Acta Math., 179 (1997), 153–196

Discrete one-dimensional quasi-periodic Schrödinger operators with pure point spectrum

by

L. H. ELIASSON

Royal Institute of Technology
Stockholm, Sweden

1. Introduction

The discrete quasi-periodic *Schrödinger operator* in one dimension

$$\mathcal{L}_\theta : l^2(\mathbf{Z}) \to l^2(\mathbf{Z})$$

is defined by

$$(\mathcal{L}_\theta u)_n = -\varepsilon(u_{n+1}+u_{n-1}) + E(\theta+n\omega)u_n,$$

where ω is a real number and where E is a smooth function on the torus $\mathbf{T}=\mathbf{R}/(2\pi\mathbf{Z})$, belonging to the Gevrey class

$$\sup_{\theta\in\mathbf{T}} |\partial^\nu E(\theta)| < C(\nu!)^2 K^\nu \quad \forall \nu \geqslant 0.$$

We shall assume that ω satisfies a *Diophantine condition*

$$\|k\omega\| := \inf_{n\in\mathbf{Z}} |k\omega - n2\pi| \geqslant \frac{\varkappa}{|k|^\tau} \quad \forall k \in \mathbf{Z}\setminus\{0\},$$

for some constants $\varkappa>0$, $\tau>1$, and that E satisfies a generic *transversality condition*

$$\begin{cases} \max\limits_{0\leqslant\nu\leqslant s} |\partial_x^\nu(E(\theta+x)-E(\theta))| \geqslant \xi > 0 & \forall\theta\ \forall x, \\ \max\limits_{0\leqslant\nu\leqslant s} |\partial_\theta^\nu(E(\theta+x)-E(\theta))| \geqslant \xi\|x\| & \forall\theta\ \forall x. \end{cases}$$

Under these two assumptions we prove the following theorem.

THEOREM. *Assume that E and ω are as above. Then there exists a constant $\varepsilon_0 = \varepsilon_0(C, K, \xi, s, \varkappa, \tau)$ such that if $|\varepsilon| < \varepsilon_0$, then $\mathcal{L}_\theta$ is pure point with a complete set of eigenvectors in $l^2(\mathbf{Z})$ for a.e. θ. Moreover, the measure of the set $[\inf E, \sup E]\setminus\sigma(\mathcal{L}_\theta)$ goes to 0 as $\varepsilon\to 0$.*

This result generalizes previous work by Fröhlich–Spencer–Wittver and by Sinai who considered the case when E has only two critical points both of which are non-degenerate (see [7], [14], [8])—such functions clearly fulfill the generic condition.

It is easy to see that the generic condition is fulfilled for any analytic function which has no period shorter than 2π. And if it has a shorter period we can scale the period to become 2π and then apply the theorem. In particular the theorem applies to any real-analytic function which is not constant.

It will also be clear (from Proposition 8) that the same result remains true if we replace in $\mathcal{L}_\theta$ the constant difference operator $u \to \varepsilon(u_{n+1}+u_{n-1})$ by any other constant finite symmetric difference operator $u \to \varepsilon\left(\sum_{i=-N}^{N} a_i u_{n+i}\right)$. Therefore this result also generalizes the work [2].

Idea of proof. The formulation we choose for the problem is to consider the operator as a symmetric infinite-dimensional matrix that depends on the parameter θ,

$$D(\theta)+\varepsilon F(\theta) = \begin{pmatrix} \ddots & & & & 0 \\ & E(\theta-\omega) & -\varepsilon & & \\ & -\varepsilon & E(\theta) & -\varepsilon & \\ & & -\varepsilon & E(\theta+\omega) & \\ 0 & & & & \ddots \end{pmatrix},$$

and that is close to a diagonal matrix. This matrix satisfies the *shift condition*

$$(D+\varepsilon F)_{m+k,n+k}(\theta) = (D+\varepsilon F)_{m,n}(\theta+k\omega),$$

a property that permits us to control for example the whole spectrum of the matrix for a fixed but arbitrary parameter value θ_0 by controlling the dependence of a single eigenvalue as a function of θ.

Our goal is to conjugate the matrix $(D+\varepsilon F)(\theta)$ to a diagonal matrix $D_\infty(\theta,\varepsilon)$ by an orthogonal matrix made up of a complete set of orthonormal eigenvectors. The way to construct such a matrix is by an iterative procedure

$$U_j^* \cdot \ldots \cdot U_1^* \cdot (D+\varepsilon F) \cdot U_1 \cdot \ldots \cdot U_j = D_{j+1} + F_{j+1}$$

that conjugates $D+\varepsilon F$ closer and closer to a diagonal matrix $D_j = \mathrm{diag}(E^j(\theta+k\omega))$. During this iteration we will have to deal with *almost multiple eigenvalues* of the main part D_j—i.e. eigenvalues that are so close so that they prevent a purely perturbative construction of the transformation U_j close to the identity. These "almost multiplicities" are of finite order and give rise to certain finite-dimensional symmetric submatrices of

D_j+F_j that will be diagonalized in a non-perturbative manner. The "almost multiplicities" of $D_1=D$ are of order

$$\leqslant 2^{s+4}\frac{C((s+1)!)^2K^{s+1}}{\xi},$$

but in the following steps this order will grow in a controlled way and we will have to handle "almost multiplicities" of higher and higher order. This is a distinct feature from the approach of [7], [14], where the "almost multiplicities" remain bounded (by two) throughout the iteration. The "almost multiplicities" will be described at the jth step by a decomposition $\bigcup_i \Lambda_i^j=\mathbf{Z}$.

In carrying out this procedure it will be essential that the order of the "almost multiplicities" do not increase too fast. Our means to control this will be through a *non-degeneracy condition* on the spectrum of D_j, which at the first step is simply the condition on $E(\theta+k\omega)$ described above—smoothness and transversality. Such a simple condition on the eigenvalues will not persist unchanged under perturbations, and we will have to consider in what way the non-degeneracy persists during the iteration. At the jth step this will be through a condition on $E^j(\theta+k\omega)$ that involves all the "almost multiplicities" of the kth eigenvalues that have been considered at the previous steps. These will be described by subsets $\Omega_i^j\supset\Lambda_i^j$. Another factor that will help us to control the "almost multiplicities" is the *exponential decay* of the perturbation off the diagonal.

The iterative construction will be carried out completely for all θ. In order to establish pure point spectrum one is left with the problem of proving that the infinite sequence $U_1\cdot\ldots\cdot U_j$ of orthogonal matrices constructed will indeed converge to an orthogonal transformation. This turns out to be the case only for a.e. θ. Our approach provides a way to analyze also the remaining θ's. (Confer in this respect [4] where a similar approach is made for another perturbation problem.)

A *technical problem* concerns the smoothness. If we insist on always having the main part D_j in diagonal form it will not be possible to control the smoothness of (the non-perturbative part of) the transformation U_j, hence the smoothness of F_{j+1}. The smoothness of the perturbation can be controlled if we give up the diagonal form of D_j and only require it to be in block-diagonal form with finite-dimensional symmetric blocks—such a matrix can of course be diagonalized. However, such a decomposition into blocks would be somewhat arbitrary and would therefore destroy the shift condition on D_j and F_j. We prefer to preserve both smoothness and shift and we will therefore have to work with a more general *normal form* which is not block diagonal but which can be block diagonalized—this requires a more complex formulation which will have to include the dimensions of the blocks, their location in $l^2(\mathbf{Z})$ and their smoothness. We describe the normal form in §4 and we give some more explanations of the idea there.

The use of Gevrey functions is also a technical point. We could work with piecewise analytic functions but would then have to count the number of pieces. This is avoided by Gevrey classes since they contain functions with compact support. There is of course nothing particular with the Gevrey class we have chosen above—we could just as well have chosen another.

Content of the paper. The paper is organized in the following way. In §2 we present some results on finite-dimensional symmetric matrices and in §3 we study the transversality condition—the proof of the lemmas in these two sections will be of no relevance for the proof of the theorem, only the statements will be of importance. In §4 we describe the normal form and a small divisor result related to it—Lemma 5 and Corollary 6. In §5 we prove the inductive lemma—Lemma 7—common for all KAM techniques and in §6 we finally prove Proposition 8 and, as a simple consequence, the Theorem. In §7 we discuss some other spectral properties of the operator. Some basic results on functions in the Gevrey class, on parameter dependence of roots of polynomials, and on parameter-dependent Gram–Schmidt orthogonalization are given in an appendix.

Notations. For any smooth function f defined on I we use the norm

$$|f|_{C^\nu} := \frac{1}{(\nu!)^2} \max_{0\leqslant k\leqslant \nu} \sup_{\theta\in I} |\partial^\nu f(\theta)|.$$

If $f=(f_1,...,f_d)$ then we let $|\partial^\nu f(\theta)|$ be the Euclidean norm

$$\sqrt{(\partial^\nu f_1(\theta))^2+...+(\partial^\nu f_d(\theta))^2},$$

and if f is a matrix then we let $|\partial^\nu f(\theta)|$ denote the operator norm.

The distance $\|\cdot\|$ on the circle $\mathbf{T}=\mathbf{R}/2\pi\mathbf{Z}$ is defined above.

If A is a subset of $\mathbf{Z}$ and $n\in\mathbf{Z}$ then

$$\operatorname{diam}(A) := \sup_{k,l\in A} |k-l| \quad \text{and} \quad \operatorname{dist}(n,A) := \inf_{k\in A} |n-k|.$$

We shall consider positive functions $\gamma(z,x)$ defined for z in some open set $U\subset\mathbf{R}^n$ and for $x\geqslant 1$ such that for some j

$$\gamma(z,x) \leqslant \underbrace{\exp\circ...\circ\exp}_{j}(A(z)x), \quad z\in U, x\geqslant 1.$$

We say that γ increases (at most) *superexponentially* in x and that $1/\gamma$ decays (at most) superexponentially in x.

If X is a subset of $\mathbf{R}$ we denote by $|X|$ the Lebesgue measure of X.

Acknowledgment. This work was begun at a visit at the Department of Mathematics at UCLA in 1994 and it was essentially completed more than a year after at a visit at the Forschungsinstitut für Mathematik at ETH, Zürich. I am grateful to both these institutions for their hospitality.

2. Block diagonalization

Let D be a $(d\times d)$-matrix, $2\leqslant d$. We let $D_{m,n}$ denote the component in the mth row and the nth column. For a subset $\Lambda\subset\{1,\dots,d\}$ we define

$$D_\Lambda=\begin{cases} D_{m,n} & \text{if } m,n\in\Lambda,\\ \delta_m^n & \text{otherwise,}\end{cases}$$
$$\mathbf{R}^\Lambda=\{x\in\mathbf{R}^d: x_i=0 \text{ if } i\notin\Lambda\}.$$

Then

$$D_\Lambda: \mathbf{R}^\Lambda+\mathbf{R}^{\Lambda^\perp}\to\mathbf{R}^\Lambda+\mathbf{R}^{\Lambda^\perp},\quad \Lambda^\perp=\{1,\dots,d\}\setminus\Lambda,$$

acting as $\mathbf{R}^\Lambda\hookrightarrow\mathbf{R}^d\xrightarrow{D}\mathbf{R}^d\xrightarrow{\perp\text{-proj}}\mathbf{R}^\Lambda$ on the first component and as the identity on the second component. When there is no risk for confusion we will use D_Λ also to denote its first component.

Let now $D(\theta)$ be symmetric and smoothly parametrized by θ on an interval I and let

$$|D|_{C^\nu}<C_1K_1^\nu\quad\forall\nu\geqslant 0,\tag{1}$$

for some C_1 and $K_1\geqslant 1$. The dependence in the parameter θ of the eigenvectors, and more generally of the invariant subspaces, is related to the spacing between the eigenvalues. This is described in the following lemma which gives estimates on "block diagonalization" of D.

LEMMA 1. *For all $0<\delta<1$, there is a smooth orthogonal matrix $Q(\theta)$ such that for all θ in I, $\tilde D(\theta)=Q^*(\theta)D(\theta)Q(\theta)$ is a product of commuting blocks $\prod_{i=1}^p\tilde D(\theta)_{\Lambda_i}$ which are such that*

$$\begin{cases} E_m(\theta) \text{ and } E_n(\theta) \text{ are eigenvalues of the same block } D(\theta)_{\Lambda_i}\\ \quad\Rightarrow |E_m(\theta)-E_n(\theta)|<4d\delta,\\ E_m(\theta) \text{ and } E_n(\theta) \text{ are eigenvalues of different blocks } D(\theta)_{\Lambda_i} \text{ and } D(\theta)_{\Lambda_j}\\ \quad\Rightarrow |E_m(\theta)-E_n(\theta)|\geqslant\delta,\end{cases}$$

and such that

$$\begin{aligned}&|Q|_{C^\nu}\leqslant K_2^\nu\quad\forall\nu\geqslant 0,\\ &K_2=\gamma_1(C_1/\delta)^{d(d+1)}K_1,\end{aligned}\tag{2}$$

where $\gamma_1=\gamma_1(\gamma_0,d)$ increases superexponentially in d, and γ_0 is defined in Lemma A1. *The block decomposition $\bigcup\Lambda_i=\{1,\dots,d\}$ depends on θ in a piecewise constant way. Moreover, the block decomposition is constant on intervals of length δ/C_1K_1.*

Proof. Choose a family of continuous eigenvalues $E_1(\theta), \ldots, E_d(\theta)$. Fix a θ, say $\theta=0$, and decompose the eigenvalues into groups

$$\begin{cases} E_1(\theta), \ldots, E_{k_1}(\theta), \\ E_{k_1+1}(\theta), \ldots, E_{k_2}(\theta), \\ \vdots \\ E_{k_{p-1}+1}(\theta), \ldots, E_d(\theta), \end{cases}$$

such that any eigenvalue of one group is separated from any eigenvalue of any other group by at least 2δ and any two eigenvalues of one group are separated by not more than $2d\delta$. It follows that on a whole symmetric neighborhood $I' \ni 0$ of length $2\delta' < \delta / C_1 K_1$, any eigenvalue of one group is separated from any eigenvalue of any other group by at least δ, and any two eigenvalues of one group are separated by not more than $(2d+1)\delta$. We now restrict the discussion to the smaller interval I'.

By Lemma A2, the characteristic polynomial $P(\lambda, \theta)$ of $D(\theta)$ satisfies

$$|P(\lambda, \cdot)|_{C^\nu} < \gamma_2 C_1^d K_1^\nu \quad \forall \nu \geqslant 0,$$

in $|\lambda| < R = C_1 + 1$; all γ-constants will depend on γ_0 and increase superexponentially in d. We can write $P(\lambda, \theta) = P_1(\lambda, \theta) \cdot \breve{P}_1(\lambda, \theta)$, where

$$\breve{P}_1(\lambda, \theta) = \prod_{i > k_1} (\lambda - E_i(\theta)) = \sum_{j=0}^{d-k_1} e_{d-k_1-j} \lambda^j.$$

Since all eigenvalues lie in $|E_i| < R-1$ we get by Lemma A4

$$|e_j|_{C^\nu} < \left(\gamma_3 R C_1^d \frac{1}{\delta^{d-1}} \right)^j \left(\gamma_3 C_1^d \frac{1}{\delta^d} K_1 \right)^\nu,$$

and the corresponding result holds of course for the other groups of eigenvalues.

Let $q^1, \ldots, q^d$ be a positively oriented orthonormal basis of eigenvectors for $D(0)$ and define

$$v^m(\theta) = \begin{cases} \breve{P}_1(D(\theta), \theta) q^m / \breve{P}_1(E_m(0), 0), & m \leqslant k_1, \\ \vdots \\ \breve{P}_p(D(\theta), \theta) q^m / \breve{P}_p(E_m(0), 0), & k_{p-1} < m \leqslant d. \end{cases}$$

Then $v^1, \ldots, v^d$ will be an orthonormal basis for $\theta = 0$ and will span p orthogonal subspaces for all $\theta \in I'$. Moreover,

$$|v^m|_{C^\nu} < \gamma_4 \left(\frac{C_1}{\delta} \right)^{d^2} \left(\gamma_4 \left(\frac{C_1}{\delta} \right)^d K_1 \right)^\nu.$$

It follows that

$$|v^m(\theta)-v^m(0)|\leqslant \gamma_4\left(\frac{C_1}{\delta}\right)^{d(d+1)} K_1|\theta| < \frac{1}{4d^2 3^{d+5}}$$

for all $\theta \in I'$ if only

$$\delta' < \frac{\delta^{d(d+1)}}{\gamma_5 C_1^{d(d+1)} K_1}.$$

Moreover, $v^1, ..., v^d$ will be a basis also on I'—indeed on this interval we get

$$|\langle v^m, v^n\rangle|_{C^0} < \frac{1}{4d^2 3^{d+4}}, \quad m\neq n,$$

$$1-\frac{1}{4d^2 3^{d+4}} < |v^m|_{C^0} < 1+\frac{1}{4d^2 3^{d+4}}.$$

Now we do Gram–Schmidt orthogonalization on this basis, obtaining an orthonormal basis $\hat{v}^1, ..., \hat{v}^d$ on I'. By Lemma A3,

$$Q=(\hat{v}^1, ..., \hat{v}^d)$$

will verify the lemma with the estimate

$$\left(\gamma_6\left(\frac{C_1}{\delta}\right)^{d(d+1)} K_1\right)^{\nu} = (K_1')^{\nu}.$$

Suppose now that we have made two such constructions of orthogonal matrices Q_1, Q_2 on two intervals I_1' and I_2' that intersect each other on an interval $\hat{I}$ of length δ', let us say $I_1'=]-\delta', \delta'[$ and $I_2'=]0, 2\delta'[$. To Q_1 and Q_2 are associated two decompositions $\bigcup \Lambda_i^1 = \bigcup \Lambda_i^2 = \{1, ..., d\}$. Let us first assume that the decomposition of Q_1 is finer than that of Q_2, so that Q_1 also has the block decomposition $\bigcup \Lambda_i^2$. We want to interpolate Q_1 on $\theta<\frac{1}{3}\delta'$ with Q_2 on $\theta>\frac{2}{3}\delta'$ by a matrix Q such that Q^*DQ is block diagonal with block decomposition $\bigcup \Lambda_i^2$. For this we consider $P=Q_1^*Q_2$. It is a block-diagonal matrix with block decomposition $\bigcup \Lambda_i^2$ and we need to interpolate it with the identity matrix by a block matrix B with the same blocks as P. Then we can take $Q=Q_1B$.

Write $P(\theta)=P(\delta')(I+\widehat{P}(\theta))$. Then

$$|\widehat{P}|_{C^\nu} < 4\delta' K_1'(\gamma_7 K_1')^{\nu}.$$

Let now $e^{S_o}=P(\delta')$ and $e^S=I+\widehat{P}$, and let $\widetilde{S}_o=\phi S_o$, $\widetilde{S}=\phi S$, where ϕ is a C^∞-cut-off function which is 0 for $\theta<\frac{1}{3}\delta'$ and 1 for $\theta>\frac{2}{3}\delta'$. Now S can be expressed as a power series in $\widehat{P}$ each term of which can be estimated by Lemma A2, ϕ is estimated by Lemma A1 and the estimate of $\widetilde{S}$ is again obtained by Lemma A2. This gives

$$|S|_{C^\nu} < 4^2\delta' K_1'\left(\frac{1}{\delta'}\right)^{\nu}, \quad |\widetilde{S}|_{C^\nu} < 4^3\delta' K_1'\left(\frac{3\gamma_0}{\delta'}\right)^{\nu}, \quad |e^{\widetilde{S}}|_{C^\nu} < 2\left(\frac{3\gamma_0}{\delta'}\right)^{\nu},$$

if $\delta'\gamma_8 K_1' < 1$—the same estimate holds for $e^{\widetilde{S}_o}$ with the only difference that the factor 2 is replaced by another constant. $e^{\widetilde{S}_o} e^{\widetilde{S}}$ is the matrix B.

If Q_1' and Q_2' are defined by decompositions of which neither one are finer than the other, then we can construct on $[0, \delta']$ a finer decomposition and an orthogonal matrix Q_3 associated to this decomposition. Using it we can now interpolate Q_1 and Q_2 via Q_3. We leave the details. □

Let F be a symmetric $(d \times d)$-matrix, smoothly parametrized by θ, and assume that

$$|F|_{C^\nu} < C_3 K_3^\nu \quad \forall \nu \geqslant 0. \tag{3}$$

LEMMA 2. *For all* $0<\delta<1$ *there are a smooth symmetric matrix* G *and a smooth anti-symmetric matrix* X *on* I *such that*

$$\langle q(\theta), G(\theta) q'(\theta)\rangle = 0 \quad \textit{if } |E(\theta) - E'(\theta)| \geqslant 4d\delta,$$

for any two eigenvalues $E(\theta), E'(\theta)$ *and corresponding eigenvectors* $q(\theta), q'(\theta)$, *and*

$$[X, D] = F - G,$$

and such that

$$\begin{aligned} \delta |X|_{C^\nu} + |G|_{C^\nu} &< \gamma_2 C_3 K_4^\nu \quad \forall \nu \geqslant 0, \\ K_4 &= \max(\gamma_2 (C_1/\delta)^{d(d+1)+1} K_1, K_3), \end{aligned} \tag{4}$$

where $\gamma_2(\gamma_0, d)$ *increases superexponentially in* d. *Moreover, if* F *has compact support in* I, *then* G *and* X *also.*

Proof. Fix an interval I' of length $\delta' = K_2^{-1}$, defined in Lemma 1, and let Q be an orthogonal matrix on I' with constant decomposition $\bigcup \Lambda_i$ as in Lemma 1. Let $\widetilde{F} = Q^* F Q$, $\widetilde{D} = Q^* D Q$ and define

$$\widetilde{G}_{m,n} = \begin{cases} \widetilde{F}_{m,n} & \text{if } m, n \text{ belongs to the same } \Lambda_i, \\ 0 & \text{otherwise.} \end{cases}$$

$\widetilde{F}$ and $\widetilde{D}$ are estimated by Lemma A2, and it follows that

$$|\widetilde{G}|_{C^\nu} < \gamma_3 C_3 \max(K_2, K_3)^\nu.$$

Let

$$\begin{aligned} \widetilde{X}_{\Lambda_i, \Lambda_j} \widetilde{D}_{\Lambda_j} - \widetilde{D}_{\Lambda_i} \widetilde{X}_{\Lambda_i, \Lambda_j} &= (\widetilde{F} - \widetilde{G})_{\Lambda_i, \Lambda_j}, \\ \widetilde{X}_{m,n} = 0 \quad &\text{if } m, n \text{ belongs to the same } \Lambda_i. \end{aligned}$$

Since the eigenvalues of $\widetilde{D}_{\Lambda_i}$ and $\widetilde{D}_{\Lambda_j}$ are more than δ apart we get

$$|\widetilde{X}|_{C^0} \leqslant \frac{\gamma_4}{\delta}|\widetilde{F}-\widetilde{G}|_{C^0}.$$

Differentiating the equality gives

$$\partial^\nu \widetilde{X}_{\Lambda_i,\Lambda_j}\widetilde{D}_{\Lambda_j}-\widetilde{D}_{\Lambda_i}\partial^\nu \widetilde{X}_{\Lambda_i,\Lambda_j} = \partial^\nu(\widetilde{F}-\widetilde{G})_{\Lambda_i,\Lambda_j}-\sum_{i=1}^{\nu}\binom{\nu}{i}(\partial^{\nu-i}\widetilde{X}_{\Lambda_i,\Lambda_j}\partial^i\widetilde{D}_{\Lambda_j}-\partial^i\widetilde{D}_{\Lambda_i}\partial^{\nu-i}\widetilde{X}_{\Lambda_i,\Lambda_j}),$$

from which the estimate

$$|\widetilde{X}|_{C^\nu} < \gamma_5\frac{1}{\delta}C_3\max\left(\gamma_5\frac{C_1}{\delta}K_2, K_3\right)^\nu$$

follows by induction.

Choose now a partition of unity $\{\phi_j\}$ with $\operatorname{supp}(\phi_j)\subset I'_j$, $|I'_j|<\delta'$. On each I'_j we have X_j and G_j such that

$$\begin{aligned}&[X_j(\theta), D(\theta)] = F(\theta)-G_j(\theta),\\ &\langle q(\theta), G_j(\theta)q'(\theta)\rangle = 0 \quad \text{if } |E(\theta)-E'(\theta)|\geqslant 4d\delta,\end{aligned}$$

with the above estimates. Now we only need to set $X=\sum\phi_jX_j$ and $G=\sum\phi_jG_j$, and we obtain the estimates from Lemmas A1 and A2. □

3. Transversality

In this section we shall analyze a kind of *transversality* condition which should be fulfilled by the spectrum of our matrices. Such a condition was considered by Pyartli in relation to Diophantine approximation [12].(1)

Let I be an interval and let u be a smooth function on I such that

$$\begin{aligned}&|u|_{C^\nu} < C_5K_5^\nu \quad \forall\nu\leqslant m+1,\\ &\max_{0\leqslant\nu\leqslant m}\left|\frac{1}{(\nu!)^2K_5^\nu}\partial^\nu u(\theta)\right|\geqslant\beta \quad \forall\theta\in I.\end{aligned} \tag{5}$$

(1) A similar condition has been used by N. Nekhoroshev under the name of steepness and by R. Krikorian under the name of Pyartli transversality.

LEMMA 3. *For all $\varepsilon>0$, there is a disjoint union of intervals $\bigcup_{i\in J} I_i$ such that*

$$\#J \leqslant 2^m\left[\frac{2C_5K_5(m+1)^2}{\beta}|I|+1\right],$$

$$\max_{i\in J}|I_i| \leqslant \frac{2}{K_5}\left(\frac{2\varepsilon}{\beta}\right)^{1/m},$$

$$|u(\theta)|\geqslant \varepsilon \quad \forall\theta\in I\setminus\bigcup I_i.$$

Notice that the estimate of $\#J$ is independent of ε.

Proof. Assume for simplicity that $0\in I$ and $|\partial^m u(0)|\geqslant\beta(m!)^2K_5^m$. Then there is an interval $\tilde I$ of length

$$l=\frac{\beta}{2C_5K_5(m+1)^2}$$

such that $|\partial^m u(\theta)|\geqslant\frac12\beta(m!)^2K_5^m$ on $\tilde I$.

Consider now $\partial^{m-1}u$ on $\tilde I$. There exists an interval I_1 such that

$$|\partial^{m-1}u(\theta)|\geqslant\tfrac12\beta\left(\frac{2\varepsilon}{\beta}\right)^{1/m}((m-1)!)^2K_5^{m-1}\quad \forall\theta\in\tilde I\setminus I_1,$$

$$|I_1|<\frac{1}{K_5}\left(\frac{2\varepsilon}{\beta}\right)^{1/m}.$$

Consider now $\partial^{m-2}u$ on $\tilde I\setminus I_1$. There exist two intervals I_2, I_3 such that

$$|\partial^{m-2}u(\theta)|\geqslant\tfrac12\beta\left(\frac{2\varepsilon}{\beta}\right)^{2/m}((m-2)!)^2K_5^{m-2}\quad \forall\theta\in\tilde I\setminus I_1\cup I_2\cup I_3,$$

$$|I_i|<\frac{1}{K_5}\left(\frac{2\varepsilon}{\beta}\right)^{1/m},$$

etc.

Hence we obtain in $\tilde I$, 2^m-1 (possibly void) intervals I_i such that

$$|u(\theta)|\geqslant\varepsilon\quad\forall\theta\in\tilde I\setminus\bigcup I_i,$$

$$|I_i|<\frac{2}{K_5}\left(\frac{2\varepsilon}{\beta}\right)^{1/m}.$$

On the whole interval I we get at most

$$2^m\times\left[\frac{2C_5K_5(m+1)^2}{\beta}|I|+1\right]$$

many such intervals. □

Let now u_j be a sequence of smooth functions defined on an open interval I and satisfying

$$\begin{aligned} &|u_j|_{C^\nu} < C_6 K_6^\nu \quad \forall \nu \leqslant N = m_1 + \ldots + m_J, \\ &\max_{0\leqslant\nu\leqslant m_j} \left| \frac{1}{(\nu!)^2 K_6^\nu} \partial^\nu u_j(\theta) \right| \geqslant \beta \quad \forall \theta \in I, \end{aligned} \tag{6}$$

with $C_6 \geqslant 1$.

LEMMA 4. *If* $v = u_1 \cdot \ldots \cdot u_J$ *then*

$$|v|_{C^\nu} < 4^{J-1} C_6^J K_6^\nu \quad \forall \nu \leqslant N = m_1 + \ldots + m_J,$$

$$\max_{0\leqslant\nu\leqslant N} \left| \frac{1}{(\nu!)^2 K_6^\nu} \partial^\nu v(\theta) \right| \geqslant \left(\frac{1}{e}\right)^{J^{2N}} \left(\frac{\beta}{C_6^2}\right)^{J^{N+1}} \quad \forall \theta \in I.$$

Proof. The first part follows from Lemma A2 in the appendix so we concentrate on the second estimate. Fix a θ, say $\theta=0$. We can assume without restriction that

$$\frac{1}{K_6^{m_i}(m_i!)^2} |\partial^{m_i} u_i(0)| \geqslant \beta \quad \forall i.$$

Then

$$\frac{1}{K_6^N (N!)^2} \partial^N (u_1 \cdot \ldots \cdot u_J)(0) = \sum_{\iota_1+\ldots+\iota_J=N} \frac{\iota_1! \cdot \ldots \cdot \iota_J!}{N!} \left[\frac{1}{K_6^{\iota_1}(\iota_1!)^2} \partial^{\iota_1} u_1(0) \cdot \ldots \cdot \frac{1}{K_6^{\iota_J}(\iota_J!)^2} \partial^{\iota_J} u_J(0) \right].$$

Consider now a vector $(\iota_1, \ldots, \iota_J)$ in $\mathbf{Z}_+^J$ such that

$$\iota_1 + \ldots + \iota_J = N \quad \text{and} \quad |m_1 - \iota_1| + \ldots + |m_J - \iota_J| = 2l.$$

Then

$$\min(m_1, \iota_1) + \ldots + \min(m_J, \iota_J) = N - l$$

and

$$\frac{1}{N^l} \leqslant \frac{m_1! \cdot \ldots \cdot m_J!}{\iota_1! \cdot \ldots \cdot \iota_J!} \leqslant N^l,$$

and there are $c_l \leqslant J^{2l}$ many such vectors.

Since at least one of the brackets in the sum is larger than β^J, the sum will be larger than $\frac{1}{2}\beta^J \cdot m_1! \cdot \ldots \cdot m_J!/N!$ unless for some $|m_1 - \iota_1| + \ldots + |m_J - \iota_J| = 2l \neq 0$,

$$\left| \left[\frac{1}{K_6^{\iota_1}(\iota_1!)^2} \partial^{\iota_1} u_1(0) \cdot \ldots \cdot \frac{1}{K_6^{\iota_J}(\iota_J!)^2} \partial^{\iota_J} u_J(0) \right] \right| \geqslant \frac{1}{2(2J^2)^l} \cdot \frac{m_1! \cdot \ldots \cdot m_J!}{\iota_1! \cdot \ldots \cdot \iota_J!} \beta^J.$$

But this implies that

$$\left|\frac{1}{K_6^{\iota_i}(\iota_i!)^2}\partial^{\iota_i}u_i(0)\right| \geqslant \left(\frac{1}{4C_6^J J^2 N}\right)^l \beta^J \quad \forall i.$$

We now have an estimate for derivatives of smaller order, $\min(m_i, \iota_i)$, and the result follows by induction. (This induction is immediate for $N \geqslant 3$ and requires a little more details when $N=1$ and 2.) □

4. Normal forms

Consider a symmetric matrix

$$D: \mathbf{Z}\times\mathbf{Z}\times\mathbf{T}\to\mathbf{R},$$

smoothly parametrized by $\theta\in\mathbf{T}$ and satisfying the shift condition

$$D_{m+k,n+k}(\theta) = D_{m,n}(\theta+k\omega),$$

where ω is a Diophantine number,

$$\|k\omega\| \geqslant \frac{\varkappa}{|k|^\tau} \quad \forall k\in\mathbf{Z}\setminus\{0\},$$

with $\varkappa>0$, $\tau>1$.

(A) Estimates.

(A_1)

$$|D_{m,n}|_{\mathcal{C}^\nu} \leqslant \begin{cases} Ce^{-|m-n|r}K^\nu & \forall \nu\geqslant 0, \\ 0, & |m-n|\geqslant N. \end{cases} \tag{7}$$

(B) Block diagonalization. *There exist a symmetric interval* $I\ni 0$ *and a disjoint decomposition* $\bigcup \Lambda_i = \mathbf{Z}$ *such that*

(B_1) $\#\Lambda_i\leqslant M$ $\forall i$;

and there exists a smooth orthogonal matrix Q *on* I *such that the conjugated matrix* $\widetilde{D}=Q^*DQ$ *is a product of commuting blocks*

$$\prod \widetilde{D}_{\Lambda_i}(\theta) \quad \forall\theta\in I,$$

and

(B_2) $Q_{m,n}\neq 0$ *only if* $|m-n|\leqslant N$, *and* $\operatorname{diam}(\Lambda_i)\leqslant MN$ $\forall i$;

(B_3) *for all* m, $Q_{m,n}\not\equiv 0$ *for at most* M *different* n;

(B_4) $|Q|_{C^\nu}\leqslant K^\nu\ \forall\nu\geqslant 0$.

The decomposition, which is independent of θ, defines an equivalence relation $k\sim l$ on the integers and for each k we denote its equivalence class by $\Lambda(k)$.

To each Λ_i there is associated an invariant subspace $Q(\theta)(\mathbf{R}^{\Lambda_i})$ of $D(\theta)$. (B_1) bounds the dimension of this space. By (B_2) its diameter is at most $(M+2)N$, and by (B_3) at most M invariant subspaces $Q(\theta)(\mathbf{R}^{\Lambda_i})$ occupy a fixed site m, i.e. the mth component of $Q(\theta)(\mathbf{R}^{\Lambda_i})$ is identically 0 except for at most M many Λ_i's.

There are of course also one-dimensional invariant subspaces, but we have no control on their regularity. The regularity of $Q(\theta)(\mathbf{R}^{\Lambda_i})$, however, is controlled by (B_4).

We can extend Q to a neighborhood $I+k\omega$ of any point $k\omega$ by the shift condition. In this way we get a piecewise smooth Q defined on all $\mathbf{T}$ and, in particular, $D(\theta)$ is pure point with a complete set of eigenvectors for all θ. The most important assumptions will be related to these eigenvalues and their eigenspaces.

(C) Eigenvalues. *There is a piecewise smooth function* $E(\theta)$ *such that*

$$\{E(k\omega)\}_{k\in\Lambda_i} \text{ are the eigenvalues of } \widetilde{D}_{\Lambda_i}(0)\ \forall i,$$

and there are sets $\Omega_i\supset\Lambda_i$ *such that*

(C_1) *for all* n, *if* $\inf_{l\in\Lambda_i}|E(l\omega)-E(n\omega)|<\alpha$, *then*

$$n\omega\in m\omega+\tfrac{1}{2}I \quad \text{for some } m\in\Omega_i,$$
$$Q(\theta)(\mathbf{R}^{\Lambda(n)})\subset\mathbf{R}^{\Omega_i+n-m} \quad \forall\theta\in I;$$

(C_2) *for all* i, *the resultant*

$$u_{\Omega_i}(x,\theta)=\operatorname{Res}(\det(D(\theta+x)_{\Omega_i}-tI),\det(D(\theta)_{\Omega_i}-tI))$$

satisfies

$$|u_{\Omega_i}|_{C^\nu}<(4MC)^{2M^2}L^\nu \quad \forall\nu\leqslant sM^2+1,$$
$$\max_{0\leqslant\nu\leqslant sM^2}\left|\frac{1}{(\nu!)^2L^\nu}\,\partial_x^\nu u_{\Omega_i}(x,\theta)\right|\geqslant\beta \quad \forall x\ \forall\theta\in\mathbf{T};$$

(C_3) $\#\Omega_i\leqslant M\ \forall i$;

(C_4) *the intervals* $\{n\omega+I\}_{\operatorname{dist}(n,\Omega_i)<N}$ *are pairewise disjoint*;

(C_5) *diam*$(\Omega_i)\leqslant(1/\lambda)^{\tau+2}\ \forall i$;

(C_6) *for all i and for all $x\in I$,*

$$|u_{\Omega_i}|_{C^\nu} < \left(4M\cdot\tfrac{1}{2}C\right)^{2M^2} L^\nu \quad \forall \nu \leqslant sM^2+1,$$

$$\max_{0\leqslant\nu\leqslant sM^2}\left|\frac{1}{(\nu!)^2L^\nu}\partial_\theta^\nu u_{\Omega_i}(x,\theta)\right| \geqslant \beta\Big(\prod_{k,l\in\Omega_i}\|x+(k-l)\omega\|\Big) \quad \forall\theta\in\mathbf{T}.$$

The sets Ω_i play a double role: by the first condition of (C_1) they restrict the possible eigenvalues $E(n\omega)$ that are equal to $E(k\omega)$, $k\in\Lambda_i$, up to an approximation α— if $n\omega\notin\Omega_i\omega+\frac{1}{2}I$ then $|E(k\omega)-E(n\omega)|\geqslant\alpha$; by the second condition of (C_1) they also describe the location of the eigenvectors corresponding to such eigenvalues.

u_{Ω_i} measures the difference of the eigenvalues of the submatrices $D(\theta+x)_{\Omega_i}$ and $D(\theta)_{\Omega_i}$, and among these eigenvalues there are precisely certain eigenvalues of $D(\theta+x)$ and $D(\theta)$ respectively, when $\theta+x$, $\theta\in I$.

The norm of u_{Ω_i} is in (C_2) with respect to the variable x and in (C_6) with respect to the variable θ.

Remark 1. Hence, a normal form is an infinite-dimensional symmetric matrix, parametrized by $\theta\in\mathbf{T}$, satisfying a shift condition with respect to a Diophantine rotation of $\mathbf{T}$ and satisfying conditions (A), (B) and (C). The normal form depends on several parameters: C, K, r, N give upper estimates on D; M, N, K give information about Q; $\alpha, C, M, L, \beta, N, \lambda$, which give lower estimates on the spectrum. It also depends on the parameters $s\geqslant 2$ and $\varkappa>0$, $\tau>1$ (through ω) which will be kept fixed.

Remark 2. Capital letters can be increased and small letters decreased, but not independently because the parameters are interrelated. For example if we increase M we must also increase L, and if we increase L we must decrease β.

Remark 3. It follows from the generalized Young inequality [5] that (A_1) implies an estimate of D in the operator norm on $l^2(\mathbf{Z})\to l^2(\mathbf{Z})$:

$$|D(\theta)|_{C^\nu} < C\frac{e^r+1}{e^r-1}K^\nu < C\frac{4}{r}K^\nu,$$

if $r\leqslant 1$. Notice that this provides an estimate of the eigenvalues of $D(\theta)$ and of their first derivatives with respect to θ.

Remark 4. $u_{\Omega_i}(x,\theta)=\det P_i(\theta, D(x+\theta)_{\Omega_i})$, where $P_i(\theta,t)=\det(D(\theta)_{\Omega_i}-tI)$. The $\det(D(\theta)_{\Omega_i}-tI)$ is a sum of $\leqslant M!$ many monomials of degree $\#\Omega_i$ in the matrix elements of $D(\theta)_{\Omega_i}-tI$. Introducing $D(\theta+x)_{\Omega_i}$ for t we get a matrix whose elements are $\leqslant M!(M)^{M-1}$ many monomials of degree $\#\Omega_i$ in the matrix elements of these two matrices and in their differences. Taking the determinant gives this number to the power $\#\Omega_i$, times $(\#\Omega_i)!$. This shows that $u_{\Omega_i}(x,\theta)$ is a sum of less than M^{2M^2} many monomials of degree less than $\#\Omega_i\times\#\Omega_i$ in the components of $D(\theta)_{\Omega_i}$ and $D(\theta+x)_{\Omega_i}$ and their differences. In particular, if $C\geqslant 1$ it follows that

$$|u_{\Omega_i}|_{C^\nu} < (4MC)^{2M^2}K^\nu \quad \forall\nu\geqslant 0.$$

Idea of proof—continued. Consider a perturbation $D(\theta)+\varepsilon F_1(\theta)$ of a normal form. Condition (B) implies that $D(\theta)$ is pure point with finite-dimensional eigenvectors (i.e. eigenvectors that only occupy finitely many sites). Conditions (A) and (B) will enable us to solve the equation

$$[X_1(\theta), D(\theta)] = F_1(\theta) - G_1(\theta)$$

for some anti-symmetric matrix X_1 and some symmetric modification G_1—with reasonable estimates. (This will be done in Lemma 5.) Then, with $U(\theta,\varepsilon)=e^{\varepsilon X_1(\theta)}$ we have

$$U(\theta,\varepsilon)^*(D(\theta)+\varepsilon F_1(\theta))U(\theta,\varepsilon) = D(\theta)+\varepsilon G_1(\theta)+\mathbf{O}^2(\varepsilon). \tag{8}$$

Conditions (C_1)–(C_3) will then guarantee that the main remaining part $D+\varepsilon G_1$ is pure point and satisfies conditions (A) and (B) (with different parameters). The conditions (C_4)–(C_5) are added in order to assure that $D+\varepsilon G_1$ also satisfies (C) (with different parameters), and hence permits an inductive construction—condition (C_6) will play a role only when we take the limit of this construction. That $D+\varepsilon G_1$ will be in normal form will be proved in Lemma 7 and the iterative construction is carried out in Proposition 8.

For technical reasons we will in Lemma 7 consider the solution of (8) not only to the first order in ε but to some much higher order—such a solution is derived in Corollary 6.

Consider now a symmetric matrix F, smoothly parametrized on $\mathbf{T}$, satisfying the shift condition and such that

$$|F_{m,n}|_{C^\nu} < \varepsilon e^{-|m-n|\varrho}K_9^\nu \quad \forall\nu\geqslant 0. \tag{9}$$

Let D be in normal form and let $q^k(\theta)$ be an eigenvector of $D(\theta)$ corresponding to the eigenvalue $E(\theta+k\omega)$. The following result will not make any use of condition (C).

LEMMA 5. *Assume $r\leqslant 1\leqslant C$. For all $0<\delta<1$, there are a smooth symmetric matrix G and a smooth anti-symmetric matrix X on $\mathbf{T}$, both satisfying the shift condition, such that*

$$\langle q^m(\theta), G(\theta)q^n(\theta)\rangle = 0 \quad \textit{if } |E(\theta+m\omega)-E(\theta+n\omega)|\geqslant\delta,$$
$$[X,D] = F-G,$$

and for all $\nu \geqslant 0$,

$$\begin{aligned} r\delta|X_{m,n}|_{C^\nu}+|G_{m,n}|_{C^\nu} &< \gamma_3 N^2 e^{6MN\varrho}\varepsilon e^{-|m-n|\varrho} K_{10}^\nu, \\ K_{10} &= \max(\gamma_3 (C/r\delta)^{2M^2(2M^2+1)+1}K, \gamma_3 N^\tau, \gamma_3/|I|, K_9), \end{aligned} \tag{10}$$

where $\gamma_3(\gamma_0, \varkappa, \tau, M)$ *increases superexponentially in* M. *Moreover, if* $F_{m,n}\equiv 0$ *when* $|m-n|\geqslant J^2-(2M+4)N$, *then both* $X_{m,n}$ *and* $G_{m,n}\equiv 0$ *for* $|m-n|\geqslant J^2$.

Proof. Let us first consider the case when F has only two non-zero "diagonals":

$$F_{m,n}\equiv 0 \quad \text{unless } n-m=\pm k.$$

By a partition of unity we may arrange so that $F_{0,k}(\theta)$ is supported in a small interval I' of length $3t$ where

$$3t=\min\left(\frac{\varkappa}{(8(M+2)N)^\tau}, |I|\right),$$

and there is no restriction in assuming that I' is contained in the middle of I. Such a localized matrix will then satisfy

$$|F|_{C^\nu} < 2^3 \varepsilon e^{-|k|\varrho} \max(K_9, \gamma_0/t)^\nu,$$

where, we recall, the norm is the operator norm. (This follows from Lemmas A1–A2.)

Due to the shift condition and the Diophantine condition, only for few entries on each of the two non-zero "diagonals" will $F_{i,j}\not\equiv 0$. Indeed two entries on a "diagonal" for which $F_{i,j}\not\equiv 0$ will be more than $8(M+2)N$ sites apart.

Consider the matrix Q given in condition (B) and let $\widetilde{F}=Q^*FQ$ and $\widetilde{D}=Q^*DQ$. By (B_4) and Lemma A2 we get

$$\begin{aligned} |\widetilde{F}|_{C^\nu} &< 2^7 \varepsilon e^{-|k|\varrho}\max(K_9, \gamma_0/t, K)^\nu, \\ |\widetilde{D}|_{C^\nu} &< \frac{2^6}{r} C K^\nu. \end{aligned}$$

Most components of $\widetilde{F}$ are identically 0. In the strip

$$S:\ k-4(M+2)N \leqslant m+n \leqslant k+4(M+2)N,$$

for example, $\widetilde{F}_{m,n}\equiv 0$ except for (m,n) belonging to the two "squares"

$$B((k,0);N)=\{|m-k|\leqslant N, |n|\leqslant N\} \quad \text{and} \quad B((0,k);N).$$

(This follows from the first part of (B_2).) Moreover, in these "squares" $\widetilde{F}_{m,n}\equiv 0$ unless

$$(m,n)\in Z_k\times Z_0 \cup Z_0\times Z_k, \quad \#Z_i\leqslant M,\ i=0,k.$$

(This follows from (B_3).)

Assume now that $|k|>(2M+2)N$ and let Λ^i be the union of all Λ_ν such that $\Lambda_\nu\cap Z_i\neq 0$, $i=0,k$. Then for $i=0,k$,

$$\#\Lambda^i\leqslant M^2 \quad \text{and} \quad \Lambda^i\subset[i-(M+1)N, i+(M+1)N].$$

(This follows from (B_1) and the second part of (B_2).) In particular $\Lambda^0\cap\Lambda^k=\varnothing$.

By Lemma 2 we can solve on I'

$$[\overline{X}_\Lambda, \widetilde{D}_\Lambda]=\widetilde{F}_\Lambda-\overline{G}_\Lambda, \quad \Lambda=\Lambda^0\cup\Lambda^k,$$

with $d=2M^2$ and with δ replaced by $\delta/8M^2$. This gives

$$r\delta|\overline{X}_{m,n}|_{C^\nu}+|\overline{G}_{m,n}|_{C^\nu}<\gamma_4\varepsilon e^{-|k|\varrho}(K')^\nu \quad \forall\nu\geqslant 0,$$
$$K'=\max(\gamma_4(C/r\delta)^{2M^2(2M^2+1)+1}K, \max(K,\gamma_0/t,K_9)),$$

and $\overline{X}_\Lambda$ and $\overline{G}_\Lambda$ has compact support in I'.

Let us now define $\widetilde{X}_{m,n}=\overline{X}_{m,n}$ for $(m,n)\in\Lambda^k\times\Lambda^0\cup\Lambda^0\times\Lambda^k$ and $\widetilde{X}_{m,n}=0$ otherwise, and define $\widetilde{G}_{m,n}$ in the same way. Then

$$([\widetilde{X},\widetilde{D}])_{m,n}=\widetilde{F}_{m,n}-\widetilde{G}_{m,n}$$

for all $(m,n)\in S$. If we now let $X'=Q\widetilde{X}Q^*$ and $G'=Q\widetilde{G}Q^*$, then $X'_{m,n}$ and $G'_{m,n}$ are $\equiv 0$ unless (m,n) belongs to the larger "squares" $B((k,0);(M+2)N)$ or $B((0,k);(M+2)N)$. Moreover, on I',

$$([X',D'])_{m,n}=F_{m,n}-G'_{m,n}$$

for all (m,n) in the smaller strip

$$S':\ k-(4M+6)N\leqslant m+n\leqslant k+(4M+6)N.$$

(Here we use again the first part of (B_2).)

In order to get a solution we now extend X' to a matrix X defined for all θ satisfying the shift condition and such that

$$X_{m,n}(\theta)=X'_{m,n}(\theta), \quad (m,n)\in S,\ \theta\in I',$$

and we do the same for G'. This gives the solution. In order to recover the exponential decay we must multiply the estimates by $e^{(2M+4)N\varrho}\leqslant e^{6MN\varrho}$.

In the case $|k|\leqslant(2M+2)N$, Λ^0 and Λ^k may not be disjoint. But then $\Lambda\times\Lambda\subset S$ and we define $\widetilde{X}_{m,n}$ to be $\overline{X}_{m,n}$ if $(m,n)\in\Lambda\times\Lambda$, and to be 0 otherwise—and the same for $\widetilde{G}_{m,n}$. With this difference the construction of the solution is the same in both cases.

Suppose now that F has two non-zero "diagonals". Take a partition of unity $\{\phi_i\}$ supported in intervals I'_i of length $<3t$ (and such that no three intervals have non-void common intersection). Then we apply the above result to each $\phi_i F$ obtaining X_i and G_i. Then $\sum X_i$ and $\sum G_i$ will be a solution with estimates that are at most $\mathrm{const}\cdot(2M+4)N\leqslant\mathrm{const}\cdot MN$ larger than those of X_i and G_i respectively.

Suppose finally that F is arbitrary and decompose it into $\sum F_k$, where

$$(F_k)_{m,n}\equiv 0 \quad \text{unless } m-n=\pm k.$$

Then we construct X_k and G_k as above, and $\sum X_k$ and $\sum G_k$ will be solutions with estimates that are at most $\mathrm{const}\cdot(2M+4)N\leqslant\mathrm{const}\cdot MN$ larger than those of X_k and G_k respectively. This completes the proof of the lemma. □

Notice that the solution looses no smoothness with respect to F but only with respect to D. This is important in the next corollary, where we shall apply this construction several times.

COROLLARY 6. *Assume* $r,\varrho\leqslant 1\leqslant C$ *and*

$$1>6a+8b, \tag{11.1}$$

$$\varepsilon<\left(\frac{\varrho^3}{2^{16}\gamma_3 N^2 e^{6MN\varrho}C}\right)^8 \quad \text{and} \quad \varepsilon^a<\left(\frac{a\varrho}{2^5 MN}\right)^4, \tag{11.2}$$

$$c>(8M^4+4M^2+2)b, \tag{11.3}$$

$$\varepsilon^c<\min\left(\frac{1}{\gamma_3^2}\left(\frac{r}{C}\right)^{8M^4+4M^2+2}, \frac{1}{\gamma_3}N^{-\tau}K, \frac{1}{\gamma_3}|I|K, \frac{K}{K_9}\right). \tag{11.4}$$

Then there are a smooth orthogonal matrix U and a smooth symmetric matrix G on $\mathbf{T}$, both satisfying the shift condition, such that

$$U^*(D+F)U=D+G+F',$$
$$\langle q^m(\theta),G(\theta)q^n(\theta)\rangle=0 \quad \text{if } |E(\theta+m\omega)-E(\theta+n\omega)|\geqslant\varepsilon^b,$$
$$G_{m,n}\equiv 0 \quad \text{if } |m-n|\geqslant(1/\varepsilon)^a,$$

and for all $\nu\geqslant 0$,

$$|(U-I)_{m,n}|_{C^\nu}<\sqrt{\varepsilon}\,e^{-|m-n|\varrho/2}((1/\varepsilon)^cK)^\nu,$$
$$|G_{m,n}|_{C^\nu}<\sqrt{\varepsilon}\,e^{-|m-n|\varrho/2}((1/\varepsilon)^cK)^\nu,$$
$$|F'_{m,n}|_{C^\nu}<\varepsilon^{(1/\varepsilon)^{a/2}/2}e^{-|m-n|\varrho/2}((1/\varepsilon)^cK)^\nu.$$

Proof. Let us first fix $\delta=\varepsilon^b$ and $J=(1/\varepsilon)^{a/2}$. Let us then define for $j\geqslant 1$,

$$\begin{aligned}[X_j, D] &= \widehat{F}_j - \mathrm{Proj}_{D,\delta}(\widehat{F}_j),\\ G_j &= \mathrm{Proj}_{D,\delta}(\widehat{F}_j),\\ F_{j+1} &= e^{-X_j}\ldots e^{-X_1}(D+F)e^{X_1}\ldots e^{X_j} - D - (G_1+\ldots+G_j),\end{aligned}$$

where $F_1=F$, $\widehat{F}_j$ is the truncation of F_j at distance $(1/\varepsilon)^a-(2M+4)N$ from the main diagonal and $\mathrm{Proj}_{D,\delta}(\widehat{F}_j)$ is the G defined in Lemma 5. We shall estimate these matrices using this lemma.

If K_{10} is defined by (10) and if

$$\varrho_1=\varrho,\quad \varrho_2=\varrho_1-\frac{\varrho}{2J},\quad \varepsilon_1=\varepsilon,\quad \varepsilon_2=\varepsilon^{3/2},$$

then

$$\begin{aligned}|(G_1)_{m,n}|_{C^\nu} &< A\varepsilon_1 e^{-|m-n|\varrho_1}K_{10}^\nu,\\ |(X_1)_{m,n}|_{C^\nu} &< \frac{1}{r\delta}A\varepsilon_1 e^{-|m-n|\varrho_1}K_{10}^\nu,\end{aligned}$$

with $A=\gamma_3 N^2 e^{6MN\varrho_1}$. In order to estimate $e^{\pm X_1}-I$ we note that an i-fold product of X_1's satisfies

$$|(X_1\cdot\ldots\cdot X_1)_{m,n}|_{C^\nu} < \frac{1}{4}\left(\frac{4^2}{\varrho_1-\varrho_2}\cdot\frac{1}{r\delta}A\varepsilon_1\right)^i e^{-|m-n|\varrho_2}K_{10}^\nu.$$

Using the power series expansion of $e^{\pm X_1}$ we get

$$|(e^{\pm X_1}-I)_{m,n}|_{C^\nu} < \frac{2^5 J}{\varrho}\cdot\frac{1}{r\delta}A\varepsilon_1 e^{-|m-n|\varrho_2}K_{10}^\nu$$

under conditions (11.1)–(11.2). Using power series expansions and estimating products, we also get

$$|(F_2)_{m,n}|_{C^\nu} < \varepsilon_2 e^{-|m-n|\varrho_2}K_{10}^\nu.$$

(Here we used the second part of (11.2) in order to estimate the difference $F_1-\widehat{F}_1$.)

Proceeding in the same way with

$$\varrho_j=\varrho-\frac{\varrho}{2J}(j-1)\quad\text{and}\quad \varepsilon_{j+1}=\varepsilon^{(j+2)/2},$$

we get, under conditions (11.1)–(11.2), that

$$\begin{aligned}|(G_j)_{m,n}|_{C^\nu} &< A\varepsilon_j e^{-|m-n|\varrho_j}K_{10}^\nu,\\ |(X_j)_{m,n}|_{C^\nu} &< \frac{1}{r\delta}A\varepsilon_j e^{-|m-n|\varrho_j}K_{10}^\nu,\\ |(e^{\pm X_j}-I)_{m,n}|_{C^\nu} &< \frac{2^5 J}{\varrho}\cdot\frac{1}{r\delta}A\varepsilon_j e^{-|m-n|\varrho_{j+1}}K_{10}^\nu,\\ |(F_{j+1})_{m,n}|_{C^\nu} &< \varepsilon_{j+1}e^{-|m-n|\varrho_{j+1}}K_{10}^\nu.\end{aligned}$$

By (11.3)–(11.4) we get $K_{10}<(1/\varepsilon)^c K$.

We now let $G=G_1+\ldots+G_J$, $U=e^{X_1}\ldots e^{X_J}$ and $F'=F_{J+1}$. These matrices will satisfy the estimates of the corollary. □

5. The inductive lemma

LEMMA 7. *Let D be in normal form on an interval I with parameters $C, K, r, M, N, \alpha, L, \beta, \lambda$, and let $a<b<c$ be numbers restricted by*

$$\frac{1}{\tau M^{3sM^3}} \leqslant a < \frac{b}{20s\tau M^4} < \frac{c}{100s^2\tau M^8}, \quad c \leqslant \frac{1}{5sM^{2sM^3}}.$$

Assume, as simplification, that

$$1 \leqslant L \leqslant K, \quad 8 \leqslant M, \quad 1<C<2, \quad r, \alpha, \beta \leqslant 1.$$

Let F be a symmetric matrix satisfying the shift condition. Assume that

$$\lambda \leqslant |I| \leqslant \beta/L,$$
$$|F_{m,n}|_{C^\nu} < \varepsilon e^{-|m-n|r} K^\nu \quad \forall \nu \geqslant 0.$$

Then there is a constant $\Gamma=\Gamma(\varkappa, \tau, s, \gamma_0, M)$, superexponentially decaying in M, such that if

$$|\varepsilon| < \Gamma \left[\frac{r^\tau \alpha \beta \lambda^{\tau^2}}{K N^\tau} e^{-N r} \right]^{e^{e^{sM^4}}},$$

then there is a smooth orthogonal matrix U, satisfying the shift condition, such that

$$|(U-I)_{m,n}|_{C^\nu} < \sqrt{\varepsilon}\, e^{-|m-n|r'} (K')^\nu$$

and

$$U^*(D+F)U = D' + F',$$

with D' in normal form on an interval I', with parameters

$$C' = (1+\varepsilon^{1/2})C, \quad K' = (1/\varepsilon)^c K, \quad r' = \tfrac{1}{2} r,$$
$$\lambda' = \left(\tfrac{1}{9}\right)^{M'} \varepsilon^b, \quad M' = M^{sM^3}, \quad N' = (1/\varepsilon)^a,$$
$$\alpha' = \varepsilon^c, \quad L' = K, \quad \beta' = \varepsilon^b K,$$

and with

$$2\lambda' \leqslant |I'| \leqslant \varepsilon^b,$$
$$|F'_{m,n}|_{C^\nu} < \varepsilon^{(1/\varepsilon)^{a/2}/2} e^{-|m-n|r'} (K')^\nu.$$

In addition,

$$|E(l\omega)-E(k\omega)| < M'\frac{K}{r}\varepsilon^b \quad \forall l\in\Lambda'(k),$$
$$Q'(\theta)(\mathbf{R}^{\Lambda'(k)}) \subset \sum_{l\in\Lambda'(k)} Q(\theta)(\mathbf{R}^{\Lambda(l)}) \quad \forall\theta\in I',$$

D' is in normal form with the same parameters also on $\frac{1}{2}I'$.

Finally, if $M\geqslant 2\tau$ then the closure of the sets

$$\{n\omega : |E(n\omega)-E(n\omega+l\omega)| < 2M'(K/r)\varepsilon^b\} \quad \textit{for all } 4(1/\lambda)^{\tau+2} < |l| \leqslant M'N',$$
$$\{n\omega : |E(n\omega)-E(n\omega+l\omega)| < 2\varepsilon^{1/8}\} \quad \textit{for all } M'N' < |l| \leqslant 4(1/\lambda')^{\tau+2}$$

are unions of, respectively, at most $\varepsilon^{-b/5sM^2}$ and ε^{-M^4b} many components, each component being of length, respectively, at most $\varepsilon^{b/4sM^2}$ and ε^{2M^4b}.

In order to avoid any confusion, let us point out that Q, Λ_i, Ω_i and $Q', \Lambda_i', \Omega_i'$ are the orthogonal matrix and sets referred to in the definition of the normal forms of D and D' respectively. $\varkappa, \tau, s$ are constants which are the same for both D and D'.

Proof. In the sequel we shall, without further mentioning, denote by $\gamma_4, \gamma_5, \ldots$ constants that only depend on $\varkappa, \tau, s, \gamma_0, M$, and that *decay* superexponentially in M.

In the proof we shall verify several inequalities involving ε and the parameters $C, K, r, \ldots$. These inequalities will be fulfilled if Γ is small enough and it is important that Γ does not depend on the parameters but only on $\varkappa, \tau, s, \gamma_0$—we will not in general stress this any further.

Construction of U and D'. Condition (11) is fulfilled with $\varrho=r$ and $K_9=K$ since

$$\varepsilon^{\tau a} < \gamma_4\left(\frac{r^\tau\lambda^\tau}{N^\tau}\right)^4. \tag{11'}$$

(Here we used $r\leqslant 1<C<2$ and $\lambda\leqslant|I|$.) So we can apply Corollary 6. Clearly U, $D'=D+G$ and F' satisfy the required estimates and it remains to verify that D' satisfies conditions (B) and (C) for the appropriate parameters.

Construction of Λ_i'. We define an equivalence relation $\sim'$ on $\mathbf{Z}$ by declaring that two integers k and l are equivalent if $|k-l|<2N'$ and

$$|E(k\omega)-E(l\omega)| < 2^9M\frac{K}{r}\varepsilon^b,$$

or if there is a sequence $k=k_1, k_2, \ldots, k_{n+1}=l$ such that each consecutive pair (k_j, k_{j+1}) is equivalent in the first sense. This equivalence relation defines the decomposition $\bigcup\Lambda_i'=\mathbf{Z}$. We shall show that condition (B) is fulfilled for any interval I' of length $\leqslant\varepsilon^b$.

Verification of (B_1). Let $k\in\Lambda_i$ be given and consider a sequence of integers $k=k_1\sim' k_2\sim' \ldots \sim' k_n$, where $|k_j-k_{j+1}|<2N'$. We shall show that the number n of such integers must be $\leqslant M'/2^9M\leqslant \frac{1}{2}M'$, so assume that $n=M'/2^9M$.

We have that

$$\eta := M'\frac{K}{r}\varepsilon^b<\alpha$$

if

$$\varepsilon^b<\gamma_5\frac{\alpha r}{K}. \tag{12}$$

(For any given constant γ_5, (12) is fulfilled if Γ is small enough. Of course, it is sufficient that $\gamma_5=1/M'$, but since we will refer to a similar condition several times we prefer to give (12) a more general form. This practice will be used in the sequel without explicit mention.)

Any $l=k_j$ must satisfy

$$|E(k\omega)-E(l\omega)|<\eta$$

so $l\omega\in m\omega+\frac{1}{2}I$ for some $m\in\Omega_i$ by (C_1). Condition (C_1) also implies that the eigenvector $q^l(\theta)\in\mathbf{R}^{\Omega_i+l-m}$ for all $\theta\in I$. In particular, $E(l\omega)$ is an eigenvalue of $D_{\Omega_i+l-m}(0)=D_{\Omega_i}((l-m)\omega)$. Since $E(k\omega)$ is an eigenvalue of $D_{\Omega_i}(0)$ it follows that $(l-m)\omega$ must belong to the set

$$\{x\in I: |u_{\Omega_i}(x,0)|<\eta(16/r)^{M^2}\};$$

the eigenvalues of D_{Ω_i} are bounded by $C\cdot 4/r<8/r$ since $C<2$. By (C_2) and by Lemma 3 there are not more than

$$M_1 := 2^{sM^2}\left[\frac{2(8M)^{2M^2}L(sM^2+1)^2}{\beta}|I|+1\right]$$

many components of this set, each component being of length at most

$$\frac{2}{L}\left(\frac{16}{r}\right)^{1/s}\left(\frac{2\eta}{\beta}\right)^{1/sM^2}.$$

Hence, if $l\sim' k$ then $l\omega$ must belong to one out of at most M_1M many intervals of this length—because $\#\Omega_i\leqslant M$ by (C_3).

Due to the size of the intervals and the distance between the k_j's, no such interval can contain more than one number $k_j\omega$ since

$$b>2s\tau M^2a, \tag{13.1}$$

$$\varepsilon^{b/2}<\gamma_6\frac{\beta}{K}r^{M^2+1}. \tag{13.2}$$

This forces n to be smaller than $M_1 M$, which is smaller than $M'/2^9 M$. (Here we use $L|I|\leqslant\beta$ and $M\geqslant 8$.)

Construction of Q'. Consider now the block matrix (written symbolically)

$$\widetilde{D}=Q^* D Q=\begin{pmatrix}\ddots & & & & \\ & \widetilde{D}_{\Lambda_{-1}} & & & \\ & & \widetilde{D}_{\Lambda_0} & & \\ & & & \widetilde{D}_{\Lambda_1} & \\ & & & & \ddots\end{pmatrix},$$

where each block is defined on I and of dimension $\leqslant M$. We now block diagonalize these blocks using Lemma 1 with $d\leqslant M$ and $\delta=(2^7 K/r)\varepsilon^b$:

$$\widetilde{\widetilde{D}}=\widetilde{Q}^*\widetilde{D}\widetilde{Q}=\begin{pmatrix}\ddots & & & & \\ & \widetilde{\widetilde{D}}_{\Lambda_{-1}} & & & \\ & & \widetilde{\widetilde{D}}_{\Lambda_0} & & \\ & & & \widetilde{\widetilde{D}}_{\Lambda_1} & \\ & & & & \ddots\end{pmatrix},$$

where the blocks now decompose into subblocks

$$\widetilde{\widetilde{D}}_{\Lambda_i}=\begin{pmatrix}\ddots & & & & \\ & \widetilde{\widetilde{D}}_{\Lambda_{i,-1}} & & & \\ & & \widetilde{\widetilde{D}}_{\Lambda_{i,0}} & & \\ & & & \widetilde{\widetilde{D}}_{\Lambda_{i,1}} & \\ & & & & \ddots\end{pmatrix}.$$

Then $\widetilde{Q}$ will be a block matrix over the decomposition $\bigcup \Lambda_i=\mathbf{Z}$.

Since

$$|D|_{C^\nu}\leqslant\frac{2^3}{r}K^\nu \quad\text{and}\quad |\widetilde{D}|_{C^\nu}\leqslant\frac{2^7}{r}K^\nu=:C_1K_1^\nu,$$

the decomposition into subblocks is constant over intervals of length

$$\leqslant\frac{\delta}{C_1K_1}=\varepsilon^b,$$

i.e. over the interval I'. Hence, two eigenvalues $E(k\omega)$ and $E(n\omega)$ of a subblock $\widetilde{\widetilde{D}}_{\Lambda_{i,j}}(0)$ will be separated by at most

$$4d\delta\leqslant 2^9 M\frac{K}{r}\varepsilon^b.$$

Since the blocks have diameter less than MN and since $MN<2N'$ if for example

$$\varepsilon^a < \gamma_7 \frac{1}{N}, \tag{14}$$

it follows that $k\sim' n$. Hence, $\tilde{\tilde{D}}(\theta)$ decomposes as blocks over the decomposition $\mathbf{Z}=\bigcup \Lambda_i'$ for all $\theta \in I'$.

Let now $E_k(\theta)$ be a continuous branch of an eigenvalue of $\tilde{D}_{\Lambda(k)}(\theta)$ starting at $E(k\omega)$ for $\theta=0$. Suppose that $\langle q^k(\theta), G(\theta)q^l(\theta)\rangle \neq 0$ for some $\theta\in I'$, where $q^k(\theta), q^l(\theta)$ are eigenvectors corresponding to $E_k(\theta)$ and $E_l(\theta)$ respectively. Then we know by Corollary 6 that

$$|E_k(\theta)-E_l(\theta)| < \varepsilon^b,$$
$$|k-l| < (1/\varepsilon)^a + 2(M+2)N.$$

By (14) this implies—recall Remark 3—that

$$|E_k(0)-E_l(0)| < \varepsilon^b + \frac{16}{r} K\varepsilon^b \quad \text{and} \quad |k-l| < 2N',$$

i.e. $k\sim' l$. Hence, also $\tilde{\tilde{G}}(\theta)$ decomposes as blocks over the decomposition $\mathbf{Z}=\bigcup \Lambda_i'$ for each $\theta\in I'$, so with $Q'=Q\tilde{Q}$ we have

$$\tilde{D}' = (Q')^* D' Q' = \prod \tilde{D}'_{\Lambda_i'}.$$

Verification of (B_2)–(B_4). If $Q_{m,l}\tilde{Q}_{l,n} \not\equiv 0$ then

$$|n-m| \leqslant |n-l| + |l-m| \leqslant MN + N \leqslant N'$$

under condition (14). This proves the first part of (B_2)—the second part is fulfilled by construction, because we have seen $\#\Lambda_i' \leqslant \frac{1}{2}M'$.

Assume $Q_{m,l}\tilde{Q}_{l,n} \not\equiv 0$. For each m there are at most M different l's such that $Q_{m,l}\not\equiv 0$ by (B_3), and for each l there are at most M different n's such that $\tilde{Q}_{l,n}\not\equiv 0$ because $\tilde{Q}$ is a block matrix with blocks of dimension $\leqslant M$. Hence, there are at most M^2 many n's—this proves (B_3).

(B_4) follows easily from Lemma 1 if

$$c > 2(M^2+M)b, \tag{15.1}$$

$$\varepsilon^b < \gamma_8. \tag{15.2}$$

We now turn to condition (C). Let us first define $E'(\theta)$ from $E(\theta)$ by analyticity in the perturbation.

Construction of Ω'_i. The set Ω'_i is certainly not unique and we shall make a construction that depends on a choice of $k\in\Lambda'_i$—we denote by $\Omega(k)$ the particular Ω_j containing $\Lambda(k)$.

Since the eigenvalues of D and D' differ by at most $\sqrt{\varepsilon}\,C\cdot 4/r$ it follows that for any $l\sim' k$,

$$|E'(l\omega)-E'(n\omega)|<\alpha' \;\Rightarrow\; \begin{cases} |E(k\omega)-E(n\omega)|<\eta+\alpha'+\sqrt{\varepsilon}\cdot 16/r<2\eta<\alpha \\ \text{and} \\ |E(l\omega)-E(n\omega)|<\alpha'+\sqrt{\varepsilon}\cdot 16/r<2\alpha', \end{cases}$$

if (12) holds and if for example $2b<c<\frac{1}{4}$. Then $n\omega\in m_n\omega+\frac{1}{2}I$ and $l\omega\in m_l\omega+\frac{1}{2}I$ for some $m_n, m_l\in\Omega(k)$, and, as in the proof of (B_1), $n\omega-(l-m_l+m_n)\omega$ will belong to the set

$$\{x\in I: |u_{\Omega(k)}(x,l-m_l)|<2\alpha'(16/r)^{M^2}\}.$$

This set has at most M_1 (defined in the proof of (B_1)) many components, each of length less than

$$\frac{2}{L}\left(\frac{16}{r}\right)^{1/s}\left(\frac{4\alpha'}{\beta}\right)^{1/sM^2}.$$

This bound is $<\lambda'$ if

$$c>2sM^2b, \tag{16.1}$$

$$\varepsilon^{c/2}<\gamma_9 r^{M^2}\beta. \tag{16.2}$$

Hence, we can cover $X:=\bigcup_{l\in\Lambda'_i}\{n\omega: |E'(l\omega)-E'(n\omega)|<\alpha'\}$ by $\leqslant M^2M_1^2$ many intervals—because $\#\Omega(k)\leqslant M$ and $\#\Lambda'_i\leqslant MM_1$—each one of length λ' and intersecting $\Omega(k)\omega+\frac{1}{2}I$. If $X_1, X_2, \ldots$ are these intervals, then the first condition of (C_1) is fulfilled if $\Omega'_i\omega+\frac{1}{2}I'\supset\bigcup X_j$. The additional statement even says that $\Omega'_i\omega+\frac{1}{4}I'\supset\bigcup X_j$, which will require no extra work.

Let now

$$\Lambda'(X_j)=\{n: n+m\in\Lambda'(m), \text{ for some } m\omega\in X_j \text{ with } \inf_{l\in\Lambda'_i}|E'(l\omega)-E'(m\omega)|<\alpha'\}.$$

It is easy to see, as in the proof of (B_1), that $\#\Lambda'(X_j)\leqslant M_1M$. In fact, if $n,n'\in\Lambda'(X_j)$ then $|E(m\omega+n\omega)-E(k\omega)|<2\eta$ and $|E(m'\omega+n'\omega)-E(k\omega)|<2\eta$ for some $m\omega, m'\omega\in X_j$. By (12) it then follows that $(m+n)\omega$ and $(m'+n')\omega$ will belong to a union of at most MM_1 many intervals, which are of length $<\lambda'$ if (16) holds. Now if $n\neq n'$, then

$$|(m+n)\omega-(m'+n')\omega|\geqslant\frac{\varkappa}{(2M'N')^\tau}-\varepsilon^b,$$

which is $>\varepsilon^b$ if

$$b > 2\tau a, \tag{17.1}$$

$$\varepsilon^b < \gamma_{10}. \tag{17.2}$$

Hence, $(m+n)\omega$ and $(m'+n')\omega$ will belong to different intervals.

Choose now k_j such that $X_j \subset k_j\omega + \frac{1}{4}I'$ for each j—the distance of $k_j\omega$ from X_j must not exceed, say, $2^{-4}\varepsilon^b$ since we want $|I'|$ to be less than ε^b—we don't require that $k_j\omega \in X_j$ but if it is then we can choose $|I'| = 8\lambda'$.

If $n \in \Lambda'(X_j)$ then there is some $m\omega \in X_j$ such that $m+n \sim' m$ and

$$|E'(l\omega) - E'(m\omega)| < \alpha' \quad \text{for some } l \in \Lambda'_i.$$

Then

$$\begin{aligned} |E(m\omega+n\omega) - E(k\omega)| \leqslant & |E(m\omega+n\omega) - E(m\omega)| + |E(m\omega) - E'(m\omega)| \\ & + |E'(m\omega) - E'(l\omega)| + |E'(l\omega) - E(l\omega)| \\ & + |E(l\omega) - E(k\omega)| \leqslant \eta + \frac{8}{r}\sqrt{\varepsilon} + \alpha' + \frac{8}{r}\sqrt{\varepsilon} + \eta, \end{aligned}$$

which is less than α under condition (12). By (C_1) it follows that

$$(m+n)\omega \in p\omega + \tfrac{1}{2}I \quad \text{for some } p \in \Omega(k).$$

Since $\|(k_j - m)\omega\| < \varepsilon^b$ and $\lambda \leqslant |I|$ it follows that $(k_j+n)\omega \in p\omega + \frac{3}{4}I$ if

$$\varepsilon^b < \gamma_{11}\lambda. \tag{18}$$

Hence, if we define $[k_j+n]_{\Omega(k)}$ to be this integer p—notice that it is unique by (C_4)—condition (C_1) implies that

$$Q(\theta)(\mathbf{R}^{\Lambda(m+n)}) \subset \mathbf{R}^{\Omega(k)+m+n-[k_j+n]_{\Omega(k)}} \quad \forall \theta \in I.$$

By the construction of $Q' = Q\widetilde{Q}$ we have

$$\begin{aligned} Q'(\theta)(\mathbf{R}^{\Lambda'(m)}) & \subset \sum_{m+n\sim' m} Q'(\theta)(\mathbf{R}^{\Lambda(m+n)}) \subset \sum_{m+n \sim' m} Q(\theta)(\mathbf{R}^{\Lambda(m+n)}) \\ & \subset \sum_{m+n\sim' m} \mathbf{R}^{\Omega(k)+m+n-[k_j+n]_{\Omega(k)}} \\ & \subset \sum_{m+n\sim' m} \mathbf{R}^{\Omega(k)+k_j+n-[k_j+n]_{\Omega(k)}+(m-k_j)} \\ & \subset \mathbf{R}^{\Omega'_i+(m-k_j)}, \end{aligned}$$

if

$$\Omega_i' \supset \bigcup_{n\in\Lambda'(X_j)} [\Omega(k)+k_j+n-[k_j+n]_{\Omega(k)}].$$

So if we choose

$$\Omega_i' = \Omega(k)+\bigcup_j \{k_j+n-[k_j+n]_{\Omega(k)} : n\in\Lambda'(X_j)\} = \Omega(k)+\Delta(k)$$

then condition (C_1) will hold.

Notice again that $\Delta(k)\omega\subset\frac{3}{4}I$—this will be used in the verification of (C_6).

Since

$$\#\Omega_i' \leqslant \#\Omega(k)\cdot\#\Delta(k) \leqslant M\cdot M^2M_1^2\cdot MM_1 < M',$$

also (C_3) will hold. (Here we used that $M\geqslant 8$.)

For $l_1\neq l_2\in\Lambda_i'$,

$$\|(l_1-l_2)\omega\| \geqslant \frac{\varkappa}{|l_1-l_2|^\tau} \geqslant \frac{\varkappa}{(2M'N')^\tau},$$

which is $>\varepsilon^b$ if for example (17) holds. Hence, each $l\omega$, $l\in\Lambda_i'$, belongs to one and only one X_j. We can therefore include among the k_j's all elements of Λ_i' assuring that

$$\Omega_i' \supset \Lambda_i'.$$

Verification of (C_5). In order to verify (C_5) we must restrict the choice of the k_j's. If $X_j\cap\Lambda_i'\omega=\{l\omega\}$, then we are forced to take $k_j=l$. If $X_j\cap\Lambda_i'\omega=\varnothing$, then we have some freedom in the choice and we decide to choose $k_j\omega\in X_j$ in such a way that k_j is as close as possible to Λ_i'. There are one or two such choices. (Indeed, suppose that there are at least two such choices, k_1 and k_2 say. Then

$$\varepsilon^b \geqslant \|(k_1-k_2)\omega\| \geqslant \frac{\varkappa}{|k_1-k_2|^\tau},$$

so by (17),

$$|k_1-k_2| \geqslant \left(\frac{\varkappa}{\varepsilon^b}\right)^{1/\tau} > M'\frac{1}{\varepsilon^a} = \operatorname{diam}(\Lambda_i').$$

This means that Λ_i' must lie in between k_1 and k_2, which excludes a third choice.) If there are two, choose any one of them. It is an easy exercise, using the Diophantine property of ω to show that for any interval I' of length $\geqslant\lambda'$ and any sequence of integers $l, l+1, ..., l+n$ of length

$$\geqslant \frac{2\pi}{\varkappa}\left(\frac{2\pi}{\lambda'}\right)^{\tau+1},$$

there is (at least) one integer k in this sequence such that $k\omega \in I'$. Hence, we can choose the k_j's so that

$$\operatorname{dist}(k_j, \Lambda_i') \leqslant \frac{2\pi}{\varkappa}\left(\frac{2\pi}{\lambda'}\right)^{\tau+1}.$$

Then

$$\operatorname{diam}(\Omega_i') \leqslant \operatorname{diam}(\Lambda_i') + 2\frac{2\pi}{\varkappa}\left(\frac{2\pi}{\lambda'}\right)^{\tau+1} + 2M'N' + 2\operatorname{diam}(\Omega(k)) < \left(\frac{1}{\lambda'}\right)^{\tau+2}$$

under conditions (17)–(18).

Verification of (C_4). This is somewhat delicate and we will have to change the choice of the k_j's—in order to guarantee that $\Omega_i' \supset \Lambda_i'$ we are not allowed to change those k_j's that belong to Λ_i'—in order to cover X_j by $k_j\omega + \frac{1}{4}I'$ with $|I'| \leqslant \varepsilon^b$, we must keep $k_j\omega$ at distance less than $2^{-4}\varepsilon^b$ from X_j—in order to keep the bound (C_5) we cannot increase the k_j's too much.

Suppose that the intervals $\{m\omega + I' : |m - k_j| \leqslant 2M'N', j \geqslant 1\}$ are not pairwise disjoint. Then there are two k_j's, k_1 and k_2 say, such that $(k_1+n_1)\omega + I'$ and $(k_2+n_2)\omega + I'$ intersect for some $|n_1|, |n_2| \leqslant 2M'N'$, i.e.

$$\|k_1\omega - (k_2+n)\omega\| < |I'|, \quad |n| = |n_2 - n_1| \leqslant 4M'N'.$$

This implies that

$$|k_1 - k_2| \geqslant \left(\frac{\varkappa}{\varepsilon^b}\right)^{1/\tau} - 4M'N',$$

which is $> 2M'N'$ under condition (17). In particular, both k_1 and k_2 cannot belong to Λ_i'—if one does, let it be k_2 say. Moreover,

$$(k_2+n)\omega + \tfrac{1}{4}\cdot 9I' \supset k_1\omega + \tfrac{1}{4}I' \supset X_1,$$

and for all $|m| < 4M'N'$, $m \neq n$,

$$\|k_1\omega - (k_2+m)\omega\| \geqslant \|(m-n)\omega\| - \lambda' \geqslant \frac{\varkappa}{(8M'N')^\tau} - \lambda' \geqslant \frac{1}{2}\cdot\frac{\varkappa}{(8M')^\tau}\varepsilon^{a\tau} \gg \varepsilon^b$$

under (17).

Hence, if we replace k_1 by k_2+n—still denoted k_1—and increase I' to $9I'$—still denoted I'—we can assume that

$$\{m\omega + I' : |m - k_j| \leqslant 2M'N', j = 1, 2\} \quad \text{are pairwise disjoint,}$$

and

$$k_1\omega + \tfrac{1}{4}I' \cup k_2\omega + \tfrac{1}{4}I' \supset X_1 \cup X_2.$$

It is now sufficient that the intervals $\{m\omega+I' : |m-k_j|\leqslant 4M'N', j\geqslant 2\}$ are pairwise disjoint. If they are not, then there are two k_j's, k_2 and k_3 say, such that $(k_2+n_2)\omega+I'$ and $(k_3+n_3)\omega+I'$ intersect for some $|n_2|,|n_3|\leqslant 4M'N'$, i.e. $\|k_1\omega-(k_2+n)\omega\|<|I'|$ for $|n|=|n_3-n_2|\leqslant 8M'N'$. Now we proceed inductively. Hence, by changing those k_j's that do not belong to Λ_i', and increasing the size of I', we can achieve that

$$\{m\omega+I'\}_{|m-k_j|\leqslant 2M'N', j\geqslant 1} \quad \text{are pairwise disjoint,}$$

and

$$\bigcup_{j\geqslant 1} k_j\omega+\tfrac{1}{4}I' \supset \bigcup_{j\geqslant 1} X_j.$$

Since the number of k_j's is less than $M^2M_1^2\leqslant M'-1$, the k_j's increase by at most

$$4M'N'+2\cdot 4M'N'+2^2\cdot 4M'N'+\ldots+2^{M'-1}\cdot 4M'N'\leqslant 2^{M'+2}\cdot M'N',$$

and therefore

$$\operatorname{dist}(k_j,\Lambda_i')\leqslant 2\frac{2\pi}{\varkappa}\left(\frac{2\pi}{\lambda'}\right)^{\tau+1}$$

under condition (17)—hence, (C_5) still holds.

Since we can start the induction by an interval I' of size $8\lambda'$ and since the interval increases at most $9^{M'-1}$ times, it will remain smaller than ε^b.

Finally, if $\operatorname{dist}(n,\Omega_i')<N'$ it follows that

$$\operatorname{dist}(n,\{k_j : j\geqslant 1\})<N'+M'N'+\operatorname{diam}(\Omega(k))\leqslant 2M'N'$$

if

$$\varepsilon^a<\gamma_{12}\lambda. \tag{19}$$

This proves (C_4).

Verification of (C_2). Let us denote $\Omega(k)$ and $\Delta(k)$ by Ω_i and Δ_i respectively, so that $\Omega_i'=\Omega_i+\Delta_i$. The construction of Ω_i together with condition (C_4) imply that Ω_i+k and Ω_i+l are separated by at least N sites, unless $k=l$. Hence, by condition (A),

$$D(\theta)_{\Omega_i'}=\prod_{k\in\Delta_i} D(\theta)_{\Omega_i+k}$$

for all θ. Hence,

$$\begin{aligned}
u_{\Omega_i'}(x,\theta) &= \operatorname{Res}(\det(D(\theta+x)_{\Omega_i'}-tI),\det(D(\theta)_{\Omega_i'}-tI))\\
&= \prod_{k,l\in\Delta_i}\operatorname{Res}(\det(D(\theta+x)_{\Omega_i+k}-tI),\det(D(\theta)_{\Omega_i+l}-tI))\\
&= \prod_{k,l\in\Delta_i} u_{\Omega_i}(x+(k-l)\omega,\theta+l\omega).
\end{aligned}$$

By Remark 4 we have, for θ fixed, that

$$|u_{\Omega_i}|_{C^\nu} < (4CM)^{2M^2}(L')^\nu \quad \forall \nu \geqslant 0,$$

and applying Lemma 4 gives

$$\max_{0\leqslant\nu\leqslant s(M')^2}\left|\frac{1}{(\nu!)^2(L')^\nu}\,\partial_x^\nu u_{\Omega_i'}(x,\theta)\right| \geqslant 2\beta',$$

since

$$\varepsilon^b < \gamma_{13}\frac{1}{K}\left(\beta\left(\frac{L}{K}\right)^{sM^2}\right)^{(M'/M)^{2(s(M')^2+1)}}. \tag{20}$$

What we want is an estimate for $u'_{\Omega_i'}$, which is defined as $u_{\Omega_i'}$ but with the matrix D' instead of D. For this we only observe that, for θ fixed,

$$|u'_{\Omega_i'} - u_{\Omega_i'}|_{C^\nu} < (4CM')^{2(M')^2}\sqrt{\varepsilon}\,(K')^\nu,$$

which is $<\varepsilon^b K^\nu$ for all $\nu \leqslant s(M')^2+1$, if

$$\tfrac{1}{4} > b + (s(M')^2+1)c, \tag{21.1}$$

$$\varepsilon^{1/4} < \gamma_{14}. \tag{21.2}$$

This implies the second part of condition (C_2).

Using the estimate of $|u_{\Omega_i}|_{C^\nu}$ obtained from Remark 4, Lemma 4 gives an estimate of $|u_{\Omega_i'}|_{C^\nu}$. The estimate of $|u'_{\Omega_i'} - u_{\Omega_i'}|_{C^\nu}$ above now gives the first part of (C_2) if for example (21.1) holds.

Verification of (C_6). The first part follows exactly as in (C_2). As for the second part we let

$$v_{\Omega_i}(y,\theta) = u_{\Omega_i}(y,\theta)\left(\prod_{m,n\in\Omega_i}\|y+(m-n)\omega\|\right)^{-1}.$$

If we notice that $u_{\Omega_i}(y,\theta)\in \mathbf{O}(|y|^{\#\Omega_i})$ then we have, for $|y|<\frac{4}{5}|I|$ fixed, that

$$\begin{aligned}|v_{\Omega_i}|_{C^\nu} &< (8M)^{2M^2}\left(\frac{5}{\lambda}\right)^{M^2}\left(\frac{(\nu+M)!}{\nu!}\right)^2(L')^M(L')^\nu \\ &< (8M)^{2M^2}\left(\frac{5}{\lambda}\right)^{M^2}(sM')^{4M}(L')^M(L')^\nu \quad \forall \nu \leqslant s(M')^2+1,\end{aligned}$$

where we have used (C_4) to estimate $\|y-(m-n)\omega\|$ from below by $\frac{1}{5}|I| \geqslant \frac{1}{5}\lambda$ when $m\neq n$.

Let now

$$v_{\Omega_i'}(x,\theta)=\prod_{k,l\in\Delta_i} v_{\Omega_i}(x+(k-l)\omega,\theta+l\omega)$$

and observe that if $x\in I'$, then

$$|x+(k-l)\omega|<|I'|+\tfrac{3}{4}|I|<\tfrac{4}{5}|I|$$

if (18) holds—recall that $\Delta_i\omega\subset\frac{3}{4}I$. By Lemma 4 we get that

$$\max_{0\leqslant\nu\leqslant s(M')^2}\left|\frac{1}{(\nu!)^2(L')^\nu}\,\partial_\theta^\nu v_{\Omega_i'}(x,\theta)\right|\geqslant 2\beta',$$

since

$$\varepsilon^b<\gamma_{15}\frac{1}{K}\left(\beta\frac{\lambda^{2M^2}}{K^{2M}}\left(\frac{L}{K}\right)^{sM^2}\right)^{(M'/M)^{2(s(M')^2+1)}} \tag{20'}$$

Hence

$$\max_{0\leqslant\nu\leqslant s(M')^2}\left|\frac{1}{(\nu!)^2(L')^\nu}\,\partial_\theta^\nu u_{\Omega_i'}(x,\theta)\right|\geqslant 2\beta'\left(\prod_{k,l\in\Omega_i'}\|x+(k-l)\omega\|\right).$$

If we notice that $u'_{\Omega_i'}(x,\theta)\in\mathbf{O}(|x|^{\#\Omega_i'})$, then we have, for x fixed, that

$$|u'_{\Omega_i'}-u_{\Omega_i'}|_{C^\nu}<(4CM')^{2(M')^2}\left(\frac{(\nu+M')!}{\nu!}\right)^2(K')^{M'}\|x\|^{\#\Omega_i'}\sqrt{\varepsilon}\,(K')^\nu$$

for all $\nu\geqslant 0$. Now this is bounded by

$$(4CM')^{2(M')^2}(sM')^{4M'}(K')^{M'}\|x\|^{\#\Omega_i'}\sqrt{\varepsilon}\,(K')^\nu\quad\forall\nu\leqslant s(M')^2+1,$$

and since

$$\min_{k\neq l\in\Omega_i'}\|(k-l)\omega\|\geqslant|I'|\geqslant\left(\tfrac{1}{9}\right)^{M'}\varepsilon^b$$

this is less than

$$\left(\prod_{k,l\in\Omega_i'}\|x+(k-l)\omega\|\right)\varepsilon^b K^\nu\quad\forall\nu\leqslant s(M')^2+1,$$

if

$$\tfrac{1}{4}>b((M')^2+1)+(s(M')^2+1+M')c, \tag{21.1'}$$

$$\varepsilon^{1/4}<\gamma_{16}(1/K)^{M'}. \tag{21.2'}$$

This gives the second part of (C_6).

The additional statement. This is just a formulation of the construction made in the proof.

The final statement. This is the only part where we use condition (C_6). Consider

$$\begin{cases} \{n\omega : |E(n\omega)-E(n\omega+l\omega)| < 2\eta = 2M'(K/r)\varepsilon^b\}, \\ 4(1/\lambda)^{\tau+2} < |l| \leqslant M'N'. \end{cases}$$

Divide $\mathbf{T}$ into intervals of length $r\alpha/16K$ and choose a $k\omega$ in each interval. For any n the minimum of $|E(n\omega)-E(k\omega)|$ over all such k is $<\frac{1}{2}\alpha$. If now

$$|E(n\omega)-E(n\omega+l\omega)| < 2\eta,$$

then by (C_1), $n\omega \in m\omega + \frac{1}{2}I$ and $(n+l)\omega \in m'\omega + \frac{1}{2}I$ for some $m, m' \in \Omega(k)$. Hence, $\theta = (n-m)\omega \in I$ belongs to the set

$$Z := \{\theta \in I : |u_{\Omega(k)}(x,\theta)| < 2\eta(16/r)^{M^2}\},$$

where we have put $x = (l-m'+m)\omega$.

For any $m_1, m_2 \in \Omega(k)$ we get by (C_5) that

$$0 < |l-m'+m+(m_1-m_2)| < M'N' + 4(1/\lambda)^{\tau+2},$$

which is less than $2M'N'$ under (19). Hence, we get that

$$\|x+(m_1-m_2)\omega\| \geqslant \frac{\varkappa}{(2M'N')^\tau}.$$

Using now (C_6) together with Lemma 3 we find that Z has at most

$$M'\left(\frac{L}{\beta}|I|\left(\frac{(2M'N')^\tau}{\varkappa}\right)^{M^2} + 1\right)$$

many components, each of length less than

$$\frac{2}{L}\left(\frac{16}{r}\right)^{1/s}\left(\frac{2\eta}{\beta}\right)^{1/sM^2}\left(\frac{(2M'N')^\tau}{\varkappa}\right)^{1/s},$$

which is $<\varepsilon^{b/4sM^2}$ under condition (13).

Since $\#\Omega(k) \leqslant M$, since the number of l's is less than $2M'N'$ and the number of k's is less than $\text{const}\cdot K/r\alpha$, we get the result if

$$b > 20s\tau M^4 a, \tag{22.1}$$

$$\varepsilon^b < \gamma_{17}(r\alpha/K)^{10sM^2}. \tag{22.2}$$

The estimate of the second set is completely analogous—it is here we use that $M \geqslant 2\tau$. □

6. Proof of theorem

PROPOSITION 8. *Let D be in normal form on an interval I with parameters $C, K, r, M, N, \alpha, L, \beta, \lambda$, satisfying the assumptions*

$$1 < L \leqslant K, \quad \max(8, 2\tau) \leqslant M, \quad \tfrac{4}{3} < C < \tfrac{5}{3}, \quad r, \alpha, \beta \leqslant 1,$$

and $rN = 1$.

Let F be a smooth symmetric matrix satisfying the shift condition. Assume that

$$\lambda \leqslant |I| \leqslant \beta / L,$$
$$|F_{m,n}|_{C^\nu} < \varepsilon e^{-|m-n|r} K^\nu \quad \forall \nu \geqslant 0.$$

Then there exists a constant $\Gamma' = \Gamma'(\varkappa, \tau, s, \gamma_0, M)$, superexponentially decaying in M, such that if

$$|\varepsilon| < \Gamma' \left(\frac{\alpha \beta r^\tau \lambda^{\tau^2}}{K N^\tau} e^{-Nr} \right)^{e^{e^{sM^4}}},$$

then there is a smooth orthogonal matrix U, satisfying the shift condition and such that

$$U^*(D+F)U = D_\infty,$$

where $D_\infty(\theta)$ is a norm limit of matrices with pure point spectrum. Moreover, $D_\infty(\theta)$ is pure point with finite-dimensional eigenvectors for a.e. θ, and the measure of $\sigma(D) \backslash \sigma(D+F)$ goes to 0 as $\varepsilon \to 0$.

Proof. Let us choose

$$\begin{aligned}
&C_{j+1} = (1 + \varepsilon_j^{1/2}) C_j, && K_{j+1} = (1/\varepsilon_j)^{c_j} K_j, && r_{j+1} = \varepsilon_j^{a_j}, \\
&\lambda_{j+1} = \left(\tfrac{1}{9}\right)^{M_{j+1}} \varepsilon_j^{b_j}, && M_{j+1} = M_j^{sM_j^3}, && N_{j+1} = (1/\varepsilon_j)^{a_j}, \\
&\alpha_{j+1} = \varepsilon_j^{c_j}, && L_{j+1} = K_j, && \beta_{j+1} = \varepsilon_j^{b_j} K_j, \\
&\varepsilon_{j+1} = \varepsilon_j^{\frac{1}{2}(\frac{1}{\varepsilon_j})^{a_j/2}},
\end{aligned}$$

where we let $C_1 = C$, $K_1 = K$, ... be the parameters of the normal form $D_1 = D$, where $\varepsilon_1 = \varepsilon$ and

$$a_j = \frac{1}{\tau} \left(\frac{1}{M_j} \right)^{3sM_j^3}, \quad b_j = 20 s \tau M_j^4 a_j, \quad c_j = \frac{1}{5s} \left(\frac{1}{M_j} \right)^{2sM_j^3}.$$

We choose

$$\frac{1}{\Gamma'(M)} = \exp \circ \ldots \circ \exp(AM^4), \quad A = A(\varkappa, \tau, s, \gamma_0),$$

in such a way that for all $j \geqslant 0$,

$$\Gamma'(M_{j+1}) \geqslant (1/e)^{(1/\Gamma'(M_j))^{a_j/2}},$$
$$\Gamma'(M_j)^{a_j/2} \leqslant (1/e)^{e^{sM_{j+1}^4}},$$
$$\Gamma'(M_j) \leqslant \Gamma(M_j),$$

where Γ is the constant in Lemma 7.

We shall show that for all $j \geqslant 1$,

$$\varepsilon_j < \Gamma'(M_j)\left(\frac{\alpha_j \beta_j r_j^\tau \lambda_j^{\tau^2}}{K_j N_j^\tau} e^{-N_j r_j}\right)^{e^{e^{sM_j^4}}}.$$

This implies that

$$\varepsilon_j < \Gamma'(M_j)\left(\left(\tfrac{1}{9}\right)^{M_j \tau^2} e^{-1}\right)^{e^{e^{sM_j^4}}},$$

so ε_j decays very rapidly.

Let us assume that the inequality holds for j and then prove it for $j+1$. If it holds for j then

$$\begin{aligned}
\Gamma'(M_{j+1})&\left(\frac{\alpha_{j+1}\beta_{j+1} r_{j+1}^\tau \lambda_{j+1}^{\tau^2}}{K_{j+1} N_{j+1}^\tau} e^{-1}\right)^{e^{e^{sM_{j+1}^4}}} \\
&\geqslant \left(\frac{1}{e}\right)^{(1/\Gamma'(M_j))^{a_j/2}} \left(\varepsilon_j^{c_j+b_j+a_j\tau+\tau^2 b_j+c_j+\tau a_j}\left(\tfrac{1}{9}\right)^{M_{j+1}\tau^2} e^{-1}\right)^{e^{e^{sM_{j+1}^4}}} \\
&\geqslant \varepsilon_j^{\frac{1}{4}\left(\frac{1}{\varepsilon_j}\right)^{a_j/2}} \left(\varepsilon_j^{1/8}\left(\tfrac{1}{9}\right)^{M_{j+1}\tau^2}\right)^{e^{e^{sM_{j+1}^4}}} \geqslant \varepsilon_j^{\frac{1}{4}\left(\frac{1}{\varepsilon_j}\right)^{a_j/2}} \varepsilon_j^{\frac{1}{4}e^{e^{sM_{j+1}^4}}},
\end{aligned}$$

which is greater than ε_{j+1}. Hence, the basic smallness assumption of Lemma 7 is fulfilled—all other assumptions of that lemma are immediate.

By Lemma 7 there is for each $j \geqslant 1$ an orthogonal matrix U_j satisfying the shift condition and such that

$$|(U_j - I)_{m,n}|_{C^\nu} < \sqrt{\varepsilon_j}\, e^{-|m-n| r_{j+1}} K_{j+1}^\nu$$

and

$$(U_1 \cdot \ldots \cdot U_j)^*(D+F)(U_1 \cdot \ldots \cdot U_j) = D_{j+1} + F_{j+1},$$

where D_{j+1} is in normal form with parameters $C_{j+1}, K_{j+1}, r_{j+1}, M_{j+1}, N_{j+1}, \alpha_{j+1}, L_{j+1}, \beta_{j+1}, \lambda_{j+1}$, and where

$$|(F_{j+1})_{m,n}|_{C^\nu} < \varepsilon_{j+1} e^{-|m-n| r_{j+1}} K_{j+1}^\nu.$$

Hence

$$U_1 \cdot \ldots \cdot U_j \to U, \quad F_j \to 0 \quad \text{and} \quad D_j \to D_\infty$$

in norm, and it remains to check that D_∞ is pure point. For this we need the additional and final statements of Lemma 7.

Clearly there is a uniform limit

$$E^j(\theta) \to E^\infty(\theta)$$

which describes the spectrum of $D_\infty(\theta)$—it is the closure of the image of E^∞. Consider now the closure Z^j of the set of all θ such that

$$|E^\infty(\theta) - E^\infty(\theta + l\omega)| < \tfrac{3}{2}\eta_j \quad \text{for some } 4(1/\lambda_j)^{\tau+2} < |l| \leqslant M_{j+1}N_{j+1}$$

or

$$|E^\infty(\theta) - E^\infty(\theta + l\omega))| < \tfrac{3}{2}\varepsilon_j^{1/8} \quad \text{for some } M_{j+1}N_{j+1} < |l| \leqslant 4(1/\lambda_{j+1})^{\tau+2},$$

where $\eta_j = M_{j+1}(K_j/r_j)\varepsilon_j^{b_j}$. According to the final statement of Lemma 7 this set is of measure less than

$$\text{const} \cdot \varepsilon_j^{b_j/20sM_j^2}.$$

From this we conclude that for a.e. θ each $\theta + k\omega$ will belong to only finitely many Z^j's. Suppose $\theta = 0$ is of this sort, i.e. for all k there is a $j_0(k)$ such that $k\omega \notin Z^j$ for $j \geqslant j_0(k)$. Hence, for such j's,

$$|E^j(k\omega) - E^j(k\omega + l\omega)| \geqslant 2\eta_j \quad \text{for all } 4(1/\lambda_j)^{\tau+2} < |l| \leqslant M_{j+}N_{j+1}$$

and

$$|E^j(k\omega) - E^j(k\omega + l\omega)| \geqslant 2\varepsilon_j^{1/8} \quad \text{for all } M_{j+1}N_{j+1} < |l| \leqslant 4(1/\lambda_{j+1})^{\tau+2}.$$

This implies that $\Lambda^j(k) \subset [k - 4(1/\lambda_{j_0(k)})^{\tau+2}, k + 4(1/\lambda_{j_0(k)})^{\tau+2}]$ for all $j \geqslant j_0(k)$. The blocks $\Lambda^j(k)$ therefore become eventually stationary:

$$\Lambda^{j+1}(k) = \Lambda^j(k) \quad \forall j \geqslant j_1(k),$$

and $D(0)$ is pure point. (Here we used the first two additional statements of Lemma 7.)

The same argument works for an arbitrary θ and shows that $D(\theta)$ is pure point for a.e. θ. One only needs to convince oneself that when we shift base point on $\mathbf{T}$ from 0 to

θ_0 then $Z^j_{\theta_0}=Z^j-\theta_0$. In order to see this, observe first that if D is in normal form on an interval I with orthogonal matrix Q, block decomposition $\mathbf{Z}=\bigcup\Lambda_i$, spectral function E and sets $\{\Omega_i\}$, then $\widetilde{D}(\theta)=D(\theta+k\omega)$ will be in normal form with corresponding objects

$$I,\quad T_k^{-1}QT_k,\quad \bigcup(\Lambda_i-k),\quad \widetilde{E}(\theta)=E(\theta+k\omega),\quad \{\Omega_i-k\},$$

where T_k is the shift by k sites. (But $\widetilde{D}(\theta)=D(\theta+\theta_0)$ may not be in normal form for arbitrary θ_0.)

The D' obtained by Lemma 7 is special however. If θ_0 is given, choose k so that $\|k\omega-\theta_0\|<\frac{1}{16}|I'|$. Then $D'_{k\omega}(\theta)=D'(\theta+k\omega)$ is in normal form on I' *and* on $\frac{1}{2}I'$—this is the third additional statement—with

$$T_k^{-1}QT_k,\quad \bigcup(\Lambda_i-k),\quad E'_{k\omega}(\theta)=E'(\theta+k\omega),\quad \{\Omega_i-k\}.$$

Then it follows that $D'_{\theta_0}(\theta)=D'(\theta+\theta_0)$ is in normal form on $\frac{3}{4}I'$ since

$$(k\omega-\theta_0)+\tfrac{1}{2}\left(\tfrac{3}{4}I'\right)\supset\tfrac{1}{4}I' \quad\text{and}\quad (k\omega-\theta_0)+\left(\tfrac{3}{4}I'\right)\subset I'.$$

Moreover, $E'(\theta+\theta_0)$ is a spectral function for D'_{θ_0}.

Hence, we can choose a sequence $k_j\omega$ such that $\|k_j\omega-\theta_0\|<\frac{1}{16}|I_j|$. Then $D^j_{\theta_0}(\theta)=D^j(\theta+\theta_0)$ will be in normal form on $\frac{3}{4}I_j$ with $E^j(\theta+\theta_0)$ as spectral function, and

$$D^j_{\theta_0}(\theta)\to D^\infty_{\theta_0}(\theta)=D^\infty(\theta+\theta_0) \quad\text{and}\quad E^j(\theta+\theta_0)\to E^\infty(\theta+\theta_0).$$

Hence $E^\infty_{\theta_0}(\theta)=E^\infty(\theta+\theta_0)$ and $Z^j_{\theta_0}=Z^j-\theta_0$.

Finally, the resolvent set of D_{j+1} consists first of all of a modification Σ'_{j+1} of all the gaps of D_j under the perturbation F_j, and secondly of newly created gaps which are contained in the image under E^j of the set of θ such that

$$|E^j(\theta)-E^j(\theta+l\omega)|<2\eta_j$$

for some $|l|\leqslant M_{j+1}N_{j+1}$. These contributions are easy to estimate and the estimate converges as $j\to\infty$ to something that goes to 0 with ε. $\square$

Proof of theorem. In order to prove the theorem we first observe that we can assume that $|E|<1$, so we can take $\frac{4}{3}<C<\frac{5}{3}$.

Now we let $D(\theta)=\operatorname{diag}(E(\theta+n\omega))$. It satisfies condition (A) for K of the theorem, for $N=1$, and for $r=1$.

If we let I be any symmetric interval around 0, $\Lambda(k)=\{k\}$ for all k, and if we let Q be the identity matrix, then also condition (B) is fulfilled for $N=1$ and any $M\geqslant 1$.

Let now

$$M=\max\left(2^{s+4}C\frac{K^{s+1}((s+1)!)^2}{\xi},2\tau,8\right).$$

By Lemma 3,

$$|E(\theta)-E(k\omega))|\geqslant\alpha$$

outside a set of at most M many intervals, each of length at most

$$l(\alpha):=\frac{2}{K}\left(\frac{2\alpha(s!)^2K^s}{\xi}\right)^{1/s}.$$

By taking α small enough we can assume that each interval is separated from any other by a distance $\geqslant 4l(\alpha_0)$. Choose then α even smaller so that

$$\frac{2\pi}{\varkappa}\left(\frac{2\pi}{l(\alpha)}\right)^{\tau+1}\leqslant\left(\frac{1}{4l(\alpha)}\right)^{\tau+2}$$

and

$$4l(\alpha)\leqslant\frac{\beta}{L},$$

where β, L will be defined below (independent of α).

We choose a $m\omega$ in each such interval with the integer m as close to k as possible,

$$|m-k|\leqslant\frac{2\pi}{\varkappa}\left(\frac{2\pi}{l(\alpha)}\right)^{\tau+1},$$

and we let $\Omega(k)$ be the set of such m's. Then (C_1) and (C_3)–(C_5) are fulfilled for M as above and for $\lambda=|I|\leqslant 4l(\alpha)$.

The first part of (C_2) and (C_6) holds with $L=K$—by Remark 4. By Lemma 4, the second part of (C_2) and (C_6) will be fulfilled for

$$\beta=\left(\frac{1}{e}\right)^{M^{4sM^2}}\left(\frac{l(\alpha_0)\xi}{(CK(sM^2+1)^2)^2K^s(s!)^2}\right)^{M^{2(sM^2+1)}},$$

as in the verification of (C_2) and (C_6) in Lemma 7.

With this choice of parameters, $D(\theta)$ will be of normal form and satisfy the assumptions of Proposition 8. The perturbation

$$F(\theta)=\begin{pmatrix}\ddots & \vdots & \vdots & \vdots & \vdots & \\ \dots & -\varepsilon & 0 & -\varepsilon & 0 & \dots\\ \dots & 0 & -\varepsilon & 0 & -\varepsilon & \dots\\ \dots & \vdots & \vdots & \vdots & \vdots & \ddots\end{pmatrix}$$

satisfies

$$|F_{m,n}|_{C^\nu}<(e\varepsilon)e^{-|m-n|^rK^\nu}.$$

Now the theorem follows from Proposition 8. □

7. Some further properties

We have not studied the spectrum as a set, nor the occurrence of gaps—these play no role in the proof. One would, however, expect that gaps occur (at least for analytic potentials) and that the spectrum is a Cantor set.

We want to discuss the continuity of the eigenvalues as functions of θ. Since $D_j(\theta)$ is smoothly conjugated to a block matrix with finite-dimensional blocks, there are continuous choices $E_k^j(\theta)$ of eigenvalues of $D_j(\theta)$ on $\mathbf{T}\setminus\{\text{point}\}$—we don't claim that they are continuous on the whole of $\mathbf{T}$. We can choose the continuous branches in such a way that $E_k^j(0)=E^j(k\omega)$ and

$$|E_k^j(\theta)-E_k^{j+1}(\theta)|\leqslant|D_j(\theta)-D_{j+1}(\theta)|<\text{const}\cdot\frac{\sqrt{\varepsilon_j}}{r_j^2}.$$

It follows that there are continuous choices of the eigenvalues of $D_\infty(\theta)$, and hence of $D(\theta)+F(\theta)$. This analysis, however, does not permit us to conclude anything about smoothness.

Another question we like to address is the multiplicity of the spectrum. The "almost multiplicities" we have been dealing with are unbounded, but in the end the true multiplicity is always one. Indeed, if $D(0)+F(0)$ had a double eigenvalue $E_k^\infty(0)=E_l^\infty(0)$ there would also be two orthogonal eigenvectors, $q_k^\infty(0)$ and $q_l^\infty(0)$. This means that

$$-\varepsilon(u_{n+1}+u_{n-1})+V_nu_n=zu_n \qquad (*_z)$$

has two independent solutions in $l^2(\mathbf{Z})$ when $z=E_k^\infty(0)=E_l^\infty(0)$. But if $(*_z)$ has two independent solutions in $l^2(\mathbf{Z})$ for one z then this will be the case for all z—see for example [3, p. 245], where the proof in the continuous case translates directly to the discrete case. But for sufficiently large $|z|$ the equation is reducible to a constant-coefficient equation so there are no solutions in $l^2(\mathbf{Z})$ at all. Hence, the spectrum of $D(0)+F(0)$ must be simple.

Since our approach permits us to analyze the system for all θ to any degree of approximation, one can use it to try to understand the equation also for θ in the exceptional measure-zero set. We expect that, for generic θ and for generic potentials (in the Gevrey class we have studied, or perhaps in some weaker Gevrey class), the blocks do not become stationary but continue to increase and give rise to eigenvectors in $l^\infty(\mathbf{Z})$ and a singular continuous spectrum for $D(\theta)+F(\theta)$. (A result of this type has been announced in [9].)

As for the exponential decay of the eigenvectors, we have not provided any proof of this. The reason why we don't immediately get such decay is that the parameter r_j, with which we measure the exponential decay, decreases in order to control the factor $e^{N_jr_j}$

appearing in Lemma 5. This cannot be avoided since r is a global parameter, measuring the minimal exponential decay of all the eigenvalues—therefore it *must* decrease to zero. However, if we measured the exponential decay of eigenvectors associated to a given block $\Lambda^j(k)$ we would probably see that the decay stops once the blocks become stationary—only a very small decay, which does not go to zero, will be required by the construction in Corollary 6.

Considered as a dynamical system on $\mathbf{T}\times \mathrm{SL}(2,\mathbf{R})$ the theorem implies, by Kotani's theorem [1, Proposition VII.3.3], that the skew product

$$\begin{pmatrix}\theta\\ u\\ u'\end{pmatrix}\to\begin{pmatrix}\theta+\omega\\ u'\\ -u+(E(\theta+\omega)-\lambda)u'/\varepsilon\end{pmatrix}$$

has non-zero Lyapunov exponents for almost every λ.

Appendix

LEMMA A1. *There is a C^∞-function $\phi\geqslant 0$ on $\mathbf{R}$, supported in $\left[-\frac{1}{2},\frac{1}{2}\right]$, such that*

$$|\phi|_{C^\nu}<\gamma_0\gamma_0^\nu \quad \forall\nu\geqslant 0,$$

for some $\gamma_0>0$.

Given two intervals $I_2\subset I_1$ with $\operatorname{dist}(I_2,\partial I_1)\geqslant\delta$. Then there is a smooth function $\phi\geqslant 0$ of compact support in I_1 which is $\equiv 1$ on I_2 and such that

$$|\phi|_{C^\nu}<\left(\frac{\gamma_0}{\delta}\right)^\nu \quad \forall\nu\geqslant 0.$$

Proof. The class of all C^∞-functions satisfying such a bound is not quasi-analytic [13, 19.11]. Hence it contains a non-zero function ψ which satisfies this bound for some $\gamma_0>0$ and which is compactly supported [13, 19.10]—by a scaling we can obviously assume that it is supported in $\left[-\frac{1}{2},\frac{1}{2}\right]$. Then if we take $\phi=\psi^2$ and replace γ_0 by a suitable larger constant we have our function.

By multiplying the function by an appropriate constant we can assume that its integral equals one. We localize then our function on the left component of $I_1\setminus I_2$. Since such a component is of length at least δ, it will satisfy

$$|\phi|_{C^\nu}<\frac{\gamma_0}{\delta}\left(\frac{\gamma_0}{\delta}\right)^\nu \quad \forall\nu\geqslant 0.$$

We now define ϕ on the right component in the same way, but negative, and then we take the primitive of this function. □

Let now u_j be a sequence of smooth functions defined on an open interval I.

LEMMA A2. *Assume that* $|u_j|_{C^\nu} < C_j K^\nu$ $\forall \nu \geqslant 0$. *Then for all* $\nu \geqslant 0$,

$$\begin{aligned} |u_1 \cdot \ldots \cdot u_J|_{C^\nu} &< 4^{J-1}(C_1 \cdot \ldots \cdot C_j) K^\nu, \\ |e^{u_1}|_{C^\nu} &< e^{4C_1} K^\nu, \\ |\sqrt{1+u_1}|_{C^\nu} &< \sqrt{1+C_1}\, K^\nu \quad \text{if } C_1 < \tfrac{1}{4}, \\ \left|\frac{1}{u_1}\right|_{C^\nu} &< \frac{1}{\delta}\left(\frac{4}{\delta} C_1 K\right)^\nu \quad \text{if } |u_1| > \delta. \end{aligned}$$

Proof.

$$|\partial^\nu(u_1 u_2)| = \left|\sum_{i=0}^{\nu} \binom{\nu}{i} \partial^{\nu-i} u_1 \partial^i u_2\right| < (C_1 C_2) K^\nu (\nu!)^2 \sum_{i=0}^{\nu} \binom{\nu}{i} \frac{(i!)^2((\nu-i)!)^2}{(\nu!)^2},$$

and the result for $J=2$ follows since $\sum_{i=0}^{\nu} \binom{\nu}{i}^{-1} < 4$ for all ν—just notice that $\binom{\nu}{i} \geqslant 2^i$ if $i \leqslant \frac{1}{2}\nu$. For $J \geqslant 3$ the result follows by induction.

We use the power series expansions of e^u and $\sqrt{1+u}$ which converge absolutely in $|u| < \infty$ and in $|u| < 1$ respectively, to prove the second and third estimates.

Notice that $|u_1| > \delta$ implies that $C_1 > \delta$. By a scaling we can suppose that $\delta = 1$ and $C_1 \geqslant 1$. The fourth statement is obvious for $\nu = 0$ so we proceed by induction. Hence

$$\begin{aligned} \left|\partial^\nu\left(\frac{1}{u_1}\right)\right| &= \left|-\frac{1}{u_1} \sum_{i=1}^{\nu} \binom{\nu}{i} \partial^{\nu-i}\left(\frac{1}{u_1}\right) \partial^i u_1\right| \\ &< 4^{\nu-1} (\nu!)^2 (C_1 K)^\nu \sum_{i=1}^{\nu} \binom{\nu}{i} \frac{(i!)^2((\nu-i)!)^2}{(\nu!)^2} \\ &< (\nu!)^2 (4 C_1 K)^\nu. \end{aligned}$$

□

Remark. The preceding lemma is valid also for matrices if we understand by $1/u$ the inverse of u, and if we use the operator norm satisfying

$$|uv| \leqslant |u| \cdot |v|.$$

LEMMA A3. *Let* $v^1, \ldots, v^d$ *be a basis of* $\mathbf{R}^d$ *for* $\theta \in I \subset \mathbf{R}$, $d \geqslant 2$, *such that*

$$\begin{aligned} |v^m|_{C^\nu} &< C K^\nu \quad \forall \nu \geqslant 0, \\ |\langle v^m, v^n \rangle|_{C^0} &< \frac{1}{4d^2 3^{d+4}}, \quad m \neq n, \\ 1 - \frac{1}{4d^2 3^{d+4}} < |v^m|_{C^0} &< 1 + \frac{1}{4d^2 3^{d+4}}, \end{aligned}$$

for all m,n and for $C\geqslant 1$. Then the orthonormal basis $\hat{v}^1,\dots,\hat{v}^d$ obtained by Gram–Schmidt orthogonalization verifies

$$|\hat{v}^m|_{C^\nu}\leqslant (d^{8^d}CK)^\nu \quad \forall\nu\geqslant 0.$$

Proof. We have

$$|v^m|_{C^\nu}<\left(1+\frac{1}{4d^2 3^{d+4}}\right)(CK)^\nu.$$

By Gram–Schmidt we first obtain an orthogonal basis $\tilde{v}^1,\dots,\tilde{v}^d$, and if we define

$$\begin{cases} w^1=\tilde{v}^1=v^1, \\ w^m=|w^1|^2\cdot\ldots\cdot|w^{m-1}|^2\tilde{v}^m, \end{cases}$$

then these vectors will satisfy

$$w^m=|w^1|^2\cdot\ldots\cdot|w^{m-1}|^2\left(v^m-\sum_{j=1}^{m-1}\langle v^m,w^j\rangle\frac{1}{|w^j|^2}w^j\right).$$

(Notice that there is no division in this formula.) From this we get by induction, using Lemma A2,

$$|w^m|_{C^\nu}<C_m(CK)^\nu,$$
$$C_m=md^m4^{2m-2}(C_1\cdot\ldots\cdot C_{m-1})^2C_1,\quad C_1=1+\frac{1}{4d^2 3^{d+4}}<2,$$

which gives

$$C_m\leqslant(4^{2m-2}md^mC_1)^{3^{m-2}}C_1^{2\cdot 3^{m-2}}\leqslant(4^{2m}md^m)^{3^{m-2}}.$$

It follows also that, for each m,

$$\sqrt{1-\tfrac{1}{4}}\leqslant\left(1-\frac{1}{3^{d+2}}\right)^{3^{m-1}}\leqslant|w^m|\leqslant\left(1+\frac{1}{3^{d+2}}\right)^{3^{m-1}}\leqslant\sqrt{1+\tfrac{1}{4}},$$

and from Lemma A2 that

$$\left|\frac{w^m}{\sqrt{\langle w^m,w^m\rangle}}\right|_{C^\nu}\leqslant(4^6dC_d^3CK)^\nu. \qquad \square$$

Polynomials. Let $P(\lambda,\theta)$ be a polynomial in λ of degree d with leading coefficient 1, smoothly parametrized by $\theta\in I$, and satisfying

$$|P(\lambda,\cdot)|_{C^\nu}<CK^\nu \quad \forall\nu\geqslant 0,$$

in a disk $|\lambda|<R$ of radius R. Assume that the roots $\{E_m(\theta)\}_1^d$ are real and lie in $|\lambda|<R-1$.

LEMMA A4. *If the roots belong to two groups,*

$$\begin{cases} E_1(\theta), \dots, E_k(\theta), \\ E_{k+1}(\theta), \dots, E_d(\theta), \end{cases}$$

such that any root of one group is separated from any root of the other group by at least $\delta<1$, *then the polynomial*

$$\prod_{m=1}^{k} (\lambda - E_m(\theta)) = \sum_{j=0}^{k} e_{k-j}(\theta)\lambda^j$$

satisfies

$$|e_j|_{C^\nu} < \left(4^{d+2}RC\frac{1}{\delta^{d-1}}\right)^j \left(2^{d+2}C\frac{1}{\delta^d}K\right)^\nu.$$

Proof. Choose a curve $\Delta(\theta)$ in $|\lambda|<R-\frac{1}{2}$, piecewise constant in θ, keeping a distance $\geqslant \frac{1}{2}\delta$ to all the roots $E_1(\theta), \dots, E_d(\theta)$ of $P(\lambda,\theta)$ and surrounding the first k of these roots. $\Delta(\theta)$ may consist of several components so we can choose it to be of length at most $d\pi\delta$. Then we have for all $\lambda\in\Delta(\theta)$,

$$|P(\lambda,\theta)| \geqslant \left(\tfrac{1}{2}\delta\right)^d = \varrho,$$

$$\left|\frac{1}{P(\lambda,\cdot)}\right|_{C^\nu} < \frac{1}{\varrho}\left(\frac{4}{\varrho}CK\right)^\nu,$$

$$\left|\frac{P_\lambda(\lambda,\cdot)}{P(\lambda,\cdot)}\right|_{C^\nu} < 4\cdot 2C\frac{1}{\varrho}\left(\frac{4}{\varrho}CK\right)^\nu,$$

where the second and third estimates follow from Lemma A2—we have used the Cauchy formula to estimate $|P_\lambda(\lambda,\cdot)|_{C^\nu}$ on $|\lambda|<R-\frac{1}{2}$.

Consider now the power symmetric functions in the first k roots:

$$p_j(\theta) = E_1(\theta)^j + \dots + E_k(\theta)^j.$$

We have the integral representation

$$p_j(\theta) = \frac{1}{2\pi}\oint_{\Delta(\theta)} \lambda^j \frac{P_\lambda(\lambda,\theta)}{P(\lambda,\theta)}\, d\lambda,$$

from which we get

$$|p_j|_{C^\nu} < 4d\delta\frac{C}{\varrho}\left(\frac{4C}{\varrho}K\right)^\nu R^j.$$

The e_j's are the elementary symmetric functions in the first k roots and we have the relation

$$e_j = \frac{(-1)^j}{j!} \det \begin{pmatrix} p_1 & 1 & 0 & 0 & \dots \\ p_2 & p_1 & 2 & 0 & \dots \\ p_3 & p_2 & p_1 & 3 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \\ p_j & p_{j-1} & p_{j-2} & p_{j-3} & \dots \end{pmatrix}$$

for $j \geqslant 1$ [11, p. 20]—of course $e_0 = 1$. So $j!\,e_j$ is a sum of at most $j!$ many products

$$p_{\iota_1} \cdot \ldots \cdot p_{\iota_l}, \quad \iota_1 + \ldots + \iota_l = j \leqslant k < d,$$

each with a coefficient that is at most $(j-1)!$. From Lemma A2 we now get the estimate we want. □

Symmetric matrices. Let $D(\theta)$ be a symmetric $(d\times d)$-matrix, smooth in the parameter $\theta \in I \subset \mathbf{R}$, and let $\{E_m(\theta)\}_1^d$ be its eigenvalues and $\{q^m(\theta)\}_1^d$ the corresponding eigenvectors. The eigenvalues and eigenvectors are smooth in θ when all eigenvalues are simple. The eigenvalues can in fact always be chosen to be continuous, and even C^1 [10, p. 122], in θ, but this is not the case for the eigenvectors. A noteworthy exception is when D is analytic in the parameter because then the eigenvalues and the eigenvectors are analytic [6]. For example, if $D = D_0 + \varepsilon D_1$ with D_0 diagonal, then the analyticity in ε gives a unique choice of the eigenvalues.

We always have estimates of the first derivative $\partial^1 E_m$ in terms of D.

LEMMA A5.

$$|E_m(\theta)| \leqslant |D(\theta)| \quad \textit{and} \quad |\partial E_m(\theta)| \leqslant |\partial D(\theta)|.$$

Proof. Let $q^m(\theta)$ be the eigenvector corresponding to $E_m(\theta)$. Then

$$(D(\theta) - E_1(\theta) I) q^1(\theta) = 0$$

and if we take the scalar product with $q^1(\theta)$ then we get an estimate of E_1.

If we differentiate the relation and take the scalar product with $q^1(\theta)$, and use that the eigenvectors are orthogonal, then we get an estimate of ∂E_1. □

For higher-order derivatives we have no such estimates. For example, if

$$D(\theta) = \begin{pmatrix} \theta & \varepsilon \\ \varepsilon & -\theta \end{pmatrix},$$

then $E_1(\theta) = \pm\sqrt{\theta^2 + \varepsilon^2}$ and, hence, $\partial^2 E_1 = \pm 1/\varepsilon$.

References

[1] CARMONA, R. & LACROIX, J., *Spectral Theory of Random Schrödinger Operators.* Birkhäuser, Boston, MA, 1990.

[2] CHULAEVSKY, V. A. & DINABURG, E. I., Methods of KAM-theory for long-range quasi-periodic operators on $\mathbf{Z}^\nu$. Pure point spectrum. *Comm. Math. Phys.*, 153 (1993), 559–577.

[3] CODDINGTON, E. A. & LEVINSON, N., *Theory of Ordinary Differential Equations.* McGraw-Hill, New York–Toronto–London, 1955.

[4] ELIASSON, L. H., Floquet solutions for the one-dimensional quasi-periodic Schrödinger equation. *Comm. Math. Phys.*, 146 (1992), 447–482.

[5] FOLLAND, G. D., *Introduction to Partial Differential Equations.* Princeton Univ. Press, Princeton, NJ, 1976.

[6] FRIEDRICHS, K. O., *Perturbation of Spectra in Hilbert Space.* Amer. Math. Soc., Providence, RI, 1965.

[7] FRÖHLICH, J., SPENCER, T. & WITTVER, P., Localization for a class of one-dimensional quasi-periodic Schrödinger operators. *Comm. Math. Phys.*, 132 (1990), 5–25.

[8] JITOMIRSKAYA, S. YA., Anderson localization for the almost Mathieu equation: a non-perturbative proof. *Comm. Math. Phys.*, 165 (1994), 49–57.

[9] JITOMIRSKAYA, S. & SIMON, B., Operators with singular continuous spectrum, III: almost periodic Schrödinger operators. *Comm. Math. Phys.*, 165 (1994), 201–205.

[10] KATO, T., *Perturbation Theory of Linear Operators.* Second edition. Springer-Verlag, Berlin–New York, 1976.

[11] MACDONALD, I. G., *Symmetric Functions and Hall Polynomials.* Clarendon Press, Oxford, 1979.

[12] PYARTLI, A. S., Diophantine approximation on submanifolds of Euclidean space. *Functional Anal. Appl.*, 3 (1969), 303–306.

[13] RUDIN, W., *Real and Complex Analysis.* McGraw-Hill, New York–Toronto–London, 1966.

[14] SINAI, YA. G., Anderson localization for the one-dimensional difference Schrödinger operator with a quasi-periodic potential. *J. Statist. Phys.*, 46 (1987), 861–909.

L. H. ELIASSON
Department of Mathematics
Royal Institute of Technology
S-10044 Stockholm
Sweden
hakane@math.kth.se

Received February 20, 1996
Received in revised form November 6, 1996