

## A Central Limit Theorem for the Disordered Harmonic Chain

A. J. O'Connor

Faculty of Mathematics, Open University, Walton Hall, Milton Keynes MK7 6AA, England

**Abstract.** Using methods introduced by Furstenberg and Tutubalin we prove a central limit theorem for the amplitudes of plane waves travelling in a semi-infinite isotopically disordered harmonic chain. This theorem is applied to the problem of heat conduction in disordered harmonic chains.

Fourier's law of heat conduction is that the steady state heat current is proportional to the applied temperature gradient. It is observed to be true over a wide range of temperatures in solids of any degree of purity. The ratio of the heat current to the temperature gradient is the thermal conductivity, and at a fixed temperature this is observed to depend only on the properties of the material. The most convincing explanation of a finite conductivity has been given by Peierls [21]. He showed that anharmonic forces play an essential role in establishing an equilibrium distribution of the energy between the various normal modes of the solid. Assuming the validity of the phonon gas picture he showed that this tendency to equilibrium produces a finite thermal conductivity even for pure crystalline solids. A rigorous formulation of these ideas is certainly the only really satisfactory solution of the problem. It is so difficult however that we look for other soluble models which contain a reflective mechanism and in particular to disordered harmonic systems.

The model we use was first studied by Lebowitz [6, 7]. We take a chain of  $N$  particles coupled to their nearest neighbours and connected at each end to heat baths. The heat baths are modelled by white noise whose covariance is proportional to the temperature of the bath and a Langevin damping term to represent the ability of a heat bath to absorb energy. If  $x_i(t)$  and  $m_i$  are the displacement from its equilibrium position and the mass of the  $i$ th particle the motion of the particles is determined by the equations

$$\begin{aligned}
 m_1 \ddot{x}_1 + 2x_1 - x_2 + \lambda m_1 \dot{x}_1 &= f_1(t) \\
 m_j \ddot{x}_j + 2x_j - x_{j-1} - x_{j+1} &= 0 \quad 1 < j < N \\
 m_N \ddot{x}_N + 2x_N - x_{N-1} + \lambda m_N \dot{x}_N &= f_N(t)
 \end{aligned} \tag{1}$$

where  $f_a$  ( $a=1, N$ ) is a Gaussian random process with variance  $2T_a \lambda m_a$ . Casher and Lebowitz show that in the steady state situation there is a steady flow of

energy along the chain and that it is given by

$$J_N(m_1 \dots m_N) = 2(T_N - T_1) \lambda^2 m_1 m_N \pi^{-1} \int_0^\infty w^2 j_N(w)^{-1} dw. \quad (2)$$

$$j_N(w) = |z_N|^2$$

$$\text{where } z_N = (1, -i\lambda m_N w) T_N \dots T_1 \begin{pmatrix} 1 \\ i\lambda m_1 w \end{pmatrix}$$

$$\text{and } T_i = \begin{pmatrix} 2 - m_i w^2 & -1 \\ 1 & 0 \end{pmatrix} = T(m_i). \quad (3)$$

The occurrence of the matrices  $T_i$  can be understood by looking at the equations of motion for an isolated chain obtained from (1) by putting  $\lambda=0$  and  $f_1 = f_N = 0$ . The plane wave solutions to these equations are solutions of the type  $x_k(t) = U_k(w) e^{iwt}$ . The equations for  $U_i$  derived in this way can be recast in the transfer matrix formalism [7]

$$\begin{bmatrix} U_{n+1}(w) \\ U_n(w) \end{bmatrix} = T_n \begin{bmatrix} U_n(w) \\ U_{n-1}(w) \end{bmatrix} = T_n \dots T_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

This suggests that we interpret (2) as summing the contributions to  $J_N$  of each frequency which is excited by the Gaussian driving terms. These individual contributions reflect the ability of the chain to carry modes of various frequencies.

Periodic chains ( $m_i = m_{i+p}$ , fixed) are unphysical because their perfect transparency to frequencies in the correct bands enables them to carry a finite heat current even as  $N$  increases. One model in which the heat current decreases with length is the disordered chain. In this the masses,  $m_1 \dots m_N$  are independent, identically distributed random variables. In this case we use the average heat current carried by all possible chains viz.

$$\langle J(m_1 \dots m_N) \rangle = J_N$$

and we want to decide if the limiting conductivity

$$\lim_{N \rightarrow \infty} N J_N (\Delta T)^{-1} \quad \text{is finite and non-zero.}$$

There is a great deal of work on this problem and although no rigorous solution exists it seems that  $N J_N$  falls off as  $N^{-\frac{1}{2}}$  [8]. The present work was motivated by an attempt to prove this rigorously. We cannot do this and have only been able to prove that contribution to  $J_N$  from frequencies above any non zero frequency  $\varepsilon$  becomes exponentially small as  $N$  increases. The real difficulty in this problem lies in treating the frequencies of order  $N^{-\frac{1}{2}}$  and smaller – we have not been able to handle these and non-trivial problems arise in this region.

From (2) we see that we must understand the asymptotic behaviour of products of random matrices in  $SL(2, \mathbb{R})$ . There is already considerable body of work on the general problem of products of random matrices. To describe this work we introduce a very useful factorization of the matrix  $T_n \dots T_1$  [this is in  $SL(2, \mathbb{R})$ ].

$$Q_n = T_n \dots T_1 = \begin{pmatrix} \sin \theta_n & -\cos \theta_n \\ \cos \theta_n & \sin \theta_n \end{pmatrix} \begin{pmatrix} t_n & r_n \\ 0 & t_n^{-1} \end{pmatrix} \quad (4)$$

with  $-\pi/2 \leq \theta_n \leq \pi/2$ ,  $t_n \in \mathbb{R}$ .  $\theta_n$  is called the phase angle at the  $n$ th mass.

Furstenberg [1] studied the general problem of random products in  $SL(m, \mathbb{R})$  and proved, among many other results, a strong law of large numbers which implies here that  $n^{-1} \log t_n^2 \rightarrow \gamma > 0$ , with probability one. His methods have a strong geometric flavour. Tutubalin [2] used more direct analytical methods to obtain a central limit theorem for the analogues of  $n^{-\frac{1}{2}}(\log t_n^2 - n\gamma)$ . However he works with the probability distribution on  $SL(m, \mathbb{R})$  which is absolutely continuous with respect to Haar measure on  $SL(m, \mathbb{R})$ . In our case this is not so; the right measure on  $SL(2, \mathbb{R})$  for our problem is concentrated on the matrices  $T(m)$ .

Although Tutubalin's result is not automatically valid here we can use his method and some ideas due to Furstenberg to obtain a central limit theorem for the random variables  $n^{-\frac{1}{2}}(\log t_n^2 - n\gamma)$ . Matsuda and Ishii were the first to realize the relevance of Furstenberg's work to this problem. Their work, and a simplification due to Yoshioka, is needed to show that  $\gamma$  is strictly positive. A comprehensive survey of this problem and other questions in disordered linear systems is to be found in the article by Ishii [9]. It is quite probable that these methods are also applicable to the disordered Kronig-Penney model and the tight binding electron model [9, 16].

In the first section we discuss the asymptotic behaviour of  $\theta_n$ . The ideas here can be traced back to Schmidt [13] and there are also strong connections with the work of Borland [14] and Halperin [15]. In the second section we prove a central limit theorem for the random variables  $n^{-\frac{1}{2}}(\log t_n^2 - n\gamma)$ . Finally we apply these results to the problem of heat conduction in disordered chains and obtain our main result, Theorem 6. This is that the contribution to  $J_N$  from frequencies above  $\varepsilon$  falls off exponentially with  $N$  for each fixed  $\varepsilon > 0$ . This applies to random chains in which the common distribution of the masses has a  $C^1$  density with compact support. Some suggestions are made for further work on this problem and in an appendix some standard results on fractional linear transformations of  $C^1$  are summarized.

### 1. Asymptotic Behaviour of the Phase Angle

We consider a semi-infinite chain of masses,  $m_j, j \geq 1$ . The masses are independent, identically distributed random variables with a distribution  $dp$ . A transfer matrix  $T_j$  is associated with the  $j$ th mass using (2) and  $Q_n$  is the product of the first  $n$  transfer matrices (this is at a fixed frequency  $w$ ). Using the equation  $Q_{n+1} = T_{n+1}Q_n$  we can obtain recurrence relations for the parameters  $\theta_n$  and  $t_n$  introduced in (4).

$$\tan \theta_{n+1} = \tan \Phi(\theta_n, m_{n+1}) \tag{5}$$

$$t_{n+1}^2 = t_n^2 [\sin^2 \theta_n + (x_{n+1} \sin \theta_n - \cos \theta_n)^2]$$

$$\text{where } \tan \Phi(\theta, m) = 2 - mw^2 - \cot \theta \tag{6}$$

and  $x_n = 2 - m_n w^2$ . The initial values are  $\theta_0 = \pi/2$  and  $t_0 = 1$ .

Since  $m_1 \dots m_n$  are random variables,  $\theta_n$  is a random variable lying in  $I = [-\pi/2, \pi/2]$ . If its distribution is  $P_n$  we have

$$P_n(A) = \int dp(m) P_{n-1}[\Psi(A, m)] \tag{7}$$

for each open set  $A$  in  $I$ . Here  $\Psi(\cdot, m)$  is the inverse function to  $\Phi(\cdot, m)$ . It satisfies

$$\tan\theta = 2 - nw^2 - \cot\Psi(\theta, m). \quad (8)$$

We introduce the space  $P$  of continuous functions  $f$  on  $I$  which are periodic with period  $\pi$  so that  $f(\pi/2) = f(-\pi/2)$ .  $P$  is a Banach space in the sup.norm. We define an operator  $T$  on  $P$  by

$$Tf(\theta) = \int dp(m) f[\Phi(\theta, m)]. \quad (9)$$

The dual space  $P^*$  is the space of Borel measures  $\mu$  on  $I$  for which  $\mu\{\pi/2\} = \mu\{-\pi/2\}$ . If  $T^*$  is the adjoint of  $T$ , (7) can be written as

$$P_n = T^*P_{n-1}.$$

For every value of  $m$ , we have  $\Phi(\pm 0, m) = \mp \pi/2$  and for each non-zero value of  $\theta$  there is a spread of values for  $\Phi(\theta, m)$ . Thus  $Tf(\theta)$  is a weighted average for  $\theta \neq 0$  and  $T^2f(\theta)$  is a weighted average for all values of  $\theta$ . In fact when we assume that  $dp$  has an integrable density,  $T^2$  is a compact operator on  $P$  and we can prove that the measures  $P_n$  converge weakly to a limiting distribution  $\nu$ .

**Theorem 1.** *We assume that  $dp(m) = \mu(m)dm$  with  $\mu \in L^1(\mathbb{R}^+)$ . Then*

(i)  *$T^2$  is a compact operator on  $P$  whose spectrum lies in the unit disc. The only eigenvalue on the unit circle is at 1 and it is simple.*

(ii) *There is a unique probability measure  $\nu$  so that*

$$\|T^{*n}\sigma - \nu\| \leq c \cdot \rho^n$$

*for any probability measure  $\sigma$  on  $I$ . Here  $\rho$  is the radius of a disc containing all the spectrum of  $T$  except 1 and  $c$  depends only on  $T$ . The norm is the total variation norm.*

(iii)  *$\nu$  is absolutely continuous with respect to Lebesgue measure on  $I$ . If  $\mu$  is a continuous function with compact support and  $d\nu = Fd\theta$ , then  $F$  is a continuous non-negative function.*

*Proof.* We begin by assuming that  $\mu$  is continuous. For each  $f$  in  $P$ ,  $Tf(\theta)$  is an average of  $f$  over a non-empty open interval when  $\theta$  is not zero. In fact  $T$  maps bounded sets of functions which are uniformly supported away from zero into compact sets in  $P$ . To prove this rewrite  $T$  as

$$Tf(\theta) = \int \mu[\Gamma(\alpha, \theta)] (\partial\Gamma/\partial\alpha) f(\alpha) d\alpha \quad (10)$$

where

$$m = \Gamma(\alpha, \theta) \quad \text{if} \quad \alpha = \Phi(\theta, m)$$

i.e.  $2 - mw^2 = \tan\alpha + \cot\theta$ .

Then

$$Tf(\theta) = \int_{\Phi(\theta, m_1)}^{\Phi(\theta, m_2)} \mu[\Gamma(\alpha, \theta)] w^{-2} \sec^2\alpha f(\alpha) d\alpha \quad (11)$$

where

$$\text{supp}\mu = [m_1, m_2].$$

For each positive  $\varepsilon$  we can find a  $\delta$  so that  $|\theta| < \delta$ ,

$$\pi/2 - \varepsilon > \Phi(\theta, m) > -\pi/2 + \varepsilon$$

for all  $m \in \text{supp} dp$ . We define

$$P_\delta = \{f \in P: f(\theta) = 0 \text{ if } |\theta| < \delta\}.$$

If  $\theta$  and  $\phi$  lie outside the interval  $(-\delta, \delta)$  and  $f \in P_\delta$  have

$$\begin{aligned} Tf(\theta) - Tf(\phi) &= \int_{\Phi(\theta, m_1)}^{\Phi(\theta, m_2)} \mu[\Gamma(\alpha, \theta)] w^{-2} \sec^2 \alpha f(\alpha) d\alpha \\ &\quad - \int_{\Phi(\phi, m_1)}^{\Phi(\phi, m_2)} \mu[\Gamma(\alpha, \phi)] w^{-2} \sec^2 \alpha f(\alpha) d\alpha \\ &= \int_{\Phi(\phi, m_2)}^{\Phi(\theta, m_2)} \mu[\Gamma(\alpha, \theta)] w^{-2} \sec^2 \alpha f(\alpha) d\alpha \\ &\quad + \int_{\Phi(\theta, m_1)}^{\Phi(\phi, m_2)} \{\mu[\Gamma(\alpha, \theta)] - \mu[\Gamma(\alpha, \phi)]\} w^{-2} \sec^2 \alpha f(\alpha) d\alpha \\ &\quad - \int_{\Phi(\phi, m_1)}^{\Phi(\theta, m_1)} \mu[\Gamma(\alpha, \phi)] w^{-2} \sec^2 \alpha f(\alpha) d\alpha. \end{aligned} \quad (12)$$

Because  $\phi$  is bounded away from  $\pm \pi/2$ ,  $\sec^2 \alpha < \sec^2 \varepsilon$  in these integrals so that

$$\begin{aligned} |Tf(\theta) - Tf(\phi)| &\leq w^{-2} \sec^2 \varepsilon \|f\| \{ \|\mu\| |\Phi(\theta, m_1) - \Phi(\phi, m_1)| + \\ &\quad \|\mu\| |\Phi(\theta, m_2) - \Phi(\phi, m_2)| + \pi \sup_{\alpha} |\mu[\Gamma(\alpha, \phi)] - \mu[\Gamma(\alpha, \theta)]| \}. \end{aligned} \quad (13)$$

So  $TP_\delta$  is an equicontinuous set of functions in  $P$ . In general  $f$  will not vanish near the origin. To handle these functions choose two functions  $g_1$  and  $g_2$  in  $P$  so that

$$\begin{aligned} g_1 + g_2 &= 1 \\ g_2 &= 0 \quad \text{when } |\theta| > 2\delta \\ &= 1 \quad \text{when } |\theta| < \delta. \end{aligned}$$

Then for any  $f$  in  $P$ ,  $f g_1$  and  $T(f g_2)$  are in  $P_\delta$  so that  $T^2 P \subseteq TP_\delta$ . So  $T^2$  is a compact operator on  $P$ .

If  $\mu \in L^2(\mathbb{R}^+)$ , then we can approximate it arbitrarily closely in  $L^1$  by a continuous function and so the corresponding  $T^2$  can be approximated in norm by compact operators. Thus  $T^2$  is compact if  $\mu \in L^1$ .

Suppose that  $T^2 f = e^{i\theta} f$  and that  $\max |f| = |f(\alpha)|$ . Then

$$|f(\alpha)| = \left| \int dp(m_1) dp(m_2) f[\Phi_{m_2} \Phi_{m_1}(\alpha)] \right|.$$

This can only be true if

$$f[\Phi_{m_2} \Phi_{m_1}(\alpha)] = e^{i\theta} |f(\alpha)| \quad \text{for all } m_1, m_2 \in \text{supp}(dp).$$

Repeating this argument implies that  $f = e^{in\theta} |f(\alpha)|$  on the set

$$S_n = \{\Phi_{m_r} \dots \Phi_{m_1}(\alpha): r = 2n, m_1 \dots m_r \in \text{supp} dp\}.$$

To study this set it is helpful to view the transformation  $\Phi(\cdot, m) = \Phi_m(\cdot)$  in another way. In terms of the variable  $z = \tan \theta$ ,  $\Phi_m$  yields a fractional linear transformation  $T(m)$  defined by

$$z' = T(m)z = 2 - mw^2 - \cot \theta = 2 - mw^2 - z^{-1}. \quad (14)$$

So products of the maps  $\Phi_m$  yield products of the fractional transformations  $T(m)$  (information on these is in the appendix).

If at least one of the transformations  $T(m)$  is hyperbolic the sequence  $\{\Phi_m^r(\alpha)\}$  converges to the attractive fixed point  $\beta$  of  $T(m)$  and so  $e^{ir\theta}|f(\alpha)| \rightarrow f(\beta)$  as  $r \rightarrow \infty$ . This is only true if  $\theta = 0$ . In this case  $f$  is real and we can apply a similar argument to the sets on which  $f = \max f$  and  $f = \min f$  and find that  $f$  is a constant. So in this case the only eigenvalue on this unit circle is simple and is at 1.

If on the other hand, all the transformations  $T(m)$  are elliptic we have for each  $m$

$$f[\Phi_m^{2r}(\alpha)] = e^{ir\theta}|f(\alpha)|.$$

The set  $S = \{\Phi_m^r(\alpha) : r \text{ even}\}$  is dense in  $I$  and so  $|f(\phi)| = |f(\alpha)|$  for all  $\phi$  in  $I$ . However  $T^2 f$  is a weighted average of  $f$  and the weighted average of vectors on a circle will only lie on the same circle if all the original vectors are parallel. Hence  $f$  is again constant on  $I$  and  $\theta = 0$ .

(ii)  $T^2$  is a positive operator on  $P$  and so we can apply Theorem 6, page 714 from Dunford and Schwartz [4] to obtain

$$\|T^n f - \langle v, f \rangle e\| \leq c \cdot \rho^n$$

$e(\theta) = 1$ , for all  $\theta$  in  $I$ .

$c$  is a constant, depending on  $T$  but not on  $f$ . So

$$\|T^{*n}\sigma - v\| = \sup_{f \in P} \langle T^{*n}\sigma - v, f \rangle \leq c\rho^n.$$

(iii) For any  $\alpha$  in  $I$  it is not hard to check that  $T^{*2}\delta_\alpha$  has a continuous density (*wod*) and so all the measures  $T^{*n}\delta_\alpha$  also have continuous densities. Convergence of the measures in the total variation norm is equivalent to convergence of the densities in  $L^1$  and so  $v$  is absolutely continuous. ( $dv = Fd\theta$ ). When  $\mu$  is continuous,  $F$  is also continuous. If one of the maps  $T(m)$  is elliptic (i.e.  $4 > m\omega^2 > 0$ ) a direct argument shows that  $F$  does not vanish on  $I$ . If all the maps  $T(m)$  are hyperbolic then  $F$  will be supported on the interval covered by the attractive fixed points of  $\{T(m) : m \in \text{supp}(dp)\}$ .  $F$  will then vanish outside this interval.

*Remarks.* (i)  $v$  is called the stationary measure for  $T^*$  since  $T^*v = v$ .

(ii) It is not clear if a unique  $v$  exists for more general mass distributions. The rather pathological behaviour of the examples given by Schmidt suggests that it need not be unique when just two masses can occur.

## 2. Asymptotic Behaviour of $\log_n^2$

We imitate the method of Tutubalin [2] and Nagaev [3] and compute the characteristic function of the random variable  $\log_n^2$ . From 4,

$$\log_{n+1}^2 = \log_n^2 + a(m_{n+1}, \theta_n) \tag{15}$$

$$\begin{aligned} a(m, \theta) &= \log \sin^2 \theta - \log \cos^2 \Phi(\theta, m) \\ &= \log[\sin^2 \theta + (x \sin \theta - \cos \theta)^2]. \end{aligned} \tag{16}$$

The asymptotic distribution of  $\theta_n$  is  $\nu$  and so the mean value of  $1/n \log t_n^2$  converges to  $\gamma = E_{p \times \nu} a(m, \theta)$ . We define new random variables by

$$X_n = \log t_n^2 - n\gamma. \tag{17}$$

Then

$$X_{n+1} = X_n + A(m_{n+1}, \theta_n) \tag{18}$$

$$A(m, \theta) = a(m, \theta) - \gamma. \tag{19}$$

If  $\Omega$  is the underlying probability space on when each  $m_j$  is defined,  $X_{n+1}$  is a random variable defined on  $\Omega^{n+1}$  and its characteristic function is

$$\begin{aligned} \phi_{n+1}(\lambda) &= E_{\Omega^{n+1}} e^{i\lambda X_{n+1}} \\ &\equiv E_{\Omega^n} E_p e^{i\lambda X_n} e^{i\lambda A(m_{n+1}, \theta_n)} \\ &\equiv E_{\Omega^n} \int dp(m) e^{i\lambda A(m, \theta_n)} e^{i\lambda X_n} (\theta_n, m_1 \dots m_n) \\ &\equiv E_{\Omega^n} [e^{i\lambda X_n} A(\lambda) e(\theta_n)] \end{aligned} \tag{20}$$

where

$$A(\lambda) f(\theta) \equiv \int dp(m) e^{i\lambda A(m, \theta)} f[\Phi(\theta, m)]. \tag{21}$$

So

$$\phi_n(\lambda) \equiv A(\lambda)^n e(\pi/2) \tag{22}$$

since

$$X_0 = 0 \quad \text{and} \quad \theta_0 = \pi/2.$$

The characteristic function of the normalized random variable  $n^{-\frac{1}{2}} X_n$  is  $\phi_n(\lambda n^{-\frac{1}{2}})$ .

**Theorem 2.** We assume that  $dp(m) = \mu(m)dm$  and that  $\mu$  is a continuous function with compact support.

(i)  $A(\lambda)$  is an analytic family of bounded operators on  $P$ .

(ii)  $A(\lambda)^2$  is a compact operator on  $P$  and when  $\lambda$  is small its dominant eigenvalue  $\theta(\lambda)$  is simple.

(iii) When  $\lambda$  is small

$$\theta(\lambda) = 1 - \frac{1}{2} \sigma^2 \lambda^2 + O(\lambda^3).$$

$$(iv) \lim_{n \rightarrow \infty} \phi_n(\lambda n^{-\frac{1}{2}}) = e^{-\frac{1}{2} \sigma^2 \lambda^2}.$$

*Proof.* (i) To show analyticity it is enough to check that if  $c_n = \sup_{\theta} \int dp(m) |a(m, \theta)|^n$  then  $c_n^{1/n}$  is bounded. In fact

$$a(m, \theta) = \log[1 + x^2/2 - (x^2/2) \cos 2\theta - x \sin 2\theta]$$

with  $x = 2 - mw^2$  and so

$$|a(m, \theta)| \leq \log(1 + m^2) + C.$$

$C$  depends only on  $w, m_1, m_2$  and so if  $m_1 \leq m_2$

$$C_n^{1/n} \leq \log(1 + m_2^2) + C.$$

(ii) We can show that  $A(\lambda)^2$  is compact by imitating Theorem 1. ( $A(0)=T$ ). Since the eigenvalues of  $A(0)^2$ , other than 1, lie inside the circle  $|z| \leq \varrho(w) < 1$  then for small  $\lambda$  the eigenvalues of  $A(\lambda)^2$  other than  $\theta(\lambda)$  are inside the circle  $|z| \leq (1 + \varrho)/2 = \varrho_1$ . There is only one eigenvalue outside this circle and if  $e(\lambda)$  and  $v(\lambda)$  are the eigenvectors of  $A(\lambda)^2$  and  $A(\lambda)^{*2}$  associated with  $\theta(\lambda)$  and  $\theta(\lambda)^*$  we can expand them in a Taylor series around  $\lambda=0$ ,

$$\begin{aligned} e(\lambda) &= e + O(\lambda) \\ v(\lambda) &= v + O(\lambda) \\ \theta(\lambda) &= 1 + i\theta'(0)\lambda - \frac{1}{2}\theta''(0)\lambda^2 + O(\lambda^3). \end{aligned} \tag{23}$$

The Feynman-Hellmann theorem [20, p. 125] gives

$$\begin{aligned} \theta'(0) &= (d/d\lambda)\langle v(\lambda), A(\lambda)e(\lambda) \rangle|_{\lambda=0} \\ &= \langle v, (dA/d\lambda)|_{\lambda=0}e \rangle = E_{p \times v} [a(m, \theta) - \gamma] = 0. \end{aligned}$$

We put  $\sigma^2 = \theta''(0)$ . (iv) follows from

$$\|A(\lambda n^{-\frac{1}{2}})^n e - \theta(\lambda n^{-\frac{1}{2}})^n e(\lambda n^{-\frac{1}{2}})\| \leq C \cdot \varrho_1^n$$

so that  $\lim_{n \rightarrow \infty} \phi_n(\lambda n^{-\frac{1}{2}}) = e^{-\frac{1}{2}\sigma^2\lambda^2}$ .

The next theorem imitates an argument of Tutubalin [2] and shows that  $\sigma^2 > 0$ , i.e., the limiting distribution of the random variable  $n^{-\frac{1}{2}}X_n$  is a non-degenerate Gaussian.

**Theorem 3.** *We assume that  $\mu$  is  $C^1$  and has compact support. Then  $\sigma^2(w)$  is strictly positive when  $w \neq 0$ .*

*Proof.* The spectral integral representation for functions of a compact operator is especially useful in this proof. We write  $A(\lambda) = \theta(\lambda)P(\lambda) + B(\lambda)$  with  $P(\lambda)$  being the projection  $e(\lambda) \otimes v(\lambda)$  associated with eigenvalue  $\theta(\lambda)$ . The spectrum of  $B(\lambda)$  lies in a disc  $|z| \leq \varrho(w) < 1$  when  $\lambda$  is sufficiently small. Then we have

$$2\pi i \phi_n(\lambda) = \int_Y z^{n-1} [z - B(\lambda)]^{-1} e(\pi/2) dz + 2\pi i \theta(\lambda)^n \langle e, v(\lambda) \rangle e(\pi/2, \lambda)$$

[ $Y$  is a closed contour lying between the circles  $|z|=1$  and  $|z|=\varrho(w)$ ]. The first term is uniformly bounded for all small  $\lambda$ . Then

$$\begin{aligned} -2\pi E(X_n) &= 2\pi i n \theta'(0) + 2\pi i e'(0)(\pi/2) \\ &+ \int_Y z^{n-1} [z - B(0)]^{-1} (dB/d\lambda)|_{\lambda=0} [z - B(0)]^{-1} e \cdot dz \\ &= \text{bounded terms.} \end{aligned} \tag{24}$$

Similarly  $Var X_n^2 = n\sigma^2 + \text{bounded terms}$  and so if  $\sigma^2 = 0$  we can choose a constant  $C$  so that for all  $n$

$$Pr(|X_n - EX_n| > c) < 1/2. \tag{25}$$

We will obtain a contradiction by showing that for large  $n$

$$Pr(|X_n - EX_n| \leq c) < 1/2, \tag{26}$$

$$\begin{aligned} E(e^{-\frac{1}{2}X_n^2}) &= \int e^{-\frac{1}{2}u^2} Pr(X_n \in du) \\ &= (2\pi)^{-\frac{1}{2}} \int \int e^{iu\lambda} e^{-\frac{1}{2}\lambda^2} d\lambda Pr(X_n \in du) \\ &= (2\pi)^{-\frac{1}{2}} \int \phi_n(\lambda) e^{-\frac{1}{2}\lambda^2} d\lambda. \end{aligned}$$

When  $\lambda \neq 0$ ,  $A(\lambda)^2$  has no eigenvalues on the unit circle and so  $\phi_n(\lambda) \rightarrow 0$  as  $n \rightarrow \infty$ . We can split the region of integration into three regions.

(i)  $|\lambda| \leq \delta$ ,  $|\phi_n(\lambda)| \leq 1$  so the integral here is less than  $2\delta$ .

(ii)  $\delta \leq |\lambda| \leq \Delta$ , the dominated convergence theorem combined with  $\phi_n \rightarrow 0$  shows that this integral goes to zero as  $n$  increases.

(iii)  $|\lambda| \geq \Delta$ , the integral is bounded by a multiple of  $e^{-\frac{1}{2}\Delta^2}$ .

So as  $n \rightarrow \infty$ ,  $E(e^{-\frac{1}{2}X_n^2}) \rightarrow 0$ , and  $Pr(|X_n| \leq C) \leq E(e^{-\frac{1}{2}X_n^2}) e^{\frac{1}{2}C^2} \rightarrow 0$  as  $n \rightarrow \infty$ . So for  $n$  large enough (26) is true and the contradiction forces  $\sigma^2$  to be strictly positive.

**Corollary to Theorem 2.** *Asymptotically  $Z_n$  and  $\theta_n$  are independent*

$$\lim_{n \rightarrow \infty} E[f(Z_n)g(\theta_n)] = \langle n, f \rangle \langle v, g \rangle$$

for any  $f \in L^2(\mathbb{R})$ ,  $g \in P$  where  $n$  is the density of Gaussian random variable with zero mean and unit variance.

*Proof.*

$$\begin{aligned} &E(e^{i\lambda Z_{n+1}} g(\theta_{n+1})) \\ &= E \int dp(m) e^{i\lambda Z_n} e^{i\lambda A(m, \theta_n)} g[\Phi(\theta_n, m)] \\ &= E \int_{\Omega^n} [e^{i\lambda Z_n} A(\lambda n^{-\frac{1}{2}}) g(\theta_n)] \\ &= A(\lambda n^{-\frac{1}{2}})^n g(\pi/2) \langle v, g \rangle \langle n, e^{i\lambda u} \rangle. \end{aligned}$$

The next theorems show that  $\gamma$  is strictly positive when  $w \neq 0$ . The first is due to Furstenberg and is just Lemma 8.9 in his paper [1].

**Theorem 4.**  $\gamma = E_{p \times v} a(m, \theta) > 0$  when  $w > 0$ .

*Proof.* We are still assuming that  $dp = \mu dm$  with  $\mu$  continuous. Hence  $dv = F d\theta$  and  $F$  is non-negative. Using (16) we can write

$$a(\theta, m) = -\log(d\Phi(\theta, m)/d\theta)$$

We define a measure  $v_m$  by  $v_m(A) = v[\Phi(A, m)]$ . Provided that  $F(\theta) > 0$  we have

$$dv_m/dv = (F[\Phi(\theta, m)]/F(\theta)) \cdot (d\Phi(\theta, m)/d\theta).$$

Since  $F$  might vanish it is possible that both sides of this equation are identically  $+\infty$  on an open interval. In this case we can replace  $v(d\theta)$  by  $v_b(d\theta) = (F(\theta) + b)d\theta$  with  $b > 0$ . Using  $v_b$  if need be instead of  $v$  the argument below applies and finally,

using the boundedness of  $F \log F$  the same result applies. With this proviso we obtain

$$\begin{aligned}
 -\gamma &= \int dv(\theta)dp(m) \log(dv_m/dv) \\
 &- \int dv(\theta)dp(m) \log(F[\Phi(\theta, m)]/F(\theta)).
 \end{aligned}$$

Using the stationarity of  $\nu$  i.e.

$$\nu(A) = \int dp(m)\nu[\Psi(A, m)]$$

means that the second term vanishes and so

$$-\gamma = \int dv(\theta)dp(m) \log(dv_m/dv).$$

Using Jensen's inequality gives

$$-\gamma \leq \log \int dp(m)dv_m(\theta) = 0 \tag{29}$$

and so  $\gamma > 0$  unless  $dv_m/dv = 1$  for all  $m \in \text{supp} \mu$ . i.e. unless  $\nu$  is invariant under all the transformations  $\Phi_m$  on  $I$ . [We will say that a measure  $\nu$  on  $I$  is invariant under a transformation  $T: I \rightarrow I$  if  $\nu(A) = \nu(TA)$  for each measurable set  $A$  in  $I$ .]

It is helpful to view this more generally. Any matrix  $g$  in  $M(2, \mathbb{R})$ ,  $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ , maps the line  $\{\lambda n: \lambda > 0\}$  into another line  $\{\lambda gn: \lambda > 0\}$  and so induces a transformation of the circle  $S^1$  into itself given by

$$\tilde{g}: \theta \rightarrow \tilde{\theta}$$

when

$$n = (\sin\theta, \cos\theta), gn = k(\sin\tilde{\theta}, \cos\tilde{\theta})$$

i.e.

$$\tan\tilde{\theta} = \frac{a \tan\theta + b}{c \tan\theta + d} \tag{30}$$

i.e.  $\tan\theta$  transforms under the fractional linear transformation associated with  $g$ . So we want to know when a measure on  $S^1$  can be invariant under the action of a subgroup of  $SL(2, \mathbb{R})$ .

The next theorem shows that there is no measure on  $S^1$  which is invariant under the action of the subgroup  $G$  of  $SL(2, \mathbb{R})$  generated by the transfer matrices  $T_m$  and so we can conclude that  $\gamma > 0$ .

**Theorem 5.**

(i) If  $H$  is a subgroup of  $SL(2, \mathbb{R})$  so that

(F1)  $H$  is not compact.

(F2) No subgroup of  $H$  with finite index in  $H$  is reducible in its action on  $\mathbb{R}^2$ ,

then there is no measure on  $S^1$  which invariant under  $H$ .

(ii) The subgroup  $G$  of  $SL(2, \mathbb{R})$  generated by at least two distinct transfer matrices obey (F1) and (F2).

*Remarks.* The criteria (F1) and (F2) are due to Furstenberg [1, Theorem 8.6]. Matsuda and Ishii [8] were the first to prove that  $G$  obeyed  $F$ . Their proof is rather difficult. Recently it has been greatly simplified by Yoshioka [11]. The proof given here is entirely due to him and is repeated for completeness.

*Proof.* (i) Any matrix  $g$  in  $SL(2, \mathbb{R})$  has a non-unique decomposition  $g = O_1 d O_2$  where  $O_1$  and  $O_2$  are orthogonal matrices and  $d$  is diagonal with positive entries  $t$  and  $t^{-1}$  (we can start with the polar decomposition  $g = so$ ,  $s$  symmetric and positive,  $o$  orthogonal and then diagonalize  $s$  by an orthogonal matrix  $O_1$ ). Thus  $g$  is specified by two angles  $\phi, \theta$  and a positive number  $t$ . Since  $H$  is not compact it is possible to choose a subsequence  $h_n$  whose parameters are  $\theta_n, \phi_n$  and  $t_n$  and as  $n$  increases,  $\theta_n \rightarrow \theta, \phi_n \rightarrow \phi$ , and  $t_n \rightarrow 0$  or  $\infty$ . In either case these two subspaces  $V_1$  and  $V_2$  so that if  $x \notin V_1$  then  $h_n x \rightarrow V_2$ , e.g. if  $t_n \rightarrow 0$  the angle between the vectors  $g_n v$  and  $(\cos \phi, -\sin \phi)$  tends to zero except perhaps if  $v$  is parallel to  $(\cos \theta, -\sin \theta)$ .

So if  $\nu$  is a measure on  $S^1$  which is invariant under  $H$  it will be concentrated at the four points of  $S^1$  determined by  $V_1$  and  $V_2$ . Any  $h \in H$  will then have to permute the subspaces  $V_1$  and  $V_2$  and since some of them will preserve  $V_1$  and  $V_2$  we can find a subgroup of  $H$  with finite index which has non-trivial invariant subspaces. This contradicts (F2). Thus if (F1) and (F2) hold there is no measure on  $S^1$  which is invariant under  $H$ .

(ii) To show that  $G$  is not compact take  $A = \begin{pmatrix} x & -1 \\ 1 & 0 \end{pmatrix}$  and  $B = \begin{pmatrix} y & -1 \\ 1 & 0 \end{pmatrix}$  with  $x \neq y$  and from  $(A^{-1}B)^n = \begin{pmatrix} 1 & 0 \\ n(x-y) & 1 \end{pmatrix}$ . Thus  $G$  is not compact. Now notice that  $AB^{-1} = \begin{pmatrix} 1 & y-x \\ 0 & 1 \end{pmatrix}$ . If  $G_0$  is a subgroup of  $G$  with finite index in  $G$  there are integers  $m$  and  $n$  so that  $(A^{-1}B)^m$  and  $(AB^{-1})^n$  are in  $G_0$ . Thus  $G_0$  contains two matrices of the type  $C = \begin{pmatrix} 1 & 0 \\ f & 1 \end{pmatrix}, D = \begin{pmatrix} 1 & g \\ 0 & 1 \end{pmatrix}$  with  $f, g \neq 0$ . The algebra generated by  $C, D$  and the identity is  $M(2, \mathbb{R})$  so any subspace invariant under  $G_0$  is also invariant under everything in  $M(2, \mathbb{R})$  and so is either  $\mathbf{0}$  or  $\mathbb{R}^2$ .

Thus  $G$  obeys (F1) and (F2) and so  $\gamma > 0$ .

Finally we remark that we have mostly suppressed the dependence of the random variables  $Z_n$  and the constants  $\gamma$  and  $\sigma$  on the frequency  $w$ . They are of course dependent on  $w$  and in future we will write  $\gamma(w)$  and  $\sigma^2(w)$  explicitly.

### 3. Conductivity of Disordered Chains

We can now use the results from the last section to show that the bulk of the heat current  $J_N$  comes from low frequencies. First we look at other ways of writing (2) which are useful. Casher and Lebowitz [7] derived the expression

$$j_N(w) = 2m_1 m_N \lambda^2 w^2 + K_{1,N}^2 + \lambda^2 w^2 (m_1^2 K_{2,N}^2 + m_N^2 K_{1,N-1}^2) + \lambda^4 w^4 m_1^2 m_N^2 K_{2,N-1}^2. \tag{31}$$

Here

$$T_N \dots T_1 = \begin{pmatrix} K_{1,N} & -K_{2,N} \\ K_{1,N-1} & -K_{2,N-1} \end{pmatrix}$$

The notation  $K_{1,N}$  comes from a representation of this function as an  $N \times N$  determinant depending on the masses  $m_1 \dots m_N$ ; in general  $K_{i,j}$  depends on  $m_i \dots m_j$ . In terms of  $t_N$  and  $\theta_N$  we have,

$$K_{1N} = t_N \sin \theta_N, \quad K_{1,N-1} = t_N \cos \theta_N. \quad (32)$$

We then have two lower bounds for  $j_N(w)$ ;

$$j_N(w) \geq 2m_1 m_N \lambda^2 w^2 \quad (33)$$

$$j_N(w) \geq t_N^2 (\sin^2 \theta_N + \lambda^2 w^2 m_N^2 \cos^2 \theta_N^2) \quad (34)$$

and so  $w^2 j_N(w)^{-1}$  is bounded either by a constant or a multiple of  $t_N(w)^{-2} \cdot \max[w^2, m_N^{-2} \lambda^{-2}]$ . We introduce the notation  $J_N(a, b)$  for the contribution to  $J_N$  from the frequency interval  $b \geq w \geq a$  and we also normalize the masses so that  $M \geq m_i \geq 1$ .

**Theorem 6.** *If the common distribution of the masses in a random chain has a  $C^1$  density with compact support then for each  $\varepsilon > 0$  and  $\lambda > 0$  there are constant  $C(\varepsilon, \lambda)$  and  $a(\varepsilon)$  so that*

$$\langle J_N(\varepsilon, \infty) \rangle \leq C(\varepsilon, \lambda) \exp[-\lambda a(\varepsilon) N^{\frac{1}{2}}] \quad (35)$$

for  $N$  sufficiently large.  $a(\varepsilon)$  is strictly positive.

*Proof.* The intervals  $[2 + \delta, \infty]$  and  $[\varepsilon, 2 + \delta]$  can be treated separately. Since each mass is greater than 1 plane waves with frequencies above 2 are exponentially attenuated by every chain. Using Rayleigh's theorem [18] we get

$$t_N(w) \geq e^{Np}$$

where  $w^2 = 2 + 2 \cosh(p)$  and so for each chain

$$J_N(2 + \delta, \infty) \leq \text{const} \exp[-\frac{1}{2} N \delta^{\frac{1}{2}}]. \quad (36)$$

At frequencies below  $2 + \delta$  the randomness of the masses is important.

We write  $T(w)$  and  $A(\lambda, w)$  to show the frequency dependence of these operators explicitly. It is only a matter of some messy algebra to check that  $T(w)^2$  can be written as an integral operator with a kernel depending continuously on  $w$ . So  $T(w)^2$  is norm-continuous with  $w$  (and so is uniformly continuous on any finite interval  $[w_0, w_1]$ ,  $w_0 > 0$ ). The same is true of  $A(\lambda, w)$  and in particular we have

(i)  $A(\lambda, w)^2$  is norm continuous in  $\lambda$  and  $w$  when  $\lambda$  lies in a small neighbourhood  $N$  of 0 and  $w \in [w_0, w_1]$ .

(ii)  $\gamma(w)$  and  $\sigma(w)$  are continuous functions of  $w$  and so have strictly positive lower bounds on  $[w_0, w_1]$ .

(iii)  $\theta(\lambda, w)$  and  $e(\lambda, w)$  and their derivatives are continuous in  $w$  and  $\lambda$ .

(iv) We can find a number  $\varrho < 1$  so that the spectra of all the operators  $B(\lambda, w)^2$  lie inside  $|z| = \varrho$ .  $\varrho$  will of course depend on  $[w_0, w_1]$ .

(v) If  $\zeta$  is a closed curve lying between the circles  $|z| = \varrho$  and  $|z| = 1$  we can find a uniform upper bound on  $\| [z - A(\lambda, w)]^{-1} \|$  when  $z \in \zeta$ ,  $\lambda \in N$  and  $w \in [w_0, w_1]$ . This upper bound  $b$  depends on  $w_0$  and  $w_1$ .

Combining these with the spectral integral representation gives

$$\begin{aligned} Ee^{\lambda Z_n(w)} &= \theta(i\lambda/(\sigma\sqrt{n}), w)^n e(i\lambda/(\sigma\sqrt{n}), w) \\ &\quad + (2\pi i)^{-1} \int_{\zeta} z^{n-1} [z - B(i\lambda/(\sigma\sqrt{n}), w)]^{-1} edz. \\ &= e^{\frac{1}{2}\lambda^2} (\text{uniformly bounded term}) + \text{error smaller than } bQ^n. \end{aligned}$$

where  $Z_n(w) = n^{-\frac{1}{2}}\sigma^{-1}X_n(w)$ .

So for any finite  $\lambda$ , it is possible to choose  $n$  so large that

$$E[e^{\lambda Z_n(w)}] \leq 2e^{\frac{1}{2}\lambda^2} \quad (37)$$

uniformly for  $w \in [w_0, w_1]$ .

If  $x$  is negative we have for negative  $\lambda$ ,

$$Pr[Z_N(w) \leq x] \leq e^{-\lambda x} 2e^{\frac{1}{2}\lambda^2} \quad (38)$$

for  $N$  sufficiently large, uniformly in  $[w_0, w_1]$ . So

$$\begin{aligned} Pr[\log t_M^2(w) < \frac{1}{2}N\gamma(w)] &= Pr[Z_N(w) < -\frac{1}{2}N^{\frac{1}{2}}\gamma(w)\sigma(w)^{-1}] \\ &\leq C(\lambda) \exp[-(\lambda/2)N^{\frac{1}{2}}a(w_0)] \end{aligned} \quad (39)$$

since where  $a(w_0) = \min\{\gamma(w)\sigma(w)^{-1} \text{ for } w \in [w_0, w_1]\}$  (by (ii), it is strictly positive).

Using (39) and the upper bound (34) gives

$$\begin{aligned} <J_N(\varepsilon, 2 + \delta)> &\leq 2 \exp[-\lambda N^{\frac{1}{2}}a(\varepsilon)] \\ &+ \int_{\varepsilon}^{2+\delta} \exp[-\frac{1}{2}N\gamma(w)] dw \leq C(\varepsilon, \lambda) \exp[-\lambda N^{\frac{1}{2}}a(\varepsilon)] \end{aligned}$$

for any fixed  $\lambda$  and  $N$  large enough.

Clearly these limit theorems must be supplemented by careful error estimates if we are extend Theorem 6 to cover frequencies which are closer to 0 and especially those of  $O(N^{-\frac{1}{2}})$ . We summarize below some accumulated intuition on disordered chains which suggests the kind of results one might try to prove.

A simple mixed view of heat conduction is that the heat baths excite a chain to vibrate in a superposition of plane wave vibrations. While only a few of these plane waves are actually normal modes of the free chain it is plausible to suppose that they resemble the normal modes whose frequencies are near. Many numerical calculations have been made on the shapes of the normal modes of a disordered chain [8, 9, 16, 19]. It seems that in most chains of length  $N$  the modes with frequencies  $w = O(N^{-\frac{1}{2}})$  fall away exponentially from a region of maximum amplitude (i.e. they are localized). Those with frequencies  $w = o(N^{-\frac{1}{2}})$  tend rather to resemble the sinusoidal forms of normal modes in periodic chains. So we expect that plane waves with frequencies  $O(N^{-\frac{1}{2}})$  will trap energy in the interior of the chain and those with frequencies  $o(N^{-\frac{1}{2}})$  will act as efficient carriers of energy along the chain, (the frequencies between  $o(N^{-\frac{1}{2}})$  and  $O(N^{-\frac{1}{2}})$  do not exhibit a sharp preference for either type).

This suggests that three steps are needed to complete this analysis.

1. The central limit Theorem 2 must be strengthened by explicit error estimates.

The spectral integral representation suggests estimating  $[z - A(\lambda)]^{-1}e$  when  $|z|$  is close to 1. A special case of this arises when we try to find the frequency dependence of  $\nu$  or  $\sigma^2$  (this involves  $[1 - T]^{-1}$ ).

We expect that  $\sigma^2(w) \sim w^2$  and since it is known that  $\gamma(w) \sim w^2$  [8, 12], we expect  $Na(w)^2 \sim Nw^2$  for small  $w$ . (This particular combination of  $N$  and  $w$  is suggested by the natural dividing line  $w \sim N^{-\frac{1}{2}}$ .)

2. For very low frequencies,  $w = o(N^{-\frac{1}{2}})$ , one must prove that the plane waves are insensitive to the individual masses but rather depend mainly on averages of the mass over long subchains.

3. To bridge the gap from  $o(N^{-\frac{1}{2}})$  to  $O(N^{-\frac{1}{2}})$  perhaps a theorem analogous to the large deviation theorems for additive random variables will suffice [17].

If (2) is true we expect that the contribution to  $J_N$  from these frequencies will be of the same order as the contribution they make to the heat current through a chain of equal masses. This is  $O(N^{-\frac{3}{2}})$  (using the method of [7], Section 3).

Finally we remark that (1) is analogous to the classical theorem of Berry and Esseen [17] and a result of Nagaev [3] for additive functions on stationary Markov chains.

### Appendix 5

Here we summarize some well known results on bilinear transformations of  $C^1$ .

The matrix  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  in  $SL(2, R)$  induces a bilinear transformation

$$T: z \rightarrow w = (az + b)/(cz + d) \quad \text{on } C^1.$$

$T$ , in general, has two fixed points  $z_+$  and  $z_-$  and with these we can rewrite  $T$  as

$$(w - z_+)/(w - z_-) = ((cz_- + d)/(cz_+ + d))((z - z_+)/z - z_-) = m((z - z_+)/z - z_-) \quad (40)$$

If  $|trA| > 2$ ,  $T$  is hyperbolic and the fixed points of the form  $\lambda$  and  $\lambda^{-1}$ ,  $\lambda$  is real and  $|\lambda| > 1$ . If  $z$  is different from  $\lambda^{-1}$ ,  $T^n z$  converges to the attractive fixed point  $\lambda$ .

If  $|trA| \leq 2$ ,  $T$  is elliptic and the fixed points are of the form  $e^{i\theta}$  and  $e^{-i\theta}$ .  $m$  is actually of the form  $e^{i\alpha}$  and if  $e^{i\alpha}$  is not an integral root of unity the successive images of any point in  $R^1$  form a dense set in  $R^1$ .

**Lemma.** Suppose  $U$  is an open set of elliptic transformations in  $SL(2, R)$ . Then for any  $x \in R^1$  the set  $S = \{g^n x : g \in U, n = 1, 2, \dots\}$  is dense in  $R^1$ .

*Proof.* Associated with each  $g$  there is a phase factor  $m = e^{i\alpha}$  (40) which depends continuously on  $g$ . Since  $\alpha$  is not a constant it cannot always be a rational multiple of  $2\pi$  and so  $S$  will be dense in  $R^1$ . In fact it is enough that there are an uncountable number of elements in  $U$ .

*Acknowledgements.* I am most grateful to Professor S. Sawyer for his painstaking reading of this paper and his many valuable comments. I also thank Professor J. T. Lewis for many helpful discussions; Professor J. L. Lebowitz for introducing me to this problem and Dr. W. G. Sullivan for helpful remarks. This work was mostly done at the Dublin Institute for Advanced Studies and I thank the Institute for its support and hospitality.

### References

1. Furstenberg, H.: Trans. A.M.S., **108**, 377 (1963)
2. Tutubalin, V. N.: Theory Probability Appl. **10**, 15 (1965); **13**, 65 (1958)
3. Nagaev, S. V.: Theory Probability Appl. **2**, 378 (1957)
4. Dunford, N., Schwartz, J.: Linear operators, pt. I. New York: Interscience 1958

5. Ahlfors, L. V.: Complex analysis. New York: McGraw Hill 1966
6. Lebowitz, J. L.: Phys. Rev. **114**, 1192 (1959)
7. Casher, A., Lebowitz, J. L.: J. Math. Phys. **12**, 1701 (1971)
8. Matsuda, H., Ishii, K.: Suppl. Prog. Theor. Phys. **45**, 231 (1970)
9. Ishii, K.: Suppl. Prog. Theor. Phys. **53**, 77 (1973)
10. Rubin, R. J., Greer, W.: J. Math. Phys. **12**, 1686 (1971)
11. Yoshioka, Y.: Proc. Japan Academy **49**, 665 (1973)
12. O'Connor, A. J., Lebowitz, J. L.: J. Math. Phys. **15**, 692 (1974)
13. Schmidt, H.: Phys. Rev. **105**, 425 (1957)
14. Borland, R. E.: Proc. Roy. Soc. (London) A. **274**, 529 (1963)
15. Halperin, B. I.: Advances in Chemical Physics **13**, 123 (1967)
16. Hori, J.: Spectral Theory of Disordered Chains and Lattices. Oxford: Pergamon, 1968
17. Feller, W.: Introduction to probability theory and its applications, II. New York: Wiley 1966
18. Maradudin, A. A., Montroll, E. W., Weiss, G. H.: Theory of lattice dynamics in the harmonic approximation. New York: Academic Press 1968
19. Bell, R. J.: Reports on Progress in Physics **35**, 1315 (1972)
20. Kato, T.: Perturbation theory for linear operators. Berlin-Heidelberg-New York: Springer 1966
21. Peierls, R. E.: Quantum Theory of Solids, Oxford 1965

Communicated by J. L. Lebowitz

*Received June 17, 1975*

