Remarks on the perturbation theory for problems of Mathieu type

1983 Russ. Math. Surv. 38 215

(http://iopscience.iop.org/0036-0279/38/4/R11)

Uspekhi Mat. Nauk 38:4 (1983), 189–203 Russian Math. Surveys 38:4 (1983), 215–233

Remarks on the perturbation theory for problems of Mathieu type

V.I. Arnol'd

Contents

Introduction

[illegible]

$$\ddot{x} + (\omega^2 + \varepsilon \cos t)x = 0$$

the width of each parametric resonance zone (or forbidden zone, in the terminology of solid state physics) decreases like a power with decreasing depth of modulation ε, and *the exponent is proportional to the number of the zone* (Fig. 1).

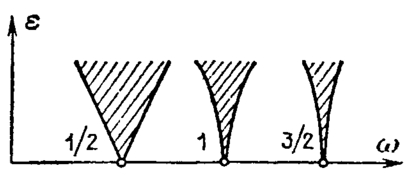


Fig. 1.

It follows, in particular, from general theorems in the present article that the zones of instability for Hill's equation

$$\ddot{x} + (\omega^2 + \varepsilon a(t))x = 0$$

have the same property when the variable coefficient a is *any trigonometric polynomial* (the exponent for the width of the zone decreases in inverse proportion to the degree of the polynomial).

For the general Hill's equation (with any periodic coefficient a) the situation is completely different; for a typical a the width of any zone decreases like the *first* power of ε.

In view of the general algebraic nature of the proof, it is applicable also to many other problems in which the perturbation is a trigonometric polynomial. Linear equations or systems with constant coefficients in the leading term and with coefficients in the form of trigonometric polynomials in the lower-order terms thus have special properties, and can be called *equations of Mathieu type.*

For example, consider the equation in the Zel'dovich problem of the existence of a steady-state kinematic magnetic dynamo:

$$H^{\cdot} = \{v, H\} + D\Delta H,$$

where H is an unknown 2π-periodic (in (x, y, z)) magnetic field of divergence zero that is carried by a 2π-periodic velocity field v of an incompressible fluid and diffuses with diffusion coefficient D ($\Delta = -\mathrm{rot\ rot}$ is the Laplacian and $\{\cdot,\cdot\} = \mathrm{rot}\ [\cdot, \cdot]$ denotes the Poisson bracket).

The equation is of Mathieu type if the components of the field v are trigonometric polynomials (for example, for the velocity field

$$\cos y + \sin z, \ \cos z + \sin x, \ \cos x + \sin y$$

considered in [1], which exponentially expands the particles of the fluid). It was in investigating this example that I observed the general properties described above of equations of Mathieu type.

In student work done in 1959 [2] I investigated the family of mappings of the circle onto itself of the form

$$x \mapsto x + a + \varepsilon \cos x$$

and established that the zone of resonance m/n in the (a, ε)-plane (Fig. 2) reaches the point $a = 2\pi m/n$ on the axis $\varepsilon = 0$ by a narrow tongue whose width decreases like ε^n as $\varepsilon \to 0$ (this is stated without proof in [2]).

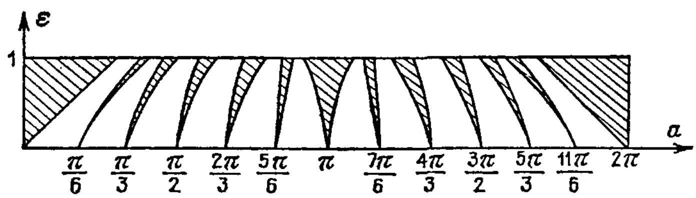


Fig. 2.

The theory of equations of Mathieu type suggests that here also the cosine can be replaced by a trigonometric polynomial of arbitrary degree p (with the corresponding replacement of ε^n by ε^r, where $r = -[-n/p]$).

Thus, the problem of zones of resonance for such special mappings of the circle also turns out to be "a problem of Mathieu type".

I did not try to formulate a single abstract theorem which would yield the above properties of resonance zones at once both for Hill's equations of Mathieu type and for mappings of the circle onto itself, although the algebraic reason for the power decrease in the width of a zone with the amplitude of modulation was obviously the same in the two cases (the conjecture that such a general theorem must exist was communicated to me by Gel'fand as far back as 1959 in a discussion of the results in [2]).

It gives me special pleasure to include a proof of the theorem from the 1959 work, carried out under the direct guidance of Kolmogorov, in the present article, which is dedicated to A.N. Kolmogorov on the occasion of his 80th birthday.

§1. Rayleigh-Schrödinger series for perturbations of a simple eigenvalue

We need some (simple) properties of the expansion of an eigenvalue of a matrix $\Lambda+\varepsilon A$, where Λ is a diagonal matrix, in a series of powers of the small parameter ε. Let λ_i denote the eigenvalues of the unperturbed operator Λ, and V_i the eigenspaces, so that

$$\Lambda = \oplus \Lambda_i, \quad \Lambda_i = \lambda_i E^i\colon\ V_i \to V_i, \quad \lambda_i \neq \lambda_j.$$

We decompose the matrix of the perturbing operator into the corresponding blocks

$$A_{ij}\colon\ V_i \to V_j.$$

Definition. The *quiver* of the family $\Lambda+\varepsilon A$ is defined to be the directed graph with vertices corresponding to the V_i (they are denoted by i) and with an edge going from i to j when the operator A_{ij} is non-zero.

Example. For the Mathieu operator

$$\Lambda + \varepsilon A = d^2/dt^2 + \varepsilon \cos t$$

defined in the space of 2π-periodic functions of t the space V_k is spanned by $e^{\pm ikt}$, and A is the operator of multiplication by $\cos t$. Therefore, the quiver has the form

$$\overset{0}{\cdot} \rightleftarrows \overset{1}{\cdot} \rightleftarrows \overset{2}{\cdot} \rightleftarrows \cdots$$

In quivers for operators of Mathieu type each vertex is directly connected with a finite (and even uniformly bounded) number of vertices. This is because a product of trigonometric polynomials is again a trigonometric polynomial.

We return to the general family $\Lambda+\varepsilon A$.

Definition. The *chronological product* of a path $\gamma = (i \to j \to \dots \to k \to l)$ along arrows of a quiver is defined to be the product of the operators corresponding to the arrows, in the order indicated by the arrows:

$$\Pi_\gamma(A) = (A_{kl}\cdot\ \dots\ \cdot A_{ij})\colon\ V_i \to V_l.$$

It is easy to prove the following result (by comparing the coefficients of ε, ε^2, ...; see §4).

***Theorem* 1.** *The perturbation of a simple eigenvalue is given by the series*

$$\lambda = \lambda_0 + \varepsilon\alpha_1 + \varepsilon^2\alpha_2 + \ldots,$$

in which the coefficient of the k-th power of the perturbation is equal to the sum

$$\alpha_k = \sum \Pi_\gamma(A)Q_\gamma(\Lambda), \quad |\gamma| = k,$$

over all loops of length k made up of arrows in the quiver and beginning and ending at the point 0.

Here $\Pi_\gamma(A)$ is the chronological product of a loop, and $Q_\gamma(\Lambda)$ is a universal (independent of A) rational function of those eigenvalues of the unperturbed operator that correspond to the vertices in the loop γ.

Remark 1. The explicit form of the rational function Q is not needed for the present, but it is given in §4.

Remark 2. If the quiver is locally finite, then there are finitely many terms in the sum for α_k, so the formal series for λ can be written. Its convergence for operators of Mathieu type is ensured by the ellipticity of the principal term, but we shall regard this series as formal (or we shall assume that the operators are finite-dimensional).

§2. Perturbations of a multiple eigenvalue

If λ_0 is a multiple eigenvalue, then under a perturbation it does not vary smoothly, but the corresponding invariant space V_0 and the restriction of the operator to this space do vary smoothly. The expansions of these objects in power series in ε are given by precisely the same formulae as the perturbations of an eigenvector and an eigenvalue when the latter is simple.

We look for the perturbed invariant space in the form of the graph of a mapping

$$\varepsilon B^1 + \varepsilon^2 B^2 + \ldots, \quad B^k\colon V_0 \to \oplus V_i, \quad i \neq 0.$$

The condition that the graph be invariant is written as the operator equation

$$(\Lambda + \varepsilon A)(E^0 \oplus \varepsilon B^1 + \ldots) = (E^0 \oplus \varepsilon B^1 + \ldots)(\lambda_0 E^0 + \varepsilon\alpha_1 + \ldots),$$

where the $\alpha_k : V_0 \to V_0$ are unknown linear operators, as are the B^k, and E^0 is the identity transformation on V_0.

The sum in the last set of parentheses plays the role of the perturbed eigenvalue. Indeed, this operator, which acts in the space V_0, is similar to the restriction of $\Lambda+\varepsilon A$ to the perturbed invariant space. Therefore, its spectrum coincides with the part of the spectrum of $\Lambda+\varepsilon A$ which arose from λ_0. We shall call it an *eigenoperator* of the family $\Lambda+\varepsilon A$.

***Theorem* 2.** *The coefficient $\alpha_k : V_0 \to V_0$ in the expansion $\lambda_0 E^0 + \varepsilon\alpha_1 + \varepsilon^2\alpha_2 + \dots$ of an eigenoperator of the family $\Lambda + \varepsilon A$ is given by the same formula in Theorem 1 as the coefficient α_k in the expansion of a simple eigenvalue.*

Indeed, this operator equation can be solved in precisely the same way as the scalar equation. For example, equating the coefficients of ε, we find that

$$\Lambda B^1 + AE^0 = E^0\alpha_1 + B^1\lambda_0,$$

from which, representing B^1 in the form $B^1 = \oplus B_i^1$, where $B_i^1 : V_0 \to V_i$ for $i \neq 0$, we find that

$$\alpha_1 = A_{00}, \quad B_i^1 = A_{0i}/(\lambda_0 - \lambda_i).$$

Equating the coefficients of ε^2, we obtain

$$\Lambda B^2 + AB^1 = E^0\alpha_2 + B^1\alpha_1 + B^2\lambda_0.$$

Equating the V_0-components of the left-hand and right-hand sides, we find that

$$\alpha_2 = \sum A_{i0}B_i^1 = \sum A_{i0}A_{0i}/(\lambda_0 - \lambda_i).$$

And equating the V_m-components for $m \neq 0$, we find that

$$(\lambda_0 - \lambda_m)\, B_m^2 = (AB^1)_m - (B^1\alpha_1)_m,$$

from which

$$\sum A. \qquad \dots_m - \quad (\lambda_0 - \lambda_m)(\lambda_0 - \lambda_i) \qquad (\lambda_0 - \lambda_m)^2$$

and so on (the general formula for the denominators of all the terms is written out in §4).

§3. Widths of forbidden zones of even Hill's equations of Mathieu type

Let us consider Hill's equation

$$d^2x/dt^2 + (\omega^2 + \varepsilon a(t))x = 0, \quad a(t + 2\pi) \equiv a(t).$$

The zones of parametric resonance (or *forbidden zones*) are defined as the regions in the plane of the parameters (ω, ε) in which the multipliers (the eigenvalues of the monodromy operator for the given equation with periodic coefficients) are greater than 1 in modulus.

***Theorem* 3.** *Assume that the coefficient a is an even trigonometric polynomial of degree p:*

$$a(t) = \sum a_s e^{ist}, \text{ the } s \text{ are integers, } a_s = 0 \text{ for } |s| > p, \text{ and } a_{-s} = a_s.$$

Then the width of the N-th forbidden zone decreases no more slowly than $C\varepsilon^r$ as $\varepsilon \to 0$, where $r = -[-N/p]$.

Proof. According to the general theory of Hill's equation, the multipliers are equal to -1 or 1 on the boundary of a zone, depending on the parity of the number of the zone. Therefore, the boundaries of the zones can be determined with the help of the eigenvalues $\lambda^{\pm}$ of the operator

$$d^2/dt^2 + \varepsilon a(t)$$

on the space of 2π-periodic functions (for even zones) or on the space of 4π-periodic functions which change sign under translation by 2π (for odd zones):

$$\omega^{\pm} = +\sqrt{-\lambda^{\pm}}.$$

For $\varepsilon = 0$ the N-th zone contracts to the point $\omega_N^{\pm} = N/2$, and the corresponding pairs of eigenfunctions have the form

$$e_n,\ e_{-n},\ \text{where}\ \ e_n = e^{int},\quad n = N/2,\quad N = 0,\ 1,\ 2,\ \ldots$$

(the eigenvalue is simple when $N = 0$).

To investigate the perturbed eigenvalues λ we apply the general formula in Theorems 1 and 2 to the diagonal operator $\Lambda = d^2/dt^2$ and to the perturbation εA (where A is multiplication by a).

Let V_m denote the space spanned by e_m and e_{-m} (this space corresponds to $\lambda_m = -m^2$, where $m = M/2$, and M runs through either all the even or all the odd non-negative integers, depending on the parity of the number N of the zone under investigation).

***Lemma* 1.** *The matrix of the operator $A_{ij} : V_i \to V_j$ in the chosen basis has the form*

$$(*)\qquad (A_{ij}) = \begin{pmatrix} a_{j-i} & a_{j+i} \\ a_{-j-i} & a_{i-j} \end{pmatrix}.$$

Proof. Multiplication by e_s carries e_i into e_{i+s}. Therefore, in the product of the trigonometric polynomial $a = \sum a_s e_s$ by e_i each term falls in an eigenspace different from V_j except for the terms with $s = j - i$ and $s = -j - i$, and similarly for ae_{-i}.

Lemma 1 gives us the next result.

***Lemma* 2.** *If $i+j > p$, then the matrix (A_{ij}) is scalar: $(A_{ij}) = a_{j-i}E$.*

Indeed, the degree of the polynomial a is equal to p, therefore, $a_{i+j} = a_{-i-j} = 0$. Since a is an even polynomial, the diagonal elements of the matrix (*) are all the same, and this proves Lemma 2.

We arrange the quiver of the family $\Lambda + \varepsilon A$ on the m-axis. Its vertices are all the non-negative integer points if the number (N) of the zone being investigated is even, and all the positive half-integer points m if N is odd. The length of each arrow in the quiver does not exceed the degree p of the perturbing polynomial a.

Definition. An arrow $i \to j$ in a quiver is said to be *remote* if the sum of the coordinates of its initial and terminal points is larger than the degree of the perturbing polynomial a:

$$i + j > p.$$

Example. A remote arrow does not contain the point 0.

The following simple lemma contains the main, "topological" point in the proof of Theorem 3.

Lemma 3. *Consider a loop of k arrows, each of length at most p, that begins and ends at the point n. If $kp < 2n$, then all the arrows in the loop are remote.*

Proof. Suppose that there is a non-remote arrow $i \to j$ in the loop, where $i+j \leqslant p$. We replace the vertex j and all the subsequent vertices and arrows of the loop by the vertices and arrows symmetric with respect to the origin. The arrow $i \to j$ itself is replaced by the arrow $i \to -j$. Its length is not greater than p, because the arrow $i \to j$ is non-remote. Accordingly, we have obtained a path from n to $-n$ made up of k arrows, each of length at most p. The length of the whole path is at most kp. But $kp < 2n$, by hypothesis. Hence, there is no non-remote arrow in the loop.

Completion of the proof of Theorem 3. Let us use Theorem 2. To compute the correction to the repeated eigenvalue $\lambda(0) = -n^2$ we form the eigenoperator (second-order matrix)

The matrix α_k can be expressed as a sum over admissible loops consisting of k arrows in our quiver and going from n to n (in the present notation the point 0 in Theorems 1 and 2 is denoted by n). The length of each arrow is at most p, since the perturbation has degree p. By Lemma 3, all the arrows in such a loop are remote when $kp < 2n$. By Lemma 2, a scalar (in our basis) matrix (A_{ij}) corresponds to a remote arrow $i \to j$. Therefore, all the chronological products in the formula for α_k consist of scalar matrices. Hence, the matrix α_k is scalar as long as $k < 2n/p$. But the difference between the two eigenvalues of the matrix $\lambda(\varepsilon)$ does not exceed $C\varepsilon^K$, where K is the number of the first non-scalar matrix among the α_k. We proved above that $K \geqslant r = -[-2n/p]$. Thus, the distance between the two eigenvalues $\lambda^{\pm}(\varepsilon)$ (and hence the width of the zone) decreases no more slowly than ε^r, which is what we needed to prove.

Remark 1. We have not used the fact that the perturbing polynomial is real: the estimate $|\lambda^+ - \lambda^-| \leqslant C\varepsilon^r$, $r = -[-2n/p]$, for the splitting of the eigenvalue $\lambda = -n^2$ under perturbation by a trigonometric polynomial of degree p has been proved for all even polynomials with complex coefficients.

Remark 2. Without the assumption that the perturbing polynomial a is even, our arguments prove only that the matrices α_k with $k < 2n/p$ are diagonal, and not that they are scalar. In reality the diagonal elements are equal even without the assumption of evenness, but to prove this we have to use not only the form of the numerators but also the denominators of the terms in the Rayleigh-Schrödinger series.

§4. Formulae for the denominators of terms in the perturbation theory series

We return to the perturbation theory series and to the notation in §§1 and 2. In order to write out explicitly the rational function Q in Theorems 1 and 2 we consider *properly formed symbols with parentheses*, of the type $((a))b(c(de))f$ ($a($ $)$, $)a($, and $(ab))$ are improperly formed symbols).

Definition. The *length* of a symbol is defined to be the number of letters and left parentheses appearing in it.

Example. Symbols of length 2:

$$ab, \ (a).$$

Symbols of length 3 occur in five forms:

$$abc, \ (a)b, \ a(b), \ (ab), \ ((a)).$$

There are 14 forms of symbols of length 4.

Remark. The numbers of forms of symbols of lengths $n = 1, 2, \ldots$ make up the *Catalan sequence* 1, 2, 5, 14, 42, 132, ...; the n-th number in this sequence is also equal to the number of simplicial decompositions of a convex $(n+2)$-gon by non-intersecting diagonals, and is given by the formula[1] $(2n)!/n!(n+1)!$

Definition. A *filling* of a symbol with parentheses is defined to be the result of substituting in it (in place of the abstract letters a, b, ...) non-zero vertices of the quiver of the family $\Lambda + \varepsilon A$ under investigation (the vertex 0 corresponds to the eigenvalue λ_0 being perturbed).

Example. Fillings of the symbols of length 2 with parentheses have the form

$$ij \ \text{ or } \ (i),$$

where $i \neq 0 \neq j$; i can coincide with j, but if $i \neq j$, then the fillings ij and ji are different.

---

[1]The Catalan numbers appear in the solution of very diverse problems, and are not always identified. For example, a general Riemann surface of genus $2k$ can be represented as a $(k+1)$-sheeted covering of the sphere (with $6k$ branch points) in finitely many ways. Griffiths and Harris [3] find this number (there are 1, 2, 5, 14, ... ways for the values of the genus 2, 4, 6, 8, ...), but do not note that they have obtained the Catalan numbers.

Definition. The *loop* of a filling of length n is defined to be the loop

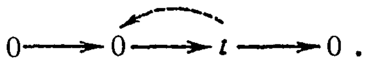
$$0 \to i_1 \to \dots \to i_n \to 0,$$

where between the end 0's we have placed in succession (from left to right) the vertices of the filling and 0's wherever there are left parentheses in the filling (the right parentheses are ignored in constructing the loop).

Example. The loop of the filling (i) is $0 \to 0 \to i \to 0$.

Definition. The *diagram of a filling* is obtained from the loop of the filling by adding a dashed arrow leading to each 0 which replaced a left parenthesis, from the vertex immediately preceding (to the left of) the right parenthesis corresponding to this left parenthesis.

Example. The diagram of the filling (i) is



Definition. The *denominator of a filling* is defined to be the product of factors of the form $\lambda_0 - \lambda_i$ (once for each solid arrow of the diagram of the filling that leads to the vertex $i \neq 0$) and factors of the form $\lambda_i - \lambda_0$ (once for each dashed arrow of the diagram that leads from the vertex $i \neq 0$).

Example. The denominator of the filling (i) is $-(\lambda_0 - \lambda_i)^2$, according to the diagram in the preceding example.

A filling is said to be *admissible* if its loop consists of arrows of the quiver.

Theorem 4. *The k-th correction to the eigenvalue (or eigenoperator) $\Lambda = \Lambda_0 E^0 + \varepsilon\alpha_1 + \varepsilon^2\alpha_2 + \dots$ under the perturbation $\Lambda + \varepsilon A$ is given by the formula*

$$\alpha_k = \sum_{|I| = k-1} \frac{\Pi_{\gamma(I)}(A)}{q_I(\Lambda)},$$

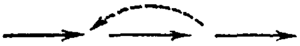
where the sum is over all admissible fillings of all the symbols of length $k-1$ with parentheses, and the contribution of each filling is equal to the ratio of the chronological product of its loop to the denominator of this filling.

Example.

$$\alpha_1 = A_{00}, \quad \alpha_2 = \sum_{i \neq 0} \frac{A_{i0}A_{0i}}{\lambda_0 - \lambda_i},$$

$$\alpha_3 = \sum_{i \neq 0 \neq j} \frac{{}_{j0}A_{ij}A_{0i}}{(\lambda_0 - \lambda_i)(\lambda_0 - \lambda_j)} - \sum_{i \neq 0} \frac{A_{i0}A_{0i}A_{00}}{(\lambda_0 - \lambda_i)^2}$$

(the sum of the contributions of the diagrams $\to \to \to$ and

).

Remark 1. Theorems 1 and 2 are obtained from this theorem by summing the contributions of the diagrams with a common loop. Thus, the rational function $Q_\gamma(\Lambda)$ in Theorem 1 is equal to the sum of the reciprocals of the denominators of all the fillings with a given loop γ.

Remark 2. If a loop does not return to the point 0 between its beginning and end, then there are no dashed arrows, and $Q = \prod (\lambda_0 - \lambda_i)^{-1}$ (the product over all internal vertices of the loop).

For example, if $A_{00} = 0$ (which is always easy to achieve by including A_{00} in λ_0), then all the loops of length < 4 in the quiver that go from 0 to 0 do not visit 0 on the way. Consequently, for $A_{00} = 0$

$$\alpha_1 = 0, \quad \alpha_2 = \sum_{i \neq 0} \frac{A_{i0}A_{0i}}{\lambda_0 - \lambda_i}, \qquad \alpha_3 = \sum_{i \neq 0 \neq j} \frac{A_{j0}A_{ij}A_{0i}}{(\lambda_0 - \lambda_i)(\lambda_0 - \lambda_j)}.$$

The formula for α_4 is even more complicated. But the *first non-zero* coefficient in α_k is always equal to the sum of the contributions of the loops that do not visit 0 and, consequently, is given by a simple formula analogous to the ones written out.

Theorem 4 can be proved at the same time as the formula for the corrections to the invariant subspace (Theorem 5). To write this formula we define for each filling I the *path* $\delta(I)$ *of the filling*, which is obtained from the loop of the filling by discarding the last (right-hand) vertex 0 along with the arrow leading to it.

A filling is said to *m-admissible* if its path goes from 0 to m along arrows of the quiver. We look for the invariant space of $\Lambda + \varepsilon A$ obtained by deformation of V_0 in the form of the graph of a mapping

$$B = \varepsilon B^1 + \varepsilon^2 B^2 + \ldots, \quad B^k\colon V_0 \to \oplus V_m, \ m \neq 0.$$

The components of the operator B^k are denoted by

$$B^k_m\colon V_0 \to V_m, \quad m \neq 0.$$

Theorem 5_k. *The operators correcting the invariant space are given by the formula*

$$B^k_m = \sum_{|I|=k} \frac{\Pi_{\delta(I)}(A)}{q_I(\Lambda)},$$

where the sum is over all the m-admissible fillings of length k, and the contribution of each filling is the chronological product of the path of the filling, divided by the deonominator of the filling.

Example 1. The formula written out in §2 for B^2_m is the sum of the contributions of the fillings *im* and *(m)*.

Example 2. If the eigenvalue λ_0 is simple and ξ_0 is an unperturbed eigenvector, then the perturbed eigenvector has the form

$$\xi_0+\varepsilon\sum\xi_m^1+\varepsilon^2\sum\xi_m^2+\ldots,\quad \xi_m^k=B_m^k\xi_0.$$

Proof of Theorems 4_k *and* 5_k. In the invariance equation

$$(\Lambda+\varepsilon A)(E^0+\varepsilon B^1+\ldots)=(E^0+\varepsilon B^1+\ldots)(\lambda_0E^0+\varepsilon\alpha_1+\ldots)$$

we equate successively the coefficients of ε, ε^2, ... For B^k we obtain the equation

$$B^k\lambda_0-\Lambda B^k=AB^{k-1}-B^{k-1}\alpha_1-\ldots-B^1\alpha_{k-1}-\alpha_k.$$

For this equation to be soluble it is necessary that

$$\alpha_k=\sum A_{m0}B_m^{k-1},\qquad m\neq 0.$$

Therefore, Theorem 4_k follows from Theorem 5_{k-1}. Choosing α_k in this way, we find that

$$(*_k)\quad B_m^k=(\lambda_0-\lambda_m)^{-1}\left(\sum A_{im}B_i^{k-1}-B_m^{k-1}\alpha_1-\ldots-B_m^1\alpha_{k-1}\right),\quad i\neq 0.$$

Computing all the B_m^k successively, we obtain the chronological products and prove Theorems 1 and 2. The explicit form of the denominators can be found by proving that Theorem 5_k follows from Theorems 4_l and 5_l, $0<l<k$.

Indeed, the first term of the expression $(*_k)$ for B_m^k is the sum of the contributions of the m-admissible fillings of length k that are obtained from [illegible]

on the right-hand side. Each of the remaining terms, say the one containing $B_m^{k-s}\alpha_s$, is the sum of the products of the contributions from the admissible fillings of length $s-1$ and the m-admissible fillings of length $k-s$, divided by $\lambda_m-\lambda_0$. From two such fillings we can make a filling of length k by appending a parenthesized m-admissible filling of length $k-s$ (from the sum for B_m^{k-s}) on the *right-hand side* of an admissible filling of length $s-1$ (from the sum for α_s). Here the contribution of the resulting filling of length k is exactly equal to the product of the contributions of the component fillings, divided by $\lambda_m-\lambda_0$ (this superfluous factor in the denominator of the filling of length k arises because of the parentheses enclosing the m-admissible filling of length $k-s$). Conversely, each m-admissible filling of length k is obtained—precisely once—under these operations, for there is a unique way to dismantle it from the right-hand end (if a letter is on the right-hand end, then it is removed, and if there is a parenthesis, then the symbol enclosed in it is detached and the corresponding outer parentheses are taken away).

Theorem 5_k thus follows from the preceding Theorems 4_l and 5_l. The assertions 4_1 and 5_1 are obvious, so Theorems 4 and 5 are proved.

Suppose now that $\alpha_1=\alpha_2=\ldots=\alpha_{k-1}=0$. In this case the answer becomes much simpler.

***Theorem* 6.** *The first non-zero correction to the eigenvalue (or eigenoperator) is given by the formula*

$$\alpha_k = \sum_{|\gamma|=k} \frac{\Pi_\gamma(A)}{\prod(\lambda_0-\lambda_i)}$$

(the summation is over the loops of length k in the quiver that start and end at 0 *but do not visit* 0, *and the multiplication in the denominator is over all the internal vertices of the loop).*

Proof. In this case the terms with $\alpha_1, \ldots, \alpha_{l-1}$ drop out in the computations according to the formula $(*_l)$ with $l < k$, therefore, $B_i^l = (\lambda_0 - \lambda_i)^{-1}\sum A_{ji}B_j^{l-1}$ $(l = 1, \ldots, k-1)$, which leads to the formula in the theorem.

§5. Conclusion of the investigation of forbidden zones

We now consider the Hill's equation of Mathieu type

$$d^2x/dt^2 + (\omega^2 + \varepsilon a(t))x = 0$$

with an arbitrary trigonometric polynomial of degree p with complex coefficients:

$$a(t) = \sum a_s e^{ist}, \text{ the } s \text{ are integers, and } |s| \leqslant p.$$

In §3 we introduced the eigenvalues $\lambda^\pm$ (which determine the ends of the forbidden zone if the polynomial a is real).

***Theorem* 7.** *The degree of splitting of the eigenvalue* $\lambda = -n^2$, *where n is a positive integer or a half-integer, admits for small* $|\varepsilon|$ *the upper estimate*

$$|\lambda^+ - \lambda^-| \leqslant C|\varepsilon|^r, \quad r = -[-2n/p].$$

Proof. Arguing as in the proof of Theorem 3, we get sums over loops of remote arrows as expressions for the coefficients α_k in the expansion of the eigenoperator in a series of powers of ε when $k < 2n/p$. Consequently, all the matrices (A_{ij}) (§3) appearing in the chronological products with such a number k of factors are diagonal and thus commute with one another. This implies that the matrices α_k are also diagonal, but we shall see that they are actually scalar due to special symmetry properties of our expansions.

Definition. Two second-order diagonal matrices are said to be *dual* if they differ only in the order of the diagonal elements.

Suppose that the arrow $i \to j$ is remote, so that the second-order matrix (A_{ij}) is diagonal. The formula given in §3 implies the next result.

***Lemma* 1.** *The diagonal elements* (a_{j-i}, a_{i-j}) *of the two matrices* (A_{ij}) *and* (A_{ji}) *differ only in their order.*

From this it is clear that any polynomial in all such matrices, in which the matrices corresponding to oppositely directed arrows appear symmetrically, is a scalar matrix.

Let us show that in the expression for α_k in Theorem 4 the matrices corresponding to oppositely directed arrows appear symmetrically. It will thereby be proved that the matrices α_k with $k < r$ are scalar, and Theorem 7 obviously follows from this.

***Lemma* 2.** *On the set of all admissible fillings of all symbols of length $k-1$ with parentheses there is an involution such that the contributions to α_k of a filling and its image under the involution are obtained from each other by replacing all the A_{ij} by A_{ji}.*

Theorem 7 follows from the formula in Theorem 4 and Lemma 2, since the matrices α_k are scalar matrices for $k < r$.

Proof of Lemma 2. We break up a filled symbol with parentheses into the fragments between the successive left parentheses. For example, the symbol $a((b)(cd))e(fg)$ is broken up into the fragments a, $b)$, $cd))e$, and $fg)$. Then inside each fragment we write all the vertices in the reverse order, regarding each right parenthesis as associated with the letter closest to it on the left. The fragments a, $b)$, $ed))c$, and $g)f$ are obtained in the example given. We next place left parentheses in the previous places between the fragments. The filling $a((b)(ed))c(g)f$ is obtained in the example.

The constructed operation carries a properly formed symbol with parentheses into the properly formed *dual symbol*, as it does not change the order of the parentheses and does not create pairs of the form ( ).

A filling dual to an admissible filling is admissible, because the matrices A_{ij} and A_{ji} are either both zero or both non-zero. We compute the contributions of dual fillings.

***Lemma* 3.** *The chronological products of mutually dual fillings appearing in the expansion of α_k for $k < r$ are mutually dual.*

For by Lemma 1, the factors are mutually dual and commute.

***Lemma* 4.** *The denominators of dual fillings coincide.*

Indeed, the denominator of a filling is composed of factors $\lambda_0 - \lambda_i$ with i running through the elements of the filling, and factors $\lambda_i - \lambda_0$ with i running through the elements preceding the right parentheses (where each factor is counted as many times as there are right parentheses immediately after the corresponding i).

Both the set of elements in the filling and the set of elements which preceded right parentheses (taking account of the number of parentheses) are preserved under the operation defined above, as Lemma 4 proves.

Lemmas 3 and 4 imply the equality of the contributions of dual fillings and, hence, Lemma 2 and Theorem 7.

§6. Widths of resonance zones for mappings of the circle

We consider an analytic mapping of the circle onto itself that depends on two parameters a and ε:

$$x \mapsto x + a + \varepsilon f(x), \quad f(x + 2\pi) \equiv f(x).$$

Definition. A point (a, ε) belongs to the *domain of resonance* m/n if some point of the circle is shifted by exactly m revolutions when the mapping is applied n times (that is, if the Poincaré rotation number of the mapping is equal to m/n).

***Theorem* 8.** *Let f be a trigonometric polynomial of degree p. Then the width of the domain of resonance m/n is at most $C\varepsilon^r$, where $-r$ is the integer part of the fraction $-n/p$.*

Moreover, the domain of resonance m/n is bounded by two curves $a = A^{\pm}(\varepsilon)$ such that the terms of degree less than r in the Taylor expansions of $A^{\pm}$ about zero coincide.

Example. Consider the mapping $x \mapsto x + a + \varepsilon \cos x$. In this case the width of the domain of resonance m/n is of order ε^n, as claimed already in [2].

The proof of the theorem is based on an investigation of the r-th variation of the k-th iterate of a rotation. Consider an analytic mapping

$$x \mapsto x + \mu + g(x), \quad g(x + 2\pi) \equiv g(x),$$

that is close to the rotation through the angle $\mu = 2\pi m/n$. Its k-th iterate is close to the rotation through the angle $k\mu$:

$$x \mapsto x + k\mu + {}_kg(x).$$

Let us consider the Taylor expansion at the point $g = 0$ for the operator carrying g into ${}_kg$. We call the term of degree r in this expansion the *r-th variation of the k-th iterate.*

Remark. In order not to have to worry about convergence we can understand the variations in the following formal way. Assume that g is a linear combination of any finite number of terms: $g = \sum \lambda_q h_q$. Then ${}_kg$ is a series in powers of λ, and the r-th variation is the sum of the terms of degree r in λ in this series.

***Lemma* 1.** *The r-th variation of the k-th iterate can be expressed as a polynomial of degree r in terms of g, the shifts of g by the angles $s\mu$ with $0 \leqslant s < k$, and their derivatives of order less than r.*

Proof. This is obvious for $k = 1$. Assume that it is true for the k-th iterate. By definition, for the $(k+1)$st iterate

$$\text{(1)} \qquad {}_{k+1}g(x) = {}_kg(x) + g(x + k\mu + {}_kg(x)).$$

Denote the shifts of g by

$$g_k(x) = g(x + k\mu).$$

Then

$$g(x + k\mu + {}_kg(x)) = g_k(x + {}_kg(x)).$$

We expand g_k in the Taylor series $\sum g_k^{(s)}({}_kg)^s/s!$, and substitute for ${}_kg$ the series of variations

$${}_kg = {}_kg^1 + {}_kg^2 + \dots$$

By assumption, ${}_kg^i$ is a polynomial of degree i in g, the shifts, and the derivatives of order less than i. The terms with $i \geqslant r$ and the terms with $s \geqslant r$ do not contribute to the r-th variation. Therefore, these transformations of the formula (1) express the r-th variation of the $(k+1)$st iterate as the sum of the r-th variation of the k-th iterate and a sum of products of derivatives of order less than r of the shift of g by the angle $k\mu$ and products of the preceding variations of the k-th iterate. The lemma is proved by induction.

Remark. The first two variations of the k-th iterate can be expressed by simple formulae:

$${}_kg^1 = g_0 + g_1 + \dots + g_{k-1},$$
$${}_kg^2 = g_1'g_0 + g_2'(g_0 + g_1) + \dots + g_{k-1}'(g_0 + \dots + g_{k-2}).$$

Consider now the n-th iterate of the mapping. Since $\mu = 2\pi m/n$, the n-th iterate of the unperturbed mapping returns all the points of the circle to their places.

On the level of the second term in the perturbation theory the theorem being proved reduces to the following.

Proposition. *Let g be a trigonometric polynomial of degree less than $n/2$, and suppose that ${}_ng^1 = 0$ (that is, the mean value of g is equal to 0). Then* ${}_ng^2 = \text{const}$.

Proof. Consider the difference

$${}_ng^2(x + \mu) - {}_ng^2(x) = g_n'(g_1 + \dots + g_{n-1}) - g_0(g_1 + \dots + g_{n-1})'.$$

Since $g_n = g_0$ and $g_1 + \dots + g_n = 0$, this difference is equal to zero. Consequently, ${}_ng^2$ is a trigonometric polynomial of degree less than n with period $2\pi/n$, that is, it is a constant. The proposition is proved. It implies our theorem for $r \leqslant 3$.

The proof in the general case uses an analogous argument, formalized by Lemma 2 below.

We consider a one-parameter analytic family of mappings of the circle onto itself:

$$x \mapsto x + \mu + g(\varepsilon, x), \quad g(\varepsilon, x + 2\pi) \equiv g(\varepsilon, x),$$

where $\mu = 2\pi m/n$ and where

$$g = \varepsilon g^1 + \varepsilon^2 g^2 + \dots$$

***Lemma* 2.** *Assume that the n-th iterate of the mapping shifts points by a distance of order ε^r, that is,*

$$_n g(\varepsilon, x) = \varepsilon^r u(x) + o(\varepsilon^r).$$

Then the principal term in the shift is invariant under the rotation through the angle μ:

$$u(x + \mu) \equiv u(x). \tag{2}$$

Proof. Consider the images x_i of a point $x_0 = x$ under the iterates of the mapping. By the condition in the lemma,

$$g(x_0) + g(x_1) + \dots + g(x_{n-1}) = \mu n + \varepsilon^r u(x_0) + o(\varepsilon^r),$$

$$g(x_1) + g(x_2) + \dots + g(x_n) = \mu n + \varepsilon^r u(x_1) + o(\varepsilon^r).$$

Consequently,

$$g(x_n) - g(x_0) = \varepsilon^r [u(x + \mu) - u(x)] + o(\varepsilon^r). \tag{3}$$

On the other hand,

$$x_n = x_0 + 2\pi m + \varepsilon^r u(x_0) + o(\varepsilon^r),$$

and, consequently,

$$g(x_n) - g(x_0) = g'(x_0)\varepsilon^r u(x_0) + o(\varepsilon^r). \tag{4}$$

Since $g' = O(\varepsilon)$, (3) and (4) imply (2).

Proof of the theorem. We consider the following two-parameter family of perturbations of the rotation through the angle $\mu = 2\pi m/n$:

$$x \mapsto x + \mu + H(\alpha, \varepsilon, x), \quad H(\alpha, \varepsilon, x) = \alpha + \varepsilon f(x),$$

where ε and α are small parameters.

Denote the n-th iterate by

$$x \mapsto x + \mu n + {}_n H(\alpha, \varepsilon, x).$$

The resonance condition is the solubility of the equation ${}_n H = 0$ with respect to x. The expansion of ${}_n H$ in a Taylor series in α and ε has the form

$${}_n H = {}_n H^1 + {}_n H^2 + \dots,$$

where each term ${}_n H^i$ is a homogeneous polynomial of degree i in α and ε (the value of the i-th variation of the n-th iterate for $g = \alpha + \varepsilon f$). For example,

$${}_n H^1 = n\alpha + \varepsilon + (f_0 + \dots + f_{n-1}), \text{ where } f_k(x) = f(x + k\alpha).$$

By the implicit function theorem, the equation ${}_nH = 0$ is soluble for $\alpha = \alpha(\varepsilon, x)$:

$$\alpha = \varepsilon v_1(x) + \varepsilon^2 v_2(x) + \dots$$

Let r be the index of the first non-constant coefficient v_i, so that $\alpha(\varepsilon, x) = A(\varepsilon) + \varepsilon^r v(x) + o(\varepsilon^r)$, where A is a polynomial of degree less than r, and v is a non-constant function.

The boundaries of the domain of resonance m/n have asymptotic expressions $\alpha = A(\varepsilon) + \varepsilon^r v_\pm + o(\varepsilon^r)$, where $v_\pm$ are the maximum and minimum of v on the circle. It remains to get a lower estimate of r.

We substitute $A(\varepsilon)$ in place of α in H and denote the resulting function by g:

$$g(\varepsilon, x) = H(A(\varepsilon), \varepsilon, x).$$

After n iterations of the mapping $x \mapsto x + \mu + g(\varepsilon, x)$ we get a rotation through a variable angle of order ε^r. Indeed,

$$\alpha(\varepsilon, x) - A(\varepsilon) = \varepsilon^r v(x) + o(\varepsilon^r),$$

therefore,

$${}_nH(\alpha(\varepsilon, x), \varepsilon, x) - {}_nH(A(\varepsilon), \varepsilon, x) = n\varepsilon^r v(x) + o(\varepsilon^r).$$

But ${}_nH(\alpha(\varepsilon, x), \varepsilon, x) = 0$ by the definition of $\alpha(\varepsilon, x)$. Accordingly,

$${}_ng(\varepsilon, x) = {}_nH(A(\varepsilon), \varepsilon, x) = \varepsilon^r u(x) + o(\varepsilon^r), \quad u = -nv.$$

By Lemma 2, the function u (hence also v) satisfies the periodicity condition $u(x+\mu) \equiv u(x)$.

On the other hand, Lemma 1 gives us the next result.

Lemma 3. *The function u can be expressed as a polynomial of degree at most r in terms of f, shifts of f by angles that are multiples of μ, and their derivatives* (*of order less than r*).

Proof. If we substitute the series without the free term $\alpha = A(\varepsilon)$ in the Taylor series of ${}_nH(\alpha, \varepsilon, x)$ with respect to α and ε, then the new coefficient of ε^r can be expressed linearly in terms of the old coefficients of terms of degree $\leqslant r$ in α and ε together. By Lemma 1, the old terms of degree i in α and ε can be expressed as polynomials of degree i in terms of shifts of the fucntion $g = \alpha + \varepsilon f$ by angles that are multiples of α and their derivatives with respect to x (for constant α and ε) of bounded order. This proves Lemma 3, because $\partial g/\partial x = \varepsilon \partial f/\partial x$.

Completion of the proof of the theorem. Assume that f is a trigonometric polynomial of degree p and that $rp < n$. Then u is a trigonometric polynomial of degree less than n, by Lemma 3. From Lemma 2 we concluded above that u has period $2\pi m/n$. A trigonometric polynomial of degree less than n with period $2\pi m/n$ is a constant. But u is not a constant (by the choice of r). Hence, $r \geqslant n/p$, and Theorem 8 is proved.

Remark. A diffeomorphism of the circle can have many cycles. Is the number of isolated cycles of a diffeomorphism given by a trigonometric polynomial bounded by a constant depending only on the degree of the polynomial?

This question can be regarded as an analogue of the question in Hilbert's 16th problem on the number of limit cycles of a polynomial differential equation. It has not even been solved for the diffeomorphisms $x \mapsto x + a + b \cos x$.

Conversely, the asymptotic expressions obtained above for the boundaries of domains of existence of cycles probably have analogues in the theory of differential equations with a polynomial right-hand side.

Supplement (30 March 1983)

While the manuscript of the present article was being edited, the author learned of some papers containing other proofs of some of the results: by C. Bloch [5] (Theorems 1, 2, 4–6), D.M. Levy and J.B. Keller [6] (Theorem 3), and H. Hochstadt [7] (Theorem 7).

Bloch labels the terms in the perturbation theory series by diagrams having the form of integer staircases joining the vertices (0, 0) and (n, n) inside the square without dropping below the diagonal.

The number d_n of such diagrams is equal to the corresponding Catalan number ($d_2 = 2$, $d_3 = 5$, ...). Bloch's work thus gives another interpretation of the Catalan numbers. The most natural definition of them is probably as the numbers of words of given lengths in a non-associative monoid with a single generator. It would be interesting to find out whether the non-associative monoid itself is connected with the perturbation theory (or with coverings of Riemann surfaces, where the Catalan numbers are also encountered).

The correspondence between symbols with parentheses and Bloch diagrams is as follows: a left parenthesis means a step upwards, a right parenthesis means a step to the right, and a letter means two steps (first upwards, and then to the right).

Moreover, Manin has communicated to the author that an approach to perturbation theory close to that presented above was worked out for different reasons in a manuscript of M. Wodzycki written in July 1982 and entitled "Variation of the ζ-function of an elliptic pseudodifferential operator".

The author thanks M. Wodzcki, I.M. Gel'fand, A.A. Kiselev, P.A. Kuchment, V.F. Lazutkin, Yu.I. Manin, and S.P. Novikov for useful discussions.

References

[1] V.I. Arnol'd, On the evolution of a magnetic field under the action of transport and diffusion, Collection of papers in memory of V.M. Alekseev, Izdat. Moskov. Gos. Univ., Moscow 1983.

[2] ———, Small denominators. I, Mappings of the circumference onto itself, Izv. Akad. Nauk Ser. Mat. **25** (1961), 21-86. MR **25** # 4113.
= Amer. Math. Soc. Transl. (2) **46** (1965), 213-254.

[3] P. Griffiths and J. Harris, Principles of algebraic geometry, Wiley, New York 1978. MR **80b** # 14001.
Translation: *Printsipy algebraicheskoi geometrii,* Mir, Moscow 1982.

[4] A.A. Kiselev, A diagram technique in the general theory of perturbations and in the adiabatic theory of molecular spectra, in: *Voprosy kvantovoi teorii atomov i molekul* (Questions in the quantum theory of atoms and molecules), No. 1, Izdat. Leningrad. Gos. Univ., Leningrad 1978, 108-143.

[5] C. Bloch, Sur la théorie des perturbations des états liés, Nuclear Physics **6** (1958), 329-347. Nuclear Science Abstracts **12** # 9529.

[6] D.M. Levy and J.B. Keller, Instability intervals of Hill's equation, Comm. Pure Appl. Math. **16** (1963), 469-476. MR **27** # 3875.

[7] H. Hochstadt, Instability intervals of Hill's equation, Comm. Pure Appl. Math. **17** (1964), 251-255. MR **28** # 2310.

Translated by H. McFaden

Moscow State University

Received by the Editors 1 February 1983