*manuscripta math.* **70** (1990) 27–37

# On the Fröhlich-Spencer Estimate in the Theory of Anderson Localization

Jürgen Pöschel

## 0 The Problem

The notion of Anderson localization refers to the appearance of pure point spectrum with exponentially localized eigenstates within the spectrum of Schrödinger operators, in particular random Schrödinger operators on integer lattices. These are defined by Hamiltonians

$$H = -\Delta + V$$

acting on the Hilbert space of square summable sequences on $\mathbf{Z}^d$ with $d \geq 1$, where $\Delta$ denotes the finite difference Laplacian on $\mathbf{Z}^d$ with

$$\Delta(x, y) = \begin{cases} 1, & |x-y| = 1 \\ 0, & |x-y| \neq 1, \end{cases}$$

and $V$ denotes a random potential on $\mathbf{Z}^d$, for example with independent, identically distributed random variables $V(x)$, $x \in \mathbf{Z}^d$.

Such Hamiltonians were introduced by Anderson [A] in the fifties to model the motion of a single quantum-mechanical electron in a random medium such as a crystal with impurities of random strength. Intuitively, sufficiently large disorder in the crystal should trap the electron in a bounded region, thus eliminating extended states and continuous spectrum and creating localized states and point spectrum instead.

In the one-dimensonal case, $d = 1$, this intuition was rigorously confirmed in the seventies by Goldscheid, Pastur and Molchanov and others — see [FS] for references. Indeed, complete localization was found to happen even for arbitrarily small disorder. The higher dimensional case, however, turned out to be more delicate. Substantial progress was made only recently following a new and fundamental estimate by Fröhlich and Spencer [FS] concerning the exponential decay properties of the associated Green's function for energy values close to the spectrum. As a result, complete localization was established for sufficiently high disorder or sufficiently low energy. See [DK] for a recent version and references.

The rational behind the Fröhlich-Spencer estimate is the following. Consider the restriction $H_\Lambda$ of $H$ to an arbitrary, usually finite, subset $\Lambda$ of the lattice $\mathbf{Z}^d$ with Dirichlet boundary conditions outside $\Lambda$. Let

$$G_\Lambda(E) = (H_\Lambda - E)^{-1}$$

denote its Green's function, where $E$ is either real valued and not in the spectrum $\sigma(H_\Lambda)$ of $H_\Lambda$ or complex valued with a nonzero but possibly small imaginary part. If $E$ is sufficiently separated from the potential values within $\Lambda$, then the coefficients of $G_\Lambda(E)$ decay exponentially fast — this is a standard result, recalled in the appendix. Otherwise, various resonances between $E$ and the potential values occur, but given the randomness of the potential, stronger resonances are usually much rarer and more sparse than milder ones. The idea is first to remove *all* those resonances from $\Lambda$ by decoupling them from the rest of $\Lambda$. Thus, in the beginning, $\Lambda$ is replaced by a suitable subset $\Lambda_0$, where again $E$ is sufficiently separated from the potential values. Then these resonance islands are recoupled *one after the other* with increasing strength using the resolvent identity and the exponential decay of the Green's function established so far. This way, estimates are obtained iteratively for an increasing sequence $\Lambda_0 \subset \Lambda_1 \subset \cdots \subset \Lambda$ until all of $\Lambda$ is retained.

The purpose of this note is to describe an approach to this procedure which avoids many of the technicalities of the proofs given so far. Its new ingredient is a family of exponentially weighted norms depending on distance functions which automatically take into account the size and location of the resonant islands. Its main result is the Coupling Lemma of the following section that quantitatively describes the effect of recoupling resonant islands.

We are going to describe estimates that are valid for certain configurations of the potential values and are thus purely *deterministic*. We leave out any questions of how probable such configurations are for certain classes of random potentials, but refer to [DK] and the references therein for these matters. Therefore we will not specify any class of potentials $V$ belongs to.



## 1 The Coupling Lemma

We are going to estimate the Green's function in terms of exponentially weighted norms of the form

$$\|S\|_{m,Z} = \sup_x \sum_y |S(x,y)|\, e^{m\, \mathrm{dist}_Z(x,y)}, \qquad m \geq 0, \quad Z \subseteq \mathbf{Z}^d$$

where $\mathrm{dist}_Z(x,y)$ denotes the distance between points $x$ and $y$ in $\Lambda$ relative to some *zero distance set* $Z \subseteq \mathbf{Z}^d$. This distance is defined as follows. Any link in $\mathbf{Z}^d$ — that is a nearest neighbor connection —is assigned zero length rather than the standard unit length, if *both* of its endpoints belong to $Z$. The $Z$-distance between two lattice points is then the length of their shortest connecting path, where path length is defined as usual as the total length of all its links.

These norms are multiplicative: $\|ST\|_{m,Z} \leq \|S\|_{m,Z}\|T\|_{m,Z}$. They are also monotone: $\|S\|_{m,Z} \leq \|S\|_{n,Z}$ for $m \leq n$ and $\|S\|_{m,Z} \leq \|S\|_{m,Y}$ for $Z \supseteq Y$. In case of an empty set $Z$ we have

$$\mathrm{dist}_\emptyset(x,y) = \mathrm{dist}(x,y) = |x-y| = \sum_{i=1}^d |x_i - y_i|,$$

the $\ell_1$-distance between $x$ and $y$ in the lattice, and $\|S\|_{m,\emptyset} = \|S\|_m$ reduces to an ordinary exponentially weighted norm.

The Coupling Lemma describes the effect of coupling a region $A$ of more resonant potential values into a region $B$ on which the Green's function is already known to show some exponential decay. These two regions are required to overlap in a sufficiently large 'collar' so that the exponential decay of $G_B$ across this 'collar' overpowers the blow up due to $G_A$. The geometric set up of this lemma was inspired by [DK].

Let

$$K_m = \|\Delta\|_m = 2d\, e^m.$$

Throughout the rest of this note all estimates refer to one and the same energy level $E$, which is therefore omitted from the notation for simplicity.

**The Coupling Lemma.** *Let $\Lambda = A \cup B$, and let $Y \subseteq A$, $Z \subseteq B$ be their respective zero distance sets. Assume that*

$$\|G_A\|_{m,Y} \leq M_A,$$

*and that around each point* $x \in \Lambda\backslash A$ *there is a neighbourhood* $V \subseteq B$ *such that*

$$\|G_V\|_{m,Z} \leq M_B, \qquad \mathrm{dist}_Z(x, \Lambda\backslash V) \geq r$$

*with fixed positive numbers* $m$, $r$ *and* $M_A, M_B \geq K_m^{-1}$. *If there is* $0 \leq \mu \leq m$ *such that*

$$\theta = K_m^2 M_A M_B e^{-\mu r} < 1,$$

*then*

$$\|G_\Lambda\|_{m-\mu, Z\cup Y} \leq \frac{2}{1-\theta} K_m M_A M_B.$$

*Proof.* For any subset $V \subseteq \Lambda$ we have $H_\Lambda = H_{V,\Lambda\backslash V} + \Gamma_V$, where $H_{V,\Lambda\backslash V}$ denotes the direct sum of the uncoupled Hamiltonians $H_V$ and $H_{\Lambda\backslash V}$, and $\Gamma_V = \Delta_\Lambda - \Delta_V - \Delta_{\Lambda\backslash V}$ represents their coupling across the boundary of $V$. It follows from the resolvent identity that

$$G_\Lambda = G_{V,\Lambda\backslash V} + G_{V,\Lambda\backslash V}\Gamma_V G_\Lambda,$$

where $G_{V,\Lambda\backslash V}$ denotes the direct sum of $G_V$ and $G_{\Lambda\backslash V}$.

Let $S^x$ denote the $x$-th row of a linear operator $S$, with

$$\|S^x\|_{m,Z} = \sum_y |S(x,y)|\, e^{m\,\mathrm{dist}_Z(x,y)},$$

so that $\|S\|_{m,Z} = \sup_x \|S^x\|_{m,Z}$. For $x \in \Lambda\backslash A$ choose a neighbourhood $V$ of $x$ as stipulated in the lemma and apply the resolvent identity to obtain $G_\Lambda^x = G_V^x + G_V^x \Gamma_V G_\Lambda$. Then, for $0 \leq n \leq m$,

$$\|G_\Lambda^x\|_{n,Z\cup Y} \leq \|G_V\|_{m,Z} + \|G_V^x \Gamma_V\|_{n,Z} \|G_\Lambda\|_{n,Z\cup Y}.$$

For $x \in A$ choose $A$ as a neighbourhood to obtain $G_\Lambda^x = G_A^x + G_A^x \Gamma_A G_\Lambda$. The product $G_A^x \Gamma_A G_\Lambda$ does not involve any elements from rows $G_\Lambda^u$ with $u \in A$. By the previous estimate and the hypotheses we thus have, for $x \in A$,

$$\begin{aligned}\|G_\Lambda^x\|_{n,Z\cup Y} &\leq \|G_A\|_{m,Y} + \|G_A\|_{m,Y}\|\Gamma_A\|_m \sup_{u\notin A} \|G_\Lambda^u\|_{n,Z\cup Y} \\ &\leq M_A + K_m M_A M_B \\ &\qquad + K_m M_A \sup_{u \notin A} \|G_V^u \Gamma_V\|_{n,Z} \|G_\Lambda\|_{n,Z\cup Y},\end{aligned}$$



where $V = V(u)$ denote the neighbourhood s assigned above to points $u$ not in $A$. Since $K_m M_A \geq 1$ by assumption, the latter estimate comprises the former. Taking the supremum over all $x \in \Lambda$ and using $K_m M_B \geq 1$ we thus obtain

$$\|G_\Lambda\|_{n,Z\cup Y} \leq 2K_m M_A M_B + K_m M_A \sup_{u\notin A} \|G_V^u \Gamma_V\|_{n,Z} \|G_\Lambda\|_{n,Z\cup Y}.$$

Now observe that for each such $V$,

$$\begin{aligned}\|G_V^u \Gamma_V\|_{n,Z} &\leq \|\Gamma_V\|_n \|G_V\|_{m,Z}\, e^{-(m-n)(\mathrm{dist}_Z(u,\Lambda\backslash V)-1)} \\ &\leq K_n M_B e^{-(m-n)(r-1)} \\ &= K_m M_B e^{-(m-n)r}.\end{aligned}$$

Choosing $n = m - \mu$ and appealing to the hypotheses of the lemma we obtain

$$\|G_\Lambda\|_{m-\mu,Z\cup Y} \leq 2K_m M_A M_B + \theta \|G_\Lambda\|_{m-\mu,Z\cup Y},$$

and the result follows. ∎

## 2 The Fröhlich-Spencer Estimate

Given a potential $V$ and a real or complex energy level $E$, let

$$S = \left\{ x \in \mathbf{Z}^d : |V(x) - E| \leq N \right\}$$

with a sufficiently large $N$ to be characterized below in terms of the parameter $m$. The set $S$ comprises the lattice sites where resonances with the energy level $E$ might occur.

We assume that there is a decomposition

$$S = \bigcup_{i\geq 1} C_i$$

of $S$ into mutually disjoint subsets $C_i$ together with a family of mutually disjoint covering sets $D_i \supseteq C_i$ for $i \geq 1$. Each of these covering sets is assumed to consist of a family of connected components $D_i^\alpha$ with 'cores' $C_i^\alpha = D_i^\alpha \cap S$ such that

$$\mathrm{dist}\left(\mathbf{Z}^d\backslash D_i^\alpha, C_i^\alpha\right) \geq r_i, \tag{1}$$

$$\mathrm{dist}\left(\sigma\left(H_{D_i^\alpha}\right), E\right) \geq 1/\Phi(r_i), \tag{2}$$

$$\mathrm{card}\, D_i^\alpha \leq \Phi(r_i). \tag{3}$$

The right hand sides of these inequalities are required to satisfy

$$\sum_{i\geq 1} r_i^{-1}\log \Phi(r_i) < \infty, \tag{4}$$

and we also assume that

$$r_{i+1} \geq 4r_i \geq 4, \qquad i \geq 1 \tag{5}$$

for convenience. Roughly speaking, the stronger the spectral resonance on $D_i^\alpha$ and the bigger this component, measured in terms of $\Phi(r_i)$, the bigger the 'overlap' $r_i$ of $D_i^\alpha$ with the complement of $S$ is required to be.

Incidentally, condition (4) is another instance of the Brjuno condition for small divisors originally arising in Siegel's problem of linearizing a complex analytic map in the plane around a neutral fixed point [B,R].

The definition of $S$ implies that for every subset $\Lambda$ of $\mathbf{Z}^d$ disjoint from $S$,

$$\|G_\Lambda\|_m \leq 1, \qquad m = \log\frac{N-1}{2d}. \tag{6}$$

Moreover, (2) and (3) imply that for all $i \geq 1$,

$$\|G_{D_i}\|_0 = \sup_\alpha \|G_{D_i^\alpha}\|_0 \leq \Psi(r_i) = \Phi^2(r_i). \tag{7}$$

The proofs are simple and given in the appendix. In the following we use (6) as a characterization of $S$, and (7) instead of (2) and (3).

A subset $\Lambda \subseteq \mathbf{Z}^d$ is called *admissible*, if

$$C_i^\alpha \cap \Lambda \neq \emptyset \quad \Rightarrow \quad D_i^\alpha \subseteq \Lambda$$

for every $C_i^\alpha$. Its *order* is $k = \max\{ i : \Lambda \cap C_i \neq \emptyset \}$, which is understood to be 0 if $\Lambda \cap S = \emptyset$.

The configurations considered here are more flexible than those in [FS]. The sets $C_i$ need not be maximal in any sense. Moreover, no assumption has to be made about the distance between different components of $D_i$. Those come into play at a later stage only. Finally, an admissible set $\Lambda$ is allowed to intersect some component $D_i^\alpha$ without containing it as long as it does not intersect with its 'core' $C_i^\alpha$.

For $k \geq 1$ set

$$Z_k = \bigcup_{1\leq i\leq k} D_i, \qquad m_k = m - \sum_{1\leq i\leq k}\mu_i,$$



with

$$\mu_k = \frac{k+1}{r_k}\log K_{m+1} + \frac{1}{r_k}\sum_{1\leq i\leq k}\log\Psi(r_i).$$

Moreover, let $Z_0 = \emptyset$ and $m_0 = m$, and let $r = r_1$.

**Theorem (Fröhlich-Spencer).** *Assume that (1) and (4-7) hold, and that*

$$m \geq m_* = \frac{a}{r-4}\Big(1 + \log 2d + \frac{r}{2}\sum_{i\geq 1} r_i^{-1}\log\Psi(r_i)\Big), \qquad a = \frac{28}{9}.$$

*If $\Lambda$ is admissible of finite order $k$, then*

$$\|G_\Lambda\|_{m_k,Z_k} \leq K_{m+1}^k \prod_{1\leq i\leq k}\Psi(r_i),$$

*where $m_k > (m-m_*)(r-a)/r \geq 0$ for all $k \geq 0$.*

*Proof.* By (5) we have

$$\sum_{k\geq 1}\frac{k+1}{r_k} \leq \frac{1}{r}\sum_{k\geq 1}\frac{k+1}{4^{k-1}} = \frac{a}{r}, \qquad a = \frac{28}{9}.$$

Moreover, $\log K_{m+1} = m + 1 + \log 2d$. Consequently, again using (5),

$$\begin{aligned}\sum_{k\geq 1}\mu_k &= \sum_{k\geq 1}\frac{k+1}{r_k}\log K_{m+1} + \sum_{k\geq 1}\sum_{1\leq i\leq k} r_k^{-1}\log\Psi(r_i)\\ &\leq \frac{a}{r}m + \frac{a}{r}(1+\log 2d) + \frac{a}{2}\sum_{i\geq 1} r_i^{-1}\log\Psi(r_i)\\ &= \frac{a}{r}m + \frac{r-a}{r}m \leq m.\end{aligned}$$

It follows that $m_k > (m-m_*)(r-a)/r$ for all $k$.

The estimate of $G_\Lambda$ obviously holds for admissible sets of order 0. So assume it proven for such sets of order up to $k \geq 0$, and let $\Lambda$ be admissible of order $k+1$.

Let $C = \Lambda \cap C_{k+1}$ and $B = \Lambda \backslash C$. Then $B$ is admissible of order $k$, so $\|G_B\|_{m_k,Z_k}$ can be bounded as stated in the theorem. Let $D$ be the smallest cover of $C$ consisting of connected components of $D_{k+1}$. By (7),

$$\|G_D\|_{m_k,D} \leq \|G_{D_{k+1}}\|_0 \leq \Psi(r_{k+1}).$$

Since $\Lambda$ is admissible, we have $D \subseteq \Lambda$ and so $\Lambda = B \cup D$. Moreover,

$$\mathrm{dist}_{Z_k}(\Lambda\backslash D, C) \geq \mathrm{dist}_{\emptyset}(\Lambda\backslash D, C) \geq r_{k+1},$$

because the shortest path representing the $Z_k$-distance between $\Lambda\backslash D$ and $C$ must lie entirely within one component of $D$ except for one endpoint, which by assumption is disjoint from $Z_k$. Whence this distance is just the ordinary distance.

We can now apply the Coupling Lemma to $\Lambda = B \cup D$, with $B$ as a neighbourhood for every point not in $D$, since

$$\begin{aligned}\theta &= K_m^2 \, \|G_D\|_{m_k,D} \|G_B\|_{m_k,Z_k} e^{-\mu_{k+1} r_{k+1}} \\ &\leq e^{-2} K_{m+1}^{k+2} \prod_{i=1}^{k+1} \Psi(r_i)\, e^{-\mu_{k+1} r_{k+1}} \\ &\leq e^{-2}\end{aligned}$$

by the very definition of $\mu_{k+1}$. Moreover, $m_{k+1} = m_k - \mu_{k+1} > 0$. Consequently,

$$\begin{aligned}\|G_\Lambda\|_{m_{k+1},Z_{k+1}} &\leq \|G_\Lambda\|_{m_k - \mu_{k+1}, Z_k \cup D} \\ &\leq K_{m+1} \|G_D\|_{m_k,D} \|G_B\|_{m_k,Z_k} \\ &\leq K_{m+1}^{k+1} \prod_{1 \leq i \leq k+1} \Psi(r_i),\end{aligned}$$

as required. ∎

As an illustration suppose (7) holds with $\Psi(r) = e^{\sqrt{r}/2}$ as in [FS]. Then

$$\sum_{i \geq 1} r_i^{-1} \log \Psi(r_i) = \frac{1}{2} \sum_{i \geq 1} \frac{1}{\sqrt{r_i}} \leq \frac{1}{\sqrt{r}} \sum_{i \geq 1} 2^{-i} = \frac{1}{\sqrt{r}}$$

by (5). It suffices to assume $r \geq 10$ and $m \geq m_o = 2 + \log 2d$ for simplicity to obtain

$$\|G_\Lambda\|_{n,Z_k} \leq K_{m+1}^k e^{\sqrt{r_k}}$$

for admissible sets $\Lambda$ of order $k$ with $n = 2(m - m_o)/3$. From this one recovers exponentially small estimates for the components of $G_\Lambda$ with $x$ and $y$ sufficiently far apart provided the reduced distance $\mathrm{dist}_{Z_k}(x, y)$ is comparable to the standard distance $|x - y|$. For instance, assume that

$$2\, \mathrm{dist}_{Z_k}(x, y) \geq |x - y| - d_k$$



for all $x$, $y$ in $\mathbf{Z}^d$ and all $k \geq 1$ with $d_k = \sup_\alpha \mathrm{diam}\, D_k^\alpha$. Assuming also that $m \geq 3m_o$ we obtain

$$|G_\Lambda(x, y)| \leq e^{-n|x-y|/4} \quad \text{for} \quad |x - y| \geq 4d_k,$$

since with these assumptions,

$$\begin{aligned}n\, \mathrm{dist}_{Z_k}(x, y) &\geq \frac{n}{2}(|x - y| - d_k) \\ &\geq \frac{n}{4}|x - y| + \frac{n}{2} d_k \\ &\geq \frac{n}{4}|x - y| + \frac{m - m_o}{3} r_k\end{aligned}$$

on one hand and

$$\log\left(K_{m+1}^k e^{\sqrt{r_k}}\right) = (m + 1 + \log 2d)\, k + \sqrt{r_k} \leq \left(\frac{m}{9} + \frac{m_o}{3}\right) r_k$$

on the other hand.

The last lemma describes a simple criterion to ensure that the reduced distance is comparable to the standard distance.

**Lemma.** *Suppose that for all $k$ and all components $D_k^\alpha$,*

$$\mathrm{dist}(D_k^\alpha, Z_k \backslash D_k^\alpha) \geq \mathrm{diam}\, D_k^\alpha,$$

*where* diam $D_k^\alpha$ *is the diameter of* $D_k^\alpha$ *measured with respect to paths lying entirely* within $D_k^\alpha$. *Then*

$$|x - y| \leq 2\, \mathrm{dist}_{Z_k}(x, y) + d_k$$

*for all $x$ and $y$ in $\mathbf{Z}^d$ and all $k \geq 1$ with $d_k = \sup_\alpha \mathrm{diam}\, D_k^\alpha$.*

*Proof.* Fix $k \geq 1$. To simplify notation we drop the subscript $k$ and write $Z$ for $Z_k$ and so on.

Consider the most disadvantagous case where both $x$ and $y$ lie in $Z$. There is a shortest path $\gamma$ between $x$ and $y$ with respect to the reduced distance. This path has a decomposition

$$\gamma = \nu_0 \cup \mu_1 \cup \nu_1 \cup \cdots \cup \mu_n \cup \nu_n,$$

where each $\nu_i$ lies entirely in $Z$, whereas each $\mu_i$ lies in $\Lambda \backslash Z$ except for its endpoints.

Thus, $|\nu_i|_Z = 0$ for $0 \le i \le n$ and $|\mu_i|_Z = |\mu_i|$ for $1 \le i \le n$, and

$$|\gamma|_Z = \sum_{1 \le i \le n} |\mu_i| = \mathrm{dist}_Z(x, y).$$

If necessary, we may also replace each $\nu_i$ by another path in the same component $D^\alpha$ and with the same endpoints so that $|\nu_i| \le \mathrm{diam}\, D^\alpha$. Here, $|\gamma|_Z$ and $|\gamma|$ denote the length of a path $\gamma$ with respect to the $Z$-distance and $\ell_1$-distance respectively.

For $\nu_0$ we clearly have $|\nu_0| \le d$. If $n \ge 1$, then $\mu_i$ must be connecting two *different* components of $Z$, since otherwise $\gamma$ were not minimal. It follows that $|\mu_i|$ is greater or equal than the diameter of the component of $Z$ to which $\mu_i$ is leading to, whence $|\mu_i| \ge |\nu_i|$ for $1 \le i \le n$ by the above choice of the $\nu_i$ and the hypotheses. Consequently,

$$\begin{aligned} |x - y| &\le |\nu_0| + \sum_{1 \le i \le n} (|\mu_i| + |\nu_i|) \\ &\le d + 2 \sum_{1 \le i \le n} |\mu_i| \\ &= d + 2\, \mathrm{dist}_Z(x, y). \ ∎ \end{aligned}$$

## Appendix

To prove (5) write formally

$$G_\Lambda = (V - E - \Delta_\Lambda)^{-1} = \sum_{k \ge 0} \left((V - E)^{-1} \Delta_\Lambda\right)^k (V - E)^{-1},$$

where the diagonal operator $V - E$ is understood to be restricted to $\Lambda$. For $\Lambda$ disjoint from $S$,

$$\|(V - E)^{-1}\|_m \le \max_{x \in \Lambda} |V(x) - E|^{-1} \le \frac{1}{N}, \qquad m \ge 0,$$

while $\|\Delta_\Lambda\|_m \le 2d\, e^m$ for any $\Lambda$. It follows that for $m < \log(N/2d)$ the formal series converges to $G_\Lambda$, and that

$$\|G_\Lambda\|_m \le \frac{1}{N - 2d\, e^m}.$$

For $m = \log((N - 1)/2d)$ the required estimate follows.



To prove (7) let $D$ be any finite subset of $\mathbf{Z}^d$ and suppose that

$$\mathrm{dist}(\sigma(H_D), E) \ge \Phi^{-1} > 0.$$

Since $H_D$ is hermitian, there exists a unitary transformation $U$ such that $U^* H_D U = \bar{H}_D$ is diagonal. Consequently, $\|\bar{H}_D^{-1}\|_0 \le \Phi$. By the Schwarz inequality, we have $\|U\|_0, \|U^*\|_0 \le \sqrt{M}$, where $M$ is the cardinality of $D$. Thus,

$$\|G_D\|_0 \le \|U^*\|_0 \|\bar{H}_D^{-1}\|_0 \|U\|_0 \le M\Phi,$$

as we wanted to show.

## References


[A] P. Anderson, *Absence of diffusion in certain random lattices*, Phys. Rev. **109** (1958), 1492.

[B] A.D. Brjuno, *Analytic form of differential equations*, Trans. Moscow Math. Soc. **25** (1971), 131–288.

[DK] H. von Dreifus, A. Klein, *A new proof of localization in the Anderson tight binding model*, Commun. Math. Phys. **124** (1989), 285–299.

[FS] J. Fröhlich, T. Spencer, *Absence of diffusion in the Anderson tight binding model for large disorder or low energy*, Commun. Math. Phys. **88** (1983), 151–184.

[R] H. Rüssmann, *On the one-dimensional Schrödinger equation with a quasi-periodic potential*, Annals of the New York Acad. Sci. **357** (1980), 90–107.