# Chapter 1
# Preliminary Facts

**Abstract** In this chapter we present some definitions and statements from the points of view of both physicists and mathematicians to be used in the next chapters. We mean especially the definitions of the lattices, periodic functions, Brillouin zones, Schrödinger operator, Bloch eigenvalues, Bloch functions, diffraction planes, band structures and Fermi surfaces. Moreover, we try to explain the transition between these notions due to the understanding of the physicists and mathematicians. Besides, we give a brief discussion of what is known from the literature and what is presented in the book about the perturbation theory of the multidimensional Schrödinger operator with a periodic potential. For this aim we consider the large Bloch eigenvalues and the corresponding Bloch functions of the one-dimensional periodic Schrödinger operator by the approach of Chap. 2, since it helps to compare the well-known one-dimensional case with the multidimensional case and to see the complexity of the results obtained in this book.

## 1.1 Lattices, Brillouin Zones, and Periodic Functions in $\mathbb{R}^d$

The structure of the crystals can be described in terms of the lattice (called in geometry and crystallography, a Bravais lattice), with a group of atoms attached to every lattice point. The Bravais lattice in

$$\mathbb{R}^d =: \left\{(x_1, x_2, \ldots, x_d) : x_1 \in \mathbb{R},\ x_2 \in \mathbb{R}, \ldots,\ x_d \in \mathbb{R}\right\},$$

where $\mathbb{R}$ is the set of all real numbers, is defined by $d$ linearly independent vectors $\omega_1, \omega_2, \ldots, \omega_d$. In the case $d = 3$ these vectors are known as fundamental translations vectors such that every atomic arrangement looks the same in every respect when viewed from the point $\mathbf{r}$ as when viewed from the point

$$\mathbf{r} + \sum_{k=1}^{3} n_k \omega_k,$$



O. Veliev, *Multidimensional Periodic Schrödinger Operator*,
Springer Tracts in Modern Physics 263, DOI 10.1007/978-3-319-16643-8_1

where $n_1, n_2$ and $n_3$ are integers. The lattice $\Omega$ generated by the vectors $\omega_1$, $\omega_2, \ldots, \omega_d$ is the set of all linear combinations of these vectors with the integer coefficients:

$$\Omega = \left\{ \omega = \sum_{k=1}^{d} n_k \omega_k : n_1 \in \mathbb{Z},\ n_2 \in \mathbb{Z}, \ldots, n_d \in \mathbb{Z} \right\}, \tag{1.1.1}$$

where $\mathbb{Z}$ is the set of all integers. The vectors $\omega_1, \omega_2, \ldots, \omega_d$ used for the generation of $\Omega$ are known as the primitive vectors or basis vectors for the lattice. The parallelotope ($d$-dimensional parallelogram)

$$F = \left\{ x = \sum_{k=1}^{d} y_k \omega_k : y_1 \in [0, 1),\ y_2 \in [0, 1), \ldots, y_d \in [0, 1) \right\} \tag{1.1.2}$$

is called the fundamental parallelotope or the primitive unit cell of the lattice. In the cases $d = 2$ and 3 the parallelotope $F$ is the parallelogram and parallelepiped, respectively. It has the origin in $\mathbb{R}^d$ as one corner and the vectors $\omega_1, \omega_2, \ldots, \omega_d$ form the sides which meet at that corner. Thus a crystal is characterized by its regular periodically repeated structure. The smallest unit of this structure is called the primitive unit cell. The primitive cells (parallelotopes) are joined together filling the entire volume and giving rise to the periodicity of the crystal lattice.

The measure $\mu(F)$ (generalized volume) of the parallelotope $F$ is equal to the absolute value of the determinant of the $d \times d$ matrix $(\omega_{i,j})$ created from the $d$ row vectors

$$\omega_1 = (\omega_{1,1}, \omega_{1,2}, \ldots, \omega_{1,d}),\ \omega_2 = (\omega_{2,1}, \omega_{2,2}, \ldots, \omega_{2,d}), \ldots,\ \omega_d = (\omega_{d,1}, \omega_{d,2}, \ldots, \omega_{d,d}).$$

Everywhere, for simplicity of notation and without loss of generality we assume that the generalized volume (measure $\mu(F)$) of the parallelotope $F$ is equal to 1. Thus

$$\mu(F) = \left|\det(\omega_{i,j})\right| = 1. \tag{1.1.3}$$

There are infinitely many choices for the basis vectors and hence for the unit cells. In other words, the set of generators for a lattice is not uniquely determined. It is well-known that the vectors $b_1, b_2, \ldots, b_d$ are the other generators of $\Omega$ if and only if there is a $d \times d$ matrix $A = (a_{i,j})$ with integer matrix elements and $|\det A| = 1$ such that

$$b_i = \sum_{j=1}^{d} a_{i,j} \omega_j$$

for $i = 1, 2, \ldots, d$. Therefore, condition (1.1.3) is not a restriction for the choices for the basis vectors of the lattice $\Omega$.

Note that when $F$ is translated through all the vectors in the lattice $\Omega$ fills all of the space $\mathbb{R}^d$ without overlapping. Therefore the fundamental domain (unit cell) $F$ of the lattice $\Omega$ can be identified with the factor space (quotient group) $\mathbb{R}^d/\Omega$ which is the set of equivalent classes, where the equivalence of two elements $x$ and $y$ of $\mathbb{R}^d$ is defined as follows: we say that $x$ and $y$ are equivalent if $x - y \in \Omega$. Thus any measurable set $M$ that contains, for each $x \in \mathbb{R}^d$, exactly one representative of the set

$$x + \Omega =: \{x + y : y \in \Omega\}$$

is called a unit cell of the lattice $\Omega$. It is also clear that $\mathbb{R}^d/\Omega$ is a d-dimensional torus (direct product of $d$ circles).

We say that a function $f : \mathbb{R}^d \to \mathbb{C}$ is periodic with respect to the lattice $\Omega$ if

$$f(x + \omega) = f(x)$$

for all $\omega \in \Omega$, where $x = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$ and $\mathbb{C}$ is the set of all complex numbers. Note that the periodic function $f$ can be regarded in this case as a function on the torus $\mathbb{R}^d/\Omega$. It is clear that the wave function $e^{i\langle\gamma,x\rangle}$ is periodic with respect to the lattice $\Omega$ if and only if

$$\langle\gamma, \omega\rangle \in 2\pi\mathbb{Z}, \tag{1.1.4}$$

for all $\omega \in \Omega$, where $\gamma \in \mathbb{R}^d$ and $\langle\cdot,\cdot\rangle$ is the inner product in $\mathbb{R}^d$. The set of all vectors $\gamma \in \mathbb{R}^d$ satisfying (1.1.4), that is,

$$\Gamma =: \{\gamma \in \mathbb{R}^d : \langle\gamma, \omega\rangle \in 2\pi\mathbb{Z}, \forall\omega \in \Omega\} \tag{1.1.5}$$

is the lattice dual to $\Omega$ and is called the reciprocal lattice. The basis vectors of the reciprocal lattice $\Gamma$ are the vectors $\gamma_1, \gamma_2, \ldots, \gamma_d$ satisfying

$$\langle\gamma_i, \omega_i\rangle = 2\pi \quad \& \quad \langle\gamma_i, \omega_j\rangle = 0 \tag{1.1.6}$$

for $i, j = 1, 2, \ldots, d$ and $j \neq i$. Thus the fundamental parallelotope of the lattice $\Gamma$ is

$$F^* = \left\{\omega = \sum_{k=1}^{d} a_k\gamma_k : a_1 \in [0, 1),\ a_2 \in [0, 1), \ldots, a_d \in [0, 1)\right\}. \tag{1.1.7}$$

As we noted above, $F^*$ can be identified with the fundamental domain $\mathbb{R}^d/\Gamma$ of the lattice $\Gamma$.

The other and famous fundamental domains (unit cells) of the reciprocal lattice $\Gamma$ are the Brillouin zones. The first Brillouin zone (called Brillouin zone) of $\Gamma$ is defined to be the set of points $x \in \mathbb{R}^d$ in reciprocal space which are nearer (not necessarily unique) to the origin than any point $x + \gamma$ with $\gamma \in \Gamma$ and $\gamma \neq 0$.

The $n$th Brillouin zone is the set of all points $x$ in the reciprocal space which have the origin as their (not necessarily unique) $n$th nearest point of the set

$$x + \Gamma = \left\{ y \in \mathbb{R}^d : \ y = x + \gamma, \gamma \in \Gamma \right\}. \tag{1.1.8}$$

Note that any interior point of the $n$th Brillouin zone is the unique $n$th nearest point. If the several points of (1.1.8) are the $n$th nearest points (i.e. are equidistant from the origin) then these points belong to the boundaries of the Brillouin zones and only one of them belongs to the $n$th Brillouin zone.

It readily follows from this definition the following properties of the Brillouin zone:

(a) All zones have equal volumes,

(b) Each zone can be translated into the first zone so as to fill it exactly by translating different pieces of the zone by appropriate reciprocal lattice-vectors.

(c) For arbitrary fixed $n$ the $n$th Brillouin zone contains unique element from any equivalent classes defined as follows: $x$ and $y$ are equivalent if $x - y \in \Gamma$. Therefore the Brillouin zones can be identified with the fundamental domain $\mathbb{R}^d / \Gamma$ of the lattice $\Gamma$.

The geometrical description of the Brillouin zones will be given in Sect. 1.3.

Now let us give the brief description of the problem discussed above. The reciprocal lattice vectors are the special wave vectors $\gamma$ for which the free electron wave function $e^{i\langle \gamma, x\rangle}$ is periodic with respect to the direct lattice. The wave vectors having this property will be said to belong to the reciprocal lattice. The primitive vectors $\gamma_1, \gamma_2, \ldots, \gamma_d$ of the reciprocal lattice can be generated from the primitive vectors $\omega_1, \omega_2, \ldots, \omega_d$ of the direct lattice by the equalities (1.1.6). A crystal is made up of a periodic arrangement of one or more atoms (the basis) repeated at each Bravais lattice point. Consequently, the crystal looks the same when viewed from any equivalent lattice point, namely those separated by the translation of one unit cell. Every periodic function is associated with a Bravais lattice. You can think of the function as being defined in a primitive unit cell and then repeating the primitive unit cell at every point of the Bravais lattice.

As we noted above the wave function $e^{i\langle \gamma, x\rangle}$ is periodic, with respect to the lattice $\Omega$, if and only if $\gamma \in \Gamma$. One can easily verify that the system

$$\left\{ e^{i\langle \gamma, x\rangle} : \ \gamma \in \Gamma \right\} \tag{1.1.9}$$

is an orthonormal basis in the Hilbert space $L_2(F)$ of square integrable functions with the inner product

$$(f, g) = \int_F f(x)\overline{g(x)}dx.$$

Indeed, by (1.1.3) we have

$$\left\| e^{i\langle \gamma, x\rangle} \right\|^2 = \int_F \left| e^{i\langle \gamma, x\rangle} \right|^2 dx = \int_F 1 dx = \mu(F) = 1,$$

where $\|\cdot\|$ is the norm in the space $L_2(F)$ defined by

$$\|f\| = \left(\int_F |f(x)|^2\, dx\right)^{1/2}.$$

The orthogonality of the system (1.1.9) means that

$$\left(e^{i\langle\gamma,x\rangle}, e^{i\langle\widetilde{\gamma},x\rangle}\right) = \int_F e^{i\langle\delta,x\rangle} dx = 0$$

for all $\widetilde{\gamma} \neq \gamma$, where $\delta = \gamma - \widetilde{\gamma} \in \Gamma$. The last integral can be calculated by using the substitution

$$(x_1, x_2, \ldots, x_d) \leftrightarrow (y_1, y_2, \ldots, y_d),$$

where $y_1, y_2, \ldots, y_d$ equal to the coefficients of the expansion $x$ in the basis $\omega_1, \omega_2, \ldots, \omega_d$ [see (1.1.2)] and by (1.1.2) this substitution transforms the parallelotope $F$ to the cube $[0, 1)^d$. Moreover the Jacobian $J$ of this substitution is nonzero since the vectors $\omega_1, \omega_2, \ldots, \omega_d$ are linearly independent. Therefore using

$$x = \sum_{k=1}^{d} y_k \omega_k \tag{1.1.10}$$

and taking into account that $\delta \in \Gamma \setminus \{0\}$, that is,

$$\delta = \sum_{k=1}^{d} n_k \gamma_k$$

where $n_1, n_2, \ldots, n_d$ are integers and at least one of them is not zero we have

$$\int_F e^{i\langle\delta,x\rangle} dx = |J| \int_0^1 \int_0^1 \ldots \int_0^1 e^{i2\pi n_1 y_1} e^{i2\pi n_2 y_2} \ldots e^{i2\pi n_d y_d} dx_1 dx_2 \ldots dx_d = 0.$$

Since the system

$$\left\{e^{i2\pi n_1 y_1} e^{i2\pi n_2 y_2} \ldots e^{i2\pi n_d y_d} : n_1 \in \mathbb{Z}, n_2 \in \mathbb{Z}, \ldots, n_d \in \mathbb{Z}\right\}$$

is complete in $L_2\left([0, 1)^d\right)$, the above substitution shows that (1.1.9) is complete in the Hilbert space $L_2(F)$ and hence is an orthonormal basis. Therefore every function $q \in L_2(F)$ has the decomposition

$$q(x) = \sum_{\gamma\in\Gamma} q_\gamma e^{i\langle\gamma,x\rangle}, \tag{1.1.11}$$

where

$$q_\gamma = \left(q, e^{i\langle\gamma,x\rangle}\right) = \int_F q(x)e^{-i\langle\gamma,x\rangle}dx$$

for $\gamma \in \Gamma$ are the Fourier coefficients of $q$ with respect to the orthonormal system (1.1.9) and the Fourier series (1.1.11) converges to $q$ in the norm of $L_2(F)$. If

$$\sum_{\gamma\in\Gamma} |q_\gamma| < \infty,$$

then the series (1.1.11) converges uniformly to the periodic function $q$.

The smoothness of $q$ depends on the Fourier coefficients. For simplicity, let us first consider the case $d = 1$. Let $\Omega = \mathbb{Z}$. Then $\Gamma = 2\pi\mathbb{Z}$ and the system

$$\{e^{i2\pi nx} : n \in \mathbb{Z}\} \tag{1.1.12}$$

is the orthonormal basis in $L_2[0, 1]$. Using the integrations by part, one can readily see that if the $s$th derivative of the periodic functions $q$ of period 1 belongs to $L_2[0, 1]$ then the Fourier coefficient $q_n^{(s)}$ of $q^{(s)}$ with respect to (1.1.12) satisfies the equality

$$q_n^{(s)} =: (2\pi n)^{-s} q_n$$

where

$$q_n = \int_0^1 q(x)e^{-i2\pi nx}dx$$

is the Fourier coefficient of $q$. Therefore the periodic function $q$ belongs to the Sobolev space

$$W_2^s[0, 1] =: \left\{f : f^{(s)} \in L_2[0, 1]\right\}$$

if and only if

$$\sum_{n\in\mathbb{Z}} |2\pi n|^{2s}|q_n|^2 < \infty$$

Similarly for arbitrary dimension $d$ the relation $q \in W_2^s(F)$ for the periodic, with respect to the lattice $\Omega$, function $q$ means that

$$\sum_{\gamma\in\Gamma} |q_\gamma|^2|\gamma|^{2s} < \infty$$

## 1.2 Schrödinger Operator and Bloch Functions

The energy operator is often referred to as the Hamiltonian and it is also called (in nonrelativistic quantum mechanics) the Schrödinger operator. The Schrödinger operator with a periodic potential arises in the quantum theory of crystals and describes the motion of a particle in the crystal. The ions forming a crystal lattice $\Omega$ actually generate a periodic field and one can examine the motion of a electron in this field. Thus if $V(x)$ is the potential seen an electron at $x$ then $V(x+\omega) = V(x)$ for all $\omega \in \Omega$. The wave function $u(x)$ of the electron placed in the periodic potential $V$ must satisfy the Schrödinger equation

$$-\frac{\hbar^2}{2m}\Delta u(x) + V(x)u(x) = Eu(x),$$

where

$$\Delta u = \sum_{j=1}^{d} \frac{\partial^2 u}{\partial x_j^2},$$

$\hbar$ is Planck's constant, $m$ and $E$ are respectively the mass of the electron and its energy eigenvalue.

In the mathematical literature the Schrödinger equation is written in the form

$$-\Delta u(x) + q(x)u(x) = \Lambda u(x), \tag{1.2.1}$$

where

$$q(x) = \frac{2m}{\hbar^2}V(x),\ \Lambda = \frac{2m}{\hbar^2}E.$$

The Schrödinger operator $L(q)$ with a real periodic, relative to a lattice $\Omega$, potential $q$ is defined in space $L_2(\mathbb{R}^d)$ as follows, where $L_2(\mathbb{R}^d)$ is the Hilbert space of square integrable functions with the inner product

$$(f,g)_{\mathbb{R}^d} = \int_{\mathbb{R}^d} f(x)\overline{g(x)}dx.$$

Let $D$ be the set of all functions $u \in L_2(\mathbb{R}^d)$ such that

(i) $u$ is compactly supported, that is, the set

$$\left\{x \in \mathbb{R}^d : f(x) \neq 0\right\}$$

is a bounded closed subset of $\mathbb{R}^d$,

(ii) $\dfrac{\partial u}{\partial x_j}$ exists and is an absolutely continuous function of $x_j$ for $j = 1, 2, \ldots, d$,

(iii) $-\Delta u + qu \in L_2(\mathbb{R}^d)$.

Let $L^0(q)$ be an operator defined in $D$ by

$$L^0(q)u = -\Delta u + qu.$$

One can readily verify that $L^0(q)$ is a symmetric operator, that is,

$$\left(L^0(q)f, g\right)_{\mathbb{R}^d} = \left(f, L^0(q)g\right)_{\mathbb{R}^d}$$

for all $f, g \in D$. The Schrödinger operator (Hamiltonian) $L(q)$ is the self-adjoint extension of $L^0(q)$. The existence and uniqueness of the extension are well-known (see [BeShu]).

Now we consider the connection of the Hamiltonian $L(q)$ with the Bloch Functions. Recall that Bloch wave or Bloch state, named after Felix Bloch, is the wave function of a particle (usually, an electron) placed in a periodic potential $q$. Bloch's theorem states that for a particle moving in the periodic potential, the wave functions $\Psi(x)$ are of the form

$$\Psi(x) = e^{i\langle t,x\rangle} p(x), \tag{1.2.2}$$

where $p(x)$ is a periodic function with the same periodicity that the potential $q$ has and $t \in \mathbb{R}^d$ is a crystal momentum (quasimomentum). The exact form of $p(x)$ depends on the potential associated with atoms (ions) that form the solid. The motion of an electron in the free space, where the potential $q$ is zero everywhere, is described by the simplest form of the Schrödinger equation

$$-\Delta u(x) = \lambda u(x)$$

and the wave function $e^{i\langle t,x\rangle}$ is the solution of this equation, since

$$-\Delta e^{i\langle t,x\rangle} = |t|^2 e^{i\langle t,x\rangle}.$$

Thus by Bloch's theorem the wave function $\Psi(x)$ of the electron in the periodic potential is the product of the wave function $e^{i\langle t,x\rangle}$ of the electron in the free space and the periodic function $p(x)$. The wave function expressed by Eq. (1.2.2) is called the Bloch wave or Bloch state.

The Bloch's theorem is very important, since by applying this theorem, the wave function in a macroscopic crystal containing as many atoms as the Avogadro number can be determined by solving the Schrödinger equation into which information from just one unit cell is inserted.

One of the often used (in mathematics) forms of Bloch's theorem is the following (see [Eas]):

*Theorem* (*Bloch*) *Let $S$ consist of the real numbers $\Lambda$ for which the* Eq. (1.2.1) *has a non-trivial bounded solution in $\mathbb{R}^d$. If $\Lambda \in S$ then* (1.2.1) *has a solution $\Psi_t(x, \Lambda)$ of the form*

$$\Psi_t(x, \Lambda) = e^{i\langle t,x\rangle} p(x), \tag{1.2.3}$$

*where $p$ is a periodic function having the same periodicity that the potential $q$* has, *the vector $t \in \mathbb{R}^d$ in* (1.2.3) *is called a crystal momentum* (*quasimomentum*) *and $S$ is said to be stability set of the* Eq. (1.2.1)

The solution of (1.2.1) of the form (1.2.3) is called the Bloch solution of (1.2.1) (see [Ku]). It readily follows from (1.2.3) that if $\omega \in \Omega$, where $\Omega$ is the period lattice of the potential $q$ and hence of $p$, then

$$\Psi_t(x+\omega, \Lambda) = e^{i\langle t, x+\omega\rangle} p(x+\omega) = e^{i\langle t,x\rangle} e^{i\langle t,\omega\rangle} p(x) = e^{i\langle t,\omega\rangle} \Psi_t(x, \Lambda)$$

Therefore the Bloch solution $\Psi_t(x, \Lambda)$ of (1.2.1) can be considered as an eigenfunction of the eigenvalue problem (1.2.1) and

$$u(x+\omega) = e^{i\langle t,\omega\rangle} u(x) \tag{1.2.4}$$

for all $\omega \in \Omega$. Conversely, if $\Psi(x, \Lambda)$ is an eigenfunction of this eigenvalue problem then by (1.2.4) we have

$$|\Psi(x+\omega, \Lambda)| = |\Psi(x, \Lambda)|$$

for all $\omega \in \Omega$. It implies that $\Psi(x, \Lambda)$ is bounded in $\mathbb{R}^d$ and by Bloch's theorem has the form (1.2.3), that is, $\Psi(x, \Lambda)$ is the Bloch solution of (1.2.1). Thus $\Psi(x, \Lambda)$ is a Bloch solution of (1.2.1) if and only if it is an eigenfunction of the eigenvalue problem (1.2.1) and (1.2.4) for some values of the quasimomentum $t \in \mathbb{R}^d$. The corresponding eigenvalue $\Lambda(t)$ is called the Bloch eigenvalue for the crystal momentum $t$. In other words, the Bloch eigenvalue $\Lambda(t)$ and Bloch function $\Psi_t(x, \Lambda)$ for fixed crystal momentum $t$ are the eigenvalue and eigenfunction of $-\Delta + q$ acting on the space

$$\left\{ u \in H^2_{loc}(\mathbb{R}^d) : u(x+\omega) = e^{i\langle t,\omega\rangle} u(x), \forall \omega \in \Omega \right\},$$

where $H^2_{loc}(\mathbb{R}^d)$ is the space of locally square integrable functions $u$ such that $\partial^\alpha u$, for $|\alpha| \le 2$, is also locally square integrable.

In the language of the operator theory the Bloch eigenvalue $\Lambda(t)$ and Bloch function $\Psi_t(x, \Lambda)$ for fixed crystal momentum $t$ are the eigenvalue and eigenfunction of the differential operator $L_t(q)$ generated in $L_2(F)$ by the differential expression

$$-\Delta u(x) + q(x)u(x) \tag{1.2.5}$$

and the boundary conditions (1.2.4), where in the writing the boundary conditions in the form (1.2.4) we take it that the eigenfunction $u$ is extended to the whole $\mathbb{R}^d$ as continuously differentiable functions. More precisely, the operator $L_t(q)$ can be defined in $L_2(\overline{F})$ as the differential operator generated by (1.2.5) and the boundary conditions

$$u(x+\omega_j) = e^{i\langle t,\omega_j\rangle} u(x),\ u_{y_j}(x+\omega_j) = e^{i\langle t,\omega_j\rangle} u_{y_j}(x) \tag{1.2.6}$$

for $x \in \overline{F}(j)$ and $j = 1, 2, \dots, d$, where $\overline{F}$ is the closure of the parallelotope (1.1.2), that is,

$$\overline{F} =: \left\{ x = \sum_{k=1}^{d} y_k \omega_k : y_1 \in [0, 1], y_2 \in [0, 1], \dots, y_d \in [0, 1] \right\} \tag{1.2.7}$$

is the closed parallelotope,

$$\overline{F}(j) =: \left\{ x = \sum_{k \in \{1,2,\dots,d\} \setminus \{j\}} y_k \omega_k : y_1 \in [0, 1], y_2 \in [0, 1], \dots, y_d \in [0, 1] \right\} \tag{1.2.8}$$

is the face of the boundary $\partial\overline{F}$ of the parallelotope $\overline{F}$ generated by $\omega_1, \omega_2, \dots, \omega_{j-1}, \omega_{j+1}, \omega_{j+2}, \dots, \omega_d$ and $u_{y_j} =: \frac{\partial u}{\partial y_j}$ is the derivative of $u$ with respect to the variable $y_j$ defined by (1.1.10) [see also (1.2.7)].

Note that the boundary conditions (1.2.6) mean that the values of $u$ and $u_{y_j}$ on the face $\omega_j + \overline{F}(j)$ of $\partial\overline{F}$ are equal to $e^{i\langle t, \omega_j \rangle}$ times of their values on opposite face $\overline{F}(j)$. The boundary conditions (1.2.6) are equivalent to the conditions (1.2.4) if, as we noted above, in the writing the boundary conditions in the form (1.2.4) we take it that the eigenfunction $u$ is extended to the whole $\mathbb{R}^d$ as continuously differentiable functions. Therefore in the next chapters for simplicity we say that the operator $L_t(q)$ is generated in $L_2(F)$ by the differential expression (1.2.5) and boundary conditions (1.2.4). Thus the operator $L_t(q)$ is defined as follows. Domain of definition $D(L_t(q))$ of $L_t(q)$ is the set of $u \in L_2(\overline{F})$ such that:

(a) $\dfrac{\partial u}{\partial x_j}$ exists and is an absolutely continuous function of $x_j$ for $j = 1, 2, \dots, d$,

(b) $-\Delta u + qu \in L_2(\overline{F})$,

(c) $u$ satisfies the boundary conditions (1.2.6).

For $u \in D(L_t(q))$ the operator $L_t(q)$ is defined by

$$L_t(q)u = -\Delta u + qu$$

It is well-known the following statements about the spectral properties of $L_t(q)$ and $L(q)$:

*Theorem* (*On the spectra of the operators $L_t(q)$ and $L(q)$*).

(a) *The spectrum $\sigma(L_t(q))$ of the operator $L_t(q)$ is discrete and consists of the eigenvalues*

$$\Lambda_1(t) \le \Lambda_2(t) \le \cdots \tag{1.2.9}$$

*such that $\Lambda_j(t) \to \infty$ as $j \to \infty$ which are the Bloch eigenvalues with the fixed quasimomentum $t$. The corresponding normalized eigenfunctions* (*Bloch functions*)

$$\Psi_{1,t}(x), \Psi_{2,t}(x), \ldots$$

*form an orthonormal basis in* $L_2(F)$.

(b) *The function* $\Lambda_n$ *is continuous with respect to* $t$ *and its range*

$$\delta_n =: \left\{\Lambda_n(t) : t \in F^*\right\},$$

*where* $F^*$ *is the fundamental parallelotope of the reciprocal lattice* $\Gamma$, *is a closed interval of* $\mathbb{R}$.

(c) *The operator* $L(q)$ *has no eigenvalue and has only the continuous spectrum. The spectrum* $\sigma(L(q))$ *of the operator* $L(q)$, *the stability set* $S$ *defined in the above formulation of Bloch's theorem, and the union of the spectra of the operators* $L_t(q)$ *for* $t \in F^*$ *are the same, that is,*

$$\sigma(L(q)) = S = \bigcup_{t\in F^*} \sigma(L_t(q)) = \bigcup_{t\in F^*} \left( \bigcup_{n=1}^{\infty} \{\Lambda_n(t)\} \right) = \bigcup_{n=1}^{\infty} \delta_n. \tag{1.2.10}$$

Thus $\sigma(L(q))$ consists of the intervals $\delta_n$ for $n = 1, 2, \ldots$, that are called the band of the spectrum of $L(q)$. The spaces between neighboring bands are called the band gaps or the gaps in the spectrum of $L(q)$. In the physical literature these bands and gaps are named as energy bands (allowed regions of energy) and forbidden regions of energy respectively

Note that the rigorous proof of this theorem can be found in [Eas] (see also [BeShu, ReSi]). First the physicists observed that the spectrum of $L(q)$ has a band structure [SomBe, Ki, Mad]. The eigenfunctions $\Psi_{1,t}(x)$, $\Psi_{1,t}(x), \ldots$, of $L_t(q)$ for all values of the quasimomentum $t$ are the Bloch waves [Bl]. For the multidimentional case Gelfand proved Parseval's relation for the Bloch waves in $L_2(\mathbb{R}^d)$ [Gel]. Oder and Keller [OdKe] proved that the spectrum of $L(q)$ is the union of all Bloch eigenvalues $\Lambda_1(t), \Lambda_2(t), \ldots$, for all $t \in F^\star$. Thomas [Th] proved that the spectrum of $L(q)$ is absolutely continuous. Wilson [Wi] studied the analytic properties of $\Lambda_n(t)$ as a function of the quasimomentum $t$.

Now let us discuss this theorem from the point of view of the mathematicians and physicists. The statement (a) follows from the fact that $L_t(q)$ is a self-adjoint operator defined in a bounded region of $\mathbb{R}^d$.

Now we discuss (b). The function $\Lambda_n$ is continuous with respect to $t$ due to the following. Let $P_n(x)$ be a function defined by

$$P_n(x) = e^{-i\langle t,x\rangle}\Psi_{n,t}(x), \tag{1.2.11}$$

where $\Psi_{n,t}(x)$ is the eigenfunction of $L_t(q)$ corresponding to the eigenvalue $\Lambda_n(t)$, that is,

$$-\Delta\Psi_{n,t}(x) + q(x)\Psi_{n,t}(x) = \Lambda_n(t)\Psi_{n,t}(x), \tag{1.2.12}$$

$$\Psi_{n,t}(x+\omega) = e^{i\langle t,\omega\rangle}\Psi_{n,t}(x), \forall\omega \in \Omega. \tag{1.2.13}$$

Using (1.2.11), (1.2.12), and (1.2.13) one can easily verify that $P_n(x)$ satisfies the following equalities

$$-\triangle P_n(x) - \langle 2it, \nabla\rangle\, P_n(x) + \langle t, t\rangle\, P_n(x) + q(x)P_n(x) = \Lambda_n(t)P_n(x)$$

and

$$P_n(x+\omega) = P_n(x) \tag{1.2.14}$$

for all $\omega \in \Omega$. Hence $\Lambda_n(t)$ is the eigenvalue of the operator generated by the operation

$$-\triangle - \langle 2it, \nabla\rangle + \langle t, t\rangle + q \tag{1.2.15}$$

and the periodic boundary conditions. Since the periodic boundary conditions do not depend on $t$ and the operation (1.2.15) continuously depends on $t$ the eigenvalue $\Lambda_n(t)$ also continuously depends on $t$. Therefore its range

$$\delta_n =: \left\{\Lambda_n(t) : t \in F^*\right\}, \tag{1.2.16}$$

where $F^*$ is the fundamental parallelotope of the reciprocal lattice $\Gamma$, is an interval of $\mathbb{R}$. The closedness of $\delta_n$ will be discussed later.

Now let us discuss $(c)$. The operator $L(q)$ is associated with the whole space $\mathbb{R}^d$ and by the Floquet theory (see [Ku]) the Schrö dinger equation (1.2.1) has no solution belonging to $L_2(\mathbb{R}^d)$. Therefore $L(q)$ has no eigenvalue. In fact, the numbers $\Lambda_n(t)$ are not the eigenvalues of the operator $L(q)$ since the corresponding Bloch solutions $\Psi_{n,t}(x)$ do not belong to $L_2(\mathbb{R}^d)$ and by definition, $\Lambda$ is an eigenvalue of the operator $L(q)$ if there exists

$$\Psi \in D(L(q)) \subset L_2(\mathbb{R}^d)$$

such that

$$L(q)\Psi = \Lambda\Psi. \tag{1.2.17}$$

Therefore Bloch eigenvalues are called the generalized eigenvalues of the operator $L(q)$. However, in some literatures $\Lambda_n(t)$ is named as an eigenvalue of $L(q)$; that is natural, say, in the following sense. Instead of the operator $L(q))$ in whole space $\mathbb{R}^d$ one can consider an operator $L(q, \mathbf{n})$ in the very large parallelotope

$$\overline{F}_{\mathbf{n}} = \left\{ x = \sum_{k=1}^{d} y_k\omega_k : y_1 \in [-n_1, n_1],\, y_2 \in [-n_2, n_2], \ldots, y_d \in [-n_d, n_d] \right\}, \tag{1.2.18}$$

with the periodic boundary conditions, where $\mathbf{n} = (n_1, n_2, \ldots, n_d)$ and $n_1$, $n_2, \ldots, n_d$ are large positive integers. Due to the fact that $\mathbb{R}^d$ is a limit of $\overline{F}_{\mathbf{n}}$ as $n_j \to \infty$ for $j = 1, 2, \ldots, d$, the eigenvalues of the operator $L(q, \mathbf{n})$ or the limit points of its eigenvalues can be named (in some sense) the eigenvalues of $L(q)$.

Moreover using this argument, it was proved that (see [Eas]) the set of limit points, of the eigenvalues of $L(q,\mathbf{n})$ as $n_j \to \infty$ for $j = 1,2,\dots,d$, coincides with $\sigma(L(q))$. On the other hand, using another argument one can see that the set of all eigenvalues of $L(q,\mathbf{n})$ and their limit points as $n_j \to \infty$ for $j = 1,2,\dots,d$, coincide with the set of Bloch eigenvalues

$$\left\{\Lambda_n(t) : t \in F^\star, n \in \mathbb{N}\right\}.$$

These arguments encourage to believe the validity of (1.2.10).

To be more precise let us define the operator $L(q,\mathbf{n})$ precisely. Moreover consideration the Schrödinger operator in the bounded and large parallelotope $\overline{F}_{\mathbf{n}}$ is interesting, since an electron in a metal must be confined in a bounded space. The effect of a finite size of a system on the motion of an electron must be taken into account. The electron wave function $u(x)$ is assumed along the parallelotope $\overline{F}_{\mathbf{n}}$. Since macroscopic crystal contains as many atoms as the Avogadro number it is interesting to consider the large parallelotope $\overline{F}_{\mathbf{n}}$ which means that $n_1, n_2, \dots, n_d$ are large numbers. Let us impose the periodic boundary conditions

$$u(x + 2n_j\omega_j) = u(x),\ u_{y_j}(x + 2n_j\omega_j) = u_{y_j}(x) \tag{1.2.19}$$

on this parallelotope for $x \in \overline{F}_{\mathbf{n}}(j)$ and $j = 1,2,\dots,d$, where $\overline{F}_{\mathbf{n}}(j)$ is the face of the boundary $\partial\overline{F}_{\mathbf{n}}$ of the parallelotope $\overline{F}_{\mathbf{n}}$ which is parallel to $\overline{F}(j)$ [see (1.2.8)] and passes through the point $-n_j\omega_j$ and the variable $y_j$ is defined by (1.1.10). Note that the boundary conditions (1.2.19) means that the values of $u$ and $u_{y_j}$ on the face $\overline{F}_{\mathbf{n}}(j)$ of the parallelotope $\overline{F}_{\mathbf{n}}$ are equal to their values on the opposite face.

Let $L(q,\mathbf{n})$ be an operator generated in $L_2(\overline{F}_{\mathbf{n}})$ by the differential expression (1.2.5) and the boundary conditions (1.2.19). Since $L(q,\mathbf{n})$ is associated with the bounded domain $\overline{F}_{\mathbf{n}}$ of $\mathbb{R}^d$ its spectrum is discrete and consists of the eigenvalues. One can readily verify that the set of the eigenvalues of $L(q,\mathbf{n})$ are the union of the Bloch eigenvalues $\Lambda_n(t)$ for $n \in \mathbb{N}$ and $t \in A(\mathbf{n})$, where

$$A(\mathbf{n}) = \left\{ t = \sum_{j=1}^{d} \frac{k_j}{2n_j}\gamma_j : k_j = 0,1,\dots,2n_j;\ j = 1,3,\dots,d \right\}. \tag{1.2.20}$$

Indeed, if $u(x)$ satisfies the first condition of (1.2.6) for $t \in A(\mathbf{n})$, then applying it $2n_j$ times and using (1.1.6) we obtain that

$$u(x + 2n_j\omega_j) = \exp\left(\sum_{j=1}^{d} \frac{k_j}{2n_j}\gamma_j, 2n_j\omega_j\right) u(x) = e^{i2\pi k_j} u(x) = u(x),$$

that is, the first condition of (1.2.19) holds. In the same way one can show that the second condition of (1.2.6) implies the second condition of (1.2.19). The proof of the converse statements are similar (see [Eas]). Thus we have

$$\sigma(L(q,\mathbf{n})) = \bigcup_{t\in A(\mathbf{n})} \sigma(L_t(q)). \tag{1.2.21}$$

Denote by $\Sigma$ the union of the spectrum $\sigma(L(q,\mathbf{n}))$ of the operators $L(q,\mathbf{n})$ for $\mathbf{n}\in\mathbb{N}^d$. It is clear that the closure $\overline{\Sigma}$ of $\Sigma$ is the set of all limit points of the spectrum $\sigma(L(q,\mathbf{n}))$ as $n_j\to\infty$ for $j=1,2,\dots,d$. Since, as we noted above, the set of these limit points is $\sigma(L(q))$, we have

$$\overline{\Sigma} = \sigma(L(q)). \tag{1.2.22}$$

On the other hand, taking into account that the set of all limit points of

$$\left\{\frac{k}{2n} : k = 0, 1, \dots, 2n\right\}$$

as $n\to\infty$ is $[0,1]$ and using the equalities (1.2.20) and (1.1.7) one can readily see that the set of all limit points of $A(\mathbf{n})$ as $n_j\to\infty$ for $j=1,2,\dots,d$ is $F^*$. Therefore (1.2.21) and the continuity of the function $\Lambda_n(t)$ on $F^*$ show that

$$\overline{\Sigma} = \bigcup_{t\in F^*} \sigma(L_t(q)). \tag{1.2.23}$$

Thus we tried to explain the reason of the well-known equalities

$$S = \bigcup_{t\in F^*} \sigma(L_t(q)) = \overline{\Sigma} = \sigma(L(q)). \tag{1.2.24}$$

Now let us discuss the well-known mathematical statements described above and some properties of the Bloch eigenvalues $\Lambda_n(t)$ and the Bloch functions $\Psi_{n,t}(x)$ from the point in view of physicists. Considering $\Lambda_n(t)$ as an eigenvalue of the boundary value problem (1.2.1) and (1.2.4) and taking into account that for any $\gamma\in\Gamma$, where $\Gamma$ is the reciprocal lattice, the equality

$$e^{i\langle t+\gamma,\omega\rangle} = e^{i\langle t,\omega\rangle} \tag{1.2.25}$$

holds, we obtain

$$\Lambda_n(t+\gamma) = \Lambda_n(t) \tag{1.2.26}$$

and

$$\Psi_{n,t+\gamma}(x) = \Psi_{n,t}(x) \tag{1.2.27}$$

for all $\gamma\in\Gamma$.

By (1.2.26) for given $n$ the energy eigenvalue $\Lambda_n(t)$ is periodic with periodicity of a reciprocal lattice. The energies $\Lambda_n(t)$ associated with the index $n$ vary continuously with the wave vector $t$ and form an energy band $\delta_n$ identified by the band index $n$.

All distinct values of $\Lambda_n(t)$ occur for $t$-values within the fundamental domain $\mathbb{R}^d/\Gamma$ of the lattices $\Gamma$, say within the first Brillouin zone or the unit cell (fundamental parallelotope $F^*$) of the reciprocal lattice. In (1.2.26) replacing $\gamma$ by $\gamma_j$ for $j = 1, 2, \ldots, d$ and using (1.1.7) we see that $\Lambda_n(t)$ takes the same values in the opposite faces of the closed parallelotope $\overline{F^*}$. Therefore the bands $\delta_n$ for $n = 1, 2, \ldots,$ of the spectrum of $L(q)$ are closed intervals, since they are the images of the closed parallelotope $\overline{F^*}$ under the continuous function $\Lambda_n(t)$. These intervals are allowed zones of energy and the spaces between the neighboring intervals are forbidden zones.

The Bloch wave energy eigenstate $\Psi_{n,t}(x)$ is written with subscripts $n$ and $t$, where $n$ is a discrete index, called the band index, which is present because there are many different Bloch waves with the same quasimomentum $t$ (each has a different periodic component $p$). Within a band (i.e., for fixed $n$), $\Psi_{n,t}(x)$ varies continuously with $t$, if its energy $\Lambda_n(t)$ is a simple eigenvalue. Since (1.2.27) holds for any reciprocal lattice vector $\gamma$, all distinct Bloch waves occur for $t$-values within the first Brillouin zone of the reciprocal lattice.

Suppose an electron is in a Bloch state $\Psi_{n,t}(x)$. It follows from (1.2.11), (1.2.14), and (1.2.25) that

$$\Psi_{n,t}(x) = e^{i\langle t,x\rangle} P_n(x) = e^{i\langle t+\gamma,x\rangle} P_{n,\gamma}(x), \tag{1.2.28}$$

where $P_n$ and $P_{n,\gamma}$ for $\gamma \in \Gamma$ are periodic with the same periodicity as the crystal lattice $\Omega$. Thus the actual quantum state of the electron is entirely determined by $\Psi_{n,t}(x)$, not $t$ or $P_n(x)$ directly, since $t$ or $P_n(x)$ are not unique. Indeed, if $\Psi_{n,t}(x)$ can be written as above using $t$, it can also be written using $t + \gamma$, where $\gamma$ is any reciprocal lattice vector [see (1.2.27)] and this replacement changes the periodic component $P_n(x)$ in (1.2.28).

Equality (1.2.27) shows that the wave vectors (quasimomenta) that differ by a reciprocal lattice vector are equivalent, in the sense that they characterize the same set of Bloch states. The first Brillouin zone is a restricted set of wave vectors with the property that no two of them are equivalent, yet every possible wave vector is equivalent to one (and only one) vector in the first Brillouin zone. Hence, if we restrict to the first Brillouin zones, then every Bloch state has a unique $t$. Therefore the first Brillouin zone is often used to depict all of the Bloch states without redundancy, for example in a band structure, and it is used for the same reason in many calculations.

## 1.3 Band Structure, Fermi Surfaces and Perturbations

In Sect. 1.2 we discussed the description of the levels of an electron in a periodic potential in terms of a family of continuous functions $\Lambda_n(t)$ called as the band functions. For each $n$, the set of electronic levels specified by $\Lambda_n(t)$ is called an energy band. The information contained in these functions for different $n$ and $t$ is referred to as the band structure of the solid. The electron in the free space corresponds

to the Schrödinger operator with zero potential. In the case $q = 0$ the eigenvalues and eigenfunctions of $L_t(q)$ are $|\gamma + t|^2$ and $e^{i\langle\gamma+t,x\rangle}$ for $\gamma \in \Gamma$, since

$$-\Delta e^{i\langle\gamma+t,x\rangle} = |\gamma + t|^2 e^{i\langle\gamma+t,x\rangle},$$

the function $e^{i\langle\gamma+t,x\rangle}$ satisfies (1.2.4) and the system

$$\left\{e^{i\langle\gamma+t,x\rangle} : \gamma \in \Gamma\right\}$$

is an orthonormal basis in $L_2(F)$.

**(i) Diffraction hyperplanes and Brillouin zones**. The eigenvalue $|\gamma + t|^2$ of $L_t(0)$ coincides with the other eigenvalue $|\gamma + t + \delta|^2$, that is, $|\gamma + t|^2$ is a multiple eigenvalue of $L_t(0)$ if and only if $\gamma + t$ belongs to the diffraction hyperplane

$$D_\delta =: \{x \in \mathbb{R}^d : |x|^2 = |x + \delta|^2\} \tag{1.3.1}$$

for some $\delta \in \Gamma$. By (1.3.1), $x \in D_\delta$ if and only if the points $x$ and $x + \delta$ have the same distance from the origin. Therefore $D_\delta$ is the hyperplane normal to the reciprocal lattice vector $-\delta$ at their midpoint. Moreover by the same reason $D_\delta$ is the boundary of the Brillouin zones defined in Sect. 1.1. The diffraction hyperplanes play a crucial role in the perturbation theory. Let us have a look the diffraction hyperplanes and Brillouin zones in the following cases:

**Case 1**. $d = 1$. Consider the case of one-dimensional Schrödinger operator $L(q)$ with a periodic, with respect to the lattice $\mathbb{Z}$, potential $q$. Then the reciprocal lattice is $2\pi\mathbb{Z}$ and the solution of the equation

$$|x|^2 = |x + 2\pi n|^2$$

in $\mathbb{R}$ is the point $\pi n$. Thus in this case the diffraction hyperplanes are the points $\pi n$ for $n = \pm 1, \pm 2, \ldots$ that are the boundaries of the Brillouin zones. The first Brillouin zone is $(-\pi, \pi]$. The second Brillouin zone is $(\pi, 2\pi] \cup (-2\pi, -\pi]$ and the $n$th Brillouin zone is $((n-1)\pi, n\pi] \cup (-n\pi, -(n-1)\pi]$.

**Case 2**. $d = 2$. Let the reciprocal lattice $\Gamma$ be the two-dimensional lattice in $\mathbb{R}^2$ and $\delta$ be a vector of the lattice. Then $x \in D_\delta$ if and only if $x$ lies in the line normal to the vector $-\delta$ at its midpoints. Thus in this case the diffraction hyperplanes are the lines normal to the reciprocal lattice vectors at their midpoints and the $n$th Brillouin zone is the union of the polygons bounded by the diffraction lines.

Similarly in the case $d = 3$ the diffraction hyperplanes are the planes normal to the reciprocal lattice vectors at their midpoints. Therefore the reciprocal space is partitioned into polyhedra bounded by the planes normal to the reciprocal lattice vectors at their midpoints. These planes are boundaries of the Brillouin zone. Hence the Brillouin zone appears in reciprocal space as an assembly of polyhedra bounded by the planes normal to the reciprocal lattice vectors at their midpoints.

**(ii) Isoenergetic surface**. The isoenergetic surface representing the momentum distribution of the electrons is also constructed in reciprocal space. Note that the isoenergetic surface $I_q(\lambda)$ corresponding to the energy $\lambda$ refers to a constant energy surface and is defined by

$$I_q(\lambda) = \{t \in F^* : \exists N, \Lambda_N(t) = \lambda\},$$

that is, $I_q(\lambda)$ is the set of quasimomenta $t$ in the primitive cell $F^*$ of the reciprocal lattice for which there exists a Bloch eigenvalue $\Lambda_N(t)$ coinciding with the constant energy $\lambda$, where the band function $\Lambda_N(t)$ is defined in Sect. 1.2. This surface for some special and important (in physics) value of $\lambda$ is called the Fermi surface. Since for the free electrons (in the case $q = 0$) the band functions are $|\gamma + t|^2$, the isoenergetic surface $I_0(\lambda)$ in this case is

$$I_0(\lambda) = \{t \in F^* : \exists \gamma \in \Gamma, |\gamma + t|^2 = \rho^2\}$$

which is the translation of the sphere $\{x \in \mathbb{R}^d : |x| = \rho\}$, where $\lambda = \rho^2$, to the primitive cell $F^*$ by the vectors of the reciprocal lattice $\Gamma$. In fact this sphere can be illustrated as the isoenergetic surfaces of the free electron.

**(iii) Perturbation of the free electron**. Now we discuss how the free-electron is perturbed by the periodic potential and then demonstrate it in the one-dimensional case (see **iv**). The effect of the periodic potential on the electron can be treated in the reciprocal space in terms of the interaction of the isoenergetic surface with the diffraction hyperplanes, that is, with the boundaries of the Brillouin zones. The isoenergetic surface begins to be distorted from a sphere before making contacts with the Brillouin zone planes. The gaps in the spectrum emerges as a result of distortion of the isoenergetic surface in the diffraction planes. Recall that the spectrum of the Schrödinger operator $L(q)$ with a periodic potential consists of the energy bands $\delta_n$ for $n = 1, 2, \ldots$, that are defined in (1.2.16) and named as the allowed bands. The gap in the spectrum is the region between the energy bands $\delta_n$ and $\delta_{n+1}$ and in the physical literature is named the forbidden band or the energy gap. This means that the electron is not allowed to take energies between the allowed bands $\delta_n$ and $\delta_{n+1}$ and, hence, there appears an energy discontinuity. Thus an energy gap appears across the Brillouin zone plane. The isoenergetic surface becomes discontinuous, being separated into pieces by the zone boundary. This means that a part of the isoenergetic surface appears in the $(n + 1)$th zone but the rest remains in the $n$th zone, leaving unoccupied states holes. It can be easily seen in the one-dimensional case [see the example below in (**iv**)].

The formation of the energy gap can also be discussed from the point of view of the diffraction phenomena of the Bloch wave. For this let us recall the Bragg reflection. The quasimomentum $\gamma + t$ is said to satisfy the Laue condition or the Bragg condition if it belongs to the diffraction plane $D_\delta$ for some $\delta$, that is,

$$|\gamma + t|^2 = |\gamma + t + \delta|^2.$$

The Bloch wave changes its direction due to the Bragg reflection.

In the following one-dimensional example we demonstrate both how the interaction of the isoenergetic surface with the diffraction hyperplanes and how the diffraction phenomena of the Bloch waves result in the energy gap. Note that the band structure calculation of a real lattice is much more complicated and this example should be looked upon as a simple demonstration.

**(iv) One-dimensional Model**. Let $H(q)$ be the one-dimensional Schrödinger operator (named as Hill's operator) generated in $L_2(-\infty, \infty)$ by the expression

$$-y'' + q(x)y, \tag{1.3.2}$$

where $q$ is a real-valued function satisfying $q(x) = q(x+1)$. Note that there are a lot of books and papers about the one-dimensional case (see [DuSch, Eas, Le, MaVi, Mar, Na, Ti] and the references on them), where the spectrum of $H(q)$ is investigated and the asymptotic formulas for the eigenvalues $\lambda$ and the eigenfunctions $\Psi$ when $\lambda \to \infty$ were obtained by different methods. Here we consider the large Bloch eigenvalues and the corresponding Bloch functions of $H(q)$ by the approach which is useful for understanding the results of Chap. 2. Moreover, it helps to compare the well-known one-dimensional case with the multidimensional case and to see the complexity of the results obtained in this book.

For simplicity assume that

$$\sup_{x\in[0,1]} |q(x)| = M < \infty \quad \& \int_0^1 q(x)dx = 0. \tag{1.3.3}$$

Note that the first condition in (1.3.3) can be replaced by $q \in L_1[0,1]$ (see [VeDe, VeDu]) and the second condition is assumed without loss of generality. Thus the period lattice of the potential $q$ is $\mathbb{Z}$ and the reciprocal lattice is $2\pi\mathbb{Z}$. As we stressed above if the reciprocal lattice is $2\pi\mathbb{Z}$ then the diffraction planes are the points $\pi n$ for $n = \pm 1, \pm 2, \dots$, since the Bragg condition holds at them. We see below that this is indeed the wave number at which the energy gap appears. Moreover we see readily the cases when the plane wave $e^{i(2\pi n+t)x}$ is reflected and when it is not reflected by the crystals.

Let us recall some well-known results about $H(q)$ that we use for the discussion of this model. The spectrum $\sigma(H)$ of the operator $H(q)$ is the union of the spectra $\sigma(H_t)$ of the operators $H_t(q)$ for $t \in [0, 2\pi)$, which are generated in $L_2[0,1]$ by the expression (1.3.2) and the $t$-periodic boundary conditions

$$y(1) = e^{it}y(0),\ y'(1) = e^{it}y'(0).$$

In the case $q = 0$ the eigenvalues and eigenfunctions of $H_t(0)$ are respectively $(2\pi n + t)^2$ and $e^{i(2\pi n+t)x}$ for $n \in \mathbb{Z}$. All eigenvalues of $H_t(0)$ for $t \neq 0, \pi$ are

simple, while the eigenvalues of $H_0(0)$, except 0, and $H_\pi(0)$ are double. Since the eigenvalues of $H_{-t}(q)$ coincide with those of $H_t(q)$, we discuss only the case $t \in [0, \pi]$. For simplicity let us investigate the case $t \in [0, \frac{\pi}{2}]$. (The case $t \in (\frac{\pi}{2}, \pi]$ can be considered in the same way). By well-known perturbation theory the eigenvalues

$$\lambda_0(t) \le \lambda_{-1}(t) \le \lambda_1(t) \le \lambda_{-2}(t) \le \lambda_2(t) \le \cdots \tag{1.3.4}$$

of $H_t(q)$ for $t \in [0, \frac{\pi}{2}]$ satisfy the inequalities

$$\left|\lambda_n(t) - (2\pi n + t)^2\right| \le M \tag{1.3.5}$$

for all $n \in \mathbb{Z}$ due to (1.3.3).

First let us give the rigorous mathematical proof of the asymptotic formulas and then discuss the band structure from the point of view of the physicists. To obtain the asymptotic formula for the eigenvalues $\lambda_n(t)$ and corresponding normalized eigenfunctions $\Psi_{n,t}(x)$ of $H_t(q)$, let us use the following relation

$$(\lambda_n(t) - (2\pi k + t)^2)(\Psi_{n,t}, e^{i(2\pi k+t)x}) = (q\Psi_{n,t}, e^{i(2\pi k+t)x}) \tag{1.3.6}$$

which can be obtained from the equation

$$-\Psi''_{n,t}(x) + q(x)\Psi_{n,t}(x) = \lambda_n(t)\Psi_{n,t}(x)$$

by multiplying $e^{-i(2\pi k+t)x}$ and integrating the resulting expression over $[0, 1]$ by parts, where $(\cdot, \cdot)$ denotes the inner product in $L_2[0, 1]$. By (1.3.3) and Schward's inequality we have

$$\left|(q\Psi_{n,t}, e^{i(2\pi k+t)x})\right| \le M. \tag{1.3.7}$$

If $t \in [0, \frac{\pi}{2}]$ then $\left|(2\pi n + t)^2 - (2\pi k + t)^2\right| \ge 2\pi(|n - k|)\,|(2\pi\,|n + k| - \pi)|$ for $k \ne \pm n$. This with (1.3.5) gives us

$$\left|\lambda_n(t) - (2\pi k + t)^2\right| > 3\pi^2\,|(n - k)(n + k)| \tag{1.3.8}$$

for $k \ne \pm n$ and for the large values of $n$.

It follows from (1.3.6)–(1.3.8) that

$$\sum_{k\in\mathbb{Z}, k\ne\pm n} \left|(\Psi_{n,t}(x), e^{i(2\pi k+t)x})\right|^2 = \sum_{k\in\mathbb{Z}, k\ne\pm n} \frac{M^2}{(3\pi^2(n-k)(n+k))^2} = O(\frac{1}{n^2}).$$

Hence

$$\left\| \sum_{k\in\mathbb{Z}, k\ne\pm n} (\Psi_{n,t}(x), e^{i(2\pi k+t)x}) e^{i(2\pi k+t)x} \right\| = O(\frac{1}{n}).$$

Therefore the expansion of $\Psi_{n,t}(x)$ by the orthonormal basis $\{e^{i(2\pi n+t)x} : n \in \mathbb{Z}\}$ has the form

$$\Psi_{n,t}(x) = u_n(t)e^{i(2\pi n+t)x} + v_n(t)e^{i(-2\pi n+t)x} + O(n^{-1}), \tag{1.3.9}$$

where $u_n(t) = (\Psi_{n,t}, e^{i(2\pi n+t)x})$, $v_n(t) = (\Psi_{n,t}, e^{i(-2\pi n+t)x})$,

$$|u_n(t)|^2 + |v_n(t)|^2 = 1 + O(n^{-2}). \tag{1.3.10}$$

Now we consider the following two cases. First let us consider the case when the quasimomentum $2\pi n + t$ is far from the diffraction points $\pi k$, that is, there exists a positive constant $c \ll 1$ such that $t \in [c, \frac{\pi}{2}]$. Then

$$\left|(2\pi n + t)^2 - (-2\pi n + t)^2\right| \geq 8\pi\,|n|\,c.$$

Therefore using (1.3.5) and (1.3.6) for $k = -n$ we obtain

$$\left|\lambda_n(t) - (-2\pi n + t)^2\right| \geq 8\pi\,|n|\,c - M$$

and

$$(\Psi_{n,t}(x), e^{i(-2\pi n+t)x}) = O(n^{-1})$$

This with (1.3.9) and (1.3.10) implies that

$$\Psi_{n,t}(x) = e^{i(2\pi n+t)x} + O(\frac{1}{n}) \tag{1.3.11}$$

for $t \in [c, \frac{\pi}{2}]$.

Now using (1.3.11) in (1.3.6), letting $k = n$ and taking into account the second relation of (1.3.3) we obtain that

$$\lambda_n(t) = (2\pi n + t)^2 + O\left(\frac{1}{n}\right). \tag{1.3.12}$$

Now let us consider the case $t \in [0, c]$, that is, the case when the quasimomentum $2\pi n + t$ is close the diffraction point $2\pi n$. In the case $t = 0$ the eigenvalues $(2\pi n)^2$ for $n \neq 0$ of the unperturbed operator $H_0(0)$ are double and the corresponding eigenfunctions are the linear combinations of $e^{i2\pi nx}$ and $e^{-i2\pi nx}$. All eigenvalues of $H_t(0)$ for $t \neq 0, \pi$ are simple. However if $t$ is very close to 0 then the eigenvalues $(2\pi n + t)^2$ and $(-2\pi n + t)^2$ are close to each other.

Let us consider the case $t = 0$. Since the eigenvalues $(2\pi n)^2$ for $n \neq 0$ of the unperturbed operator $H_0(0)$ are double, by (1.3.4) and (1.3.5) the perturbed operator $H_0(q)$ has two eigenvalues (counting multiplicity) denoted by $\lambda_n =: \lambda_n(0)$ and $\lambda_{-n} =: \lambda_{-n}(0)$ such that

$$\lambda_{-n} \le \lambda_n, \left|\lambda_{\pm n} - (2\pi n)^2\right| \le M.$$

First let us prove that the eigenvalues $\lambda_n$ and $\lambda_{-n}$ are simple if

$$|nq_{2n}|^{-1} = o(1), \tag{1.3.13}$$

that is, $|q_{2n}| \gg \frac{1}{n}$, where

$$q_{2n} = \left(q, e^{i4\pi nx}\right) = \int_0^1 q\,(x)\, e^{-i4\pi nx} dx.$$

Suppose to the contrary that $\lambda_n$ is a double eigenvalue, that is, $\lambda_n = \lambda_{-n}$. Then by (1.3.9) and (1.3.10) the corresponding eigenspace is close to the span of the plane waves $e^{i2\pi nx}$ and $e^{-i2\pi nx}$, and there exists an eigenfunction of the form $e^{-2\pi nx} + O\left(n^{-1}\right)$. Using this eigenfunction instead of $\Psi_{n,0}(x)$ in the formula

$$(\lambda_n - (2\pi n)^2)(\Psi_{n,0}, e^{i2\pi nx}) = (q\Psi_{n,0}, e^{i2\pi nx}), \tag{1.3.14}$$

obtained from (1.3.6) by taking $t = 0$ and $k = n$, we get $O\left(n^{-1}\right) = q_{2n} + O\left(n^{-1}\right)$ which contradicts (1.3.13). Thus the eigenvalues $\lambda_n$ and $\lambda_{-n}$ are simple for large values of $n$ if (1.3.13) holds.

Now, for simplicity, let us consider the case when $q$ is an even function. Then

$$q_{2n} = \int_0^1 q(x) \cos 4\pi nx dx \in \mathbb{R} \tag{1.3.15}$$

and without loss of generality it can be assumed that $q_{2n} > 0$. Moreover in the case of even potential $q$, it is well-known that (see [Eas, MaVi]) the periodic solutions and hence the eigenfunction $\Psi_n(x) =: \Psi_{n,0}(x)$ is either even or odd function. Therefore, by (1.3.9) either $v_n = u_n + O(n^{-1})$ or $v_n = -u_n + O(n^{-1})$, where

$$u_n = (\Psi_n, e^{i2\pi nx}), v_n = (\Psi_n, e^{-i2\pi nx}).$$

In the first case from (1.3.9) and (1.3.10) one can easily obtain that

$$\Psi_n(x) = u_n e^{i2\pi nx} + u_n e^{-i2\pi nx} + O(n^{-1}) = \sqrt{2}\cos 2\pi nx + O(n^{-1}). \tag{1.3.16}$$

Using this and taking into account that $(\Psi_n, \Psi_{-n}) = 0$,where $\Psi_{-n} =: \Psi_{-n,0}(x)$, we obtain

$$\Psi_{-n}(x) = u_n e^{i2\pi nx} - u_n e^{-i2\pi nx} + O(n^{-1}) = \sqrt{2}\sin 2\pi nx + O(n^{-1}). \tag{1.3.17}$$

Now using (1.3.16) and (1.3.17) in (1.3.14) and taking into account that $\lambda_n - (2\pi n)^2$ is a real number,

$$(\cos 2\pi nx)^2 = \frac{1}{2}(1 + \cos 4\pi nx),\ (\sin 2\pi nx)^2 = \frac{1}{2}(1 - \cos 4\pi nx),$$

and then using (1.3.15) we get

$$\lambda_n = (2\pi n)^2 + q_{2n} + O\left(n^{-1}\right) \tag{1.3.18}$$

and

$$\lambda_{-n} = (2\pi n)^2 - q_{2n} + O\left(n^{-1}\right) \tag{1.3.19}$$

respectively. Note that the condition $\lambda_{-n} \leq \lambda_n$ and formulas (1.3.18) and (1.3.19) show that we have to take $v_n = u_n + O(n^{-1})$ if $q_{2n} > 0$ and therefore (1.3.16) and (1.3.17) hold.

In the same way one can show that the eigenvalues

$$\mu_{-1} \leq \mu_1 \leq \mu_{-2} \leq \mu_2 \leq \cdots$$

of $H_\pi(q)$ and the corresponding eigenfunction $\Phi_{-1}, \Phi_1, \Phi_{-2}, \Phi_2, \ldots$ satisfy the following asymptotic formulas

$$\mu_n = (2n\pi - \pi)^2 + q_{2n-1} + O\left(n^{-1}\right) \tag{1.3.20}$$

$$\mu_{-n} = (2\pi n - \pi)^2 - q_{2n-1} + O\left(n^{-1}\right) \tag{1.3.21}$$

and

$$\Phi_n(x) = \sqrt{2}\cos(2\pi n - \pi)x + O(\frac{1}{n}). \tag{1.3.22}$$

$$\Phi_{-n}(x) = \sqrt{2}\sin(2\pi n - \pi)x + O(\frac{1}{n}). \tag{1.3.23}$$

It is well-known that [Eas, MaVi, Ti] the spectrum of $H(q)$ consists of the intervals

$$[\lambda_0, \mu_{-1}], [\mu_1, \lambda_{-1}], [\lambda_1, \mu_{-2}], [\mu_2, \lambda_{-2}], \ldots, [\lambda_{j-1}, \mu_{-j}], [\mu_j, \lambda_{-j-1}], \tag{1.3.24}$$

where $j = 3, 4 \ldots$, that are the energy bands. Therefore the gaps in the spectrum (energy gaps) of the Hill's operator $H(q)$ consist of the intervals

$$\Delta_1 = (\mu_{-1}, \mu_1), \Delta_2 = (\lambda_{-1}, \lambda_1), \ldots, \Delta_{2j-1} = (\mu_{-j}, \mu_j), \Delta_{2j} = (\lambda_{-j}, \lambda_j), \tag{1.3.25}$$

where $j = 2, 3, \ldots$, that are the forbidden zones. Then (1.3.18)–(1.3.21) imply that the length $|\Delta_n|$ of the $n$th forbidden zone $\Delta_n$ (gap of the spectrum) satisfies the asymptotic formula

$$|\Delta_n| = 2\,|q_n| + O\left(\frac{1}{n}\right). \tag{1.3.26}$$

From the point of view of mathematicians the gaps arise as follows. For any real periodic potential $q$ the spectrum of the Hill's operator $H(q)$ consists of the intervals (1.3.24). The ends of the intervals are periodic and antiperiodic eigenvalues. In the case of unperturbed operator $H(0)$ these intervals are

$$[0, \pi^2], [\pi^2, (2\pi)^2], \ldots, [(2n\pi)^2, ((2n+1)\pi)^2], [((2n+1)\pi)^2, ((2n+2)\pi)^2] \tag{1.3.27}$$

for $n = 1, 2, \ldots$ The right end of the $n$th band coincides with the left end of the $(n+1)$th band and these ends are the double eigenvalues $(n\pi)^2$ of periodic (if $n$ is an even number) or antiperiodic (if $n$ is an odd number). Under the perturbation $q$ these double eigenvalues (double eigenvalue can be considered as two coinciding eigenvalues) are separated and one eigenvalue goes to the left and becomes the right end $\lambda_{-j}$ of the $n$th band (if $n = 2j$) of the perturbed operator $H(q)$ and the other eigenvalue goes to the right and becomes the left end $\lambda_j$ of the $(n+1)$th band of $H(q)$. The space $\Delta_{2j}$ between these ends $\lambda_{-j}$ and $\lambda_j$ can not be occupied by the Bloch eigenvalues $\lambda_{-j}(t)$ and $\lambda_j(t)$, since for $t \in [0, c]$, where $c \ll 1$, the eigenvalues $\lambda_{-j}(t)$ and $\lambda_j(t)$ together with $\lambda_{-j}(0) =: \lambda_{-j}$ and $\lambda_j(t) =: \lambda_j$ go to the left and right respectively and hence arise gaps in the spectrum.

Now we summarize the discussed statements about the one-dimensional Schr ödinger operator $H(q)$ with a periodic potential $q$ in the language of physicists. In the above example, we rigorously constructed the Bloch waves in the high energy region by asymptotic method that is very similar to the two-wave approximation. As we noted above the Bragg condition is satisfied at $\pm\pi n$, since the reciprocal lattice is $2\pi\mathbb{Z}$. The isoenergetic surfaces $I_0((\pi n)^2)$ corresponding to the energy $(\pi n)^2$ consist of two points $-\pi n$ and $\pi n$ and these points are the diffraction planes of the reciprocal lattice. Under the perturbations the isoenergetic surfaces are separated into pieces by the zone boundary and part of the isoenergetic surface appears in the $(n+1)$th zone but the rest remains in the $n$th zone, leaving unoccupied states holes.

Formula (1.3.11) means that the plane wave $e^{i(2\pi n + t)x}$ is almost not reflected by the crystals if the wave number $2\pi n + t$ is far from the diffraction planes $\pi n$. Formulas (1.3.16) and (1.3.17) show that under perturbation $q$ the plane waves $e^{i2\pi nx}$ and $e^{-i2\pi nx}$ interface each other. The standing waves $\sqrt{2}\cos 2\pi nx$ and $\sqrt{2}\sin 2\pi nx$ are the results of the interference between two waves $e^{i2\pi nx}$ and $e^{-i2\pi nx}$ traveling in the opposite directions. On the other hand, it is well-known that the eigenvalues of $H_t(q)$ for $t \neq 0, \pi$ are simple. Therefore if $\lambda_n(0)$ is a simple eigenvalue, then $\Psi_{n,t}(x)$ continuously depend on $t \in [0, \pi)$. This situation with (1.3.16) and (1.3.17) shows that if $t$ is close to zero then under perturbation $q$ the plane waves $e^{i(2\pi n + t)x}$ and $e^{-i(2\pi n + t)x}$ interface each other. Moreover, these situations with (1.3.22) and (1.3.23)

show the same result when $t$ is close to $\pi$. Thus the electrons in the crystal are arranged in the energy bands separated by the forbidden regions, called energy gaps or band gaps, in the energy for which no wavelike electron orbitals exist. The band gap is a result of the interference between two waves traveling in the opposite directions. The plane wavefunction $e^{i(2\pi n+t)x}$ represents the running wave and carries the momentum $k = 2\pi n + t$. If $t \neq 0, \pi$ then this wave function is the travelling wave. However, the wave function at $t = 0$ is not wave $e^{i2\pi nx}$ or $e^{-i2\pi nx}$ travelling to the right or left, respectively. Namely when the Bragg reflection condition $t = 0$ is satisfied by the wave vector $2\pi n + t$ a wave travelling to the right is Bragg-reflected to travel to the left and vice versa. As a result the standing waves $\sqrt{2}\cos 2\pi x$ and $\sqrt{2}\sin 2\pi x$ are obtained from the travelling waves $e^{i2\pi nx}$ and $e^{-i2\pi nx}$. The two standing waves $\sqrt{2}\cos \pi x$ and $\sqrt{2}\sin \pi x$ pile up the electrons at the different regions. Therefore the two waves have different values of the potential energy which is the origin of the energy gap. It is well-known and we can see from the above example that the magnitude of the energy gap depends on the Fourier coefficients of the periodic potential. Thus the effect of the periodic potential is to produce an energy gap in the band structure of the one-dimensional case and the energy gap appears when the Bragg condition is satisfied at $\pm\pi n$. In other words, when the wave vector is near to these diffraction planes the Bloch wave is expressed by a linear combination of the unperturbed plane waves $e^{i(2\pi n+t)x}$ and $e^{-i(2\pi n+t)x}$ perturbed by the lattice planes. The running wave $-\pi n$ is reflected to the wave $\pi n$ by receiving the crystal momentum $2\pi n$ from the lattice planes and the reflected wave $-\pi n$ is again reflected to the wave $\pi n$ by receiving the crystal momentum $2\pi n$ from the lattice planes. This process is infinitely repeated, resulting in a cosine- or sine-type stationary wave. Under the above condition on the potential, the energy of the sine-type Bloch wave is lowered and the energy of the cosine-type Bloch wave is raised. Thus, the difference in the energy between these two stationary states must be responsible for the formation of the energy gap.

## 1.4 Some Discussions of the Perturbation Theory

In this section we discuss the perturbation theory and isoenergetic surfaces for the multidimensional Schrödinger operator $L(q)$ in the high energy region. This case, for the first time, was investigated in the papers [Ve1, Ve2, Ve3, Ve4, Ve5, Ve6]. In Chap. 2 we consider it in detail. Now we only describe briefly the crucial points and complexity of this theory. For this, first let us recall that, in general, the perturbation theory is easy if the potential $q$ is smaller than the distance between the eigenvalues of the unperturbed operator $L(0)$. In other words, as well-known from the quantum mechanics, if the perturbation is small compared to the energy difference between the states, then we can use the regular perturbation theory to calculate the wave functions and energy levels. The perturbation theory breaks down, however, in those cases when the potential cannot be considered as a small perturbation. This happens when the magnitude of the potential becomes comparable with the energy separation. To be

more precise let us define a constant $h$ for the energy separation, named as the energy separation constant, as follows. One can readily see from the **One-dimensionel model** (see Sect. 1.3) that there are two cases:

Case 1. **Isolated eigenvalue**. An eigenvalue $\lambda$ is isolated if all other eigenvalues are far from $\lambda$ (see the case $t \in [c, \frac{\pi}{2}]$). Then the energy separation constant $h$ is a distance from $\lambda$ to the set of all other eigenvalues.

Case 2. **Isolated pair of eigenvalues**. If the two eigenvalues $\lambda_1$ and $\lambda_2$ are close to each other and the others are far from these eigenvalues (see the case $t \in [0, c)$), then the energy separation constant $h$ is a distance from the set $\{\lambda_1, \lambda_2\}$ to the set of all other eigenvalues.

If $\|q\| \ll h$ then the perturbation theory is easy and well-known, since in Case 1 and Case 2 one can use the regular perturbation theory and two wave approximations, respectively, where the relation $h \gg 1$ means that $h$ is a sufficiently large number. The inequality $\|q\| \ll h$ which easifies the perturbation theory occurs in the following two cases:

**First case**: The perturbation $q$ is bounded or $\|q\| = O(1)$ and the energy separation constant $h$ tends to infinity as the eigenvalues go to infinity. This case is the one-dimensional case in the high energy region and we demonstrated it in the **One-dimensionel model** (see Sect. 1.3) and noted that this case was investigated very well, there are a lot of books and papers about it.

**Second case**: The energy separation constant $h$ is greater than some constant and the potential $q$ is replaced by $\varepsilon q$, where $\varepsilon$ is a small parameter, that is, $\|\varepsilon q\| \ll h$. This case can be used for the small eigenvalues of the multidimensional operator $L(\varepsilon q)$ to obtain the formulas for $\varepsilon \to 0$. Indeed if the eigenvalue $|\gamma + t|^2$ has a distance greater than some constant from the other eigenvalues then the small perturbation $\varepsilon q$ can be investigated by the regular perturbation theory. Moreover if $|\gamma + t|^2$ coincides with (or it is near to) the eigenvalue $|\gamma + t + \delta|^2$ but has a distance greater than some constant from the other eigenvalues, that is, if $\gamma + t$ lies in (or it is near to) only one Bragg plane $D_\delta$, then a weak periodic potential $\varepsilon q$ has its major effect on those free electron levels whose wave vectors are close to ones at which the Bragg reflection can occur. In this case, in order to find the energy levels and the wave functions one can use, for example, the two wave approximations. We will discuss this case in detail in Chap. 5.

Thus in the **first and second case**, we can use the regular perturbation theory to calculate the wave functions and energy levels.

Now we are ready to discuss the multidimensional operator $L(q)$ in the high energy region. In this case, in the big contrary of the **first and second case** (see above) we meet with the situation $h \ll \|q\|$ instead of $\|q\| \ll h$, since the denseness of the Bloch eigenvalues of the free operator increases infinitely with the increasing energy and hence the distance between the eigenvalues tends to zero or the multiplicity of the eigenvalues tend to infinity. To describe this case more precisely, let us introduce some notations. The relation $a(\rho) \sim b(\rho)$ as $\rho \to \infty$ means that $a(\rho) = O(b(\rho))$ and $b(\rho) = O(a(\rho))$, that is, there exist constants $c_1$ and $c_2$ such that

$$c_1 b(\rho) < a(\rho) < c_2 b(\rho).$$

In this case we say that $a(\rho)$ is of order $b(\rho)$. Let $E(\rho)$ be the number of the Bloch eigenvalues (counting multiplicity) of the unperturbed operator $L_t(0)$ lying in the interval $[\rho^2, \rho^2 + 1)$. The number $E(\rho)$ depends on $t \in F^*$, however, in average, $E(\rho) \sim \rho^{d-2}$, since $|\gamma + t|^2 \in [\rho^2, \rho^2 + 1)$ if and only if

$$\gamma + t \in \{x \in \mathbb{R}^d : \rho^2 \le |x|^2 < \rho^2 + 1\} =: W(\rho). \tag{1.4.1}$$

On the other hand, the spherical washer $W(\rho)$ is filled with the translations of $F^*$ by the vectors $\gamma$ of the reciprocal lattice $\Gamma$, and

$$\mu(W(\rho)) \sim \rho^{d-2}\mu(F^*),$$

where $\mu(A)$ denotes the volume of the set $A$. Thus in the interval $[\rho^2, \rho^2+1)$ of length 1 there are, in average, $E(\rho)$ Bloch eigenvalues $|\gamma + t|^2$ of the free operator, where $E(\rho) \sim \rho^{d-2}$. It means that the eigenvalues are densely situated in the high energy region $[\rho^2, \rho^2 + 1)$ and for the energy separation constant $h(\rho)$ (now it depends on $\rho$) one can write the equality

$$h(\rho) = O(\rho^{2-d}). \tag{1.4.2}$$

Hence in the multidimensional case in the high energy region the bounded potential $q$ cannot be considered as a small perturbation, since

$$\|q\| \sim \rho^{d-2}h(\rho) \gg h(\rho) \tag{1.4.3}$$

for $d > 2$ and $\rho \gg 1$. Therefore the regular perturbation theory is ineffective in this case. In Chap. 2 we consider this case in detail. Now we only describe briefly the following three problems **(a), (b) and (c)** which are the crucial and remarkable points of the perturbation theory of the multidimensional operator $L(q)$ in the high energy region.

**(a) Simplicity problem**. *Determine the set of quasimomenta $\gamma + t$ such that the corresponding Bloch eigenvalues $\Lambda(\gamma + t) \in [\rho^2, \rho^2 + 1)$ of $L_t(q)$ are simple.*

The complexity of this problem is the following. The eigenvalue $\Lambda(\gamma + t) \in [\rho^2, \rho^2 + 1)$ is a result of moving of the Bloch eigenvalues $|\gamma + t|^2$ of the free electron under the perturbation $q$. In the interval $[\rho^2, \rho^2 + 1)$ of length 1 there are, in average, $E(\rho)$ Bloch eigenvalues $|\widetilde{\gamma} + t|^2$ of $L_t(0)$, where $\widetilde{\gamma} \in \Gamma$ and $E(\rho) \sim \rho^{d-2}$. After the periodic perturbation $q$ all these eigenvalues move and some of them move of order 1 and hence each of the resulting eigenvalues $\Lambda(\widetilde{\gamma} + t)$ of $L_t(q)$ may coincide with $\Lambda(\gamma + t)$. Thus we need to control the moving of all eigenvalues $|\widetilde{\gamma} + t|^2 \in [\rho^2, \rho^2 + 1)$ for some values of $t$ in order that all resulting eigenvalues $\Lambda(\widetilde{\gamma} + t)$ do not coincide with $\Lambda(\gamma + t)$ and hence $\Lambda(\gamma + t)$ becomes a simple eigenvalue. Therefore it seems that it is impossible to find the values of the quasimomenta $\gamma + t$ for which the corresponding Bloch eigenvalues $\Lambda(\gamma + t)$ of $L_t(q)$ are simple. The importance of the simplicity of $\Lambda(\gamma + t)$ is the following. The simplicity of $\Lambda(\gamma + t)$ is necessary for the investigation of the corresponding Bloch wave $\Psi_{\gamma+t}(x)$ and for proving that it is close to the plane wave $e^{i\langle\gamma+t,x\rangle}$ that is, satisfies the formula

$$\Psi_{\gamma+t}(x) = e^{i\langle\gamma+t,x\rangle} + O(|\gamma+t|^{-\alpha}), \tag{1.4.4}$$

where $\alpha > 0$. The last equality means that the plane wave $e^{i\langle\gamma+t,x\rangle}$ goes through the crystal almost without the diffraction. On the other hand, it is well known that the plane wave $e^{i\langle\gamma+t,x\rangle}$ is reflected by the crystal if $\gamma + t$ belongs to (or it is near to) a diffraction hyperplane $D_\delta$ for some $\delta \in \Gamma$. Then the reflected wave $e^{i\langle\gamma+\delta+t,x\rangle}$ interferes with the initial wave $e^{i\langle\gamma+t,x\rangle}$ (see [BS], [Ki, Mad]) and (1.4.4) does not hold. As we noted above there are, in average, $E(\rho)$ eigenvalues

$$|\gamma+t|^2, |\gamma+t+\delta_1|^2, |\gamma+t+\delta_2|^2, \ldots, |\gamma+t+\delta_n|^2,$$

where $n = E(\rho) \sim \rho^{d-2}$, lying in the interval $[\rho^2, \rho^2+1)$. On the other hand, by choosing the coordinate axis so that the direction of $\delta$ coincides with the direction of $(1, 0, 0, \ldots, 0)$, we can easily verify that if

$$|\gamma+t|^2 - |\gamma+t+\delta|^2 = c$$

then the quasimomentum $\gamma + t$ lies on the distance $\frac{|c|}{|\delta|}$ from the diffraction plane $D_\delta$. Therefore all the diffraction planes $D_{\delta_1}, D_{\delta_2}, \ldots, D_{\delta_n}$, may reflect the wave $e^{i\langle\gamma+t,x\rangle}$ with the fixed quasimomentum $t$. If we do not fix $t$, then all diffraction planes passing through the washer $W(\rho)$ may reflect the wave $e^{i\langle\gamma+t,x\rangle}$ if the corresponding eigenvalue $|\gamma+t|^2$ lies in the interval $[\rho^2, \rho^2+1)$. On the other hand, the number $D(\rho)$ of the diffraction planes having nonempty intersection with the sphere

$$S(\rho) = \{x \in \mathbb{R}^d : |x| = \rho\}$$

and hence with $W(\rho)$ is of order $\rho^d$, that is, $D(\rho) \sim \rho^d$. Thus the second problem is the following.

**(b) Bragg Reflection Problem**. *Determine the set of quasimomenta $\gamma+t \in W(\rho)$ for which the plane wave $e^{i\langle\gamma+t,x\rangle}$ under the periodic perturbationq goes through the crystal without the essential influence of the $D(\rho)$ diffraction hyperplanes, where $D(\rho) \sim \rho^d$.*

That is why the mathematical difficulties of the perturbation theory of the multidimensional operator $L(q)$ in the high energy region have a physical nature—a complicated picture of diffraction inside the crystal.

As we explained above in one-dimensional case it is very easy to explain the arising of the gaps in the spectrum. Briefly speaking, there are only two Bloch eigenvalues $(-n\pi)^2$ and $(n\pi)^2$ of the free operator lying at the point $\lambda = (n\pi)^2$ and the isoenergetic surface $I_0((n\pi)^2)$ consists only of the two points $-n\pi$ and $n\pi$ which are the diffraction planes. Under the perturbation $q$ one eigenvalue goes to the left and one to the right and the gap in the neighborhood of $(n\pi)^2$ emerges as a result of these movings.

In the big contrary of the one-dimensional case, in the multidimensional case the set of all Bloch eigenvalues $|\gamma+t|^2$ of the unperturbed operator $L(0)$ lying at the same point $\rho^2$ as much as the points of the sphere $S(\rho)$, since $|\gamma+t|^2 = \rho^2$ if and only

if $(\gamma+t) \in S(\rho)$. Some of these eigenvalues $|\gamma+t|^2$ are multiple. Recall that $|\gamma+t|^2$ is multiple if $\gamma+t$ lies in the intersection of the sphere $S(\rho)$ and diffraction planes and the all other eigenvalues are simple. If the sphere is large, then after the perturbation $q$ the probability that all these eigenvalues go away from the point $\rho^2$ and the other Bloch eigenvalues do not come to this point and hence the isoenergetic surface $I_q(\rho^2)$ becomes an empty set is very small. (Hence the probability of the validity of the Bethe-Sommerfeld conjecture is close to 1). However as we noted above there are the $D(\rho)$ diffraction planes intersecting $S(\rho)$ for large $\rho$, where $D(\rho) \sim \rho^d$, and the isoenergetic surface begins to be distorted from a sphere before making contacts with the diffraction planes. Thus the isoenergetic surface is divided into a lot of pieces. Therefore the rigorous mathematical investigation of the perturbations of all these eigenvalues and to prove that the isoenergetic surface $I_q(\rho^2)$ can not become an empty set are extremely complicated. Thus the third problem is the following:

**(c) Isoenergetic Surfaces Problem**. *Determine the shape and measure of the isoenergetic surface $I_q(\rho^2)$ of $L(q)$ which emerges from the isoenergetic sphere $S(\rho)$ of $L(0)$ as a result of its distortion and separation into very small pieces by the $D(\rho)$ diffraction planes intersecting $S(\rho)$, where $D(\rho) \sim \rho^d$.*

To answer all these **three problems (a), (b) and (c),** in Chap. 2 we develop a new mathematical approach to this problem. The momentum space is divided into two domains: $U$ (non-resonance domain) and $V$ (resonance domain) and the eigenvalues $|\gamma+t|^2$, for large $\gamma \in \Gamma$, are divided into two groups: non-resonance ones if $\gamma+t \in U$ and resonance ones if $\gamma+t \in V$ and various asymptotic formulae are obtained for the perturbations of each groups. (The precise definitions of $U$ and $V$ are given in the introduction of Chap. 2). For the first time in the papers [Ve1, Ve2, Ve3, Ve4] we constructed the set $B \subset U$, called as a simple set, such that if $\gamma+t \in B$, then the corresponding Bloch eigenvalue $\Lambda(\gamma+t)$ is simple and satisfies

$$\Lambda(\gamma+t) = |\gamma+t|^2 + O(|\gamma+t|^{-\alpha}),$$

where $\alpha > 0$ and the Bloch function $\Psi_{\gamma+t}(x)$, corresponding to the eigenvalue $\Lambda(\gamma+t)$ satisfies (1.4.4). Moreover we proved that the simple set $B$ has the asymptotically full measure on $\mathbb{R}^d$ and constructed a part of the isoenergetic surface $I_q(\rho^2) \subset B$ for large $\rho$ which is a union of the smooth surfaces and has the measure asymptotically close to the measure of the sphere $S(\rho)$. Thus, we constructed the set $B \subset U$ that positively solves all the problems **(a), (b) and (c)** described above. Therefore the main difficulty and the crucial point of the investigations of the Bloch functions and isoenergetic surfaces and hence of the perturbation theory of $L(q)$ is the construction and estimation of the set $B$. We discuss it in detail in the introduction of Chap. 2. Note that, in Chap. 2, we construct the simple set in the non-resonance domain $U$ so that it contains a big part of the isoenergetic surfaces of $L(q)$. However in the case of the resonance domain $V$ we construct the simple set so that it can be easily used for the constructive determination (in Chap. 3) a family of the spectral invariants by the given Bloch eigenvalues. Then in Chap. 4, we constructively determine the

potential $q$ by these spectral invariants. We will continue these discussions at the end (in Chap. 5) of this book after the construction a perturbation theory (Chap. 2) and its applications (Chaps. 3 and 4).

## References


[BeShu] Berezin, F.A., Shubin, M.A.: The Schrödinger Equation. Kluwer Academic Publishers, Dordrecht (1991)

[BS] Bethe, G., Sommerfeld, A.: Elektronentheorie der Metalle. Springer-Verlag, Berlin (1967)

[Eas] Eastham, M.S.P.: The Spectral Theory of Periodic Differential Equations. Scotting Academic Press, Edinburgh (1973)

[Ku] Kuchment, P.: Floquet Theory for Partial Differential Equations. Operator Theory: Advances and Applications. Birkhäuser Verlag, Basel (1993)

[ReSi] Reed, M., Simon, B.: Methods of Modern Mathematical Physics. Academic Press, New York (1987)

[SomBe] Sommerfeld, A., Bethe, H.: Elektronentheorie der Metalle. Handbuch der Physik. 2nd. ed. Bd. XXIV/2. Springer, Berlin (1933)

[Ki] Kittel, Ch.: Introduction to Solid State Physics. Wiley, New York (1976)

[Mad] Madeling, O.: Introduction to Solid State Theory. Springer, New York (1978)

[Bl] Bloch, F.: Über die quantenmechanik der electronenin kristallgittern. Z. Phys. **52**, 555–600 (1928)

[Gel] Gelfand, I.M.: Expansions in series of eigenfunctions of an equation with periodic coefficients. Sov. Math. Dokl. **73**, 1117–1120 (1950)

[OdKe] Odeh, F., Keller, J.B.: Partial differential equation with periodic coefficients and Bloch waves in crystals. J. Math. Phys **5**, 1499–1504 (1964)

[Th] Thomas, L.E.: Time dependent approach to scattering from impurities in crystal. Commun. Math. Phys. **33**, 335–343 (1973)

[Wi] Wilcox, E.: Theory of Bloch waves. J. d'Analyse Mathematique **33**, 146–167 (1978)

[DuSch] Dunford, N., Schwartz, J.T.: Linear Operators, Part 3. Spectral Operators. Wiley-Interscience, New York (1988)

[Le] Levitan, B.M.: Inverse Sturm-Liouville Problems. VNU Science Press, Utrecht (1987)

[MaVi] Magnus, W., Winkler, S.: Hill's Equation. Interscience Publishers, New York (1966)

[Mar] Marchenko, V.A.: Sturm-Liouville Operators and Applications. Birkhauser Verlag, Basel (1986)

[Na] Naimark, M.A.: Linear Differential Operators. George G. Harap & Company, London (1967)

[Ti] Titchmarsh, E.C.: Eigenfunction Expansion (Part II). Oxford University Press, London (1958)

[VeDe] Dernek, N., Veliev, O.A.: On the Riesz basisness of the root functions of the non-self-adjoint Sturm-Liouville operator. Isr. J. Math. **145**, 113–123 (2005)

[VeDu] Veliev, O.A.: Toppamuk Duman, the spectral expansion for a nonself-adjoint Hill operators with a locally integrable potential. J. Math. Anal. Appl. **265**(1), 76–90 (2002)

[Ve1] Veliev, O.A.: On the spectrum of the Schrödinger operator with periodic potential. Sov. Math. Dokl. **27**(1), 234–237 (1983)

[Ve2] Veliev, O.A.: Asymptotic formulas for the eigenvalues of the multidimensional Schrödinger operator and periodic differential operators. Preprint Inst. Phys. AN Azerb. SSR no 156 (1985)

[Ve3] Veliev, O.A.: Asymptotic formulas for the eigenvalues of the periodic Schrödinger operator and the Bethe-Sommerfeld conjecture. Funct. Anal. Appl. **21**, 87 (1987)

[Ve4] Veliev, O.A.: The spectrum of multidimensional periodic operators (in Russian). Teor. Funkstiĭ Funktsional. Anal. i Prilozhen **49**, 17–34 [English trans. in J. Sov. Math. **49**, 1045–1058 (1990)] (1988)
[Ve5] Veliev, O.A.: Asymptotic formulas for Block function of multidimensional periodic Schrödinger operators and some of their application. J. Spectr. Theory Oper. Appl. (Elm, Baku) **9**, 59–76 (1989)
[Ve6] Veliev, O.A.: Asymptotic formulas for the eigenvalues and periodic differential operators. Doctor of Sciences Thesis. Razmadze Mathematical Institute, Tbilisi (1989)

http://www.springer.com/978-3-319-16642-1

Multidimensional Periodic Schrödinger Operator
Perturbation Theory and Applications
Veliev, O.
2015, X, 242 p., Hardcover
ISBN: 978-3-319-16642-1