SMOOTHNESS OF THE INTEGRATED DENSITY OF STATES FOR RANDOM SCHRÖDINGER OPERATORS ON MULTIDIMENSIONAL STRIPS

Thesis by

Clemens H. Glaffig

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

(Submitted June 7, 1988)

ACKNOWLEDGEMENTS

I wish to express my gratitude to my advisor, Prof. B. Simon, for suggesting this thesis problem, for his support, advice and his patience during its solution.

I would also like to thank Prof. F. Gesztesy for his great support and encouragement and for being a patient listener, Prof. J. Lacroix for some deeper insight in other implications of some of this work and A. Hardarson for also being a patient listener.

Finally I would like to thank all my friends, they know who they are, for keeping me company and providing me with amusement and distraction over the last several years.

ABSTRACT

We investigate smoothness properties of the integrated density of states (ids) for random Schrödinger operators on a multidimensional strip lattice, where only the potentials on the "top surface" of this lattice have a distribution with some regularity.

We view the eigenvalue equation on the strip as the action of an abstract group on some homogeneous space, from where we derive a representation of the ids in terms of a distinguished measure on that homogeneous space.

This representation allows us to conclude that using minimal smoothness of the potential distribution on the "top surface", combined with a negative moment condition for the distribution of all other potentials, is enough to obtain smoothness of the ids. This includes the original Anderson model.

We also discuss cases, where the distribution of the potentials below the "top surface" is Bernoulli, satisfying this negative moment condition.

TABLE OF CONTENTS

§ 1 INTRODUCTION

Let H be the discrete random Schrödinger operator defined on $l^2(\mathbb{Z}^d)$ by

$$H_\omega u\ (n) = \sum_{\|n-m\|=1} u(m) + V_\omega(n)\ u(n) \quad n,m \varepsilon\ \mathbb{Z}^d \tag{1.1}$$

where $\{ V_\omega(n),\ n \varepsilon\ \mathbb{Z}^d \}$ is a family of independent random variables and we will assume that the potential is bounded, that is, the distribution of the $V_\omega(n)$ has compact support. It is well known, that H_ω is a selfadjoint operator and if V_ω is an i.i.d. family with distribution $d\mu$, that the spectrum is almost surely the set

$$\sigma(H_\omega) = [-2d,2d] + \text{supp}\ d\mu.$$

In addition, if d=1, the spectrum is pure point with exponentially decaying eigenfunctions.

We will be interested in random Schrödinger operators on a 'strip', that is on $l^2(S_m)$, where $S_m = \mathbb{Z} \times \{1,...,m\}^d$, a horizontal, (d+1)-dimensional strip of width m. The reason for studying the strip is that one can observe phenomena not occurring in one dimension, giving some hints to what might happen in higher dimensions. At the same time, while techniques to study the higher dimensional case have not yet proven to be powerful enough to obtain conjectured results, the strip can be studied by one dimensional techniques.

We are mainly interested in regularity results of the integrated density of states on the strip, defined below, where we will focus on C^∞ results.

For the finite box $\Lambda_l = (n \varepsilon \mathbb{Z}^d, |n_i| \leq l \}$, let $H_{\omega,l}$ denote the restriction of H to Λ_l with Dirichlet boundary conditions. This makes $H_{\omega,l}$ a $|\Lambda_l| \times |\Lambda_l|$ *matrix.*

The integrated density of states , (ids), is defined by

$$k(E) = \lim_{l \to \infty} |\Lambda_l|^{-1} \# \{ \text{ of eigenvalues of } H_{\omega,l} \leq E \}. \tag{1.2}$$

As a consequence of the ergodic theorem, (see e.g. Carmona [1]), one obtains

Theorem:

For all E, the limit in (1.2) exists for a.e. ω and is independent of the boundary condition chosen for $H_{\omega,l}$.

If $\mathbb{Z}^d$ is replaced by S_m, the definition of k (E) remains the same, if the boxes Λ_l cover the whole width and are horizontaly restrained to -l,. . . ,l.

The ids and in particular regularity properties thereof have been studied extensively in recent years. One of the first results was Pastur's proof of the continuity of k in one dimension ([1]). For arbitrary dimension, Simon-Craig, [1], showed that k is log-Hölder continuous. Other results in higher dimensions had to assume large disorder or high energies, as the result of Constantinescu, Fröhlich and Spencer ,[1], who showed, in the i.i.d. case, that if the potential distribution has an analytic extension around the real axis, then so does k. More recent results include the proof of Bovier et al.,[1], showing smoothness of k for a class of potential distributions that includes the uniform distribution, again at high disorder.

The first C^∞ result for k(E) with minimal regularity assumptions for the distribution of the potential was obtained by Simon-Taylor, ([1]), in the one dimensional case with i.i.d potential distribution. The result was later extended by Klein-Speis,([1]), to i.i.d potential distributions having a Fourier transform with all derivatives bounded and decaying to 0 at ∞.

Recently, Klein - Speis, (K.-S. [1]), proved k(E) to be C^∞ on S_m for d = 1 and i.i.d potential distributions $d\mu$ having first two moments and satisfying

$$(1+ t^2)^{\frac{\alpha}{2}}\, d\hat{\mu} \;\varepsilon\; L^1, \;\; \alpha > \tfrac{1}{2}$$

While they used a 'supersymmetric replica trick' to prove their results, we are going to follow Simon-Taylor and use the group action approach to obtain regularity results for the ids. In higher dimensions, it is expected that the ids is smooth for all i.i.d. potentials, in particular for potentials with a pure point potential, including the ones for which the one dimensional ids is singular continuous (Carmona et al., [1]).

The main object of this thesis is, to not only generalize one dimensional regularity results for the ids to the multidimensional strip, but also to obtain results that can be considered as natural "interpolated results" between known facts in one dimension and conjectured results in higher dimensions. As such a result, we derive smoothness of the ids by only using minimal regularity of the potentials on the "top surface", while the potential below the "top surface" won't have to satisfy any regularity conditions. The extension of the one dimensional result in Simon-Taylor, [1], to the multidimensional strip is a consequence of this. In particular, we investigate what happens, if the variables below the "top surface" have pure point distribution.

For our purpose, it will be convenient to redefine H_ω as an operator on $\mathbb{R}^{m^d}$ - vectors, acting in the following way:

$$H_\omega u\,(n) = u(n+1) + u(n-1) + V_\omega(n)\,u(n)\,,\quad n \in \mathbb{Z} \tag{1.3}$$

where $u(i) \in \mathbb{R}^{m^d}$ and

$$V_\omega(n) = \begin{bmatrix} W_{n1} & id & & & \\ id & W_{n2} & id & 0 & \\ & \ddots & \ddots & \ddots & \\ & 0 & \ddots & \ddots & id \\ & & & id & W_{nm} \end{bmatrix} \tag{1.4}$$

where the blocks are symmetric $m^{d-1} \times m^{d-1}$ matrices and W_{ni} have again the same structure as in (1.4), just with $m^{d-2} \times m^{d-2}$ blocks. The m x m diagonal blocks of (1.4) are of the form

$$V_\omega^m(n) = \begin{bmatrix} x_{nk_i} & 1 & & & \\ 1 & \ddots & 1 & 0 & \\ & \ddots & \ddots & \ddots & \\ & 0 & \ddots & \ddots & 1 \\ & & & 1 & x_{nk_i} \end{bmatrix},\ n \in \mathbb{Z},\ k_i \in \{1, \ldots, m\}^d,\ i=1, \ldots, m \tag{1.5}$$

x_{nk_i} is the potential at site (n,k_i).

The V_ω are just d- dimensional Schrödinger operators, restricted to Λ_m.

For d = 1, i.e., for the 2-dimensional strip, (1.4) takes the form of (1.5).

Associated to (1.3) are, for each energy E, the so called transfer matrices (see more in § 2), that are of the form

$$g_n(\omega,E) := \begin{bmatrix} E\text{-}V_\omega(n) & -\mathrm{id} \\ \mathrm{id} & 0 \end{bmatrix}$$

where the blocks are $m^d \times m^d$.

These matrices are elements of the symplectic group $Sp(m^d)$, (again, more in § 2). The condition we require for the potential distribution below the "top surface", will be stated in terms of products of transfer matrices. More precisely, consider, for fixed E, the map ψ: $\mathbb{R}^{nm^d} \rightarrow Sp(m^d)$ given by

$$\psi(x_{11}, \ldots, x_{nm^d}) := g_n \ldots g_1$$

where n will be specified later.

We will show in § 3, that ψ is for a.e. fixed realization of the potentials below the "top surface" a diffeomorphism a.e. in the "top surface" - variables. If ψ_c denotes the restriction of ψ for such a typical realization of the below-"top surface" variables, then $\psi_c : \mathbb{R}^{nm^{d-1}} \rightarrow Sp(m^d)$ is a diffeomorphism a.e. and the Jacobian determinant will locally be of the form

$$\det D\psi_c = \alpha(c)\ \beta(c,x)\ ,\ c\varepsilon\ \mathbb{R}^{n(m-1)m^{d-1}},\ x\varepsilon\ \mathbb{R}^{nm^{d-1}}, \qquad (1.6)$$

with α and β polynomial.

For our results, we require the following condition on the distribution of the potentials below the "top surface":

(C1) If $d\eta$ denotes the distribution of the x_{ik} for $i = 1, \ldots, n$ and $k \varepsilon \{1, \ldots, m\}^d$ with $k(d) < m$ and $n = 2m^{2m^d} + m^d$ ($=$ dim $Sp(m^d)$), then there exists $t > 0$ such that for all E

$$\int \alpha^{-t}(x-E)\, d\eta(x) < \infty$$

where α is a polynomial in x, given by (1.6)

Theorem 1:

Let x_{ij} for $i \varepsilon \mathbb{Z}$ and $j \varepsilon \{1, \ldots, m\}^d$ with $j(d) = m$, (the top surface variables), be i.i.d. random variables with distribution $F(x)dx$ with supp F compact and $F \varepsilon L^1_\alpha$ for some $\alpha > 0$, L^1_α the usual Sobolev space.

Let x_{ij} for $i \varepsilon \mathbb{Z}$ and $j \varepsilon \{1, \ldots, m\}^d$ with $j(d) < m$, (the variables below the top surface), be i.i.d. random variables, mutually independent of the top surface variables, with distribution satisfying (C1) and having compact support.

Then $k(E)$ is a C^∞ function.

Remarks: (i) If the x_{ij}, $i \varepsilon \mathbb{Z}$, $j \varepsilon \{1, \ldots, m\}^d$, $j(d) < m$ are i.i.d. with absolutely continuous distribution, (C1) is satisfied.

(ii) If the family of all potentials is an i.i.d. family with distribution $F(x)dx$, $F \varepsilon L^1_\alpha$, some $\alpha > 0$, then (C1) is satisfied. In particular, if all potentials have uniform distribution over some interval, k is C^∞.

To emphasize that only smoothness on the "top surface" is needed to smoothen the ids, we are particularly interested in the case when the potentials below the "top surface" can only take two values a or b with complementary probabilities, i.e., when those variables have a Bernoulli (a,b) distribution with Prob $(X = a) = p$ and Prob $(X = b) = 1-p$, $0 < p < 1$. We have

Theorem 2:

(i) *Let the distributions of the "top surface" potentials be as in Theorem 1 and all potentials independent. If supp* $F \subset K$ *for some fixed compact* K, *there exists* $b_0 > 0$ *such that if the variables below the "top surface" have Bernoulli* (a,b) *distribution with* $|b-a| < b_0$, *then* $k(E)$ *is a* C^∞ *function.*

(ii) *Let the distributions of the "top surface" potentials be as in Theorem 1 and all potentials independent. If the variables below the "top surface" have Bernoulli - (a,b) distribution, then, for any* $\epsilon > 0$ *and a.e. vector* $c = (c_{ik}) \varepsilon B_\epsilon$, *where* B_ϵ *is the ball of radius* ϵ *around* 0 *in* $\mathbb{R}^{n(m^d - m^{d-1})}$, *let* $\tilde{x}_{ik} = x_{ik} + c_{ik}$, $i = i \bmod n$, $k \varepsilon \{1, \ldots m\}^d$, $k(d) < m$. *That is, the revised potential is the former one with an arbitrarily small periodic perturbation added to it. Then, for the perturbed model,* $k(E)$ *is a* C^∞ *function.*

(iii) *For* $d = 1$, *i.e. the 2-dimensional strip, and* $m = 2$, *let the distribution of the "top surface" potentials be as in Theorem 1 and all potentials independent. If the variables below the "top surface" have Bernoulli* (a,b) *distribution,* $k(E)$ *is a* C^∞ *function.*

Remarks: (i) The spectrum of the perturbed operator in part (ii) is pure point.

(ii) Other results that follow easily the same way as (i) and (ii) are for all but finitely many Bernoulli (a,b), that is for all but finitely many values of $|b-a|$, and all but finitely many energies E.

Along the way, we also get a localisation result "for free" in the case where the potentials are constant along each horizontal line, if the top surface is perturbed by any i.i.d. random sequence with absolutely continuous distribution on $\mathbb{R}$, that is, if such a pertubation is added, the spectrum of H_{pert} becomes pure point with exponentially decaying eigenfunctions. We will comment on this at the end of section 4.

For our goal, to show smoothness of the ids, (1.2) is not a very useful expression to analyze. We will therefore derive a different representation for the ids, which will relate the smoothness of the ids to smoothness of quantities, that are easier to analyze. Therefore, we will relate the number of eigenvalues as described in (1.2), with the number of solutions of a first order recurrence equation for symmetric matrices (see 1.7 below). Randomizing the initial condition of this recurrence equation and viewing it as the action of an abstract group G on the symmetric matrices, we obtain a relation between the ids and a measure $d\nu_E$ on the set of symmetric matrices, which will satisfy $d\mu_E * d\nu_E = d\nu_E$ for a given measure $d\mu_E$ on the group G. First we will "generate" smoothness by taking convolution powers of $d\mu_E$, using that this measure has a certain smoothness locally in one variable. By taking higher convolution powers, we obtain any desired smoothness of those powers. We will then obtain suitable E-smoothness of $d\nu_E$ by perturbation theory, from where we finally conclude the smoothness of the ids.

The setup of this will be done next and in section 2 we outline the group action approach. The smoothness of the convolution powers of $d\mu_E$ will be the purpose of section 3. In section 4, we relate this smoothness to smoothness of the invariant measure and make the final conclusions for Theorem 1. The proof of Theorem 2 will be given in section 5.

The following negative eigenvalue theorem allows us to find a representation for k(E) in terms of quantities which's studies will be the object of the following sections.

Let U(n) be the square matrix of order m^d, the j-th column of which is the solution of

$$H_\omega u(n) = E u(n) \qquad n = -l, \ldots, l$$

with initial conditions u(-l-1) = (0,. . .,0) and u(-l) = (0,. .,1,. .,0), where the one is in the j-th coordinate.

Then the U(n) satisfy

$$U(n+1) + U(n-1) + (V_\omega(n) - E) U(n) = 0 , n = -l,\ldots,l \qquad (1.7)$$

If U(n) is invertible, define $X(n+1) := U(n+1) U(n)^{-1}$ for n = -l, . . ,l. X(n) is symmetric and depends on ω and E. If U(n) is not invertible, X(n) is not invertible. In this case let $X(n)^{-1}$ denote the matrix with the same spectral decomposition as X(n), just with the reciprocal eigenvalues and define X(n+1) by

$$X(n+1) := E - V_\omega(n) - X(n)^{-1}$$

accordingly. Then the X(n) solve

$$X(n) = (E - V_\omega(n-1)) - X(n-1)^{-1}, \qquad n = -l+1,\ldots,l \qquad (1.8)$$

with $X(-l)^{-1} := 0$.

Theorem (Dean- Martin [1]) :

Let $X(n)$ solve (1.8) for the given choice of $X(-l)$. Then:

$$\#\{ \text{ of eigenvalues of } H_{\omega,l} \leq E \} = \sum_{i=-l+1}^{l+1} \#\{ \text{ of positive eigenvalues of } X(i) \}.$$

proof: For notational convenience, we will drop the indices ω and l. H is then the following matrix:

$$H := \begin{bmatrix} V(-l) & id & & \\ id & \ddots & id & \\ & \ddots & \ddots & \ddots \\ & & id & V(l) \end{bmatrix}, \text{ all empty entries are 0,} \tag{1.9}$$

where the diagonal blocks are m^d x m^d, given by (1.4).

Then E - H can be written as a product of two block triangular matrices:

$$E - H = \begin{bmatrix} id & & & \\ Y(-l+1) & id & & \\ & \ddots & \ddots & \\ & & Y(l) & id \end{bmatrix} \begin{bmatrix} X(-l+1) & id & & \\ & \ddots & id & \\ & & \ddots & \ddots \\ & & & X(l+1) \end{bmatrix}, \tag{1.10}$$

where all empty entries are 0 and $Y(i+1) = X(i)^{-1}$ and the X(i) are defined in (1.8).

It follows from (1.10), that

$$\det(E\text{-}H) = \prod_{i=-l}^{l} \det X(i).$$

From (1.8), we can see that for E negative and $|E|$ large, all eigenvalues of the X(i) are negative. If E is increased through an eigenvalue of H such that det (E-H) changes sign, then at least one of the eigenvalues of the X(i) will change sign. Suppose X(i) has an eigenvalue x with $|x|$ small. Then X(i+1) has an eigenvalue close to $-x^{-1}$. If x changes sign through 0, $-x^{-1}$ changes sign through ∞, so the total number of positive eigenvalues is unchanged if $-l < i < l$. For the same reason, if a large eigenvalue of X(i+1) changes sign through ∞, a small eigenvalue of X(i) with opposite sign will change sign through 0. Therefore, the total number of positive eigenvalues can only change if:

(i) An eigenvalue of X(l) changes sign through 0

(ii) An eigenvalue of X(-l) changes sign through ∞

Since det X(-l) = det (E-V(-l)) has no poles, only (i) is possible. Therefore, the total number of positive eigenvalues of the X(i)'s will only change if E increases through an eigenvalue of H and the difference in the number will equal the multiplicity of the eigenvalue of H. Since for E large, all the eigenvalues of the X(i)'s will be positive, the statement follows. □

Since the limit in the definition of the ids was obtained for a.e. ω and was independent of the boundary condition chosen for $H_{\omega,l}$, it follows from the theorem above that for all choices of symmetric X_{-l}

$$k(E) = \lim_{l \to \infty} \; | \Lambda_l |^{-1} \sum_{n=-l}^{l} \sum_{i=1}^{m^d} \mathrm{Exp} \, [\, \chi_{(0,\infty]} \, (\, \lambda_i(X(n,\omega,X_{-l}))) \,] \qquad (1.11)$$

where $X(n,\omega, X_{-l})$ is the solution of (1.8) with initial condition X_{-l}, λ_i is the i-th eigenvalue of $X(n,\omega)$ and Exp is integration with respect to the potential distribution.

Because the right hand side of (1.11) is bounded, this relation also holds if X_{-l} is random, independent of the potentials and the expectation with respect to this random variable is taken. With X_{-l} random, (1.11) defines a recurrence equation for a Markov chain of symmetric random matrices. It will be advantageous for the further anlysis to view (1.11) as the action of elements of an abstract group on elements of a homogeneous space, as outlined in the next section.

§2 THE GROUP ACTION APPROACH

We are going to outline an abstract view of (1.8), based on work done by Furstenberg, ([1],[2]). For all the following considerations, let G be a locally compact, semisimple Lie group and M a Borel space.

Definition:

M is called a G-space, if there is a continuous action, $(g,x) \rightarrow gx$ *of* $G \times M \rightarrow M$, *satisfying*

$$(g_1\ g_2)\ x = g_1\ (\ g_2\ x) \tag{2.1}$$

If in addition the equation, $gx = y$, *has a solution in G for every x,y in M, the action is said to be transitive and M is called a homogeneous space of G.*

In the following we are going to recall some definitions and facts from Furstenberg [1],[2].

For measures $d\mu$ on G and $d\nu$ on M, the convolution $d\mu * d\nu$ is defined by

$$\int f(x)\ d\mu * d\nu(x) = \int f(gx)\ d\mu(g)\ d\nu(x) \tag{2.2}$$

where f is continuous and vanishes at infinity.

The convolution of two measures on G is defined by (2.2), if G itself is considered a G-space. If $d\mu_1$, $d\mu_2$ are measures on G, $d\nu$ is a measure on M, (2.1) implies that

$$(\ d\mu_1 * d\mu_2\ ) * d\nu = d\mu_1 * (\ d\mu_2 * d\nu\ ) \tag{2.3}$$

If $d\mu$, $d\nu$ are probability measures on G, M respectively, then $d\mu * d\nu$ is the distribution of gx, if g and x have distribution $d\mu$ and $d\nu$ respectively and are independent.

Definition:

(i) *If M is compact and for any probability measure* $d\eta$ *on M there exists a sequence,* (g_n), *in G with* $\delta_{g_n} * d\eta$ *converging to a point measure, where* δ_g *is the point measure on G concentrated on g, M is called a boundary of G.*

(ii) *If* M_1, M_2 *are homogeneous spaces of G, a map* $\varphi : M_1 \rightarrow M_2$ *is called equivariant if* $g\varphi(x_1) = \varphi(gx_1)$ *for all* $g \varepsilon G$, $x_1 \varepsilon M_1$.

One of the results in Furstenberg [2] is:

Proposition:

All boundaries of G are equivariant images of one of them, the maximal boundary.

For any $x_0 \varepsilon M$, let $L = \{ g: gx_0 = x_0 \}$, the stability group of x_0. If G acts transitively on M, G/L is homeomorphic to M.

Let $d\mu$ be a probability measure on G.

Definition: *A probability measure on M is called an invariant measure for* $d\mu$, *if*

$$d\mu * d\nu = d\nu.$$

The following result of Furstenberg is fundamental to our further analysis.

<u>Theorem</u>: (Furstenberg [1])

(i) If M is compact, there exists an invariant measure for $d\mu$ on M, independent of $d\mu$.

(ii) If M is a boundary of G and $d\mu$ an absolutely continuous (w.r. to Haar measure) probability measure, then there is one and only one invariant measure for $d\mu$ on M.

The special case we are interested in, is when G equals the symplectic group, $G = Sp\ (m^d)$, i.e., the set of $2m^d \times 2m^d$ matrices g, satisfying

$$g^T J g = J, \tag{2.4}$$

where

$$J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$ the 1 stands for the $m^d \times m^d$ identity matrix

M will be the set of Lagrangian subspaces of $\mathbb{R}^{2m^d}$, that is the set of m^d-dimensional subspaces of $\mathbb{R}^{2m^d}$ satisfying

$$< u, Jv > = 0 \quad \forall u,v \in x,\ x \in M \tag{2.5}$$

where $< , >$ is the usual inner product.

For given $x \in M$, let (u_i) be a basis for x and let $\bar{x}$ denote the $2m^d \times m^d$ matrix, whose columns are the vectors (u_i), $i = 1,\ldots,m^d$.

The maximal boundary of G is well known (see Lacroix [2] for this special ,case and Furstenberg [2] for the general case) to be the following flag manifold:

Let x_i be an i-dimensional isotropic subspace of $\mathbb{R}^{2m^d}$, i.e., an i-dimensional subspace satisfying (2.5). The maximal boundary of G is the set

$$\{ (x_1, \ldots, x_{m^d}, x_i \subset x_{i+1}, i<m^d, x_i \text{ isotropic} \}.$$

From this it is also easy to see that M is an equivariant image of this set.

Lemma 2.1:

M is a homogeneous space under the action

$$(g,x) \rightarrow g\,\bar{x}$$

proof: Recall that the (u_i) span x. For any u_i, u_j

$$< gu_i, Jgu_j> = < u_i, g^T Jgu_j> = < u_i, Ju_j> = 0 ,$$

so the columns of $g\,\bar{x}$ span again a Lagrangian subspace. The associativity of the action is also easily verified. To see the transitivity, let x_1, $x_2 \varepsilon$ M with bases (u_i) and (w_i) respectively. The systems $u_{m^d+i} := Ju_i$, $w_{m^d+i} := Jw_i$, $i = 1, \ldots, m^d$, extend the previous bases to ones for $\mathbb{R}^{2m^d}$ with $< w_{i+m^d}, Jw_{j+m^d} > = < w_i, Jw_j > = 0 \;\forall\; i,j \leq m^d$ and therefor also $< u_{i+m^d}, Ju_{j+m^d} > = 0 \;\forall\; i,j \leq m^d$. Then there exists a $2m^d$ x $2m^d$ matrix g such that

$$gu_i = w_i \;\forall i\leq 2m^d \text{ and } < u_i, g^T Jgu_j> = 0 \;\;\forall i,j\leq m^d.$$

Thus, for fixed but arbitrary $i\leq 2m^d$,

$$< w_j, (Jg - g^{-1^T} J)u_i > = 0 \;\;\forall\; j\leq 2m^d.$$

Therefore $(Jg - {g^{-1}}^T J)u_i = 0 \; \forall i \leq 2m^d$, so $g^T Jg = J$ and g is symplectic, mapping x_1 to x_2. □

Lemma 2.2:

$$M \simeq G/H \simeq \mathcal{K}/\mathcal{B} \quad , \; \textit{where}$$

$$H = \left\{ \begin{bmatrix} A & B \\ 0 & (A^{-1})^T \end{bmatrix}, \; A, \, B \textit{ are } m \times m \textit{ matrices with } AB^T = BA^T \right\}.$$

$$\mathcal{K} = \left\{ \begin{bmatrix} A & B \\ -B & A \end{bmatrix} \right\} \cap G = SO \cap G \simeq U(n),$$

$U(n)$ the complex unitary group.

$$\mathcal{B} = \left\{ \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}, \; A \textit{ orthogonal} \right\}$$

proof: Since G acts transitively on M, $M \simeq G/L$, where L is the stabilizer for $e \varepsilon M$ with $\overline{e} = (\text{id}, 0)^T$.

If for $g \varepsilon G$,

$$g = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

$g\overline{e} = \overline{e}$ yields $C = 0$, $D = (A^{-1})^T$ and (2.3) yields $AB^T = BA^T$. Any g satisfying these properties is in L and thus $M \simeq G/H$.

To see the second relation, we note that for G one has the Iwasawa decomposition $G = \mathcal{K}\,\mathcal{A}\,\mathcal{N}$, where

$\mathcal{K}$ is defined in the statement of the Lemma,

$$\mathcal{A} = \left\{ \begin{bmatrix} B & 0 \\ 0 & B^{-1} \end{bmatrix}, \; B = \mathrm{diag}\,(b_1, \ldots, b_n)\,, \;\; \text{all } b_i > 0 \right\}$$

$$\mathcal{N} = \left\{ \begin{bmatrix} A & B \\ 0 & (A^{-1})^T \end{bmatrix}, \; AB^T = BA^T, \text{ A upper triangular with diagonal all one's} \right\}$$

Clearly $\mathcal{A}$, $\mathcal{N}$ are in L.

An element $\begin{bmatrix} A & B \\ -B & A \end{bmatrix}$ in $\mathcal{K}$ leaves $\bar{e}$ invariant iff $B = 0$ and A is orthogonal.

Therefore, the subgroup of $\mathcal{K}$ leaving $\bar{e}$ invariant is $\mathcal{B}$ and

$$H = \mathcal{B}\mathcal{A}\mathcal{N}$$

from where $G/H \simeq \mathcal{K}/\mathcal{B}$ and the second relation follows. □

Remark: $\mathcal{K}$ is a maximal compact subgroup of G.

In order to employ Furstenberg's result, we need to show that M is a boundary. To simplify the argument, we are going to use another result of Furstenberg:

Proposition: (Furstenberg [2])

If M is a compact homogeneous space of G, M is a boundary if, for some smooth probability measure $d\nu_0$ *on M, there is a sequence of probability measures on G with* $d\mu_n * d\nu_0$ *convergent to a point measure.*

Remarks: (i) A measure $d\nu$ on M is smooth, if it is locally equivalent to Lebesgue measure, that is, for any coordinate neighborhood U in M with coordinate map α, there is a function $\varphi(x) > 0$ for $x \varepsilon \alpha(U)$ such that for f with supp $f \subset U$

$$\int f(m)\, d\nu(m) = \int f(x)\, \varphi(x)\, dx.$$

(ii) The set of Lagrangian subspaces that cannot be spanned by the columns of $[A,B]^T$, where $\det A \neq 0$, is a zero set for any smooth measure on M, since $\det A = 0$ cannot be one of the restrictions on A,B imposed by (2.5).

Corollary 2.1:

M is a boundary of G.

proof: That M is a compact homogeneous G-space follows from Lemma 2.1 and the remark thereafter.

Let g_n be the diagonal matrix with $g_n(i,i) = n$ for $i \leq m^d$ and $g_n(i,i) = \frac{1}{n}$ for $i = m^d+1, \ldots, 2m^d$. Obviously all g_n are in G. If $m \varepsilon M$ is spanned by the columns of $[A,B]^T$, then $g_n m \rightarrow e$ (e as in Lemma 2.2) for those m with $\det A \neq 0$. But then $g_n m \rightarrow e$ for almost all m with respect to any smooth measure on M. Therefore, taking $d\mu_n = d\delta_{g_n}$, $d\mu_n * d\nu \rightarrow d\delta_e$ for any smooth $d\nu$ and thus, by the preceding proposition, M is a boundary of G. □

Furstenberg's theorem allows us to conclude

Corollary 2.1:

If $d\mu$ is an absolutely continuous probability measure on G, there exists a unique probability measure $d\nu$ on M such that

$$d\mu * d\nu = d\nu.$$

We want to reconsider (1.11) in this framework. Let us remind that the transfer matrices associated to H_ω are of the form

$$g_n(\omega, E) := \begin{bmatrix} E - V_\omega(n) & -1 \\ 1 & 0 \end{bmatrix}, \tag{2.6}$$

where the one's denote the m^d x m^d identity matrix and $V_\omega(n)$ is given by (1.4).

The evolution of a solution to (1.7) is given by

$$\begin{bmatrix} U(n+1) \\ U(n) \end{bmatrix} = g_n(\omega, E) \begin{bmatrix} U(n) \\ U(n-1) \end{bmatrix}, \quad n = -l+1, \ldots, l$$

and thus (1.8) is equivalent to

$$\begin{bmatrix} X(n) \\ 1 \end{bmatrix} = g_{n-1}\,(\omega,E) \begin{bmatrix} 1 \\ X(n-1)^{-1} \end{bmatrix}, \qquad n = -l+1, \ldots, l \tag{2.7}$$

For any symmetric m^d x m^d matrix A, the columns of $(A, 1)^T$ span a Lagrangian subspace. If A is invertible, the same subspace is spanned by the columns of $(1, A^{-1})^T$. If $\tilde{X}(n)$ denotes the Lagrangian subspace associated to X(n), (2.4) becomes

$$\tilde{X}\,(n) = g_{n-1}\,(\omega,\,E)\;\;\tilde{X}\,(n-1), \qquad n = -l+1, \ldots, l \tag{2.8}$$

If $d\nu_0$ denotes the distribution of X (-l), (2.7) defines a Markov chain in M with initial distribution $d\nu_0$, if one views the elements $g_n\,(\omega,\,E)$ as random variables with the induced distribution of the potential- matrix (1.4) , denoted by $d\mu_E$. The distribution of X (-l+n), n < 2l+1, is then given by

$$d\nu_{-l+n} \;=\; \Big((*)^n\; d\mu_E\Big) * \; d\nu_0 \tag{2.9}$$

If $d\nu_0$ is an invariant distribution for $d\mu_E$, call it $d\nu_E$, then

$$d\nu_n = d\nu_E \qquad \forall\; n = -l, \ldots, l$$

In this case, it follows from (1.11) that

Corollary 2.2:

$$k(E) = \int \kappa(x)\, d\nu_E(x) \tag{2.10}$$

where

$$\kappa(x) = \sum_{i=1}^{m^d} \chi_{(0,\infty]}(\lambda_i(x)) \text{ and } \lambda_i(x) \text{ is the } i\text{-th eigenvalue of } x.$$

The smoothness of k (E) will now follow from appropriate smoothness with respect to E of the invariant measure $d\nu_E$. Since $\kappa(x) \varepsilon L^\infty(M)$, we have

Corollary 2.3:

If $d\nu_E$ *is* C^∞ *in E as an element of the dual space of* $L^\infty(M)$, $k(E)$ *is also* C^∞.

Remark: It follows from Furstenberg's theorem that if there exists some n such that $(*)^n d\mu_E$ is absolutely continuous, then $d\mu_E$ has a unique invariant measure. We will show the existence of such n in the next section.

§3 SMOOTHNESS ON THE GROUP

The desired smoothness in E of the invariant measure will follow from the smoothness in E of a number of convolutions of $d\mu_E$ by perturbation theory. Recall that $d\mu_E$ denotes the measure on G, that is concentrated on matrices of the form (2.6). $d\mu_E$ itself does not have any smoothness but we can generate smoothness by convolution, using fractional smoothness of only the "top surface potential", that is, we only use some fractional smoothness of the distributions of the potentials at sites (i,k), $i \varepsilon \mathbb{Z}$ and $k \varepsilon \{1, \ldots ,m\}^d$ with $k(d) = m$. More precisely, we are going to show that for some n with $nm^{d-1} \geq \dim G = 2m^{2d}+m^d =: l$, (equality for $d = 1$), and E fixed,

$$(*)^n \, d\mu_E \, (g) = G_n \, (g,E) \, dg, \tag{3.1}$$

where locally on $\mathbb{R}^l$, $G_n \varepsilon L^p_\alpha$, for some $p>1, \alpha>0$, which is equivalent to saying that G_n is in the same Sobolev space on the group as indicated in the next section. In this section all considerations are locally, so all Sobolev spaces will be on $\mathbb{R}^l$.

$(*)^n d\mu_E$ (g) is the distribution of the product $g_n(\omega,E) \ldots g_1(\omega,E)$, where the g_i are i.i.d. with distribution $d\mu_E$.

(3.1) will only be possible if $(*)^n d\mu_E$ has l-dimensional support. G_n dg is, as we will see, the push forward map of the joint distribution of l potential variables. Since we only want to use the differentiability of the potential distributions on the top surface of S_m, it is clear that those l variables have to be on the top surface to ensure that $G_n \varepsilon L^p_\alpha$, so we see that n has to be such that nm^{d-1} is at least dim G to achieve (3.1).

For the following, we are going to fix the potentials not lying on the "top surface", by setting $x_{ik} = c_k$ = const., for $i \varepsilon \mathbb{Z}$ and $k \varepsilon \{1, . . ,m\}^d$ with $k(d) < m$.

We will denote the measure on the transfer matrices resulting from this restriction by $d\mu_{E,c}$. Then we have

Lemma 3.1:

$< \text{supp } d\mu_{E,c} > = G$, where $< \bullet >$ denotes the generated subgroup.

Remark: This, of course, also proves that $< \text{supp } d\mu_E > = G$, since $< \text{supp } d\mu_{E,c} > \subset < \text{supp } d\mu_E >$, independent of the distribution of the x_{ik}'s for all $i, k \varepsilon \{1, . . ,m\}^d$, $k(d) < m$.

proof: We will show that the Lie algebra of $< \text{supp } d\mu_{E,c} >$ coincides with $\mathfrak{G}$, the Lie algebra of G, which is the set

$$\mathfrak{G} = \left\{ \begin{bmatrix} X_1 & X_2 \\ X_3 & -X_1^T \end{bmatrix} \quad X_2, X_3 \text{ symmetric} \right\}$$

It is easy to see that the system $\{ (X_{ij}), (Y_{ij}), (Z_{ij}) \}$ is a basis for $\mathfrak{G}$, if these elements are defined as follows:

Let E_{ij} be the m^d x m^d matrix with a one in the (i,j) entry and zeros everywhere else.

Define

$$X_{ij} = \begin{bmatrix} 0 & E_{ij} + E_{ji} \\ 0 & 0 \end{bmatrix},$$

$$Y_{ij} = X_{ij}^T$$

$$Z_{ij} = \begin{bmatrix} E_{ij} & 0 \\ 0 & -E_{ji} \end{bmatrix} \tag{3.2}$$

We will need the following commutator relations:

$$[\, X_{ii}, Y_{ij} \,] = (1+\delta_i(j)\,)\, Z_{ij}, \qquad [\, Z_{ij}, X_{jk} \,] = (1+\delta_j(k)\,)\, X_{ik}$$

$$[\, Z_{ij}, Y_{jk} \,] = (-1-\delta_j(k)\,)\, Y_{ik} \tag{3.3}$$

Let g_1, g_2 be matrices of the form (2.6) with (1,1) entries $g_1(1,1) = x_1$ and $g_2(1,1) = x_2$ with $x_1 \neq x_2$, all other entries the same. Then $g_1^{-1}g_2$ and $g_1g_2^{-1}$ are in $S := <\text{supp } d\mu_{E,c}>$ and

$$g_1^{-1}g_2 = \text{id} + (x_1 - x_2)\, X_{11} = \exp\,(\,(x_1 - x_2)\, X_{11})$$

$$g_1g_2^{-1} = \text{id} + (x_1 - x_2)\, Y_{11} = \exp\,(\,(x_1 - x_2)\, Y_{11})$$

where x_1-x_2 can take values over some interval around 0, as the distribution of these potentials was assumed to be absolutely continuous. Therefore we conclude that X_{11}, Y_{11} are contained in the Lie algebra of $<\text{supp } d\mu_{E,c}>$,which we denote by $\mathfrak{A}$, and from (3.3), $Z_{11} \in \mathfrak{A}$.

By doing the same with the other "top surface" varibles, i.e., choose g_1, g_2 in G with $g_1(i,i) \neq g_2(i,i)$ for $i = 2, \ldots, m^{d-1}$, we see that X_{ii}, Y_{ii} and $Z_{ii} \in \mathfrak{A}$ for all $i = 1, \ldots, m^{d-1}$.

We are going to prove inductively that the whole system (3.2) is contained in $\mathfrak{A}$, using the following sublemma:

<u>Sublemma</u>:

Let $g_1, g_2 \varepsilon$ *G as before, that is with* $g_1(1,1) \neq g_2(1,1)$. *Then, for* $k < m$,

(*i*) $g_1 X_{k1} g_1^{-1}$ - $g_2 X_{k1} g_2^{-1}$ *is a linear combination of* Z_{1k}, $X_{k-1,1}$, $X_{k-m,1}, \ldots, X_{k-m^{d-1},1}, X_{k1}, X_{k+1,1}, X_{k+m,1}, \ldots, X_{k+m^{d-1},1}$.

(*ii*) $g_1^{-1} Y_{k1} g_1$ - $g_2^{-1} Y_{k1} g_2$ *is a linear combination of* Z_{k1}, $Y_{k-1,1}$, $Y_{k-m,1}, \ldots, Y_{k-m^{d-1},1}, Y_{k1}, Y_{k+m,1}, \ldots, Y_{k+m^{d-1},1}$.

We continue with the proof of the Lemma. Let us assume for simplicity that $d = 2$. Since $X_{11} \varepsilon\ \mathfrak{A}$, it follows from the sublemma, that $X_{12} + X_{m+1,1} \varepsilon\ \mathfrak{A}$. With the same reasoning, we also conclude that $Y_{12} + Y_{m+1,1} \varepsilon\ \mathfrak{A}$ and from (3.3) , $Z_{12} + Z_{1,m+1}$, $Z_{21} + Z_{m+1,1}\ \varepsilon\ \mathfrak{A}$. Then $[\ Z_{12} + Z_{1,m+1}, X_{22}\ ] = [\ Z_{12}, X_{22}\ ] = X_{12} \varepsilon\ \mathfrak{A}$. With the same reasoning $Y_{12} \varepsilon\ \mathfrak{A}$. Since now also $X_{m+1,1}$ and $Y_{m+1,1} \varepsilon\ \mathfrak{A}$, we conclude, using (3.3) again, that $Z_{1,m+1}, Z_{m+1,1} \varepsilon\ \mathfrak{A}$. It follows that $X_{m+1,m+1}$ and $Y_{m+1,m+1} \varepsilon\ \mathfrak{A}$. Using the same argument for all other $k \leq m$, we conclude that X_{1i}, Y_{1i}, X_{ii} and Y_{ii} are in $\mathfrak{A}$ for all $i \leq 2m$ and therefor, again by (3.3), X_{ij}, Y_{ij} and Z_{ij} are in $\mathfrak{A}$ for all $i, j \leq 2m$. Repeating this argument for all other k, shows that X_{1j}, Y_{1j}, X_{jj} and Y_{jj} are in $\mathfrak{A}$ for all $j \leq m^d$ and using (3.3) the same way as before shows that the whole system (3.2) is contained in $\mathfrak{A}$.

For general d, the exact same argument can be used if m is replaced by m^{d-1} in the previous steps ($X_{ii}, Y_{ii} \varepsilon\ \mathfrak{A}$ for $i = 1, \ldots, m^{d-1}$). □

proof of the sublemma:

(i) $$g_1 X_{k1} g_1^{-1} - g_2 X_{k1} g_2^{-1} =$$

$$\begin{bmatrix} (V_\omega(2)-V_\omega(1))\,\tilde{E}_{k1} & V_\omega(1)\tilde{E}_{k1}V_\omega(1)-V_\omega(2)\tilde{E}_{k1}V_\omega(2) \\ 0 & \tilde{E}_{k1}(V_\omega(1)-V_\omega(2) \end{bmatrix} \quad (3.4)$$

where $\tilde{E}_{k1} = E_{k1} + E_{1k}$.

$V_\omega(2) - V_\omega(1) = (x_2-x_1)\, E_{11}$, so

$(\, V_\omega(2) - V_\omega(1)\,)\, \tilde{E}_{k1} = (x_2-x_1)\, E_{11}\, \tilde{E}_{k1} = (x_2-x_1)\, E_{1k}$ for $k > 1$.

Also for $k > 1$:

$V_\omega(1)\, \tilde{E}_{k1}\, V_\omega(1) = A + A^T$, where

$$A = \begin{bmatrix} 0 & 0 & 0 & \\ \vdots & \vdots & \vdots & \\ x_1 & 1 & 1 & \\ \vdots & \vdots & \vdots & \\ x_1 & 1 & 1 & \\ x_1 c_k & c_k & c_k & \cdots \\ x_1 & 1 & 1 & \\ \vdots & \vdots & \vdots & \\ x_1 & 1 & 1 & \\ \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & \end{bmatrix}$$

where the c_k - terms appear in the k-th row, the columns 2, m+1, . . , $m^{d-1}+1$ are equal, the rows k-1, k-m, . . , $k\text{-}m^{d-1}$, k+1, k+m, . . , $k+m^{d-1}$ are equal.

A similiar formula holds with $V_\omega(2)$ in place of $V_\omega(1)$.

From there it follows that (3.4) becomes:

$$(x_2\text{-}x_1)\ Z_{1k} + (x_1\text{-}x_2) \sum_{i=0}^{d-1} (\ X_{k-m^i,1} + X_{k+m^i,1}\ ) \quad + (x_1\text{-}x_2)c_k\ X_{k1}$$

(ii) follows by the same arguments, just using transposes, since

$$g_1^{-1} Y_{k1} g_1 = [\ g_1 X_{k1} g_1^{-1}\ ]^T. \qquad \square$$

Fix E. Define a map ψ: $\mathbb{R}^{nm^d} \to G$, for given n, by

$$\psi\ (x_{11}, \ldots, x_{nm^d}) = \psi_n\ (x_{n1}, \ldots, x_{nm^d}) \cdots \psi_1\ (x_{11}, \ldots, x_{1m^d})$$

where

$$\psi_i\ (x_{i1}, \ldots, x_{im^d}) = g_\omega\ (i,E) =: g_i$$

with potentials $(x_{i1}, \ldots, x_{im^d})$.

Clearly, $\psi_i\ (x_{i1} + t, \ldots, x_{im^d}) = \exp\ (tX_{11})\ \psi_i\ (x_{i1}, \ldots, x_{im^d})$, so that

$$\frac{\partial}{\partial x_{i1}}\ \psi_i = X_{11}\ \psi_i$$

and a similiar formula for all other x_{ik}.

Let again $x_{ik} = c_k$ for all i and $k \varepsilon \{1, \ldots, m\}^d$, $k(d) < m$. Define ψ_c: $\mathbb{R}^{nm^{d-1}} \to G$

$$\psi_c\ (x_1, \ldots, x_{nm^{d-1}}) = \psi\ (x_{(n-1)m^{d-1}}, \ldots, x_{nm^{d-1}}, c_1, \ldots, c_h, \ldots, x_1, \ldots, x_{m^{d-1}}, c_1, \ldots, c_h)\ ,$$

$h := m^{d-1}(m\text{-}1)$.

Lemma 3.2:

(i) For all $c \varepsilon \mathbb{R}^{m^{d-1}(m-1)}$, $\exists$ *n such that* ψ_c *has maximal rank almost everywhere on* $\mathbb{R}^{nm^{d-1}}$.

(ii) If $d = 1$, *n of part (i) can be taken to be dim G.*

proof: (i) (see also Lacroix [1]) $\frac{\partial}{\partial x_{ij}} \psi_c = g_n \,.\,.\, X_{jj} g_i \,.\,.\, g_1 \;\varepsilon\; G_\psi$, the tangent space at $\psi = \psi_c(x_1, \,.\,., x_d)$, where $j = 1, \,.\,.\, , m^{d-1}$, $i = 1, \,.\,.\, , n$.

Multiplying by $g_1^{-1} \,.\,.\, g_n^{-1}$ maps this vector into

$$g_1^{-1} \,.\,.\, g_i^{-1} X_{jj} g_i \,.\,.\, g_1 = \mathrm{Ad}_{g_1 \,.\,.\, g_i} X_{jj} \;\varepsilon\; G_e \simeq \mathfrak{G}.$$

So we have to show that there exists some n such that the system

$$\{ < \mathrm{Ad}_{g_1} X_{jj}, \,.\,.\,.\, , \mathrm{Ad}_{g_1 \,.\,.\, g_n} X_{jj} > ; \; j = 1, \,.\,.\, , m^{d-1} \} = \mathfrak{A} \text{ for a.e. } x \varepsilon \mathbb{R}^{nm^{d-1}}.$$

Assume, that for $g_1, \,.\,.\, , g_k$, $\{ < \mathrm{Ad}_{g_1} X_{jj}, \,.\,.\, , \mathrm{Ad}_{g_1 \,.\,.\, g_k} X_{jj} > \} =: H$ is maximal in the sense that $\{ < H, \mathrm{Ad}_{g_1 \,.\,.\, g_{k+1}} X_{jj}, \,.\,.\, , \mathrm{Ad}_{g_1 \,.\,.\, g_n} X_{jj} > \} = H \neq \mathfrak{G}$, for all choices of $g_{k+1}, \,.\,.\, g_n$ and any n.

Let $H_1 = \{ < \mathrm{Ad}_{g_{k+1}} X_{jj}, \,.\,.\, , \mathrm{Ad}_{g_{k+1} \,.\,.\, g_n} X_{jj} > \}$. Then $< H, \mathrm{Ad}_{g_1 \,.\,.\, g_k} H_1 > = H$, so $\mathrm{Ad}_{g_1 \,.\,.\, g_k} H_1 \subset H$, for all H_1 of the predefined type. In particular with H in place of H_1. But then also $\mathrm{Ad}_g H \subset H$, for all transfer matrices g with potentials as prescribed. The set of those g generate the group G and therefor we conclude $\mathrm{Ad}_g H = H$ for all $g \varepsilon G$. This means that H is an ideal in $\mathfrak{G}$ and because H is nonempty $H = \mathfrak{G}$ follows from the simplicity of $\mathfrak{G}$, which proves (i), as the entries of $g_n \,.\,.\, g_1$ are polynomial.

(ii) To show that n can be taken to be dim G if d = 1, we note that if for l = dim G

$$\dim < \partial_1\psi_c, \ldots, \partial_l\psi_c > \neq \dim G, \quad \forall\, x \varepsilon\, \mathbb{R}^n$$

then, by the first part, there exists k such that for k ≥ l and almost all choices of $x \varepsilon\, \mathbb{R}^n$

$$\dim G > \dim < \partial_1\psi_c, \ldots, \partial_k\psi_c > = \dim < \partial_1\psi_c, \ldots, \partial_{k+1}\psi_c > - 1 \tag{3.5}$$

where the ∂_i stand for x_i derivatives and ψ_c is the n-fold product of transfer matrices, such that (i) holds.

Thus, for all choices of $x \varepsilon\, \mathbb{R}^n$, after multiplying by $g_{k+2}^{-1} \cdot\cdot g_n^{-1}$,

$$g_{k+1}\, X_{11} g_k \cdot\cdot g_1 \;\varepsilon\; < g_{k+1} X_{11} g_k \cdot\cdot g_1, \ldots, g_{k+1} \cdot\cdot X_{11} g_1 > \;=: H_0$$

$$\Leftrightarrow \qquad g_{k+1} X_{11} g_k \cdot\cdot g_1 g_0 \;\varepsilon\; H_0 g_0 \qquad \forall\, g_0$$

$$\Leftrightarrow \qquad X_{11} g_{k+1} \cdot\cdot g_1 \;\varepsilon\; < g_{k+1} \cdot X_{11} g_{k-1} \cdot\cdot g_1, \ldots, g_{k+1} \cdot\cdot X_{11} g_1 > \}$$

for all choices of $x \varepsilon\, \mathbb{R}^n$.

But this contradicts (3.5) and therefore, there exists $x \varepsilon\, \mathbb{R}^n$ such that

$$\dim < \partial_1\psi_c, \ldots, \partial_l\psi_c > = \dim G,$$

so n can be taken to equal l. □

The last result tells us, that there is, for fixed $c \varepsilon\, \mathbb{R}^{m^d - m^{d-1}}$, a distinguished index set $i_1k_1, \ldots, i_lk_l$ for $i_j \varepsilon \{1, \ldots, l\}$, l = dim G, $k_j \varepsilon \{1, \ldots, m\}^d$, $k_j(d) = m$, such that ψ is a diffeomorphism in those variables:

Corollary 3.1:

For a. e. $y_{ik} \varepsilon \mathbb{R}^{nm^d - l}$, $i = 1, \ldots, n$ $k \varepsilon \{1, \ldots, m\}^d$, $(i,k) \notin \{ (i_1,k_1), \ldots, (i_l,k_l) \}$, *the map* $\psi_y: \mathbb{R}^l \to G$, *defined by*

$$\psi_y (x_{i_1}, \ldots, x_{i_l}) := g_n \ldots g_1$$

where all variables in the g_i *different from* $x_{i_1}, \ldots, x_{i_l}$ *are denoted by* y_{ik}, *is a diffeomorphism in some neighbourhood of a.e. point in* $\mathbb{R}^l$.

proof: It follows from the preceding lemma, that for some fixed $x = (x_{i_1}, \ldots, x_{i_l}) \varepsilon \mathbb{R}^l$ and $y_{ik} = c_k$, $k \varepsilon \{1, \ldots, m\}^d$, $k(d) < m$ and $y_{ik} = c_{ik}$ for some c_{ik}, fo the "top surface" variables different from $x_{i_1}, \ldots, x_{i_l}$, ψ_y has maximal rank. Since the coordinates of ψ_y are polynomial in all variables, it follows that for a.e. $y \varepsilon \mathbb{R}^{nm^d - l}$, ψ_y has maximal rank at that fixed x. Therefore, for those y, ψ_y has maximal rank a.e. in $\mathbb{R}^l$. But then, for those y again, a.e. $x \varepsilon \mathbb{R}^l$ has a neighbourhood such that ψ_y is of maximal rank in this neighbourhood. □

Define now $B \subset \mathbb{R}^{nm^d - l}$ to be the set

$$B := \{ y_{ik} \varepsilon \mathbb{R}^{nm^d - l}, \psi_y \text{ has maximal rank a.e. in } \mathbb{R}^l \}.$$

Corollary 3.1 states that B has full measure, that is, for any bounded $A \subset \mathbb{R}^{nm^d - l}$, $A \cap B$ has the same Lebesgue measure as A. With the notation as in (1.6), the complement of B, B^c is the set $\{\alpha(y) = 0\}$ and in particular, (C1) implies that B^c and all translates along the diagonal have zero $d\eta$ - measure.

In the following, we will keep E fixed and therefor drop the index in $d\mu_E$ for notational convenience. If $d\eta$ denotes the joint distribution of the y_{ik} for all $(i,k) \notin \{ (i_1,k_1), \ldots, (i_l,k_l) \}$, then

$$(*)^n \, d\mu = \int (*)^n \, d\mu_y \, d\eta \, (y) \tag{3.6}$$

where

$$(*)^n d\mu_y = d\mu_{y_{nk}} * \ldots * d\mu_{y_{1k}}$$

and $d\mu_{y_{ik}}$ is the distribution of the transfer matrix with potentials, not having any of the distinguished indices, fixed. Let, for notational convenience, $x_j := x_{i_j k_j}$, $j = 1, \ldots, l$. Corollary 3.1 then says, that for $y \varepsilon B$,

$$(*)^n d\mu_y = G\,(g,y) dg$$

where dg is the Haar measure on G. Locally

$$dg = S(x)\, dx,$$

where S is C^∞ and $S \neq 0$ and thus, again locally,

$$G(g,y)\, dg = S(x)\, J_y^{-1}\,(x)\, F\,(\,\psi_y^{-1}\,(x)\,)\, dx\,, \tag{3.7}$$

where J_y is the Jacobian determinant of the diffeomorphism ψ_y and F is the joint distribution of the x_i.

We would like to see that for each $y \varepsilon$ B, G(g,y) is in some Sobolev space L^p_α, and that

$$\int \| G(g,y) \|_{L^p_\alpha} d\eta(y) < \infty.$$

It will be sufficient, since S is C^∞, to prove that J_y^{-1} F o $\psi_y^{-1}$is in some Sobolev space on $\mathbb{R}^l$, with

$$\int \| J_y^{-1} \text{ F o } \psi_y^{-1} \|_{L^p_\alpha} d\eta(y) < \infty.$$

Taking full derivatives with respect to any of the x_i of J_y^{-1} F o ψ_y^{-1}, will result in an additional J_y^{-1} factor, which we might not be able to control. Taking fractional derivatives should correspond to an additional fractional factor of J_y^{-1}, which is for small fractions integrable, since J_y is polynomial. This is the reason why we only take fractional derivatives of $(*)^n d\mu$, which we will later on add up by convolving in n-fold chunks.

Lemma 3.3:

For $A \subset \mathbb{R}^l$ and compact, χ_A the indicator function of A, there exists some $t > 0$, such that

$$\int \| \chi_A \, J_y^{-t}(x) \|_1 \, d\eta(y) \; < \; M$$

for some constant M. The 1-norm is on $\mathbb{R}^l$.

Note: The t in the Lemma above is the same for all E.

<u>proof</u>: J_y is polynomial in x and y, if $y \varepsilon$ B. Then, for fixed $y \varepsilon$ B, t sufficiently small

$$| \int \chi_A \, J_y^{-t}(x) \, dx \, | \; \leq \; \phi \, (y),$$

where $dx = dx_1 \, . \, . \, dx_n$ and ϕ is the reciprocal of a polynomial in y.

To see this, we note that, in general, also for $y \not\in$ B, $J_y = \tilde{\alpha}(y) \, \tilde{\beta}(x,y)$, both $\tilde{\alpha}$ and $\tilde{\beta}$ polynomial and there exists no y_0 such that $\tilde{\beta}(x,y_0) \equiv 0$ in x. It is then not hard to see that

$$| \int \chi_A \, \tilde{\beta}^{-t} \, dx| \leq c,$$

where c only depends on the support of $d\eta$ and on A.

It follows then readily that $\phi \, (y) = \tilde{c} \, \tilde{\alpha}^{-t}(y)$. The last step to reach the statement is then provided by our condition on $d\eta$. □

<u>Remark</u>: For $y \varepsilon$ B, $| \, \partial_i \, \psi_y^{-1} \, | \leq c \, (\, 1 + |\psi_y^{-1}|^d \,) \, J_y^{-1}$, $d = l^2$. (3.8)

This follows from the fact that the inverse of the Jacobian matrix is a polynomial of the Jacobian matrix, divided by the Jacobian determinant.

We are now going to use interpolation to see that G(g,y) is in some fractional Sobolev space, if F is.

Definition:

$X_0 := L^1\,(dx)$

$X_1 := \{\ F\varepsilon\ L^1\,(dx),\ \int ||\ \nabla_x\,[J_y^{-1}(x)\ F\circ\psi_y^{-1}(x)]\ ||_p\ d\eta(y) < \infty\ \}$,

where p is given in the context and the spaces are endowed with the obvious norms.

Denote, for $\alpha > 0$, by L^p_α the usual Sobolev space on $\mathbb{R}^k$, with k specified in the context, i.e., $F\varepsilon\ L^p_\alpha$ iff $(1+|t|^2)^{\frac{\alpha}{2}}\ \hat{F}\varepsilon\ L^q$, q the dual index to p.

Define for $y\varepsilon$ B the maps

$$T_y : F \rightarrow J_y^{-1}\ F\circ\ \psi_y^{-1},$$

where F is some function on $\mathbb{R}^l$. Then,

Lemma 3.4:

(*i*) *For all* $y\varepsilon\ B,\ T_y : X_0 \rightarrow L^1\,(dx\,)$ *and* $\int ||\ T_y\ F\ ||_1\ d\eta(y) < \infty$ *for* $F\varepsilon\ X_0$.

(*ii*) *For all* $y\varepsilon\ B,\ T_y : X_1 \rightarrow L^p_1\,(dx\,)$ *and* $\int ||\ T_y\ F\ ||_{L^p_1}\ d\eta(y) < \infty$ *for* $F\varepsilon\ X_1$

proof: Both statements follow immediately from the definition of X_0 and X_1. Indeed, if $X_{0,y} := T_y^{-1}\,(\,L^1)$ and $X_{1,y} := T_y^{-1}\,(\,L^p_1)$, then each f in X_i, i = 0,1, corresponds to some $\tilde{f}$ in $X_{i,y}$, such that

$$||\ f\ ||_{X_i} = \int ||\ \tilde{f}\ ||_{X_{i,y}}\ d\eta(y). \qquad \square$$

If F is in some fractional Sobolev space on $\mathbb{R}^l$ and has compact support, we can use

Lemma 3.3 to see that F will be in some of the interpolation spaces (X_0, $X_1)_t$, which we are going to identify next. We will then be able to use Calderon's interpolation theorem to show that our density G(y,g) is in some fractional Sobolev space. Before we identify the interpolation spaces, we state one result of Calderon, which will be useful to us.

Let us remark first that an interpolation pair (B_0, B_1) is a pair of Banach spaces B_0, B_1, continuously embedded in a topological vector space V.

Proposition 3.1: (Calderon [1])

If (B_0, B_1) and (C_0, C_1) are two interpolation pairs and L is a bounded, linear map from $B_0 \cap B_1$ to $C_0 \cap C_1$ for $i = 0, 1$, then L is a bounded, linear map from (B_0, $B_1)_t =: B_t$ to $C_t := (C_0, C_1)_t$.

In particular, if L is invertible and $L\ B_i = C_i$ for $i = 0, 1$, then $L\ B_t = C_t$.

Lemma 3.5:

$$(X_0, X_1)_t = \{ F\varepsilon\ L^1(dx),\ T_y F\varepsilon\ L_t^{p_t} \text{ for } y\varepsilon\ B \text{ and } \int \| T_y\ F \|_{L_t^{p_t}} d\eta\ (y) < \infty \}$$

where $\frac{1}{p_t} = (1\text{-}t) + t\,\frac{1}{p}$.

proof: The integrability with respect to $d\eta$ remains for all interpolation spaces, so it suffices to show that for each $y\varepsilon$ B, { T_y $F\varepsilon$ $L_t^{p_t}$ } are the interpolation spaces for { T_y $F\varepsilon$ L^1 } and { T_y $F\varepsilon$ L_1^p }

For $y\varepsilon$ B, T_y is an invertible map, with bounded inverse on the following spaces: $T_y^{-1}\ L^1 =: X_{0,y}$ and $T_y^{-1}\ L_1^p =: X_{1,y}$. Therefore, by Proposition 3.1, the interpolation spaces for the just defined spaces are

$X_{t,y} = T_y^{-1} L_t^{pt}$, since $L_t^{pt} = (L^1, L_1^p)_t$ (see, e.g., Taylor [1]). □

Note: $L_\alpha^1 \subset L_\beta^r$ holds for $r > 1$ and $\beta \leq \alpha$ as long as $1-\frac{1}{r} < (\alpha - \beta) \frac{1}{l}$.

For the following, we are going to fix β, r such that this relation holds for the α chosen in Theorem 1.

Lemma 3.6:

There exists $t > 0$, such that for $F \varepsilon L_\alpha^1 \subset L_\beta^r$, with supp $F \subset K$, K compact

$$F \varepsilon X_t.$$

proof: For $t < \beta$, $\hat{g}(\lambda) := (1+|\lambda|^2)^{\frac{t}{2}} \hat{F}(\lambda)$ with $g \varepsilon L^r$, supp g compact and

$$\| g \|_r = \| F \|_{L_t^r} .$$

For fixed $y \varepsilon B$,

$$| \partial_i J_y^{-1} F(\psi_y^{-1}) | \leq | (\partial_i J_y^{-1}) F \circ \psi_y^{-1} | + | J_y^{-1} \partial_i \psi_y^{-1} (\nabla F) \circ \psi_y^{-1} |$$

$$\leq | (\partial_i J_y) J_y^{-2} F \circ \psi_y^{-1} | + | J_y^{-1} (1+|\psi_y^{-1}|^d) J_y^{-1} (\nabla F) \circ \psi_y^{-1} |$$

$$\leq | (1+|\psi_y^{-1}|^d) J_y^{-2} F \circ \psi_y^{-1}| + | (1+|\psi_y^{-1}|^d) J_y^{-2} (\nabla F) \circ \psi_y^{-1} |$$

The product of the gradient of F and the derivative of ψ_y^{-1} in the first inequality is the inner product, the second inequality follows by (3.8).

Integrating with respect to dx and doing a change of variables, we obtain for fixed $y \varepsilon B$:

$$\| \nabla T_y F \|_p \leq \| (1+|x|^d) J_y^{-2+\frac{1}{p}} F \|_p + \| (1+|x|^d) J_y^{-2+\frac{1}{p}} \nabla F \|_p \tag{3.9}$$

If we let the operator S(z) be defined by

$$(S(z)\, F)^{\hat{}}\, (u) = (1+|u|^2\,)^{\frac{z}{2}}\, \hat{F}(u),$$

we can use interpolation between the spaces $\{\ F,\ S(1)\ F\varepsilon\ a^{-1}\ L^p\ \} =: Y_1$, where $a = (1+|x|^d)\ J_y^{-2+\frac{1}{p}}$ and $\{\ F,\ S(0)\ F\varepsilon\ L^1\ \} = L^1 =: Y_0$ again to conclude that

$Y_t = \{\ F,\ S(t)\ F\varepsilon\ a^{-t}\ L^{p_t}\ \}$

and therefore, $F\varepsilon\ X_t$ is implied by

$$\int ||\ [\ (1+|x|^d)\ J_y^{-2+\frac{1}{p}}(x)\ ]^t\ g\ ||_{p_t}\ d\eta\ (y) < \infty \tag{3.10}$$

Set $\frac{1}{p_t} = \frac{1}{s} + \frac{1}{r}$ for some s, and adjust p accordingly. Then

$$(3.10) \leq \int ||\ [\ (1+|x|^d)\ J_y^{-2+\frac{1}{p}}(x)\ ]^{ts}\ \chi_{supp\ g}(x)\ ||_1\ \ ||\ g\ ||_r\ d\eta(y)$$

$\leq c\ ||\ g\ ||_r\ \leq c\ ||\ F\ ||_{L^r_\beta}$, where the second to last inequality holds for small t by Lemma 3.3 □

We have now reached our goal, which follows in two ways after:

<u>Lemma 3.7</u>:

The norms on X_0 and X_1 are consistent on X_1.

proof: We have to show, that if $f_n \varepsilon X_0$ with $|| f_n ||_{X_0} \to 0$ and f_n is Cauchy in the X_1 norm, then $|| f_n ||_{X_1} \to 0$ and vice versa.

Clearly, for $f \varepsilon X_1$, $|| f ||_{X_0} \leq || f ||_{X_1}$, so if $|| f_n ||_{X_1} \to 0$, then also $|| f_n ||_{X_0} \to 0$. If f_n is Cauchy in the X_1 norm and $|| f_n ||_{X_0} \to 0$, then there exists f such that $|| f_n - f ||_{X_1} \to 0$. But then $|| f_n - f ||_{X_0} \to 0$, so $f = 0$. □

Corollary 3.2:

For all $y \varepsilon B$,

$$T_y: (X_0, X_1)_t \to L_t^{p_t}$$

and

$$\int || T_y F ||_{L_t^{p_t}} d\eta(y) < \infty,$$

where $(X_0, X_1)_t$ are the interpolation spaces of X_0 and X_1, $\frac{1}{p_t} = (1-t) + t \frac{1}{p}$ and $F \varepsilon X_t$.

proof: This is an immediate consequence of Lemma 3.5 or the Calderon- interpolation theorem (see, e.g., Reed- Simon [2] , p.37). □

Remark: It follows from the proof of Lemma 3.6, that for supp $F \subset K$

$$|| F ||_{X_t} \leq c || F ||_{L_\beta^r} \tag{3.11}$$

Lemma 3.6, combined with Corollary 3.2, tells us that for $y \varepsilon B$, F the distribution of the "top surface" potentials $x_1, \ldots, x_n$

$$T_y\ F = J_y^{-1}\ F \circ \psi_y^{-1} \varepsilon\ L_\alpha^p \text{ for some } \alpha > 0,\ p > 1.$$

Therefore $(*)^n d\mu_y \varepsilon\ L_\alpha^p$ locally for the same α, p.

Also, since $\| T_y\ F \|_{L_\alpha^p} = \| F \|_{X_{\alpha,y}} \leq c_0\ \tilde{\alpha}(y)^{-\alpha}\ \| F \|_{L_\beta^r}$, we see that

$$\int \| (*)^n\ d\mu_y \|_{L_\alpha^p}\ d\eta\ (y) \leq c_1\ \| F \|_{L_\beta^r}$$

for some c_1 by the condition on $d\eta$.

Since also, $\quad \| \int (*)^n\ d\mu_y\ \ d\eta(y)\ \|_{L_\alpha^p} \leq \int \| (*)^n\ d\mu_y \|_{L_\alpha^p}\ d\eta(y)$

(using Jensen's inequality), where derivatives are with respect to x only,we conclude

Corollary 3.3:

If $F \varepsilon\ L_\beta^r$, *with supp* $F \subset K$, K *compact, then*

$(*)^n d\mu \varepsilon\ L_\alpha^p$ *locally and*

$$\| (*)^n\ d\mu \|_{L_\alpha^p} \leq c\ \| F \|_{L_\beta^r}$$

for some $\alpha > 0$ *and* $p > 1$ *and constant* c.

Note: The α and p of the last Corollary are different from before. Here $\alpha \leq \beta$ and $\frac{1}{p} = (1\text{-}\alpha) + (1\text{-}\alpha)\ \frac{1}{r}$. Also, α and p are the same for all E, c depends on E.

§4 SMOOTHNESS OF THE INVARIANT MEASURE

The goal of this section is to show that for given k, there exists an l, depending on k, such that $\frac{\partial^k}{\partial E^k} (*)^l \, d\mu_E$ is a finite signed measure of bounded total variation. From this we will be able to conclude that $d\nu_E$ is smooth in $L^\infty(M)^*$, which proves Theorem 1 through Corollary 2.2. For simplicity, let n = dim G.

First we will need several results about Sobolev spaces on G. To define those, we are going to follow Simon - Taylor [1] and define the Laplace operator in the following way:

Let l_g and r_g denote left translation by g^{-1} and right translation by g respectively, i.e., $l_g(h) = g^{-1}h$, $r_g(h) = hg$ and acting on functions: l_g F (h) = F(l_gh), r_g F (h) = F(r_gh). A vector field is called left, right invariant, if it commutes with l_g, r_g respectively. Let $X_1, \ldots, X_n$ be a basis of G_e, the tangent space at the origin. Let $(\tilde{X}_i)_i$ be the unique set of left invariant vector fields with $\tilde{X}_i(e) = X_i$.

Then define

$$\Delta_l := \sum_{i=1}^{n} \tilde{X}_i^2 \tag{4.1}$$

The same procedure can be used with right translation to obtain right invariant vector fields and to define Δ_r.

If d_0 is any Riemannian metric on G, there exists a unique left invariant metric d_l on

G that coincides with d_0 at the identity, obtained by setting

$$d_l\,(X,X) = d_0\,(\widetilde{X},\widetilde{X}) \qquad \forall \text{ vector fields } X,$$

where $\widetilde{X}$ is the unique left invariant vector field with $\widetilde{X}$ (e) = X (e). Then

$$d_l\,(X(e),X(e)) = d_0\,(\widetilde{X}(e),\widetilde{X}(e)) = d_l\,(\tilde{X}(e),\tilde{X}(e))$$

and

$$d_l\,(dl_gX, dl_gX) = d_0\,(\widetilde{X},\widetilde{X}) = d_l\,(X,X),$$

so d_l is the desired left invariant metric.

If d_0 is the metric under which $(X_i)_i$ is an orthonormal basis of G_e, Δ_l is the Laplace operator associated to d_l. Similarly, we can find a unique right invariant metric d_r, such that Δ_r is the Laplace operator associated to d_r.

It was proven in Strichartz [1], that the Laplace operator is essentially selfadjoint on $C_0^\infty(G)$ if the underlying metric is complete. That we also have this property follows from

Lemma 4.1:

The metrics d_l, d_r are complete.

proof: For g,h ε G, $\text{dist}_l\,(g,h) = \inf\{\, L(\gamma) = \text{length of } \gamma \text{ in } d_l,\ \gamma: g \to h \,\}$

$$= \inf\{\, L(g^{-1}\gamma),\ g^{-1}\gamma: e \to g^{-1}h \,\}$$

$$= \text{dist}_l\,(e,g^{-1}h)$$

and the goedesics run for the same time, from where it follows that d_l is geodesically complete. The argument for d_r is identical. □

Lemma 4.2:

If $\Phi(g) = g^{-1}$, *then* $\Delta_{\#}\,\Phi = \Phi\,\Delta_{\#}$, *where* $\# = l$ *or* r.

proof: If γ is the geodesic between e and g with $\gamma(0) = e$ and $\gamma(1) = g$, then $\tilde{\gamma} := g^{-1}\gamma(1\text{-}t)$ is the geodesic between e and g^{-1}, from where it follows that Φ is an isometrie. Therefor Φ commutes with $\Delta_{\#}$, (see e.g. Helgason [1], p.246). □

After this preparation, we define Sobolev spaces on G the same way as on $\mathbb{R}^n$, namely as the image of L^p- spaces under the action of the Bessel potential $(1\text{-}\Delta)^{-\frac{\alpha}{2}}$.

Definition:

(*i*) *For* $1<p<\infty$, $\alpha>0$, L^p_α *is the set of all* $F \varepsilon\ L^p$ *such that*

$$F = (1\text{-}\Delta_l)^{-\frac{\alpha}{2}}\, G \text{ for some } G\varepsilon\ L^p, \text{ with norm}$$

$$\| F \|_{L^p_\alpha} = \| G \|_p$$

(*ii*) *For* $\alpha < 0$, *set* $L^q_\alpha := (L^p_{-\alpha})^*$, *where* $\frac{1}{p} + \frac{1}{q} = 1$

(*iii*) $H_s := L^2_s$

Remarks: (i) Integrations are with respect to Haar measure on G.

(ii) The L^p_α- norm will differ if we take Δ_r instead of Δ_l, however, no problems occur, if we stay in compact regions, see below.

Lemma 4.3:

(i) If F has support in a fixed compact, $F \varepsilon\ L^p_\alpha$ *iff* $F \varepsilon\ L^p_\alpha(\mathbb{R}^n)$ *in a local coordinate system about each point.*

(ii) If $d\mu(g) = F(g)\,dg$, $d\mu \varepsilon\ L^p_\alpha$ *iff* $F \varepsilon\ L^p_\alpha$

proof: (i) This follows immediately from the fact that on $\mathbb{R}^n$, L^p_α is the image of L^p under the action of the Bessel potential $(1-\Delta)^{\frac{\alpha}{2}}$ and

$$\Delta_l = \sum_i \widetilde{X}_i^2 = \sum_i \Big(\sum_k a_{ik} \frac{\partial}{\partial x_k}\Big)^2 \quad \text{locally, where the } a_{ik} \text{ are } C^\infty,$$

by exploiting a partition of unity. Therefore, since the a_{ik} are bounded on the compact support of F, $\|\,(1-\Delta_l)^{\frac{\alpha}{2}} F\,\|_p$ and $\|\,F\,\|_{L^p_\alpha(\mathbb{R}^n)}$ are equivalent norms for all F with support in a fixed compact.

(ii) is obvious. □

Lemma 4.4:

On fixed compacts, the norms

$$\|\,(1-\Delta_r)^{\frac{\alpha}{2}} F\,\|_p, \qquad \|(1-\Delta_l)^{\frac{\alpha}{2}} F\,\|_p,$$

$$\|\,(1-\Delta_r)^{\frac{\alpha-\beta}{2}} F\,\|_{L^p_\beta}, \qquad \|(1-\Delta_r)^{\frac{\beta}{2}} (1-\Delta_l)^{\frac{\alpha-\beta}{2}} F\,\|_p$$

are equivalent for $\alpha \geq \beta \geq 0$.

proof: Again,

$$\Delta_r = \sum_i \Big(\sum_k b_{ik} \frac{\partial}{\partial x_k} \Big)^2 \text{ , locally, where all } b_{ik} \text{ are } C^\infty.$$

Thus, it follows from the proof of Lemma 4.3, that for F with support in a fixed compact

$$\| (1 - \Delta_r)^{\frac{\alpha}{2}} F \|_p \leq c_0 \| F \|_{L^p_\alpha(\mathbb{R}^n)} \leq c_1 \| (1 - \Delta_l)^{\frac{\alpha}{2}} F \|_p$$

and a similiar statement with Δ_l and Δ_r interchanged. The statement for the other two norms follows the same way, using the additivity of the Bessel potential in the powers on $\mathbb{R}^n$. □

We will have to use the equivalence of the norms in Lemma 4.4 to show that convolution is smoothening. It is for that reason, that we restrict everything to compact domains.

Next we will see how much smoothness is lost by taking E- derivatives of the measure $d\mu_E$:

Lemma 4.5:

$$\frac{\partial^k}{\partial E^k} d\mu_E \, \varepsilon \, H_s \qquad \forall \; s < -\tfrac{n}{2} - k.$$

<u>proof</u>: Since $d\mu_E$ and all E- derivatives of $d\mu_E$ have compact support, it is enough to operate locally. Then, for $G \varepsilon\ C_0^\infty(U)$, U some coordinate neighborhood

$$\left| \frac{\partial^k}{\partial E^k} \int G\, d\mu_E \right| \leq c_0 \sum_{|\alpha| \leq k} \| D^\alpha G \|_\infty \leq c_1 \| (1 + |t|^2)^{\frac{k}{2}} \hat{G}(t) \|_1 \qquad (4.2)$$

by Hausdorff- Young. Then,

$$(4.2) \leq c_1 \| (1 + |t|^2)^{\frac{k+s}{2}} \|_2 \| (1 + |t|^2)^{-\frac{s}{2}} \hat{G}(t) \|_2$$

$$\leq c_2 \| G \|_{H_{-s}} \qquad \text{if } k+s < -\tfrac{n}{2}, \text{ or } s < -\tfrac{n}{2} - k \qquad \square$$

For the convolution to have a smoothing character, it will be necessary to pull the Laplace operator into the convolution. To illustrate how this can be done and where Lemma 4.4 comes into play, let us recall that for $F, G \varepsilon\ L^1$, the convolution is defined by

$$F * G(g) = \int F(gh^{-1})\, G(h)\, dh$$

$$= \int F(h)\, G(h^{-1}g)\, dh \qquad (4.3)$$

since the Haar measure dh is invariant under $h \rightarrow h^{-1}$. These equations can also be written as

$$F * G(g) = \int \left(r_{h^{-1}} F(g) \right) G(h)\, dh$$

$$= \int F(h) \Big(l_h G(g) \Big) dh \tag{4.4}$$

Therefore, since Δ_l, Δ_r are left-, right- invariant respectively, for $F, G \varepsilon\ C_0^\infty$

$$(1 - \Delta_r)^{\frac{\alpha}{2}} \Big(F * G \Big) = \Big((1 - \Delta_r)^{\frac{\alpha}{2}} F \Big) * G$$

$$(1 - \Delta_l)^{\frac{\alpha}{2}} \Big(F * G \Big) = F * \Big((1 - \Delta_l)^{\frac{\alpha}{2}} G \Big) \tag{4.5}$$

As we are going to take E- derivatives of convolutions of the measure $d\mu_E$, it will be necessary, having Lemma 4.5 in mind, to have convolutions of distributions in $\bigcup_s H_s$, where the union goes over all $s \varepsilon\ \mathbb{R}$, defined.

Noting that

$$\int F_1(g)\ F_2 * F_3(g)\, dg = \iint F_1(g)\, F_2(gh^{-1})\, F_3(h)\, dg\, dh$$

$$= \iint \tilde{F}_2 * F_1(g)\, F_3(g)\, dg$$ and also, using the second relation of (4.3)

$$= \iint F_1 * \tilde{F}_3(g)\, F_2(g)\, dg, \quad \text{where } \tilde{F}(g) := F(g^{-1}) \tag{4.6}$$

we can define the convolution of a compactly supported distribution T with a C_0^∞ function F by

$$(F * T)(G) = T(\tilde{F} * G)$$

and $$(T * F)(G) = T(G * \tilde{F}),$$

which allows us to define the convolution of any compactly supported distributions

S,T by

$$S * T (F) = T (\tilde{S} * F), \quad F \varepsilon C_0^\infty,$$

where $\tilde{S} (F) = S (\tilde{F})$.

Remarks:(i) If T is a compactly supported distribution, $F \varepsilon C_0^\infty$, then $T * F$ and $F * T$ are in C_0^∞.

(ii) If $d\mu_1 = F_1 \, dg$ and $d\mu_2 = F_2 \, dg$, then $d\mu_1 * d\mu_2 = (F_1 * F_2) \, dg$.

(iii) $|| \tilde{F} ||_{L_s^p} = || F ||_{L_s^p}$ follows from Lemma 4.2.

Lemma 4.6:

If $d\mu = F \, dg$, supp F compact, $F \varepsilon L_s^p$, $0 < s < 1$, then

$$|| d\mu * G ||_{H^{s+t}} \leq c \, || d\mu ||_{L_s^p} \; || G ||_{H^t} \qquad \textit{if supp } G \textit{ compact},$$

i.e. convolution with $d\mu$ defines a bounded map from H_c^t to H_c^{s+t}, $\forall$ t, where

$$H_c^t := \{ G \varepsilon H^t, \textit{ supp } G \textit{ compact}\}.$$

proof: $d\mu * G = F * G \, dg$ and $|| d\mu ||_{L_s^p} = || F ||_{L_s^p}$, so this is just Theorem A.2.2 of Simon-Taylor [1]. □

Lemma 4.7:

For arbitrary s,t and compactly supported $F \varepsilon\ H^s$, $G \varepsilon\ H^t$:

$$\| F * G \|_{H^{s+t}} \leq c \| F \|_{H^s} \| G \|_{H^t}$$

i.e. convolution by compactly supported H^s *distributions defines a bounded map*

$$H_c^s \colon H_c^t \rightarrow H_c^{s+t} \quad \forall\ s,t$$

proof: This is Theorem A.2.3 of Simon-Taylor [1]. □

These two results will enable us to see, that taking E- derivatives up to a given order k of $(*)^l d\mu_E$, with l large and depending on k, will give us a well behaved distribution, namely a signed measure with a compactly supported C^1 density.

Lemma 4.8:

For given k,s > 0, there exists an l such that

$$\frac{\partial^i}{\partial E^i} (*)^l d\mu_E\ \varepsilon\ H_s \qquad \forall i \leq k\,.$$

proof: We know from section 3 that $(*)^n d\mu_E \varepsilon\ L_\alpha^p$ for some $\alpha > 0$ and $p > 1$.

For any $l \geq 2$:

$$\frac{\partial}{\partial E}(*)^l d\mu_E = \sum_{j=0}^{l-1} (*)^j d\mu_E * \frac{\partial}{\partial E} d\mu_E\ (*)^{l-j-1} d\mu_E \tag{4.8}$$

By Lemma 4.5, $\frac{\partial}{\partial E} d\mu_E$ decreases the Hölder index by at most $\frac{n}{2}+1$. If $j < n$, the first factor in (4.7) may not give any smoothness, but the Hölder index will be decreased by at most $(n-1)\frac{n}{2}$ by those j factors, thus the first $(j+1)$ factors in the sum of (4.8) decrease the Hölder index by at most $(n-1)\frac{n}{2}+\frac{n}{2}+1 = \frac{n^2}{2} + 1$. Therefore, by adding up the worst case scenario for the Hölder indices,

$$\frac{\partial}{\partial E}(*)^l \, d\mu_E \, \varepsilon \, H_s \quad \text{if } l > \frac{n}{\alpha}(s+\frac{n^2}{2}+1)+n+1.$$

For higher order derivatives, repeat this argument for the inividual terms in the sumands of (4.8) to conclude the statement. □

Corollary 4.1:

For given k and s, there exists l, large enough, such that

$$\frac{\partial^i}{\partial E^i} (*)^l d\mu_E(g) \;=\; G_i(g,E) \, dg \qquad i = 1, \ldots ,k$$

with $G_i \, \varepsilon \, C_0^s$

proof: By the previous Lemma there exists an l such that

$$\frac{\partial^i}{\partial E^i} (*)^l d\mu_E \, \varepsilon \, H_{\frac{n}{2}+s} \, , \quad \forall \, i \leq k.$$

Thus $\frac{\partial^i}{\partial E^i}(*)^l d\mu_E(g) = G_i (g,E) \, dg$, $i \leq k$, with $G_i \, \varepsilon \, H_{\frac{n}{2}+s}$ by Lemma 4.3 (ii) . By

Sobolev's Lemma, the G_i are C^s- functions in the usual sense. □

Corollary 4.2:

$d\mu_E$ has a unique invariant measure $d\nu_E$ on M.

proof: Take l such that $(*)^l d\mu_E(g) = G_E(g)dg$, with G a C^1 function. It follows from Furstenberg's theorem, that $(*)^l d\mu_E$ has a unique invariant measure. Thus $d\mu_E$ has a unique invariant measure. □

This last result will enable us to view the l- fold convolution of $d\mu_E$ as a differentiable, (in E), operator on the dual space of C(M), the set of continuous functions on M, which is a Banach space under the $|| \bullet ||_\infty$ - norm, as M is compact. The order of differentiability will depend on l.

The set $C(M)^*$ is just the set of finite signed measures on M (Riesz- Markov theorem). Let now $T_E: C(M)^* \to C(M)^*$ be defined by

$$T_E d\nu := d\mu_E * d\nu.$$

T_E defines clearly a bounded operator on $C(M)^*$.

<u>Lemma 4.9</u>:

T_E^n is a compact operator on C(M), ∀ E.*

<u>proof</u>: We know that $(*)^n d\mu_E(g) = G(g,E)\, dg =: d\mu_E^n$. $x \in M$ can be identified with a coset $k\,a\,n\,H = k\,H$, with $k \in \mathcal{K}$, $a \in \mathcal{A}$, $n \in \mathcal{N}$ (see Lemma 2.2). For $g \in G$, gx is the the coset $g\,k\,H = \tilde{k}(g,x)\,\tilde{u}(g,x)\,H = \tilde{k}(g,x)\,H$ with uniquely determined $\tilde{k} \in \mathcal{K}$ and $\tilde{u} \in \mathcal{AN}$. The Haar measure dg on G can be written as $dg = dk\, d_l u$ where dk is the Haar measure on $\mathcal{K}$ and $d_l u$ is the left- invariant measure on $\mathcal{AN}$ (see Helgason [1], p.94). Therefor we have for $f \in C(M)$

$$\left(T_E^n \right)^* f(x) = \int f(gx)\, d\mu_E^n(g) = \int\int f(gx)\, G(g,E)\, dg$$

$$= \int\int f(\tilde{k}\,H)\, G(\tilde{k}\,\tilde{u}\,k^{-1})\, d_l\tilde{u}\, d\tilde{k}$$

$$= \int \left(\int G(\tilde{k}\tilde{u}k^{-1})\, d_l\tilde{u} \right) f(\tilde{k}H)\, d\tilde{k}$$

so $\left(T_E^n \right)^*$ has, w.l.o.g., a bounded integral kernel with a compact space as its domain and is therefor compact (Reed- Simon [1], Thm. VI.23). Thus T_E^n is compact. □

Because the integrand in (2.9) is not in C(M) but in $L^\infty(M)$, we will restrict T_E to the dual of $L^\infty(M)$. $L^\infty(M)^*$ is the set of all signed measures with bounded total variation. In particular $L^\infty(M)^*$ is contained in $C(M)^*$. Since T_E is convolution by a probability measure, T_E maps $L^\infty(M)^*$ into itself.

Corollary 4.3:

T_E^n restricted to $L^\infty(M)^*$ has 1 as an isolated eigenvalue.

proof: It follows from Corollary 2.1 and the remark at the end of section 2, that T_E has eigenvalue 1 if restricted to $L^\infty(M)^*$, since $d\nu_E$ is a probability measure and therefor in $L^\infty(M)^*$. If 1 would not be an isolated eigenvalue, it wouldn't be isolated for T_E^n on $C(M)^*$, as $L^\infty(M)^*$ is contained in $C(M)^*$, contradicting Lemma 4.9. □

If we want to conclude smoothness in E for $d\nu_E$, which will be done by showing smoothness of the eigenprojections, we will have to show that the eigenvalue 1 has geometric and algebraic multiplicity one on $L^\infty(M)^*$. So far we only know there is a unique normalized positive functional that is an eigenvector for 1.

Lemma 4.10 :

(i) The geometric multiplicity of 1 is one, i. e., $d\nu_E$ is the unique normalized eigenvector for 1 in $L^\infty(M)^*$.

(ii) The algebraic multiplicity of 1 as an eigenvalue of T_E^n is one.

proof: (i) $d\nu \in L^\infty(M)^*$ has a unique Jordan decomposition $d\nu = d\nu_+ - d\nu_-$, where $d\nu_+, d\nu_-$ are both finite positive measures. If $T_E\, d\nu = d\nu$, then

$T_E\, d\nu = T_E\, d\nu_+ - T_E\, d\nu_- = (T_E\, d\nu)_+ - (T_E\, d\nu)_-$.

Because of the uniqueness of the decomposition and the fact that T_E applied to a positive measure will give a positive measure, we have

$$T_E\, d\nu_+ = d\nu_+ = c_+\, d\nu_E \text{ and } T_E\, d\nu_- = d\nu_- = c_-\, d\nu_E$$

for positive constants c_+, c_-.

Therefor $d\nu = (c_+ - c_-)\, d\nu_E$, so $d\nu$ equals $d\nu_E$ after normalizing.

(ii) $L^\infty(M)^* = F + N$, the topological direct sum of a closed F and a finite dimensional N, such that T_E^n leaves F and N invariant (Dieudonné [1], Thm. 11.4.1). If the algebraic multiplicity of 1 is k, k is the smallest integer such that $(T_E^n - 1)^k$ restricted to N is 0 and $(T_E^n - 1)^{k-1} N = E(1)$, where E (1) is the eigenspace of 1. This means in this case, if $k > 1$, that there is a $d\eta \varepsilon L^\infty(M)^*$ (actually in $(T_E^n - 1)^{k-2} N \supset E(1)$) with $T_E\, d\eta - d\eta = c\, d\nu_E$, for some constant $c \neq 0$. But $\int T_E^n\, d\eta - d\eta = 0 \neq \int c\, d\nu_E$, so the statement follows. □

For given k, it follows from Corollary 4.1 that for l large enough $\frac{\partial^i}{\partial E^i}(*)^l d\mu_E(g) = G_i(g,E)\, dg$ with $G_i \varepsilon C_0$ in g for $i \leq k$. Convolution with these signed measures map $L^\infty(M)^*$ into itself. Therefor T_E^l is C^k in E as an operator on $L^\infty(M)^*$ with $\frac{\partial^i}{\partial E^i} T_E^l$ the convolution by $G_i(g,E)\, dg$ for $i \leq k$.

<u>Lemma 4.11</u>:

$d\nu_E$ is C^∞ in E as an element of $L^\infty(M)^*$.

<u>proof</u>: For given k, T_E^l is C^k in E for l large enough and therefor

$$P_E := -\frac{1}{2\pi i} \int_\Gamma R(\xi, E)\, d\xi$$

is C^k in E, where Γ is a circle around 1, only containing this one spectral point of T_E^l

and $R(\xi, E) = (T^l_E - \xi)^{-1}$, which is bounded for $\xi \varepsilon \Gamma$ and C^k in E. P_E is the projection onto the subspace N of the last Lemma, so onto the eigenspace of 1, which is spanned by $d\nu_E$. Therefor $d\nu_E$ is C^k in E as an element of $L^\infty(M)^*$. Since k was arbitrary and the final conclusion is independent of l, the statement follows. □

This finally allows us to employ Corollary 2.2 to conclude the proof of Theorem 1.

As we mentioned in the introduction, we also get a localisation result. This is a consequence of Lemma 3.1, for the case where $x_{ik} = c_k$, $i \varepsilon \mathbb{Z}$ and $k \varepsilon \{1, \ldots, m\}^d$ with $k(d) < m$ and x_{ik}, i and k as before, just $k(d) = m$, are i.i.d. random variables with absolutely continuous distribution. This follows from the positivity of the smallest Lyapunov exponent associated to the transfer matrices, which is true if a certain number of convolutions of the measure on the transfer matrices is absolutely continuous with respect to Haar measure on G.

More precisely, since

$$\int \log \| g \| \, d\mu_c(g) < \infty,$$

it follows from Oseledec's theorem, (see e.g. Walters [1]), that there are, for almost all ω, real numbers $\gamma_1, \ldots, \gamma_{2m^d}$, numbered in increasing order, with $\gamma_i = -\gamma_{2m^d-i+1}$ and a sequence of subspaces W_i, $i = 1, \ldots, 2m^d$ of $\mathbb{R}^{2m^d}$ with $W_i \subset W_{i+1}$, $W_{2m^d} = \mathbb{R}^{2m^d}$, such that for $u \varepsilon W_{i+1} \setminus W_i$

$$\lim_{n\to\infty} \tfrac{1}{n} \log \| g_n \cdots g_0 u \| = \gamma_i$$

and $W_{i+1} = W_i$ iff $\gamma_i = \gamma_{i+1}$.

The γ_i depend on E.

The following result, extracted out of different works on the asymptotic behaviour of products of random matrices, will enable us, supplemented by Lemma 3.1, to obtain localisation for our special case.

Proposition 4.1: (see Furstenberg [1], Bougerol - Lacriox [1], Delyon et al. [1], Guivarc'h - Raugi [1], Sazanov - Tutubalin [1])

(*i*) *If there exists an integer n, such that* $(*)^n d\mu_c$ *has a* C^0*- density with respect to Haar measure on G, then all Lyapunov exponents are different.*

(*ii*) *If* $\gamma_{m^d} \neq 0$, *for all E,* H_ω *has almost surely pure point spectrum with exponentially decaying eigenfunctions. The rate of this decay is given by* $|\gamma_{m^d}|$.

The consequence of Lemma 3.1 is that the condition of part (i) is satisfied for some n, since the distribution of $g_n \ . \ . \ g_1$ is just the push forward of the distribution $F(x_{m^{d-1}}) \ . \ . \ F(x_1) \ dx_1 \ . \ . \ dx_{m^{d-1}}$ under ψ_c, for any $c \varepsilon \mathbb{R}^{m^d - m^{d-1}}$. Therefore, we conclude that $\gamma_{m^d} \neq 0$, for all E, in this case and the conclusion of part (ii) hold in our case.

§5 A PROOF OF THEOREM 2

§5.1 Part (iii):

The main part of the proof of (iii) of Theorem 2 is the "construction" of Lemma 3.2 . Denote for simplicity the top potential in g_n by x_n, the bottom potential by b_n. Since the measure on the b_i's is now pure point, but the result of Corollary 3.1 is only true a.e. in the b_i's, we need to prove that for all realizations of b-variables, the map $\psi_b(x)$ defined in Corollary 3.1 has maximal rank a.e. in x. For this we will need more than 10 = dim Sp(2) convolutions; to be precise, our proof uses 58. After this, the result follows essentially from the corresponding results of Section 3. Again, for any n and fixed E,

$$(*)^n d\mu = \int d\eta(b)\ (*)^n d\mu_b\ , \quad \text{with the notation as in Section 3}$$

$$= \sum_{(b)} c_b\ (*)^n d\mu_b$$

where the sum goes over all 2^n possible Bernoulli states $(b_1, \ldots, b_n)$ and c_b are the corresponding probabilistic weights.

Therefor, if we can show that the result of Corollary 3.1 holds for some n for all b and a.e. x, we can employ all the results thereafter to obtain Theorem 2.

Lemma 5.1:

There exists an integer n such that for each realization of $(b_1, \ldots, b_n)$, *the map* ψ_b *has maximal rank at a.e. point.*

Note: This also implies the same statement for $(E\text{-}b_1, \ldots, E\text{-}b_n)$ for all values of the (b_i) and all E. For that reason, we don't have to consider the E-dependence anymore, all conclusions are valid for all E.

We will postpone the proof of this result for a while.

It follows from this Lemma that for each realization of $(b_1, \ldots, b_n)$, there exists a set of indices $i_1, \ldots, i_{10}$ such that

$$(\,\partial_{i_k}\,\psi_b\,)_k \quad k = 1, \ldots, 10$$

is an independent set for a.e. $x \varepsilon\ \mathbb{R}^n$. We will now fix the x_i's with $i \notin \{i_{i_1}, \ldots, i_{i_{10}}\}$ and go through the same considerations as in Section 3. Denote the map resulting from keeping b and all the x_i, $i \notin \{i_{i_1}, \ldots, i_{i_{10}}\}$ fixed by $\tilde{\psi}_b$: $\mathbb{R}^{10} \rightarrow$ G. Then

<u>Corollary 5.1</u>:

For all $b \varepsilon\ \mathbb{R}^n$ *and a.e.* $x \varepsilon\ \mathbb{R}^{n-10}$ *where* $x = (x)_i$ *with indices i restricted as above, the map* $\tilde{\psi}_b$ *is a diffeomorphism for a.e.* $(x_{i_1}, \ldots, x_{i_{10}})$.

<u>Note</u>: We fix $b \varepsilon\ \mathbb{R}^n$ first. The distinguished index set then depends on b.

<u>proof</u>: The argument is identical to the one in the proof of Corollary 3.1. □

We can write:

$$(*)^n d\mu_b = \int d\tilde{F}(x)\ (*)^n d\tilde{\mu}_b \tag{5.1}$$

where $d\tilde{\mu}_b$ is the measure on the transfer matrices with x_i fixed, i as above, and $d\tilde{F}$ is the joint distribution of those variables x_i, with $d\tilde{F} = F^{n-10}dx$.

This is of the same form as the measure described in (3.6) with $d\tilde{F}$ in place of $d\eta$. Locally

$$(*)^n d\tilde{\mu}_b = S(\tilde{x})\ \tilde{J}_b^{-1}(\tilde{x})\ F \circ \tilde{\psi}_b^{-1}(\tilde{x})$$

where $\tilde{J}_b = \det D\tilde{\psi}_b$ and $\tilde{x} = (x_{i_1}, \ldots, x_{i_{10}})$.

Since $d\tilde{F}$ satisfies (C1), the exact same arguments as in Section 3 apply now to (5.1) for each b and a.e. x (the x now takes the role of the y in Section 3 for the proofs) and we can conclude

Corollary 5.2:

(i) $(*)^n d\mu_b \varepsilon\ L^p_\alpha$ *for some* $\alpha>0$, $p>1$ *and for all b*

(ii) For all E, $(*)^n d\mu \varepsilon\ L^p_\alpha$ *for some* $\alpha>0$, $p>1$

Since also all the results of Section 4 apply, we conclude Theorem 2, part (iii).

Let us now turn to the proof of Lemma 5.1. The strategy will be to show constructively that $Y_{11}+Y_{12}+Y_{22}$ and $Y_{11}-Y_{12}+Y_{22}$ with the notation as in Section 2, are contained in

$$\mathfrak{A} := \langle Y_{11}, \ldots, Ad_{g_1 \ldots g_n} Y_{11} \rangle \text{ for some n,}$$

from where we conclude that all $Y_{ij} \varepsilon\ \mathfrak{A}$. Then we do the same with the X_{ij} and finally, we show that all Z_{ij} are in $\mathfrak{A}$, which means $\mathfrak{A} = \mathfrak{G}$, the desired result. We need to do some preliminary work:

Let us first note some useful facts for computations:

$$\text{If } g \varepsilon\ G,\ g = \begin{bmatrix} A & B \\ C & D \end{bmatrix}, \text{ then } g^{-1} = \begin{bmatrix} D^T & -B^T \\ -C^T & A^T \end{bmatrix} \tag{5.2}$$

$$g^{-1}Y_{11}g = \begin{bmatrix} -B^TE_{11}A & -B^TE_{11}B \\ A^TE_{11}A & A^TE_{11}B \end{bmatrix} \tag{5.3}$$

In proofing Lemma 5.1 the way we outlined it, we have to show that equations of the form $Ad_H\ Y_{11} = X$ with H a product of transfer matrices are solvable in H for certain $X\varepsilon\ \mathfrak{G}$. The next four results prepare for this.

<u>Lemma 5.2</u>:

If $g = g_4\ .\ .\ g_1$, then the first row of g consists of 4 independent polynomials in the x_i's for all b_i, $i = 1, . . ,4$.

<u>proof</u>: Let $p_1, . . ,p_4$ denote the entries of the first row of g. If for α_i not all 0, $\sum\alpha_i\ p_i \equiv 0$ in x then $\alpha_1 = 0$, since it is the only factor of $x_1\ . .x_4$.

Also:

$$x_1x_2x_3(\alpha_2-\alpha_3)+x_1(\alpha_2b_3b_4-\alpha_2-\alpha_4b_3)+x_1x_2(\alpha_2b_4-\alpha_4)-\alpha_2b_4+\alpha_2b_2b_3b_4-\alpha_2b_2+\alpha_4$$
$$-\alpha_4-\alpha_4b_2b_3 = 0,$$

from where $\alpha_i = 0$ for all i follows by comparing coefficients, independent of the b_i. □

<u>Lemma 5.3</u>:

For constant α_i, $i = 1, . . ,4$ with $\alpha_1 \neq 0$, let, for $g = g_4\ .\ .g_1$,

$$q_k := \sum \alpha_i\ g(k,i)\ ,\ \textit{i.e., a weighted row sum.}$$

Then the q_k are independent polynomials in the x_i, for all b_i.

proof: The argument is the same as above, just comparing coefficients.

Note: It also follows easily that any nontrivial linear combination of the q_i's defined above is nonconstant in the x_i for all b_i.

Corollary 5.3:

Let H be a product of transfer matrices with the sum of the 3rd and 4th column equal to $(\alpha_1, \ldots, \alpha_4)^T$ *and* $\alpha_1 \neq 0$. *Then, the entries of the sum of the 3rd and 4th row of gH, with g as before, are independent polynomials in the* x_i *for all* b_i.

proof: Noting that the i-th entry in the sum of the 3rd and 4th row of gH is given by α_1 g(i,1) + . . + α_4 g(i,4), this follows from Lemma 5.3. □

Corollary 5.4:

Let g, H be as above. Any given nontrivial linear combination of the entries of the sum of the 3rd and 4th row of gH is nonconstant in the x_i *independent of the* b_i.

proof: This follows from the Note after Lemma 5.3 and the specific form of the entries given in the previous proof. □

We are now in the position to build up $\mathfrak{G}$ by going to higher and higher products of transfer matrices.

Lemma 5.4:

Given H, a product of transfer matrices, with $H(3{+}4,1) \neq 0$, *where* $H(3{+}4,1)$ *denotes the first entry of the sum of rows* 3 *and* 4 *of H.*

Then there exist g_i, $i = 1, . . ,8$ *such that for* $g = g_8 \, . \, . \, g_1$

$$Y_0 := Y_{11} + Y_{12} + Y_{22} = Ad_{Hg} Y_{11} = H^{-1} g^{-1} Y_{11} g H$$

The b-values in both H and g are arbitrary.

proof: Let $\tilde{H} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ and assume for Y_0 that

$$Ad_{\tilde{H}^{-1}} Y_0 = Y_{11}, \tag{5.4}$$

from where we will solve backwards for $\tilde{H}$. Then we will show that we can solve for $gH = \tilde{H}$ from where the statement follows. (5.4) leads to the conditions

$$B(1,1)+B(1,2) = 0, \;\; B(2,1)+B(2,2) = 0$$

$$D(1,1)+D(1,2) = 1, \;\; D(2,1)+D(2,2) = 0$$

which means that the sum of the 3rd and 4th row of $\tilde{H} = (0,0,1,0)^T$.

Solving then for $\tilde{H} = gH$ leads to

$$(g_6 \, . \, . \, g_1 H)(3{+}4,\bullet) = (g_8 g_7)^{-1} \, (0,0,1,0)^T = \text{3rd row of } (g_8 g_7)^{-1}$$

$$= (-x_7,-1,-x_7 x_8,-x_7-b_8) \tag{5.5}$$

Set $H_0 := g_4 \, . \, . \, g_1 H$. Then

$(g_6 g_5 H_0)(3+4,\bullet) =$

$$\begin{bmatrix} x_5 x_6 h_1^0 + (x_5 + b_6) h_2^0 - x_5 h_3^0 - h_4^0 \\ x_6 h_1^0 + b_5 h_1^0 + b_5 b_6 h_2^0 - h_3^0 - b_5 h_4^0 \\ x_6 \; h_1^0 + h_2^0 - h_3^0 \\ h_1^0 + b_6 h_2^0 - h_4^0 \end{bmatrix} \tag{5.6}$$

where the h_i^0 are the entries of $H_0(3+4,\bullet)$. Setting (5.4) equal to (5.5) yields

$$x_6 = \frac{-1 - b_5 h_1^0 - b_5 b_6 h_2^0 + h_3^0 + b_5 h_4^0}{h_1^0}, \quad x_5 = \frac{b_8 - b_6 h_2^0 + h_1^0 + b_8 h_2^0}{-1 - b_5 h_1^0 - b_5 b_6 h_2^0 + h_3^0 + b_5 h_4^0 + h_2^0 - h_3^0}$$

$$x_7 = -(h_1^0 + b_6 h_2^0 - h_4^0) - b_8, \qquad x_8 = -\frac{1}{x_7} \left(x_6 h_1^0 + h_2^0 - h_3^0 \right) \tag{5.7}$$

From Corollary 5.3, we can conclude that all the denominators in (5.6) are nonzero at appropriate $x_1, . . , x_4$ for given b-values. □

<u>Remark</u>: If $g = g_4 \, . \, . \, g_1$, the first row of g is a set of independent polynomials in the x_i by Lemma 5.2. If H is any product of transfer matrices, $(gH)\ (3\pm4,1)$ is a nontrivial linear combination of the first row entries of g, and so nonzero for an appropriate choice of x_i, $i = 1, . . ,4$.

Whenever a condition like $H(3\pm 4,1)\neq 0$ is required, we will multiply H by an appropriate g to satisfy this condition. The same is valid for the sum or difference of rows 1 and 2.

Lemma 5.5:

Given H, a product of transfer matrices with $H(3\text{-}4,1) \neq 0$, there exist, for any given set of b-values, $g_1, \ldots, g_8$ such that

$$Y_{11}-Y_{12}+Y_{22} = Ad_{Hg_1 \,..\, g_8}\, Y_{11}$$

proof: With a similiar notation as in the previous proof, we have to solve for $gH = \tilde{H}$ with $\tilde{H}(3\text{-}4,\bullet) = (0,0,1,0)^T$. The proof is then virtually identical. □

Lemma 5.6:

Given H, a product of transfer matrices with $H(1\pm 2,1) \neq 0$, there exists, for any given set of b-values, $g_1, \ldots, g_8$ such that

$$X_{11} \pm X_{12} + X_{22} = Ad_{Hg_1 \,..\, g_8}\, Y_{11}$$

proof: Let $X_0 := X_{11} + X_{12} + X_{22}$ and assume that

$$Ad_{\tilde{H}^{-1}}\, X_0 = Y_{11}.$$

This leads to $\tilde{H}(1+2,\bullet) = (0,0,1,0)^T$ and one can use the same arguments as in the proof of Lemma 5.4 to obtain the result. The case with the '-' sign is identical. □

To get the Z_{ij} , we need the following

Lemma 5.7:

There exist $g_1, \ldots, g_4$ *such that the top left blocks of* $Ad_{g_1} Y_{11}, \ldots, Ad_{g_1 \cdots g_4} Y_{11}$ *viewed as elements of* $\mathbb{R}^4$, *are independent for any given* $b_1, \ldots, b_4$.

proof: Let $x_3 = x_4 = 0$. Then the top left blocks of the prescribed matrices are of the following form (using (5.3)):

$$Ad_{g_1} Y_{11} : \begin{bmatrix} x_1 & 1 \\ 0 & 0 \end{bmatrix}, \quad Ad_{g_1 g_2} Y_{11} : \begin{bmatrix} x_2 x_1 + x_2 & x_2^2 + b_1 x_2 \\ x_2 x_1 + 1 & x_2 + b_1 \end{bmatrix}$$

$$Ad_{g_1 g_2 g_3} Y_{11} : \begin{bmatrix} -x_1 x_2 b_2 & -x_1 x_2 b_2 (b_1 b_2 - 1) \\ (-x_1 - b_2)(b_2) & (-x_1 - b_2)(b_1 b_2 - 1) \end{bmatrix}$$

$$Ad_{g_1 g_2 g_3 g_4} Y_{11} : \begin{bmatrix} b_2(b_3 x_1 + b_3 b_2 + 1) & b_2(b_1 b_2 b_3 - b_1) \\ (b_1 b_2 - 1)(b_3 x_1 + b_2 b_3 + 1) & (b_1 b_2 - 1)(b_1 b_2 b_3 - b_1) \end{bmatrix}$$

Viewing these matrices as $\mathbb{R}^4$ - vectors and taking the determinant over those vectors, we see that this determinant depends on x_1 and x_2 and thus, for any given $b_1, \ldots, b_4$, these vectors are independent for a.e. $x_1, \ldots, x_4$. □

Corollary 5.5:

$Z_{kl} = \alpha_1 \, Ad_{g_1} \, Y_{11} + \ldots + \alpha_4 \, Ad_{g_1 \, . \, . \, g_4} \, Y_{11} + \sum \alpha_{ij} \, Y_{ij} + \sum \beta_{ij} \, X_{ij}$

for any given $b_1, \ldots, b_4$ *and suitable choice of* $x_1, \ldots, x_4$, *where the* α, β*'s depend on* k *and* l.

proof: This follows immediately from Lemma 5.7. □

After this preparation, we can now start with the actual proof.

Assume that we are given the values for $b_1, \ldots, b_{58}$. We are now going to show that for this choice of b-values and $n = 58$ that $\mathfrak{A} = \mathfrak{G}$.

Assume $< Y_{11}, Ad_{g_1} Y_{11}, \ldots, Ad_{g_1 \, . \, . \, g_9} Y_{11} > \neq \mathfrak{G}$ for all $x_1, \ldots, x_9$.

Then :

Step 1:

$X_{11} \varepsilon < Y_{11}, Ad_{g_1} Y_{11}, \ldots, Ad_{g_1 \, . \, . \, g_k} Y_{11} > =: \mathfrak{A}_k, \; \forall \; x_1, \ldots, x_k$, *some* $k < 10$.

proof: Let $H = g_8 \ldots g_1$. Then $Ad_{H^{-1}} X_{11} = Y_{11}$ leads to the condition:

First column of $H = (0,0,1,0)^T$, which can be handled as in Lemma 5.4. □

Step 2:

For $k = 34$, Y_{ij} , $X_{11} \; \varepsilon \; \mathfrak{A}_k \quad i,j = 1,2$.

proof: From Step 1 we have that $X_{11} \varepsilon \; \mathfrak{A}_k$ for any $k > 9$. For an appropriate choice of

$x_{11}, \ldots, x_{14}$, we know by the remark after Lemma 5.4 that for $H = g_1 \ldots g_{14}$ the conditions of Lemma 5.4 are satisfied. Therefor we conclude the existence of x's such that $Y_0 \varepsilon\ \mathfrak{A}_{22}$. With the same argument, i.e., satisfying the condition of Lemma 5.5 first, we get that $Y_{11} - Y_{12} + Y_{22} \varepsilon\ \mathfrak{A}_{34}$, from where we conclude that Y_{12} and $Y_{22} \varepsilon\ \mathfrak{A}_{34}$ □

Step 3:

$$X_{ij},\ Y_{ij} \varepsilon\ \mathfrak{A}_{58}\ ,\ i,j = 1,2.$$

proof: Set $H = g_1 \ldots g_{34}$ and follow the same arguments as before, using Lemma 5.6. □

And finally:

Step 4:

$$\mathfrak{A}_{58} = \mathfrak{G}$$

proof: It remains to show that $Z_{ij} \varepsilon\ \mathfrak{A}_{58}$. But this follows from the previous steps and Corollary 5.5. □

Since the choice of $b_1, \ldots, b_{58}$ was arbitrary, Lemma 5.1 now follows.

§5.2 Parts (i), (ii):

Let us again denote the potentials with Bernoulli distribution by b_{ik}. We will show that the distribution of the (b_{ik}) as stated in the Theorem satisfy a condition similiar to (C1), which suffices to use the proofs in § 3, if we take i = 1, . . , n+1, that is if we take n+1 instead of n products of transfer matrices. Again, we have

$$(*)^{n+1} d\mu = \sum_{(b)} c_b \, (*)^{n+1} d\mu_b$$

with the same notation as in § 5.1.

We will show that for each fixed E, Lemma 3.3 is satisfied for a specified map and its Jacobian, from where we conclude the proof by following the same arguments for the proof of Theorem 1, applied to all the individual $(*)^{n+1} d\mu_b$. We saw in §3 that there is a distinguished index set $I := \{ (i_1,k_1), \ldots, (i_d,k_d) \}$ such that for fixed y_{ik} with $(i,k) \notin I$ and fixed b_{ik}, (i = 1, . . ,n), the map

$$\psi_{y,b} := g_n \cdot \cdot g_1$$

is a diffeomorphism in the index set I - variables $x_1, \ldots, x_d$ and

$$\det D\psi_{y,b} = \alpha(b\text{-}E)\, \beta(b\text{-}E, y\text{-}E, x\text{-}E)$$

where α, β are polynomial and there is no $b^0 = (b^0_{ik})$ such that $\beta\,(b^0\text{-}E, y\text{-}E, x\text{-}E) \equiv 0$ in the x, y variables.

Let us define the map

$$\tilde{\psi} := g_{n+1}\, g_n \,..\, g_1$$

$\tilde{\psi}$ has maximal rank, if one of the two maps $\psi_1 := g_n \,..\, g_1$, $\psi_2 := g_{n+1} \,..\, g_2$ has maximal rank. Associated to ψ_1 and ψ_2 are two sets of distinguished indices I_1 and I_2 such that for fixed y_{ik}, $(i,k) \notin I_s$ and fixed b_{ik}, $\psi_{s,y,b}$ is a diffeomorphism a.e. in the (x_i), $i \varepsilon\, I_s$, $s = 1, 2$. Also,

$$\det D\psi_{s,y,b} = \alpha_s\,(b\text{-}E)\; \beta_s\,(b\text{-}E, y\text{-}E, x\text{-}E),$$

where $\alpha_1(b_{ik}) = \alpha_2(b_{i+1,k})$ for $i = 1, \ldots, n$. Therefor, for fixed E and b, $\tilde{\psi}$ will have maximal rank at a.e. $(y,x) \varepsilon\, \mathbb{R}^{(n+1)(m^d - m^{d-1})}$ unless

$$\alpha_1(b\text{-}E) = \alpha_2(b\text{-}E) = 0 \tag{5.8}$$

and thus, if for all $b \varepsilon$ supp $d\eta$ and E fixed, where $d\eta$ is the joint distribution of the b_{ik} with $i = 1, \ldots, n+1$, $\alpha_s\,(b\text{-}E) \neq 0$ for $s = 1$ or 2, we can use the corresponding ψ_s in place of ψ in § 3 to prove that $(*)^{n+1} d\mu_E\, \varepsilon\, L^p_\alpha$ for some α,p. The statement then follows as before. The same argument can be used for all E.

To see that the distributions specified in the statement of the Theorem don't have any b satisfying (5.8) in its support, let us define

$$N_s\,(E) := \{\, b \mid \alpha_s\,(b\text{-}E) = 0 \,\}\,,\quad s = 1, 2$$

and

$$N := \bigcup_E \{\, N_1(E) \cap N_2(E) \,\}.$$

Then N has dimension at most $(n+1)\,(m^d - m^{d-1}) - 1$. Since we also know from Lemma 3.1, that the diagonal of $\mathbb{R}^{(n+1)(m^d - m^{d-1})}$ does not meet N, we have dist (0,N) $= \epsilon > 0$. If b_{ik} has Bernoulli (0,b) distribution with $|b| < \epsilon$, then supp $d\eta \cap N = \emptyset$, so

(5.8) cannot be satisfied for any $b \varepsilon$ supp $d\eta$. This means that for all E, the map $\tilde{\psi}$ has maximal rank for all b-variables and a.e. in the "top surface"-variables. For fixed b and any E, det $D\psi_{s,y,b}$ is not identical 0 in the (x,y) for s = 1 or 2 and Lemma 3.3 is satisfied with det $D\psi_{s,y,b}$ in place of J_y. So we can use $\psi_{s,y,b}$ instead of ψ_y in § 3 and follow the same arguments to conclude that k(E) is a C^∞ function for this model. Since smoothness of k(E) is invariant under translations of the potential, (k just gets translated as well), (i) follows for $b_0 = \epsilon$.

To prove part (ii), we note that if a particular realization (b_{ik}) of the Bernoulli variables is contained in N, it follows from the fact that N is a lower dimensional manifold, that for any given $\epsilon > 0$ and a.e. vector $c = (c_{ik}) \varepsilon B_\epsilon$, where B_ϵ is the ball of radius ϵ around 0 in $\mathbb{R}^{(n+1)(m^d - m^{d-1})}$, $(b_{ik} + c_{ik}) \notin N$. Then the same argument as above applies, as (5.8) cannot be satisfied for the perturbed potential, which proves part (ii).

REFERENCES

Bougerol, P., Lacroix,J. [1] : *Products of random matrices with applications to Schrödinger operators.* Progress in Probability and Statistics, Vol. 8, Birkhauser, Boston (1985).

Bovier, A., Campanino, M., Klein, A. and Perez, J. [1] : *Smoothness of the density of states in the Anderson model at high disorder.* Comm. Math. Phys., to appear.

Calderon, A. [1] : *Intermediate spaces and interpolation, the complex method,* Studia Math., 24, (1964), 113.

Campanino, M. and Klein, A. [1] : *A supersymmetric transfer matrix and differentiability of the density of states in the one dimensional Anderson model,* Comm. Math. Phys., 104, (1984), 227.

Carmona, R. [1] : *Random Schrödinger Operators.* Proceedings of the Ecole de probability XIV Saint-Flour 1984. Lecture Notes in Math. Vol. 1180. Springer Verlag, (1986).

Carmona, R., Klein, A. and Martinelli, F. [1] : *Anderson localization for Bernoulli and other singular potentials,* preprint.

Constantinescu, F., Fröhlich, J. and Spencer, T. [1] : *Analyticity of the density of states and replica method for random Schrödinger operators on a lattice*, Journal of Stat. Physics, 34, (1984), 571.

Craig, W. and Simon, B. [1] : *Log Hölder continuity of the integrated density of states for stochastic Jacobi matrices*, Comm. Math. Phys., 75, (1983), 207.

Dean, P. and Martin, J. [1] : *A method of determining the frequency spectra of disordered lattices in two dimensions*, Proc. Phys. Soc., 259, (1961), 409.

Delyon, F., Lèvy, Y.-E. and Souillard, B. [1] : *Approach à la Borland to multidimensional localization*, Physical Review Letters, 55, No. 6, (1985), 618.

Dieudonnè, J. [1] : *Foundations of modern analysis*, Academic Press, (1969).

Furstenberg, H. [1] : *Non commuting random products*, Trans. Amer. Math. Soc., 108, (1963), 377.

[2] : *A Poisson formula for semi simple Lie groups*, Ann. of Math., 77, (1963), 335.

Guivarc'h, Y. and Reugi, A. [1] : *Frontière de Furstenberg. Propriètès de contraction et thèorèmes de convergence*, Sèminaire de Probabilitès, Rennes, (1981).

Helgason, S. [1] : *Groups and geometric analysis*, Academic Press, (1984).

[2] : *Differential geometry, Lie groups and symmetric spaces*, Academic Press, (1978).

Hori, J. [1] : *Spectral properties of disordered chains and lattices*, Pergamon Press, (1968).

Kato, T. [1] : *Perturbation theory for linear operators*, Springer Verlag, (1976).

Klein, A., Speis, A. [1]: *Smoothness of the integrated density of states in the Anderson model on a one dimensional strip*, preprint (1988).

Lacroix, J. [1] : *Singularitè du spectre de l'opèrateur de Schrödinger alèatoire dans un ruban ou un demi - ruban*, Annales de l'Institute Henri Poincarè, Section A - Vol. 38, (1983), 385.

[2] : *Localisation pour l'opèrateur de Schrödinger alèatoire dans un ruban*, Annales de l'Institute Henri Poincarè - Physique theorique, 40, (1984), 97.

Pastur, L. [1] : *Spectral properties of disordered systems in one body approximation*, Comm. Math. Phys., 75, (1980), 179.

Reed, M. and Simon, B. [1] : *Methods of modern mathematical physics vol.1*, Academic Press, (1980).

[2] : *Methods of modern mathematical physics vol.2*, Academic Press, (1975).

Sazanov, V. and Tutubalin, V. [1] : *Probability distributions on topological groups*, Theory of Prob. and its Appl., 11, (1966), 1.

Strichartz, R. [1] : *Analysis of the Laplacian on the complete Riemannian manifold*, Journal of Funct. Anal., 52, (1983), 48.

Taylor, M. [1] : *Pseudodifferential operators*, Princeton University Press, (1981).

[2] : *Noncommutative harmonic analysis*, AMS, (1986).

Tutubalin, V. [1] : *On limit theorems for the product of random matrices*, Theory of Prob. and its Appl., X, No. 1, (1965), 15.

Walters, P. [1] : *An introduction to ergodic theory*, Springer Verlag, (1975).

Warner, F. [1] : *Foundations of differentiable manifolds and Lie groups*, Springer Verlag, (1971).