y) & \text{if } \eta(x) = 1,\ \eta(y) = 0 \\ p(y,x) & \text{if } \eta(x) = 0,\ \eta(y) = 1 \\ 0 & \text{otherwise.} \end{cases}$$

When $p(x,y) = p(y,x)$: symmetric simple exclusion.

Symmetric simple exclusion. This process is self-dual in that

$$P^A(F \subset A_t) = P^F(F_t \subset A)$$

for all $A \subset S$ and finite subsets F of S, where A_t and F_t are exclusion processes with initial sets $A_0=A$, $F_0=F$ (the sets represent the positions of the 1's). This selfduality leads to a complete categorization of the invariant probability measures of such a process through the following theorem.

Theorem. Let H be the set of functions $\alpha : S \to [0,1]$ which are p-harmonic in that

$$\alpha(x) = \sum_y p(x,y)\, \alpha(y) \quad \text{for all } x.$$

For $\alpha \in H$, let μ_α be product measure on X with $\mu_\alpha\{ \eta : \eta(x)=1 \} = \alpha(x)$. Then

$\nu_\alpha = \lim_{t\to\infty} \mu_\alpha S(t)$ exists,
$I_e = \{ \nu_\alpha : \alpha \in H \}$. □

Let X_t and Y_t be independent continuous-time random walks on S with generator $\{ p(x,y) : x \neq y \}$. Let

$$g(x,y) = P^{x,y}(X_t = Y_t \text{ for some } t),$$

where $x=X_0$ and $y=Y_0$.

Case 1: $g(x,y) = 1$ for all x and y. In this case H contains constant functions only, and I_e contains only "uniform" product measures. This is the situation when X is recurrent.
Case 2: $g(x,y) < 1$ for some x and y. In this case H may contain non-constant functions. This is the situation when X is transient, but there exist also recurrent X with this property.

General simple exclusion. The general process is not self-dual. Under certain circumstances straight calculations show some invariant measures.
Doubly-stochastic. If p is doubly-stochastic then the product measures μ_α with α constant are invariant.
Time-reversible. If there exists $\pi \geq 0$ such that $\pi(x)p(x,y) = \pi(y)p(y,x)$ for all x and y, then $\alpha(x) = \pi(x)/(1+\pi(x))$ defines an invariant product measure μ_α.

Coupling techniques are available also, to show for example the next theorem.

Theorem. If p is irreducible, positive recurrent and time-reversible, then

$$I_e = \{ \mu_n : 0 \leq n \leq \infty \},$$

where μ_n is the unique invariant measure of the n-particle exclusion process and $\mu_\infty = \delta_S$. □

Motion of a tagged particle. Suppose that p is translation invariant and $S=Z^d$. A particle is placed at the origin at time t=0, and its position at time t is denoted by X_t. The sites in $S\backslash\{0\}$ are occupied at t=0 according to product measure μ_α with α constant. The motion of the tagged particle is retarded by the presence of the other particles, but by how much?

Theorem. Seen from X_t, the landscape is statistically invariant. Hence

$$EX_t = (1-\alpha)\ (\sum_y yp(0,y))\ t$$

and $\lim_{t\to\infty} t^{-1}X_t$ exists a.s. □

Theorem. If d=1, p(x,x+1)=p and p(x,x-1)=1-p, then

$$(X_t - EX_t)\ /\ \mathrm{Var}^{\frac{1}{2}} X_t \longrightarrow N(0,1).$$

If p=1, then X_t is a Poisson process with rate $1-\alpha$. Furthermore, Var $X_t \simeq (2t/\pi)^{\frac{1}{2}}(\alpha^{-1}-1)$ if $p=\frac{1}{2}$, while Var $X_t \sim t$ if $\frac{1}{2}<p\leq 1$. □

II. First-passage percolation.

References:

(1) Smythe R. T., Wierman J. C., "First-passage percolation on the square lattice", Springer Lecture Notes in Mathematics 671.

(2) Grimmett G. R., Kesten H., "First-passage percolation, network flows and electrical resistances", ZWG 66 (1984) 335-366.

(3) Kesten H., "Aspects of first-passage percolation", to appear in Proc. St. Flour meeting, 1984, Springer Lecture Notes in Mathematics.

Underlying lattice Z^d, d≥1, each edge e having an associated time coordinate t(e), thought of as the time required by disease/rumour/fluid to traverse e (in either direction). We assume that $\{ t(e) : e\in Z^d \}$ are independent non-negative random variables with common distribution function F. The passage time between two sites x and y is

$$a(x,y) = \inf \{ \sum_{e\in\pi} t(e) : \text{paths } \pi \text{ from } x \text{ to } y \}.$$

Basic question: What is the behaviour of a(x,y) when |x-y| is large?

Theorem. If d=2, there exists a constant $\mu=\mu(F)$ such that

$$n^{-1}\ a((0,0),(n,0)) \longrightarrow \mu(F) \quad \text{as } n\to\infty,$$

the convergence being in probability. Convergence with probability 1 turns out to hold if and only if

$$\int (1-F(x))^4 dx < \infty. \qquad \square$$

We discuss:

(a) the relationship between this theorem and the ergodic theory of subadditive stochastic processes;

(b) limit theorems for other passage times, such as:

b_{0n} = infimum passage time from (0,0) to line x=n,

λ_n = infimum passage time accross opposite sides of a nxn square;

(c) large and small deviation theorems for passage times;

(d) properties of the time constant $\mu(F)$ as a function of the underlying distribution function F;

(e) limiting shape and height problems;

(f) the spatial dual to first-passage percolation, viz. network flows through randomly capacitated media.

III. Random electrical networks.

References:

(1) Chayes J. T., Chayes L., "Percolation and random media", Les Houches Lecture Notes, 1984.

(2) Grimmett G. R., "Random flows", in Surveys in Combinatorics, L.M.S. Lecture Notes 103, 1985 (and references therein).

Electrical network: finite graph G with source at 0 and sink at ∞, each edge e having some resistance r(e). We suppose that $\{ r(e) : e \in G \}$ is a collection of independent random variables with some common distribution function.

Basic question: What can be said about the effective resistance R(G) of the network?

We discuss limit theorems for R(G) for large networks G which are trees, complete graphs, or subsections of crystalline lattices.

Complete graph. Suppose G is the complete graph on n vertices and each edge-resistance r has distribution

$$P(\, r \le x\,) = p(n)F(x),\ P(\, r=\infty\,) = 1-p(n),$$

where F is a prescribed distribution function concentrated on $[0,\infty)$ and $0<p(n)<1$. Let $\gamma(n) = np(n)$, having the order of magnitude of the mean number of conducting edges incident to any given vertex.

Theorem. If $\lim_{n\to\infty} \gamma(n) = \infty$, then the effective resistance R_n between a pair of sites on the graph satisfies

$$\gamma(n)R_n \longrightarrow 2\left(\int_{[0,\infty)} x^{-1} dF(x) \right)^{-1} \text{ in probability.}$$

If $\lim_{n\to\infty} \gamma(n) = \gamma$, then

(i) if $\gamma<1$: $\lim_{n\to\infty} P(R_n=\infty) = 1$,

(ii) if $\gamma\ge 1$: $R_n \to R' + R''$ in distribution, where R' and R'' are independent random variables, each being distributed as the resistance between the root and ∞ of a branching process with Poisson-distributed family sizes with mean γ. □

Lattices. In two dimensions, let G be a nxn section of Z^2, with the left- and right-hand sides as source and sink. Let R_n be the effective resistance when the edge-resistances are i.i.d.

Theorem. If each edge-resistance r satisfies $P(\alpha<r<\beta) = 1$ for some real α and β satisfying $0<\alpha\beta<\infty$, then $\lim_{n\to\infty} R_n$ exists in mean square. □

Theorem. If $P(r=\alpha) = P(r=\beta) = \frac{1}{2}$ for some α and β satisfying $0<\alpha\beta<\infty$, then $\lim_{n\to\infty} R_n = (\alpha\beta)^{\frac{1}{2}}$ in mean square. □

Theorem. In the special (percolation) case when

$$P(r=1) = p,\ P(r=\infty) = 1-p,$$

for some given $p\in[0,1]$, we have the following:

(a) $p<\frac{1}{2}$: $\lim_{n\to\infty} P(R_n=\infty) = 1$;

(b) $p=\frac{1}{2}$: $P(\lim_{n\to\infty} R_n = \infty) = 1$;

(c) $p>\frac{1}{2}$: $0 < \liminf R_n \le \limsup R_n < \infty$ a.s. □

We discuss extensions of this result to higher dimensions, together with the relationship of part (c) to the existence of "open paths" in the percolation model.

STATISTICAL COOLING: A GENERAL APPROACH TO COMBINATORIAL OPTIMIZATION PROBLEMS

E. Aarts, Philips Research Laboratorium, Eindhoven

Many combinatorial optimization problems belong to the class of NP-hard problems. Consequently, for large-size problems exact solutions require prohibitive computational efforts. Less time-consuming optimization algorithms can be constructed by applying tailored heuristics striving for near-optimal solutions. These tailored algorithms, however, often depend strongly on the structure of the problem to be solved. This is a major drawback since it prohibits fast and flexible implementation and application. Furthermore, there is a growing number of combinatorial problems, originating from different fields (e.g. the design of computer hard- and software), for which no adequate (heuristic) optimization methods are known at all. It is for the above reasons that the need arises for a generally applicable and flexible optimization method. Statistical cooling[*] can be viewed as such a method[1]: it is a general optimization technique for solving combinatorial problems.

The method originates from the analogy between the annealing of solids, as described by the theory of statistical physics, and the optimization of large combinatorial problems. Randomization techniques are used to model this analogy. The salient features of the algorithm are its general applicability and its ability to obtain near-optimal solutions. The quality of the final solution is determined by the convergence, which is governed by a set of parameters: the cooling schedule. The most important parameter in the cooling schedule is the cooling control parameter, which plays the role of temperature in the physical annealing process. The statistical cooling algorithm can be formulated in terms of a set of (time-homogeneous Markov chains (the homogeneous algorithm) or in terms of a single inhomogeneous Markov chain (the inhomogeneous algorithm)[2,3]. It can be shown that with probability 1 the algorithm converges asymptotically to globally optimal solutions, provided the transition matrix associated with the Markov chain satisfies a number of conditions. Additional conditions on the control parameter are required if the formulation based on the inhomogeneous Markov chain is used.

The asymptotic convergence of the homogeneous algorithm to a global minimum requires a number of transitions that is exponential in the size of the input of the problem[1], resulting in an exponential-time execution. For the inhomogeneous algorithm we conjecture a similar complexity result[2,3]. On the other hand, near-optimal solutions can be obtained in polynomial time by choosing an appropriate (time-)dependence for the parameters that control the convergence[1].

The relation between combinatorial optimization and statistical mechanics has been addressed by many authors, because of interest either in a phenomenological analogy or in a possible framework to model the convergence and the corresponding control of the algorithm. Most authors focus on global aspects such as entropy, ensemble averages, phase transitions, and the relation between spin-glass Hamiltonians and cost functions. More specific studies, based on relations with physics, are e.g. studies on ultra metricity and on the replica analysis.

The statistical cooling algorithm has been applied to diverse combinatorial optimization problems, both practical and theoretical, in a wide range of disciplines[3]. The list of applications includes optimization problems related to VLSI design, image processing, code design and neural networks. The general conclusion emerging from these applications is that near-optimal solutions can be obtained in nearly all cases, but that some problems require large computational efforts. For randomly generated problem instances the performance of the algorithm is found to be especially excellent.

Researchers believe that the statistical cooling algorithm may evolve more and more into a technique that can be applied successfully to a wide range of combinatorial optimization problems. There is also scepticism, however, especially when the computational effort becomes a critical factor. For some problems there exist worthy competing tailored algorithms and it is therefore believed that the strength of the statistical cooling algorithm lies primarily in application to problem areas where no such tailored algorithms are available. Further research on large-scale applications and on means to speed up the algorithm by parallel execution, possibly using dedicated hardware, has to be carried out in order to provide more insight into the practical use of the method.

(*) Footnote: Other names to denote the method are "simulated annealing, "Monte Carlo annealing" and "probabilistic hill climbing".

References:

(1) E. H. L. Aarts and P. J. M. van Laarhoven, "Statistical Cooling: A General Approach to Combinatorial Optimization Problems", Philips J. of Research 40 (1985) 193-226.

(2) E. H. L. Aarts and P. J. M. van Laarhoven, "A Pedestrian Review of the Theory and Applications of the Simulated Annealing Algorithm", Proc. Heidelberg Colloquium on Glassy Dynamics and Optimization, Heidelberg, 1986 (Springer, 1987).

(3) P. J. M. van Laarhoven and E. H. L. Aarts, "Simulated Annealing: A Review of the Theory and Applications"(submitted to Acta Applicandae Mathematicae).

ON A NONCOMMUTATIVE APPROACH TO EQUILIBRIUM STATES IN THE q-STATE POTTS MODEL

R. Kuik, Instituut voor Theoretische Natuurkunde, RU Groningen

In classical statistical mechanics one is interested in properties of probability measures (states) which describe equilibrium for a particular type of interaction. The interaction depends on parameters with a physical interpretation, such as temperature and coupling constants.

The method we will discuss shortly is useful when one considers the set of equilibrium states at a fixed interaction, i.e., at fixed temperature and coupling constants. The method consists of a variant of the standard transfer-matrix method. By this method one identifies classical states in d dimensions with states for noncommutative systems in d-1 dimensions. Decomposition theory of states on noncommutative algebras is used to provide decompositions of the classical states into states which describe pure thermodynamic phases. There are two main steps:

(1) a Markov chain construction leading (together with (2)) to a global Markov property for equilibrium states;

(2) a reformulation of the standard transfer-matrix formalism which can deal with all boundary conditions.

The main point in the first step is to show that the construction yields equilibrium states. The proof uses the variational principle. In the second step we consider only translationally invariant nearest-neighbor interactions on Z^d.

Let f_k be a function depending on finitely many variables in the plane

$$P_k=\{ x\in Z^d:\ x_1=k \},\ k\in Z.$$

The transfer-matrix formalism tells us that the expectation value $\langle . \rangle_N$ in a Gibbs ensemble on the volume

$$\Lambda_N = \{ x\in Z^d:\ |x_i| \leq N,\ i=1,\dots d \}$$

can be written, e.g. for a product $f_{-1}f_0f_1$, as

$$\langle f_{-1}f_0f_1\rangle_N = L_n(\ (V_N)^N\ T^{(N)}(f_{-1}f_0f_1)\ (V_N)^N\ ).$$

Here

$$T^{(N)}(f_{-1}f_0f_1) = V_N{}^{-1}\,\hat{f}_{-1}\,V_N\,\hat{f}_0\,V_N\,\hat{f}_1\,V_N{}^{-1},$$

where V_N and $\hat{f}_k$ are matrices and L_N is a linear functional on the space of matrices. The matrix V_N is the transfer matrix and the matrix $\hat{f}_k$ depends only on the function f_k. Write $l_N(.) = L_N((V_N)^N(.)(V_N)^N)$. (Note that when one considers periodic boundary conditions, $l_N(.)$ is proportional to Trace$\{(.)(V_N)^{2N+1}\}$.) We extend $T^{(N)}$ linearly to all C_0 (= set of functions depending on finitely many variables). The map $T^{(N)}$ takes values in A_0 (= algebra of matrices) and, moreover, $T^{(N)}(f)$ for $f \in C_0$ becomes independent of N for N sufficiently large. The value for large N is denoted by T(f). Hence, if μ is the thermodynamic limit of the sequence $\{\langle.\rangle_N\}_{N\geq 0}$, one obtains

$$(*) \qquad \mu = l_\mu \circ T .$$

Here l_μ is the linear functional on A_0 that is the limit of l_N, $N\to\infty$. The important observation to make is that <u>T does not depend on μ</u> and, moreover, for every Gibbs state μ we can find a unique l_μ such that (*) holds.

In case μ is, for instance, translationally invariant the functional l_μ is a state on A_0 and one can use step (1) to go from l', a linear functional on A_0 which appears in a state decomposition of l_μ, to a Gibbs measure μ' which satisfies (*), i.e., $\mu' = l' \circ T$.

For a translationally and reflectionally invariant nearest-neighbor interaction the above is applicable whenever the transfer matrix is invertible. Invertibility of the transfer matrix holds for the standard q-state Potts model whenever there is nontrivial coupling.

<u>Some results:</u>

(A) Every Gibbs state is uniquely determined by its restrictions to neighboring planes.

(B) Every Gibbs state that is extremal among the $(Z_{ev})^d$-invariant states is strongly clustering of all orders for nontrivial subgroups of $(Z_{ev})^d$ which are generated by one element parallel to a coordinate axis. Here $(Z_{ev})^d$ is the group of all translations over an even number of lattice spacings along each coordinate axis.

(C) There exists an infinite-volume transfer operator when μ is Z^d-invariant, also in the multiple-phase domain. This operator is in the von Neumann algebra generated by the observables and by the unitaries which implement translations orthogonal to the transfer direction.

A transfer-matrix method, like the one discussed, is used by Fredenhagen, Commun. Math. Phys. 101 (1985), to prove the existence of dynamics in lattice gauge theories.

RANDOM FIELDS ON TREES

C. Scheffer and M. Dekking, Faculteit Wiskunde en Informatica, TU Delft

Part I. Correlations of a Markov random field on a tree (C. Scheffer)

Let T be the set of vertices of a tree with nearest-neighbour relation $\langle .,. \rangle$. An E-valued random field on T is a probability distribution on the set E^T of (E-) configurations on T. Here we specialize to $E = \{-1,+1\}$ so that

$$\Omega = E^T$$

is a compact group (under coordinatewise multiplication) with dual

$$\hat{\Omega} = \{ D \subset T \mid |D| < \infty \}$$

(and with duality $(\omega|D) = \prod_{t\in D} \omega(t)$ $(\omega\in\Omega,\ D\in\hat{\Omega})$). Thus, if P is a random field, the correlations

$$\int \prod_{t\in D} \omega(t)\ P(d\omega)$$

are nothing but the values of the Fourier transform $\hat{P}$ of P and hence determine P completely.

If P is a Markov random field specified by a 2x2 matrix

$$Q = \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix} \qquad (\,p,q \in [0,1]\,)\ ,$$

i.e. P is determined by its marginals P_S for any finite connected subset S of T as follows:

$$P[\ \omega(s)=\epsilon(s),\ s\in S\ ] = \pi_{\epsilon(s_0)} \prod_{\substack{s,s'\in S \\ \langle s,s'\rangle}} Q_{\epsilon(s),\epsilon(s')} \qquad (\epsilon\in E^S,\ s_0\in S)\ ,$$

where

$$\pi = (\pi_{-1},\pi_1) = (\,\frac{1-q}{2-p-q}, \frac{1-p}{2-p-q}\,)\ ,$$

which we take, moreover, to be symmetric:

$$p = q\ ,$$

in which case $\pi = (\frac{1}{2},\frac{1}{2})$, then we can compute $\hat{P}$ explicitly.

Theorem. There is a function $D \mapsto k_D : \hat{\Omega} \to N$ with the following properties:
(1) k_D depends only on the least subtree $\bar{D}$ of T containing D;
(2) $k_D = 0$ if $|D|$ is odd;
(3) if D' is the union of D with a two-point set $\{s,t\}$ for which there exists $u \in D$ such that $\langle u,s \rangle$ and $\langle s,t \rangle$, then $k_{D'} = k_D+1$;
(4) if D'' is the union of D with a two-point set $\{s,t\}$ for which there exists $u \in D$ such that $\langle u,s \rangle$ and $\langle u,t \rangle$, then $k_{D''} = k_D+2$;
(5) $\hat{P}(D) = (2p-1)^{k_D}$.

Corrolary 1. The product (in terms of the group operation on Ω) of two independent symmetric Markov random fields P_p and $P_{p'}$ on Ω is again a Markov random field with parameter $p'' = p\,p' + (1-p)\,(1-p')$.

Corrolary 2. A symmetric Markov random field with parameter p is infinitely divisible iff $p \geq \frac{1}{2}$.

Part II. Bernoulli percolation on the b-ary tree and branching processes in random environments (M. Dekking)

The b-ary tree is the graph with a distinguished node ρ, the root, which is connected to b nodes, the first level nodes; generally, there are b^n nodes at level n, each connected to a single node at level n-1 and to b nodes at level n+1 ($b \geq 2$ is an integer). A Bernoulli process on the b-ary tree is a process of i.i.d. random variables X_ν, indexed by the nodes ν of the tree. Let $(\nu_0,\nu_1,\ldots,\nu_n)$ be a chain on the tree, i.e., $\nu_0=\rho$, and the ν_i are i^{th}-level nodes such that ν_i and ν_{i+1} are connected. We consider the random set

$$W_n = \{ X_{\nu_0} X_{\nu_1} \ldots X_{\nu_n} : (\nu_0,\nu_1,\ldots\nu_n) \text{ a chain} \}$$

of all words of length n+1 occurring on the tree, starting at the root.

Theorem A. Let Card W_n be the number of different words of length n+1 occurring on the b-ary tree with Bernoulli process (X_ν) given by $P[X_\nu=0]=1-p$, $P[X_\nu=1]=p$,

$p \in [0,1]$. Then $\alpha_p = \lim_{n\to\infty} \{ E_p \text{ Card } W_n \}^{1/n}$ exists, and $\alpha_0 = \alpha_1 = 1$ and

$$\alpha_p = \begin{cases} 2 & \text{if } p(1-p) \geq b^{-2} \\ \exp H(\beta) & \text{if } 0 < p(1-p) < b^{-2} \end{cases}$$

where $H(\beta) = -\beta \log \beta - (1-\beta) \log (1-\beta)$ and $\beta = \{ 1 - \log (bp)/\log (b-bp) \}^{-1}$.

Note that α_p has the interesting property that $d\alpha_p/dp = \infty$ at $p=0$ and $p=1$. Theorem A can be proved as an application of the following theorem on the extinction probability of branching processes in a two state i.i.d. random environment.

<u>Theorem B</u>. Let (Z_n) be a critical or subcritical B.P.R.E. in a two state i.i.d. environment with offspring distributions f_0 and f_1 having finite second moments, and environmental probabilities p_0 and $p_1 = 1-p_0$, $p_0 \in (0,1)$. Let $m_0 = f_0'(1)$ and $m_1 = f_1'(1)$. Suppose that: (a) $0 < m_0 < 1 < m_1$, (b) $p_0 m_0 \log m_0 + p_1 m_1 \log m_1 > 0$. Then

$$\lim_{n\to\infty} P[Z_n > 0]^{1/n} = (p_1/\gamma)^{\gamma} (p_0/(1-\gamma))^{1-\gamma},$$

where $\gamma = \{ 1 - \log m_1/\log m_0 \}^{-1}$.

References:

(1) K. B. Athreya and P. Ney, "Branching Processes", Springer, Berlin, 1970.

(2) F.M. Dekking, "Subcritical branching processes in a two state random environment and a percolation problem on trees" (J. Appl. Prob., december 1987).

(3) F.M. Dekking, "On the survival probability of a branching process in a finite state i.i.d. environment" (to appear in Stoch. Proc. and their Appl.).

(4) F.M. Dekking and G. Grimmett, "Superbranching processes and projections of random Cantor sets" (preprint 1987).

Remark (November 1987). Theorem B has been generalized to finite state i.i.d. environments (see ref.3). Theorem A has the following extension: $(\text{Card } W_n)^{1/n}$ converges to α_p as $n\to\infty$ a.s. and in L_1 (see ref.4).

A MODEL FOR CRYSTALLIZATION: A VARIATION ON THE HUBBARD MODEL

E. Lieb, Departments of Mathematics and Physics, Princeton University (USA)

(Abstract made by the organizers)

A deep unsolved problem in mathematical physics is to derive crystallization from fundamental principles. A modest aim going in this direction is to construct some caricature model based on free particles only, and to prove the onset of some spatial ordering as temperature is lowered. One of the models believed to possess these characteristics (without having been proved to do so) is the Gutzwiller-Hubbard-Kanamori model(1-3). We have succeeded in proving that a still further simplified version of this model does display crystallization(4,5).

The Hubbard model itself is defined by the Hamiltonian

$$H_{Hub} = \sum_{\sigma=\pm 1} \sum_{x,y\in\Lambda} t_{xy}\, c_{x\sigma}^{*} c_{y\sigma} + 2U \sum_{x\in\Lambda} n_{x,1}\, n_{x,-1} \,.$$

Here Λ is a finite lattice and $c_{x\sigma}$ is the (fermion) annihilation operator for an electron at site x with spin σ. The operator $n_{x\sigma} := c_{x\sigma}^{*} c_{x\sigma}$ counts the number (0 or 1) of spin-σ electrons at x. The symmetric $\Lambda x \Lambda$ matrix $T = (t_{xy})$ gives the amplitudes for quantum hopping between the points of Λ. The positive number U is the potential due to the electric repulsion between electrons at the same site. The crucial assumption will be made that Λ is the disjoint union of two sublattices A and B such that $t_{xy} = 0$ as soon as x and y are in the same sublattice. The simplification which we shall make is the assumption that one kind of electrons (say $\sigma = -1$) does not hop. The Hamiltonian then becomes (we may now omit the index σ)

$$H = \sum_{x,y\in\Lambda} t_{xy}\, c_x^{*} c_y + 2U \sum_{x\in\Lambda} W(x)\, n_x .$$

The function W takes the values 0 and 1, indicating the locations of the fixed electrons. Since $n_x = c_x^{*} c_x$, H is now a quadratic form in the creation and annihilation operators, so that the analysis reduces to the single particle space, where the single particle Hamiltonian h' = T+V+UI governs the dynamics. Here V is the diagonal matrix with entries U(2W(x)-1) and I is the identity matrix. Let us put h = h'-UI.

Henceforth we shall call the movable particles "electrons" and the fixed ones "nuclei". This terminology becomes most appropriate if we also put $U < 0$, thus replacing repulsion by attraction. (Note, however, that the Hamiltonian with the sign of U changed is related to the original one through the symmetry operations $c_x \to c_x^{\star}$, $T \to -T$ and $W \to 1-W$.) The ground state energy $E(N_e,N_n)$, when there are N_e electrons and N_n nuclei, is defined as

$$E(N_e,N_n) = \min_{\sum_x W(x)=N_n} \{ \langle\psi,H\psi\rangle : \psi \in \mathcal{I}(N_e) \}$$

where $\mathcal{I}(N_e)$ is the N_e-th eigenspace of the electron number operator $\sum_x n_x$.

Our results are of two kinds. The first concerns the ground state, which we prove always has perfect crystalline ordening (and an energy gap). The second concerns the positive temperature grand canonical state. For low temperature the long range order persists in dimension $d \geq 2$. For high temperature it disappears, and there is an exponential clustering of the nuclear correlation functions.

Here we give only the first result in some detail.

Theorem. (a) For $U<0$ the minimum

$$\min_{N_e+N_n \leq 2|\Lambda|} E(N_e,N_n)$$

is taken for $N_e = N_n = |A|$ in the nuclear configuration $W = 1_A$. The same holds for B. If Λ is connected, these are the only ground states, i.e., if $|A| > |B|$ the ground state is unique, while if $|A| = |B|$ it is doubly degenerate.

(b) For $U>0$ the minimum

$$\min_{N_e+N_n \leq |\Lambda|} E(N_e,N_n)$$

is taken for two ground states: $N_e= |A|$, $N_n=|B|$, $W=1_B$, and $N_e=|B|$, $N_n=|A|$, $W=1_A$. If Λ is connected, these are the only ground states.

Proof. Let $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_{|\Lambda|}$ be the eigenvalues of h. Then

$$E(N_e,N_n) - UN_e = \sum_{j=1}^{N_e} \lambda_j \geq \sum_{\lambda_j<0} \lambda_j = \tfrac{1}{2} (\operatorname{tr} h - \operatorname{tr} |h|).$$

Now, tr h = tr V = $U(2N_n-|\Lambda|)$ and $|h| = (T^2+U^2+UJ)^{\frac{1}{2}}$, with J the matrix with elements $J_{xy} = 2t_{xy}(W(x)+W(y)-1)$. Since $x \to x^{\frac{1}{2}}$ is concave on $[0,\infty)$, the function $f(y) = \mathrm{tr}(T^2+U^2+yUJ)^{\frac{1}{2}}$ is concave on $[0,1]$. But $f(-1) = f(1)$ (since the spectrum of h is invariant under $T \to -T$), so $f(1) \le f(0)$ with equality if and only if J=0. Thus

$$E(N_e,N_n) \ge U(N_e+N_n) - \tfrac{1}{2}\, U\, |\Lambda| - \tfrac{1}{2}\, \mathrm{tr}(T^2+U^2) \ .$$

If Λ is connected, the only way to have J=0 is either $W=1_A$ or $W=1_B$. If U<0 and $N_e+N_n \le 2|A|$, then from the above,

$$E(N_e,N_n) \ge U\,(\, 2|A| - \tfrac{1}{2}|\Lambda|\,) - \tfrac{1}{2}\, \mathrm{tr}(T^2+U^2)^{\frac{1}{2}} \ .$$

If $W=1_A$, then h has precisely |A| negative and |B| positive eigenvalues. Thus, if $W=1_A$ and $N_e=|A|$, equality holds. The case U>0 is similar.

References:

(1) M. C. Gutzwiller, Phys. Rev. Lett. 10 (1963) 159-162, Phys. Rev. A134 (1964) 923-941, A137 (1965) 1726-1735.

(2) J. Hubbard, Proc. Roy. Soc. (London), Ser. A276 (1963) 238-257, A277 (1964) 237-259.

(3) J. Kanamori, Prog. Theor. Phys. 30 (1963) 275-289.

(4) T. Kennedy and E. H. Lieb, in: Proceedings of the Conference on Statistical Mechanics and Field Theory, Groningen, 1985 (Springer Lecture Notes in Physics).

(5) T. Kennedy and E. H. Lieb (to appear in Physica).

The theory of large deviations and its applications in statistical mechanics

J.L. van Hemmen

Sonderforschungsbereich 123 der Universität Heidelberg
6900 Heidelberg
Federal Republic of Germany

This is a summary of two extensive talks presented at the Mark Kac Seminar (Amsterdam, 1986). As such, it cannot be a substitute of the lectures themselves. I, therefore, concentrate on the main ideas and refer the reader to the literature at the end of this paper for details and further references.

Contents:
I. Theoretical considerations
II. Neural nets and spin glasses

I. Theoretical considerations

The philosophy of large deviations has contributed substantially to the solution of certain spin-glass models [1-4] and the understanding of a large class of nonlinear neural nets [5-12]. The underlying mathematics can be described succinctly as follows.

Let $\{W_N, \mu_N\}$ be a sequence of pairs of random variables W_N and probability measures μ_N. We require the sequence $\{ W_N, \mu_N \}$ to be such that the corresponding c-function,

$$c(t) = \lim_{N \to \infty} N^{-1} \ln \mu_N (e^{tW_N}), \tag{1.1}$$

exists and is differentiable. It will be shown that under these conditions

$$\text{Prob}\, \{N^{-1} W_N \approx \varepsilon \} \sim e^{-Nc^*(\varepsilon)} \tag{1.2}$$

where $c^*(\varepsilon) = \sup_t \{\varepsilon t - c(t)\}$ is the Legendre transform of c(t). To verify (1.2) we use a key lemma.

Lemma [13]: If c(t) in (1.1) is differentiable at t = 0, then whatever $\delta > 0$ there exists a constant $a(\delta) > 0$ such that, as $N \to \infty$,

$$\text{Prob}\, \{ | N^{-1} W_N - c'(0) | \geq \delta \} \leq e^{-Na(\delta)}. \tag{1.3}$$

Let us now consider a small open interval (a,b) surrounding $[\varepsilon, \varepsilon + d\varepsilon)$. We have

$$\mu_N (a < N^{-1} W_N < b) = \mu_N \left(\underbrace{\frac{e^{tW_N}}{\mu_N (e^{tW_N})}}_{\downarrow} e^{-tW_N} \mu_N (e^{tW_N}); a < N^{-1} W_N < b \right)$$

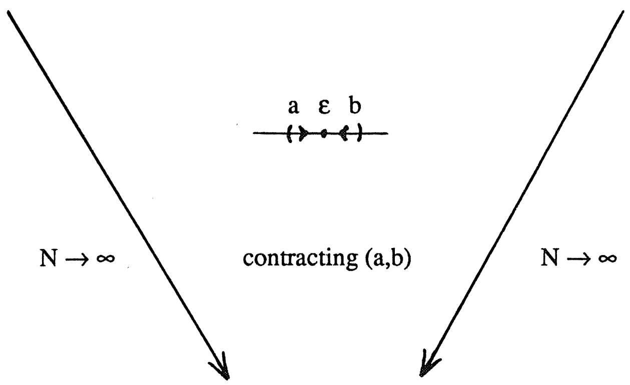
$$\equiv \mu^*_N \left(e^{-tW_N} \mu_N (e^{tW_N}) ; a < N^{-1} W_N < b \right). \tag{1.4}$$

Since $Na < W_N < Nb$, we can easily estimate $\exp(-tW_N)$ in (1.4) and obtain

$$e^{-tNb}\,\mu_N(e^{tW_N})\,\mu^*_N(a < N^{-1}W_N < b) \leq \ldots \leq e^{-tNa}\,\mu_N(e^{tW_N})\,\mu^*_N(a < N^{-1}W_N < b) \qquad (1.5)$$

where ... comprises the probability (1.2) we are interested in. We will show shortly that $\mu^*_N(a < N^{-1}W_N < b)$ converges to one as $N \to \infty$. Taking this for granted, applying logarithms, and dividing by N, we then get from (1.5)

$$tb - N^{-1}\ln\mu_N(e^{tW_N}) \geq N^{-1}\ln \text{Prob}\{a < N^{-1}W_N < b\} \geq ta - N^{-1}\ln\mu_N(e^{tW_N}) \qquad (1.6)$$



$$t\varepsilon - c(t) = c^*(\varepsilon) \qquad (1.7)$$

where (1.7) holds provided t is chosen in such a way that $c'(t) = \varepsilon$. By the lemma, $\mu^*_N(a < N^{-1}W_N < b) \to 1$. To see this, we calculate the c-function corresponding to $\{W_N, \mu^*_N\}$,

$$c_1(s) = \lim_{N\to\infty} N^{-1}\ln\mu^*_N(e^{sW_N}) = c(s+t) - c(t), \qquad (1.8)$$

and note that $c'_1(0) = c'(t) = \varepsilon$. So we are done.

The upshot of the above considerations is that to estimate the probability of a large deviation which is a rare event with respect to the original measure, you shift or "translate" this measure so that your large deviation gets probability one with respect to the new measure. This is the *translation principle.*

A typical application of (1.2) is the evaluation of a free energy f(β),

$$-\beta f(\beta) = \lim_{N \to \infty} N^{-1} \ln \mu_N \left(\exp \{\beta N F(N^{-1} W_N)\}\right), \tag{1.9}$$

for some reasonably smooth function F. For instance, μ_N may be a normalized trace over N Ising spins - or the like. By (1.2) and a Laplace argument we can write

$$\begin{aligned} -\beta f(\beta) &= \lim_{N \to \infty} N^{-1} \ln \int dm \exp \left\{ N \left(\beta F(m) - c^*(m)\right) \right\} \\ &= \sup_m \{ \beta F(m) - c^*(m) \}, \end{aligned} \tag{1.10}$$

which solves the problem. An extension of this argument to vector spins is immediate. Furthermore [5], one need not evaluate the Legendre transform $c^*(m)$ explicitly.

II. Neural nets and spin glasses

Some years ago I proposed an exactly soluble spin-glass [1-4], whose predictions turned out to be in excellent argument with both experiments and large-scale Monte Carlo simulations [14]. As a random-site model with frustration it is a prototype of a large class of neural network models. Since "neural" computing has aroused a great deal of attention, I would like to concentrate here on this new development in computational and biophysics.

One of the most fascinating aspects of the brain is its function as an associative memory with a surprising *fault tolerance* to both input data errors and internal failures. Since this fault tolerance is global (not just a response to a local modification), collective features are expected to play a dominant role. The study of collective aspects of the brain dates back at least to the mid-forties (von Neumann, McCulloch and Pitts) but only recently a true revival of the subject was initiated by a seminal paper of John Hopfield [15] . Two of the main issues are: How do we recall a memory- a question which can be reformulated in terms of pattern recognition - and how can we solve certain problems more efficiently by using the notions of neural computing [16] ? For the moment we will concentrate on pattern recognition aspects.

A neuron can be modeled formally - à la McCulloch and Pitts - by a two-state element or, equivalently, by an Ising spin S(i) with S(i) = + 1 for firing and S(i) = - 1 for quiescent. In this language patterns are specific Ising spin configurations.

The data are stored in the synapses, which transmit information between neurons - and *not* in the neurons themselves. Under suitable conditions, the dynamics of the network can be represented as a Monte Carlo dynamics or down-hill motion in a (free-) energy landscape associated with a Hamiltonian

$$H_N = -\frac{1}{2} \sum_{i,j} J_{ij} \; S(i) \, S(j), \tag{2.1}$$

which is nothing but a function on the Ising phase space. See Fig. 1.

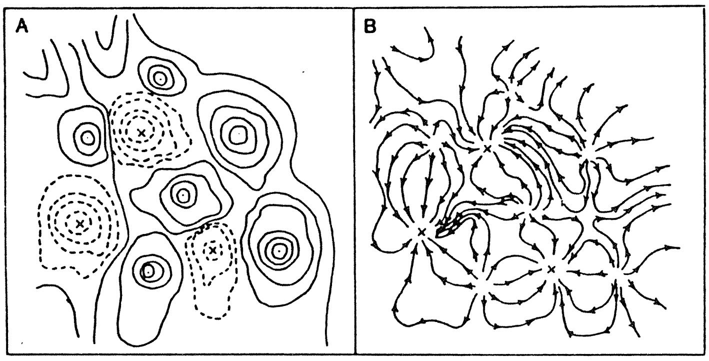


Fig. 1: (A) Energy landscape. In the picture, the phase space is two-dimensional and the contour lines indicate the height of the energy surface. Solid lines indicate a hill and dashed lines a valley.
(B) A dynamics which only allows that the energy decreases induces a flow ("down-hill motion") in the landscape of the previous figure A.

So the (free-) energy landscape induces a flow. The art of neural computing partially consists in choosing the J_{ij} in such a way that the stored patterns are (free-) energy *minima* of the Hamilton function (2.1).

Hopfield's original choice [15] was

$$J_{ij} = N^{-1} \sum_{\alpha=1}^{q} \xi_{i\alpha} \, \xi_{j\alpha} \equiv N^{-1} \, \underset{\sim}{\xi}_i \cdot \underset{\sim}{\xi}_j \tag{2.2}$$

where the $\{\xi_{i\alpha} \, ; \, 1 \le i \le N\}$ are the patterns, i.e., N-bit words labeled by $1 \le \alpha \le q$, and $\underset{\sim}{\xi}_i = (\xi_{i\alpha} \, ; \, 1 \le \alpha \le q)$ is a q-vector which represents the local information to which neuron i has been exposed. The $\xi_{i\alpha}$ are independent, identically distributed random variables which assume the values ± 1 with a probability which is usually 1/2.

Locality is an important ingredient of (2.2). For the type of model under consideration locality means that J_{ij} is determined by $\underset{\sim}{\xi}_i$ and $\underset{\sim}{\xi}_j$ *only* and, thus [6],

$$J_{ij} = N^{-1}\, Q\, (\underset{\sim}{\xi}_i\, ;\, \underset{\sim}{\xi}_j) \tag{2.3}$$

for some function $Q\,(\mathbf{x}\,;\,\mathbf{y}) = Q\,(\mathbf{y}\,;\,\mathbf{x})$ on $\mathbf{R}^q \times \mathbf{R}^q$. As before, q denotes the number of patterns. For example,

$$Q\,(\mathbf{x}\,;\,\mathbf{y}) = \sqrt{q}\;\;\phi\,(\mathbf{x}\cdot\mathbf{y}\,/\,\sqrt{q}) \tag{2.4}$$

defines the important class of inner-product models [6, 10 - 12]. The original Hopfield model has $\phi(x) = x$ and is therefore called *linear* . The so-called "clipped" synapses have $\phi(x) = \mathrm{sgn}\,(x)$. Clipping is extremely important in hardware versions of (2.4). It is highly *non*linear.

For finite q the free energy of the model (2.1) - (2.3) can be computed exactly [6,10], whatever the nonlinearity. Hence the (free-) energy valleys (ergodic components) can be obtained explicitly [11]. Using a replica Ansatz [8,17] and exploiting a weak invariance condition [6,10] I have derived [12] the free energy and thus determined the (stability of the) ergodic components related to the stored patterns in the case of *extensively* many patterns where $q = \alpha N$.

Specializing to inner-product models, I could show [12] that, *whatever* the synaptic function ϕ in (2.4), there exists a critical $\alpha_c \le \alpha_c^{\text{Hopfield}}$ which is such that for $\alpha > \alpha_c$ the system has lost its memory completely. Physiologically this does not seem completely satisfying. However, the class of soluble models also comprises memories which gradually forget. Their synaptic efficacies are defined via an iterative procedure which allows enough nonlinearity to be physiologically plausible,

$$J_{ij}\,(\alpha) := \phi\;\,(\varepsilon_N\, \xi_{i\alpha}\, \xi_{j\alpha} \;+ J_{ij}\,(\alpha - 1)\,). \tag{2.5}$$

The $J_{ij}\,(\alpha - 1)$ contains the information related to the patterns 1,2,..., α - 1. Furthermore, $\phi\,(x)$ is a function of the type $\phi\,(x) = \tanh\,(x)$, and ε_N is at our disposal as a function of the system size N. The special case

$$\phi\;(x) = \begin{cases} -1 \;, & x \le -1 \\ x \;, & -1 \le x \le 1 \\ 1 \;, & x \ge 1 \end{cases} \tag{2.6}$$

with $\varepsilon_N = \varepsilon/\sqrt{N}$ was proposed by Hopfield and studied numerically by Parisi. The present theory has opened up the way to an understanding of forgetful memories with general synaptic function ϕ.

References

1. J. L. van Hemmen, "Classical spin-glass model", Phys. Rev. Lett. **49** (1982) 409.

2. J. L. van Hemmen, A.C. D. van Enter and J. Canisius, "On a classical spin-glass model", Z. Phys. B **50** (1983) 311.

3. J. L. van Hemmen, "Equilibrium theory of spin glasses: Mean-field theory and beyond", in: *Heidelberg Colloquium on Spin Glasses* , edited by J. L. van Hemmen and I. Morgenstern, Springer Lecture Notes in Physics **192** (1983) 203 - 233, esp. the Appendix.

4. J. L. van Hemmen, " What is a true mean-field spin glass?", J. Phys. C **19** (1986) L 379.

5. J. L. van Hemmen, "Spin-glass models of a neural network", Phys. Rev. A **34** (1986) 3435.

6. J. L. van Hemmen and R. Kühn, "Nonlinear neural networks", Phys. Rev. Lett. **57** (1986) 913.

7. J. L. van Hemmen, D. Grensing, A. Huber and R. Kühn, "Elementary solution of classical spin-glass models", Z. Phys. B **65** (1986) 53.

8. J. L. van Hemmen and V. A. Zagrebnov, "Storing extensively many weighted patterns in a saturated neural network", J. Phys. A: Math. Gen. **20** (1987) to appear.

9. J. L. van Hemmen, "Nonlinear neural networks: Efficient storage and retrieval of information", in: *Heidelberg Colloquium on Glassy Dynamics* , edited by J. L. van Hemmen and I. Morgenstern, Springer Lecture Notes in Physics **275** (1987) 547 - 568.

10. J. L. van Hemmen, D. Grensing, A. Huber and R. Kühn, "Nonlinear neural networks: I. General theory", SFB-Preprint **411** (1987).

11. J. L. van Hemmen, D. Grensing, A. Huber and R. Kühn, "Nonlinear neural networks: II. Information processing", SFB-Preprint **412** (1987)

12. J. L. van Hemmen, "Nonlinear neural networks near saturation", ITP Preprint, University of California at Santa Barbara (1987).

13. J. L. van Hemmen and R. G. Palmer, "The thermodynamic limit and the replica method for short-range random systems" J. Phys. A.: Math. Gen. **15** (1982) 3881.

14. A. T. Ogielski and I. Morgenstern, "Critical behaviour of three-dimensional Ising spin-glass model", Phys. Rev. Lett. **54** (1985) 928 and J. Appl. Phys. **57** (1985) 3382.

15. J. J. Hopfield, "Neural networks and physical systems with emergent collecitve computational abilities", Proc. Natl . Acad. Sci. USA **79** (1982) 2554.

16. J. J. Hopfield and D. W. Tank, "Computing with neural circuits : A model ", Science **233** (1986) 625.

17. J. L. van Hemmen and R. G. Palmer, "The replica method and a solvable spin-glass model", J. Phys. A. : Math. Gen. **12** (1979) 563.

ON THE UNIQUENESS OF THE INFINITE CLUSTER IN PERCOLATION

A. Gandolfi, Department of Mathematics, Delft University of Technology
(Report made by the organizers)

Consider the probabilistic situation in which each of the sites of the lattice Z^d, $d \geq 1$, is either occupied or vacant according to a probability measure μ on the set of all such configurations. If nearest-neighbour sites are regarded as connected, then the set of occupied sites in any given configuration falls apart into maximal connected components which are called occupied clusters. The theory of percolation deals with the description of these clusters.

In this lecture we are concerned with the number N of occupied clusters which contain an infinite number of sites. Clearly, N is a random variable invariant under the group of "rigid" transformations of Z^d, i.e. translations, rotations and axis reflections. Therefore, if μ is ergodic under this group, then N is constant with probability one. If $N>0$ we say that percolation occurs.

The simplest example is Bernoulli percolation, which arises when each site is occupied with probability p and vacant with probability 1-p, independently of the other sites. It is known that for this case there is a critical value $p_c \in (0,1)$ for the parameter p such that $N=0$ when $p<p_c$ and $N>0$ when $p>p_c$. For d=2 it has been known since the 1960's that N=1 above p_c. For $d \geq 3$, on the other hand, the same was anticipated for a long time, but a proof was given only recently by Aizenman, Kesten and Newman [1]. Their proof is one of the major breakthroughs that percolation theory has seen in recent years. The situation at p_c is still open: N=0 in d=2, but it is not known whether N=0 also for $d \geq 3$.

With these facts established one may inquire what the value of N is for non-Bernoulli measures. Newman and Schulman [2] have shown that under minor additional restrictions on μ the only possible values for N are 0,1 or ∞. This suggests that for all "reasonable" μ one should have N=1 whenever there is percolation, but attempts to exclude $N=\infty$ on general grounds have failed so far. For d=2, Coniglio et al. [3] proved that indeed N=1 when μ describes a percolating Ising model with no external field. This result was extended in a paper by Gandolfi, Keane and Russo [4], which was the subject of the talk. The following theorem describes the conditions on μ.

Theorem. Let d=2. Suppose that:

(1) μ is invariant under horizontal and vertical translation and reflection.

(2) μ is ergodic with respect to horizontal and vertical translation separately.

(3) μ has the FKG-property (i.e. increasing events are positively correlated).

(4) N>0.

Then N=1 with probability one.

In the lecture an outline of the proof was given. The main idea is to show that any finite set of sites is surrounded by an occupied circuit with probability one. An important element in the proof is a 3-mixing theorem by Furstenberg [5]. The proof depends heavily on topological properties of Z^2 and cannot be extended to d≥3.

None of the conditions of the theorem is redundant: there are counterexamples as soon as any of the conditions is dropped.

Note: In a more recent paper [6], the speaker has extended the uniqueness result to arbitrary dimension for measures μ which are stationary Gibbs states with an interaction that is either finite range or decreases sufficiently rapidly with distance. The techniques used in this setting are based on large deviation theory.

References:

(1) M. Aizenman, H. Kesten and C.M. Newman, Uniqueness of the infinite cluster and continuity of connectivity functions for short and long range percolation, Comm. Math. Phys. (1987).

(2) C.M. Newman and L.S. Schulman, Infinite clusters in percolation models, Stat. Phys. 26 (1981) 613-628.

(3) A. Coniglio, C.R. Nappi, F. Peruggi and L. Russo, Percolation and phase transitions in the Ising model, Comm. Math. Phys. 51 (1976) 315-323.

(4) A. Gandolfi, M. Keane and L. Russo, On the uniqueness of the infinite occupied cluster in dependent two-dimensional percolation (to appear in Ann. Probab.).

(5) H. Furstenberg, Recurrence in ergodic theory and combinatorial number theory, Princeton University Press, Princeton, 1981.

(6) A. Gandolfi, Uniqueness of the infinite cluster for stationary Gibbs states, Preprint 1987.

APPLICATIONS OF DYNAMICAL SEMIGROUPS

A. Verbeure, Instituut voor Theoretische Fysica, Universiteit Leuven (Belgium)

Dynamical semigroups are semigroups of completely positive, unity preserving maps from an algebra (generally a C^*-algebra) into itself.

States of a dynamical sytem are positive linear forms on the algebra of observables which are normalized to one.

Dynamical semigroups have the interesting property of mapping, by duality, the state space of a system into itself. Therefore they can be used to describe perturbations of a state or they can be used to describe irreversible evolutions of the state of a system. An illustration of both these applications will be given below.

1. Thermodynamic stability.

An equilibrium state ω_β at inverse temperature β for a system described by a Hamiltonian H satisfies the variational principle of statistical mechanics, i.e.

$$F(\omega_\beta) \leq F(\omega) \qquad \text{for all states } \omega, \tag{1}$$

where F is the free energy given by $F(\omega)=\omega(H)-\beta^{-1}S(\omega)$, with S the entropy function. Using Lindblad's explicit formula for a generator L of a dynamical semigroup $\{\exp \lambda L \mid \lambda>0\}$:

$$L(Y) = X^*[Y,X] + [X^*,Y]X \quad ,$$

where X is an arbitrary observable, one derives from the variational principle in a straightforward way the following correlation inequality:

$$\beta\omega_\beta(X^*[H,X]) \geq \omega_\beta(X^*X) \log (\omega_\beta(X^*X)/\omega_\beta(XX^*)) \ . \tag{2}$$

This expresses the energy-entropy balance of an equilibrium state. Conversely, one may show that if a state ω satisfies (2) for all observables then the state ω also satisfies (1), i.e. it is an equilibrium state. Therefore one gets a characterization of the equilibrium states by means of the energy-entropy expression (2). Furthermore, the correlation inequality (2) turns out to be a powerful technique for proving or disproving the occurrence of spontaneous symmetry breaking.

Here we want to mention another application of (2), namely an equipartition theorem for quantum systems. One proves that

$$\lim_{V\to\infty} \frac{\langle P_V{}^2\rangle}{2m} = 3kT\rho \quad , \tag{3}$$

provided that the equilibrium state satisfies certain clustering conditions. In (3), $\langle\cdot\rangle$ stands for the thermal average, P_V is the bulk momentum of the system, V the volume, and ρ denotes the density of the particles. Equation (3) yields an intrinsic definition of the absolute temperature which neither depends on the type of interaction nor on the statistics (Bose or Fermi) governing the system. However, contrary to the classical mechanical equipartition theorem the above equality breaks down in the presence of a phase transition, which is always accompanied by long range order, and thus violates the clustering condition.

2. Critical slowing down.

Dynamical semigroups can also be used to describe irreversible time evolutions of quantum systems. In particular one might consider those semigroups $\{\exp tL \mid t\geq 0\}$ which satisfy the so-called detailed balance condition, i.e. for all observables X and Y the following symmetry property holds:

$$\omega_\beta(e^{tL}(X)Y) = \omega_\beta(Xe^{tL}(Y)) \quad . \tag{4}$$

For all evolutions $\{\exp tL \mid t\geq 0\}$ satisfying (4) as well as a locality property one proves that the spectral gap of the operator L is majorized by the inverse of the fluctuation of a certain observable. This implies that if the fluctuations diverge the spectral gap of L vanishes, in other words the lifetime diverges. This is essentially the phenomenon of critical slowing down.

References.

[1] M. Fannes, A. Verbeure: J. Math. Phys. 19 (1978) 558.
[2] M. Fannes, Ph. Martin, A. Verbeure: J. Phys. A16 (1983) 4293.
[3] R. Alicki, M. Fannes, A. Verbeure: J. Stat. Phys. 41 (1985) 263.

Renormalization and the Continuum Limit *

T. C. Dorlas.

Dublin Institute for Advanced Studies,
10 Burlington Road, Dublin 4, Ireland.

Abstract

It is explained how the renormalization transformation can be used to take the continuum limit of a lattice field. It is shown that, by rescaling, the problem can be formulated on a fixed lattice $\mathbf{Z}^d$. The procedure is illustrated by two examples: the one-dimensional Euclidean free field and a hierarchical model with ϕ^4 -interaction.

* Talk given at the Mark Kac Seminar, Amsterdam, 5 December 1986.

1. The general renormalization procedure

In the following we shall adopt as a definition of a (Euclidean) scalar field theory a generalized random field on some space $\mathcal{F}$ of functions, i.e. a *linear* mapping $\phi : \mathcal{F} \longrightarrow L^0(E,\mu)$, where $L^0(E,\mu)$ is the set of random variables on a topological space E with probability measure μ . We do not concern ourselves here with the Osterwalder-Schrader axioms. In the case of a d-dimensional lattice field $\mathcal{F}$ is a class of functions $f : \mathbf{Z}^d \longrightarrow \mathbf{C}$, and in the case of a continuum field $\mathcal{F}$ is a class of functions $f : \mathbf{R}^d \longrightarrow \mathbf{C}$.

We want to study the continuum limit of a lattice field ϕ on $\mathbf{Z}^d$. We therefore rescale the lattice $\mathbf{Z}^d$ with a factor $\delta > 0$ to obtain fields φ_δ on finer and finer lattices $\delta\mathbf{Z}^d$,and hope to be able to give a meaning to the limiting field $\varphi = \lim_{\delta\downarrow 0}\varphi_\delta$. In general it will be necessary to rescale the parameters defining the fields φ_δ in order to obtain a meaningful limit. This defines a transformation of parameters which is called the renormalization transformation. As a function of $\delta > 0$ these transformations obviously form a multiplicative 1-parameter semigroup, which is (erroneously) called the renormalization group.

In order to arrive at a suitable procedure to obtain a continuum limit let us assume for the moment that the continuum field φ is already given. Then we can obtain lattice fields φ_δ by coarse-graining, i.e. by averaging over lattice blocks

$$\Box_\delta(\underline{x}) = \{\underline{x}_i - \frac{1}{2}\delta \leq \underline{u}_i < \underline{x}_i + \frac{1}{2}\delta, i = 1, ..., d\} \tag{1.1}$$

for $\underline{x} \in \delta\mathbf{Z}^d$.
Explicitly,

$$\varphi_\delta(\underline{x}) = \delta^{-d}\varphi(1_{\Box_\delta(\underline{x})}) \tag{1.2}$$

where 1_A is the indicator function of the set A.
For $f \in \mathcal{F}_\delta$ this becomes

$$\varphi_\delta(f) = \sum_{\underline{x}\in\delta\mathbf{Z}^d} f(\underline{x})\varphi(1_{\Box_\delta(\underline{x})}) \tag{1.3}$$

The lattice fields φ_δ satisfy

$$\varphi_{L\delta}(\underline{x}) = L^{-d} \sum_{\underline{y}\in\delta\mathbf{Z}^d\cap\Box_{L\delta}(\underline{x})} \varphi_\delta(\underline{y}). \tag{1.4}$$

Conversely, given a sequence of lattice fields φ_n on $L^{-n}\mathbf{Z}^d$, we can define φ by the limit

$$\varphi(f) = \lim_{n\to\infty}\varphi_n(f) = \lim_{n\to\infty} \sum_{\underline{x}\in L^{-n}\mathbf{Z}^d} L^{-d}f(\underline{x})\varphi_n(\underline{x}). \tag{1.5}$$

If the φ_n satisfy (1.4) then this limit certainly exists for functions f of the form

$$f = \sum_{\underline{x} \in L^{-m}\mathbf{Z}^d} f(\underline{x}) 1_{\square_m(\underline{x})} \tag{1.6}$$

for some positive integer m.
Although it is desirable that this limit exist for a class of smooth functions f , we do not go into that problem here and concentrate on obtaining a sequence φ_n satifying (1.4). Rescaling we can reduce the latter problem to a fixed unit lattice $\mathbf{Z}^d$. Indeed, given fields ϕ_n on $\mathbf{Z}^d$ satisfying

$$\phi_n(x) = L^{-d} \sum_{y \in B_L(x)} \phi_{n+1}(y), \tag{1.7}$$

with

$$B_L(x) = \{y \in \mathbf{Z}^d | -\frac{L}{2} \leq y_i - Lx_i < \frac{L}{2}, i = 1, ..., d\}, \tag{1.8}$$

we can put

$$\varphi_n(\underline{x}) = \phi_n(L^n \underline{x}) \qquad (\underline{x} \in L^{-n}\mathbf{Z}^d). \tag{1.9}$$

We shall obtain ϕ_n by averaging a field ϕ on $\mathbf{Z}^d$ with a restricted number of parameters that can vary with n. The eventual continuum field then also depends on a finite number of parameters and thus, in principle, has predictive power. We define the averaging procedure by

$$(M\phi)_x = L^{-d+\sigma} \sum_{y \in B_L(x)} \phi_y, \tag{1.10}$$

where σ is an adjustable parameter, to be fixed later.
We denote the original field ϕ, but with parameters depending on m, by $\phi_{(m)}$. The fields ϕ_n can be defined by

$$(\phi_n)_x = L^{n\sigma} \lim_{m \to \infty} (M^{m-n} \phi_{(m)}). \tag{1.11}$$

One easily checks that these fields satisfy (1.7). Notice also that the limit depends on σ ,so that this is not a superfluous parameter. In fact only one particular choice of σ leads to a non-trivial limit.

2. The one-dimensional Euclidean free field.

Let us now illustrate the above procedure by a simple example : a one-dimensional Euclidean free field. We can define a Gaussian measure γ_M on $\mathcal{S}'(\mathbf{Z})$ by its covariance

$$C_{xy} = \int \delta_x . \delta_y \; d\gamma, \tag{2.1}$$

where $\delta_x(\phi) = \phi(x)$. We put

$$C_{xy} = (-\Delta_1 + M^2)^{-1}_{xy}, \tag{2.2}$$

where Δ_1 is the lattice Laplacian,

$$(\Delta_1 f)(x) = f(x+1) + f(x-1) - 2f(x). \tag{2.3}$$

By Fourier transformation we find

$$\begin{aligned} C_{xy} &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{e^{ip(x-y)}}{4\sin^2\frac{p}{2} + M^2} dp \\ &= \frac{e^{-\omega|x-y|}}{2M\sqrt{1+(M/2)^2}} \qquad if \quad x \neq y, \end{aligned} \tag{2.4}$$

with

$$\omega = 2 \ \text{arsinh}\frac{M}{2}. \tag{2.5}$$

The field ϕ is simply given by

$$\phi(f)(F) = \langle F, f\rangle \quad ; f \in \mathcal{S}(\mathbf{Z}), F \in \mathcal{S}'(\mathbf{Z}), \tag{2.6}$$

where $\langle .,. \rangle$ are the duality brackets.
Heuristically,

$$\gamma(d\phi) = \frac{1}{Z}\exp[-\frac{1}{2}\langle \phi, (-\Delta_1 + M^2)\phi\rangle] \prod_{x\in\mathbf{Z}} d\phi_x. \tag{2.7}$$

(Because of the simple defining relation for the field we write ϕ instead of F with an abuse of notation). Z is a normalization factor.

Let us calculate the covariance C' of the renormalized field $\phi' = M\phi$.

$$\begin{aligned} C'_{xy} &= \int \phi'_x . \phi'_y \ \gamma(d\phi) \\ &= L^{-2+2\sigma} \sum_{u\in B_L(x)} \sum_{v\in B_L(y)} C_{uv} \\ &= L^{-2+2\sigma} \frac{e^{-\omega L|x-y|}}{2M\sqrt{1+(M/2)^2}} \left\{ \frac{\sinh\frac{\omega L}{2}}{\sinh\frac{\omega}{2}} \right\}^2 \qquad (x \neq y). \end{aligned} \tag{2.8}$$

Clearly $(M_L)^n = M_{L^n}$, so that

$$C^{(n)}_{xy} = L^{-2n+2n\sigma} \frac{e^{-\omega L^n|x-y|}}{2M\sqrt{1+(M/2)^2}} \left\{ \frac{\sinh\frac{\omega L^n}{2}}{\sinh\frac{\omega}{2}} \right\}^2 . \tag{2.9}$$

Obviously $C_{xy}^{(n)} \to 0$ unless we let $\omega = \omega_n$ depend on n so that $\omega_n L^n \to const.$ We shall take $M_n = L^{-n} M_0$, so that, by (2.5), $\omega_n L^n \to M_0$. Formula (2.9) then contains a factor $L^{-2n+2n\sigma+n+2n}$,so that we have to take $\sigma = -\frac{1}{2}$.
The covariance C_n of the field ϕ_n defined by (1.11) thus becomes

$$\begin{aligned} C_{n;xy} &= L^{2n\sigma} \lim_{m\to\infty} \left(C_{M_n}^{(m-n)} \right)_{xy} \\ &= \frac{e^{-M_0 L^{-n}|x-y|}}{2M_0} \left\{ \frac{\sinh \frac{1}{2} M_0 L^{-n}}{\frac{1}{2} M_0} \right\}^2 . \end{aligned} \tag{2.10}$$

Rescaling we find the covariance of the fields φ_n on $L^{-n}\mathbf{Z}$,

$$\int \varphi_n(\underline{x})\varphi_n(\underline{y})\gamma(d\varphi_n) = \frac{e^{-M_0|\underline{x}-\underline{y}|}}{2M_0} \left\{ \frac{\sinh \frac{1}{2} M_0 L^{-n}}{\frac{1}{2} M_0} \right\}^2 . \tag{2.11}$$

Taking the limit $n \to \infty$ of

$$\int \varphi_n(f)\varphi_n(g)\gamma_n(d\varphi_n) = \sum_{\underline{x},\underline{y}\in L^{-n}\mathbf{Z}} L^{-2n} f(\underline{x})g(\underline{y}) \int \varphi_n(\underline{x})\varphi_n(\underline{y})\gamma_n(d\varphi_n)$$

we find the covariance of the continuum field,

$$\mathcal{C}(f,g) = \int_{\mathbf{R}} d\underline{x} \int_{\mathbf{R}} d\underline{y} f(\underline{x}) g(\underline{y}) \frac{e^{-M_0|\underline{x}-\underline{y}|}}{2M_0}. \tag{2.12}$$

The kernel

$$\mathcal{C}(\underline{x},\underline{y}) = \frac{e^{-M_0|\underline{x}-\underline{y}|}}{2M_0} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{ip(\underline{x}-\underline{y})}}{p^2 + M_0^2} dp \tag{2.13}$$

is the usual Euclidean free field "propagator" or Green's function.

In d dimensions one finds $\sigma = \frac{d-2}{2}$. This is called the canonical dimension of the scalar field φ. If one considers fields with (self-)interaction it may be necessary to change σ. One then speaks of an anomalous dimension.

All this may seem like a complicated way of replacing Δ_1 by

$$\Delta_n f(\underline{x}) = L^{2n} [f(\underline{x} + L^{-n}) + f(\underline{x} - L^{-n}) - 2f(\underline{x})]$$

and taking the limit

$$\sum_{\underline{x}\in L^{-n}\mathbf{Z}^d} L^{-n} f(\underline{x})((-\Delta_n + M_0^2).g)(\underline{x}) \longrightarrow \int d\underline{x} \quad f(\underline{x})((-\Delta + M_0^2).g)(\underline{x}).$$

For models with interaction, however, i.e. models with non- quadratic terms in the Hamiltonian or in other words a non- Gaussian measure μ instead of γ , serious singularities appear. By the renormalization group method described above these singularities are broken up into contributions from different scales, which all have approximately the same form. We illustrate this by a hierarchical model in which the massless quadratic part of the Hamiltonian is replaced by a hierarchical analogue.

3. The hierarchical model.

We replace the massless quadratic part $-\Delta_1$ of the Hamiltonian by the expression

$$H_0 = \sum_{k=0}^{\infty} L^{-(2+d)k} \Gamma_k \tag{3.1}$$

The kernels Γ_k are defined by

$$\Gamma_k(x,y) = \Gamma(x^{(k)}, y^{(k)}) \tag{3.2}$$

with

$$\begin{aligned} \Gamma(x,y) &= 1 - L^{-d} && \text{if } x = y \\ &= -L^{-d} && \text{if } x \neq y \text{ but } x^{(1)} = y^{(1)} \\ &= 0 && \text{if } x^{(1)} \neq y^{(1)} \end{aligned} \tag{3.3}$$

$x^{(1)}$ is the label of the block containing x , i.e. $x \in B_L(x^{(1)})$, and more generally, $x^{(k+1)} \in B_L(x^{(k)})$.
The corresponding covariance is

$$C^H_{xy} = \sum_{k=0}^{\infty} L^{-2\sigma k} \Gamma(x^{(k)}, y^{(k)}) \tag{3.4}$$

with $\sigma = \frac{d-2}{2}$.
One easily shows that

$$\left(C^H\right)'_{xy} = L^{-2\sigma} C^H_{x^{(1)}y^{(1)}} + \Gamma(x,y). \tag{3.5}$$

Heuristically one can define a measure μ by

$$\begin{aligned} \mu(d\phi) &= \frac{1}{Z} \exp[-\frac{1}{2}\langle \phi, H_0 \phi\rangle - \sum_{x \in \mathbf{Z}^d} v(\phi_x)] \prod_{x \in \mathbf{Z}^d} d\phi_x \\ &= \frac{1}{Z'} \exp[-\sum_{x \in \mathbf{Z}^d} v(\phi_x)] \gamma^H(d\phi) \end{aligned} \tag{3.6}$$

with the quartic interaction potential

$$v(\phi_x) = \frac{1}{2} r \phi_x^2 + \frac{1}{4} g \phi_x^4. \tag{3.7}$$

It can be shown that there actually exists a Gibbs measure for the Hamiltonian

$$H(\phi) = \frac{1}{2} \langle \phi, H_0 \phi \rangle + \sum_{x \in \mathbf{Z}^d} v(\phi_x) \tag{3.8}$$

if $g \geq 0$. For details see [1]. Alternatively one can work in a finite volume Λ_N but prove the convergence independently of N. (Cf.[2]).

The simplifying property of the hierarchical model is that ϕ' can be described on a measure space with a measure μ' of the same form as μ :

$$\mu'(d\phi') = \frac{1}{Z'} \exp[- \sum_{x \in \mathbf{Z}^d} v'(\phi'_x)] \gamma^H(d\phi') \tag{3.9}$$

with

$$e^{-v'(\phi'_x)} = \frac{\int \exp[- \sum_{y \in B_L(x)} v(L^{-\sigma} \phi'_x + \xi_y)] \gamma_\Gamma(d\xi)}{\int \exp[- \sum_{y \in B_L(x)} v(\xi_y)] \gamma_\Gamma(d\xi)}. \tag{3.10}$$

The integrals appearing in formula (3.10) are finite dimensional. Nevertheless this transformation is far from simple. However, one can make a Taylor expansion in g. Replacing v by the Wick-ordered expression

$$v(\phi) = \frac{1}{2} r : \phi^2 : + \frac{1}{4} g : \phi^4 : \tag{3.11}$$

and retaining terms up to second order one finds

$$v'(\phi') = \frac{1}{2} r' : \phi'^2 : + \frac{1}{4} g' : \phi'^4 : + O_3 \tag{3.12}$$

with

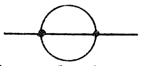
$$\begin{aligned} r' &= L^2(r - 3arg - 6cg^2) \\ g' &= L^{4-d}(g - 9ag^2) \end{aligned} \tag{3.13}$$

where

$$\begin{aligned} a &= 1 - L^{-d} \\ c &= 3L^{2-d}(1 - L^{-d}) \langle \phi^2 \rangle + (1 - L^{-d})(1 - 2L^{-d} + 2L^{-2d}) \end{aligned} \tag{3.14}$$

For $d = 3$ one finds that $v_{(m)}^{(m-n)}$ converges as $m \longrightarrow \infty$ if we put

$$\begin{aligned} r_m &= L^{-2m}(r_0 + 6mcg_0^2) \\ g_m &= L^{-m} g_0 \end{aligned} \tag{3.15}$$

Apart from the scaling factors L^{-2m} and L^{-m} there appears a non-trivial *mass-renormalization* term $6mcg_0^2$. It corresponds to the primitive divergent Feynman diagram 

A closer look at (3.13) shows that the substitution (3.15) might well be sufficient to all orders of perturbation theory. Indeed the mass-renormalization term appears because $g_m = O(L^{-m})$ while $r_m = O(L^{-2m})$. Using techniques developed by Gawedzki and Kupiainen [2] , this can actually be proved to be the case: see [1].

REFERENCES

[1]. T. C. Dorlas : On some aspects of renormalization group theory and hierarchical models, Thesis University of Groningen, 1987

[2]. K. Gawedzki & A. Kupiainen : Triviality of ϕ_4^4 and all that in a hierarchical model approximation, J.Stat.Phys. 29 (1982), 683.

ALGEBRAIC DUALITY OF MARKOV PROCESSES

Wim Vervaat
Catholic University, Nijmegen

1. Introduction

The following notion of duality has been exploited successfully in specific cases of interaction processes (see Liggett (1985) + references therein). Let $\eta = (\eta_t)_{t \geqslant 0}$ and $\zeta = (\zeta_t)_{t \geqslant 0}$ be Markov processes with measurable state spaces Y and Z, and let H be a real- or complex-valued measurable function on $Y \times Z$. Then η and ζ are said to be *dual* with respect to H if

(1.1) $\quad \mathsf{E}^{\eta_0=y} H(\eta_t, z) \; = \; \mathsf{E}^{\zeta_0=z} H(y, \zeta_t) \quad \text{for } y \in Y,\, z \in Z,\, t \geqslant 0.$

If Ω_η and Ω_ζ are the infinitesimal generators of η and ζ, so

$$\Omega_\eta f(y) \; := \; \lim_{t \downarrow 0} t^{-1}(\mathsf{E}^{\eta_0=y} f(\eta_t) - f(y)),$$

then (1.1) implies

(1.2) $\quad \Omega_\eta H(\cdot, z)(y) \; = \; \Omega_\zeta H(y, \cdot)(z),$

provided that $H(\cdot, z)$ and $H(y, \cdot)$ belong to the domains of the generators.

The notion of duality is useful because of the following result, a special case of a more general theorem of Cox & Rösler (1983) on entrance and exit laws.

Theorem 1. *Let η and ζ be dual Markov processes with respect to H. Suppose that H is non-negative or uniformly bounded, and that $(H(y, \cdot))_{y \in Y}$ determines probabilities on Z. Let ν be a probability measure on Z and set $h := \int_Z H(\cdot, z)\nu(dz)$. Then ν is an invariant distribution for ζ iff h is a harmonic function for η.*

Proof. By (1.1) we have

$$\begin{aligned} \mathsf{E}^{\eta_0=y} h(\eta_t) \; &= \; \int_Z \mathsf{E}^{\eta_0=y} H(\eta_t, z)\nu(dz) \\ &= \; \int_Z \mathsf{E}^{\zeta_0=z} H(y, \zeta_t)\nu(dz) \; = \; \mathsf{E}^{\mathrm{law}\,\zeta_0=\nu} H(y, \zeta_t). \end{aligned}$$

Now h is η-harmonic iff $\mathsf{E}^{\eta_0=y} h(\eta_t) = h(y)$ for $y \in Y$ iff $\mathsf{E}^{\mathrm{law}\,\zeta_0=\nu} H(y, \zeta_t) \; = \; \int_Z H(y, z)\nu(dz)$ for $y \in Y$ iff $\mathrm{law}^{\mathrm{law}\,\zeta_0=\nu} \zeta_t \; = \; \nu$ iff ν is ζ-invariant. □

Corollary. z_0 is an absorbing state for ζ iff the Dirac measure at z_0 is ζ-invariant iff $H(\cdot, z_0)$ is η-harmonic. In particular, z_0 is absorbing for ζ if $H(\cdot, z_0) \equiv 1$.

This result enables us to relate ergodicity of η to recurrence of ζ in case ζ has an absorbing state. In this way, duality has been applied successfully in specific cases.

2. Interpretation of duality

A question that arises naturally with the notion of duality is whether we can interpret (1.1) intuitively, for instance by constructing η and ζ on one probability space.

Although dual η and ζ have been constructed on one probability space in specific cases (Liggett (1985, §III.6), Clifford & Sudbury (1985)), it seems hard to find such a construction in general.

To understand what duality means, let us start with a Markov process η with state space Y, and let $(H(\cdot,z))_{z\in Z}$ be a 'base' in some sense of a sufficiently rich space of functions on Y. To determine thoughts, let $Z = \mathsf{N}$ and $(H(\cdot,z))_{z\in Z}$ be an orthonormal base of $\mathsf{L}^2(Y)$ (w.r.t. some σ-finite measure on Y). Suppose the mapping $y \mapsto \mathsf{E}^{\eta_0=y} H(\eta_t,z)$ belongs to the function space generated by $(H(\cdot,z))_{z\in Z}$, i.e., there is some signed bounded measure $\nu^t(z,\cdot)$ on Z such that

$$\mathsf{E}^{\eta_0=y} H(\eta_t,z) \;=\; \int_Z H(y,z')\nu^t(z,dz') \quad \text{for } y \in Y.$$

From the Chapman-Kolmogorov identity applied to the left-hand side it follows that

$$\nu^{t+u}(z,\cdot) \;=\; \int_Z \nu^t(z,dz')\,\nu^u(z',\cdot).$$

If there is a $y_0 \in Y$ with $H(y_0,\cdot) \equiv 1$, then it follows that $\nu^t(z,Z) = 1$. So $(\nu^t)_{t\geq 0}$ turns out to be the transition kernel of a Markov process ζ with state space Z, *provided that the* $\nu^t(z,\cdot)$ *are positive measures*.

This interpretation fails if the measures $\nu^t(z,\cdot)$ are not positive. However, in certain cases positive $\nu^t(z,\cdot)$ can be obtained by first enlarging the space Z in a suitable way (cf. Section 9).

3. The choice of the auxiliary function H

Another question is the choice of the auxiliary function H. Is there any recipe? In view of the previous section, one desired property is that $(H(\cdot,z))_{z\in Z}$ and $(H(y,\cdot))_{y\in Y}$ generate sufficiently rich classes of functions on Y and Z respectively.

One general recipe could be: $((H(\cdot,z))_{z\in Z}$ is a system of linearly independent vectors in a topological vector space of functions on Y, whose linear span is dense in this vector space. We did not pursue research in this direction, because we discovered that many examples in the literature have the following form (occasionally after simple transformations like $H \to 1-H$ which do not affect duality).

The spaces Y and Z are commutative semigroups. We denote the product in these semigroups by $+$, which need not be addition. The function H is a bihomomorphism from $Y \times Z$ into R or C with multiplication, i.e.,

$$\begin{aligned} H(y_1+y_2,z) &= H(y_1,z)\,H(y_2,z), \\ H(y,z_1+z_2) &= H(y,z_1)\,H(y,z_2). \end{aligned} \tag{3.1}$$

Whenever this situation occurs we say that H is a *duality between the semigroups Y and Z*. We call duality of Markov processes with respect to such an H *algebraic duality*.

Convenient additional properties of H are that H separates points in Y and Z, i.e., if $y_1 \neq y_2$, then $H(y_1,z) \neq H(y_2,z)$ for some $z \in Z$, and vice versa. We assume that Y and Z have zeros (if not, they can be added to them), so that $H(0,\cdot) \equiv 1$ and $H(\cdot,0) \equiv 1$. From the corollary at the end of Section 1 it follows that the zeros of Y and Z are absorbing states for η and ζ.

4 Examples of semigroups in duality

In (semi)group theory, a *character* of a commutative semigroup G is a homomorphism χ from G into $\mathbb{C}$ with multiplication: $\chi(x+y) = \chi(x)\chi(y)$ for $x,y \in G$. We exclude $\chi \equiv 0$. The characters form itself a semigroup under pointwise multiplication. Note that in the situation of (3.1) the $H(\cdot,z)$ are characters of Y and the $H(y,\cdot)$ are characters of Z.

If G is a topological group, then all bounded characters have values in the unit circle of $\mathbb{C}$, are continuous if they are measurable, and form a group on their own. This group is indicated by $G^{\wedge}$, the dual group. If G is locally compact with countable base, then $G^{\wedge\wedge} = G$ (Pontryagin's duality theorem). The theory of characters of semigroups is less smooth and less established. The best reference for probabilists is Berg et al. (1984).

A semigroup G is called *idempotent* if $x+x = x$ for all $x \in G$. In this case we can define a partial order $\leqslant$ in G by $x \leqslant y :\Leftrightarrow x+y = y$. It turns out that $0 \leqslant x$ for all x and that $x+y = x \vee y := \sup\{x,y\}$. We introduce the notation

$$\downarrow x \ := \ \{y \in G: y \leqslant x\}.$$

From $\chi(x) = \chi(x+x) = (\chi(x))^2$ it follows that all characters of G have values in $\{0,1\}$, so $\chi = 1_A$ for some $A \subset G$. One easily checks that 1_A is a character of G iff A is an *ideal* in G, i.e., A is decreasing ($\downarrow x \subset A$ for $x \in A$) and closed for finite suprema (see Berg et al. (1984, p.119), and Gierz et al. (1980) for ideals in lattices). In particular, the sets $\downarrow x$ are ideals, and the corresponding characters $1_{\downarrow x}$ separate the elements of G. Suppose that G is a lattice, i.e., the order allows an infimum besides the supremum. Then $1_{\downarrow x}1_{\downarrow y} = 1_{\downarrow x \cap \downarrow y} = 1_{\downarrow (x \wedge y)}$. So the collection of characters $1_{\downarrow x}$ is a semigroup isomorphic to G, but with $+ = \wedge$ rather than $\vee$. We have found the duality $Y = (G,\vee)$, $Z = (G,\wedge)$, $H(y,z) = 1_{\downarrow z}(y)$.

Here is a list of examples of semigroups in duality.

#	Y	+	Z	+	$H(y,z)$
1	$\mathbb{R}$	$+$	$\mathbb{R}$	$+$	e^{iyz}
2	$\mathbb{Z}$	$+$	$[0,1)$	$+ \bmod 1$	$e^{2\pi iyz}$
3	$\{0,1\}$	$+ \bmod 1$	$\{0,1\}$	$+ \bmod 1$	-1 if $y = z = 1$ 1 else
4	$[0,\infty]$	$+$	$[0,\infty]$	$+$	e^{-yz}
5	$\{0,1,\ldots,\infty\}$	$+$	$[0,1]$	$\times$	z^y
6	lattice L	$\vee$	L	$\wedge$	$1_{\downarrow z}(y)$
7	$\mathcal{M}(E)$	$+$	$C_k^+(E)$	$+$	$\exp -\int_E z\, dy$
8	$\bigcup_{m=0}^{\infty}\mathcal{M}(E^m)$	$\otimes$	$\bigcup_{n=0}^{\infty}C_k^+(E^n)$	$\otimes$	$\int_{E^{mn}} z^{\otimes n} d(y^{\otimes m})$ if $y \in \mathcal{M}(E^m)$ and $z \in C_k^+(E^n)$

Comments and explanations.

1. If H_j is a duality between the semigroups Y_j and Z_j for $j = 1,2$, then $H_1 \otimes H_2$ is a duality between $Y_1 \times Y_2$ and $Z_1 \times Z_2$. Here

$$H_1 \otimes H_2((y_1,y_2),(z_1,z_2)) \;:=\; H_1(y_1,z_1)\, H_2(y_2,z_2).$$

We write $H^{\otimes d} := H \otimes H \otimes \ldots \otimes H$ (d times). Then $H^{\otimes d}$ is a duality between Y^d and Z^d if H is between Y and Z.

2. Groups occur in rows 1, 2 and 3.

3. Important examples of lattices are $[0,1]$ with $\vee$, $[0,1]^d$ with $\vee$ componentwise, the closed subsets of a topological space E with $\cup$, the upper semicontinuous functions on E with $\vee$ pointwise.

4. In row 7, $\mathcal{M}(E)$ is the set of (positive) Radon measures on a locally compact space E with countable base; $C_k^+(E)$ is the set of positive continuous real-valued functions on E with compact support.

5. Notations in row 8. If $\mu \in \mathcal{M}(E^m)$ and $\nu \in \mathcal{M}(E^k)$, then $\mu \otimes \nu$ is the product measure and belongs to $\mathcal{M}(E^{m+k})$. If $f \in C_k^+(E^n)$ and $g \in C_k^+(E^l)$, then $f \otimes g(s,t) := f(s)g(t)$ for $s \in E^n$ and $t \in E^l$, and $f \otimes g$ belongs to $C_k^+(E^{k+l})$. Finally, $\mu^{\otimes n} := \mu \otimes \mu \otimes \ldots \otimes \mu$ (n times) and $f^{\otimes m} := f \otimes f \otimes \ldots \otimes f$ (m times). The spaces Y and Z with their duality can be extended in an obvious way to linear combinations: $\sum_{m=0}^{\infty} a_m \mu_m$ with $\mu_m \in \mathcal{M}(E^m)$ and $a_m \in \mathbb{R}_+$, and $\sum_{n=0}^{\infty} b_n f_n$ with $f_n \in C_k^+(E^n)$ and $b_n \in \mathbb{R}_+$.

5. Examples in the literature

Here we list which pairs of state spaces Y and Z and auxiliary functions H occur in the literature about Markov processes in duality. For examples treated in Liggett (1985) we do not mention the original source. We ignore possible extensions to larger state spaces in order to make speed functions positive (cf. end of Section 2). They will be considered in Section 9.

1. Holley & Stroock (1979): $Y = \mathbf{Z}^d$, $+$ = addition componentwise; $Z = [0,1)^d$, $+$ = addition mod 1 componentwise; $H(y,z) = e^{2\pi iyz}$ (cf. row 2).

2. Liggett (1985, §III.5): $Y = [0,1]$, $+ = \vee$; $Z = [0,1]$, $+ = \wedge$; $H(y,z) = 1$ if $y \leqslant z$, 0 else (cf. row 6).

3. Liggett (1985, §III.4, case H_1): S is a countable set; $Y = \{\text{subsets of } S\}$, $+ = \cup$; $Z = Y$, $+ = \cap$; $H(y,z) = 1$ if $y \subset z$, 0 else (cf. row 6 and comment 3).

4. Liggett (1985, §III.4, case H_2): S is a countable set; $Y = \{0,1\}^S$, $+$ = addition mod 1 componentwise;

$$Z = \{(z_s)_{s\in S} \in Y\colon z_s = 0 \text{ for all but finitely many } s\},$$

$+$ = addition mod 1 componentwise; $H(y,z) = (-1)^{\#\{s\in S:\, y_s = z_s = 1\}}$ (cf. row 3).

5. Dawson & Hochberg (1982): Y, Z and H as in row 8, or as in the last clause of the previous section.

6. Characteristic functions

In the next three sections we assume H to be a bounded duality between semigroups Y and Z (so $|H| \leqslant 1$: consider $H(my,nz)$ for natural m and n). Provide Y with the smallest σ-field $\mathcal{Y}$ for which all functions $H(\cdot,z)$ are measurable, and Z with the smallest σ-field $\mathcal{Z}$ for which all $H(y,\cdot)$ are measurable.

If μ is a bounded measure on $\mathcal{Y}$, then its *characteristic function* is the function $\hat{\mu}\colon Z \to \mathbf{C}$ defined by $\hat{\mu}(z) := \int_Y H(y,z)\mu(dy)$. Similarly, we set $\hat{\nu}(y) := \int_Z H(y,z)\nu(dz)$ for bounded measures ν on $\mathcal{Z}$. We do not know to which extent μ is determined by $\hat{\mu}$, but such results are known (including intrinsic characterizition of characteristic functions by positive definiteness) under additional conditions, for instance that Y consists of all bounded characters of a semigroup (cf. Berg et al. (1984)).

If ξ is a random variable with values in $(Y,\mathcal{Y})$ and with probability distribution μ, then $\hat{\mu} = \mathsf{E}H(\xi,\cdot)$ is called the characteristic function of ξ. If ξ_1 and ξ_2 are independent random variables in $(Y,\mathcal{Y})$, then

$$\mathsf{E}H(\xi_1+\xi_2,z) \;=\; \mathsf{E}H(\xi_1,z)H(\xi_2,z) \;=\; \mathsf{E}H(\xi_1,z)\,\mathsf{E}H(\xi_2,z),$$

so the characteristic function of $\xi_1+\xi_2$ is the product of the characteristic functions of ξ_1 and ξ_2.

Checking the examples in the table of Section 4 we find in row 1 for $Y = \mathbf{R}$ with addition the classical characteristic function $\mathsf{E}e^{iz\xi}$, in row 4 for $Y = [0,\infty]$ with addition the moment generating function $\mathsf{E}e^{-z\xi}$, in row 5 for $Y = \mathbf{N} \cup \{\infty\}$ the probability generating function $\mathsf{E}z^{\xi}$, in row 7 for Y = Radon measures on E with addition the Laplace transform $\mathsf{E}\exp -\int_E z\,d\xi$, and in row 6 for lattices Y with supremum the distribution function $\mathsf{P}[\xi \leq z]$.

7. Random walks in semigroups

In the present section we introduce random walks in semigroups. A modification of them will embody in the next section our major example of Markov processes in duality.

Let $(\xi_t)_{t=0}^{\infty}$ be a sequence of independent identically distributed random variables in $(Y,\mathcal{Y})$. Set

$$\eta_t^y := y + \textstyle\sum_{k=1}^{t} \xi_k \quad \text{for } y \in Y \text{ and } t = 0, 1, \ldots .$$

Then $(\eta_t^y)_{t=0}^{\infty}$ is a discrete-time random walk in Y with steps distributed as ξ_1, starting at y. Let φ be the characteristic function of ξ_1, and suppose that characteristic functions determine probability measures. Then the system $(\eta_t^y)_{t=0}^{\infty}$ for varying y is a Markov chain with state space Y and transitions determined by

$$\text{(7.1)} \qquad \mathsf{E}^{\eta_0=y} H(\eta_t, z) = H(y,z)\,(\varphi(z))^t .$$

A continuous-time random walk in Y is a Markov process η that satisfies (7.1) for real $t \geq 0$. Then the characteristic function φ must be *infinitely divisible,* i.e., φ^t is a characteristic function for all real $t > 0$. For general results about infinitely divisible characteristic functions on semigroups, see Berg et al. (1984, Ch.4).

From now on all random walks are continuous-time. We call φ the characteristic function of the steps. The random walk η has infinitesimal generator determined by

$$\text{(7.2)} \qquad \Omega_\eta H(\cdot,z)(y) = H(y,z)\log\varphi(z),$$

provided that the functions $H(\cdot,z)$ belong to the domain of Ω_η.

Well-known examples of random walks in semigroups that are not groups, are random walks in idempotent semigroups, for instance $[-\infty,\infty]$ with supremum. These random walks are known as *extremal processes.*

8. Tempered random walks

We now modify random walks by introducing a state-dependent clock $c(y)$ that moderates the speed of transitions from state y. This means that we replace (7.3) by

$$\text{(8.1)} \qquad \Omega_\eta H(\cdot,z)(y) = H(y,z)\,c(y)\log\varphi(z)$$

for some nonnegative function c on Y. We call η a *tempered random walk.* Suppose that there is an infinitely divisible characteristic function ψ of a probability measure on Z such that $c(y) = -\log\psi(y)$. Then (8.1) takes the form

$$\text{(8.2)} \qquad \Omega_\eta H(\cdot,z)(y) = -H(y,z)\log\psi(y)\log\varphi(z).$$

We now obtain the duality in (1.2) with ζ a random walk in Z with ψ as characteristic function of the steps and $-\log\varphi$ as clock speed, provided that $-\log\varphi$ is nonnegative and finite-valued. We have found:

Theorem 2. *If φ and ψ are infinitely divisible real characteristic functions without zeros of probability distributions on Y and Z repectively, then the tempered random walk η in Y with step characteristic function φ and clock speed $-\log\psi$ and the tempered random walk ζ in Z with step characteristic function ψ and clock speed $-\log\varphi$ are algebraically dual.*

From $\varphi(0) = \psi(0) = 1$ we find, in accordance with the corollary in Section 1, that the clock speeds at state 0 are 0, so the zeros of Y and Z are absorbing states.

Examples.

1. $Y = Z = \mathsf{R}$ with addition, $\varphi(y) = \psi(y) = \exp -\frac{1}{2}y^2$. We find that tempered Brownian motion with clock speed $\frac{1}{2}y^2$ is algebraically dual to itself.

2. $Y = \mathsf{Z}$ with addition, $Z = [0,1)$ with addition mod 1, $\varphi(z) = \exp(\cos 2\pi z - 1)$, $\psi(y) = \exp -y^2$. This example, or rather its d-dimensional version, is treated in Holley & Stroock (1979).

3. $Y = [0,1]$ with supremum, $Z = [0,1]$ with infimum, F and G are distribution functions on $[0,1]$, $\varphi(z) = F(z)$, $\psi(y) = 1-G(y)$. The tempered sup-extremal process with step distribution function F and clock speed $-\log(1-G)$ and the tempered inf-extremal process with step distribution function G and clock speed $-\log F$ are algebraically dual.

The major examples related to spin processes in Liggett (1984, §III.4) cannot be brought into this form.

9. Extension of state space

In the present section we only assume that Z is a semigroup and that $H(y,\cdot)$ is a character of Z for each $y \in Y$. Consequently, Y is a subset of the semigroup of all characters of Z, but we do not assume that Y is a semigroup itself.

Suppose the infinitesimal generator Ω_η has the form

$$(9.1) \qquad \Omega_\eta H(\cdot,z)(y) = \int_Y c(y,dy')(H(y',z)-H(y,z)),$$

where $c(y,\cdot)$ is a positive bounded measure for each $y \in Y$. In view of (1.2) one may try, in specific cases, to bring the right-hand side into the form

$$(9.2) \qquad \int_Z b(z,dz')(H(y,z')-H(y,z)) - V(z)H(y,z) =: \Omega_\zeta H(y,\cdot)(z),$$

often to find that the measures $b(z,\cdot)$ and the numbers $V(z)$ are not positive, as they should be. We ignore here additional problems of boundedness and convergence.

It may be helpful to extend Z to the semigroup $\tilde{Z} := Z\times\{0,1\}$, and H to the collection of characters of $\tilde{Z}$

$$\tilde{H}(y,(z,\varepsilon)) := H(y,z)(1-2\varepsilon)$$

in case H is real-valued, or to $\tilde{Z} := Z\times\mathsf{R}$ and

(9.3) $\quad \tilde{H}(y,(z,r)) \ := \ H(y,z)\,e^{ir}$

in case H is complex-valued.

Suppose we are given a class Σ of measurable function σ: $Y \to Y$ such that η has no transition from y to outside $\{\sigma(y)\colon \sigma \in \Sigma\}$, i.e., (9.1) can be rewritten as

(9.4) $$\Omega_\eta H(\cdot,z)(y) \ = \ \int_\Sigma \lambda(d\sigma)\gamma(\sigma,y)(H(\sigma(y),z)-H(y,z)),$$

where λ is some fixed σ-finite measure on Σ and γ is a nonnegative function. Suppose, moreover, that

(9.5) $$H(\sigma(y),z) - H(y,z) \ = \ \alpha(\sigma,y)\beta(\sigma,z)H(y,\tau(z))$$

for certain functions α: $\Sigma \times Y \to \mathbf{C}$, β: $\Sigma \times Z \to \mathbf{C}$ and τ: $Z \to Z$. Instances of (9.5) occur in Liggett (1985, p.159), where Σ consists of the flip operations in spin processes, and in dualities H of groups, where Σ consists of the translations of Y:

$$H(y+y',z) - H(y,z) \ = \ (H(y',z)-1)\,H(y,z).$$

Continuing with (9.4), (9.3) and (9.5) we find

$$\begin{aligned}
&\Omega_\eta \tilde{H}(\cdot,(z,r)) \ = \ e^{ir}\,\Omega_\eta H(\cdot,z)(y) \\
&= \ e^{ir} \int_\Sigma \lambda(d\sigma)\,\gamma(\sigma,y)\,(H(\sigma(y),z)-H(y,z)) \\
&= \ e^{ir} \int_\Sigma \lambda(d\sigma)\gamma(\sigma,y)\,\alpha(\sigma,y)\beta(\sigma,z)\,H(y,\tau(z)).
\end{aligned}$$

Suppose that $(H(\cdot,z'))_{z'\in Z}$ generates a sufficiently rich class of functions in the sense that there is a complex bounded measure ν_σ on Z such that

$$\gamma(\sigma,y)\,\alpha(\sigma,y) \ = \ \int_Z H(y,z')\,\nu_\sigma(dz').$$

Then we can continue the calculation by

$$\begin{aligned}
&\Omega_\eta \tilde{H}(\cdot,(z,r))(y) \ = \ \dots \\
&= \ e^{ir} \int_\Sigma \lambda(d\sigma) \int_Z \nu_\sigma(dz')\beta(\sigma,z)H(y,z')H(y,\tau(z)) \\
&= \ \int_\Sigma \lambda(d\sigma) \int_Z \nu_\sigma(dz')\,|\beta(\sigma,z)|\,\tilde{H}(y,(z'+\tau(z),r+\arg\beta(\sigma,z))).
\end{aligned}$$

Let $|\nu_\sigma|$ be the total variation measure of ν_σ. By Hahn decomposition there is a version of $d\nu_\sigma/d|\nu_\sigma|$ with values in the unit circle, say

$$\frac{d\nu_\sigma}{d|\nu_\sigma|}(z') \ = \ e^{ih_\sigma(z')}.$$

Then

$$\begin{aligned}
\Omega_\eta \tilde{H}(\cdot,(z,r))(y) \ = \ &\int_\Sigma \lambda(d\sigma) \int_Z |\nu_\sigma|(dz') \cdot \\
&\cdot\ |\beta(\sigma,z)|\,\tilde{H}(y,(z'+\tau(z),r+\arg\beta(\sigma,z)+h_\sigma(z'))).
\end{aligned}$$

Here the general analysis stops, and actually starts going the wrong way, since the weights of $\tilde{H}$ should not be positive, but occasionally positive and occasionally negative, in order to arrive at something like (9.2). This is a delicate choice depending on the case in question.

For a successful extension to $\tilde{Z} = Z \times \{0,1\}$ with Σ the flip operators, see Liggett (1985, §III.4). It would be interesting to know wether tempered random walks can be handled this way.

If also Y happens to be a semigroup, then it is tempting to extend Y as well, to $\tilde{Y} := Y \times \mathbb{R}$, and to extend H to the duality $\tilde{H}((y,s),(z,r)) := H(y,z)e^{isr}$ between the semigroups $\tilde{Y}$ and $\tilde{Z}$. However, if Markov processes η and ζ with state spaces $\tilde{Y}$ and $\tilde{Z}$ would be algebraically dual, then so would be their restrictions to $Y \times \{0\} \simeq Y$ and $Z \times \{0\} \simeq Z$. So the whole extension operation would have been unnecessary from the beginning. This is why we relaxed our assumptions in the first lines of this section.

References

[1] BERG, C., J.P.R. CHRISTENSEN & P. RESSEL (1984): *Harmonic Analysis on Semigroups*. Springer.

[2] CLIFFORD, P. & A. SUDBURY (1985): A sample path proof of the duality for stochastically monotone Markov processes. *Ann. Probab.* **13** 558-565.

[3] COX, J.T. & U. RÖSLER (1983): A duality relation for entrance and exit laws for Markov processes. *Stochastic Process. Appl.* **16** 141-156.

[4] DAWSON, D.A. & K.J. HOCHBERG (1982): Wandering random measures in the Fleming-Viot model. *Ann. Probab.* **10** 554-580.

[5] GIERZ, G., K.H. HOFMANN, K. KEIMEL, J.D. LAWSON, M. MISLOVE & D.S. SCOTT (1980): *A Compendium of Continuous Lattices*. Springer.

[6] HOLLEY, R. & D. STROOCK (1979): Dual processes and their applications to infinite interacting systems. *Adv. Math.* **32** 149-174.

[7] LIGGETT, T.M. (1985): *Interacting Particle Systems*. Springer.

Mathematisch Instituut
Katholieke Universiteit
Toernooiveld 5
6525 ED Nijmegen
The Netherlands

e-mail:
vervaa%kunivv1.mcvax@cwi.nl.UUCP

Complexity, randomness and unpredictability

Michiel van Lambalgen
Dept. of mathematics and computer science
Technical University Delft
Delft, Netherlands

1 Introduction It is one of the distinguishing features of a chaotic dynamical system that the time over which it can be predicted is rather limited, the exact time depending for one dimensional maps on the Lyapunov exponent and, generally, on the Kolmogorov entropy. For one dimensional maps of the unit interval the connection between Lyapunov exponent and the time over which the system can be predicted, can be stated informally as follows. Suppose we have a map $f: [0,1] \rightarrow [0,1]$ with positive Lyapunov exponent λ. It is assumed that we can determine the initial state of the system with finite precision only, say within an interval of length h; each iteration of f magnifies this margin of uncertainty, until it finally surpasses 1, which implies that we have completely lost track of the system. Calculation then shows that predictions about the system are no longer possible for times larger than the "unpredictability time" $t_U := -\lambda^{-1}\log h$.
In the general case, the role of the Lyapunov exponent is played by the Kolmogorov entropy H and we have that the unpredictability time t_U is proportional to $-H^{-1}\log h$, where h is the volume of a cell in phase space which contains the system initially (see Farmer [7]). It should be emphasized that t_U is an *average* and a property of the system, not of individual trajectories.
Below, we investigate a different concept of unpredictability, which results from taking the determinism of the system explicitly into account. We may now talk about the trajectory x of the system, consisting, say, of outcomes of measurements with finite precision; thus x may be taken to be an infinite sequence of numbers in $\{0,...,k-1\}$, for some natural number k.
Considering a trajectory x, we may try to measure the *difficulty* of computing a segment x(n) of x of length n, given an initial segment x(m), where $m < n$. We shall say that x(n) *is unpredictable given* x(m) if a minimal program that would compute x(n) given input x(m) is bigger than c, where c is some number too large for practical purposes.
The minimal $n_U(m)$ such that x(n) is unpredictable given x(m) is expressible in terms of m, c and the Kolmogorov entropy of the system; but in this case the resulting condition is (at least for almost all trajectories) more stringent than the one abtained above for t_U. Roughly speaking, n_U is equal to $m + {}^c/_H$, whereas the average unpredictability time t_U would be equal to ${}^m/_H$, which is usually larger.

2 Notation 2^ω is the set of infinite binary sequences. If $x \in 2^\omega$, then x(n) is the intial segment of x of length n, and x_n is the n^{th} term (also called coordinate) of x. The mapping

$T: 2^\omega \to 2^\omega$ (called the left shift) is defined by $(Tx)_n = x_{n+1}$.
$2^{<\omega}$ is the set of all finite binary sequences. An finite binary sequence is alternatively called a *word* or a *string*. The length of a word w is denoted $|w|$. 2^n is the set of all strings w such that $|w| = n$. If $m \le |w|$, then w(m) is the initial segment of w of length m, and w_m is the m^{th} term of w. If B is a set, 1_B denotes the characteristic function of B. Let $2 = \{0,1\}$ have the discrete topology and form the product topology on 2^ω. The open sets in this topology are then unions of *cylinders* [w] defined by $[w] := \{x \in 2^\omega | x(|w|) = w\}$. The topology on spaces of the form $(2^\omega)^m$ is constructed analogously.

3 Recursion theory It will be clear from the informal description of the contents of this article given above, that we shall state our results in terms of the theory of computable, or *recursive*, functions. It is out of the question to give here even the fundamental definitions, so an intuitive description must suffice. We shall take as primitive the notion of an algorithm operating on natural numbers, which yields as output natural numbers. It is understood that an algorithm need not terminate on every input. A *partial recursive function* $f: \omega \to \omega$ is a function which can be computed by an algorithm. With this intuitive description it is more or less clear that there exists an effective procedure which associates to each partial recursive function a natural number, its *Gödelnumber*. A consequence of the existence of a Gödelnumbering is that there exist *universal machines*, i.e. partial recursive functions which simulate every partial recursive function. A *recursive function* is a partial recursive function which is in fact total. More formal definitions of (partial) recursive function and Gödelnumber are possible; see Rogers [23]. The connection between the informal concept of an algorithm and the formal definition of a partial recursive function is provided by *Church's Thesis*, which states that every algorithm computes a partial recursive function. Although we defined partial recursive functions to have the natural numbers as domain and range, this restriction is not as severe as may seem, since many objects can be coded into the natural numbers. In particular, this is true for $\mathbb{Q}$ and $2^{<\omega}$. The following concepts thus make sense. A function $f: \omega \to \mathbb{R}$ is called *computable* if there exists a recursive function $g: \omega \times \omega \to \mathbb{Q}$ such that for all n,k: $|f(n) - g(n,k)| < 2^{-k}$. A measure μ on 2^ω is computable if there exists a recursive function $g: 2^{<\omega} \times \omega \to \mathbb{Q}$ such that for all w, n: $|\mu[w] - g(w,n)| < 2^{-k}$.

4 Complexity of finite strings There is a vast literature on random sequences ("typical" outcome sequences of some stochastic process), starting with von Mises' attempt [20] to use infinite random sequences as a mathematical foundation for probability theory. For a while it was thought that Kolmogorov's axiomatization of probability theory [10] dispensed with the need to provide a mathematical characterization of randomness. But Kolmogorov himself, in [11], emphasized that such a characterization is necessary to say precisely what the frequency interpretation of probability means. Kolmogorov's own attempt at a definition of randomness involved a measure of *complexity* of finite sequences. This complexity measure will be our main tool in investigating unpredictability properties of dynamical systems.
The intuition behind the definition of complexity of finite strings can be stated in various ways. One might say that if a sequence exhibits a regularity, it can be written as the output of a (simple) rule applied to a (simple) input. Another way to express this idea is to say that a sequence exhibiting a regularity can be *coded* efficiently, using the rule to produce the sequence from its code. Taking *rules* to be partial recursive functions from $2^{<\omega}$ to $2^{<\omega}$, we may define the *complexity* of a word w with respect to a rule A to be the length of a shortest

input p such that A(p) = w. Sequences with low complexity (with respect to A) are then supposed to be fairly regular (with respect to A). In order to take account of all possible rules (i.e. partial recursive functions), we then use a *universal* machine. One obtains different concepts of complexity by imposing additional restrictions on the functions A. We begin with Kolmogorov–complexity, where no such restrictions are imposed.

4.1 Definition Let $A: 2^{<\omega} \to 2^{<\omega}$ be a partial recursive function with Gödelnumber $\ulcorner A \urcorner$. The *complexity* $K_A(w)$ of w with respect to A is defined to be

$$K_A(w) = \begin{cases} \infty & \text{if there is no p such that } A(p) = w \\ |p| & \text{if p is a shortest input such that } A(p) = w. \end{cases}$$

A universal machine U is said to be *asymptotically optimal* if it is specified by the requirement that on inputs of the form $q = 0^{\ulcorner A \urcorner}1p$ (i.e. a sequence of $\ulcorner A \urcorner$ zeroes followed by a one, followed by a string p), U simulates the action of A on p. Fix a Gödelnumbering and an asymptotically universal machine U and put $K(w) := K_U(w)$. K is called the *Kolmogorov–complexity* of w (Kolmogorov [12]). Inputs will also be called *programs*. Clearly, we have

4.2 Lemma (a) For any partial recursive $A: 2^{<\omega} \to 2^{<\omega}$ and for all w, $K(w) \leq K_A(w) + \ulcorner A \urcorner + 1$; (b) for some constant c and for all w, $K(w) \leq |w| + c$.

Before we put the above definition to work, let us remark that complexity measures are not restricted to finite words over the alphabet {0,1}; *any* alphabet n = {0,.., n–1} will do. We only have to replace the functions $A: 2^{<\omega} \to 2^{<\omega}$ by functions which have as their range n^{ω}. Identifying a natural number with its binary representation, it makes sense to speak of the complexity of natural numbers. Similarly, given some recursive bijection $2^{<\omega} \to 2^{<\omega} \times 2^{<\omega}$, it makes sense to speak of the complexity of a *pair* of binary strings.

We now embark upon the promised definition of regular and irregular sequences. First suppose that $K(w) << |w|$; then for some algorithm A and input p such that both A and |p| are small compared to |w|, A(p) = w. In this case, we say that w exhibits a (simple) regularity. How small K(w) has to be is a matter of taste; we shall not be precise here. On the other hand, it is worthwhile to develop a theory of *irregularity* for finite sequences. Recall that for some c, $K(w) \leq |w| + c$. We wish to say that w is irregular if it is maximally complex. Formally:

4.3 Definition Fix some natural number m. A binary string w is called *irregular* if $|w| > m$ and $K(w) > |w| - m$.

The definition of irregularity is relative to the choice of m, but this is inessential for our (highly theoretical) purposes.
A simple counting argument will show that infinitely many irregular sequences exist. In the sequel, the expression "#A" always stands for the cardinality of the (finite) set A.

4.4 Lemma (a) $\#\{w \in 2^n | K(w) \leq n-m\} \leq 2^{n-m+1}-1$; (b) $\#\{w \in 2^n | K(w) > n-m\} > 2^n \cdot (1 - 2^{-m+1})$

Proof (a) The number of programs on U of length $\leq n-m$ is $\leq 2^{n-m+1}-1$. Hence (b) at least $2^n - 2^{n-m+1} = 2^n \cdot (1 - 2^{-m+1})$ sequences in 2^n satisfy $K(w) > n-m$. □

Note the extreme simplicity of the argument: it can be formalized in any formal system capable of handling finite sets of integers. This is to be contrasted with the following

Fact (Chaitin [3;6]) For any consistent formal system S, there exists a constant c_S, such that S proves *no* statement of the form "$K(w) > c_S$".

In other words, any consistent formal system can prove the irregularity of at most a finite number of irregular sequences. See van Lambalgen [16] for a critical discussion of this and related results.

While definition 4.1 captures the basic idea of a complexity measure for sequences, it is open to dispute whether it is really the most satisfactory definition. The intuition behind the definition is supposed to be that if p is a minimal program (on U) for w (i.e. a program of shortest length), then the *bits* of p contain all information necessary to reproduce w on U. But this might well be false: U might begin its operation by scanning all of p to determine its length, only then to read the contents of p bit for bit. In this way, the information p is really worth $|p| + \log_2|p|$ bits, so it's clear we have been cheating in calling $|p|$ the complexity of p. Chaitin [4;5] and Levin independently observed that we may circumvent this problem if we take seriously the idea of *coding* used to motivate the definition of K. The essence of a coding procedure is that a coded message be uniquely decipherable. In other words, there must be a unique way to chop the message up into code words. One way to ensure this property is to require that the set of code words is *prefixfree*: no code word may be a prefix (i.e. initial segment) of another.We may now introduce

4.5 Definition A *prefix algorithm* is a partial recursive function $A: 2^{<\omega} \to 2^{<\omega}$ which has a prefixfree domain.

We may now define a universal prefix algorithm as in definition 4.1: on inputs of the form q = $0^{\ulcorner A \urcorner}1p$, U simulates the action of A on p, where A is a prefix algorithm. We put

4.6 Definition Let A: $2^{<\omega} \rightarrow 2^{<\omega}$ be a prefix algorithm with Gödelnumber $\ulcorner A \urcorner$. The *complexity* (also called *information)* $I_A(w)$ of w with respect to A is defined to be

$$I_A(w) = \begin{cases} \infty & \text{if there is no p such that } A(p) = w \\ |p| & \text{if p is a shortest input such that } A(p) = w. \end{cases}$$

If U is the universal prefix algorithm constructed above, we let $I(w) := \min \{|p| \mid U(p) = w\}$.

This definition is due to Chaitin [4;5]; the notation "I(w)" derives from the formal similarities of this complexity measure with Shannon's measure of information. Indeed, the complexity measure I is not only conceptually cleaner than K, it has also a number of technical advantages, as will become gradually clear in the sequel. We first state some fundamental properties, parallel to those of K.

4.7 Lemma For some constant c and for all w: $I(w) \leq |w| + I(|w|) + c$.

4.8 Lemma (a) for some constant c: $\#\{w \in 2^n \mid I(w) \leq n + I(n) - m\} \leq 2^{n-m} \cdot c$; (b) for some constant c: $\#\{w \in 2^n \mid I(w) > n + I(n) - m\} > 2^n \cdot (1 - 2^{-m} \cdot c)$.

A proof of this lemma may be found in Chaitin [4,337] . It should be noted that, whereas the corresponding result for K was trivial, the proof of 4.8 is rather involved. This fact may add fuel to a nagging suspicion on the reader's part, that Chaitin's definition introduces only gratuitious complications. This impression, however, is mistaken; although proofs are sometimes more difficult, theorems and formulae generally take on a pleasanter aspect. An example will be given below; we shall meet another instance of this phenomenon in 4.13 where we define *conditional* complexity.

4.9 Example The main technical advantage of I lies in the fact that desirable results which hold for K only with logarithmic error terms, are now true within O(1). E.g. for K we have only: $K(\langle v,w \rangle) \leq K(v) + K(w) + \min[\log_2 K(v), \log_2 K(w)] + O(1)$, but the formula for I is more intuitive:

For some constant c, for all v,w: $I(\langle v,w \rangle) \leq I(v) + I(w) + c$.

One immediate application of the above formula for the complexity of a pair will illustrate its force: if T is the leftshift on 2^ω, we have for some constant c and all x in 2^ω,

$$I(x(n+m)) \leq I(x(n)) + I(T^n(x(n+m))) + c.$$

The sequence of functions $f_n(x) := I(x(n))$ thus forms a *subadditive* sequence and by the subadditive ergodic theorem[1], we have that for any ergodic measure μ there exists a constant H such that

$$\lim_{n\to\infty} \frac{I(x(n))}{n} = H \quad \mu\text{–a.e.}$$

It is, however, notoriously difficult to identify the limit of a subadditive process; eventually, we shall show that H equals the metric entropy of μ, but via an entirely different route. These considerations justify calling the property of I stated above *subadditivity*. End of the example.

Parallel to definition 4.3 we have

4.10 Definition Fix a natural number m. A binary word w is *irregular* if $I(w) > |w| + I(|w|) - m$.

By lemma 4.8, the great majority of binary strings is irregular in the new sense.
A very useful property of prefix algorithms, is given by

4.11 Lemma (a) If A is a prefix algorithm, then $\sum_{A(p)\text{ defined}} 2^{-|p|} \leq 1$. (b) $\sum_{w\in 2^{<\omega}} 2^{-I(w)} \leq 1$.

Proof (a) The cylinders in $\{[p] \mid A(p) \text{ defined}\}$ are pairwise disjoint. (b) Apply (a) to the universal prefix algorithm. □

The following result on the relation between K and I is due to Solovay [24]. Obviously, for all w: $K(w) \leq I(w)$.

4.12 Lemma For all w, $I(w) = K(w) + K[K(w)] + O(\log_2 K[K(w)])$.

The intuitive meaning of this expression is, that it takes $K[K(w)] + O(\log_2 K[K(w)])$ bits to turn a minimal program for w into a self delimiting program.

4.13 Conditional complexity In Chaitin's set–up, conditional complexity comes in two varieties. The most straightforward definition is the following. We consider algorithms B(p,q) in two arguments p and q. Such an algorithm is called a *prefix algorithm* if for each q, the set $\{p \mid B(p,q) \text{ defined}\}$ is prefixfree. We shall use U interchangeably for both the one–argument and the two–argument universal prefix algorithm.

4.14 Definition $I_0(w|v) := \min\{|p| \mid U(p,v) = w\}$.

For the second variant, denoted $I(w|v)$, we demand that U is presented, not with v itself, but rather with a minimal program for v.

4.15 Definition $I(w|v) := \min\{|p| \mid U(p,v^*) = w\}$, where v^* is some minimal program for v.

It will be seen in the sequel that both notions are useful. Some easy facts (given in Chaitin [4]):

4.16 Lemma For some constant c and all w: $I_0(w||w|) \leq |w| + c$.

4.17 Lemma For some constant c and for all w: $I(w||w|) \leq I_0(w||w|) + c$.

The main difference between I and I_0, however, is that the former satisfies

4.18 Lemma For some constant c, for all v,w: $|I(w|v) + I(v) - I(\langle w,v\rangle)| \leq c$.

This formula is proved in Chaitin [4,336] and is desirable if we think of I as giving the *information* of a string.

Conditional complexity is useful in stidying the problem of *prediction* of trajectories. Let x be a trajectory. Suppose we have observed x(m) and wish to predict the state of the system at instants m+1,...,n. We then look for a program p such that U(p,x(m)) = x(n), where U is the universal algorithm defined above. Typically, p contains the instructions for an algorithm and the instruction to compute the trajectory at instants m+1,...,n. There is a practical limit to the length of programs, say c. We then say that x(n) is not predictable (in the sense of: *not potentially predictable*) given x(m) if $I(x(n)|x(m)) > c$. We shall see below that, for typical trajectories x, the smallest n such that x(n) is not predictable given x(m) can be expressed in terms of m, c and the (metric) entropy of the system.

We shall need below the following upper bounds on I:

4.19 Lemma Let μ be a computable measure on 2^ω. Then for some c and all w:

$$I_0(w||w|) \leq [-\log_2\mu[w]] + c \text{ and } I(w||w|) \leq [-\log_2\mu[w]] + c.$$

As a consequence, for some c and all w:

$$I(w) \leq [-\log_2\mu[w]] + I(|w|) + c.$$

4.20 Lemma (Kolmogorov [12]) Let $x \in 2^\omega$. Fix an integer k and denote by $q_i(n)$ the relative frequency of the i^{th} word of length k in $x(n{\cdot}k)$. Then

$$I(x(n{\cdot}k)) \leq -n\cdot\sum_{i=1}^{2^k} q_i(n)\log_2 q_i(n) + I(n{\cdot}k) + O(\log_2 n).$$

It is instructive to compare the preceding lemma with lemma 4.19. Both determine an upper bound on I(w) in terms of probabilities; but in 4.20 these probabilities are the relative frequencies of small words in w, whereas in 4.19 the upper bound is derived using the frequency of w itself

5 Complexity and randomness for infinite sequences The obvious extension of definition 4.3: w is irregular if $K(w) > |w| - m$, to *infinite* sequences is of course: x is irregular if $\exists m \forall n (K(x(n)) > n - m)$. However, this attempt foundered upon the following obstacle:

5.1 Theorem (Martin-Löf [18]) For all x and for all m, there are infinitely many n such that $K(x(n)) \leq n - m$. More precisely, if $f: \omega \rightarrow \omega$ is a total recursive function such that $\sum_n 2^{-f(n)} = \infty$, then for all x there are infinitely many n such that $K(x(n)) \leq n - f(n)$.

Martin-Löf's theorem was considered to be a surprising result. In retrospect, it is somewhat difficult to understand why Martin-Löf's theorem should be surprising. After all, results indicating that total chaos in infinite binary sequences is impossible were known already. One example is van der Waerden's theorem (from 1928), which states that if the natural numbers are partitioned into two classes, then at least one of these classes contains arithmetic progressions of arbitrary lengths. Another example is a theorem in Feller [8,210] which states that if $a \in (0,1)$, then for μ_p–a.a. x, for infinitely many n, x_n is followed by a run of $[a{\cdot}\log_q n]$ 1's, where $q = p^{-1}$. In fact, the latter theorem can be used to give a new proof of Martin–Löf's theorem. We now sketch yet another proof (details are given in [16]).

Proofsketch Suppose for some m, $\{x | \ \forall n\ K(x(n)) > n - m\} \neq \emptyset$. This set is closed, hence can be represented as the set of infinite paths through a (recursive) binary tree T. T has (at least) two distinguished branches, the leftmost and the rightmost infinite path. Both branches are "simply definable", i.e. with an arithmetical formula involving just two quantifiers. But if x is "simply definable", its complexity is fairly low; in fact $n - K(x(n))$ is unbounded. Contradiction.

Similar results on oscillations hold for I, in this case with lower bound $n - m$ replaced by $n + I(n) - m$. Nevertheless, by weakening the proposed condition of irregularity to: x is irregular with respect to μ if $\exists m \forall n\, I(x(n)) > [-\log_2\mu[x(n)]] - m$, we get a non–vacuous condition, as the following theorem shows.

5.2 Theorem Let μ be a probability measure on 2^ω. Then $\mu\{x \mid \exists m \forall n\, I(x(n)) > [-\log_2\mu[x(n)]] - m\} = 1$.

Proof It suffices to show that $\mu\{x \mid \exists n\, I(x(n)) \leq [-\log_2\mu[x(n)]] - m\} \leq 2^{-m}$ for each m. We may write

$$\mu\{x \mid \exists n\, I(x(n)) \leq [-\log_2\mu[x(n)]] - m\} \leq \sum \{\mu[w] \mid w \in 2^{<\omega}, I(w) \leq [-\log_2\mu[w]] - m\};$$

however, since $I(w) \leq [-\log_2\mu[w]] - m$ iff $\mu[w] \leq 2^{-m}\cdot 2^{-I(w)}$, the right hand side of the above inequality is less than or equal to

$$\sum \{2^{-m}\cdot 2^{-I(w)} \mid w \in 2^{<\omega}, I(w) \leq [-\log_2\mu[w]] - m\} \text{ and since } \sum_{w\in 2^{<\omega}} 2^{-I(w)} \leq 1, \text{ this is } \leq 2^{-m}.$$

5.3 Remark Sequences irregular with respect to μ satisfy all "effective" probabilistic laws for μ; e.g. if μ is a Bernoulli distribution, a sequence irregular with respect to μ satisfies the law of large numbers, the law of the iterated logarithm etc. We shall not pursue this matter; see Martin–Löf [19].

6 Complexity, entropy and unpredictability The purpose of this section is to link I, which is a measure of disorder for *sequences*, with a more traditional measure of chaotic behaviour, defined for *dynamical systems*, namely metric entropy. This problem has received some attention in the physics literature (see Ford [9], Lichtenberg and Lieberman [17], Alekseev and Yakobson [1] and Brudno [2]), in connection with research on chaotic dynamical systems. It is shown here that if μ is an ergodic measure, then μ-a.a. x satisfy

$$\lim_{n\to\infty} \frac{I(x(n))}{n} = H(\mu),$$

where $H(\mu)$ is the metric entropy of μ. We use the theorem o elucidate the meaning of I in terms of (un)predictability.

6.1 Dynamical systems Our set–up is as follows. A *symbolic dynamical system* on a set of symbols $n = \{0,\ldots,n\text{-}1\}$ is a set $X \subseteq n^\omega$ (or $n^{\mathbb{Z}}$, as the case may be), together with the left–shift (or two–sided shift) T. We assume that X is closed under the action of T. Symbolic dynamical systems arise naturally in the study of general dynamical systems, in the following way.

Suppose (Γ,S) is a dynamical system, where Γ can be thought of as a phase space, equipped with a σ–algebra of measurable sets, and S is a measurable transformation on Γ,

which represents the evolution of the system, considered in discrete time. A measurement with finite accuracy on (Γ,S) is represented (ideally) by a measurable partition $A_0,\dots A_{n-1}$ of Γ, corresponding to "pointer readings" $0,\dots,n-1$.
Define a mapping $\psi: \Gamma \to n^\omega$ by $\psi(\gamma)_k = i$ iff $S^k(\gamma) \in A_i$; then $\psi(\gamma)$ represents the sequence of pointer readings obtained upon repeatedly measuring $\{A_0,\dots A_{n-1}\}$ on a system which is in state γ at time $t = 0$.
If the system (Γ,S) is also equipped with a probability distribution P, this distribution generates a measure μ on n^ω by $\mu A := P\psi^{-1}A$.
One may now study the dynamical system (Γ,S,P) by means of its symbolic representative $(\psi[\Gamma],T,\mu)$. In particular, the question whether, and to what extent, (Γ,S,P) displays chaotic behaviour can be investigated in this way. Below, we introduce various measures of disorder directly for *symbolic* dynamical systems, where for notational convenience we assume that the alphabet consists of just two symbols, 0 and 1. For an overview of the theory of dynamical systems, the reader may consult Petersen [21].

6.2 Metric entropy Let μ be a *stationary* measure on 2^ω; that is, for all Borel sets A, μ satisfies $\mu T^{-1}A = \mu A$. In other words, T conserves μ. For such measures, we may define the *metric entropy* $H(\mu)$ as follows:

$$H(\mu) := \lim_{n\to\infty} -\frac{1}{n}\sum_{w\in 2^n} \mu[w]\log_2\mu[w]. \quad \text{(Petersen [21,240])}$$

The interpretation of $H(\mu)$ is roughly as follows. $w \in 2^n$ is a possible series of outcomes if we perform n experiments upon the system under consideration. The probabilistic information present in w is (by definition) $-\log_2\mu[w]$; then

$$-\frac{1}{n}\sum_{w\in 2^n} \mu[w]\log_2\mu[w]$$

is the average amount of information gained per experiment if we perform n experiments. $H(\mu)$ is obtained if we let n go to infinity. A positive value of $H(\mu)$ indicates that each repetition of the experiment provides a non–negligable amount of information; systems with this property may be called random. Obviously, $H(\mu)$ is a global characteristic of the system $(2^\omega,T,\mu)$; it depends only on μ and T and reflects the the randomness of the system as a whole. We must now investigate how this global characteristic is related to randomness properties of individual sequences.

The measures occurring in 6 will be assumed to be *ergodic*; that is, if $T^{-1}A = A$, μA is either 0 or 1. If μ is ergodic, then $\mu[w]$ can be interpreted as the limiting relative frequency of w in a typical sequence x:

6.3 Ergodic theorem (see Petersen [21,30]) Let μ be a stationary measure on 2^ω, $f: 2^\omega \to \mathbb{R}$ integrable. Then

$$f^*(x) = \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f(T^k x)$$

exists μ-a.e., f^* is T-invariant and $\int f d\mu = \int f^* d\mu$. In addition, if μ is ergodic then f^* is constant μ-a.e. As a consequence, if μ is ergodic, then for any $w \in 2^{<\omega}$:

$$\mu\{x \mid \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} 1_{[w]}(T^k x) = \mu[w]\} = 1.$$

Below, we shall also need the Shannon – McMillan – Breiman theorem:

6.4 Theorem (see Petersen [21,261]) Let μ be an ergodic measure on 2^ω, $H(\mu)$ its entropy. Then for μ–a.a. x: $\lim_{n\to\infty} -\frac{\log_2 \mu[x(n)]}{n} = H(\mu)$.

The main result of this section is the computation of the constant H such that

$$\lim_{n\to\infty} \frac{I(x(n))}{n} = H \quad \mu\text{–a.e.}$$

We saw in 4 that this constant exists, due to the subadditivity of I; but we couldn't compute it.

6.5 Theorem Let μ be an ergodic measure, $H(\mu)$ its entropy. Then for μ–a.a. x:
$\lim_{n\to\infty} \frac{I(x(n))}{n} = H(\mu)^2$.

Proof Theorem 5.3 says that $\mu\{x \mid \forall m \exists n \; I(x(n)) > [-\log_2 \mu[x(n)]] - m\} = 0$. Using the Shannon – McMillan – Breiman theorem,

it follows that $\liminf_{n\to\infty} \frac{I(x(n))}{n} \geq H(\mu)$, for μ–a.a. x. To get $\limsup_{n\to\infty} \frac{I(x(n))}{n} \leq H(\mu)$ for μ–a.a. x, we remark first that, for each x and for each k, $\limsup_{n\to\infty} \frac{I(x(n))}{n} = \limsup_{n\to\infty} \frac{I(x(n\cdot k))}{n\cdot k}$.

Indeed, by the subadditivity of I, there exists a constant c such that for all k: $I(x(n)) = I(x(n_0 \cdot k + r)) \leq I(x(n_0 \cdot k)) + I(x_{n_0\cdot k+1}, \ldots, x_{n_0 \cdot k+r}) + c$.

Clearly, then, $\limsup_{n\to\infty} \frac{I(x(n))}{n} \leq \limsup_{n\to\infty} \frac{I(x(n\cdot k))}{n\cdot k}$; the converse inequality is trivial.

We now use lemma 4.20, slightly rephrased:

$$I(x(n\cdot k)) \leq n\cdot\left[-\sum_{w\in 2^k}\left(\frac{1}{n}\sum_{j=1}^{n}1_{[w]}(T^{j\cdot k}x)\log_2\left(\frac{1}{n}\sum_{j=1}^{n}1_{[w]}(T^{j\cdot k}x)\right)\right) + \frac{O(\log_2 n)}{n}\right],$$

which implies

$$(*)\quad \frac{I(x(n\cdot k))}{n\cdot k} \leq -\frac{1}{k}\sum_{w\in 2^k}\left(\frac{1}{n}\sum_{j=1}^{n}1_{[w]}(T^{j\cdot k}x)\log_2\left(\frac{1}{n}\sum_{j=1}^{n}1_{[w]}(T^{j\cdot k}x)\right)\right) + \frac{O(\log_2 n)}{n\cdot k}.$$

Since μ is stationary (although not necessarily ergodic) with respect to the T^k, the ergodic theorem implies that $f_w(x) := \lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}1_{[w]}(T^{j\cdot k}x)$ exists μ–a.e. and that $\int f_w d\mu = \mu[w]$.

Taking limsups (with respect to n) and integrals (with respect to μ) on the left hand side and right hand side of (*), we get, for all k:

$\int \limsup_{n\to\infty}\frac{I(x(n\cdot k))}{n\cdot k}d\mu \leq -\frac{1}{k}\sum_{w\in 2^k}\int f_w \log_2 f_w d\mu$; hence by Jensen's inequality

$$\int \limsup_{n\to\infty}\frac{I(x(n\cdot k))}{n\cdot k}d\mu \leq -\frac{1}{k}\sum_{w\in 2^k}\int f_w d\mu \log_2 \int f_w d\mu = -\frac{1}{k}\sum_{w\in 2^k}\mu[w]\log_2\mu[w].$$

Since $\limsup_{n\to\infty}\frac{I(x(n\cdot k))}{n\cdot k} = \limsup_{n\to\infty}\frac{I(x(n))}{n}$, we have, for each k: $\int \limsup_{n\to\infty}\frac{I(x(n))}{n}d\mu \leq$

$-\frac{1}{k}\sum_{w\in 2^k}\mu[w]\log_2\mu[w]$. Letting k go to infinity, we see that $\int \limsup_{n\to\infty}\frac{I(x(n))}{n} \leq H(\mu)$ and

the desired result follows since $\limsup_{n\to\infty}\frac{I(x(n))}{n}$ is T–invariant, hence constant μ–a.e. ▯

6.6 Remark Use of Solovay's formula (4.13) immediately gives $\lim_{n\to\infty}\frac{K(x(n))}{n} =$

$= H(\mu)$ μ–a.e., but employing I instead of K reduces one half of the proof to a triviality.

We now interpret the preceding theorem as a result on the amount of computer power necessary to predict the outcome sequence x(n), given x(m), where m < n. This problem arises for instance in the study of dynamical systems (Γ,S) on which we perform a measurement given by the partition $A_0,\ldots A_{k-1}$: we have observed the state of the system (i.e. one of the numbers 0,...,k–1) at instants t = 1,...,m and we wish to predict the state at instants t = m+1,...,n. For notational convenience we take k = 2, but the general case is treated analogously.

To calculate x(n) from x(m) we may use the evolution S, but other algorithms are also allowed. We impose but one restriction: the algorithm should not be too large. So we fix some constant c (representing the size of a program too large for practical purposes) and we call x(n) *unpredictable* given x(m) if $I(x(n)|x(m)) > c$.

We now show that there exists a close connection between entropy and unpredictability. Since c has been chosen so large, we may use the fundamental properties of I given in 4 to write the following chain of equivalent inequalities:

$$I(x(n)|x(m)) > c \Leftrightarrow$$

$$I(x(n)|x(m)) + I(x(m)) > c + I(x(m)) \Leftrightarrow$$

$$I(\langle x(n),x(m)\rangle) > c + I(x(m)) \Leftrightarrow$$

$$I(x(n)) + I(m) > c + I(x(m)) \Leftrightarrow$$

$$(*) \quad I(x(n)) > c + I(x(m)) - I(m).$$

Since $I(x(m)) \leq m + I(m) + d$, with $d << c$, (*) surely holds if $I(x(n)) > c + m$.

Now let μ be an ergodic measure with entropy $H(\mu)$ and suppose $\lim_{n\to\infty} \frac{I(x(n))}{n} = H(\mu)$.

Assume $H(\mu) > 0$, choose $\varepsilon > 0$ small compared to $H(\mu)$ and let n_0 be so large that $I(x(n)) > n(H(\mu) - \varepsilon)$ for $n \geq n_0$.

Then (*) is surely satisfied if $n > \frac{c+m}{H(\mu)-\varepsilon}$, an inequality which can thus be taken as a sufficient condition for unpredictability. If $\varepsilon << H(\mu)$, then this condition is approximately the same as the one given in the introduction for the *average* unpredictability time t_U.

But note that this condition can be significantly improved if we assume in addition that μ is computable. In this case we may replace the upper bound $I(x(m)) \leq m + I(m) + d$ by $I(x(m)) \leq [-\log_2\mu[x(m)]] + I(m) + d$. By the Shannon–McMillan–Breiman theorem (6.4), there is $m_0(\varepsilon)$ such that for $m \geq m_0(\varepsilon)$: $[-\log_2\mu[x(m)]] \leq m(H(\mu) + \varepsilon)$. For suitable choices of n and m the above sufficient condition for unpredictability can thus be sharpened to:

$$n > \frac{c + m(H(\mu) + \varepsilon)}{H(\mu) - \varepsilon}.$$

If $\varepsilon << H(\mu)$, then this boils down to: $n > m + \frac{c}{H(\mu)}$.

We thus see that investigating unpredictability properties of dynamical systems via their trajectories leads to sharper estimates.

6.6 Remark The above, sharper, estimate was derived under the assumption that the measure μ is computable. Some constructions of invariant measures, for instance those using (weak) limitpoints, do not automatically lead to *computable* measures. In such cases the result obtained above can be upheld using the technique of relative recursion.

Notes

1. For the subadditive ergodic theorem, see e.g. Y. Katznelson, B. Weiss, A simple proof of some ergodic theorems, *Isr. J. Math.*42 (1982) 291 – 300.
2. Brudno [2,132] proved: if μ is an ergodic measure, then for μ – a.a. x:

$$\limsup_{n\to\infty} \frac{K(x(n))}{n} = H(\mu).$$

The present formulation is theorem 5.5.2.5 in van Lambalgen [16].

Ackowledgement The author gratefully acknowledges support by the Netherlands Foundation for Scientific Research ZWO, under grant 22 – 110.

References

1. V.M. Alekseev, M.V. Yakobson, Symbolic dynamics and hyperbolic dynamical systems, *Physics Reports*, **75** (1981), 287-325.
2. A.A. Brudno, Entropy and the complexity of the trajectories of a dynamical system, *Trans. Mosc. Math. Soc.* **44** (1983), 127-151.
3. G.J. Chaitin, Information theoretic limitations on formal systems, *J. Ass. Comp. Mach.* **21** (1974), 403-424.
4. G.J. Chaitin, A theory of program size formally identical to information theory, *J. Ass. Comp. Mach.* **22** (1975), 329-340.
5. G.J. Chaitin, Algorithmic information theory, *IBM J. Res. Dev.* **21** (1977), 350-359;496.
6. G.J. Chaitin, Gödel's theorem and information, *Int. J. Theor. Phys.* **21** (1982), 941-954.
7. J.D. Farmer, Information dimension and the probabilistic structure of chaos, *Z. Naturforsch.* **37a**, 1304
8. W. Feller, Introduction to probability theory and its applications, Volume 1, 3rd ed.

Wiley (1968).

9. J. Ford, How random is a coin toss?, *Physics Today*, April 1983.
10. A.N. Kolmogorov, Grundbegriffe der Wahrscheinlichkeitsrechnung, *Ergebnisse der Mathematik und ihrer Grenzgebiete*, J. Springer (1933).
11. A.N. Kolmogorov, On tables of random numbers, *Sankhya* **25** (1963), 369-376.
12. A.N. Kolmogorov, Three approaches to the definition of the concept of "amount of information", *Sel. Transl. Math. Stat. and Prob.* **7**, AMS, Providence, R.I. (1968).
13. A.N. Kolmogorov, The logical basis for information theory and probability theory, *IEEE Trans. Inf. Theory* **IT-14** (1968), 662-664.
14. A.N. Kolmogorov, Combinatorial basis of information theory and probability theory, *Russ. Math. Surveys* **38** (1983), 29-40.
15. A.N. Kolmogorov, On logical foundations of probability theory, in: K. Itô, J.V. Prokhorov (eds), Probability theory and mathematical statistics, *Lecture Notes in Mathematics*, **1021**, Springer-Verlag (1984).
16. M. van Lambalgen, Random Sequences, Ph.D. dissertation, Dept. of Mathematics and Computer Science, University of Amsterdam, 1987
17. A.J. Lichtenberg, M.A. Lieberman, Regular and stochastic motion, Springer-Verlag (1983).
18. P. Martin-Löf, Algorithmen und zufällige Folgen, lecture notes, Universität Erlangen (1966). See also: Complexity oscillations in infinite binary sequences, *Z.Wahrsch. verw. Geb.* **19** (1971), 225-230.
19. P. Martin-Löf, The definition of random sequences, *Inf. and Control* **9** (1966), 602-619.
20. R. von Mises, Grundlagen der Wahrscheinlichkeitsrechnung, *Math. Z.* **5** (1919), 52-99.
21. K. Petersen, Ergodic theory, Cambridge University Press (1983).
23. H. Rogers, Theory of recursive functions and effective computability, McGraw Hill (1967).
24. R.M. Solovay, On random r.e. sets, in: A. Arruda, N. da Costa, R. Chuaqui (eds.), Non-classical logics, model theory and computability, North-Holland (1977).

Lecture I: The Large Deviation Principle in Statistical Mechanics

§1 Introduction

In May, 1971, Mark Kac visited Oxford; he gave me a Xerox copy of some pages from his note-book (see [1]) in which he had calculated correlation functions for the free-boson gas. He asked me about the relationship between his results and those of Araki and Woods [2]. With the help of Joe Pulé, then a graduate student at Oxford, I set to work. (The answer to Kac's question appeared in [3] and the full proofs were given in [4].) This was my initiation into statistical mechanics viewed from a probabilistic standpoint. There began my collaboration with Pulé which, I am glad to say, is still flourishing; in 1978 we were joined by Michiel van den Berg. Much of what I will say in these lectures will be a report on joint work with Pule and van den Berg; all three of us wish to acknowledge a debt of gratitude to Mark Kac for his warm encouragement, his generosity and his stimulating conversation.

§2 The Grand Canonical Pressure

Since the pioneering work of van Hove [7], the importance of proving the existence of thermodynamic functions in the thermodynamic limit has been recognized. We recall the definition of the grand canonical pressure: consider a sequence $\{\Lambda_\ell : \ell = 1, 2, \ldots\}$ of regions of Euclidean space R^d, and denote the volume of Λ_ℓ by V_ℓ; associated with each region Λ_ℓ is a countable set Ω_ℓ, the set of configurations of particles in Λ_ℓ; on Ω_ℓ are defined random variables $H_\ell : \Omega_\ell \to \mathbb{R}$ and $N_\ell : \Omega_\ell \to \mathbb{N}$; $H_\ell(\omega)$ is interpreted as the energy of the configuration ω and $N_\ell(\omega)$ as the number of particles in ω. The grand cononical measure $\mathbb{P}_\ell^\mu$ with chemical potential μ is defined on subsets of Ω_ℓ by

$$\mathbb{P}_\ell^\mu[A] = \Xi_\ell(\mu)^{-1} \sum_{\omega \in A} e^{\beta\{\mu N_\ell(\omega) - H_\ell(\omega)\}} \quad ; \tag{2.1}$$

here $\beta = 1/kT$ is the inverse temperature and $\Xi_\ell(\mu)$ is the grand canonical partition function given by

$$\Xi_\ell(\mu) = \sum_{\omega \in \Omega_\ell} e^{\beta\{\mu N_\ell(\omega) - H_\ell(\omega)\}} \quad . \tag{2.2}$$

The grand canonical pressure $p_\ell(\mu)$ is defined by

$$p_\ell(\mu) = (\beta V_\ell)^{-1} \ln \Xi_\ell(\mu) . \tag{2.3}$$

It is closely related to the cumulant generating function for the particle number density $X_\ell = N_\ell / V_\ell$; a straightforward calculation yields the formula

$$\int_{[0,\infty)} e^{\beta V_\ell x t} \mathbb{K}_\ell^\mu[dx] = e^{\beta V_\ell \{p_\ell(\mu+t) - p_\ell(\mu)\}} , \tag{2.4}$$

where $\mathbb{K}_\ell^\mu$ is the distribution function for X_ℓ defined by $\mathbb{K}_\ell^\mu = \mathbb{P}_\ell^\mu \circ X_\ell^{-1}$. We have introduced the concepts associated with the grand canonical pressure in the simplest case, namely, when Ω_ℓ is a countable set. But the formula (2.4) linking the distribution function $\mathbb{K}_\ell^\mu$ with the grand canonical pressure holds in wider contexts; for example, when Ω_ℓ is an aribtrary measure space carrying a pair of random variables H_ℓ and N_ℓ from which a grand canonical Gibbs measure can be defined, or when H_ℓ and N_ℓ are commuting self-adjoint operators on some hilbert space $\mathcal{H}_\ell$ such that trace$(e^{\beta\{\mu N_\ell - H_\ell\}})$ is finite. Our first assumption in the remainder of this note is that the distribution function $\mathbb{K}_\ell^\mu$ satisfies (2.4) for some function $p_\ell(\mu)$. Our second assumption is that $p(\mu) = \lim_{\ell\uparrow\infty} p_\ell(\mu)$ exists. From these two assumptions, much follows: the large deviation upper bound holds and, with it, the Berezin-Sinai criterion for a first-order phase-transition; the large deviation lower bound holds in the complement of first-order phase-transition segments; if the limit function $\mu \mapsto p(\mu)$ is differentiable at some μ_0 then $\{\mathbb{K}_\ell^{\mu_0} : \ell = 1,2,\ldots\}$ converges in distribution to the degenerate distribution concentrated at $p'(\mu)$.

In probability theory, it is natural to prove large deviation results for sums of independent, or weakly dependent, random variables: the Markov chain condition is an example of a condition of weak dependence. I claim that, in statistical mechanics, the natural condition of weak dependence is the existence of the pressure. In this lecture, we explore the consequences for the distribution of the particle number density of the existence of the pressure.

§3 The General Setting

For each μ in some open interval D of the real line, let $\{\mathbb{K}_\ell^\mu : \ell = 1,2,\ldots\}$

be a sequence of probability measures on $[0,\infty)$ satisfying

$$\text{(P1)} \qquad \int_{[0,\infty)} e^{V_\ell t x}\, \mathbb{K}_\ell^\mu[dx] = e^{V_\ell \{p_\ell(\mu+t) - p_\ell(\mu)\}} < \infty$$

where $\{V_\ell : \ell = 1, 2, \dots\}$ is a sequence of positive constants diverging to $+\infty$.

(P2) The limit $p(\mu) = \lim_{\ell\uparrow\infty} p_\ell(\mu)$ exists for all values of μ in the interval of definition, D.

The first consequence makes use of Hölder's Inequality and we omit the proof:

Lemma 1

Assume that (P1) holds; then $\mu \mapsto p_\ell(\mu)$ is convex. Assume, in addition, that (P2) holds; then $\mu \mapsto p(\mu)$ is convex.

Lemma 2

For all μ and $\mu+\alpha$ in the domain of definition of $\mathbb{K}_\ell^\mu$, the measures $\mathbb{K}_\ell^\mu$ and $\mathbb{K}_\ell^{\mu+\alpha}$ are mutually absolutely continuous:

$$\mathbb{K}_\ell^{\mu+\alpha}[dx] = e^{V_\ell c_\ell^\mu(x;\alpha)}\, \mathbb{K}_\ell^\mu[dx] \tag{3.1}$$

where

$$c_\ell^\mu(x;\alpha) = \alpha x + p_\ell(\mu) - p_\ell(\mu+\alpha) . \tag{3.2}$$

Proof:

$$\int_{[0,\infty)} e^{xt} . e^{V_\ell x\alpha}\, \mathbb{K}_\ell^\mu[dx] = e^{V_\ell\{p_\ell(\mu+\alpha+t/V_\ell) - p_\ell(\mu)\}}$$

by (P1); again, by (P1), we have

$$\frac{e^{V_\ell p_\ell(\mu+\alpha)}}{e^{V_\ell p_\ell(\mu)}} \int_{[0,\infty)} e^{xt}\, \mathbb{K}_\ell^{\mu+\alpha}[dx] = e^{V_\ell\{p_\ell(\mu+\alpha+t/V_\ell) - p_\ell(\mu)\}} .$$

The claim follows from the uniqueness theorem for Laplace transforms ■

Theorem 1

Assume that (P1) and (P2) hold and that p is differentiable at μ ; then

(1) the limit $\rho = \lim_{\ell \uparrow \infty} \int_{[0,\infty)} x \, \mathbb{K}_\ell^\mu[dx]$ exists.

(2) the sequence $\{\mathbb{K}_\ell^\mu : \ell = 1,2,\dots\}$ converges weakly to the degenerate distribution δ_ρ concentrated at ρ .

The proof utilises the convexity of the functions p_ℓ , $\ell = 1,2,\dots$ established in Lemma 1. This enables us to apply

Griffith's Lemma

Let $\{f_\ell : \ell = 1,2,\dots\}$ be a sequence of convex functions defined on a common open interval G converging pointwise to a function f. Let $\{x_\ell : \ell = 1,2,\dots\}$ be a sequence of points of G converging to a point x of G. Then

$$f'_-(x) \leq \liminf_{\ell \to \infty} (f_\ell)'_-(x_\ell) \leq \limsup_{\ell \to \infty} (f_\ell)'_+(x_\ell) \leq f'_+(x) .$$

(See [5] and references contained therein.)

Proof of Theorem 1:

By an elementary computation, $\int_{[0,\infty)} x \, \mathbb{K}_\ell^\mu[dx] = p'_\ell(\mu)$ since, by (P1), the moment generating function $s \mapsto \int_{[0,\infty)} e^{sx} \mathbb{K}_\ell^\mu[dx]$ of $\mathbb{K}_\ell^\mu$ is finite on a neighbourhood of zero. Since p is assumed to be differentiable at μ , it follows from Griffith's Lemma that $\{p'_\ell(\mu) : \ell = 1,2,\dots\}$ converges to $p'(\mu)$ since, by (P2), $\{p_\ell : \ell = 1,2,\dots\}$ converges pointwise to p on D. Thus (1) holds with $\rho = p'(\mu)$.

By (P1), we have

$$\int_{[0,\infty)} e^{sx} \mathbb{K}_\ell^\mu[dx] = e^{s\{p_\ell(\mu + s/V_\ell) - p_\ell(\mu)\}/(s/V_\ell)} \tag{3.3}$$

for s in a neighbourhood of zero. Fix s and put $\mu_\ell = \mu + s/V_\ell$; for ℓ sufficiently large, μ_ℓ is in D; moreover,

$\lim_{\ell\uparrow\infty} \mu_\ell = \mu$. By the convexity of $\mu \mapsto p_\ell(\mu)$, we have

$$(p_\ell)'_+(\mu) \leqslant \{p_\ell(\mu + s/V_\ell) - p_\ell(\mu)\}/(s/V_\ell) \leqslant (p_\ell)'_-(\mu_\ell) . \quad (3.4)$$

Since p is differentiable at μ , it follows from Griffith's Lemma that both $\{(p_\ell)'_+(\mu)\}$ and $\{(p_\ell)'_-(\mu_\ell)\}$ converge to $p'(\mu)$. Thus we have

$$\lim_{\ell\uparrow\infty} \int_{[0,\infty)} e^{sx}\, \mathbb{K}^\mu_\ell[dx] = e^{s\rho} . \quad (3.5)$$

But $\int_{[0,\infty)} e^{sx}\, \delta_\rho[dx] = e^{s\rho}$ so that (2) follows by the continuity and uniqueness theorems for the Laplace transform∎

Thus we have established that if the pressure p exists and is differentiable at μ then the sequence $\{\mathbb{K}^\mu_\ell : \ell = 1, 2, \dots\}$ satisfies the weak law of large numbers.

§4 The Heuristics of Large Deviations

Returning to the context of §2, we have the following reformulation of Theorem 1:

Suppose that the pressure $p = \lim_{\ell\uparrow\infty} p_\ell$ exists pointwise on the interval D on which the p_ℓ are defined and that p is differentiable at μ ; then

(1) the limit $\rho = \lim_{\ell\uparrow\infty} \mathbb{E}^\mu_\ell[X_\ell]$ exists.

(2) Let $g : [0,\infty) \longrightarrow \mathbb{R}$ be a bounded function which is continuous at ρ ; then $\lim_{\ell\uparrow\infty} \mathbb{E}^\mu_\ell[g(X_\ell)] = g(\rho)$.

Proof: (1) is a straight translation:

$$\mathbb{E}^\mu_\ell[X_\ell] = \sum_{\omega\in\Omega} X_\ell(\omega)\, \mathbb{P}^\mu_\ell[\omega] = \int_{[0,\infty)} x\, \mathbb{K}^\mu_\ell[dx] .$$

To prove (2), choose $\epsilon > 0$; by the continuity of g at ρ , there exists a neighbourhood I_ρ of ρ on which $|g(x) - g(\rho)| < \epsilon$. Now

$$g(\rho) - \mathbb{E}_\ell^\nu[g(X_\ell)] \;=\; \int_{[0,\infty)} (g(\rho) - g(x))\, \mathbb{K}_\ell^\nu[dx] \;. \tag{4.1}$$

Thus

$$\begin{aligned} \left| g(\rho) - \mathbb{E}_\ell^\nu[g(X_\ell)] \right| &\leq \int_{I_\rho} |g(\rho) - g(x)|\, \mathbb{K}_\ell^\nu[dx] \\ &\quad + \int_{I_\rho^c} |g(\rho) - g(x)|\, \mathbb{K}_\ell^\nu[dx] \\ &\leq \epsilon + 2M\, \mathbb{K}_\ell^\nu[I_\rho^c]\,, \end{aligned} \tag{4.2}$$

where $M = \sup_{[0,\infty)} g(x)$. But by Theorem 1, $\mathbb{K}_\ell^\nu \to \delta_\rho$; this means that there exists ℓ_0 such that, for all $\ell > \ell_0$, we have $\mathbb{K}_\ell^\nu[I_\rho^c] < \epsilon$. Hence, for all $\ell > \ell_0$,

$$\left| g(\rho) - \mathbb{E}_\ell^\nu[g(X_\ell)] \right| \leq \epsilon\,(1 + 2M)\;; \tag{4.3}$$

but ϵ was an arbitrary positive number, so that

$$\lim_{\ell\to\infty} \mathbb{E}_\ell^\nu[g(X_\ell)] \;=\; g(\rho) \quad \blacksquare \tag{4.4}$$

It sometimes happens in statistical mechanics that we find it interesting to introduce a perturbed grand canonical measure $\tilde{\mathbb{P}}_\ell^\nu$ which is conveniently defined via its expectation functional $\tilde{\mathbb{E}}_\ell^\nu$:

$$\tilde{\mathbb{E}}_\ell^\nu[A] \;=\; \frac{\mathbb{E}_\ell^\nu[A\, e^{\beta V_\ell u(X_\ell)}]}{\mathbb{E}_\ell^\nu[e^{\beta V_\ell u(X_\ell)}]} \tag{4.5}$$

where u is a continuous function on $[0,\infty)$ which is bounded above. Can we say anything about

$$\lim_{\ell\to\infty} \tilde{\mathbb{E}}_\ell^\nu[g(X_\ell)] \;?$$

If the V_ℓ factor were absent from the exponent, we could conclude that the limit would be the same as before: $g(\rho)$. However, the presence of the factor V_ℓ causes the fluctuations in X_ℓ to contribute to the limiting value; we expect that the answer will be $g(\tilde{\rho})$ where $\tilde{\rho} \neq \rho$, in general. There are two ways of proving this; they are closely related, as we might expect.

We can introduce a perturbed Hamiltonian $\tilde{H}_\ell(\omega) = H_\ell(\omega) + V_\ell(u \circ X_\ell)(\omega)$ and use it to define a perturbed pressure $\tilde{p}_\ell(\mu)$. A straightforward manipulation gives

$$\tilde{p}_\ell(\mu) = p_\ell(\mu) + \frac{1}{\beta V_\ell} \ln \mathbb{E}_\ell^\mu [e^{\beta V_\ell u(X_\ell)}] . \tag{4.6}$$

Since

$$\mathbb{E}_\ell^\mu [e^{\beta V_\ell u(X_\ell)}] = \int_{[0,\infty)} e^{\beta V_\ell u(x)} \mathbb{K}_\ell^\mu [dx] , \tag{4.7}$$

proof of the existence of the pressure

$$\tilde{p}(\mu) = \lim_{\ell \to \infty} \tilde{p}_\ell(\mu)$$

amounts to proving the existence of the limit

$$\lim_{\ell \to \infty} \frac{1}{\beta V_\ell} \ln \int_{[0,\infty)} e^{\beta V_\ell u(x)} \mathbb{K}_\ell^\mu [dx] .$$

conditions on $\{ \mathbb{K}_\ell^\mu : \ell = 1, 2, \ldots \}$ sufficient to ensure this were given by Varadhan [6] in a general setting, and we will give a precise statement of them in the next section. Roughly speaking, they are that there exists a function $I^\mu(\cdot) : [0,\infty) \to [0,\infty]$ such that $\mathbb{K}_\ell^\mu[dx] \sim e^{-\beta V_\ell I^\mu(x)} dx$ in our case, the only zero of I^μ is at $x = \rho = p'(\mu)$ so that $I^\mu(\cdot)$ determines the rate at which $\mathbb{P}_\ell^\mu[A]$ goes to zero if ρ is not in A. Intuitively, one would expect that

$$\lim_{\ell\to\infty} \frac{1}{\beta V_\ell} \ln \int_{[0,\infty)} e^{\beta V_\ell u(x)} \mathbb{K}_\ell^\mu[dx] = \sup_{[0,\infty)} \{u(x) - I^\mu(x)\}; \quad (4.8)$$

this is the conclusion of Varadhan's First Theorem. It follows that the perturbed pressure $\tilde{p}(\mu)$ exists and is given by

$$\tilde{p}(\mu) = p(\mu) + \sup_{[0,\infty)} \{ u(x) - I^\mu(x) \} , \quad (4.9)$$

provided that u and $\{\mathbb{K}_\ell^\mu\}$ satisfy the hypotheses of the theorem. If $\tilde{p}$ is differentiable at μ, it follows from our previous argument that

$$\lim_{\ell\to\infty} \tilde{\mathbb{E}}_\ell^\mu [g(X_\ell)] = g(\tilde{\rho}) \quad (4.10)$$

where now $\tilde{\rho} = \tilde{p}'(\mu)$.

Another way of computing this limit is via Varadhan's Second Theorem: if the supremum $\sup_{[0,\infty)} \{u(x) - I^\mu(x)\}$ is attained at an isolated point x*, then

$$\lim_{\ell\to\infty} \tilde{\mathbb{E}}_\ell^\mu [g(X_\ell)] = g(x^*).$$

We shall see, in our case, that the supremum is attained at an isolated point if and only if $\tilde{p}$ is differentiable at μ and then $x^* = p'(\mu)$.

We have seen that large deviations from the mean (deviations on the scale of V_ℓ) are of importance in the evaluation of

$$\lim_{\ell\to\infty} \frac{\mathbb{E}_\ell^\mu [g(X_\ell) e^{\beta V_\ell u(X_\ell)}]}{\mathbb{E}_\ell^\mu [e^{\beta V_\ell u(X_\ell)}]} .$$

It is for this reason that we are interested in the rate at which the degenerate distribution is approached; those distributions which approach the degenerate distribution exponentially fast are said to satisfy the large deviation principle. Next, we turn to the precise definition of this concept.

§5 Varadhan's Theorems

Donsker initiated the study of singular perturbations of partial differential equations by means of functional integration; he showed how, in the case of Burger's equation, a transformation introduced by Hopf can be used to convert the equation to a linear equation which can be solved as a function space integral. The perturbation problem can then be studied by an analysis of the asymptotic behaviour of function space integrals; in the case of Burger's equation, the asymptotic analysis was carried out by Schilder [8]. Varadhan [6] showed how a class of such problems can be treated using more general families of measures on function space whose asymptotic behaviour is to be investigated; in §3 of [6], Varadhan gave an account of the asymptotic analysis in an abstract setting. Subsequently, in a sequence of papers, Donsker and Varadhan applied these methods to a wide variety of problems involving stochastic processes (a full bibliography can be found in Varadhan's monograph [9]). The method is a far-reaching generalization of the saddle-point method (Laplace's method) for one-dimensional integrals. It is our experience that whenever, in statistical mechanics, an author claims to use the saddle-point method, an efficient way of giving a rigorous proof is to check that the hypotheses of Varadhan's Theorems are verified. For that reason, we summarize here the results proved in §3 of [6].

Let E be a complete separable metric space; let $\{\mathbb{K}_\ell : \ell = 1,2,\dots\}$ be a sequence of probability measures on the σ- field of Borel subsets of E and let $\{V_\ell : \ell = 1,2,\dots\}$ be a sequence of non-negative numbers such that $V_\ell \to \infty$. We say that $\{\mathbb{K}_\ell\}$ obeys the large deviation principle with constants $\{V_\ell\}$ and rate-function $I(\cdot)$ if there exists a function $I: E \to [0,\infty]$ satisfying:

(LD1): $I(\cdot)$ is lower semi-continuous on E.

(LD2): For each finite m, $\{x : I(x) \leq m\}$ is compact.

(LD3): For each closed subset C of E, $\limsup_{\ell\to\infty} \frac{1}{V_\ell} \ln \mathbb{K}_\ell[C] \leq -\inf_C I(x)$.

(LD4): For each open subset G of E, $\liminf_{\ell\to\infty} \frac{1}{V_\ell} \ln \mathbb{K}_\ell[G] \geq -\inf_G I(x)$.

For example, if $I(\cdot)$ is a lower semi-continuous function whose level sets are compact and m is a σ- finite measure on E such that $x \mapsto e^{-I(x)}$ is integrable with respect to m, and $\{V_\ell\}$ is a sequence of non-negative numbers such that $V_\ell \to \infty$, then the sequence $\{\mathbb{K}_\ell\}$ of probability measure defined by

$$\mathbb{K}_\ell[A] = \frac{\int_A e^{-V_\ell I(x)} m(dx)}{\int_E e^{-V_\ell I(x)} m(dx)} \tag{5.1}$$

satisfies the large deviation principle with constants $\{V_\ell\}$ and rate-function $I(\cdot)$. The definition above has the advantage that it does not require the existence of a reference measure such as m. We are now in a position to state

Varadhan's First Theorem

Let $\{\mathbb{K}_\ell : \ell = 1, 2, \ldots\}$ be a sequence of probability measures on E obeying the large deviation principle with constants $\{V_\ell\}$ and rate-function $I(\cdot)$. Then, for any continuous function G on E which is bounded above, we have

$$\lim_{\ell \to} \frac{1}{V_\ell} \ln \int_E e^{V_\ell G(x)} \mathbb{K}_\ell[dx] = \sup_E \{G(x) - I(x)\} .$$

The condition that G be bounded above can be weakened; it is enough to suppose that $\sup\{G(x) : x \in \bigcup_{\ell \geq 1} \operatorname{supp} \mathbb{K}_\ell\}$ is finite. The theorem can be extended to cover the situation where the function G is replaced by a sequence of functions $\{G_\ell : \ell = 1, 2, \ldots\}$; this is Theorem 3,4 of [6].

Let $\tilde{\mathbb{K}}_\ell$ be defined by

$$\tilde{\mathbb{K}}_\ell[A] = \frac{\int_A e^{V_\ell G(x)} \mathbb{K}_\ell[dx]}{\int_E e^{V_\ell G(x)} \mathbb{K}_\ell[dx]} ;$$

Varadhan's Second Theorem gives sufficient conditions for the sequence of perturbed measures to converge weakly to a degenerate distribution.

Varadhan's Second Theorem

Suppose that $\Lambda = \sup_E \{G(x) - I(x)\}$ is attained at a point x^* of E and that

$$\sup_{\{x : d(x, x^*) \geq \epsilon\}} \{G(x) - I(x)\} < \Lambda$$

for every $\epsilon > 0$; if g is a bounded function on E which is continuous at x^* then

$$\lim_{\ell \to \infty} \int_E g(x)\, \tilde{\mathbb{K}}_\ell[dx] = g(x^*) .$$

§6 The Upper Bound in the General Setting

In this section we return to the programme, begun in §3, of exploring the consequences of the existence of the pressure in the thermodynamic limit.

Theorem 2

Let $\{\mathbb{K}_\ell : \ell = 1, 2, \cdots\}$ be a sequence of probability measures on $[0, \infty)$ satisfying (P1) and (P2); then (LD3) holds with rate-function $I^\mu(\cdot)$ given by

$$I^\mu(x) = p(\mu) + f(x) - \mu x ,$$

where $f(\cdot)$ is the free-energy, the Legendre transform of $p(\cdot)$: $f(x) = \sup\{\mu x - p(\mu)\}$

Proof:

First consider an interval $I_1 = [0, \rho_1]$ with $\rho_1 < p'_-(\mu)$ (since $\mu \mapsto p(\mu)$ is convex, the left-hand derivative $p'_-(\mu)$ and the right-hand derivative $p'_+(\mu)$ exist for all μ in D). For each ℓ and each $\alpha < 0$, we have

$$\mathbb{K}^\mu_\ell[I_1] = \int_{[0,\infty)} 1_{[0,\rho_1]}(x)\, \mathbb{K}^\mu_\ell[dx] \leq \int_{[0,\infty)} e^{V_\ell \alpha(x-\rho_1)}\, \mathbb{K}^\mu_\ell[dx]$$

$$= e^{V_\ell\{p_\ell(\mu+\alpha) - p_\ell(\mu) - \alpha\rho_1\}} . \qquad (6.1)$$

Thus

$$\limsup_{\ell \to \infty} \frac{1}{V_\ell} \ln \mathbb{K}^\mu_\ell[I_1] \leq p(\mu+\alpha) - p(\mu) - \alpha\rho_1 , \quad \alpha < 0 . \qquad (6.2)$$

It follows that

$$\limsup_{\ell\to\infty} \frac{1}{V_\ell} \ln \mathbb{K}_\ell^\mu[I_1] \leqslant \inf_{\alpha<0} \{p(\mu+\alpha) - p(\mu) - \alpha\rho_1\}$$

$$= -\left(p(\mu) + \sup_{\alpha'<\mu}\{\alpha'\rho_1 - p(\alpha')\} - \mu\rho_1\right) . \qquad (6.3)$$

But

$$\sup_{\alpha<\mu}\{\alpha\rho_1 - p(\alpha)\} = \sup_{\alpha}\{\alpha\rho_1 - p(\alpha)\}$$

since $\rho < p'_-(\mu)$; hence

$$\limsup_{\ell\to\infty} \frac{1}{V_\ell} \ln \mathbb{K}_\ell^\mu[I_1] \leqslant -I^\mu(\rho_1) . \qquad (6.4)$$

Next consider $I_2 = [\rho_2, \infty)$ where $\rho_2 > p'_+(\mu)$. It follows in analogous fashion that

$$\limsup_{\ell\to\infty} \frac{1}{V_\ell} \ln \mathbb{K}_\ell^\mu[I_2] \leqslant -I^\mu(\rho_2) . \qquad (6.5)$$

Now let C be an arbitrary closed subset of $[0,\infty)$; if $C \cap [p'_-(\mu), p'_+(\mu)]$ is non-empty then $\inf_C I^\mu(x) = 0$ and the inequality holds trivially since, for an arbitrary Borel set A, we have $\mathbb{K}_\ell^\mu[A] \leqslant 1$; on the other hand, if $C \cap [p'_-(\mu), p'_+(\mu)]$ is empty, let (ρ_1, ρ_2) be the largest open interval containing $[p'_-(\mu), p'_+(\mu)]$ which does not intersect C so that $C \subset [0,\rho_1] \cup [\rho_2,\infty)$ and

$$\limsup_{\ell\to\infty} \frac{1}{V_\ell} \ln \mathbb{K}_\ell^\mu[C] \leqslant \limsup_{\ell\to\infty} \frac{1}{V_\ell} \ln\{\mathbb{K}_\ell^\mu[I_1] + \mathbb{K}_\ell^\mu[I_2]\}$$

$$\leqslant \left(\limsup_{\ell\to\infty} \frac{1}{V_\ell} \ln \mathbb{K}_\ell^\mu[I_1]\right) \vee \left(\limsup_{\ell\to\infty} \frac{1}{V_\ell} \ln \mathbb{K}_\ell^\mu[I_2]\right)$$

$$= (-I^\mu(\rho_1)) \vee (-I^\mu(\rho_2))$$

$$= -\inf_C I^\mu(x) , \qquad (6.6)$$

since $x \mapsto I^\mu(x)$ is decreasing on $[0, p'_-(\mu)]$ and increasing on $[p'_+(\mu), \infty)$ ∎

§7 The Berezin-Sinai Criterion for a First-Order Phase-Transition

We call an interval $[x_1,x_2]$ of the positive real axis a first-order phase-transition segment if the free-energy function $x \mapsto f(x)$ is linear for $x_1 \leqslant x \leqslant x_2$. Since $f(x) = \sup_{\mu}\{\mu x - p(\mu)\}$, each first-order phase-transition segment corresponds to a point μ at which the grand canonical pressure is non-differentiable: there exists μ such that

$$[x_1, x_2] = [p'_-(\mu), p'_+(\mu)] .$$

On such an interval, the pressure as a function of the density is constant. In [10], Berezin and Sinai established a criterion for the existence of a first-order phase-transition segment; Dobrushing [11] simplified the proof considerably, pointing out that the criterion reduces the question of the existence of a phase-transition to the question of a "violation of the law of large numbers" in the grand canonical ensemble. Here we point out that the proof of the Berezin-Sinai criterion makes use only of the large-deviation upper bound and this, as we have seen, holds whenever the pressure exists.

The Berezin-Sinai Criterion

Suppose that, for some μ_0 of μ, the rate function is symmetric about some point x_0:

$$I^{\mu_0}(x_0+y) = I^{\mu_0}(x_0-y)$$

for all y. Suppose also that for some $\delta > 0$:

$$\mathbb{P}_\ell^{\mu_0}[\,|X_\ell - x_0| \geqslant \delta\,] \geqslant c > 0$$

for all ℓ sufficiently large. Then there is a first-order phase-transition at μ_0 and the interval $[x_0 - \delta, x_0 + \delta]$ is contained in the phase-transition segment $[p'_-(\mu_0), p'_+(\mu_0)]$.

Proof

Let $C = (-\infty, x_0-\delta] \cup [x_0+\delta, \infty)$; then, by hypothesis

$$\liminf_{\ell \to \infty} \frac{1}{V_\ell} \ln \mathbb{K}_\ell^{\mu_0}[C] \geqslant 0 \; ; \tag{7.1}$$

by Theorem 2,

$$\limsup_{l\to\infty} \frac{1}{V_l} \ln \mathbb{K}_l^{\mu_0}[C] \leq -\inf_C I^{\mu_0}(x) \leq 0 \ . \tag{7.2}$$

Hence $\inf_C I^{\mu_0}(x) = 0$; by the symmetry of $I^{\mu_0}(\cdot)$ about x_0,

$$\inf\{ I^{\mu_0}(x) \ : \ x \in (-\infty, x_0-\delta] \}$$

$$= \inf\{ I^{\mu_0}(x) \ ; \ x \in [x_0+\delta, \infty)\} = 0 \tag{7.3}$$

so that $x_0 - \delta$ and $x_0 + \delta$ must lie in $[p_-'(\mu_0), p_+'(\mu_0)]$ and therefore

$$[x_0-\delta, x_0+\delta] \subset [p_-'(\mu_0), p_+'(\mu_0)] \quad \blacksquare$$

§8 The Lower Bound in the General Setting

We define the first-order phase-transition set F to be the union of the first-order phase-transition segments:

$$F = \bigcup_{\mu\in S} [p_-'(\mu), p_+'(\mu)]$$

where

$$S = \{\mu : p_-'(\mu) \neq p_+'(\mu)\}$$

Theorem 3

Let $\{\mathbb{K}_l : l = 1, 2, \dots\}$ be a sequence of probability measures on $[0,\infty)$ satisfying (P1) and (P2); let G be an open subset of ran $\partial p \setminus F$ where $\partial p(\mu)$ is the sub-differential of p at μ; then

$$\liminf_{l\to\infty} \frac{1}{V_l} \ln \mathbb{K}_l^{\mu}[G] \geq -\inf_G I^{\mu}(x) \ .$$

Proof:

Let y be an arbitrary point of G; choose δ so that the neighbourhood

$B_y^\delta = (y-\delta, y+\delta)$ is contained in G; then

$$\mathbb{K}_\ell^\nu[G] \geqslant \mathbb{K}_\ell^\nu[B_y^\delta] = \int_{B_y^\delta} \mathbb{K}_\ell^\nu[dx] = \int_{B_y^\delta} e^{-V_\ell c_\ell^\nu(x;\alpha)} \mathbb{K}_\ell^{\nu+\alpha}[dx] \quad (8.1)$$

for all ν in the domain D of $\mathbb{K}_\ell^\nu$ (by Lemma 2 of §3). Now choose α so that $y = p'(\nu+\alpha)$; this is possible because, by hypothesis, y is in ran $\partial p \setminus F$. Then

$$\mathbb{K}_\ell^\nu[B_y^\delta] = e^{-V_\ell c_\ell^\nu(y;\alpha)} \int_{B_y^\delta} e^{-V_\ell \alpha(x-y)} \mathbb{K}_\ell^{\nu+\alpha}[dx]$$

$$\geqslant e^{-V_\ell c_\ell^\nu(y;\alpha)} e^{-V_\ell \delta|\alpha|} \mathbb{K}_\ell^{\nu+\alpha}[dx] \quad (8.2)$$

and, by Theorem 1, $\{\mathbb{K}_\ell^{\nu+\alpha}\} \underset{w}{\rightarrow} \delta_y$ so that $\mathbb{K}_\ell^{\nu+\alpha}[B_y^\delta] > \frac{1}{2}$ for all ℓ sufficiently large. Hence

$$\liminf_{\ell\to\infty} \frac{1}{V_\ell} \ln \mathbb{K}_\ell^\nu[G] \geqslant -I^\nu(y) - \delta|\alpha| ;$$

but δ was an arbitrary positive number and y an arbitrary point of G; it follows that

$$\liminf_{\ell\to\infty} \frac{1}{V_\ell} \ln \mathbb{K}_\ell^\nu[G] \geqslant \sup_G(-I^\nu(y)) = -\inf_G I^\nu(y) \blacksquare \quad (8.3)$$

§9 The Large Deviation Principle in the General Setting

In this section we put together the results of §6 and §8. First, we note some properties of the rate-function stemming from the convexity of the pressure $\nu \mapsto p(\nu)$. The free-energy f(·) is the Legendre transform of p(·):

$$f(x) = \sup_\nu \{\nu x - p(\nu)\} . \quad (9.1)$$

Hence

$$I^\nu(x) = p(\nu) + f(x) - \nu x \geqslant 0 .$$

We may regard $I^{\mu}(\cdot)$ itself as the Legendre transform of the convex function $\alpha \mapsto p(\mu+\alpha) - p(\mu)$; it follows that $x \mapsto I^{\mu}(x)$ is a closed convex function and hence lower semi-continuous, so that (LD1) holds. Since $I^{\mu}(x) + p(\mu+\alpha) - p(\alpha) - \alpha x \geqslant 0$, it follows that on the level set

$$L_m = \{ x : I^{\mu}(x) \leq m \} \tag{9.2}$$

we have

$$\alpha x \leqslant m + p(\mu+\alpha) - p(\mu); \tag{9.3}$$

hence, for a > 0

$$\begin{aligned} a x &= \sup_{\alpha \in [-a,a]} \alpha x \\ &\leqslant m + \sup_{\alpha \in [-a,a]} p(\mu+\alpha) - p(\mu) < \infty . \end{aligned} \tag{9.4}$$

It follows that L_m is bounded; since $x \mapsto I^{\mu}(x)$ is lower semi-continuous, L_m is closed; hence L_m is compact and so (LD2) holds.

In §6 we saw that (LD3) holds whenever the pressure exists in the thermodynamic limit; on the other hand, it is clear from §8 that more is required for (LD4) to hold since it asserts the lower bound for all open sets while the existence of the pressure suffices to establish the lower bound only for open subsets of ran $\partial p \backslash F$. A sufficient condition for (LD4) to hold is that p exists and is differentiable on the whole of R and that ran $p' = [0,\infty)$; this is far from being necessary, however, as can be seen from the case of the free boson gas. Nevertheless, this condition is satisfied sufficiently often to make the following theorem useful:

Theorem 4

Let $\{\mathbb{K}_\ell : \ell = 1,2,\ldots\}$ be a family of sequences of probability measures on $[0,\infty)$ defined for all values of μ in $\mathbb{R}$ and satisfying (P1) and (P2). Suppose that p(·) is differentiable and that ran $p' = [0,\infty)$: then, for each value of μ , the sequence $\{\mathbb{K}^{\mu}_\ell : \ell = 1,2,\ldots\}$ satisfies the large deviation principle with constants $\{V_\ell\}$ and rate-function $I^{\mu}(\cdot)$ given by

$$I^{\mu}(x) = p(\mu) + f(x) - \mu x$$

where

$$f(x) = \sup_{\mu} \{ \mu x - p(\mu) \} .$$

§10 Remarks

I have attempted in this lecture to set out the results described in [5] in the framework established by Varadhan [6], thus showing the probabilistic consequences of the existence of the grand canonical pressure in the thermodynamic limit. Independently of [5], Ellis [12] proved a large deviation result for vector-valued random variables; his basic hypothesis is the existence of the limit of a sequence of cumulant generating functions, and Theorem 4 can be deduced from his theorem.

REFERENCES

[1] Kac, M.,: O Idealnym Gazie Bosego-Einsteina.
Statistical Theories for Solids, Fluids and Gases
(Proc. Working Conf., Jabłonna, 1972) 432-449
Ossolineum-Wydawn Polskiej Akad. Nauk, Wrocław (1974).

[2] Araki, H., Woods, E.J.: Representations of the Canonical Commutation Relations Describing a Non-Relativistic Free Boson Gas, J. Math. Phys. 4, 637-662 (1963).

[3] Lewis, J.T. : The Free Boson Gas. Proceedings of the LMS Instructional Conference, Bedford College 1971, Mathematical Problems in Mathematical Physics, New York: Academic Press 1972.

[4] Lewis, J.T., Pulé, J.V.: The Equilibrium States of the Free Boson Gas, Commun. Math. Phys. 36, 1-18 (1974).

[5] J.T. Lewis and J.V. Pulé: The Equivalence of Ensembles in Statistical Mechanics, in Stochastic Analysis and its Applications, Proceedings, Swansea 1983, ed. A. Truman and D. Williams, LNM 1095, Springer: Heidelberg 1984.

[6] S.R.S. Varadhan: Asymptotic Probabilities and Differential Equations, Comm. Pure Appl. Math. 19, 261-286 (1966).

[7] L. van Hove: Quelques Propriétés Générales de l'Intégrale de Configuration d'Un Système de Particules à Une Dimension, Physica 15, 951-961 (1949).

[8] M. Schilder: Some Asymptotic Formulae for Wiener Integrals, Trans. Amer. Math. Soc. 125, 63-85 (1966).

[9] S.R.S. Varadhan: Large Deviations and Applications, Philadelphia: Society for Industrial and Applied Mathematics 1984.

[10] F.A. Berezin and Ya. G. Sinai: Existence of a Phase-Transition of the First Kind in a Lattice Gas with Attraction between Particles, Trudy Mosk. Mat. Obshch 17, 197-212 (1967).

[11] R.L. Dobrushin: Existence of a Phase-transition in Models of a Lattice Gas, Proc. Fifth Berkeley Smyposium, III, 73-87 (1967).

[12] R.S. Ellis: Large Deviations for a General Class of Random Vectors, Ann. Probab. 12, 1-12 (1984).

Lecture II: Large Deviations and the Boson Gas

§1 Introduction

In this lecture we review some large deviation results for probability distributions associated with the free boson gas and discuss briefly their application to models of an interacting boson gas. In §2 we describe the probabilistic setting; in §3 we review results on the free boson gas which we shall require; in §4 and §5 we summarize some large deviation results; in §6 we sketch an application.

§2 The Probabilistic Setting

Our ultimate aim is to compute thermodynamic functions for certain models of an interacting boson gas. The physical relevance of these calculations will not be discussed here; we shall concentrate on the probabilistic aspects of the investigation.

The probability space Ω on which the models are defined is the space of terminating sequences of non-negative integers: an element ω of Ω is a sequence

$$\{\omega(j) \in \mathbb{N} : j = 1, 2, \dots\}$$

satisfying $\sum_{j \geq 1} \omega(j) < \infty$.

The basic random variables, the occupation numbers, are the evaluation maps $\sigma_j : \Omega \to \mathbb{N}$ given by

$$\sigma_j(\omega) = \omega(j) . \tag{2.1}$$

The sequence $\{H_\ell : \ell = 1, 2, \dots\}$ of free-gas hamiltonians is defined by

$$H_\ell(\omega) = \sum_{j \geq 1} \lambda_\ell(j) \sigma_j(\omega) , \tag{2.2}$$

where $\{\lambda_\ell(j) : j = 1, 2, \dots\}$ is a ordered sequence of real numbers associated with a region Λ_ℓ of some Euclidean space R^d :

$$0 = \lambda_\ell(1) \leq \lambda_\ell(2) \leq \cdots \tag{2.3}$$

The total number of particles $N(\omega)$ is defined by

$$N(\omega) = \sum_{j \geq 1} \sigma_j(\omega) \ . \tag{2.4}$$

We define, for $\mu < 0$, the grand canonical measure $\mathbb{P}_\ell^\mu[\cdot]$ on Ω and the grand canonical pressure $p_\ell(\mu)$:

$$\mathbb{P}_\ell^\mu[\omega] = \frac{e^{\beta\{\mu N(\omega) - H_\ell(\omega)\}}}{e^{\beta V_\ell p_\ell(\mu)}} \ , \tag{2.5}$$

where

$$e^{\beta V_\ell p_\ell(\mu)} = \sum_{\omega \in \Omega} e^{\beta\{\mu N(\omega) - H_\ell(\omega)\}} \ . \tag{2.6}$$

Because of (2.3), both (2.5) and (2.6) hold for $\mu < 0$. The mean particle number density $\mathbb{E}_\ell^\mu[X_\ell]$, where $X_\ell = N/V_\ell$ and $\mathbb{E}_\ell^\mu[\cdot]$ denotes the expectation with respect to the probability measure $\mathbb{P}_\ell^\mu[\cdot]$, is given by

$$\mathbb{E}_\ell^\mu[X_\ell] = p_\ell'(\mu) \ . \tag{2.7}$$

Using an identity known to Euler, we have

$$e^{\beta V_\ell p_\ell(\mu)} = \prod_{j \geq 1} (1 - e^{\beta(\mu - \lambda_\ell(j))})^{-1} , \tag{2.8}$$

so that we write

$$p_\ell(\mu) = V_\ell^{-1} \sum_{j \geq 1} p(\mu | \lambda_\ell(j)) \tag{2.9}$$

where the partial pressure $p(\mu|\lambda)$ is given by

$$p(\mu|\lambda) = \beta^{-1} \ln (1 - e^{\beta(\mu - \lambda)})^{-1} . \tag{2.10}$$

Lemma 1. For each $\mu < 0$, the occupation numbers are independent, geometrically distributed random variables:

$$\mathbb{P}_\ell^\mu[\sigma_j \geq m] = e^{m\beta(\mu - \lambda_\ell(j))} \ . \tag{2.11}$$

Proof: For $\alpha_j \leq 0$, $j = 1, 2, \ldots$ we have

$$\mathbb{E}_\ell^\mu [e^{\beta \sum_{j\geq 1} \alpha_j \sigma_j}] = \prod_{j \geq 1} \frac{(1 - e^{\beta(\mu - \lambda_\ell(j))})}{(1 - e^{\beta(\mu + \alpha_j - \lambda_\ell(j))})} \blacksquare$$

It is convenient to introduce the distribution function

$$F_\ell(\lambda) = V_\ell^{-1} \# \{ j : \lambda_\ell(j) \leq \lambda \} \; ; \tag{2.12}$$

with respect to this, (2.9) can be rewritten as

$$p_\ell(\mu) = \int_{[0,\infty)} p(\mu|\lambda) \, dF_\ell(\lambda) \; ; \tag{2.13}$$

the mean particle density is given by

$$\mathbb{E}_\ell^\mu [X_\ell] = \int_{[0,\infty)} p'(\mu|\lambda) \, dF_\ell(\lambda) \; . \tag{2.14}$$

We note that, for each ℓ , $\mu \mapsto p_\ell(\mu)$ is a convex function defined on $(-\infty, 0)$; we define

$$p_\ell(0) = \lim_{\mu \uparrow 0} p_\ell(\mu) = +\infty \tag{2.15}$$

and

$$p_\ell(\mu) = +\infty \; , \quad \mu > 0 \; . \tag{2.16}$$

Then each p_ℓ is a closed convex function defined on the whole of $\mathbb{R}$; its essential domain is

$$\mathrm{dom}\, p_\ell = (-\infty, 0).$$

In order to prove the existence of the pressure in the thermodynamic limit, it is necessary to make some assumptions about the $\lambda_\ell(j)$ and the V_ℓ; putting $\phi_\ell(\beta) = \int_{[0,\infty)} e^{-\beta\lambda} dF_\ell(\lambda)$, we formulate conditions:

(S1) $$\phi(\beta) = \lim_{\ell\to\infty} \phi_\ell(\beta)$$

exists for all β in $(0,\infty)$.

(S2) $\phi(\beta)$ is non-zero for at least one value of β in $(0,\infty)$.

These conditions are weak restrictions on the sequences; their verification in a particular instance can involve some hard analysis.

§3 Results Concerning the Free Boson Gas

In this section we review some results on the general theory of the free boson gas; the proofs can be found in [1].

Proposition 1. Suppose that (S1) and (S2) hold; then the following limits exist.

(1) $$p(\mu) = \lim_{\ell\to\infty} p_\ell(\mu), \qquad \mu < 0,$$

(2) $$F(\lambda) = \lim_{\ell\to\infty} F_\ell(\lambda).$$

They are related by

$$p(\mu) = \int_{[0,\infty)} p(\mu|\lambda)\, dF(\lambda).$$

Moreover, we have

$$p'(\mu) = \int_{[0,\infty)} p'(\mu|\lambda)\, dF(\lambda).$$

The standard example is the following one: let $h_\ell = -\frac{1}{2}\Delta$ in Λ_ℓ with Dirichlet conditions on $\partial\Lambda_\ell$ where $\{\Lambda_\ell : \ell = 1, 2, \dots\}$ is a sequence of dilations of a convex region in R^d which eventually fills out the whole of $\mathbb{R}^d$;

let $\varepsilon_\ell(1) \leq \varepsilon_\ell(2) \leq \cdots$ be the eigenvalues of h_ℓ and put $\lambda_\ell(j) = \varepsilon_\ell(j) - \varepsilon_\ell(1)$; then (S1) and (S2) hold and $F(\lambda) = C_d \lambda^{d/2}$.

Next we define the critical density ρ_c :

if $\lambda \mapsto p'(0|\lambda)$ is integrable on $[0,\infty)$ with respect to F, put

$$\rho_c = \int_{[0,\infty)} p'(0|\lambda)\, dF(\lambda)\,; \tag{3.1}$$

put $\rho_c = \infty$ otherwise.

It follows from the dominated convergence principle that if ρ_c is finite then

$$\rho_c = \lim_{\mu \uparrow 0} \int_{[0,\infty)} p'(\mu|\lambda)\, dF(\lambda) = \lim_{\epsilon \downarrow 0} \int_{[\epsilon,\infty)} p'(0|\lambda)\, dF(\lambda). \tag{3.2}$$

Clearly, if $F(\lambda) \sim \lambda^\sigma$ with $\sigma > 1$ then ρ_c is finite; if ρ_c is finite then $F(\lambda) \to 0$ as $\lambda \downarrow 0$. (In fact, we have the more precise estimate: for $\epsilon > 0$, $F(\epsilon) < \beta\epsilon e^{\beta\epsilon}\rho_c$. Note that in the standard example, ρ_c is finite if and only if d > 2.

Again it is convenient to follow the standard conventions for convex functions in extending p to the whole of R: we define $p(0) = \lim_{\mu \uparrow 0} p(\mu)$ and put $p(\mu) = +\infty$, $\mu > 0$. Since p is convex and differentiable for $\mu < 0$, $p'_-(0) = \lim_{\mu \uparrow 0} p'(\mu) = \rho_c$. Define $p'_+(0)$ to be $+\infty$ and $p'_-(\mu) = p'_+(\mu) = +\infty$ for $\mu > 0$. Then p is a closed convex function on the whole of R.

The sub-differential ∂p is given by

$$(\partial p)(\mu) = \begin{cases} p'(\mu)\,, & \mu < 0\,, \\ [\rho_c,\infty)\,, & \mu = 0\,. \end{cases} \tag{3.3}$$

For fixed ℓ, the function $\mu \mapsto p'_\ell(\mu)$ is strictly increasing on $(-\infty,0)$ and $p'(\mu) \to 0$ as $\mu \to -\infty$ while $p'_\ell(\mu) \to \infty$ as $\mu \to 0$ since $\lambda_\ell(1) = 0$. It follows that the equation

$$p'_\ell(\mu) = \rho \tag{3.4}$$

has a unique solution $\mu_\ell(\rho)$ in $(-\infty, 0)$ for each ρ in $(0,\infty)$. On the other hand, for $\rho_c < \infty$, the function $\mu \mapsto p'(\mu)$ increases from zero to ρ_c as μ ranges through $(-\infty, 0)$. It is convenient to define $\mu(\rho)$ for ρ in $(0,\infty)$ to be the unique root of

$$p'(\mu) = \rho \tag{3.5}$$

if $\rho < \rho_c$ and to be zero if $\rho \geqslant \rho_c$.

Defining

$$\pi_\ell(\rho) = (p_\ell \circ \mu_\ell)(\rho)$$

so that $\pi_\ell(\rho)$ is the pressure at mean density ρ and $\pi = p \circ \mu$, we have

Proposition 2. Suppose that (S1) and (S2) hold; then

(1) $$\lim_{\ell\to\infty} \mu_\ell(\rho) = \mu(\rho),$$

(2) $$\lim_{\ell\to\infty} \pi_\ell(\rho) = \pi(\rho),$$

(3) $$f(x) \stackrel{\text{def}}{=} \sup_{\mu<0} \{\mu x - p(\mu)\} = x\mu(x) - \pi(x).$$

Thus we have a first-order phase-transition when ρ_c is finite; the first-order phase-transition segment is $[\rho_c, \infty)$.

§4 Large Deviations of the Particle Number Density

Let $\mathbb{K}_\ell^\mu = \mathbb{P}_\ell^\mu \circ X_\ell^{-1}$ be the distribution function of the particle number density $X_\ell = N/V_\ell$. It follows from Theorem 1 that, for $\mu < 0$, $\{\mathbb{K}_\ell^\mu\}$ converges weakly to the degenerate distribution δ_ρ concentrated at $\rho = p'(\mu)$. It follows from Theorem 2 that the Large Deviation upperbound (LD3) holds for $\mu<0$ with rate-function $I^\mu(\cdot)$ given by

$$I^\mu(x) = p(\mu) + f(x) - \mu x . \tag{4.1}$$

However, the existence of the pressure is not sufficient to ensure that the Large Deviation lowerbound holds for an arbitrary open subset of $[0,\infty)$ when ρ_c is finite; although ran $\partial p = [0,\infty)$ the existence of the first-order phase-transition segment $[\rho_c,\infty)$ prevents an application of Theorem 3 to the whole of $[0,\infty)$. Nevertheless, as we shall see, special features of the free boson gas enable us to establish the Large Deviation lowerbound (LD4).

Theorem 1 Suppose that (S1) and (S2) hold; then, for $\mu<0$, the sequence

$$\{ \mathbb{K}^\mu_\ell = \mathbb{P}^\mu_\ell \circ X_\ell^{-1} : \ell = 1,2,\dots \}$$

satisfies the Large Deviation Principle with constants $\{ V_\ell : \ell = 1,2,\dots\}$ and rate function $I^\mu(\cdot)$ given by

$$I^\mu(x) = \begin{cases} p(\mu) + f(x) - \mu x, & x \geq 0, \\ \infty & , x<0. \end{cases} \qquad (4.2)$$

Proof:

It was proved in §9 of Lecture 1 that (LD1), (LD2) hold and in §6 that (LD3) holds; it remains to prove that, for each open subset G of $[0,\infty)$:

$$\liminf_{\ell\to\infty} \frac{1}{\beta V_\ell} \ln \mathbb{K}^\mu_\ell[G] \geq -\inf_G I^\mu(x). \qquad (4.3)$$

Let y be an arbitrary point of G; choose $\delta > 0$ so that $B^\delta_y = (y-\delta, y+\delta) \subset G$ and t_ℓ such that $p'_\ell(\mu+t_\ell)=y$. Then, as in §8, we have

$$\mathbb{K}^\mu_\ell[G] \geq e^{\beta V_\ell\{p_\ell(\mu+t_\ell) - p_\ell(\mu) - t_\ell y - \delta|t_\ell|\}} \mathbb{K}^{\mu+t_\ell}_\ell[B^\delta_y].$$

By Proposition 2, $\mu+t_\ell \to \mu(y)$ and $p_\ell(\mu+t_\ell) \to p(\mu(y))$ so that

$$\liminf_{\ell\to\infty} \frac{1}{\beta V_\ell} \ln \mathbb{K}^\mu_\ell[G] \geq p(\mu(y)) - p(\mu) - (\mu(y)-\mu)y$$
$$- \delta|\mu(y)-\mu| + \liminf_{\ell\to\infty} \frac{1}{\beta V_\ell} \ln \mathbb{K}^{\mu+t_\ell}_\ell[B^\delta_y]$$

We now have to distinguish two cases: if $y < \rho_c$ then $\mu(y) < 0$ and we make use of the fact that $\mu \mapsto p(\mu)$ is differentiable for $\mu < 0$; then, by Theorem 1, for all ℓ sufficiently large $\mathbb{K}_\ell^{\mu+t_\ell}[B_y^\delta] > \frac{1}{2}$ and hence

$$\liminf_{\ell \to \infty} \frac{1}{\beta V_\ell} \ln \mathbb{K}_\ell^{\mu+t_\ell}[B_y^\delta] = 0 \; ; \tag{4.4}$$

on the other hand, if $y \geq \rho_c$ we have $\mu(y) = 0$ and we must proceed differently.

Lemma2 Let N_1 and N_2 be independent non-negative integer-valued random variables with means m_1 and m_2 respectively. Suppose that N_1 is geometrically distributed and that $\delta \geq 1$; then

$$\mathbb{P}[N_1 + N_2 \in B^\delta_{m_1+m_2}] \geq \frac{1}{m_1+m_2}\left(\frac{m_1}{m_1+1}\right)^{m_1+m_2+1} . \tag{4.5}$$

Proof:

The interval $B^1_{m_1+m_2} = (m_1+m_2-1, m_1+m_2+1)$ contains a unique integer $n_0 \geq m_1+m_2$. Now

$$\mathbb{P}[N_1 + N_2 \in B^\delta_{m_1+m_2}] = \sum_{m \in B^\delta_{m_1+m_2}} \sum_{n=0}^{m} \mathbb{P}[N_1 = m-n]\,\mathbb{P}[N_2 = n]$$

$$\geq \sum_{n=0}^{n_0} \mathbb{P}[N_1 = n_0 - n]\,\mathbb{P}[N_2 = n] . \tag{4.6}$$

Since N_1 is geometrically distributed, $n \mapsto \mathbb{P}[N_1 = n]$ is a decreasing function so that

$$\sum_{n=0}^{n_0} \mathbb{P}[N_1 = n_0 - n]\,\mathbb{P}[N_2 = n] \geq \mathbb{P}[N_1 = n_0]\,\mathbb{P}[n_2 = n] . \tag{4.7}$$

Now

$$\mathbb{P}[N_1 = n_0] = \frac{1}{m_1+1}\left(\frac{m_1}{m_1+1}\right)^{n_0}$$

$$\geq \frac{1}{m_1+1}\left(\frac{m_1}{m_1+1}\right)^{m_1+m_2+1} \tag{4.8}$$

and

$$\mathbb{P}[N_2 \leq n_0] \geq \mathbb{P}[N_2 \leq m_1+m_2] \geq \frac{m_1}{m_1+m_2}, \tag{4.9}$$

by Markov's Inequality. Hence

$$\mathbb{P}[N_1 + N_2 \in B^{\delta}_{m_1+m_2}] \geq \frac{1}{m_1+m_2}\left(\frac{m_1}{m_1+1}\right)^{m_1+m_2+2} \tag{4.10}$$

■

Returning to the proof of Theorem 1., it follows from Lemma 1 that σ_1 is geometrically distributed; applying the Lemma with $N_1 = \sigma_1$ and $N_2 = N - \sigma_1$, we have $\frac{m_1}{m_1+1} = e^{\beta(\mu+t_\ell)}$ and $m_1 + m_2 = V_\ell y$; thus

$$\mathbb{K}^{\mu+t_\ell}_{\ell}[B^{\delta}_{y}] \geq \frac{1}{V_\ell y}\, e^{\beta(\mu+t_\ell)(V_\ell y+2)}, \tag{4.11}$$

for $V_\ell \geq \delta^{-1}$. It follows that

$$\liminf_{\ell\to\infty} \frac{1}{\beta V_\ell} \ln \mathbb{K}^{\mu+t_\ell}_{\ell}[B^{\delta}_{y}] = 0 \tag{4.12}$$

since, for $y \geq \rho_c$, $\mu + t_\ell \to 0$. Thus we have, in both cases,

$$\liminf_{\ell\to\infty} \frac{1}{\beta V_\ell} \ln \mathbb{K}^{\mu}_{\ell}[G] \geq -p(\mu) - f(y) + \mu y$$

$$= - I^{\mu}(y) \tag{4.13}$$

for all y in G, since δ was arbitrary. Hence

$$\liminf_{l\to\infty} \frac{1}{\beta V_l} \ln \mathbb{K}_l^{\mu}[G] \geqslant \sup_{G}(-I^{\mu}(y))$$

$$= -\inf_{G} I^{\mu}(y) \quad \blacksquare \qquad (4.14)$$

§5. A Large Deviation Result for the Occupation Measure

We introduce a measure-valued random variable

$$L_l(\omega; B) = \frac{1}{V_l} \sum_{j\geq 1} \sigma_j(\omega)\, \delta_{\lambda_l(j)}[B]$$

where $\delta_\lambda[B] = 1$ if λ is in B and is zero otherwise. Then L_l maps Ω into the space $E = M_b^+(\mathbb{R}^+)$ of positive bounded measures on the positive real line. Let $\mathbb{K}_l^\mu = \mathbb{P}_l^\mu \circ L_l^{-1}$ be the induced probability measure on E; in terms of this we can express the expectation of a functional of L_l as an integral over E. For example,

$$\mathbb{E}_l^\mu[e^{-\beta a N^2/2V_l}] = \int_E e^{\beta V_l G(m)} \mathbb{K}_l^\mu[dm]$$

where $G(m) = -\frac{a}{2}\|m\|^2$ and $\|m\| = \int_{[0,\infty)} m(d\lambda)$. We state the large deviation result without proof:

Theorem 2 Suppose that (S1) and (S2) hold and that the density of states is strictly positive on $(0,\infty)$; then, for $\mu < 0$, the sequence $\{\mathbb{K}_l^\mu = \mathbb{P}_l^\mu \circ L_l^{-1}\}$ of probability measures on $M_b^+(\mathbb{R}^+)$ satisfies the Large Deviation Principle with constants $\{V_l\}$ and rate-function $I^\mu[\cdot]$, given by

$$I^\mu[m] = \sup_{t\in \mathscr{C}^\mu} \{ \langle m,t\rangle - C^\mu[t]\} \qquad (5.1)$$

where

$$C^\mu[t] = \int_{[0,\infty)} \{p(\mu + t(\lambda)|\lambda) - p(\mu|\lambda)\}\, dF(\lambda) \qquad (5.2)$$

and

$$\mathscr{C}^\mu = \{t \in C(\mathbb{R}^+) : \sup_{[0,\infty)} \{t(\lambda) - \lambda\} < -\mu\}. \qquad (5.3)$$

§6 An Application

In this section, we sketch an application to the statistical mechanics of a model of the interacting boson gas. We consider the diagonal model [4] for which the hamiltonian is

$$H_\ell^D(\omega) = H_\ell(\omega) + \frac{a}{2V_\ell}\{2N(\omega)^2 - \sum_{j=1}^{\infty}\sigma_j(\omega)^2\}$$

$$+ \frac{1}{2V_\ell}\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} u(\lambda_\ell(i),\lambda_\ell(j))\,\sigma_i(\omega)\sigma_j(\omega)\,. \qquad (6.1)$$

The last two terms in this hamiltonian have different asymptotic behaviour for large ℓ. The hamiltonian

$$H_\ell^{HYL}(\omega) = H_\ell(\omega) + \frac{a}{2V_\ell}\{2N(\omega)^2 - \sum_{j=1}^{\infty}\sigma_j(\omega)^2\}$$

was introduced by Huang, Yang and Luttinger [3]; the model has been investigated, using large deviation methods, by van den Berg, Lewis and Pulè [2] Here we report a result on the perturbed mean field model obtained in [5]; the hamiltonian is given by

$$H_\ell^{PMF}(\omega) = H_\ell(\omega) + \frac{1}{2V_\ell}\sum_{=1}^{\infty}\sum_{j=1}^{\infty} v(\lambda_\ell(i),\lambda_\ell(j))\,\sigma_i(\omega)\sigma_j(\omega). \qquad (6.2)$$

If we assume that $v : \mathbb{R}^2 \to \mathbb{R}$ is continuous, bounded and positive then we can use Theorem 2 of §5 to obtain the following result (see [5]):

Theorem Suppose that (S1) and (S2) hold and that the density of states is strictly positive on $(0,\infty)$; then the pressure $p^{PMF}(\mu)$ corresponding to the sequence of hamiltonians $\{H_\ell^{PMF}\}$ is given by

$$p^{PMF}(\mu) = \sup_{M_b^+(R^+)} \{\mu\|m\| - f^{PMF}[m]\}\,, \qquad (6.3)$$

where

$$f^{PMF}[m] = \int_{[0,\infty)} \lambda \, m(d\lambda) + \frac{1}{2}\int_{[0,\infty)}\int_{[0,\infty)} v(\lambda,\lambda')\, m(d\lambda)\, m(d\lambda') - \beta^{-1}\int_{[0,\infty)} (s \circ p)(\lambda)\, dF(\lambda) \tag{6.4}$$

and

$$s(x) = (1+x)\ln(1+x) - x\ln x \; ; \tag{6.5}$$

here

$$m(d\lambda) = m_s(d\lambda) + p(\lambda)\,dF(\lambda) \tag{6.6}$$

is the Lebesgue decomposition of m with respect to $dF(\lambda)$.

REFERENCES

[1] M. van den Berg, J.T. Lewis and J.V. Pulè: A General Theory of Bose-Einstein Condensation, Helv. Phys. Acta, 59, 1271-1288 (1986).

[2] M. van den Berg, J.T. Lewis, J.V. Pulè: The Large Deviation Principle and Some Models of an Interacting Boson Gas, to appear in Commun. Math. Phys. .

[3] K. Huang, C.N. Yang, J.M. Luttinger: Imperfect Bose gas with Hard-sphere Interactions, Phys. Rev., 105, 776-784 (1957).

[4] D.J. Thouless: The Quantum Mechanics of Many-Body Systems, New York: Academic Press 1966.

[5] M. van den Berg, T.C. Dorlas, J.T. Lewis, J.V. Pulè: (in preparation).

Lecture III: The Large Deviation Principle for the Kac Distribution

§1 Introduction

In 1971, Kac discovered that, for the free boson gas, the canonical and grand canonical ensembles are not strictly equivalent although they give rise to the same equation of state. This manifests itself in the fact that, above the critical density, the grand canonical distribution of the particle number density is not asymptotically degenerate; in the standard example (where the single-particle hamiltonian is a constant multiple of the Laplacian with Dirichlet boundary conditions in a star-shaped region and the thermodynamic limit is taken by dilating the region about an interior point, holding the mean number density fixed) the distribution is exponential; in general, when the distribution converges, it converges to an infinitely-divisible distribution. For a full discussion of these aspects of the free boson gas, see [1] and [2]; using their terminology, we shall refer to the grand canonical distribution of the number density as the Kac distribution.

Kac conjectured that the lack of equivalence of ensembles in the strict sense was a pathology of the free gas which would disappear in the presence of a repulsive interaction, however weak. To test this idea, Davies [3] studied in great detail a mean-field model of an interacting boson gas. He proved that, if the mean-field potential is strictly convex, the Kac distribution is asymptotically degenerate. In [4], a general result was proved, from which it was deduced that the Kac density is asymptotically degenerate whenever the free-energy exists and is strictly convex. In §2 of this lecture, we adapt the arguments of [4] to prove that the Kac distribution satisfies Varadhan's Large Deviation Principle whenever the free-energy density exists in the thermodynamic limit; in §3, we provide an alternative proof to that given in Lecture II of the result that the Kac distribution for the free boson gas satisfies Varadhan's Large Deviation Principle. The result proved in §2 applies also to mean-field models, even when the mean-field potential is non-convex; in such cases, it is possible for the free-energy density in the thermodynamic to be non-convex; nevertheless the Kac density satisfies the Large Deviation Principle, albeit with a non-convex rate function (examples of such rate-functions were given by Ellis [6], see also [7].)

To illustrate the situation which can arise with a non-convex rate-function, we investigate in §3 of this paper the model discussed by Davies in [3]. In §4, we describe in detail possible asymptotic distributions for the Kac distribution in this example.

The results described in this lecture are based on joint work with Zagrebnov and Pulè [11].

§2. A Large Deviation Result

Varadhan's Theorem [5] concerns the asymptotic behaviour of integrals with respect to a sequence of probability measures satisfying the Large Deviation Principle and extends Laplace's Method to infinite dimensional spaces. Even in the case of a one-dimensional space, it has advantages over Laplace's Method: it applies to a wider class of measures and to a wider class of integrands.

Let E be a complete separable metric space and $\{\mathbb{K}_\ell : \ell = 1, 2, \dots\}$ a sequence of probability measures on the Borel subsets of E: let $\{V_\ell : \ell = 1, 2, \dots\}$ be a sequence of positive constants such that $V_\ell \to \infty$. We say that $\{\mathbb{K}_\ell\}$ obeys the Large Deviation Principle with constants $\{V_\ell\}$ and rate-function $I(\cdot)$ if there exists a function $I : E \to [0, \infty]$ satisfying:

(LD1): $I(\cdot)$ is lower semi-continuous on E.

(LD2): For each $m < \infty$, the set $\{x : I(x) \leq m\}$ is compact.

(LD3): For each closed subset C of E,

$$\limsup_{\ell \uparrow \infty} \frac{1}{V_\ell} \ln \mathbb{K}_\ell[C] \leq -\inf_C I(x).$$

(LD4): For each open subset G of E,

$$\liminf_{\ell \uparrow \infty} \frac{1}{V_\ell} \ln \mathbb{K}_\ell[G] \geq -\inf_G I(x).$$

A version of Varadhan's Theorem adequate for many applications in statistical mechanics is the following:

Varadhan's Theorem

Let $\{\mathbb{K}_\ell\}$ be a sequence of probability measures on E satisfying the Large Deviation Principle with constants $\{V_\ell\}$ and rate-function $I(\cdot)$. Let $G : E \to R$ be a continuous function which is bounded above on the set $\bigcup_{\ell \geq 1} \operatorname{supp} \mathbb{K}_\ell$. Then

$$\lim_{\ell \to \infty} \frac{1}{V_\ell} \ln \int_E e^{V_\ell G(x)} \mathbb{K}_\ell[dx] = \sup_E \{G(x) - I(x)\}.$$

In [4], Pulè and I proved a large deviation result whose main hypothesis was the existence of the free-energy in the thermodynamic limit; at the time we were not aware of Varadhan's work and so our result was not formulated within that frame-work. Here we reorganize the proof given in [4] to establish a result within the Varadhan scheme; this enables us to give a simpler proof of the free-boson gas result proved in Lecture II.

Let $\{f_\ell : \ell = 1, 2, \cdots\}$ be a sequence of functions $f_\ell : [0,\infty) \to \mathbb{R}$ satisfying $f_\ell(0) = 0$; the grand canonical pressure $p_\ell(\mu)$ determined by f_ℓ is defined by

$$p_\ell(\mu) = \frac{1}{\beta V_\ell} \ln \int_{[0,\infty)} e^{\beta V_\ell \{\mu x - f_\ell(x)\}} m_\ell[dx]$$

where, for each Borel subset A of $[0,\infty)$

$$m_\ell[A] = \sum_{n \geq 0} \delta_{\frac{n}{V_\ell}}[A]$$

and

$$\delta_x[A] = \begin{cases} 1 & , \quad x \in A , \\ 0 & , \quad x \in A^c , \end{cases}$$

and $\{V_\ell\}$ is a sequence of positive constants, $V_\ell \to \infty$.
For each μ for which $p_\ell(\mu)$ is finite, the Kac distribution $\mathbb{K}_\ell^\mu$ determined by f_ℓ is defined on the Borel subsets of $[0,\infty)$ by

$$\mathbb{K}_\ell^\mu[A] = e^{-\beta V_\ell p_\ell(\mu)} \int_A e^{\beta V_\ell \{\mu x - f_\ell(x)\}} m_\ell[dx] .$$

We prove the following theorem:

Theorem 1

Suppose that, on each compact, the sequence $\{f_\ell : \ell = 1, 2, \ldots\}$ is bounded below and converges uniformly to a lower semi-continuous function f. Let μ_∞ be defined by

$$\mu_\infty = \lim_{\ell \uparrow \infty} \left(\liminf_{x \uparrow \infty} \left(\frac{1}{x} \inf_{k \geq \ell} f_k(x) \right) \right) .$$

Then for each $\mu < \mu_\infty$, the grand canonical pressure $p(\mu) = \lim_{\ell \to \infty} p_\ell(\mu)$ exists and is given by the Legendre-Fenchel transform of f:

$$p(\mu) = f^*(\mu) \stackrel{def}{=} \sup_{x \geq 0} \{ \mu x - f(x) \} .$$

Moreover, the sequence $\{\mathbb{K}^{\mu}_{\ell} : \ell = 1, 2, \ldots\}$ satisfies the Large Deviation Principle with constants $\{V_\ell\}$ and rate-function $I^{\mu}(\cdot)$ given by

$$I^{\mu}(x) = p(\mu) + f(x) - \mu x .$$

Proof:

Put $g_\ell(x) = \mu x - f_\ell(x)$ and $g(x) = \mu x - f(x)$ so that we can write

$$p_\ell(\mu) = \frac{1}{\beta V_\ell} \ln \int_{[0,\infty)} e^{\beta V_\ell g_\ell(x)} m_\ell(dx) .$$

Choose A such that $\mu < A < \mu_\infty$; choose m such that

$$\liminf_{x \uparrow \infty} \left(\frac{1}{x} \inf_{k \geq m} f_k(x) \right) > A ;$$

then there exists x_1 such that $f_\ell(x) > Ax$ for all $x > x_1$ and all $\ell > m$; hence $g_\ell(x) < -(A-\mu)x$ for all $x > x_1$ and $\ell > m$, and $g(x) \leq -(A-\mu)x$ for all $x > x_1$. But g(0) = 0 so that

$$\sup_{[0,\infty)} g(x) = \sup_{[0,x_1]} g(x)$$

Now g is upper semi-continuous and bounded above on compacts, so that the supremum of g on $[0, x_1]$ is attained at some point x_0 in $[0, x_1]$; hence

$$f^*(\mu) \stackrel{def}{=} \sup_{[0,\infty)} g(x) = g(x_0) < \infty .$$

Furthermore, given $\epsilon > 0$, there exists $\delta > 0$ such that $g(x_0) - g(x) < \frac{\epsilon}{2}$ for x in $[x_0 - \delta , x_0 + \delta]$; by the uniformity of convergence on compacts which was postulated for $\{f_\ell\}$, there exists m' such that, for all $\ell > m'$, $g(x) - g_\ell(x) < \frac{\epsilon}{2}$ for all x in $[x_0 - \delta , x_0 + \delta]$. Thus we have, for all ℓ sufficiently large,

$$\int_{[0,\infty)} e^{\beta V_\ell g_\ell(x)} m_\ell[dx] > \int_{[x_0-\delta, x_0+\delta]} e^{\beta V_\ell g_\ell(x)} m_\ell[dx] > e^{\beta V_\ell (g(x_0) - \epsilon)},$$

since eventually $[x_0 - \delta, x_0 + \delta]$ contains at least one point of $\{\frac{n}{V_\ell} : n = 0, 1, 2, \ldots\}$. Since $\epsilon > 0$ was arbitrary, we have

$$\liminf_{\ell \to \infty} \frac{1}{\beta V_\ell} \ln \int_{[0,\infty)} e^{\beta V_\ell g_\ell(x)} m_\ell[dx] \geq g(x_0).$$

On the other hand ,

$$\int_{[0,\infty)} e^{\beta V_\ell g_\ell(x)} m_\ell[dx] < e^{\beta V_\ell \{g(x_0) + \epsilon\}} \int_{[0,x_1]} m_\ell[dx] + \int_{(x_1,\infty)} e^{-\beta V_\ell (A-\mu)x} m_\ell[dx]$$

$$\leq e^{\beta V_\ell \{g(x_0) + \epsilon\}} \left\{ \frac{e^{-\beta V_\ell \{(A-\mu)x_1 + g(x_0) + \epsilon\}}}{1 - e^{-\beta(A-\mu)}} + (V_\ell x_1 + 1) \right\}$$

for all ℓ sufficiently large; hence, since $g(x_0) \geq 0$ and $(A - \mu) x_1 > 0$, we have

$$\limsup_{\ell \to \infty} \frac{1}{\beta V_\ell} \ln \int_{[0,\infty)} e^{\beta V_\ell g_\ell(x)} m_\ell[dx] \leq g(x_0) + \epsilon$$

and the statement concerning the pressure is proved, since $\epsilon > 0$ was arbitrary.

We turn to the proof of the assertion concerning the sequence $\{\mathbb{K}_\ell^\mu : \ell = 1, 2, \ldots\}$.

(LD1) holds by the hypothesis that $x \mapsto f(x)$ is lower semi-continuous. It follows that, for each $m < \infty$, the set $L_m = \{x : I^\mu(x) \leq m\}$ is closed; for $x \in L_m$, we have

$$f(x) \leq m - p(\mu) + \mu x ;$$

on the other hand, we have shown that, for $x > x_1$, $f(x) \geq Ax$. Hence, either $x \leq x_1$ or $x \leq \frac{m - p(\mu)}{A - \mu}$ so that L_m is bounded and (LD2) holds.

For a closed set C, we have, given $\epsilon > 0$,

$$\mathbb{K}_\ell^\mu [C] \leq e^{-\beta V_\ell p_\ell(\mu)} e^{\beta V_\ell \{\sup_{C \cap [0, x_2]} g(x) + \epsilon\}} \int_{C \cap [0, x_2]} m_\ell [dx] + \int_{C \cap [x_2, \infty)} e^{-\beta V_\ell (A - \mu) x} m_\ell [dx]$$

for all ℓ sufficiently large and all x_2 in $(0, \infty)$.

Since $\sup_{C \cap [0, x_1]} g(x) \leq \sup_C g(x)$ we have

$$\mathbb{K}_\ell^\mu [C] \leq e^{\beta V_\ell \{p_\ell(\mu) + \sup_C g(x) + \epsilon\}} \left\{ (V_\ell x_2 + 1) + \frac{e^{-\beta V_\ell \{\sup_C g(x) + \epsilon + (A - \mu) x_2\}}}{1 - e^{-\beta (A - \mu)}} \right\}$$

for all ℓ sufficiently large. Now choose $x_2 \geq x_1$ such that

$$\sup_C g(x) + (A - \mu) x_2 \geq 0 .$$

Hence

$$\limsup_{\ell \uparrow \infty} \frac{1}{\beta V_\ell} \ln \mathbb{K}_\ell^\mu[C] \leqslant -p(\mu) + \sup_C g(x) = -\inf_C I^\mu(x)$$

and (LD3) holds.

Let G be an arbitrary open subset of $[0,\infty)$; given $\epsilon > 0$ and y in G, choose δ such that $(y-\delta, y+\delta) = B_y^\delta \subset G$ and $g(y) - g(x) < \frac{\epsilon}{2}$ for all x in B_y^δ (this is possible since g is upper semi-continuous). Thus, for ℓ sufficiently large,

$$\mathbb{K}_\ell^\mu[G] \geqslant \mathbb{K}_\ell^\mu[B_y^\delta] = e^{-\beta V_\ell p_\ell(\mu)} \int_{B_y^\delta} e^{\beta V_\ell g_\ell(x)} m_\ell[dx]$$

$$\geqslant e^{-\beta V_\ell p_\ell(\mu)} e^{\beta V_\ell (g(y)-\epsilon)} m_\ell[B_y^\delta] .$$

But eventually B_y^δ contains at least one point of $\{\frac{n}{V_\ell} : n = 1, 2, \dots\}$ so that $m_\ell[B_y^\delta] \geqslant 1$ and hence

$$\liminf_{\ell \uparrow \infty} \frac{1}{\beta V_\ell} \ln \mathbb{K}_\ell^\mu[G] \geqslant -p(\mu) + g(y) = -I^\mu(y) ;$$

since this inequality holds for each point of G, we have

$$\liminf_{\ell \uparrow \infty} \frac{1}{\beta V_\ell} \ln \mathbb{K}_\ell^\mu[G] \geqslant \sup_G \{-I^\mu(y)\} = -\inf_G I^\mu(y)$$

and (LD4) holds.

Let $N(\omega) = \sum_{j \geqslant 1} \sigma_j(\omega)$ denote the total number of particles in the configuration ω ; then the canonical partition function $Z_\ell(n)$ is defined by

$$Z_\ell(n) = \sum_{\{\omega \in \Omega : N(\omega) = n\}} e^{-\beta H_\ell(\omega)} , \quad n \geqslant 1 . \tag{3}$$

It is convenient to put $Z_\ell(0) = 1$. The free-energy density $f_\ell : [0,\infty) \to \mathbb{R}$ is defined first on the

set $\{ \frac{n}{V_\ell} : n = 0, 1, 2, \ldots \}$ by $f_\ell(\frac{n}{V_\ell}) = -\frac{1}{\beta V_\ell} \ln Z_\ell(n)$,

then extended to the whole of $[0,\infty)$ by linear interpolation. Using the methods of [8] and the results of [1] one may prove in the case of the free boson gas, (see the Appendix):

Theorem 2

Suppose that (S 1) and (S 2) hold; then on each compact subset of $[0,\infty)$ the sequence $\{f_\ell\}$ is bounded and converges uniformly on compacts to a convex function f satisfying $f(0) = 0$; moreover, $\mu_\infty = 0$.

Putting together this result with Theorem 1 of the previous section, we have

Corollary

Suppose that (S 1) and (S 2) hold; then, for each $\mu < 0$, the sequence $\{\mathbb{K}_\ell^\mu : \ell = 1, 2, \ldots\}$ satisfies the Large Deviation Principle with constants $\{V_\ell\}$ and rate-function $I^\mu(\cdot)$ given by

$$I^\mu(x) = f^*(\mu) + f(x) - \mu x , \tag{3.4}$$

and $$p(\mu) = f^*(\mu) .$$

Remark

It may seem surprising that no use appears to have been made of the special features of the free boson gas, while the earlier proof of this result made fairly delicate use of the fact that the occupation numbers $\{\sigma_j : j = 1, 2, \ldots\}$, in the grand canonical ensemble, are independent geometrically distributed random variables. It is worthwile, perhaps, to examine this point further. The grand canonical measure $\mathbb{P}_\ell^\mu$ is defined by

$$\mathbb{P}_\ell^\mu[\omega] = e^{-\beta V_\ell p_\ell(\mu)} e^{\beta\{\mu N(\omega) - H_\ell(\omega)\}}$$

for each $\mu < 0$; an easy calculation, see [1] for instance, yields

$$\mathbb{P}_\ell^\mu[\sigma_j \geqslant m] = e^{m\beta(\mu - \lambda_\ell(j))} . \tag{3.5}$$

By expressing N as $\sigma_1 + (N-\sigma_1)$ and using (3.5), a lower bound was obtained for $\mathbb{P}_\ell^\mu[X_\ell \in B_y^\delta]$, where $X_\ell = N/V_\ell$ and $y = \mathbb{E}_\ell^\mu[X_\ell]$; this was required for the proof in Lecture II that (LD4) holds. However, the use of (3.5) can be detected in the proof of Theorem 2: the convexity of the functions $x \mapsto f_\ell(x)$ was used to prove the uniform convergence of the sequence $\{f_\ell\}$ on compacts; the proposition that f_ℓ is convex is equivalent to the proposition that the inequality $Z_\ell(n)^2 \geq Z_\ell(n+1)Z_\ell(n-1)$ holds for each $n \geq 1$, but this is equivalent to the proposition that $n \mapsto \mathbb{P}_\ell^\mu[\sigma_j \geq m \mid N = n]$ is an increasing function in view of the result, proved in [8], that

$$\mathbb{P}_\ell^\mu[\sigma_j \geq m \mid N = n] = \begin{cases} e^{-m\beta\lambda_\ell(j)} \dfrac{Z_\ell(n-m)}{Z_\ell(n)} , & m \leq n , \\ 0 , & m > n . \end{cases}$$

Following Davies [3], we consider the hamiltonian

$$\tilde{H}_\ell = H_\ell + V_\ell w(X_\ell) , \tag{3.6}$$

where H_ℓ is the free-gas hamiltonian of (3.1) and $X_\ell = N/V_\ell$ is the particle number density; unlike Davies, we require only that $w : [0,\infty) \to \mathbb{R}$ be lower semi-continuous and satisfy

$$w(0) = 0 , \qquad \liminf_{x\to\infty} \frac{w(x)}{x} = +\infty . \tag{3.7}$$

Define $\tilde{f}_\ell$ by $\tilde{f}_\ell(x) = f_\ell(x) + w(x)$; it then follows that

$$\tilde{f}_\ell\left(\frac{n}{V_\ell}\right) = -\frac{1}{\beta V_\ell} \ln \sum_{\{\omega : N(\omega) = n\}} e^{-\beta \tilde{H}_\ell(\omega)} . \tag{3.8}$$

Using Theorem 2 and (3.7), we verify that $\{\tilde{f}_\ell : \ell = 1, 2, \cdots\}$ satisfies the conditions of Theorem 1 with $\nu_\infty = \infty$; we conclude that the following result follows from Theorem 1:

Theorem 3

Suppose that (S 1) and (S 2) hold and that the mean-field hamiltonian (3.6) satisfies (3.7); then for each $\nu < \infty$ the grand canonical pressure $\tilde{p}(\nu) = \lim_{\ell \to \infty} \tilde{p}_\ell(\nu)$ exists and is given by the Legendre-Fenchel transform $\tilde{f}^*$ of $\tilde{f}$, where

$$\tilde{f}(x) = f(x) + w(x) \qquad (3.9)$$

and f is the free-energy density of the free boson gas, and the sequence $\{\tilde{\mathbb{K}}^\nu_\ell : \ell = 1, 2, \cdots\}$ of Kac distributions determined by $\{\tilde{f}_\ell : \ell = 1, 2, \cdots\}$ satisfies the Large Deviation Principle with constants $\{V_\ell\}$ and rate-function $\tilde{I}(\nu)$ given by

$$\tilde{I}^\nu(x) = \tilde{p}(\nu) + \tilde{f}(x) - \nu x . \qquad (3.10)$$

It follows that $\tilde{p}^* = \tilde{f}^{**} = \text{conv}\ \tilde{f}$, where conv g denotes the convex envelope of g; hence the intervals $[\tilde{p}'_-(\nu), \tilde{p}'_+(\nu)]$ of discontinuity of the derivative of $\tilde{p}$ correspond to the linear segments in the convex envelope of $\tilde{f}$. We conclude once more that $\{\tilde{\mathbb{K}}^\nu_\ell\}$ is asymptotically degenerate whenever w is strictly convex.

§4 The Asymptotics of the Kac Distribution

In this section we examine the consequences of the non-convexity of $\tilde{f}$ for the asymptotics of the Kac distribution. We consider the case where conv $\tilde{f}$ has precisely one linear segment $[\rho_-, \rho_+]$ and $\tilde{f}(x) > \text{conv}\,\tilde{f}(x)$ for x in (ρ_-, ρ_+) the general situation should be clear from this discussion. We recall that the asymptotic Kac distribution $\mathbb{K}^\nu = \lim_{\ell \to \infty} \mathbb{K}^\nu_\ell$ gives the decomposition of the grand canonical limiting state $\langle \cdot \rangle^\nu$ into extremal (canonical) limiting states $\langle \cdot \rangle_\rho$:

$$\langle \cdot \rangle^\nu = \int_{[0,\infty)} \langle \cdot \rangle_\rho \, \mathbb{K}^\nu[d\rho] \quad . \qquad (4.1)$$

In general, if $\mathbb{K}^{\mu} = \lim_{\ell\to\infty} \mathbb{K}^{\mu}_{\ell}$ exists its support is contained in the set $\{x : I^{\mu}(x) = 0\}$; however, if this set consists of more than one point there is no guarantee that the sequence $\{\mathbb{K}^{\mu}_{\ell} : \ell = 1, 2, \ldots\}$ converges. Nevertheless, by the Helly selection principle, $\{\mathbb{K}^{\mu}_{\ell} : \ell = 1, 2, \ldots\}$ contains at least one convergent subsequence; in the case under consideration, where conv $\tilde{f}$ has precisely one linear segment $[\rho_-, \rho_+]$ and $\tilde{f}$ is non-convex, we have three cases determined by μ_c which is defined by $\tilde{p}'_-(\mu_c) = \rho_-$ (and hence $\tilde{p}'_+(\mu_c) = \rho_+$) so that μ_c is the slope of the linear segment of conv $\tilde{f}$:

$$\mathrm{I}: \quad \mu < \mu_c \; ; \quad \tilde{\mathbb{K}}^{\mu}_{\ell} \to \tilde{\mathbb{K}}^{\mu} = \delta_{\rho(\mu)} \quad , \quad \rho(\mu) = \tilde{p}'(\mu) \; . \tag{4.2}$$

II : $\mu = \mu_c$; there exists $\{\ell_j : j = 1, 2, \ldots\}$ such that $\lim_{j\to\infty} \tilde{\mathbb{K}}^{\mu}_{\ell_j} = \tilde{\mathbb{K}}^{\mu}$ exists and

$$\tilde{\mathbb{K}}^{\mu} = \alpha\, \delta_{\rho_-} + (1-\alpha)\, \delta_{\rho_+} \quad , \quad 0 \leq \alpha \leq 1 . \tag{4.3}$$

$$\mathrm{III}: \quad \mu > \mu_c \; ; \quad \tilde{\mathbb{K}}^{\mu}_{\ell} \to \tilde{\mathbb{K}}^{\mu} = \delta_{\rho(\mu)} \quad , \quad \rho(\mu) = \tilde{p}'(\mu) \; . \tag{4.4}$$

We sketch the proof of II; the proof of the remaining cases should then be clear. Choose ρ_0 such that $\rho_- < \rho_0 < \rho_+$;

let $A_- = [0, \rho_0)$ and $A_+ = [\rho_0, \infty)$. Then $\{\tilde{\mathbb{K}}^{\mu_c}_{\ell}[A_-]\}$ is a bounded sequence of real numbers and hence contains a convergent subsequence

$$\{\tilde{\mathbb{K}}^{\mu_c}_{\ell_k}[A_-] : k = 1, 2, \ldots\} \; .$$

Consider the case in which

$$\lim_{k\to\infty} \tilde{\mathbb{K}}^{\mu_c}_{\ell_k}[A_-] = \alpha \; , \quad 0 < \alpha < 1 \; .$$

Then

$$\int_{[0,\infty)} e^{-tx}\, \tilde{\mathbb{K}}^{\mu_c}_{\ell_k}[dx] = \tilde{\mathbb{K}}^{\mu_c}_{\ell_k}[A_-] \int_{A_-} e^{-tx}\, L^-_k[dx]$$

$$+ (1 - \tilde{\mathbb{K}}^{\mu_c}_{\ell_k}[A_-]) \int_{A_+} e^{-tx}\, L^+_k[dx] \quad ,$$

where

$$L_k^-[A] = \frac{\tilde{\mathbb{K}}_{\ell_k}^{\mu_c}[A \cap A_-]}{\tilde{\mathbb{K}}_{\ell_k}^{\mu_c}[A_-]} \quad , \quad L_k^+[A] = \frac{\tilde{\mathbb{K}}_{\ell_k}^{\mu_c}[A \cap A_+]}{\tilde{\mathbb{K}}_{\ell_k}^{\mu_c}[A_+]}$$

Now $\{L_k^-\}$ and $\{L_k^+\}$ satisfy the large deviation principle with rate-functions I^- and I^+ respectively, where $\tilde{I}^-$ ($\tilde{I}^+$) is the restriction of $\tilde{I}$ to $A^-(A^+)$. Now $\tilde{I}^-$ has a unique minimum at ρ_- and $\tilde{I}^+$ has a unique minimum at ρ_+, hence

$$\int_{[0,\infty)} e^{-tx}\, \tilde{\mathbb{K}}_{\ell_k}^{\mu_c}[dx] \to \alpha\, e^{-t\rho_-} + (1-\alpha)\, e^{-t\rho_+} ,$$

so that $\{\tilde{\mathbb{K}}_{\ell_k}^{\mu_c}\}$ converges weakly to $\alpha\, \delta_{\rho_-} + (1-\alpha)\, \delta_{\rho_+}$.

The remaining cases are clear.
It remains to consider case II in more detail: we investigate the possible dependence of α on the subsequence $\{\ell_j : j = 1,2, \dots\}$. We remark, in passing, that if we adopt the quasi-average approach of Bogoliubov [9], we get

$$\lim_{\mu \uparrow \mu_c} \lim_{\ell \to \infty} \tilde{\mathbb{K}}_\ell^\mu = \delta_{\rho_-} \quad , \quad \lim_{\mu \downarrow \mu_c} \lim_{\ell \to \infty} \tilde{\mathbb{K}}_\ell^\mu = \delta_{\rho_+} . \qquad (4.5)$$

On the other hand, the generalized quasi-average procedure [10] enables us to scan the whole interval $0 \leq \alpha \leq 1$: here we put $\mu_\ell = \mu_c + \frac{\delta}{\beta V_\ell^\gamma}$, $\gamma \geq 1$, and get the following limiting values:

$$\tilde{\mathbb{K}}^{\mu_c}_{\gamma,\delta} = \lim_{k\to\infty} \tilde{\mathbb{K}}^{\mu_{\ell_k}}_{\ell_k} .$$

$$\gamma = 1 \quad : \quad \tilde{\mathbb{K}}^{\mu_c}_{\gamma,\delta} = \lambda_\delta \, \delta_{\rho_-} + (1-\lambda_\delta)\, \delta_{\rho_+} , \tag{4.6}$$

where
$$\lambda_\delta = \frac{\alpha\, e^{\delta\rho_-}}{\alpha\, e^{\delta\rho_-} + (1-\alpha)\, e^{\delta\rho_+}} . \tag{4.7}$$

$$\gamma > 1 \quad : \quad \tilde{\mathbb{K}}^{\mu_c}_{\gamma,\delta} = \lambda_\delta \, \delta_{\rho_-} + (1-\lambda_\delta)\, \delta_{\rho_+} . \tag{4.8}$$

Note that

$$\int_{[0,\infty)} e^{-tx}\, \tilde{\mathbb{K}}^{\mu_\ell}_{\ell}[dx] = \frac{\int_{[0,\infty)} e^{-(t-\delta V_\ell^{(1-\gamma)})x}\, \tilde{\mathbb{K}}^{\mu_c}_{\ell}[dx]}{\int_{[0,\infty)} e^{\delta V^{(1-\gamma)}x}\, \tilde{\mathbb{K}}^{\mu_c}_{\ell}[dx]} .$$

Now we can choose B and ℓ_0 sufficiently large so that

$$\frac{\delta}{V_\ell^{\gamma-1}} x + \beta V_\ell \{ \mu_c x - w(x) - f_\ell(x) \} < - \beta V_\ell x$$

for $x \geqslant B$ and $\ell \geqslant \ell_0$. This follows from the fact that $\lim_{x \to \infty} \frac{w(x)}{x} = \infty$ and the fact that $f_\ell(x) > -x$ for $x > B$ and $\ell > \ell_0$. Then for $t \geqslant 0$,

$$\int_{[B,\infty)} e^{-(t - \delta V_\ell^{(1-\gamma)})x} \tilde{\mathbb{K}}_\ell^{\mu_c}[dx]$$

$$= e^{-\beta V_\ell \tilde{\rho}_\ell(\mu_c)} \int_{[B,\infty)} e^{-(t - \delta V_\ell^{(1-\gamma)})x} e^{\beta V_\ell \{ \mu_c x - w(x) - f_\ell(x) \}} m_\ell[dx]$$

$$\leqslant \int_{[B,\infty)} e^{-\beta V_\ell x} m_\ell[dx]$$

$$< \frac{e^{-\beta V_\ell B}}{1 - e^{-\beta}} \qquad \text{for } \ell \geqslant \ell_0 .$$

Therefore

$$\lim_{k \to \infty} \int_{[0,\infty)} e^{-t \delta V^{(1-\gamma)} x} \tilde{\mathbb{K}}_\ell^{\mu_c}[dx]$$

$$= \lim_{k \to \infty} \int_{[0,B)} e^{-(t - \delta V^{(1-\gamma)})x} \tilde{\mathbb{K}}_\ell^{\mu_c}[dx]$$

and the proof of (4.7) and (4.8) follows as for (4.2), (4.3) and (4.4).

We end with some remarks on boson condensation in this situation; it is only necessary to comment on the case in which the Bose-Einstein critical density ρ_c lies between ρ_- and ρ_+. In this case we have, in the standard example described in the introduction perturbed by the mean-field term $V_\ell w(X_\ell)$ the following result for the occupation of the ground state.

$$\lim_{\ell\to\infty} \tilde{\mathbb{E}}_\ell^\mu \left[\frac{\sigma_1}{V_\ell} \right] = \begin{cases} 0 , & \mu < \mu_c , \\ \rho_+ - \rho_c , & \mu = \mu_c , \\ \rho(\mu) - \rho_c , & \mu > \mu_c . \end{cases} \tag{4.9}$$

<u>Appendix</u>: <u>The free-energy density of the free boson gas</u>

Here we prove the results about the free-energy density of the free boson gas which we used in the body of the paper. First, we remark that, as

$$\lambda_\ell(1) = 0 , \quad p_\ell(\mu) > (\beta V_\ell)^{-1} \ln(1 - e^{+\beta\mu})^{-1}$$

so that $p_\ell(\mu) \to \infty$ as μ increases to zero. Moreover, it was proved in [1] that when (S 1) and (S 2) hold we have $p(\mu) = \lim_{\ell\to\infty} p_\ell(\mu)$ exists for $\mu < 0$ and is given by $p(\mu) = \int_{[0,\infty)} p(\mu|\lambda)dF(\lambda)$ where $p(\mu|\lambda) = \beta^{-1}\ln(1 - e^{-\beta(\lambda-\mu)})^{-1}$.
It was proved also that

$$p^*(x) = \sup_{\mu<0} \{ \mu x - p(\mu) \}$$

is given by

$$p^*(x) = x\mu(x) - p(\mu(x))$$

where $\mu(x) = 0$ for $x > \rho_c$ and $\mu(x)$ is the unique real root of $x = p'(\mu)$ for $x < \rho_c$; the function $p(\mu)$, defined on $(-\infty, 0)$; is extended defining

$$p(0) = \lim_{\mu\uparrow 0} p(\mu) = \int_{[0,\infty)} p(0|\lambda)\, dF(\lambda) .$$

Lemma 1

The function $x \mapsto f_\ell(x)$ is convex.

Proof: It is enough to prove that, for each n,

$$f_\ell\left(\frac{n}{V_\ell}\right) \leq \frac{1}{2} f_\ell\left(\frac{n-1}{V_\ell}\right) + \frac{1}{2} f_\ell\left(\frac{n+1}{V_\ell}\right) \quad ;$$

that is, that

$$Z_\ell(n)^2 \geqslant Z_\ell(n-1)\, Z_\ell(n+1) \qquad (*)$$

where

$$Z_\ell(n) = \sum_{\{\omega : N(\omega) = n\}} e^{-\beta\{\lambda_\ell(1)\omega_1 + \lambda_\ell(2)\omega_2 + \cdots\}} \quad .$$

We proceed by induction on the number of levels: .et

$$Z_\ell^k(n) = \sum_{\{\omega : N(\omega) = n\}} e^{-\beta\{\lambda_\ell(1)\omega_1 + \cdots + \lambda_\ell(k)\omega_k\}} \quad .$$

For k = 1, the result (*) holds trivially.
Assume that

$$Z_\ell^k(n)^2 \geqslant Z_\ell^k(n-1)\, Z_\ell^k(n+1) \qquad \text{for all } n \geqslant 1 \ ,$$

so that

$$Z_\ell^k(n)\, Z_\ell^k(m) \geqslant Z_\ell^k(n+1)\, Z_\ell^k(m-1) \quad .$$

Now
$$Z_\ell^{k+1}(n) = \sum_{m=0}^{n} z^{n-m} Z_\ell^k(m)$$

where

$$z = e^{-\beta \lambda_\ell(k+1)} \quad \text{, so that}$$

$$(Z_\ell^{k+1}(n))^2 = S + Z_\ell^k(n) \sum_{m=0}^{n} z^{n-m} Z_\ell^k(m) ,$$

where

$$S = \sum_{m_1=0}^{n-1} \sum_{m_2=0}^{n} z^{2n-m_1-m_2} Z_\ell^k(m_1) Z_\ell^k(m_2)$$

while

$$Z_\ell^{k+1}(n-1) Z_\ell^{k+1}(n+1) = S + Z_\ell^k(n+1) \sum_{m=1}^{n} z^{n-m} Z_\ell^k(m-1) .$$

Thus

$$(Z_\ell^{k+1}(n))^2 - Z_\ell^{k+1}(n-1) Z_\ell^{k+1}(n+1)$$

$$= z^n Z_\ell^k(n) + \sum_{m=1}^{n} z^{n-m} \{ Z_\ell^k(n) Z^k(m) - Z_\ell^k(n+1) Z_\ell^k(m-1) \}$$

$$\geqslant 0 \quad , \text{ by } (*) \quad \blacksquare$$

Lemma 2

For the free boson gas, the finite-volume free-energy density is a decreasing function:

$$f_\ell(x) \leq f_\ell(y) \quad \text{for all} \quad x \geq y\,.$$

Proof:

Since $x \mapsto f_\ell(x)$ is convex, it has a line of support at each point: for each y, there exists $a_\ell(y)$ such that

$$f_\ell(x) - f_\ell(y) \geq a_\ell(y)(x-y) \qquad \text{for all } x.$$

Suppose there exists a point x_0 where $a_\ell(x_0) > 0$; then, for each $\mu < 0$, we have

$$e^{\beta V_\ell p_\ell(\mu)} = \int_{[0,\infty)} e^{-\beta V_\ell \{f_\ell(x) - \mu x\}} m_\ell[dx]$$

$$\leq \int_{[0,x_0)} e^{-\beta V_\ell f_\ell(x)} m_\ell[dx] + e^{-\beta V_\ell f_\ell(x_0)} \int_{[x_0,\infty)} e^{-\beta V_\ell a_\ell(x_0)(x-x_0)} m_\ell[dx]$$

$$< \infty\,, \quad \text{since } a_\ell(x_0) > 0\,.$$

But $p_\ell(\mu) \to \infty$ as $\mu \uparrow 0$; contradiction. Hence $a_\ell(y) \leq 0$ for all y and

$$f_\ell(x) - f_\ell(y) \leq 0 \qquad \text{for all} \qquad x \geq y \quad \blacksquare$$

Lemma 3

For all $x \geqslant 0$, $\liminf_{\ell \to \infty} f_\ell(x) \geqslant p^*(x) = \sup_{\mu < 0} \{ \mu x - p(\mu) \}$.

Proof: We have

$$e^{\beta V_\ell p_\ell(\mu)} = \int_{[0,\infty)} e^{-\beta V_\ell \{f_\ell(x) - \mu x\}} m_\ell[dx]$$

$$\geqslant e^{-\beta V_\ell \{f_\ell(x) - \mu x\}} \quad \text{for } 0 < x < \infty .$$

$$e^{\beta V_\ell p_\ell(\mu)} = \sum_{n \geqslant 0} e^{-\beta V_\ell \{ f_\ell(\frac{n}{V_\ell}) - \mu \frac{n}{V_\ell} \}} .$$

Thus $p_\ell(\mu) \geqslant -f_\ell(\frac{n}{V_\ell}) + \frac{n}{V_\ell}$ for each n.

Since $f_\ell(x)$ is defined for all x in $(0,\infty)$ by linear interpolation it follows that

$$p_\ell(\mu) \geqslant -f_\ell(x) + \mu x \qquad \text{for all x in } (0,\infty) \;;$$

thus

$$f_\ell(x) \geqslant \mu x - p_\ell(\mu)$$

so that

$$\liminf_{\ell \to \infty} f_\ell(x) \geqslant \mu x - p(\mu) .$$

Hence

$$\liminf_{\ell \to \infty} f_\ell(x) \geqslant \sup_{\mu < 0} \{ \mu x - p(\mu) \} = p^*(x) \blacksquare$$

Lemma 4 For all $x < \rho_c$, $\limsup_{\ell \to \infty} f_\ell(x) \leqslant p^*(x)$.

Proof: For the measure $\mathbb{K}_\ell^\mu$ defined in §2 we have

$$\int_{[0,\infty)} e^{-sx}\,\mathbb{K}_\ell^\mu[dx] = \exp\left\{-\frac{s\,(p_\ell(\mu) - p_\ell(\mu-\delta_\ell))}{\delta_\ell}\right\}, \quad \delta_\ell = \frac{s}{V_\ell}.$$

Now for the free boson gas if $\mu < 0$, (see [1]),

$$\lim_{\ell\to\infty} \frac{p_\ell(\mu) - p_\ell(\mu-\delta_\ell)}{\delta_\ell} = p'(\mu).$$

Therefore $\{\mathbb{K}_\ell^\mu\}$ converges weakly to $\delta_{p'(\mu)}$.
Let $x \in [0,\rho_c]$ and $\delta \in (0,\infty)$;

then $\lim_{\ell\to\infty} \mathbb{K}_\ell^{\mu(x)-\delta}[(p'(\mu(x)-2\delta), x]] = 1$.

But, by Lemma 2,

$$\frac{1}{\beta V_\ell} \ln \mathbb{K}_\ell^{\mu(x)-\delta}[(p'(\mu(x)-2\delta), x]]$$

$$\leq \frac{1}{\beta V_\ell} \ln V_\ell\,(x - p'(\mu(x)-2\delta) + 1) - f_\ell(x)$$
$$+ (\mu(x)-\delta)\,p'(\mu(x)-2\delta) - p_\ell(\mu(x)-\delta).$$

Thus $\limsup_{\ell\to\infty} f_\ell(x) \leq (\mu(x)-\delta)\,p'(\mu(x)-2\delta) - p(\mu(x)-\delta)$.

Since p and p' are continuous ([1]) and δ is arbitrary, this proves the lemma.

Lemma 5 For all $x \geq \rho_c$, $\limsup_{\ell\to\infty} f_\ell(x) \leq p^*(\rho_c)$.

Proof: By Lemma 2, for every $\epsilon > 0$ and $x \geq \rho_c$, we have

$$f_\ell(x) \leq f_\ell(\rho_c - \epsilon);$$

hence

$$\limsup_{\ell \to \infty} f_\ell(x) \leq \limsup_{\ell \to \infty} f_\ell(\rho_c - \epsilon) .$$

But, by lemma 4, we have

$$\limsup_{\ell \to \infty} f_\ell(\rho_c - \epsilon) < p^*(\rho_c - \epsilon) ,$$

so that

$$\limsup_{\ell \to \infty} f_\ell(x) \leq p^*(\rho_c)$$

since ϵ is arbitrary and p is continuous ■

Since $p^*(x) = p^*(\rho_c)$ for $x \geq \rho_c$, we have by Lemma 4 and Lemma 5 that

$$\limsup_{\ell \to \infty} f_\ell(x) \leq p^*(x) \text{ for } x \geq 0 .$$

Combining this with Lemma 3, we establish Theorem 2: $\lim_{\ell \to \infty} f_\ell(x) = p^*(x)$ and $\{f_\ell\}$ is bounded on compacts, by Lemma 2.

Since $f_\ell(x) \leq 0$, $\mu_\infty \leq 0$. From the inequality $f_\ell(x) \geq \mu x - p_\ell(\mu)$ for $\mu < 0$ in Lemma 3, we get $\mu_\infty \geq 0$ ■

REFERENCES

[1] van den Berg M., Lewis J.T., Pulè J.V., Helv. Phys. Acta 59, 1271-1288 (1986).

[2] Zagrebnov V.A., Papoian, Vl. V., Theor. and Math. Phys. 69, 420-438 (1986).

[3] Davies E.B., Comm. Math. Phys. 28, 69-86 (1972).

[4] Lewis J.T., Pulè J.V., LNM 1095, 25-35 (1984).

[5] Varadhan S.R.S., Comm. Pure and Appl. Math. 19 261-286 (1966).

[6] Ellis R., Entropy, Large Deviations and Statistical Mechanics, New York: Springer 1985

[7] Eisele Th., Ellis R.S., Multiple Phase Transitions in the Generalized Curie-Weiss Model.

[8] Buffet E., Pulè J.V., J. Math. Phys. 24, 1608-1616 (1983).

[9] Bogoliubov N.N., Lectures on Quantum Statistics, Vol. II: Quasi-Averages. New York: Gordon and Breach, 197.

[10] Angelescu N., Zagrebnov V.A., J. Stat. Phys. 41, 323-334 (1985).

[11] Lewis J.T., Pulè J.V., Zagrebnov V.A., The Large Deviation Principle for the Kac Distribution (submitted to Helv. Phys. Acta).

Lecture IV: Limit Theorems for Stochastic Processes Associated with a Boson Gas

§1 Introduction

In this lecture, we discuss the density of particles having energy less than λ in a boson system as a stochastic process indexed by λ. Recall that the hamiltonian for the free boson gas is given by

$$H_\ell(\omega) = \sum_{j\geq 1} \lambda_\ell(j)\, \sigma_j(\omega) \quad , \tag{1.1}$$

where $0 = \lambda_\ell(1) \leq \lambda_\ell(2) \leq \cdots$. For a system in a region of volume V_ℓ, the grand canonical pressure $p_\ell(\mu)$ is defined for $\mu < 0$ by

$$p_\ell(\mu) = \frac{1}{\beta V_\ell} \ln \Big\{ \sum_{\omega \in \Omega} e^{\beta(\mu N(\omega) - H_\ell(\omega))} \Big\} \quad . \tag{1.2}$$

In Lecture II, we recalled results (proved in [3]) on the existence of the pressure in the thermodynamic limit:

$$p(\mu) = \lim_{\ell \to \infty} p_\ell(\mu) . \tag{1.3}$$

In order to discuss the phenomenon of boson condensation, we introduced in [3] the family of random variables $\{X_\ell(\cdot\,;\lambda) : \lambda \geq 0\}$ defined by

$$X_\ell(\omega;\lambda) = \frac{1}{V_\ell} \sum_{\{j : \lambda_\ell(j) \leq \lambda\}} \sigma_j(\omega) \quad . \tag{1.4}$$

For the free boson gas, we have the following result:

Theorem 1

Suppose that (S 1) and (S 2) hold: then, for ρ_c finite,

$$\lim_{\lambda \downarrow 0} \lim_{\ell \to \infty} \mathbb{E}_\ell^\rho [X_\ell(\lambda)] = (\rho - \rho_c)^+ , \tag{1.5}$$

Here $\mathbb{E}_\ell^\rho[\cdot]$ denotes the expectation taken with respect to the grand canonical probability measure $\mathbb{P}_\ell^\mu[\cdot]$ with $\mu = \mu_\ell(\rho)$; it is the expectation at fixed mean density.

Proof:

From the definition of $X_\ell(\lambda)$, we have

$$\mathbb{E}_\ell^\rho[X_\ell(\lambda)] = \int_{[0,\lambda)} p'(\mu_\ell(\rho)|\lambda)\, dF_\ell(\lambda) = \rho - \int_{[\lambda,\infty)} p'(\mu_\ell(\rho)|\lambda)\, dF_\ell(\lambda). \quad (1.6)$$

But, for $\mu < \lambda$, the sequence

$$\left\{ \int_{[\lambda,\infty)} p'(\mu|\lambda)\, dF_\ell(\lambda) : \ell = 1, 2, \cdots \right\} \quad (1.7)$$

converges uniformly in μ on compacts to

$$\int_{[\lambda,\infty)} p'(\mu|\lambda)\, dF(\lambda) . \quad (1.8)$$

Hence, by Proposition 2 of [2], we have for $\lambda > 0$:

$$\lim_{\ell\to\infty} \mathbb{E}_\ell^\rho[X_\ell(\lambda)] = \rho - \int_{[\lambda,\infty)} p'(\mu|\lambda)\, dF(\lambda) . \quad (1.9)$$

But, by hypothesis, ρ_c is finite so that we may invoke the dominated convergence principle to conclude that

$$\lim_{\lambda\downarrow 0} \int_{[\lambda,\infty)} p'(\mu(\rho)|\lambda)\, dF(\lambda) = \int_{[0,\infty)} p'(\mu(\rho)|\lambda)\, dF(\lambda) = \begin{cases} \rho , & \rho < \rho_c , \\ \rho_c , & \rho \geq \rho_c . \end{cases}$$

Thus we have

$$\lim_{\ell \to \infty} \mathbb{E}_\ell^\rho[X_\ell(\lambda)] = (\rho - \rho_c)^+ \blacksquare$$

In the free boson gas there is a second effect, discovered by M. Kac in 1971. We saw in Lecture II that the free-energy has a first-order phase-transition segment $[\rho_c, \infty)$; it follows that for $\rho > \rho_c$ there is no guarantee that the weak law of large numbers wil hold for the distribution $\mathbb{K}_\ell^\rho = \mathbb{P}_\ell^\rho \circ X_\ell^{-1}$ of the number density $X_\ell = N/V_\ell$. In fact, there is no guarantee that, for $\rho > \rho_c$, the sequence $\{\mathbb{K}_\ell^\rho : \ell = 1, 2, \dots\}$ will converge; nevertheless, by the Helly Selection Principle, a subsequence will converge, but the limit distribution will depend on the detailed behaviour of the corresponding subsequence of the sequence $\{\lambda_\ell(\cdot) : \ell = 1, 2, \dots\}$. In other words, it is possible to have two sequences, $\{\lambda_\ell(\cdot) : \ell = 1, 2, \dots\}$ and $\{\hat{\lambda}_\ell(\cdot) : \ell = 1, 2, \dots\}$, each satisfying (S 1) and (S 2) and having the same integrated density of states F(•) but having limit distributions $\mathbb{K}^\rho$ and $\hat{\mathbb{K}}^\rho$ which are distinct for $\rho > \rho_c$. (For $\rho < \rho_c$, they must both be equal to δ_ρ, the degenerate distribution concentrated at ρ, by Theorem 1 of Lecture I.) For example, Kac showed that in the standard example the limit distribution is the exponential distribution supported on $[\rho_c, \infty)$ with mean ρ for $\rho > \rho_c$; other examples are investigated in detail in [3]. We shall see in the next section that, in the mean-field model, this phenomenon disappears: there is no first-order phase-transition segment, the grand canonical pressure exists for all values of μ and is a differentiable function; the weak law of large numbers holds for X_ℓ for all values of the mean density ρ; nevertheless, condensation persists. In these circumstances it becomes interesting to regard $\lambda \mapsto X_\ell(\cdot;\lambda)$ as a stochastic process and to enquire about the convergence in distribution of a re-scaled, centred version of it. This we do in §3.

§2 The Mean Field-Model

To describe the mean-field model, we define a sequence of hamiltonians $\{H_\ell^{MF} : \ell = 1, 2, \dots\}$ by

$$H_\ell^{MF}(\omega) = H_\ell(\omega) + \frac{a}{2V_\ell} N(\omega)^2 \tag{2.1}$$

with $a > 0$. The term $\frac{a}{2V_\ell} N^2$, which provides a crude caricature of the interaction, can be understood classically: it arises in an "index of refraction" approximation in which we imagine each particle to move through the system as if it were moving in a uniform optical medium and so receiving an increment of energy proportional to the density $X_\ell = N/V_\ell$; since a is positive, the interaction is repulsive.

First, we compute the pressure $p_\ell^{MF}(\mu)$, as explained in §4 of Lecture I; writing $u(x) = (\mu - \alpha)x - \frac{a}{2}x^2$, a straight-forward manipulation gives

$$p_\ell^{MF}(\mu) = p_\ell(\alpha) + \frac{1}{\beta V_\ell} \ln \mathbb{E}_\ell^\alpha [e^{\beta V_\ell u(X_\ell)}] = p_\ell(\alpha) + \frac{1}{\beta V_\ell} \ln \int_{[0,\infty)} e^{\beta V_\ell u(x)} \mathbb{K}^\alpha[dx] \quad (2.2)$$

for each $\alpha < 0$, where $\mathbb{K}_\ell^\alpha = \mathbb{P}_\ell^\alpha \circ X_\ell^{-1}$. But $x \mapsto u(x)$ is continuous and bounded above and $\{\mathbb{K}_\ell^\alpha : \ell = 1, 2, \dots\}$ satisfies the Large Deviation Principle with rate-function $I^\alpha(x) = p(\alpha) + f(x) - \alpha x$, by Theorem 1 of Lecture II. Hence, by Varadhan's First Theorem, $p^{MF}(\mu) = \lim_{\ell \to \infty} p_\ell^{MF}(\mu)$ exists and is given by

$$p^{MF}(\mu) = p(\alpha) + \sup_{[0,\infty)} \{u(x) - I^\alpha(x)\} = \sup_{[0,\infty)} \{u(x) - f^{MF}(x)\}, \quad (2.3)$$

where the mean-field free-energy $f^{MF}(x)$ is given by

$$f^{MF}(x) = f(x) + \frac{a}{2} x^2 . \quad (2.4)$$

Thus we have proved:

Theorem 2

Suppose that (S 1) and (S 2) hold; then the mean-field pressure exists for all real μ and is given by

$$p^{MF}(\mu) = \sup_{[0,\infty)} \{ \mu x - f^{MF}(x) \} , \quad (2.5)$$

where $x \mapsto f^{MF}(x)$ is the mean-field free energy, given by $f^{MF}(x) = f(x) + \frac{a}{2}x^2$.

Next, we introduce the mean-field expectation functional $\widetilde{\mathbb{E}}_\ell^\mu[\cdot]$ defined by

$$\tilde{\mathbb{E}}_\ell^\mu[\cdot] = \frac{\mathbb{E}_\ell^\alpha[\cdot \, e^{\beta M_\ell}]}{\mathbb{E}_\ell^\alpha[e^{\beta M_\ell}]} \tag{2.6}$$

and the associated probability measure $\tilde{\mathbb{P}}_\ell^\mu[\cdot]$, where

$$M_\ell = V_\ell u(X_\ell). \tag{2.7}$$

Corollary

The mean-field pressure $\mu \mapsto p^{MF}(\mu)$ is differentiable for all values of μ. The sequence of distribution functions $\{\tilde{\mathbb{K}}_\ell^\mu = \tilde{\mathbb{P}}_\ell^\mu \circ X_\ell^{-1}\}$ converges weakly to the degenerate distribution δ_ρ concentrated at $\rho = p'(\mu)$ and satisfies the Large Deviation Principle with constants $\{V_\ell\}$ and rate-function $\tilde{I}^\mu(x) = p^{MF}(\mu) + f^{MF}(x) - \mu x$.

Proof:

Since $x \mapsto f(x)$ is strictly convex for $0 \leq x < \rho_c$ and constant for $\rho_c \leq x < \infty$ and $x \mapsto \frac{a}{2}x^2$ is strictly convex for $0 \leq x < \infty$, the function $x \mapsto f^{MF}(x) = f(x) + \frac{a}{2}x^2$ is strictly convex for $0 \leq x < \infty$, hence there is no first-order phase-transition segment; equivalently, $\mu \mapsto p^{MF}(\mu)$ the Legendre transform of $x \mapsto f^{MF}(x)$, is differentiable for $\mu < \infty$. It follows from Theorem 1 of [1] that $\tilde{\mathbb{K}}_\ell^\mu \to \delta_\rho$, where $\rho = p^{MF\prime}(\mu)$, and from Theorem 4 of [1] that $\{\tilde{\mathbb{K}}_\ell^\mu : \ell = 1, 2, \dots\}$ satisfies the Large Deviation Principle with constants $\{V_\ell\}$ and rate-function $\tilde{I}^\mu(\cdot)$ ∎

Although the first-order phase-transition segment, which was present in the free energy function of the free-gas, has disappeared, the phenomenon of condensation persists:

Theorem 3

Suppose that (S 1) and (S 2) hold: then, for ρ_c finite, we have

$$\lim_{\lambda \downarrow 0} \lim_{\ell \to \infty} \tilde{\mathbb{E}}_\ell^\mu[X_\ell(\lambda)] = (\rho - \rho_c)^+, \tag{2.8}$$

where $\tilde{\mathbb{E}}_\ell^\mu[\cdot]$ is the mean-field expectation functional and $\rho = (p^{MF})'(\mu)$.

Proof:

First, we remark that an elementary exercise yields the following alternative formula for the mean-field pressure

$$p^{MF}(\mu) = \inf_{\alpha<0} \left\{ \frac{(\mu-\alpha)^2}{2a} + p(\alpha) \right\} \tag{2.9}$$

where $p(\alpha)$ is the free-gas pressure. The idea of the proof of (2.8) is that we compute the cumulant generating function of $X_\ell(\lambda)$; since

$$V_\ell X_\ell(\lambda) = V_\ell X_\ell - \sum_{\{j:\, \lambda_\ell(j)>\lambda\}} \sigma_j \tag{2.10}$$

we get

$$\tilde{\mathbb{E}}_\ell^\mu [e^{\beta s V_\ell X_\ell(\lambda)}] = \tilde{\mathbb{E}}_\ell^{(s,\lambda),\mu} [e^{\beta s V_\ell X_\ell}]$$

where $\tilde{\mathbb{E}}_\ell^{(s,\lambda),\mu}[\,\cdot\,]$ is the mean-field expectation functional for which the free-gas hamiltonian has been modified by the addition of the term $\sum_{\{j:\lambda_\ell(j)>\lambda\}} s\sigma_j$. These considerations yield the formula

$$\lim_{\ell\to\infty} \tilde{\mathbb{E}}_\ell^\mu [X_\ell(\lambda)] = \frac{\partial}{\partial s} p^{MF}(\mu+s; s,\lambda) \Big|_{s=0} \tag{2.11}$$

where

$$p^{MF}(\mu+s;s,\lambda) = \inf_{\alpha<0} \left\{ \frac{(\mu+s-\alpha)^2}{2a} + p(\alpha;s,\lambda) \right\} ,$$

and

$$p(\alpha;s,\lambda) = \int_{[0,\lambda)} p(\alpha\,|\,\lambda)\, dF(\lambda) + \int_{[\lambda,\infty)} p(\alpha\,|\,s+\lambda)\, dF(\lambda) .$$

A standard argument , using Griffith's Lemma, yields the result ∎

§3 Fluctuations in the Mean-Field Model

Fluctuations in $X_\ell = N/V_\ell$ in the mean-field model in the thermodynamic limit were studied for the standard example by Davies [4], Wreszinski [5], Fannes and Verbeure [6] and Buffet and Pule [7]. The mean-field model in the general situation, where the only assumptions about the single-particle spectrum are that (S 1) and (S 2) hold, was investigated in [8] we have summarized the results of [8] in §2 and now go on to investigate the fluctuations in X_ℓ. In fact, we do rather more; we regard $\lambda \mapsto X_\ell(\lambda)$ as a stochastic process and prove a central limit theorem:

Theorem 4

Let $Z_\ell(\lambda) = V_\ell^{1/2}\{X_\ell(\lambda) - \tilde{\mathbb{E}}_\ell^\mu[X_\ell(\lambda)]\}$ then, for $\mu < a\rho_c$, $Z_\ell(\lambda) \to Z(\lambda)$ where $Z(\lambda)$ is gaussian with mean zero and covariance $\Gamma(\lambda_1, \lambda_2)$ given by

$$\Gamma(\lambda_1, \lambda_2) = J^\mu_{\lambda_1 \wedge \lambda_2} - \frac{a J^\mu_{\lambda_1} J^\mu_{\lambda_2}}{1 + a J^\mu_\infty}, \tag{3.1}$$

where

$$J^\mu_\lambda = \int_{[0,\lambda)} p''(\alpha(\mu) | \lambda)\, dF(\lambda), \tag{3.2}$$

and $\alpha(\mu)$ is the value of α at which $\inf\limits_{\alpha < 0} \left\{ \frac{(\mu - \alpha)^2}{2a} + p(\alpha) \right\}$ is attained.

Sketch of Proof:

The result follows from a routine, but somewhat tedious, calculation of

$$\lim_{\ell \to \infty} \tilde{\mathbb{E}}^\mu_\ell [e^{\beta(s_1 Z_\ell(\lambda_1) + s_2 Z_\ell(\lambda_2))}]$$

along the lines of the proof of Theorem 3 ∎

It is interesting to identify the process $Z(\cdot)$ in terms of a standard process.

Theorem 5

Let $B(\cdot)$ be a BM(1), a brownian motion in $\mathbb{R}^1$ starting at zero; then, for $\mu < a\rho_c$,

$$Z(\lambda) \overset{(d)}{=} B(J^\mu_\lambda) - \frac{a J^\mu_\lambda}{1 + a J^\mu_\infty} B(J^\mu_\infty + 1/a) . \tag{3.3}$$

Proof:

A routine computation shows that the mean of the right-hand side of (3.3) is zero and the covariance is the same as that of $Z(\cdot)$, given by (3.1). Hence the two gaussian processes are equal in distribution. ∎

The process (3.3) is a modification of a time-changed brownian bridge; it never reaches the point at which it is tied-down but, as a increases, that point comes closer to J^μ_∞ . This shows how, as the strength of the interaction increases, the fluctuations in $Z(\infty)$ are damped down.

It is a little more difficult to deal with the case $\mu > a\rho_c$; we introduce

$$W_\ell(\lambda) = Z_\ell(\infty) - Z_\ell(\lambda) \tag{3.4}$$

and prove in anal gous fashion:

Theorem 6

For $\mu > a\rho_c$, $W_\ell(\lambda) \overset{(d)}{\to} W(\lambda)$ where $W(\cdot)$ is a gaussian process with mean zero and covariance $\Gamma(\lambda_1,\lambda_2)$ given by

$$\Gamma(\lambda_1,\lambda_2) = K^\mu_{\lambda_1 \vee \lambda_2} , \tag{3.5}$$

where

$$K^\mu_\lambda = \int_{[\lambda,\infty)} p''(0|\lambda)\, dF(\lambda) . \tag{3.6}$$

In this case,

$$W(\lambda) \overset{(d)}{=} K_\lambda^\nu B\left(\frac{1}{K_\lambda^\nu}\right) \quad . \tag{3.7}$$

The method by which we discovered the representations may be of some interest. The stochastic differential equation satisfied by a process $(X_t)_{t \geq 0}$ with filtration $(\mathcal{F}_t)$ is discussed by Nelson [9]; see also McGill [10].

Suppose that a process $(X_t, \mathcal{F}_t)$ satisfies the stochastic differential equation

$$X_t = X_s + \int_s^t \sigma(u, X_u)\, dB(u) + \int_s^t \tau(u, X_u)\, du \tag{3.8}$$

then

$$\tau(s, X_s) = \lim_{t \downarrow s} \frac{1}{t-s} \mathbb{E}[X_t - X_s | \mathcal{F}_s] \tag{3.9}$$

and

$$\sigma^2(s, X_s) = \lim_{t \downarrow s} \frac{1}{t-s} \mathbb{E}[(X_t - X_s)^2 | \mathcal{F}_s] \quad . \tag{3.10}$$

Assuming that the processes $Z(\cdot)$, $W(\cdot)$ satisfy stochastic differential equations, the corresponding coefficients σ and τ can be computed using (3.9) and (3.10); this is a routine exercise starting from the expressions (3.1) and (3.5) for the covariances since the processes are gaussian. Obvious time-changes then give the stochastic differential equations for a brownian bridge and a brownian motion respectively.

REFERENCES

[1] Lewis,J.T.,: The Large Deviation Principle in Statistical Mechanics (see Lecture I)

[2] van den Berg, M., Lewis J.T., Pulè J.V.,: Large Deviations and the Boson Gas (see Lecture II)

[3] van den Berg, M., Lewis J.T., Pulé J.V.,: A General Theory of Bose-Einstein Condensation, Helv. Phys. Acta, 59, 1271-1288 (1986).

[4] Davies E.B.: The Thermodynamic Limit for an Imperfect Boson Gas, Commun. Math. Phys. 28, 69-86 (1972).

[5] Wreszinski W.F.,: Normal Fluctuations in some Mean-Field Models in Quantum Statistical Mechanics, Helv. Phys. Acta, 46, 844-868 (1974).

[6] Fannes M., Verbeure A.,: The Imperfect Boson Gas, J. Math. Phys. 21, 1809-1818 (1980).

[7] Buffet E., Pulé J.V.,: Fluctuation Properties of the Imperfect Boson Gas, J. Math. Phys. 24, 1608-1616 (1983).

[8] van den Berg, M., Lewis J.T., de Smedt P.,: Condensation in the Imperfect Boson Gas, J. Stat. Phys., 37, 697-707 (1984).

[9] Nelson E.,: Dynamical Theories of Brownian Motion, New Jersey: Princeton University Press, 1967.

[10] McGill P.,: Seminaires de Probabilités, XX, LNM 1204, Springer: Heidelberg 1986.

DYNAMICS OF THE DISSIPATIVE TWO-STATE SYSTEM

H. Dekker, Physics & Electronics Laboratory FEL-TNO, Den Haag

1. The two-state quantum system is of fundamental importance both in physics and in chemistry, for example, in the study of magnetic impurities in metals or molecular transitions in liquids. It also presents an interesting mathematical problem, both in the weak dissipation regime (where the destruction of quantum coherence can be studied in its purest form) and in the strong dissipation regime (where at zero temperature a symmetric - "freezing" - transition occurs[1]). In some cases it can be considered as a Hilbert subspace description of a more complicated system, such as for a quantum mechanical particle moving in a double-well potential[2]. The recent interest in this model has originated, in particular, from progress in superconducting technology (quantum interference devices: SQUIDs). Indeed, in the deterministic limit the dynamics of the magnetic flux Φ threading a superconducting ring interrupted by a Josephson-type weak link (e.g. a thin oxide layer) can be described in terms of a multistable potential $U(x)$ (with coordinate $x=\Phi$) which is experimentally realised by virtue of the nonlinear current-voltage characteristics of the weak link plus an externally applied magnetic field[3]. In addition, the Josephson junction is an inherently dissipative element of ohmic type, so that the classical equation of motion may be written as $m\ddot{x}+2\lambda\dot{x}+U'(x)=0$. However, in view of its small mass (i.e. its capacitance) the system is essentially quantum mechanical rather than classical. Also, thermal effects may not be ignored.

2. From a fundamental point of view dissipation arises from interactions with the environment, for example the phonon system of the solid state device. If one considers[4] the microscopic model of two pieces of BCS-type superconducting material separated by a thin insulating layer, described by means of a standard tunneling Hamiltonian, and one investigates the imaginary time functional integral for the partition function, then it turns out that there is in fact only one dynamically important degree of freedom, namely the phase jump ϕ of the "Cooper pair wave function" across the junction. If this system is incorporated in an electrically closed loop, then the Aharonov-Bohm relation $\phi+(2e/\hbar)\Phi=2\pi m$ ($m=0,1,2,\cdots$) provides the connection between the phase jump and the embraced flux.

Tracing out all "irrelevant" degrees of freedom, one is left with a relatively simple functional integral over ϕ (or Φ) only, which involves an effective action with a Feynman influence functional effectively describing dissipative shot noise. One then observes that in the small BCS energy gap and Gaussian noise limit this effective action coincides with that obtained from a less detailed model, where the potential system is coupled to an infinite set of harmonic oscillators[5] (labelled by $k=0,1,2,\cdots$). The interaction Hamiltonian is taken to be linear both in the particle coordinate x and in the bath coordinates x_k, i.e. $H_{int} = x\Sigma_k c_k x_k$. The coupling strengths define a spectral density $J(\omega) = \frac{1}{2}\pi\Sigma_k(c_k^2/m_k\omega_k)\delta(\omega-\omega_k)$, which suffices for a reduced quantum statistical description of the particle. For ohmic dissipation one must have $J(\omega) = 2\lambda\omega f_c(\omega/\omega_c)$, where f_c is a cutoff function and $\omega_c \rightarrow \infty$.

3. Now let U(x) indeed be a perfectly symmetric bistable potential, with minima at $x=\pm a$ and a high barrier ar $x=0$ (i.e. $U_b >> \hbar\omega_0$, ω_0 being the harmonic frequencies at the local minima). In that case the energy levels group into distinct doublets with an exponentially small splitting $\Delta E_{2n}=E_{2n+1}-E_{2n}$ ($n=0,1,2,\cdots$). In particular, for the ground state doublet the associated tunneling frequency becomes $\Delta_0=(\bar{\omega}_0/\pi)\exp(-S_0/\hbar)$, where $\bar{\omega}_0=2a\omega_0(\pi m\omega_0/\hbar)^{\frac{1}{2}}\exp\omega_0\tau_0$ is called the attempt frequency, and where S_0 and $2\tau_0$ are the instanton action and lifetime, respectively [6,7]. If (i) the pertinent initial condition does not involve the excited state doublets and if (ii) the temperature is sufficiently small ($k_BT << \hbar\omega_0$), then a simplifying two-level description is suggested on the basis of this lowestlying doublet. It has, however, been indicated[2] that one should rather require $k_BT << \hbar\omega_c'$, where ω_c' is a (somewhat arbitrary) adiabatic cutoff with $\Delta_0 << \omega_c' << \omega_0$. (See also the next section.) In the usual spin representation one then arrives at the so-called spin-boson Hamiltonian

$$H = -\tfrac{1}{2}\hbar\Delta_0\sigma_x + a\sigma_z \textstyle\sum_k c_k x_k + \tfrac{1}{2}\sum_k (p_k^2/m_k + m_k\omega_k^2 x_k^2) . \qquad (1)$$

The polarisation σ_z is taken to have the eigenstates $|\pm\rangle$ with eigenvalues ± 1, and $a\sigma_z$ now represents the particle's coordinate.

4. In the displaced-bath oscillators basis $|n_k,\pm\rangle$ the explicit interaction term in (1) vanishes at the expense of dressing the bare tunneling matrix element Δ_0. The adiabatically renormalised value reads $\Delta'(\omega_c') = \Delta_0\Pi_{k>k'}\ \langle n_k,+|n_k,-\rangle$,

where $\omega_{k'} \equiv \omega_c' \gtrsim \Delta'(\omega_c')$. Momentarily confining ourselves to the ground states $|0_k,\pm\rangle = \exp(\pm\frac{1}{2}i\Omega_k)|0\rangle$, where $\Omega_k = (2ac_k/\hbar m_k\omega_k^2)p_k$, we find the following results depending on the infrared nature of the spectral density $J(\omega) \simeq 2\lambda_s\omega^s$. For a subohmic spectrum $(0<s<1)$, letting $\omega_c' \to 0$ always yields $\Delta'(0)=0$, i.e. localisation of the particle (absence of any tunneling). For superohmic spectra $(s>1)$ one finds $\Delta'(0)=\Delta_0\exp(-FC)$, where FC is a finite Franck-Condon factor. Finally, for ohmic dissipation $(s=1)$ the result reads $\Delta'(\omega_c')=(\omega_c'/\omega_c)^\alpha\Delta_0$, where $\omega_c \gg \omega_c'$ is the physical ultraviolet cutoff of the bath spectrum[8] and $\alpha \equiv 4\lambda a^2/\pi\hbar$. Hence, if $\alpha>1$ one has $\Delta_0'(0)=0$, but if $\alpha<1$ the limit $\omega_c \to 0$ cannot be taken adiabatically (as ω_c' catches up with $\Delta'(\omega_c')$ in the renormalisation process) and one rather obtains the finite result $\Delta'(\Delta')=(\Delta_0/\omega_c)^{\alpha/(1-\alpha)}\Delta_0$. That is, there appears to be an ohmic dissipative phase transition (at $T=0$) from a localised to a delocalised state at $\alpha=1$[9,10]. The dynamics generated by (1) yields a further renormalisation of the tunnel frequency (in consequence of Kramers-Kronig causality) which neatly fits into the above picture. Namely, with the adiabatic cutoff ω_c' as (a somewhat arbitrary) ultraviolet cutoff for the remaining tunneling dynamics, one finds a dynamic renormalisation $\Delta''(\omega_c'')= (\omega_c''/\omega_c')^\alpha\Delta'(\omega_c')$ with (at least in the weak coupling limit[11]) $\omega_c'' \to \Delta''$. Notice that in effect $\Delta''(\omega_c'')=(\omega_c''/\omega_c)^\alpha\Delta_0$, so that the intermediate ω_c' becomes irrelevant. In the sequel we define $\Delta \equiv \Delta''(\Delta'')$.

5. The dynamics of the two-state model (1) has been discussed by Leggett et al. using instanton-type path integral techniques[1]. Integration over the environmental coordinates in the real-time double path integral for the density matrix yields a Feynman-Vernon influence functional, the remaining integrations being over particle trajectories only. The problem of two paths connecting two states $|\pm\rangle$ can be reformulated in terms of a single path connecting four states A= (+,+), B=(+,-), C=(-,+) and D=(-,-). The pertinent trajectories then switch instantaneously between diagonal (A and D) and off-diagonal (B and C) states. Calling a path in a diagonal state a "sojourn", and the excursions to off-diagonal states "blips", one then observes a blip-selfinteraction that tends to shorten them relative to the sojourns. If this selfshortening is sufficiently effective, the blips form a dilute "gas", and one introduces the so-called "noninteracting-blip approximation", its justification being the main subject of [1]. The analysis is subtle, and the conclu-

sions are given here only in a rather abbreviated form. That is, the noninteracting-blip approximation (for (1), i.e. at zero bias) can be justified (i) in the extreme weak coupling limit for all spectral densities, (ii) for superohmic spectra in general ($1<s<2$ only at zero temperature) and (iii) in the strong coupling regime in the "Golden rule" limit (incoherent relaxation) provided the latter exists. Presumably excluded is the ohmic case with $\frac{1}{2}<\alpha<1$.

6. The noninteracting-blip formulae can also be obtained in a much simpler way, namely directly from the Heisenberg dynamics in the displaced oscillator basis [12]. In this basis (1) reads

$$H = -\tfrac{1}{2}\hbar\Delta_0(e^{-i\Omega}\sigma_+ + e^{i\Omega}\sigma_-) + \tfrac{1}{2}\sum_k(p_k^2/m_k + m_k\omega_k^2 x_k^2) \quad , \tag{2}$$

where $\sigma_\pm \equiv \frac{1}{2}(\sigma_x \pm \sigma_y)$. Note that (2) is diagonal when $\Delta_0=0$ (or in the strong coupling limit). The Heisenberg equations of motion for $\sigma_\pm(t)$ are easily obtained. Substitution of their formal solutions into the equation of motion for $\sigma_z(t)$ and approximation of the boson displacement operators by the free bath dynamics, yields a first-order differential-integral operator equation for $\sigma_z(t)$ only. The next step is to average this equation with respect to the bath, decorrelating the spin and the free boson $\exp(\pm i\Omega)$. The thermal average of the latter is most easily found invoking the simple Baker-Hausdorff theorem for operators which commute with their commutator, and using a cluster (or cumulant) expansion. Since the free bath is a linear Gaussian system, only the second cumulant contributes. The convolution-type result is conveniently Laplace transformed to yield

$$\langle\bar{\sigma}_z(s)\rangle = 1/(s+\bar{f}(s)) \quad , \tag{3}$$

where $\bar{f}(s)$ is the transform of the memory kernel $f(t)$, which may be specified as

$$f(t) = \Delta_0^2 \cos[Q_1(t)/\pi\hbar]\, \exp[-Q_2(t)/\pi\hbar] \quad , \tag{4}$$

where

$$Q_1(t) = \int_0^\infty d\omega\, \sin\omega t\; J(\omega)/\omega^2 \quad , \tag{5}$$

and

$$Q_2(t) = \int_0^\infty d\omega\, \coth(\tfrac{1}{2}\beta\omega)\; (1-\cos\omega t)\; J(\omega)/\omega^2 \quad , \tag{6}$$

with $\beta=1/k_BT$. These formulae are identical to the "noninteracting-blip" results. Apparently the underlying physical assumption is that the tunneling process does not dynamically disturb the environment too strongly. The above analysis is rather easily extended to include a nonzero bias energy (asymmetry in the double-well potential). See [13].

7. In the case of the ohmic environment $Q_1(t)$ and $Q_2(t)$, and hence $f(t)$, can be evaluated analytically. Disregarding terms of order $\Delta_0^2/\omega_c^2 \ll 1$, the Laplace transform then becomes

$$\bar{f}(s) = \Delta_{eff}\,\{\Delta_{eff}/\nu\}^{1-2\alpha}\,\Gamma(\alpha+s/\nu)/\Gamma(1-\alpha+s/\nu), \qquad (7)$$

where $\Delta_{eff} \sim \Delta$ (apart from factors of order unity[1]) and $\nu \equiv 2\pi k_BT/\hbar$. A special case, namely $\alpha=\frac{1}{2}$ (corresponding to the known solution of the Toulouse-Hamiltonian, describing a localised resonance level at the Fermi energy, related to the Kondo problem) is easily spotted to yield $\langle\sigma_z(t)\rangle=\exp(-\Delta_{eff}\,t)$, with $\Delta_{eff} = \frac{1}{2}\pi\Delta$ and $\Delta=\Delta_0^2/\omega_c$. At high temperatures, such that $\alpha\nu \gg s$, it suffices to employ $\bar{f}(0)$, which leads to the Golden-rule result $\langle\sigma_z(t)\rangle = \exp(-\bar{f}(0)t)$. Internal consistency then requires that $\alpha\nu \gg \bar{f}(0)$, so that the Golden rule appears to be jusitfied for all T if $\alpha>1$, but only for $\alpha k_BT/\hbar\Delta \gg 1$ if $\alpha<1$. At strictly zero temperature one has $\bar{f}(s) = \Delta_{eff}^{2(1-\alpha)}s^{2\alpha-1}$ if $\alpha<1$, and essentially $\bar{f}(s)\equiv0$ if $\alpha>1$. This once again demonstrates the symmetry breaking transition ("freezing" if $\alpha>1$). For $\alpha<1$ the result for $\langle\sigma_z(t)\rangle$ can even be presented as a standard Mittag-Leffler function, which for $\alpha<\frac{1}{2}$ involves a damped oscillatory part with an actual frequency $\Delta_{eff}\cos\theta$ and decay rate $\Delta_{eff}\sin\theta$, where $\theta \equiv \frac{1}{2}\pi\alpha/(1-\alpha)$, and a power-like long time behaviour $\sim (\Delta_{eff}\,t)^{-2(1-\alpha)}$. However, this result has been questioned recently (at least for $\alpha \ll 1$) by second order perturbation theory[14], suggesting a quartic rather than a quadratic tail. Finally, at intermediate temperatures there is a crossover from oscillatory behaviour to pure relaxation at $T^{\star}(\alpha)$; one has $k_BT^{\star}(\alpha\to0)/\hbar\Delta_{eff} \simeq 1/\pi\alpha$ and $k_BT^{\star}(\alpha\to\frac{1}{2})/\hbar\Delta_{eff} \sim 1/\pi$, while $T^{\star}(\alpha>\frac{1}{2}) \equiv 0$.

8. If a bias energy $\pm\frac{1}{2}\epsilon$ is added to the localised states $|\pm\rangle$, by means of a term $\frac{1}{2}\epsilon\sigma_z$ in (1), then the above neglected response of the boson system to the actual presence of the spin must be carried along[13]. The noninteracting-blip result (3) generalises to $\langle\bar{\sigma}_z(s)\rangle = [1-\bar{h}(s)/s]/[s+\bar{g}(s)]$, with $g(t) \equiv f(t)\cos(\epsilon t/\hbar)$

and $h(t) \equiv f(t)\sin(\epsilon t/\hbar)$. For the long time behaviour it suffices to consider $\bar{h}(0)$ and $\bar{g}(0)$, which can be shown to have the equilibrium property $\bar{h}(0)/\bar{g}(0) = \tanh(\epsilon/2k_BT)$ for all spectral densities. Since $\langle\sigma_z(\infty)\rangle = -\bar{h}(0)/\bar{g}(0)$, this predicts symmetry breaking at T=0 for any infinitesimal bias, even for $\alpha<1$ where it is known to be absent. That is, the noninteracting-blip approximation gives qualitatively incorrect behaviour at long time for $\alpha<1$, when ϵ is nonzero but small. Only if $|\epsilon|>>\hbar\Delta_{eff}$, but of course still $|\epsilon|<<\hbar\omega_0$ for the two-state model to apply to the double-well problem, the polarisation σ_z essentially determines the total energy of the spin, and hence can be expected to relax to the thermal equilibrium value. Despite such apparent limitations, it is generally felt that the above outlined analysis provides a proper picture of the typical dynamics of the dissipative two-state system. Unfortunately, cryogenic experiments to date are not yet in the appropriate regime to test the predictions.

References:

[1] A.J. Leggett, S. Chakravarty, A.T. Dorsey, M.P.A. Fisher, A. Garg, W. Zwerger: Rev. Mod. Phys. 59 (1987) 1.

[2] A.T. Dorsey, M.P.A. Fisher, M.S. Wartak: Phys. Rev. A33 (1986) 1117.

[3] R. de Bruyn Ouboter: Physica 126B (1984) 423.

[4] U. Eckern, G. Schön, V. Ambegaokar: Phys. Rev. B30 (1984) 6419.

[5] A.O. Caldeira, A.J. Leggett: Ann. Phys. 149 (1983) 374.

[6] S. Coleman: "Aspects of symmetry", Cambridge Univ. Press 1985, Ch.7.

[7] H. Dekker: Phys. Lett. 114A (1986) 295.

[8] H. Dekker; Phys. Rev. A31 (1985) 1067.

[9] A.J. Bray, M.A. Moore: Phys. Rev. Lett. 49 (1982) 1546.

[10] S. Chakravarty, A.J. Leggett: Phys. Rev. Lett. 52 (1984) 5.

[11] H. Dekker: J. Phys. C (1987), to appear.

[12] H. Dekker: Phys. Rev. A35 (1987) 1436.

[13] H. Dekker: Physica 141A (1987) 570.

[14] H. Dekker: Physica A (1987), to appear.

THE QUANTUM LANGEVIN EQUATION AND ITS STATIONARY STATE

Ph. de Smedt, Instituut voor Theoretische Fysica, K.U. Leuven, (Belgium)

The classical Langevin Equation (LE) is a phenomenological equation which describes the motion of a particle interacting with its environment. The particle can dissipate energy to this environment, but also experiences a random force from it. The LE reads:

$$M \frac{d^2q}{dt^2}(t) + \eta \frac{dq}{dt}(t) + V'(q(t)) = W(t) \quad , \qquad (1)$$

where q(t) denotes the position of the particle, M is its mass, η is the friction constant (η>0) and V(x) is some external potential. W(t) denotes the classical white noise, i.e. it is a Gaussian process with covariance given by:

$$\langle W(t)W(s)\rangle = (2\eta/\beta)\, \delta(t-s) \quad ,$$

β being the inverse temperature of the bath.

The LE has been studied in detail in the literature. In particular, we know its stationary state. Whatever the initial distribution of position and momentum of the particle, when time tends to infinity the distribution of q(t) tends to the Gibbs distribution of the independent particle:

$$C^{st} \exp(-\beta\{ P^2/2M + V(q) \}) \quad .$$

Note that this state does not depend on η.

The following questions are now natural:

- Can this phenomenological equation be obtained from an underlying microscopic model ?
- What is the quantum-mechanical equivalent of (1) ?

The answer to these questions was given in a famous paper by Ford, Kac and Mazur (FKM : J. Math. Phys. 6 (1965) 504). They consider a model with Hamiltonian H, describing a particle in an external potential V and interacting with a set of harmonic oscillators (These oscillators form the "bath".) Moreover, they assume that at time t=0 the system is in thermal equilibrium with respect to the Hamiltonian $H_0 = H - V(q)$. Let us denote this equilibrium state by ω. FKM show that in a certain limit, where the number of particles in the bath is infinite while the couplings between the particles are rather singular, one obtains for the equation of motion of the particle the LE by projecting out the degrees of freedom of the bath. This can also be done in the quantum-mechanical case, and they thus obtained what is now called the quantum Langevin equation (QLE). This equation, however, is in some respects less well-behaved than its classical couterpart. In particular, one faces a problem in defining the momentum operator because

$$\omega(p(t)^2) = +\infty \qquad \text{for all } t.$$

If we define $q(f) = \int q(t)f(t)dt$ for all f in Schwartz's class with support contained in R^+, then the QLE is alternatively written as:

$$q(Mf'' - \eta f') + \int dt\, f(t)\, V'(q(t)) = W(f) \quad , \tag{2}$$

where the noise W(f) is now an operator with the following properties:

$$[W(f),W(g)] \ = \ -2\eta\hbar \int_{-\infty}^{+\infty} dk\, \overline{\hat{f}(k)}\, \hat{g}(k)\, k$$

and

$$\omega(\exp i\lambda\, W(f)) = \exp\{ -\lambda^2 \eta\hbar \int dk\, |\hat{f}(k)|^2 \frac{k}{\exp(\beta\hbar k) - 1}\} \ ,$$

where

$$\hat{f}(k) = (2\pi)^{-\frac{1}{2}} \int dt\, f(t)\, e^{ikt} \ .$$

The following question remained unanswered.

A. What is the stationary state of the QLE ? Until now, this was only known in the case where V is quadratic and there it depends explicitly on η, contrary to the classical case.

B. What is the precise meaning of (2) ? In particular, does it have a solution when the potential V is not quadratic ? And related to this is the following question: How do the initial conditions show up in (2) ? The FKM paper is rather vague on this point. An answer to this question has already been provided by Maassen (J. Stat. Phys. 34 (1984) 239), but our approach is somewhat different, being essentially more clear in the role played by the initial conditions.

Together with D. Dürr, J. Lebowitz and C. Liverani, I have tried to answer these questions, and the results have been written down in a preprint. Here I shall just review the principle result of this paper.

We consider a simple model consisting of a heavy particle and 2N light particles evolving according to the Hamiltonian

$$H^V_{N,L} = P_0^2/2M + \sum_{\substack{i=-N \\ i\neq 0}}^{N} P_i^2/2m_L + \tfrac{1}{2} k_L \sum_{i=-N}^{N} (q_{i+1} - q_i)^2 ,$$

$(q_{N+1}=q_{-N})$ with $m_L=m^*/L$ and $k_L=k^*L$. At time t=0 we thermalize the system at temperature β^{-1} with respect to the Hamiltonian $H^K_{N,L}=H^V_{N,L}-V(q_0)+\frac{1}{2}Kq_0^2$. We denote the corresponding Gibbs state by $\omega^K_{N,L}$ (similarly we write $\omega^V_{N,L}$ for the Gibbs state with respect to $H^V_{N,L}$).

In the classical case, one obtains the LE in the limit $N\to\infty$ followed by the limit $L\to\infty$. It is therefore natural to study this limiting procedure also in the quantum mechanical case.

With regard to question A, we remark that, since the QLE is obtained for this system in the limit $N\to\infty$, $L\to\infty$ (at least for initial states $\omega^K_{N,L}$), we conjecture that the stationary state of the QLE is given by the restriction of the state $\omega^V_\infty \equiv \lim_{L\to\infty}\lim_{N\to\infty} \omega^V_{N,L}$ to the heavy particle. (This is indeed the case when V is quadratic.)

Using the Feynman-Kac formula, we find the following result:

i) $\omega^V_\infty(\exp(i\lambda p_0)) = \delta_{\lambda,0}$;

ii) $\omega^V_\infty(\exp(i\lambda q_0)) = \int dx \exp(i\lambda x)\, \rho^V(x)$.

where

$$\rho^V(x) = \frac{E^G\{\exp(-\int_0^\beta V(\omega(s)+x)\, ds\}}{\int_{-\infty}^{\infty} dx_0\, E^G\{\exp(-\int_0^\beta V(\omega(s)+x_0)\, ds\}}$$

and E^G denotes the expectation with respect to the Gaussian process G living on the paths $\omega(t)$ $(0\le t\le\beta)$ with $\omega(0) = \omega(\beta) = 0$, having mean zero and covariance:

$$(-\hbar^{-2} M\Delta + \hbar^{-1}\eta\,(-\Delta)^{1/2})^{-1}$$

Here Δ denotes the 1-dimensional Laplacian with Dirichlet boundary conditions on the interval $[0,\beta]$.

With respect to question B, we write an integral equation for the operator $q_{N,L}(t)$ defined by:

$$q_{N,L}(t) \equiv \exp(\, i\hbar^{-1} t\, H^V_{N,L}\,)\; q_0\; \exp(\, -i\hbar^{-1} t\, H^V_{N,L}\,)\;.$$

This equation can be written in the form:

$$Mq_{N,L}(t) \;=\; B(f^t_{N,L}(s)) - \int_0^t ds\, f^t_{N,L}(s)\, V_0'(q_{N,L}(s)) \quad , \tag{3}$$

where we have assumed that $V(x) = \frac{1}{2}\overline{K}x^2 + V_0(x)$ with $V_0' \in C^0_\infty(R)$. Here $f^t_{N,L}$ is a specific function depending on $\overline{K}$ which has its support on $[0,t]$. $B(\psi)$ (with ψ in a certain set of functions Φ) is an operator depending on the function ψ and on the position and momentum of all the particles at time zero only. In this operator $B(\psi)$, we can distinguish a part which will play the role of the noise $W(\psi)$ as well as a part pertaining to the position and momentum operator of the heavy particle at time zero.

We next construct an abstract von Neumann algebra M generated by the operators $\exp(i\mu B(\psi))$, $\psi\in\Phi$. The program we carry out consists of the following steps.

(i) We show that on this von Neumann algebra $\exp(\, i\lambda q_{N,L}(t)\,)$ converges strongly to $\exp(i\lambda q_\infty(t))$, where $q_\infty(t)$ is the solution of (3) with $f^t_{N,L}$ replaced by $\lim_{L\to\infty}\lim_{N\to\infty} f^t_{N,L}$. This limiting equation is equivalent to the QLE.

(ii) Next, we show that for all A in this von Neumann algebra :

$$\omega^K_{N,L}(A) \to \omega^K_\infty(A).$$

(iii) From this we conclude that:

$$\lim_{L\to\infty}\lim_{N\to\infty} \omega^K_{N,L}(\, \exp(\, i\lambda q_{N,L}(t)\,)\,) \;=\; \omega^K_\infty(\, \exp(\, i\lambda q_\infty(t)\,)\,)\;.$$

I conclude this summary by remarking that the main merit of this result is to make quite explicit (contrary to the FKM paper) what we mean by the statement that our microscopic model yields the QLE in the limit $N\to\infty$, $L\to\infty$. It also shows that for a relatively large class of potentials V, the QLE has a solution.

References

1. G.W. Ford, M. Kac and P. Mazur, J. Math. Phys. **6** (1965) 504.
2. H. Maassen, J. Stat. Phys. **34** (1984) 239.
3. P. de Smedt, D. Dürr, J.L. Lebowitz and C. Liverani (preprint 1987).

LARGE SYSTEMS WITH LOCALLY INTERACTING COMPONENTS

A. Greven, Institut fur Angewandte Mathematik, Universität Heidelberg (BRD)
(joint work with T. Cox, Department of Mathematics, Syracuse University, USA)

1. The models.

We want to study the time evolution of large systems with locally interacting components. The focus is on the behaviour for large time. More precisely, consider a Markov process with:

state space: S^{E_N}, with $E_N=[-N,N]^d$ and $S=\{0,1\}$, N or R^+,
transitions: $\eta \to \eta'$ at rate $f_x(\eta)$, where $x \in E_N$, $\eta(z)=\eta'(z)$ for all $z \in E_N \setminus \{x\}$,
$f_x(\eta)$ depends on $\eta(z)$ only for z in a "neighbourhood" of x.

We always start our system off in a measure μ on S^{Z^d} or its restriction to S^{E_N}, and μ is translation invariant and shift ergodic (with $\int \eta^2(x) d\mu < \infty$ usually).

The process on Z^d is denoted by $(\eta^\mu_t)_{t \in R^+}$, the finite system by $(\eta^N_t)_{t \in R^+}$. The parameter N is considered to be large (which in computer simulations means typically N>>30). Examples are spin-flip systems such as the stochastic Ising model, voter model, branching random walks, contact process etc.

2. The problem.

The fundamental question for large N is: How do we describe the process for large t? A reasonable description could run as follows. The finite system has a unique equilibrium state, denoted by Π^N. If the infinite system also has a unique equilibrium, say Π, then one usually has $\Pi^N \Rightarrow \Pi$ as $N \to \infty$ (see Spitzer [4]). Therefore a good description of a finite system could be:

(*) $\quad \mathcal{L}((\eta^N_{t+s})_{s \in R^+}) \simeq \mathcal{L}((\eta^\Pi_s)_{s \in R^+}) \quad$ N,t large.

This turns out to be reasonable indeed. However, many interesting systems do not have a unique equilibrium for the infinite system (phase transition is the key word here!) and we shall be concerned with such systems here. In this report it will be pointed by what to replace (*) in that situation.

Since the answer depends on the particular model one considers, let me first describe the problem informally: An observer watches a simulation of our finite process on a computer screen. The simulation runs for a few weeks, and each day he watches the system for a couple of minutes during his lunch break. He compares his observations from different days. What does he see? In other words, we want to follow the system both on a microscopic and on a macroscopic scale. Mathematically speaking, for some appropriate time scale T(N):

$$\mathcal{I}\left((\eta^N_{sT(N)+t})_{t\in R^+} \right) \underset{N\to\infty}{\Rightarrow} \ ? \qquad \text{microscopic picture}$$

$$\mathcal{I}\left((\eta^N_{sT(N)})_{s\in R^+} \right) \underset{N\to\infty}{\Rightarrow} \ ? \qquad \text{macroscopic picture.}$$

3. The recipe.

First we shall introduce the objects needed to describe our system and then we shall give the analogue of (*). Consider a system where the infinite system has several phases (so it is non-ergodic).

Definitions. We associate with the system the following objects:

$$T(N) \ ; \ I \ ; \ [Q_s(.,.),(Y_s)_{s\in R^+}] \ ; \ [\hat{\theta}^N{}_t,\hat{\theta}].$$

Here:

(1) $T(N)$ is the scale function $T: N\to R^+$, $T(N)$ increasing to ∞.

(2) I describes the ergodic components of the system. That is, the infinite system has a set $(\nu_\theta)_{\theta\in I}$ of extremal invariant measures.

(3) $\hat{\theta}: S^{Z^d} \to I$ is the conserved quantity or first integral of motion of the infinite system, and we assume that $\mu(\{\eta|\hat{\theta}(\eta)=\theta\})=1$ implies $\mathcal{I}(\eta^\mu{}_t)\Rightarrow\nu_\theta$ as $t\to\infty$.

(4) $\hat{\theta}^N{}_t = \hat{\theta}^N(\eta^N{}_t)$ is the slowly varying variable of the system, $\hat{\theta}^N: S^{E_N}\to I$, and we assume that $\hat{\theta}^N\Rightarrow\hat{\theta}$ as $N\to\infty$ for any μ.

(5) $(Y_s)_{s\in R^+}$ is the macroscopic observable and its law is $Q_s(.,.)$, i.e. Y_s is a Markov process on I with transition kernel $Q_s(.,.)$.

With these definitions we can now state the analogue of (*).

On a microscopic scale:

$$(**) \qquad \mathcal{I}\left((\eta^N_{sT(N)+t})_{t\in R^+} \right) \underset{N\to\infty}{\Rightarrow} \mathcal{I}\left((\eta^{\nu(s)}{}_t)_{t\in R^+} \right),$$

where $\nu(s) = \int Q_s(\theta',d\theta)\nu_\theta$ with θ' determined by $\mathcal{I}(\eta^\mu{}_t) \Rightarrow \nu_{\theta'}$ as $t\to\infty$.

On a macroscopic scale:

(***) $\mathcal{I}((\hat{\theta}^N_{sT(N)})_{s\in R^+}) \underset{N\to\infty}{\Rightarrow} \mathcal{I}((Y_s)_{s\in R^+})$

$$\mathcal{I}(\eta^N_{sT(N)} \mid \hat{\theta}^N(\eta^N_{tT(N)}) = \theta'_N) \underset{N\to\infty}{\Rightarrow} \int Q_{s-t}(\theta',d\theta)\nu_\theta \text{ if } \theta'_N \to \theta'.$$

In words: On a microscopic scale the system looks like the infinite system in a certain "randomly chosen" equilibrium. This equilibrium is determined by the path of the slowly varying observable, which describes the evolution on a macroscopic scale. Note that this means that we can detect a phase transition for the infinite system by observing the finite system on a macroscopic scale!

Let f be a local function on configuration space and τ_x the shift by x. Let

$$M_N(f)(\eta) = |E_N|^{-1}\Sigma_{x\epsilon E_N} (f.\tau_x)(\eta).$$

Then for any t(N) increasing to ∞ such that t(N)=o(T(N)) we have:

(****) $t^{-1}(N) \int_0^{t(N)} M_N(f)(\eta_{sT(N)+u})du \underset{N\to\infty}{\longrightarrow} \int_I Q_s(\theta',d\theta)\langle\nu_\theta,f\rangle$ in probability,

where $\langle\nu_\theta,f\rangle = \int f \, d\nu_\theta$ and θ' is as above. The results (**) and (***) in most cases imply (****). The latter deals with measurements of global observables.

4. The results.

In section 3 we have outlined a general recipe for describing the large time behaviour of large but finite systems. We believe this recipe should work for very general models, but at present we have been able to verify its correctness only for a few special examples [1]. Some of these will be discussed below. There are two main classes: models with I continuous and models with I discrete.

A. Models with I continuous.

Here we shall be concerned with critical branching and with the voter model. Let us write down the definitions for these processes.

critical branching: state space N^{Z^d}

transition $\eta \to \eta+\delta_x$ at rate $\eta(x)$

$\eta \to \eta-\delta_x$ at rate $\eta(x)$

$\eta \to \eta+\delta_y-\delta_x$ at rate $\eta(x)p(x,y)$, where $p(x,y)$ is an irreducible translation invariant transition kernel on $Z^d x Z^d$ with finite range.

voter model: state space $\{0,1\}^{Z^d}$

transition $\eta \to \eta+\delta_x$ at rate $\Sigma_y p(x,y)\eta(y)$ if $\eta(x)=0$

$\eta \to \eta-\delta_x$ at rate $\Sigma_y p(x,y)(1-\eta(y))$ if $\eta(x)=1$.

As finite systems we consider the restrictions of these to the torus of size N, i.e. E_N with identification of boundary points. It is well known (see Liggett [2]) that both infinite systems for $d\geq 3$ have a set $(\nu_\theta)_{\theta\in I}$ of extremal equilibria, with $I=R^+$, respectively $I=[0,1]$, and that $\int\eta(x)d\nu_\theta=\theta$.

Theorem 1. For critical branching in $d\geq 3$ the relations (**) and (***) hold with:

$T(N) = N^d$, $I = R^+$,

$\hat{\theta}^N = |E_N|^{-1}\Sigma_{x\in E_N}\eta(x)$, $\hat{\theta}^N \Rightarrow \hat{\theta}$,

Y_s, $Q_s(.,.)$: diffusion on R^+ with absorption at 0 and generator $2x(\partial/\partial x)^2$.

Theorem 2. For the voter model in $d\geq 3$ the relations (**) and (***) hold with:

$T(N) = N^d$, $I = [0,1]$,

$\hat{\theta}^N = |E_N|^{-1}\Sigma_{x\in E_N}\eta(x)$, $\hat{\theta}^N \Rightarrow \hat{\theta}$,

Y_s, $Q_s(.,.)$: diffusion on $[0,1]$ with absorption at 0 and 1 and generator $2G^{-1}x(1-x)(\partial/\partial x)^2$, with G = expected # of visits to 0 of the symmetrized random walk with transition kernel $\frac{1}{2}[p(x,y)+p(y,x)]$.

Theorem 3. In $d\leq 2$ the only invariant measure of critical branching is $\delta_{\{\eta=0\}}$ and of the voter model are $\delta_{\{\eta=0\}}$ and $\delta_{\{\eta=1\}}$. We have for every $T(N) \to \infty$ that:

$$\mathcal{L}(\eta^N_{T(N)}) \underset{N\to\infty}{\Rightarrow} \delta_{\{\eta=0\}}\text{, respectively } \theta'\, \delta_{\{\eta=1\}} + (1-\theta')\, \delta_{\{\eta=0\}}\,.$$

Remark: Theorems 1 and 2 tell us that for observation times $\gg N^d$ we will see the ergodicity of the finite sytem. For times $\ll N^d$ the system behaves very much like the infinite system, and for times around N^d we begin to see its finiteness because the associated slowly varying observables begin to fluctuate.

B. A model with $|I|=2$.

Here we consider the contact process in d=1, defined as follows:

state space $\{0,1\}^{Z}$,

transition $\eta \to \eta-\delta_x$ at rate 1 if $\eta(x)=1$

$\eta \to \eta+\delta_x$ at rate $\lambda\ \#\{ z \mid \eta(x+z)=1, |z|=1 \}$ if $\eta(x)=0$.

It is known [2] that there is a critical value $\lambda_c \in (1,2]$ such that for $\lambda > \lambda_c$ the system has two extremal invariant measures ν and $\delta_{\{\eta=0\}}$, while for $\lambda < \lambda_c$ the only invariant measure is $\delta_{\{\eta=0\}}$. Moreover, ν is the limit of the evolution started in $\delta_{\{\eta=1\}}$. The following theorem uses results of Schonmann [3] (and of Durrett and Schonmann, private communication).

Theorem 4. For the contact process on Z with $\lambda > \lambda_c$ the relations (**) and (***) hold with:

T(N) is the expected extinction time of η^N_t, $N^{-1}\log T(N) \to$ constant,

$I = \{0,1\}$, $\hat{\theta}^N(\eta) = 1_{\{\eta(x)=1 \text{ for some } x \in E_N\}}$, $\hat{\theta}(\eta) = 1_{\{\eta(x)=1 \text{ for some } x \in Z^d\}}$,

Y_s, $Q_s(.,.)$: Markov jump process on $\{0,1\}$ with absorption at 0.

For $\lambda < \lambda_c$, on the other hand, we have: $\mathcal{L}(\eta^N_{T(N)}) \Rightarrow \delta_{\{\eta=0\}}$ for all $T(N) \to \infty$.

5. Outlook.

The main open problem at present is to prove (**) and (***) for the stochastic Ising model below critical temperature. We expect $T(N)=e^{\gamma N}$, γ constant, and Y_s a jump process on $\{0,1\}$. The hardest part is to prove an ergodic theorem for the infinite system and to identify the slowly varying variables. It would also be interesting to analyze examples where one has several conserved quantities.

References:

(1) J.T. Cox and A. Greven, The long time behaviour of large but finite particle systems (in preparation).

(2) T. Liggett, Interacting particle systems, Grundlehren der mathematischen Wissenschaften 276 (1985), Springer.

(3) R. Schonmann, Metastability for the contact process, J. Stat. Phys. 41 (1985).

(4) F. Spitzer, Interactions of Markov processes, Adv. Math. 5 (1970).

MC SYLLABI

1.1 F. Göbel, J. van de Lune. *Leergang besliskunde, deel 1: wiskundige basiskennis.* 1965.

1.2 J. Hemelrijk, J. Kriens. *Leergang besliskunde, deel 2: kansberekening.* 1965.

1.3 J. Hemelrijk, J. Kriens. *Leergang besliskunde, deel 3: statistiek.* 1966.

1.4 G. de Leve, W. Molenaar. *Leergang besliskunde, deel 4: Markovketens en wachttijden.* 1966.

1.5 J. Kriens, G. de Leve. *Leergang besliskunde, deel 5: inleiding tot de mathematische besliskunde.* 1966.

1.6a B. Dorhout, J. Kriens. *Leergang besliskunde, deel 6a: wiskundige programmering 1.* 1968.

1.6b B. Dorhout, J. Kriens, J.Th. van Lieshout. *Leergang besliskunde, deel 6b: wiskundige programmering 2.* 1977.

1.7a G. de Leve. *Leergang besliskunde, deel 7a: dynamische programmering 1.* 1968.

1.7b G. de Leve, H.C. Tijms. *Leergang besliskunde, deel 7b: dynamische programmering 2.* 1970.

1.7c G. de Leve, H.C. Tijms. *Leergang besliskunde, deel 7c: dynamische programmering 3.* 1971.

1.8 J. Kriens, F. Göbel, W. Molenaar. *Leergang besliskunde, deel 8: minimaxmethode, netwerkplanning, simulatie.* 1968.

2.1 G.J.R. Förch, P.J. van der Houwen, R.P. van de Riet. *Colloquium stabiliteit van differentieschema's, deel 1.* 1967.

2.2 L. Dekker, T.J. Dekker, P.J. van der Houwen, M.N. Spijker. *Colloquium stabiliteit van differentieschema's, deel 2.* 1968.

3.1 H.A. Lauwerier. *Randwaardeproblemen, deel 1.* 1967.

3.2 H.A. Lauwerier. *Randwaardeproblemen, deel 2.* 1968.

3.3 H.A. Lauwerier. *Randwaardeproblemen, deel 3.* 1968.

4 H.A. Lauwerier. *Representaties van groepen.* 1968.

5 J.H. van Lint, J.J. Seidel, P.C. Baayen. *Colloquium discrete wiskunde.* 1968.

6 K.K. Koksma. *Cursus ALGOL 60.* 1969.

7.1 *Colloquium moderne rekenmachines, deel 1.* 1969.

7.2 *Colloquium moderne rekenmachines, deel 2.* 1969.

8 H. Bavinck, J. Grasman. *Relaxatietrillingen.* 1969.

9.1 T.M.T. Coolen, G.J.R. Förch, E.M. de Jager, H.G.J. Pijls. *Colloquium elliptische differentiaalvergelijkingen, deel 1.* 1970.

9.2 W.P. van den Brink, T.M.T. Coolen, B. Dijkhuis, P.P.N. de Groen, P.J. van der Houwen, E.M. de Jager, N.M. Temme, R.J. de Vogelaere. *Colloquium elliptische differentiaalvergelijkingen, deel 2.* 1970.

10 J. Fabius, W.R. van Zwet. *Grondbegrippen van de waarschijnlijkheidsrekening.* 1970.

11 H. Bart, M.A. Kaashoek, H.G.J. Pijls, W.J. de Schipper, J. de Vries. *Colloquium halfalgebra's en positieve operatoren.* 1971.

12 T.J. Dekker. *Numerieke algebra.* 1971.

13 F.E.J. Kruseman Aretz. *Programmeren voor rekenautomaten; de MC ALGOL 60 vertaler voor de EL X8.* 1971.

14 H. Bavinck, W. Gautschi, G.M. Willems. *Colloquium approximatietheorie.* 1971.

15.1 T.J. Dekker, P.W. Hemker, P.J. van der Houwen. *Colloquium stijve differentiaalvergelijkingen, deel 1.* 1972.

15.2 P.A. Beentjes, K. Dekker, H.C. Hemker, S.P.N. van Kampen, G.M. Willems. *Colloquium stijve differentiaalvergelijkingen, deel 2.* 1973.

15.3 P.A. Beentjes, K. Dekker, P.W. Hemker, M. van Veldhuizen. *Colloquium stijve differentiaalvergelijkingen, deel 3.* 1975.

16.1 L. Geurts. *Cursus programmeren, deel 1: de elementen van het programmeren.* 1973.

16.2 L. Geurts. *Cursus programmeren, deel 2: de programmeertaal ALGOL 60.* 1973.

17.1 P.S. Stobbe. *Lineaire algebra, deel 1.* 1973.

17.2 P.S. Stobbe. *Lineaire algebra, deel 2.* 1973.

17.3 N.M. Temme. *Lineaire algebra, deel 3.* 1976.

18 F. van der Blij, H. Freudenthal, J.J. de Iongh, J.J. Seidel, A. van Wijngaarden. *Een kwart eeuw wiskunde 1946-1971, syllabus van de vakantiecursus 1971.* 1973.

19 A. Hordijk, R. Potharst, J.Th. Runnenburg. *Optimaal stoppen van Markovketens.* 1973.

20 T.M.T. Coolen, P.W. Hemker, P.J. van der Houwen, E. Slagt. *ALGOL 60 procedures voor begin- en randwaardeproblemen.* 1976.

21 J.W. de Bakker (red.). *Colloquium programmacorrectheid.* 1975.

22 R. Helmers, J. Oosterhoff, F.H. Ruymgaart, M.C.A. van Zuylen. *Asymptotische methoden in de toetsingstheorie; toepassingen van naburigheid.* 1976.

23.1 J.W. de Roever (red.). *Colloquium onderwerpen uit de biomathematica, deel 1.* 1976.

23.2 J.W. de Roever (red.). *Colloquium onderwerpen uit de biomathematica, deel 2.* 1977.

24.1 P.J. van der Houwen. *Numerieke integratie van differentiaalvergelijkingen, deel 1: eenstapsmethoden.* 1974.

25 *Colloquium structuur van programmeertalen.* 1976.

26.1 N.M. Temme (ed.). *Nonlinear analysis, volume 1.* 1976.

26.2 N.M. Temme (ed.). *Nonlinear analysis, volume 2.* 1976.

27 M. Bakker, P.W. Hemker, P.J. van der Houwen, S.J. Polak, M. van Veldhuizen. *Colloquium discretiseringsmethoden.* 1976.

28 O. Diekmann, N.M. Temme (eds.). *Nonlinear diffusion problems.* 1976.

29.1 J.C.P. Bus (red.). *Colloquium numerieke programmatuur, deel 1A, deel 1B.* 1976.

29.2 H.J.J. te Riele (red.). *Colloquium numerieke programmatuur, deel 2.* 1977.

30 J. Heering, P. Klint (red.). *Colloquium programmeeromgevingen.* 1983.

31 J.H. van Lint (red.). *Inleiding in de coderingstheorie.* 1976.

32 L. Geurts (red.). *Colloquium bedrijfsystemen.* 1976.

33 P.J. van der Houwen. *Berekening van waterstanden in zeeën en rivieren.* 1977.

34 J. Hemelrijk. *Oriënterende cursus mathematische statistiek.* 1977.

35 P.J.W. ten Hagen (red.). *Colloquium computer graphics.* 1978.

36 J.M. Aarts, J. de Vries. *Colloquium topologische dynamische systemen.* 1977.

37 J.C. van Vliet (red.). *Colloquium capita datastructuren.* 1978.

38.1 T.H. Koornwinder (ed.). *Representations of locally compact groups with applications, part I.* 1979.

38.2 T.H. Koornwinder (ed.). *Representations of locally compact groups with applications, part II.* 1979.

39 O.J. Vrieze, G.L. Wanrooy. *Colloquium stochastische spelen.* 1978.

40 J. van Tiel. *Convexe analyse.* 1979.

41 H.J.J. te Riele (ed.). *Colloquium numerical treatment of integral equations.* 1979.

42 J.C. van Vliet (red.). *Colloquium capita implementatie van programmeertalen.* 1980.

43 A.M. Cohen, H.A. Wilbrink. *Eindige groepen (een inleidende cursus).* 1980.

44 J.G. Verwer (ed.). *Colloquium numerical solution of partial differential equations.* 1980.

45 P. Klint (red.). *Colloquium hogere programmeertalen en computerarchitectuur.* 1980.

46.1 P.M.G. Apers (red.). *Colloquium databankorganisatie, deel 1.* 1981.

46.2 P.G.M. Apers (red.). *Colloquium databankorganisatie, deel 2.* 1981.

47.1 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60: general information and indices.* 1981.

47.2 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 1: elementary procedures; vol. 2: algebraic evaluations.* 1981.

47.3 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 3A: linear algebra, part I.* 1981.

47.4 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 3B: linear algebra, part II.* 1981.

47.5 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 4: analytical evaluations; vol. 5A: analytical problems, part I.* 1981.

47.6 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 5B: analytical problems, part II.* 1981.

47.7 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 6: special functions and constants; vol. 7: interpolation and approximation.* 1981.

48.1 P.M.B. Vitányi, J. van Leeuwen, P. van Emde Boas (red.). *Colloquium complexiteit en algoritmen, deel 1.* 1982.

48.2 P.M.B. Vitányi, J. van Leeuwen, P. van Emde Boas (red.). *Colloquium complexiteit en algoritmen, deel 2.* 1982.

49 T.H. Koornwinder (ed.). *The structure of real semisimple Lie groups.* 1982.

50 H. Nijmeijer. *Inleiding systeemtheorie.* 1982.

51 P.J. Hoogendoorn (red.). *Cursus cryptografie.* 1983.

CWI SYLLABI

1 Vacantiecursus 1984: *Hewet - plus wiskunde.* 1984.

2 E.M. de Jager, H.G.J. Pijls (eds.). *Proceedings Seminar 1981-1982. Mathematical structures in field theories.* 1984.

3 W.C.M. Kallenberg, et al. *Testing statistical hypotheses: worked solutions.* 1984.

4 J.G. Verwer (ed.). *Colloquium topics in applied numerical analysis, volume 1.* 1984.

5 J.G. Verwer (ed.). *Colloquium topics in applied numerical analysis, volume 2.* 1984.

6 P.J.M. Bongaarts, J.N. Buur, E.A. de Kerf, R. Martini, H.G.J. Pijls, J.W. de Roever. *Proceedings Seminar 1982-1983. Mathematical structures in field theories.* 1985.

7 Vacantiecursus 1985: *Variatierekening.* 1985.

8 G.M. Tuynman. *Proceedings Seminar 1983-1985. Mathematical structures in field theories, Vol.1 Geometric quantization.* 1985.

9 J. van Leeuwen, J.K. Lenstra (eds.). *Parallel computers and computations.* 1985.

10 Vacantiecursus 1986: *Matrices.* 1986.

11 P.W.H. Lemmens. *Discrete wiskunde: tellen, grafen, spelen en codes.* 1986.

12 J. van de Lune. *An introduction to Tauberian theory: from Tauber to Wiener.* 1986.

13 G.M. Tuynman, M.J. Bergvelt, A.P.E. ten Kroode. *Proceedings Seminar 1983-1985. Mathematical structures in field theories, Vol.2.* 1987.

14 Vacantiecursus 1987: *De personal computer en de wiskunde op school.* 1987.

15 Vacantiecursus 1983: *Complexe getallen.* 1987.

16 P.J.M. Bongaarts, E.A. de Kerf, P.H.M. Kersten. *Proceedings Seminar 1984-1986. Mathematical structures in field theories, Vol.1.* 1988.

17 F. den Hollander, H. Maassen (eds.). *Mark Kac seminar on probability and physics Syllabus 1985-1987.* 1988.