that the reciprocal voltage amplification is

$$D = \frac{1}{e^{\frac{e}{kT}(U_c - U_e)} - 1} \approx e^{-\frac{e}{kT}(U_c - U_e)} \ll 1$$

This extremely small reciprocal voltage amplification is the reason for the strong similarity between the pentode and the *n-p-n* transistor characteristics. The similarity is, incidentally, not tied to the foregoing special case (see Fig. V.3.6) and results from the saturated character of the collector current.[1]

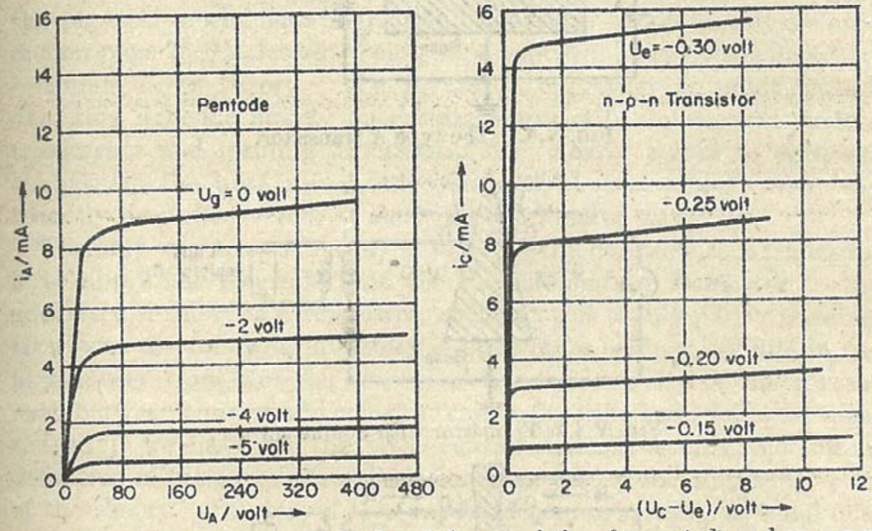

Fig. V.3.6. Comparison of the family of characteristics of a pentode and an *n-p-n* transistor.

§4. The Point-contact Transistor

a. Qualitative Treatment

Historically, the first transistor is the so-called type A transistor shown in Fig. V.4.1, whose discovery in 1948 by John Bardeen and W. H. Brattain[2] created, rightly, a sensation. The transistor with double surface of J. N. Shive[3] (see Fig. V.4.2) and the "coaxial tran-

---

[1] Concerning the comparison of the *n-p-n* transistor and the vacuum tube, see also L. J. Giacoletto, *Proc. IRE*, **40**: 1490 (1952). Figure V.3.6 is based on data from this paper.

[2] J. Bardeen and W. Brattain, *Phys. Rev.*, **74**: 230 and 231 (1948); **75**: 1208 (1949).

[3] J. N. Shive, *Phys. Rev.*, **75**: 689 (1949).

sistor" of W. E. Kock and R. L. Wallace[1] (see Fig. V.4.3) are, as was found shortly afterward, only geometric modifications of the type A transistor and fall within the general concept of the "point-contact transistor." The physically significant mechanism is the same for all these transistors. It consists in the fact that the boundary layer of a point contact, biased in the reverse direction, represents the current path which is influenced by carrier injection and that the emitter is also a point contact, biased, however, in the forward direction. A transistor of *n* germanium shall be considered for the qualitative description of the mechanism. A current I_e (e.g., 0.75 ma) enters the *n* germanium from a positively biased (e.g., + 0.15 volt) emitter. A small fraction $1 - \gamma$ of this current consists of electrons which come from the large area nonrectifying base and flow into the emitter after passing through the body of the germanium. The major part γI_e

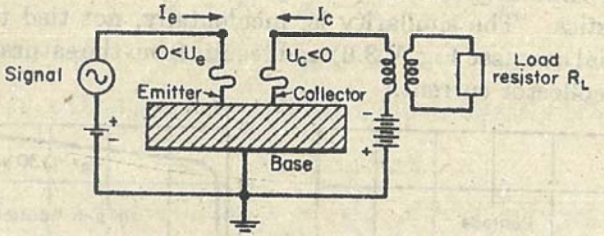

FIG. V.4.1. The type A transistor.

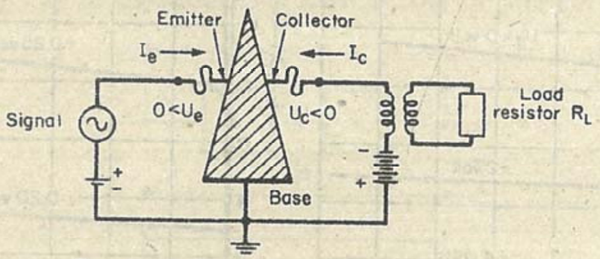

FIG. V.4.2. Transistor with double surface.

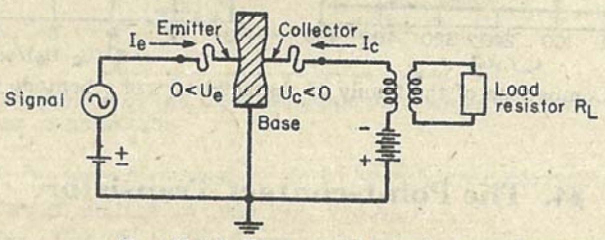

FIG. V.4.3. The coaxial transistor.

---

[1] W. E. Kock and R. L. Wallace, *Elec. Eng.*, **68**: 222 (1949).

of the emitter current, however, consists of holes which are naturally attracted by the negatively biased collector in the immediate vicinity, $U_c < 0$ (e.g., -20 volts). Hence the major portion β of the hole current γI_e is captured by the collector and thus modifies the conductance of the collector boundary layer which is depleted of carriers as a result of the reverse bias. In this manner the current path of the battery in the collector circuit, which includes the collector boundary layer, can be modulated with very little power in the emitter circuit.

An attempt to develop these concepts quantitatively fails because the physical mechanism of the point contacts—as previously pointed out on page 73—is described only very approximately by the Schottky boundary-layer theory. The favorable operation of point-contact detectors depends largely on certain empirically determined surface treatments and forming processes. This clearly points to chemical changes in the pure metal-semiconductor contact which have only recently been discovered, at least in a qualitative way.[1]

The only way to arrive at a "theory" of the point-contact transistor is to apply the concepts used for the filamentary transistor to the geometry of the type A transistor, although this is only partly possible. However, the following difficulty is even more serious. While in the filamentary transistor the control of the conductivity of the current path between emitter and collector can be described and followed theoretically,[2] we have for the point-contact transistor only the formal definition of the current-amplification factor α_i for this important part of the theory. In view of this unsatisfactory situation, we shall deal with this subject only briefly in the next section.

b. The Quantitative "Theory" of the Type A Transistor

We begin with a discussion of the emitter, and we consider first the properties of the emitter alone, ignoring the influence of the collector. For this purpose we must keep the collector in the zero-current condition and measure the emitter current-voltage characteristics:

$$I_e = f_{e0}(U_e) \tag{V.4.01}$$

With positive collector current I_c, the potential of the germanium surrounding the emitter would be $R_b \cdot I_c$ higher than before. According

---

[1] R. Thedieck, *Physik. Verhandl.*, **3:** 31 and 212 (1952). L. B. Valdes, *Proc. IRE*, **40:** 445 (1952).

[2] The pertinent theoretical concepts have also been experimentally verified. See W. Shockley, G. L. Pearson, and J. R. Haynes, *Bell System Tech. J.*, **288:** 344 (1949).

to Fig. V.4.4, R_b is the resistance between the equipotential surface through the emitter and the grounded base, assuming the current to spread out radially from the collector.

A positive collector current I_c creates in the germanium with the conductivity σ a hemispherical potential distribution

$$V(r) = \frac{I_c}{2\pi\sigma r} \tag{V.4.02}$$

With the collector as center point, the resulting field strength

$$E(r) = -V'(r) = +\frac{I_c}{2\pi\sigma r^2} \tag{V.4.03}$$

then produces with the conductivity σ exactly the required current I_c through each hemisphere $2\pi r^2$.

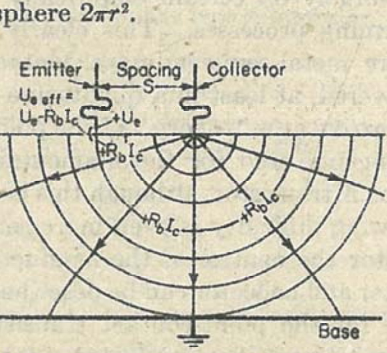

FIG. V.4.4. The origin of the effective emitter voltage $U_{e\,\mathrm{eff}} = U_e - R_b I_c$

If the emitter is at a distance S from the collector, the equipotential surface intersecting the emitter has the potential

$$V(S) = \frac{I_c}{2\pi\sigma S} \tag{V.4.04}$$

The base is identified with the infinitely distant hemisphere and therefore has the potential

$$V(\infty) = 0 \tag{V.4.05}$$

The voltage between the equipotential surface through the emitter and the base is, therefore,

$$V(S) - V(\infty) = \frac{I_c}{2\pi\sigma S} \tag{V.4.06}$$

and the resistance between the two surfaces in question is

$$R_b = \frac{V(S) - V(\infty)}{I_c} = \frac{1}{2\pi\sigma S} \qquad \text{(V.4.07)}$$

For $\sigma = 1/10$ ohm^{-1} cm^{-1} and $S = 50 \cdot 10^{-4}$ cm, by way of example, the resistance is

$$R_b = \frac{10 \text{ ohm cm}}{6.28 \cdot 5 \cdot 10^{-3} \text{ cm}} = 318 \text{ ohms} \qquad \text{(V.4.08)}$$

Inasmuch as with a collector current I_c the potential of the germanium near the emitter exceeds that for zero collector current by $R_b \cdot I_c$, it seems reasonable to regard $U_e - R_b \cdot I_c$ as the effective emitter voltage and to write in place of (V.4.01)

$$I_e = f_{e0}(U_e - R_b I_c) \qquad \text{(V.4.09)}$$

We have here assumed, it is true, that potentials created by the emitter current and the collector current can be superposed without interaction. This assumption is certainly not strictly correct because the conductivity of the germanium is modified by the injected holes. Taking this effect into account would lead to appreciable complications and will not be attempted here.[1]

After having treated the conditions on the emitter side, though in a somewhat incomplete manner, and having obtained Eq. (V.4.09) for the emitter current I_e, we shall now consider the collector current I_c. Without the influence of the emitter, I_c is given by the unmodified collector characteristic

$$I_c = f_{c0}(U_c) \quad \text{for } I_e = 0 \qquad \text{(V.4.10)}$$

In addition to this component, arising directly from the collector voltage U_c, a further component is introduced by the injection from the emitter. A fraction γ of the emitter current I_e consists of holes, a fraction β of which, in turn, is captured by the collector. The hole current $\beta\gamma I_e$ arriving at the collector releases a collector-current component $\alpha_i\beta\gamma I_e = \alpha I_e$ which is larger by the true current-amplification factor α_i and which is added to (V.4.10) though with negative sign, since according to our definition a positive emitter current traverses the collector circuit in a negative direction:

$$I_c = f_{c0}(U_c) - \alpha I_e \qquad \text{(V.4.11)}$$

Equations (V.4.09) and (V.4.11) are the two equations which relate

---

[1] L. B. Valdes, *Proc. IRE*, **40**: 1429 (1952), has also neglected this effect. Valdes considers the finite thickness of the Ge wafer in his calculation, which leads to a deformation of the hemispherical propagation of the collector current, in contrast to the above treatment.

the currents I_e and I_c and the voltages U_e and U_c. If we consider small deviations i_e, i_c, u_e, u_c from an operating point I_e, I_c, U_e, U_c, we obtain from (V.4.09) and (V.4.11) by linearization

$$i_e = \frac{1}{r_e}(u_e - R_b i_c)$$

$$i_c = \frac{1}{r_c}u_c - \alpha i_e$$

and

$$u_e = r_e i_e + R_b i_c \tag{V.4.12}$$

$$u_c = \alpha r_c i_e + r_c i_c \tag{V.4.13}$$

Here r_e and r_c are the differential resistances which can be calculated from the unperturbed characteristics (V.4.01) and (V.4.10) for the operating points I_e and I_c, respectively. R_b is given approximately by

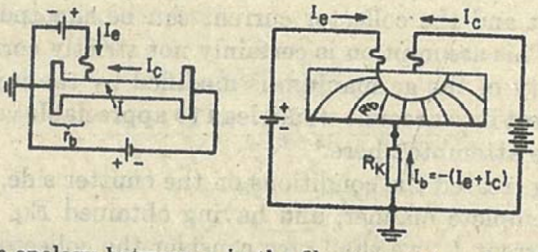

FIG. V.4.5. The analogy between r_b in the filamentary transistor and the *external* coupling resistor R_K in the point-contact transistor. The *internal* base resistance R_b does not have an analogue in the filamentary transistor.

Eq. (V.4.07). The current-amplification factor α of the point-contact transistor is formally expressed by the product

$$\alpha = \alpha_i \beta \gamma \tag{V.4.14}$$

but it is not possible to relate any of the three factors α_i, β, and γ to the geometrical dimensions and the material constants of the device.[1] Thus the theory of the point-contact transistor has a rather formal character compared with the theories of the *n-p-n* transistor and particularly the filamentary transistor.

We shall add a few remarks concerning the obvious comparison with the current-voltage characteristics of the filamentary transistor:

$$u_e = (r_b + r_e)i_e + r_b \cdot i_c \tag{V.2.43}$$

$$u_c = (r_b + \alpha_e r_c)i_e + (r_b + r_c)i_c \tag{V.2.44}$$

As can be seen, the resistance R_b of the point-contact transistor does not act in a manner analogous to the base resistance r_b in the filamen-

---

[1] See further, J. Bardeen, *Bell System Tech. J.*, **29**: 469 (1950).

tary transistor (resistance between point I and the base electrode in Fig. V.4.5). The latter acts rather as a feedback resistance R_K which can be added between base electrode and ground in the point-contact transistor. In view of

$$U_e^* = U_e - R_K I_b = U_e + R_K(I_e + I_c) \qquad (V.4.15)$$
$$U_c^* = U_c - R_K I_b = U_c + R_K(I_e + I_c) \qquad (V.4.16)$$

and the corresponding equations for the deviations u_e^* . . . from the operating point, we obtain from (V.4.12) and (V.4.13)

$$u_e^* = (R_K + r_e)i_e + (R_K + R_b)i_c \qquad (V.4.17)$$
$$u_c^* = (R_K + \alpha r_c)i_e + (R_K + r_c)i_c \qquad (V.4.18)$$

Comparison with (V.2.43) and (V.2.44) shows the analogy between r_b in the filamentary transistor and R_K in the point-contact transistor, whereas there is no analogue in the filamentary transistor to the resistance R_b of the point-contact transistor.

In conclusion we shall refer very briefly to an extensive field which is extremely important for practical applications. On page 124 we already touched upon the internal feedback in a transistor which can lead to instability.[1] In such cases the input impedance, measured from the primary or secondary side, of a transistor is negative, indicated by falling characteristics (see Fig. V.4.6). This phenomenon is utilized in switching circuits and has led to numerous applications of the transistor in pulse, switching, and electronic computing techniques.[2]

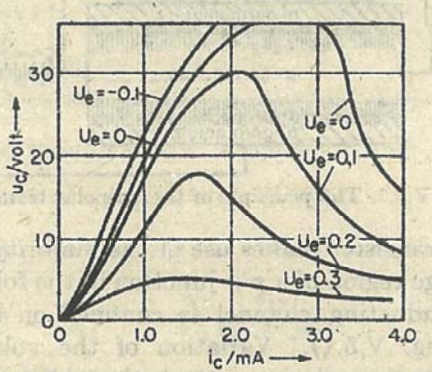

FIG. V.4.6. Family of characteristics of a transistor with regions of negative resistance.

---

[1] See in this connection also p. 148.

[2] See in this connection, for instance, B. A. E. Anderson, *Proc. IRE*, **40:** 1541 (1952).

§5. The Unipolar Transistor

In contrast to the previously discussed three transistor types, the injection of *minority* carriers plays no part in the unipolar transistor; instead use is made of displacement effects of *majority* carriers in the transition regions of *p-n* junctions. We may recall that there is a potential step V_D in a *p-n* junction even for zero current (see Fig. IV.6.1). The establishment of this step requires space charges which stem from the fact that, on both sides of the transition region, the respective majority carrier concentration is smaller than the corresponding impurity density. If, by application of a reverse voltage U, the potential step is increased to $V_D + U$, the space-charge regions which are depleted of carriers must increase in width (see Fig. IV.7.1). All this has been discussed in detail in Chap. IV, §6 and §7, in connection with the rectifying action of a *p-n* junction.

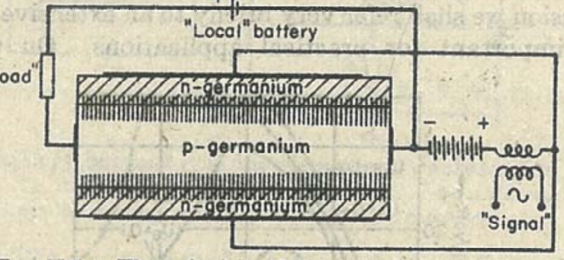

FIG. V.5.1. The principle of the unipolar transistor.

The unipolar transistor makes use of the majority carrier depletion in the space-charge region of a *p-n* junction in the following way. For instance, a *p*-conducting channel is confined on both sides by *p-n* junctions (see Fig. V.5.1). Variation of the voltage between the *p* channel and the *n* boundaries controls the width of the space-charge regions so that a varying width of the boundary strip on each side of the *p* channel is depleted of carriers. This boundary strip is subtracted from the conducting cross section for a current in the *longitudinal direction* of the channel. In the limit the entire channel width can be made carrier-free by applying sufficiently high reverse voltages between the *p* channel and the *n* boundaries. The channel is then apparently "pinched off," and the longitudinal current is blocked.

Thus the over-all effect is that the width of the channel, and with it the conductance for the longitudinal current, is controlled by the reverse voltage. The control requires only very little power, because the control voltages are in the reverse direction and the reverse currents of p-n junctions are extremely small. At the time of writing, no experimental realization of this unipolar transistor has been published, although it has already been announced by Shockley.[1]

§6. Appendix: The Amplification of Current, Voltage, and Power for a Transducer Element with Output Load

We employ the linearized current-voltage relations (V.2.18) and (V.2.19) for the treatment of these equations:

$$u_e = r_{11}i_e + r_{12}i_c \qquad \text{(V.2.18)}$$
$$u_c = r_{21}i_e + r_{22}i_c \qquad \text{(V.2.19)}$$

According to Fig. V.6.1 we have

$$u_c = -u_L = -R_L i_c \qquad \text{(V.6.01)}$$

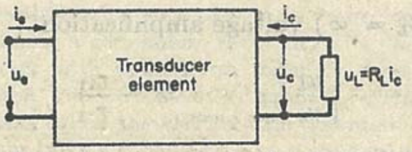

FIG. V.6.1. Transducer element with a load resistor at the output side.

where R_L is the load resistor in the secondary circuit:

$$0 = r_{21}i_e + (r_{22} + R_L)i_c \qquad \text{(V.6.02)}$$

From the preceding we obtain for the current amplification

$$\frac{i_c}{i_e} = -\frac{r_{21}}{r_{22}}\frac{r_{22}}{r_{22} + R_L} \qquad \text{(V.6.03)}$$

For the voltage amplification[2] we obtain with (V.6.03), (V.6.01), and (V.2.18)

---

[1] W. Shockley, *Proc. IRE*, **40**: 1365 (1952), particularly end of p. 1376. See also G. L. Pearson, *Phys. Rev.*, **90**: 336 (1953); and G. C. Dacey and I. M. Ross, *Proc. IRE*, **41**: 970 (1953).

[2] We define the ratio of the secondary terminal voltage u_c to the primary terminal voltage u_e as the voltage amplification. The maximum voltage which the primary voltage source is capable of producing without transducer element is often taken as the reference quantity instead of u_e. This would be the emf e_g of the primary *generator*. In this case we would have to introduce an internal resistance r_g of the generator in Fig. V.6.1, and we would have to distinguish between the emf e_g and the terminal voltage u_e. For our discussion, which is directed at the physical essentials, we can refrain from considering these more detailed differences.

$$\frac{u_L}{u_e} = -\frac{r_{21}}{r_{11}}\frac{R_L}{r_{22}+R_L}\frac{1}{1-\dfrac{r_{21}}{r_{11}}\dfrac{r_{12}}{r_{22}+R_L}} \tag{V.6.04}$$

Equations (V.6.04) and (V.6.03) together lead to the power amplification[1]

$$\frac{u_L i_c}{u_e i_e} = \frac{r_{21}^2}{r_{11}\cdot r_{22}}\frac{r_{22}R_L}{(r_{22}+R_L)^2}\frac{1}{1-\dfrac{r_{21}}{r_{11}}\dfrac{r_{12}}{r_{22}+R_L}} \tag{V.6.05}$$

In both of these equations, the second factor $r_{22}/(r_{22}+R_L)$, $R_L/(r_{22}+R_L)$, and $r_{22}R_L/(r_{22}+R_L)^2$ represents the effect of the voltage division between the circuit element and the load resistance on the secondary side.[2] The physically important information is obtained from the short-circuit ($R_L = 0$) current amplification

$$\left[\frac{i_c}{i_e}\right]_{\text{short circuit}} = -\frac{r_{21}}{r_{22}} \tag{V.6.06}$$

the open-circuit ($R_L = \infty$) voltage amplification

$$\left[\frac{u_L}{u_e}\right]_{\text{open circuit}} = -\frac{r_{21}}{r_{11}} \tag{V.6.07}$$

and the power amplification under matched ($R_L = r_{22}$) conditions[3]

$$\left[\frac{u_L i_c}{u_e i_e}\right]_{\text{matched}} = \frac{1}{4}\frac{r_{21}^2}{r_{11}r_{22}}\frac{1}{1-\dfrac{1}{2}\dfrac{r_{21}r_{12}}{r_{11}r_{22}}} \tag{V.6.08}$$

The factor on the right of (V.6.08) is significant. It shows the feedback effect of the resistance r_{12} and the resulting tendency toward

---

[1] We take the power $u_e \cdot i_e$ which is delivered to the four-pole element by the primary voltage source as the reference quantity for the definition of the power amplification. Instead, the maximum power $\frac{1}{4} \cdot (e_g^2/r_g)$ which can be taken out of the primary generator e_g, r_g without the transducer element is often used as reference quantity for the power amplification.

[2] In this connection we may recall the, possibly trivial, fact that for an a-c short circuit ($R_L = 0$) the voltage amplification and for an a-c open circuit ($R_L = \infty$) the current amplification is zero. Whether a circuit element exhibits mostly current amplification or voltage amplification depends *also* very strongly on the magnitude of the load resistance and should, therefore, not be used for the characterization of the physical mechanism of the particular circuit element.

[3] If we designate $R_L = r_{22}$ as "matching," we neglect the feedback resistance r_{12} and its effect on the input resistance $r_{22} - r_{12}r_{21}/(r_{11} + r_g)$ on the secondary side. This, too, is not essential for our discussion which is aimed at the basic principles.

instability which can occur if the magnitude of r_{12} is such that the denominator of this factor can become zero so that the power amplification becomes ∞, leading to destruction of the transducer element. Much as an increase of the power amplification is desirable, the utilization of the feedback r_{12} is risky because of the danger of instability, so that one tries in general to make r_{12} as small as possible.[1]

§7. Problems

1. What is the average transit time required for the injected minority carriers to diffuse through a base region of the thickness w, assuming $w \ll L$? Give the numerical values for germanium transistors of both polarities with $w = 10^{-4}$ cm.

2. Compare the diffusion transit time through the base region of a transistor (see Prob. 1) with the drift transit time under the influence of an electric field E. The drift transit time can be defined as the transit time due to the field only. What potential difference across the base region would be necessary in a transistor so that the drift transit time becomes equal to the diffusion transit time? For lower electric fields, the transport process is mostly diffusion-determined, for higher fields mostly field-determined.

3. Calculate the drift transit time in an n-type germanium point-contact transistor with an emitter-collector distance of 5×10^{-3} cm, for a collector electron current of 0.1 mA, assuming a germanium resistivity of 1 ohm-cm.

4.* In a semiconductor with nonuniform impurity density, the free carrier density will be nonuniform, too. On the other hand, the carrier density in thermal equilibrium can vary only when the electrostatic potential varies. This is true not only for p-n junctions but for inhomogeneous n-type or p-type doping as well. In all cases, the equilibrium carrier densities and the potential are connected by a pair of equations analogous to (IV.6.10) and (IV.6.11):

$$p(x) = p(0) \exp \left\{ -\frac{e}{kT} [V(x) - V(0)] \right\} \tag{V.7.01}$$

$$n(x) = n(0) \exp \left\{ +\frac{e}{kT} [V(x) - V(0)] \right\} \tag{V.7.02}$$

a. What is the potential distribution inside the base layer of a transistor if the donor density decreases exponentially from the emitter to the collector, according to, say,

$$n_D(x) = n_D(0) \cdot e^{-\frac{x}{\lambda}} \tag{V.7.03}$$

b. What is the drift transit time for holes (see Prob. 2) in a p-n-p germanium transistor with $n_D(0) = 1.5 \cdot 10^{17}$ cm^{-3}, $\lambda = 6.25 \cdot 10^{-5}$ cm, and $w = 5 \cdot 10^{-4}$ cm?

5. Calculate the current amplification factor α_i for an n-p-n germanium transistor with the following physical parameters: $n_D = 10^{17}$ cm^{-3}, $n_A = 2 \cdot 10^{15}$ cm^{-3}, $\tau_p = 10^{-6}$ sec, $\tau_n = 5 \cdot 10^{-5}$ sec, $w = 5 \cdot 10^{-4}$ cm. How much of the deviation of α_i from unity is due to recombination losses, and how much is due to incomplete injection?

---

[1] In contrast to the feedback resistance r_{12}, the coupling resistance r_{21} must never be zero. For $r_{21} = 0$ there would be no effect of the primary circuit on the secondary circuit, much less an *amplification* of the primary signal.

6. When the injected minority-carrier density in the base region increases with increasing emitter current, the majority-carrier density increases by the same amount in order to neutralize the charge of the injected carriers, as described in §2*a*. The result is the same as though the increase in the majority carrier density had been caused by an increase in the impurity density. In the case of an *n-p-n* transistor, this means that the hole current flowing out of the base into the emitter is not proportional to the injected electron current, as would be the case for a constant hole density in the base. Instead, this "lost" hole current rises faster with increasing emitter voltage than the electron current, so that the emitter efficiency γ_e decreases.[1] Calculate the dependence of γ_e alone and of the total current amplification factor α_i on the current. Plot the variation of the current-amplification factor for the transistor of Prob. 5.

7. With increasing collector voltage the width of the collector junction increases, thereby reducing the width of the base region.[2] If the emitter voltage is held constant, the diffusion gradient in the base region becomes steeper and the current increases. This means that the collector has a low, but finite, conductance, as already shown in Fig. V.3.6. If, however, the emitter current rather than the emitter voltage is held constant, the emitter voltage will drop with increasing collector voltage by such an amount that the diffusion gradient in the base region remains constant. This means that there exists a negative feedback from the collector to the emitter. Calculate the collector conductance for constant emitter voltage $(\partial i_c/\partial U_c)u_e$ and the feedback factor for constant emitter current $(\partial U_e/\partial u_c)i_e$ for an *n-p-n* transistor with an abrupt collector junction. For simplicity, make the following additional assumptions: $n_{D,\text{collector}} \gg n_{A,\text{base}}$, $\alpha_i = 1$. Give numerical values for germanium, assuming $W = 10^{-4}$ cm, $n_{A,\text{base}} = 5 \cdot 10^{15}$ cm^{-3}, $i_e = 1$ mA, $U_c = 3$ volts.

[1] W. M. Webster, *Proc. IRE*, **42:** 914 (1954).

[2] J. M. Early, *Proc. IRE*, **40:** 1401 (1952).

Fundamentals of Semiconductor Physics

Detailed deductions of concepts and theorems from the fundamentals of semiconductor physics have been avoided in the preceding five chapters, since such deductions would not have been in the interest of the application-minded reader. However, the physicist who wishes to delve deeper in the semiconductor field will want to distinguish between the elements in modern solid-state physics which are based on fundamental physical laws and those which rest on hypotheses and simplifying assumptions. Chapters VI to X, on fundamentals, may here be helpful.

A solid body, with its enormous number of atomic nuclei and electrons, represents a many-body problem whose exact treatment is out of the question. Since a crystal is, in a sense, a single giant molecule, it seems reasonable to transfer to crystals the approximation methods found useful for molecules. This has, in fact, taken place. For instance, the approximation method of Heitler and London has been applied by Heisenberg[1] to the theory of ferromagnetism and by Hylleraas[2] and Landshoff[3] to the theory of the cohesive forces in ionic crystals. Thus a wave-mechanical basis is provided for the *atomistic* picture, which is employed mainly by crystal and physical chemists and is best suited for insulators and ionic crystals. In the last two decades this picture has unjustly been regarded as antiquated and as unduly influenced by the corpuscular viewpoint. This may arise in part from the fact that the other approximation method, the band model, is particularly suited to metals and that, for historical reasons, the long-known and well-defined phenomenon of metallic conduction

[1] W. Heisenberg, *Z. Physik*, **49**: 619 (1928).

[2] E. A. Hylleraas, *Z. Physik*, **63**: 771 (1930).

[3] H. Landshoff, *Z. Physik*, **102**: 201 (1936); *Phys. Rev.*, **52**: 246 (1937).

was investigated first. By comparison, semiconductor physics, initially so confused, has expanded its scope only in the last three decades. In fact, in the minds of most physicists it lost a slightly scurrilous character only a few years ago.

In the field of molecular structure, the band model has its starting point in the methods of Hund and Mulliken. To demonstrate clearly the equal standing[1] of the band model and the atomistic picture, it seemed appropriate to compare, in Chap. VI, the methods of London and Heitler on the one hand and of Hund and Mulliken on the other as applied to the simplest molecule, namely, the hydrogen molecule. In Chap. VII we shall discuss the band model in detail. We cannot enter upon the atomistic picture to the same extent. Quantitative results have been derived from the atomistic picture only for the theory of ferromagnetism and the theory of cohesive forces in ionic crystals. These topics lie, however, outside the scope of this book.

Instead, Chap. VIII covers the quantitative results of the Fermi statistics of electrons in crystals within the scope of the band model. We return here to the concepts of lattice-defect reactions and the laws of mass action introduced in Chap. II. In Chap. II wep ointed out that this approach was proper only for sufficient dilution of the participants in the reaction. A quantitative formulation of this assumption becomes possible only with the aid of Fermi statistics in Chap. VIII. These statistics also provide a basis for the value of the mass-action constant which, in Chap. II, could be given only without proof.

The law of mass action, which is so important for semiconductor physics, is discussed once more in Chap. IX, this time from the kinetic standpoint. Here we find relations between the mass-action constant, the so-called recombination coefficient (or the effective cross section), and the lifetime. These matters are of importance in the theory of phosphors and may be of significance in the high-frequency behavior of rectifiers and detectors.

Finally, in Chap. X, the phenomena are discussed which occur when two different solids are brought into contact. Here we obtain a representation of the concepts of the Galvani voltage, Volta potential difference (= contact potential), work function, photoelectric activation energy, and diffusion voltage.

In this connection we shall also enter into a discussion of surface states and their possible effect on the contact between two solids.

---

[1] See F. Stöckmann, *Z. physik. Chem.*, **198:** 215 (1951).

CHAPTER VI

Approximation Methods in the Quantum Mechanics of the Hydrogen Molecule

§1. Introduction

Our concern with the theory of the hydrogen molecule is not an end in itself. Instead, we wish to recognize in the approximation methods used in this two-electron problem those characteristic traits which we shall meet again in the multielectron problem of a solid.[1] In such an approximation method,[2] the Schrödinger function $\psi(\mathbf{r}_1, \mathbf{r}_2)$ of the complete molecule, which contains the position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ of both electrons 1 and 2, is initially approximated by a product of two functions $u(\mathbf{r}_1)$ and $v(\mathbf{r}_2)$, which each contain the position vector of only one electron:

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = u(\mathbf{r}_1) \cdot v(\mathbf{r}_2) \tag{VI.1.01}$$

A nonidentical[3] eigenfunction with the same total energy would be

---

[1] Hence we need not consider the refined approximation methods of S. C. Wang, *Phys. Rev.*, **31**: 579 (1928); of E. Hylleraas, *Z. Physik*, **71**: 739 (1931); or of H. M. James and A. S. Coolidge, *J. Chem. Phys.*, **1**: 325 (1933).

[2] We are not dealing here with Schrödinger's original perturbation method in which the problem to be treated differs by only a slight perturbation from an accurately solvable problem and where the solution of the perturbed problem is expanded in the eigenfunctions of the unperturbed problem. In principle, this is an infinite process which, at least theoretically, is capable of arbitrary refinement and, in case of convergence, will yield arbitrarily precise results. The methods to be described are, rather, specific realizations of a finite procedure. The historical starting point for them is the treatment of the hydrogen molecule by W. Heitler and F. London, *Z. Physik*, **44**: 455 (1927), which is to be discussed next. The general scheme was indicated by J. C. Slater, *Phys. Rev.*, **38**: 1109 (1931). See also F. Hund in H. Geiger and K. Scheel, "Handbuch der Physik," vol. XXIV, part 1, pp. 572ff., Springer-Verlag OHG, Berlin, 1933.

[3] The eigenfunctions of a square membrane clamped on the edge

$$\cos x \cdot \sin 2y$$
$$\sin 2x \cdot \cos y$$

indicate how, by the exchange of two coordinates or degrees of freedom, a nonidentical eigenfunction with the same characteristic frequency can be obtained (see Fig. VI.1.1).

that in which electrons 1 and 2 have exchanged their role

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = v(\mathbf{r}_1) \cdot u(\mathbf{r}_2) \tag{VI.1.02}$$

since, for a measurable quantity such as the energy, it is a matter of indifference which electron is denoted by 1 and which is denoted by 2. The two functions (VI.1.01) and (VI.1.02) are hence degenerate and, consequently, the correct eigenfunction will, in general, be neither of the two, but rather a linear combination of them:

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = c \cdot u(\mathbf{r}_1)v(\mathbf{r}_2) + d \cdot v(\mathbf{r}_1)u(\mathbf{r}_2) \tag{VI.1.03}$$

The values of the coefficients c and d and of the total energy E are obtained in effect by solving a variation problem replacing the Schrödinger differential equation within the limits imposed by restricting the choice of functions to the manifold (VI.1.03). We need not enter upon this here in greater detail.[1] It is merely important for us that this procedure yields as correct eigenfunctions

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = u(\mathbf{r}_1)v(\mathbf{r}_2) - v(\mathbf{r}_1)u(\mathbf{r}_2) \tag{VI.1.04}$$

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = u(\mathbf{r}_1)v(\mathbf{r}_2) + v(\mathbf{r}_1)u(\mathbf{r}_2) \tag{VI.1.05}$$

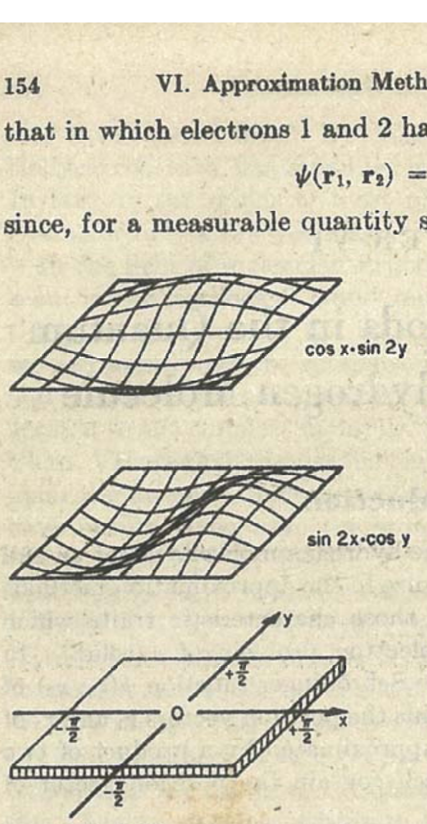

FIG. VI.1.1. Two degenerate characteristic vibrations of the square membrane with clamped boundaries.

An exchange of the two electrons 1 and 2 does not change the "symmetric" eigenfunction (VI.1.05) at all and changes only the sign of the "antisymmetric" eigenfunction (VI.1.04).[2] This was to be expected since the measurable quantities deducible from the Schrödinger function such as energy, probability density, etc., cannot depend on the numbering of the electrons. A change of sign of ψ with an exchange of

---

[1] The interested reader is referred to Geiger and Scheel, *op. cit.*, vol. XXIV, part 1, pp. 572ff.

[2] The fact that the eigenfunction which is *symmetric* in the *position* coordinates of the electrons is retained and not eliminated from the consideration is only apparently a violation of the Pauli principle. The restriction to eigenfunctions antisymmetric in the electron coordinates required by this principle demands the simultaneous consideration of the electron spin, which may occur at least formally by the introduction of the spin variables σ and the spin functions $\alpha(\sigma)$ and $\beta(\sigma)$. See H. A. Bethe in Geiger and Scheel, *op. cit.*, vol. XXIV, part 2, pp. 587–598.

electrons 1 and 2 is permissible since the Schrödinger function always occurs quadratically in these quantities.

The approximation methods of Hund and Mulliken on the one hand and of Heitler and London on the other, to be discussed below, differ first of all in the formulation of the one-electron functions $u(\mathbf{r})$ and $v(\mathbf{r})$.

§2. The Approximation Method of Hund[1] and Mulliken[2]

We proceed here from the case of closely adjoining hydrogen nuclei. Since the fields of the two nuclei overlap to a large extent, it becomes meaningless to distinguish between the electron being in the field of one nucleus a or its being in the field of the other nucleus b. From this standpoint the eigenfunctions of one electron in the field of two nuclei a and b, i.e., the eigenfunctions of the hydrogen molecular ion (see Fig. VI.2.1), appear appropriate for $u(\mathbf{r})$ and $v(\mathbf{r})$. The exact treatment of this one-electron–two-center problem is possible with the use of elliptical coordinates. However, the utilization of the resulting exact eigenfunctions of the hydrogen molecular ion in the procedure of Hund and Mulliken is awkward. Hence we shall be content with approximate expressions for the eigenfunction of an electron in the field of the two nuclei a and b obtained when the separation ab is relatively large:[3]

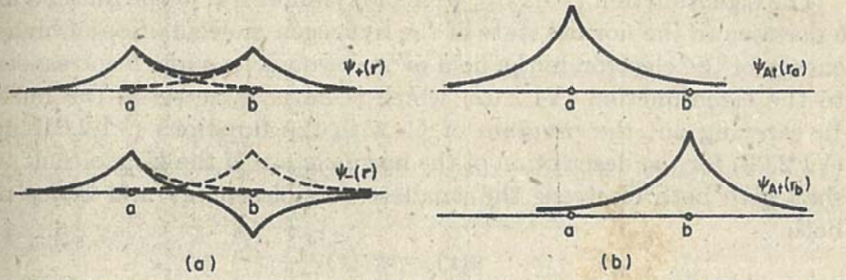

Fig. VI.2.1. The starting points of the procedures of Hund and Mulliken and of Heitler and London. (*a*) Hund and Mulliken: The eigenfunctions $\Psi_+(r)$ and $\Psi_-(r)$ of the molecular ion H_2^+ are chosen as one-electron functions. (*b*) Heitler and London: The atomic eigenfunctions $\Psi_{\mathrm{at}}(r_a)$ and $\Psi_{\mathrm{at}}(r_b)$ are chosen as one-electron functions.

$$\psi_+ (\mathbf{r}) = \psi_{\mathrm{at}}(r_a) + \psi_{\mathrm{at}}(r_b) \qquad \text{(VI.2.01)}$$

$$\psi_- (\mathbf{r}) = \psi_{\mathrm{at}}(r_a) - \psi_{\mathrm{at}}(r_b) \qquad \text{(VI.2.02)}$$

---

[1] F. Hund, *Z. Physik*, **51**: 759 (1928); **63**: 719 (1930).

[2] R. S. Mulliken, *Phys. Rev.*, **32**: 186, 761 (1928); **33**: 730 (1928).

[3] This is unquestionably a logical difficulty in our procedure, since it was stated at the start that the Hund-Mulliken approximation corresponded to the case of

Here, according to Fig. VI.2.2, $\mathbf{r}$ is the position vector of the electron referred to an arbitrary origin. r_a and r_b are the distances of the electron from the nucleus a and b, respectively. Hence the atomic eigenfunction[1]

$$\psi_{at}(r_a) = \frac{1}{\sqrt{\pi a_0^3}}\, e^{-\frac{r_a}{a_0}} \qquad \text{(VI.2.03)}$$

is centered about the nucleus a. A corresponding statement applies of course for $\psi_{at}(r_b)$. a_0 is the radius of the first Bohr orbit

$$a_0 = \frac{h^2}{4\pi^2 m e^2} = 0.5281 \cdot 10^{-8}\ \text{cm} \qquad \text{(VI.2.04)}$$

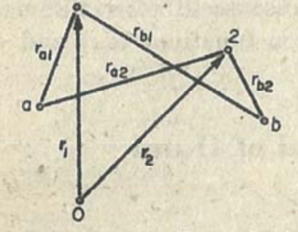

FIG. VI.2.2. Two electrons 1 and 2 in the field of two hydrogen nuclei a and b.

The eigenfunction (VI.2.01) which is symmetrical in the nuclei a and b pertains to the normal state of the hydrogen molecular ion; a higher[2] energy of the electron in the field of the two nuclei a and b corresponds to the eigenfunction (VI.2.02) which is antisymmetric in the nuclei. In carrying out the program of §1 with the functions (VI.2.01) and (VI.2.02) for the description of the normal state of the H_2 molecule, we shall give both electrons the smallest possible energy and hence put both

$$u(\mathbf{r}) = \psi_+(\mathbf{r})$$

and

$$v(\mathbf{r}) = \psi_+(\mathbf{r})$$

Since $u \equiv v$ we obtain for the normal state of the hydrogen molecule only the eigenfunction (VI.1.05) which is symmetrical in electrons 1

---

closely adjoining nuclei. In this limiting case, however, no simple relation is obtained between the eigenfunction of the molecular ion and the atomic eigenfunctions $\psi_{at}(r_a)$ and $\psi_{at}(r_b)$. In view of the intended comparison with the results of Heitler and London, such a relation is more important to us here than the attainment of the highest possible accuracy in the numerical results. Hence we make use of the approximate expressions (VI.2.01) and (VI.2.02) for the eigenfunction of the molecular ion which apply for large separation of the nuclei.

[1] Hydrogen problem, normal state!

[2] This can be recognized intuitively, even without calculation. The plane of symmetry between the two nuclei is a region of (relatively) high potential. For the antisymmetric eigenfunction, the probability density of the electron in the plane of symmetry is zero (because of $r_{a1} = r_{b1}$), and it is small in the neighborhood of the plane of symmetry. The large negative contributions to the potential energy made by this region for the symmetrical eigenfunction drop out for the antisymmetric eigenfunction, so that the total energy is reduced less. Hence the normal state with the lowest energy corresponds to the symmetrical eigenfunction, whereas an excited state with higher (or less negative) energy corresponds to the antisymmetric eigenfunction. The symmetric eigenfunction (VI.2.01) thus represents a "binding state," the antisymmetric eigenfunction (VI.2.02) a "loosening state."

and 2, whereas the antisymmetric eigenfunction (VI.1.04) is identically equal to zero. With (VI.2.01) we obtain[1]

$$\psi_+(\mathbf{r}_1) \cdot \psi_+(\mathbf{r}_2) = [\psi_{at}(r_{a1}) + \psi_{at}(r_{b1})] \cdot [\psi_{at}(r_{a2}) + \psi_{at}(r_{b2})]$$

With the readily understood abbreviations $a1$, $b1$, $a2$, $b2$ for the atomic eigenfunctions $\psi_{at}(r_{a1})$, $\psi_{at}(r_{b1})$, $\psi_{at}(r_{a2})$, $\psi_{at}(r_{b2})$ we obtain for the eigenfunction of the normal state

$$\begin{aligned}&[a1 + b1] \cdot [a2 + b2] \\ &\qquad = (a1 \cdot a2 + b1 \cdot b2) + (a1 \cdot b2 + b1 \cdot a2) \qquad \text{(VI.2.05)}\end{aligned}$$

If, next, we wish to consider an excited state of the hydrogen molecule, we must give only one of the electrons the lowest possible energy. Accordingly, we substitute in the scheme of §1

$$\begin{aligned} u(\mathbf{r}) &= \psi_+(\mathbf{r}) \\ v(\mathbf{r}) &= \psi_-(\mathbf{r}) \end{aligned}$$

Then (VI.1.04) and (VI.1.05) lead to the following two molecular eigenfunctions for two excited states:[2]

$$\begin{aligned}&\psi_+(\mathbf{r}_1) \cdot \psi_-(\mathbf{r}_2) - \psi_-(\mathbf{r}_1) \cdot \psi_+(\mathbf{r}_2) \\ &\qquad = [a1 + b1] \cdot [a2 - b2] - [a1 - b1] \cdot [a2 + b2] \\ &\qquad\qquad \sim (a1 \cdot b2 - b1 \cdot a2) \qquad \text{(VI.2.06)} \\ &\psi_+(\mathbf{r}_1) \cdot \psi_-(\mathbf{r}_2) + \psi_-(\mathbf{r}_1) \cdot \psi_+(\mathbf{r}_2) \\ &\qquad = [a1 + b1] \cdot [a2 - b2] + [a1 - b1] \cdot [a2 + b2] \\ &\qquad\qquad \sim (a1 \cdot a2 - b1 \cdot b2) \qquad \text{(VI.2.07)}\end{aligned}$$

Finally, accommodating both electrons 1 and 2 in the antisymmetric state (VI.2.02) and carrying out, correspondingly, the program of §1 with

$$\begin{aligned} u(\mathbf{r}) &= \psi_-(\mathbf{r}) \\ v(\mathbf{r}) &= \psi_-(\mathbf{r}) \end{aligned}$$

leads, as for the normal state, to a single molecular function which is symmetrical in the electrons:

$$\begin{aligned}\psi_-(r_1) \cdot \psi_-(r_2) &= [a1 - b1] \cdot [a2 - b2] \\ &= (1a \cdot a2 + b1 \cdot b2) - (a1 \cdot b2 + b1 \cdot a2) \qquad \text{(VI.2.08)}\end{aligned}$$

Altogether, the procedure of Hund and Mulliken thus leads to the following four eigenfunctions:

Normal state:

$$(a1 \cdot a2 + b1 \cdot b2) + (a1 \cdot b2 + b1 \cdot a2) \qquad \text{(VI.2.05)}$$

---

[1] We omit the factor 2 in front of the complete eigenfunction which follows from (VI.1.04) for $u \equiv v \equiv \psi_+$ since it would be changed again at any rate in carrying out the normalization procedure.

[2] Factors -2 and $+2$, respectively, have been omitted from the final expressions.

Excited states:

$$(a1 \cdot b2 - b1 \cdot a2) \qquad \text{(VI.2.06)}$$

$$(a1 \cdot a2 - b1 \cdot b2) \qquad \text{(VI.2.07)}$$

$$(a1 \cdot a2 + b1 \cdot b2) - (a1 \cdot b2 + b1 \cdot a2) \qquad \text{(VI.2.08)}$$

§3. The Original Procedure of Heitler and London

In contrast with the program of Hund and Mulliken, we proceed here from the case of widely separated hydrogen nuclei. We see immediately that the state of neutral H atoms is energetically much more favorable than the state $H^+ + H^-$. Starting from two neutral H atoms we reach the state $H^+ + H^-$ by the ionization of one of the hydrogen atoms (energy required: 13.54 ev)[1] and capture of the electron thus freed by the other neutral H atom (energy liberated: 0.76 ev).[2] This would require the very material expenditure of energy of $(13.54 - 0.76)$ ev $= 12.78$ ev. From this point of view it appears appropriate to approximate at least the normal state of the H_2 molecule by regarding one electron as bound to one nucleus and the other bound to the other nucleus. Accordingly, we substitute in the relations of §1

$$u(\mathbf{r}) = \psi_{at}(r_a) = \frac{1}{\sqrt{\pi a_0^3}} e^{-\frac{r_a}{a_0}}$$

$$v(\mathbf{r}) = \psi_{at}(r_b) = \frac{1}{\sqrt{\pi a_0^3}} e^{-\frac{r_b}{a_0}}$$

Then (VI.1.04) and (VI.1.05) yield as eigenfunctions

$$\psi_{at}(r_{a1}) \cdot \psi_{at}(r_{b2}) - \psi_{at}(r_{b1}) \cdot \psi_{at}(r_{a2}) = a1 \cdot b2 - b1 \cdot a2$$

$$\psi_{at}(r_{a1}) \cdot \psi_{at}(r_{b2}) + \psi_{at}(r_{b1}) \cdot \psi_{at}(r_{a2}) = a1 \cdot b2 + b1 \cdot a2$$

The sign of the so-called exchange integral determines for which of the two functions the energy is lower and which, hence, represents the normal state. For most molecules, including the hydrogen molecule, the exchange integral is negative, so that the normal state corresponds to the eigenfunction which is symmetrical in the position coordinates of the electrons:[3]

Normal state: $a1 \cdot b2 + b1 \cdot a2$ (VI.3.01)
Excited state: $a1 \cdot b2 - b1 \cdot a2$ (VI.3.02)

---

[1] See J. D'Ans and E. Lax, "Taschenbuch für Chemiker und Physiker," p. 113, Springer-Verlag OHG, Berlin, 1943.

[2] *Ibid.*, p. 116.

[3] See, e.g., H. A. Bethe in Geiger and Scheel, *op. cit.*, vol. XXIV, part 2, p. 592.

§4. Extension of the Procedure of Heitler and London by the Inclusion of Polar States. Comparison of the Approximations of Hund and Mulliken and Those of Heitler and London

In a comparison of the results (VI.3.01) and (VI.3.02) of the original procedure of Heitler and London with those of the procedure of Hund and Mulliken, the omission of polar states with H^+ and H^- in the former, resulting in the absence of the products $a1 \cdot a2$ and $b1 \cdot b2$, proves disturbing. Furthermore, an application of the original procedure of Heitler and London to a crystal could never lead to conductivity. The transition of an electron from an atom a to an atom b must invariably be accompanied by a compensating transition of another electron from atom b to atom a if we limit ourselves to states in which every atom has at all times one electron, and only one electron. Thus we must supplement the results (VI.3.01), (VI.3.02) of the original procedure of Heitler and London by two additional excited states $a1 \cdot a2 - b1 \cdot b2$ and $a1 \cdot a2 + b1 \cdot b2$. The comparison now takes the form:

Heitler and London

$$\begin{array}{lllr}
\text{Normal state:} & & a1 \cdot b2 + b1 \cdot a2 & \text{(VI.4.01)} \\
\multirow{3}{*}{\text{Excited states:}} & \left\{\begin{array}{l} \\ a1 \cdot a2 - b1 \cdot b2 \\ a1 \cdot a2 + b1 \cdot b2 \end{array}\right. & \begin{array}{l} a1 \cdot b2 - b1 \cdot a2 \\ \\ \\ \end{array} & \begin{array}{r} \text{(VI.4.02)} \\ \text{(VI.4.03)} \\ \text{(VI.4.04)} \end{array}
\end{array}$$

Hund and Mulliken

$$\begin{array}{llr}
\text{Normal state:} & (a1 \cdot a2 + b1 \cdot b2) + (a1 \cdot b2 + b1 \cdot a2) & \text{(VI.4.05)} \\
\text{Excited states:} & \left\{\begin{array}{r} a1 \cdot b2 - b1 \cdot a2 \\ a1 \cdot a2 - b1 \cdot b2 \qquad\qquad\qquad\quad \\ (a1 \cdot a2 + b1 \cdot b2) - (a1 \cdot b2 + b1 \cdot a2) \end{array}\right. & \begin{array}{r} \text{(VI.4.06)} \\ \text{(VI.4.07)} \\ \text{(VI.4.08)} \end{array}
\end{array}$$

We now see that the two approaches differ in the estimation of the normal state and of what is here the highest excited state; according to Hund and Mulliken these are quite definite linear combinations of the two corresponding states as adjudged by London and Heitler. According to Hund and Mulliken the heteropolar states $a1 \cdot a2$ and $b1 \cdot b2$ participate with equal weight in the normal state as the homopolar states $a1 \cdot b2$ and $b1 \cdot a2$. Simply expressed, the electrons are, according to Hund and Mulliken, as frequently together at one and the same nucleus as distributed between the two nuclei. By contrast, according to London and Heitler, the electrons are always distributed between the two nuclei in the normal-state and never at one and the same nucleus.

At the same time it is intuitively obvious that the actual circumstances are reflected in the procedure of Heitler and London only for very widely separated nuclei and in the procedure of Hund and Mulliken only for the converse case of very closely adjoining nuclei. These are, indeed, the starting points of the two procedures.

We can conclude, hence, that neither procedure describes the actual circumstances correctly for intermediate nuclear separations. The foregoing comparison of the results of the approximation of Hund and Mulliken and the extended approximation of Heitler and London indicates, however, how the approximation might be improved without the introduction of functions more complex than the atomic eigenfunction (VI.2.03). The scheme of §1 or, more precisely, the range of functions admitted to comparison will have to be extended and Slater's perturbation theory will have to be carried out with the formulation

$$\begin{aligned}\psi(\mathbf{r}_1, \mathbf{r}_2) = {} & A[a1 \cdot b2 + b1 \cdot a2] + B[a1 \cdot b2 - b1 \cdot a2] \\ & + C[a1 \cdot a2 - b1 \cdot b2] + D[a1 \cdot a2 + b1 \cdot b2] \quad \text{(VI.4.09)}\end{aligned}$$

Considerations of symmetry lead to the conclusion that there will be three types of eigenfunctions:

Nuclei antisymmetric; electrons antisymmetric:

$$a1 \cdot b2 - b1 \cdot a2 \qquad \text{(VI.4.10)}$$

Nuclei antisymmetric; electrons symmetric:

$$a1 \cdot a2 - b1 \cdot b2 \qquad \text{(VI.4.11)}$$

Nuclei symmetric; electrons symmetric:

$$A(a1 \cdot b2 + b1 \cdot a2) + D(a1 \cdot a2 + b1 \cdot b2) \qquad \text{(VI.4.12)}$$

Carrying through the perturbation calculation again leads for the two intermediate states to the types (VI.4.10) and (VI.4.11) as in the procedure of Heitler and London [(VI.4.02) and (VI.4.03), respectively] and that of Hund and Mulliken [(VI.4.06) and (VI.4.07), respectively]. However, eigenfunctions of the type (VI.4.12) are obtained for the normal state and the highest excited state. The heteropolar components vanish for the normal state and the homopolar components for the highest excited state only for the limiting case of very widely separated nuclei (Heitler and London). Correspondingly, homopolar and heteropolar components participate equally in the normal state and the highest excited state only in the limiting case of very closely adjoining nuclei (Hund and Mulliken). In any case, the heteropolar component of the normal state is surprisingly large even for the finite nuclear separation of the actual H_2 molecule. The probability of finding both electrons at one nucleus is 37 per cent, whereas according to Hund and Mulliken it would be 50 per cent and according to Heitler and London, 0 per cent.[1]

What conclusions may be drawn for a giant molecule such as a crystal from the relationships just described in the hydrogen molecule? The approximation of Heitler and London with the inclusion of polar states corresponds to the atomistic picture in solid-state physics. Here every electron is ascribed to a particular atom in the normal state, and excited states are created by ion formation (by the formation of "pairs" and "vacancies"). This picture is generally preferred by physical chemists and then frequently described somewhat misleadingly as a corpuscular picture. The considerations of Hund and Mulliken correspond to the band model in solid-state physics. Here even in the normal state every electron is distributed over all the atoms of the crystal. The normal state of the entire system is formed by the distribution of the electrons over the states of the one-electron approximation in accord with the Fermi distribution (any one state occupied by only two electrons). Quite incorrectly, the band model is frequently regarded as the only approach consistent with wave mechanics.

The relationships for the hydrogen molecule show us, however, that the atomistic picture has the drawback of a complete disregard of the polar states in the description of the normal state. The band model, on the other hand, exaggerates the importance of the polar states. Both pictures or models are thus equally imperfect and may be employed with the same justification, and the same care, supplementing each other mutually.[2] There is no analogy to a perturbation calculation with the broader formulation (VI.4.09) in solid-state theory.

Since the foregoing considerations proceed from the homopolar hydrogen molecule, they apply only for atomic lattices with atomic or valence bonds. However, if the difference between the atomistic picture and the band model is seen to rest in the fact that the atomistic picture ascribes the same state of charge to every building stone of the lattice without exception in the normal state, whereas the band model exaggerates the importance of charge transfers between the building stones of the lattice in the description of the normal state, the following extension to the ionic lattice appears appropriate:

In an ionic lattice the atomistic picture ascribes, in the normal state, the same state of charge to all building stones of a sublattice. Thus according to the atomistic picture, in an NaCl crystal all Na are positively charged and all Cl are negatively charged as long as the crystal

---

[1] See H. A. Bethe in Geiger and Scheel, *op. cit.*, vol. XXIV, part 1, p. 541.

[2] See F. Stöckmann, *Z. physik. Chem.*, **198:** 215 (1951).

remains in its normal state. Any deviation from this would correspond to an excited state of the crystal. The band model, on the other hand, would include deviations from the arrangement Na^+Cl^- even in the normal state and would, in fact, go much too far in this direction.

§5. Problems

1. Solve the Schrödinger equation in a one-dimensional delta function potential defined by

$$U(x) = -S \cdot \delta(x) \tag{VI.5.01}$$

where S is a constant while

$$\delta(x) = \begin{cases} 0 & \text{for } x \neq 0 \\ \infty & \phantom{\text{for }} x = 0 \end{cases} \tag{VI.5.02}$$

and

$$\int_{-\varepsilon}^{+\varepsilon} \delta(x)\, dx = 1 \qquad \text{for all } \varepsilon > 0 \tag{VI.5.03}$$

For positive S values, there exist bound states which satisfy the boundary conditions

$$\psi(\pm\infty) = 0 \tag{VI.5.04}$$

and which can be normalized

$$\int_{-\infty}^{+\infty} |\psi(x)|^2\, dx = 1 \tag{VI.5.05}$$

Calculate the energy of these bound states as a function of S.

2. Write down the two normalized Heitler-London functions $\psi_\pm(x_1, x_2)$, for a "one-dimensional molecule" whose potential is given by two delta functions of the type described in Prob. 1:

$$U(x) = -S[\delta(x-d) + \delta(x+d)] \tag{VI.5.06}$$

Give the mathematical expression for both functions and make a schematic drawing showing the probability distribution $|\psi|^2$ for *one* electron when the other electron is held fixed at the position of one of the two "atoms" and halfway between the two, that is, at $x = \pm d$ and at $x = 0$.

3. For the molecular potential of Eq. (VI.5.06), calculate the two Hund-Mulliken type one-electron wave functions which represent bound one-electron states in that molecule. Show that the energy values are the solutions of the transcendental equations

$$\varkappa[1 + \tanh \varkappa d] = \frac{2mS}{\hbar^2} \qquad \varkappa[1 + \coth \varkappa d] = \frac{2mS}{\hbar^2} \tag{VI.5.07}$$

where

$$\varkappa = \sqrt{\frac{2m}{\hbar^2}(-E)} \tag{VI.5.08}$$

Which of the two equations belongs to the symmetrical and the antisymmetrical function, respectively? Which function represents the ground state? Between what limits do the ground state and the excited state vary if the interatomic distance varies from zero to infinity?

From (VI.5.07) and (VI.5.08), derive explicitly approximation formulas for the energies of the two states as functions of d for large and for small d's, respectively.

CHAPTER VII

The Band Model

§1. Introduction

The band model results from approximation methods which represent an application of the method developed by Hund and Mulliken for ordinary molecules to the giant molecule of the crystal. Strictly the crystal, with N nuclei having m electrons apiece, constitutes an $(N + N \cdot m)$-body problem. If, to begin with, the mobility of the N nuclei is neglected, the problem is reduced to an $N \cdot m$-body problem, whose solution is again built up from solutions of a one-electron problem. The characteristic common trait of the procedure of Hund and Mulliken on the one hand and of the band model on the other rests in the choice of the one-electron problem and the manner in which the various solutions of the one-electron problem are combined.

If, in analogy to the procedure of Heitler and London, the motion of an electron in the field of a single atom core were chosen as the one-electron problem, we would obtain the atomistic picture.[1] For the band model we choose, instead, the behavior of an electron in the field of all atom cores and all remaining electrons as the one-electron problem. It is true that only the force actions of the atomic cores, regarded as immobile, are considered in greater detail. The force actions of the remaining $N \cdot m - 1$ mobile electrons are considered most inadequately only in so far as they screen, more or less, the fields of the atomic cores. Incidentally, these details play a role only as the band model is applied to specific substances.[2] The assumption that the electron

---

[1] See Fig. VI.2.1.

[2] J. C. Slater, *Phys. Rev.*, **45:** 794 (1934) (sodium). J. Millman, *Phys. Rev.*, **47:** 286 (1935) (lithium). F. Seitz, *Phys. Rev.*, **47:** 400 (1935) (lithium). H. M. Krutter, *Phys. Rev.*, **48:** 664 (1935) (copper). F. Hund and B. Mrowka, *Ber. Verhandl. sächs. Akad. Wiss. Leipzig, Math.-phys. Kl.*, **87:** 185 and 325 (1935) (diamond). G. E. Kimball, *J. Chem. Phys.*, **3:** 560 (1935) (diamond). W. Shockley, *Phys. Rev.*, **50:** 754 (1936) (NaCl). D. H. Ewing and F. Seitz, *Phys. Rev.*, **50:** 760 (1936) (LiF and LiH). C. Herring and A. G. Hill, *Phys. Rev.*, **58:** 132 (1940) (beryllium). F. G. von der Lage and H. A. Bethe, *Phys. Rev.*, **71:** 612 (1947) (sodium).

considered moves in a periodic field suffices for a series of very important general theorems.

It must be emphasized that this is an assumption. As already mentioned, a part of the force actions on the electron is exerted by the remaining electrons. These are mobile and influenced in their behavior by the reference electron, resulting, in turn, in changes in their force actions on the reference electron.[1] If we assume that the electrons dodge each other with infinite ease and that the space vacated by one electron is very rapidly occupied by another electron taking its place, an electron passing from one atomic core to the next meets identical force fields and the description by a fixed periodic potential appears sensible. This need not always be the case, and it is as yet an open question whether the peculiarities of the so-called open-band semiconductors result from an inadequacy of the assumption of a periodic potential in this instance.[2]

On the other hand, the same difficulties arise already in the treatment of a single atom with several or a large number of electrons. Here Hartree's method of the self-consistent field has demonstrated that stepwise approximations permit the determination of potential distributions which take account of the interactions of the electrons quite well, even in a one-electron treatment.

In the band model we proceed, in any case, from a one-electron problem describing the motion of an electron in a fixed periodic potential field.

Accordingly, we must seek those energy eigenvalues E of the Schrödinger equation

$$\left[-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)+(-e)\cdot U(x, y, z)\right]\psi(x, y, z) = E\cdot\psi(x, y, z) \qquad \text{(VII.1.01)}$$

with the lattice-periodic potential $U(\mathbf{r} + l_1\mathbf{a}_1 + l_2\mathbf{a}_2 + l_3\mathbf{a}_3) \equiv U(\mathbf{r})$, for which the solution $\psi(\mathbf{r})$ remains finite throughout the crystal lattice. However, if we should regard a crystal with finite macro-

---

[1] Thus, strictly speaking, the reference electron does not move in a fixed potential, but in a force field such as occurs in the treatment of the image force in the emission of electrons from an incandescent solid. See, e.g., the article by W. Schottky in Wien and Harms, "Handbuch der Experimentalphysik," vol. XIII, part 2, p. 254, Akademische Verlagsgesellschaft, Leipzig, 1928.

[2] J. H. de Boer and E. J. W. Verwey, *Proc. Phys. Soc. (London)*, **49:** 59 (1935). W. Schottky, *Z. Elektrochem.*, **45:** 33 (1939), in particular p. 57. H. Dressnandt, *Z. Physik*, **115:** 369 (1940). C. Wagner, *Physik. Z.*, **36:** 721 (1935); reprinted in *Z. tech. Phys.*, **16:** 327 (1935). C. Wagner and E. Koch, *Z. physik. Chem.*, **B32:** 439 (1936). R. Peierls, *Proc. Phys. Soc. (London)*, **49:** 72 (1937).

scopic dimensions as an "entire" crystal lattice, we would introduce complexities in the solution which occur naturally at the crystal surface and have nothing to do with the behavior of the electron deep in the interior of the crystal. To dissect out neatly this behavior deep within the crystal, we apply an artifice also employed elsewhere in the theory of crystal lattices. We assume an infinite crystal and demand of ψ periodicity in an arbitrarily chosen fundamental domain G, which contains a large number of unit cells. This procedure is justified by the fact that it permits a large number of conclusions in which the arbitrary size of the fundamental domain does not enter.

There are certain specific formulations of the periodic potential $U(\mathbf{r})$ for which an exact solution of the problem under consideration is possible.[1] However, the approximation methods of Bloch and Brillouin, to be discussed next (§2 and §3), are more informative physically.[2]

In §4 we summarize the general properties, which do not depend on specific approximations, of the one-electron problem with periodic potential field and discuss briefly the cell method (Wigner and Seitz).

Then in §5 to §9 we outline additional properties of the one-electron approximation. In §5, momentum, velocity, and current contribution of an electron are discussed. In §6 and §7 we are concerned with the action of an external static field on an electron in a crystal, whereas §8 covers the case of an optical alternating field. Finally, in §9, we discuss the effect of material and thermal deviations of a real lattice from an ideally ordered lattice, leading to the concepts of mean free path, mean time between collisions, and mobility.

Up to this point we are still concerned with obtaining a one-electron approximation and familiarizing ourselves with it. The many-electron problem is first attacked in §10, though only in the relatively primitive manner corresponding to the approximation of Hund and Mulliken discussed in Chap. VI. It was shown there how the many-electron eigenfunction is built up as sum of products of solutions of the one-electron problem. Such a many-electron eigenfunction cannot tell

---

[1] R. de L. Kronig and W. G. Penney, *Proc. Roy. Soc. (London)*, **130**: 499 (1931). P. M. Morse, *Phys. Rev.*, **35**: 1310 (1930). M. J. O. Strutt, "Lamésche, Mathieusche und verwandte Funktionen in Physik und Technik," Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 1, No. 3, Springer-Verlag OHG, Berlin, 1932.

[2] The situation is similar to the two-center–one-electron problem discussed in §1 of the preceding chapter. This can be solved exactly if the two centers are hydrogen nuclei with a simple coulomb potential. Nevertheless, we preferred an approximate solution, since it was more fruitful for a comparison with the procedure of Heitler and London.

more regarding a particular selected electron than the one-electron eigenfunction which pertains to the electron in question in the many-electron eigenfunction. At this stage of the theory, the many-electron eigenfunction has really only formal significance and may be replaced by a statement regarding the distribution of the available electrons over the quantum states of the one-electron problem. This makes it plausible why in many descriptions of the band model of solid-state theory the many-electron eigenfunctions (the so-called Slater determinants[1]) are not even written down and why, instead, the electrons are simply distributed over the band system of the one-electron problem. We shall follow this example in §10 of this introductory representation.

In §11 we shall discuss, finally, the conclusions derived from the band model regarding the conductivity properties of a particular crystal.

§2. The Bloch Approximation for Strongly Bound Electrons

a. Construction of a Wave Function

In the method of Bloch,[2] an approximate solution of the one-electron problem (VII.1.01) just outlined is made possible by proceeding from the case of strongly bound electrons. This has the following significance. The energy of the electron under consideration is to be so low that, energetically, it lies deep within the funnel of potential energy created for an electron by an atom core of the lattice. In order to pass over to a neighboring atom core, the electron must tunnel through a high potential hill (see Fig. VII.2.1). Such a transition will, hence, occur relatively rarely. Under these circumstances, the ψ function of the electron will have approximately the same form in the neighborhood of the lth atom core as though this core were completely isolated and the electron were the valence electron in the field of a single atom core. In formulating these ideas mathematically, we shall restrict ourselves to a one-dimensional atomic lattice. The essential relation-

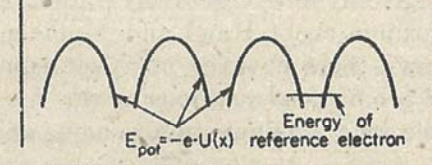

FIG. VII.2.1. Position of a strongly bound electron on the energy diagram.

---

[1] J. C. Slater, *Phys. Rev.*, **34:** 1293 (1929).

[2] F. Bloch, *Z. Physik.*, **52:** 555 (1929).

ships are brought out more clearly in this fashion. The transfer of the results to the three-dimensional lattice will lead to no difficulties. We shall discuss molecular lattices toward the end (see page 182).

For $x \approx la$ with $l = 0, \pm 1, \pm 2, \ldots$ and a = lattice constant we obtain

$$\psi(x) \approx \psi_{\text{at}}(x - la) \qquad \text{(VII.2.01)}$$

The atomic eigenfunction $\psi_{\text{at}}(x - la)$ is centered about the lth lattice point and falls off rapidly with increasing distance from this lattice point (see Figs. VII.2.2 and VII.2.3).

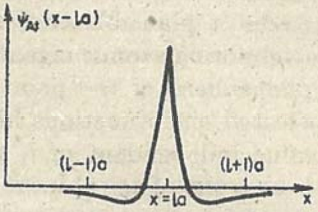

Fig. VII.2.2. Concentration of an atomic eigenfunction ψ_{at} about the lth lattice point.

The relation (VII.2.01) will apply only in the neighborhood of the lth lattice point and cannot be extended to all of space. Instead, in the neighborhood of another lattice point, the ψ function of the electron in question will be equal to the eigenfunction centered about this other lattice point. Thus we are led to the formulation

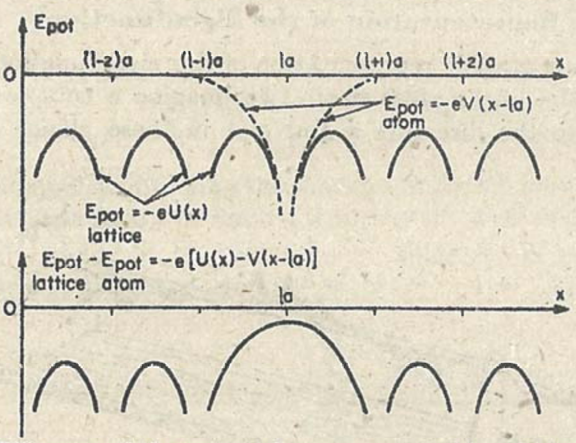

Fig. VII.2.3. Difference of the potential energy supplied by the lattice and that supplied by the lth atom.

$$\psi(x) = \sum_{l=-\infty}^{l=+\infty} c_l \psi_{\text{at}}(x - la)$$

Symmetry considerations alone make it possible to make a guess regarding the essential features of the coefficients c_l. Since all atoms in the lattice are completely equivalent, the participation of a particular atomic eigenfunction cannot be larger or smaller than that of any other eigenfunction. Accordingly, the coefficients c_l must all have the same absolute value:

$$c_l = c\, e^{j\varphi_l}$$

Furthermore, the regular arrangement of the lattice points in space makes it plausible that the phase difference $\varphi_{l+1} - \varphi_l$ between two neighboring atomic eigenfunctions $\psi_{at}(x - (l + 1)a)$ and $\psi_{at}(x - la)$ is independent of the point of the linear atom chain at which the two selected eigenfunctions happen to lie. Hence $\varphi_{l+1} - \varphi_l$ must have a value independent of l, which we can designate with ka. We shall discuss the physical meaning of the quantity k, which, for the present, may take on any real value, at a later point. We have thus arrived at the following form[1] for $\psi(x)$:

$$\psi(x) = \psi(x; k) = c \sum_{l=-\infty}^{l=+\infty} e^{jkla}\psi_{at}(x - la) \qquad \text{(VII.2.02)}$$

b. Graphic Representation of the Eigenfunction

We obtain a graphic representation of this eigenfunction if, at every point x of the linear atom chain, we imagine a complex plane perpendicular to the direction x and plot in these planes $\operatorname{Re}\psi(x)$ and

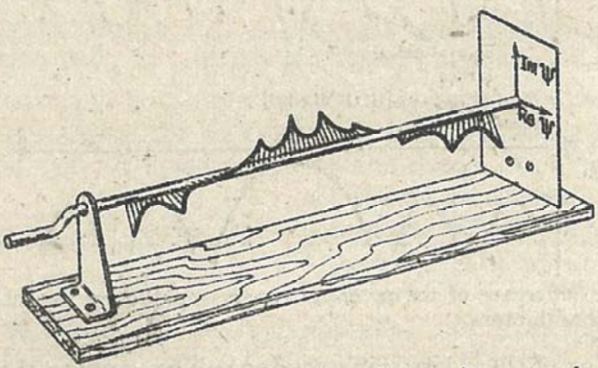

FIG. VII.2.4. Eigenfunction of an electron in a crystal.

$\operatorname{Im}\psi(x)$ (see Fig. VII.2.4). We see that, in every unit cell, the total eigenfunction $\psi(x)$ corresponds to the atomic eigenfunction but that it is rotated progressively from unit cell to unit cell by a phase angle ka. The absolute value of $\psi(x)$ is thus the same in every unit cell, i.e., equal to the absolute value of the atomic eigenfunction ψ_{at} [apart from the as yet undetermined factor c in (VII.2.02)], so that the probability density is the same in all unit cells.

---

[1] The correctness of (VII.2.02) and (VII.2.03) will be further confirmed below, see p. 177.

c. Normalization of the Eigenfunctions in the Fundamental Domain

We utilize this remark for determining the undetermined factor c in (VII.2.02). The probability of finding the electron in the lth unit cell is

$$c^2 \cdot \int_{x=la-\frac{1}{2}a}^{x=la+\frac{1}{2}a} |\psi_{\mathrm{at}}(x - la)|^2 \, dx$$

In the case here assumed of strongly bound electrons, the atomic eigenfunction $\psi_{\mathrm{at}}(x - la)$ is concentrated about the point $x = la$. We may hence neglect the contributions of the regions $-\infty < x < la - \frac{1}{2}a$ and $la + \frac{1}{2}a < x < 1 + \infty$ to the "atomic" normalization integral

$$\int_{x=-\infty}^{x=+\infty} |\psi_{\mathrm{at}}(x - la)|^2 \, dx = 1$$

and set

$$\int_{x=-\infty}^{x=+\infty} |\psi_{\mathrm{at}}|^2 \, dx = \int_{x=+la-\frac{1}{2}a}^{x=+la+\frac{1}{2}a} |\psi_{\mathrm{at}}|^2 \, dx$$

Then the probability of finding the electron in the lth unit cell becomes $c^2 \cdot 1$ and the probability of finding it in any of the G unit cells of the fundamental domain, $G \cdot c^2$. Since $\psi(x)$ is supposed to be normalized in the fundamental domain, we must put $G \cdot c^2 = 1$, so that $c = 1/\sqrt{G}$ and

$$\psi(x) = \psi(x; k) = \frac{1}{\sqrt{G}} \sum_{l=-\infty}^{l=+\infty} e^{jkla} \psi_{\mathrm{at}}(x - la) \qquad \text{(VII.2.03)}$$

d. Fulfillment of the Requirement of Periodicity in the Fundamental Domain

Apart from the requirement of normalization, we must also fulfill the requirement of periodicity in the fundamental domain, as stated on page 165. In the one-dimensional case this takes the form

$$\psi\left(-\frac{G}{2}a\right) = \psi\left(+\frac{G}{2}a\right) \qquad \text{(VII.2.04)}$$

If we substitute (VII.2.03) in this equation, the strong concentration of the atomic eigenfunctions about their lattice point makes the term with $l = -G/2$ the only one of importance on the left and that with $l = +G/2$ the only one of importance on the right. Equation

(VII.2.04) thus reduces to

$$\frac{1}{\sqrt{G}}\,e^{jk\left(-\frac{G}{2}\right)a}\,\psi_{at}(0) = \frac{1}{\sqrt{G}}\,e^{jk\left(+\frac{G}{2}\right)a}\,\psi_{at}(0)$$

or

$$e^{jkGa} = 1$$

which is satisfied by

$$k = \frac{2\pi}{a}\cdot\frac{n}{G} \qquad n = \pm\text{ integer} \qquad \text{(VII.2.05)}$$

One quantum state claims[1] a segment $2\pi/aG$ on the k scale.

e. Free and Reduced k Values

Of the values of k which are consistent with periodicity in the fundamental domain as given by (VII.2.05), $k = 0$ has the smallest absolute value. For it the eigenfunction $\psi(x) = \psi(x; 0)$ is real since the individual atomic eigenfunctions may be assumed to be real and since these are added without any phase rotation if $k = 0$:

$$\psi(x; 0) = \frac{1}{\sqrt{G}}\{\cdots + \psi_{at}(x + 2a) + \psi_{at}(x + 1a) + \psi_{at}(x) + \psi_{at}(x - 1a) + \psi_{at}(x - 2a) + \cdots\} \qquad \text{(VII.2.06)}$$

For the next k values $\pm 2\pi/aG$, the ψ functions are no longer real since the individual atomic eigenfunctions are rotated by the angle $ka = \pm 2\pi/G$, which is finite though small.

For the next two k values $\pm 4\pi/aG$, the rotation $ka = \pm 4\pi/G$ is already twice as large and ψ is again complex. Then ψ becomes once again real only for the $(G/2)$th pair of values

$$k = \pm\frac{2\pi\,(G/2)}{aG} = \pm\frac{\pi}{a}$$

Here the rotation between successive atomic eigenfunctions is just $\pm\pi$, which, for the two cases $k = +\pi/a$ and $k = -\pi/a$, leads to one and the same eigenfunction

$$\psi\left(x; \pm\frac{\pi}{a}\right) = \frac{1}{\sqrt{G}}\{\cdots + \psi_{at}(x + 2a) - \psi_{at}(x + 1a) + \psi_{at}(x) - \psi_{at}(x - 1a) + \psi_{at}(x - 2a) - \cdots\} \qquad \text{(VII.2.07)}$$

---

[1] Here we have typical examples of statements which, by themselves, have no physical meaning since they contain the arbitrary number of cells G in the fundamental domain. In combination with other considerations they lead, however, to statements which no longer contain the arbitrary G and which are then of great physical significance. See, e.g., footnote 1 on p. 179, Eqs. (VII.10.07) to (VII.10.09) on p. 264, or the derivation of Eq. (VIII.1.07) on p. 286.

Altogether there are thus

$$1 + 2\left(\frac{G}{2} - 1\right) + 1 = 1 + G - 2 + 1 = G$$

different eigenfunctions in the k interval

$$-\frac{\pi}{a} < k \leq +\frac{\pi}{a} \tag{VII.2.08}$$

This exhausts the totality of eigenfunctions. The next pair of values consistent with (VII.2.05), $k = \pm(\pi/a + 2\pi/aG)$, yields the phase rotations $\pm(\pi + 2\pi/G)$. These, however, are identical with

$$\pm\left(\pi + \frac{2\pi}{G} - 2\pi\right) = \mp\left(\pi - \frac{2\pi}{G}\right)$$

since in a phase rotation the addition or subtraction of a complete revolution 2π is immaterial. The two eigenfunctions with

$$k = \pm\left(\frac{\pi}{a} + \frac{2\pi}{aG}\right)$$

thus are identical with the two with $k = \mp(\pi/a - 2\pi/aG)$. However, the last two values occurred already as the pair preceding $k = \pm\pi/a$. We see thus that we can confine attention to the interval (VII.2.08) in the k values and may regard this as periodicity interval on the k scale, since outside this interval the eigenfunctions repeat themselves periodically. Thus, we obtain the same eigenfunctions for k and $k + 2\pi h/a$ with $h = \pm 1, \pm 2, \ldots$. If k is limited to the interval (VII.2.08) we also speak of a reduced k value.

f. The Interpretation of $\psi(x; k)$ as a Running Electron Wave Modulated with the Periodicity of the Lattice. The Special Cases of the Standing Waves $\psi(x; 0)$ and $\psi(x; \pm\pi/a)$

The special real cases (VII.2.06) and (VII.2.07) are particularly well suited for graphic representation (see Fig. VII.2.5). As we pass over to the time-dependent Schrödinger function

$$\psi(x, t) = \psi(x)\, e^{-iEt/\hbar} \tag{VII.2.09}$$

we find that these real cases are, furthermore, distinguished by the fact that they represent standing waves. In general, however, the ψ function represents a running wave. We may obtain an idea of the nature of this wave by the following transformation of (VII.2.03) and (VII.2.09):

$$\psi(x,\, t) = \frac{1}{\sqrt{G}} \sum_{l=-\infty}^{l=+\infty} \mathrm{e}^{jkla} \psi_{\mathrm{at}}(x - la) \cdot \mathrm{e}^{-j\frac{E}{\hbar}t}$$

$$= \frac{1}{\sqrt{G}} \sum_{l=-\infty}^{l=+\infty} \mathrm{e}^{-jk(x-la)} \psi_{\mathrm{at}}(x - la) \cdot \mathrm{e}^{j\left(kx - \frac{E}{\hbar}t\right)}$$

$$\psi(x,\, t) = u(x;\, k) \cdot \mathrm{e}^{j(kx-\omega t)} \tag{VII.2.10}$$

with

$$\omega = \frac{E}{\hbar} \tag{VII.2.11}$$

and

$$u(x;\, k) = \frac{1}{\sqrt{G}} \sum_{l=-\infty}^{l=+\infty} \mathrm{e}^{-jk(x-la)} \psi_{\mathrm{at}}(x - la) \tag{VII.2.12}$$

If we compare (VII.2.10) with the usual representations of a plane wave

$$a\, \mathrm{e}^{j(kx-\omega t)} = a\, \mathrm{e}^{j2\pi\left(\frac{x}{\lambda} - ft\right)} = a\, \mathrm{e}^{jk(x-ct)} \tag{VII.2.13}$$

we see that the amplitude of the electron wave in the periodic potential field is not independent of position, but is modulated with the periodicity of the lattice. Thus, (VII.2.12) shows that the amplitude $u(x;\, k)$ has the periodicity of the lattice

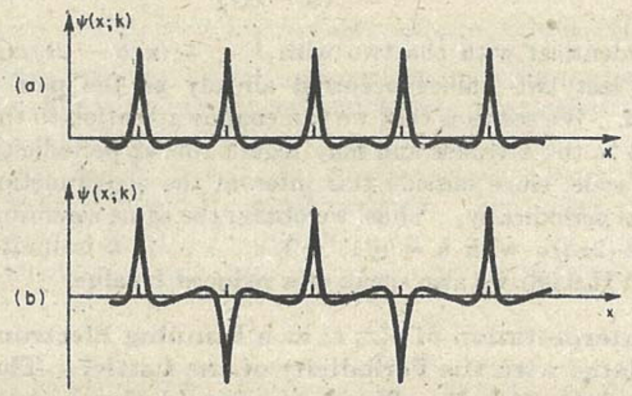

FIG. VII.2.5. The eigenfunctions are real at the band edges (in the one-dimensional atomic lattice). (*a*) $k = 0$: state at the lower edge of the band if the exchange integral is negative. (*b*) $k = \pi/a$: state at the upper edge of the band if the exchange integral is negative.

$$u(x + a;\, k) \equiv u(x;\, k) \tag{VII.2.14}$$

Equation (VII.2.14) is proved by substituting the definition of $u(x;\, k)$ given by (VII.2.12) and introducing a new summation index

$l' = l - 1$. We see, furthermore, that the k values, which have been rather colorless so far, represent the wave number $k = 2\pi/\lambda$.

It was shown on page 171 that the two k values k and $k + 2\pi h/a$ lead to the same eigenfunction $\psi(x)$. This is possible only if the modulation factors with lattice periodicity $u(x; k)$ and $u(x; k + 2\pi h/a)$ are related by

$$u\left(x; k + 2\pi\frac{h}{a}\right) = u(x; k)\, e^{-j2\pi\frac{h}{a}x}.$$

A check with (VII.2.12) confirms this. Although the substitution $k \rightarrow k + 2\pi h/a$ does not change the eigenfunction itself,[1] this does not apply to the modulation factor $u(x; k)$ by itself.

g. Perturbation Calculation

So far we have dealt in detail with the setting up and the character of an eigenfunction for an electron in a periodic potential field. However, since we have based the eigenfunction (VII.2.03) only on plausible considerations of symmetry, it is now high time to confirm the form (VII.2.03) by a perturbation calculation. Furthermore, only such a perturbation calculation can supply the energy eigenvalues which belong to the guessed eigenfunctions, and these eigenvalues are always of greatest importance in quantum-mechanical considerations.

The procedure which we shall employ for the approximate solution of the Schrödinger equation for an electron in a crystal

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi - (E + eU(x))\psi = 0 \qquad \text{(VII.1.01)}$$

will become plainer if we forget about our past considerations and start once again from the beginning.

Accordingly, we write for the unknown eigenfunction $\psi(x)$ of the electron in the crystal

$$\psi(x) = \sum_{l=-\infty}^{l=+\infty} c_l\psi_{at}(x - la)$$

and seek to determine the unknown coefficients c_l. Here we proceed just as though the atomic eigenfunctions $\psi_{at}(x - la)$ centered about the several lattice points constituted a complete orthogonal system or as in the determination of the Fourier coefficients in a Fourier expan-

---

[1] Nor does it change the energy eigenvalue E, as we shall deduce from Eq. (VII.2.22).

sion.[1] Thus we substitute the expansion in the Schrödinger equation (VII.1.01):

$$\sum_{l=-\infty}^{l=+\infty} c_l \left\{ -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi_{at}(x-la) - (E+eU(x))\psi_{at}(x-la)\right\} = 0 \tag{VII.2.15}$$

Now the atomic eigenfunction $\psi_{at}(x - la)$ satisfies the Schrödinger equation of the isolated atom with the atomic eigenvalue E_{at} and the potential $V(x - la)$ of the lth atom core

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi_{at}(x-la) - (E_{at}+eV(x-la))\psi_{at}(x-la) = 0 \tag{VII.2.16}$$

If this is utilized in (VII.2.15), we find

$$\sum_{l=-\infty}^{l=+\infty} c_l\{(E-E_{at})\psi_{at}(x-la) + e(U(x)-V(x-la))\psi_{at}(x-la)\} = 0$$

We now multiply with an arbitrary atomic eigenfunction $\psi_{at}^*(x - ra)$ and integrate over the fundamental domain $-Ga/2 < x < +Ga/2$:

$$\sum_{l=-\infty}^{l=+\infty} c_l \left\{(E-E_{at})\cdot \int_{x=-\frac{G}{2}a}^{x=+\frac{G}{2}a} \psi_{at}^*(x-ra)\psi_{at}(x-la)\,dx \right.$$

$$\left. + e \int_{x=-\frac{G}{2}a}^{x=+\frac{G}{2}a} \psi_{at}^*(x-ra)(U(x)-V(x-la))\psi_{at}(x-la)\,dx\right\} = 0 \tag{VII.2.17}$$

Bloch now puts

$$\int_{x=-\frac{G}{2}a}^{x=+\frac{G}{2}a} \psi_{at}^*(x-ra)\cdot\psi_{at}(x-la)\cdot dx \approx \delta_{rl} = \begin{cases} 1 & \text{for } r = l \\ 0 & \text{for } r \neq l \end{cases} \tag{VII.2.18}$$

since the factors in the integrand which are concentrated about the

---

[1] The deeper justification of this procedure rests in the fact that we obtain in this manner the best solution of a variation problem equivalent to the Schrödinger equation, as in the Ritz procedure. See, e.g., F. Hund in H. Geiger and K. Scheel, "Handbuch der Physik," vol. XXIV, part 1, p. 575, Springer-Verlag OHG, Berlin, 1933.

points $x = ra$ and $x = la$ coincide only for $r = l$ and since the integrand can attain appreciable values only in these cases. For $r \neq l$ one of the two factors has dropped to such a low value at every point x that no appreciable contributions can result. On the other hand, for $r = l$ the integral is practically[1] the normalization integral of the atomic eigenfunction ψ_{at}, and hence $= 1$. For the second type of integral $\int_{x=-\frac{G}{2}a}^{x=+\frac{G}{2}a} \psi^*_{at}(x - ra)[U(x) - V(x - la)]\psi_{at}(x - la)\,dx$ in (VII.2.17) Bloch retains the three cases $l = r + 1$, $l = r$, and $l = r - 1$ (see Fig. VII.2.6). He does not content himself with the case $l = r$ since, although the factors $\psi_{at}(x - ra)$ and $\psi_{at}(x - la)$ coincide here, the central factor $[U(x) - V(x - la)]$ becomes very small in the very region $x \approx ra = la$ in which these two factors contribute materially (see Fig. VII.2.3). Hence the integrals with $l = r + 1$ and with

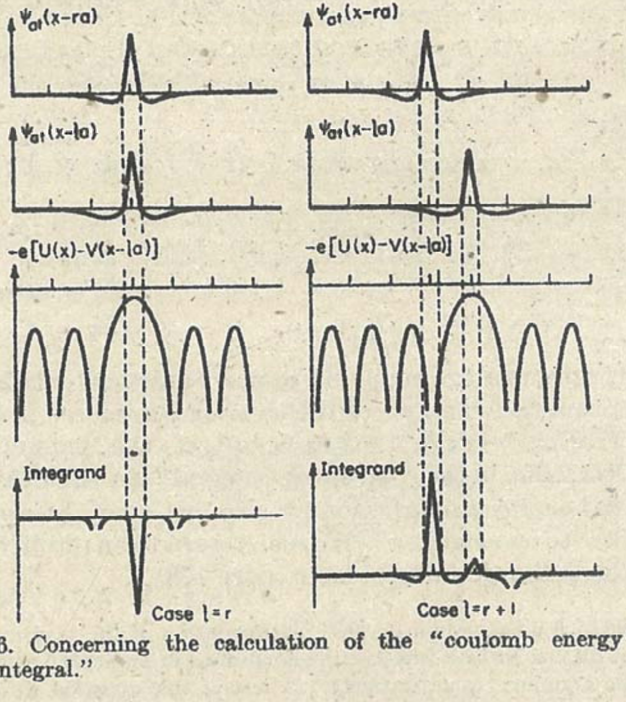

FIG. VII.2.6. Concerning the calculation of the "coulomb energy" and the "exchange integral."

---

[1] In the ranges $-\infty < x < -Ga/2$ and $+Ga/2 < x < +\infty$ which are missing as compared with the normalization integral, the contributions are negligible because of the rapid falling off of $\psi_{at}(x)$.

$l = r - 1$ must be retained in addition to the integral with $l = r$, since in these two instances also at least two of the three factors in the integrand, namely, now $\psi_{at}(x - ra)$ and $[U(x) - V(x - la)]$, are appreciably large in the region $x \approx ra$. Incidentally, the symmetry of the atomic eigenfunctions about "their" lattice point and the symmetry of the potential variations U and V lead to equal values for the integrals[1] with $l = r + 1$ and $l = r - 1$.

We put

$$-e \cdot \int_{x=-\frac{G}{2}a}^{x=+\frac{G}{2}a} \cdots dx \approx \int_{x=-\infty}^{x=+\infty} \psi^*_{at}(x - ra) \cdot (-e)[U(x) - V(x - ra)] \cdot \psi_{at}(x - ra) = C \qquad \text{(VII.2.19)}$$

$$-e \cdot \int_{x=-\frac{G}{2}a}^{x=+\frac{G}{2}a} \cdots dx \approx \int_{x=-\infty}^{x=+\infty} \psi^*_{at}(x - ra) \cdot (-e) \cdot [U(x) - V(x - (r + 1)a)] \cdot \psi_{at}(x - (r + 1)a)\, dx = A$$

$$-e \cdot \int_{x=-\frac{G}{2}a}^{x=+\frac{G}{2}a} \cdots dx \approx \int_{x=-\infty}^{x=+\infty} \psi^*_{at}(x - ra) \cdot (-e) \cdot [U(x) - V(x - (r - 1)a)] \cdot \psi_{at}(x - (r - 1)a)\, dx = A \qquad \text{(VII.2.20)}$$

These designations are supposed to point out that (VII.2.19) represents the "coulomb energy" of the electron considered in the field of the perturbation $-e[U(x) - V(x - ra)]$ of the potential energy, whereas (VII.2.20) is the "exchange integral" (*Austauschintegral*) of this potential-energy perturbation for a transition of the electron from one atom core to its neighbor. We shall return to an intuitive interpretation of the exchange integral later (page 178).

---

[1] Figure VII.2.6 gives only a graphic representation of the reasoning of Bloch. We shall not discuss further whether the conditions in an actual case are such as to justify the argument quantitatively. A test of this question would lead to a rather difficult study of the variation of the lattice potential and would also, e.g., have to treat separately the case in which the reference electron is a valence electron in the outermost shell and that in which it is strongly bound in the inner shells. In the last instance in particular, it is questionable whether the approximation of Heitler and London is not to be preferred to the procedure of Hund and Mulliken.

It should be emphasized at this point that the Bloch approximation represents, historically, the first treatment of electrons in a crystal. Here qualitative traits are of primary importance; the results can scarcely be regarded as quantitatively significant.

With (VII.2.18), (VII.2.19), and (VII.2.20), Eq. (VII.2.17) leads to

$$\cdots + 0 + 0 \cdots + c_{r-1}\{(E - E_{\mathrm{at}}) \cdot 0 - A\} + c_r\{(E - E_{\mathrm{at}}) \cdot 1 - C\} \\ + c_{r+1}\{(E - E_{\mathrm{at}}) \cdot 0 - A\} + 0 + \cdots = 0$$

or

$$-Ac_{r-1} + \{E - (E_{\mathrm{at}} + C)\}c_r - Ac_{r+1} = 0 \qquad \text{(VII.2.21)}$$

Here the subscript r may be given any value $0, \pm 1, \pm 2, \ldots$ leading to an infinite number of linear Eqs. (VII.2.21) for the determination of an infinite number of unknowns c_r.

For their solution we may note the following: Since no terms occur which are free from the unknowns c_r, we are dealing with a homogeneous system of equations. A solution which is not identically equal to zero exists only if the determinant of this system of equations vanishes. This leads to an equation for determining the energy parameter E. If E is one of the roots of this secular equation, a nonvanishing solution becomes possible. It is determined except for a common factor which is fixed by the normalization requirement for

$$\psi(x) = \sum_{l=-\infty}^{l=+\infty} c_l \psi_{\mathrm{at}}(x - la)$$

However, we shall not carry through a systematic solution of Eqs. (VII.2.21) but shall utilize instead the previously (pages 168 and 169) guessed values of the c_r, namely,

$$c_r = \frac{1}{\sqrt{G}}\, e^{jkra}$$

and shall verify that these satisfy the infinite system of linear equations (VII.2.21) if the energy parameter E has the value[1]

$$E = E_{\mathrm{at}} + C + 2A \cos ka \qquad \text{(VII.2.22)}$$

For the left side of (VII.2.21), we obtain

$$-A\,\frac{1}{\sqrt{G}}\, e^{jk(r-1)a} + \{2A \cos ka\} \cdot \frac{1}{\sqrt{G}}\, e^{jkra} - A\,\frac{1}{\sqrt{G}}\, e^{jk(r+1)a} \\ = \frac{A}{\sqrt{G}}\, e^{jkra}[-e^{-jka} + 2 \cos ka - e^{jka}]$$

The bracketed term makes this in fact equal to zero independently of the running index r.

---

[1] It is of course decisive for the verification that the running index r no longer appears in the value (VII.2.22), i.e., that with a single value of E all equations of the infinite system (VII.2.21) can be satisfied by the system of values

$$c_r = \frac{1}{\sqrt{G}}\, e^{jkra} \qquad r = 0, \pm 1, \pm 2, \ldots$$

So far we have confirmed the values of the coefficients guessed previously as well as the form (VII.2.02) or (VII.2.03) of the eigenfunction $\psi(x)$ of the electron in the crystal. In addition to this, the energy spectrum of the electron in the crystal has been determined as given by (VII.2.22).

h. Physical Interpretation of the Formula for the Energy (VII.2.22)

The result (VII.2.22) is readily understood physically. The atomic eigenvalue E_{at} is changed first of all by the fact that the potential energy $-eV(x-la)$ of a single atom core is supplemented by a perturbation $-e[U(x)-V(x-la)]$. In the total energy there appears, accordingly, the mean value C of this potential-energy perturbation weighted with the probability density $|\psi_{at}(x-ra)|^2$. C is negative since, in the integrand of (VII.2.19), the central factor $-e[U(x)-V(x-ra)]$ is negative (see Fig. VII.2.3) and the product $|\psi_{at}(x-ra)|^2$ of the other two factors is always positive. The negative term C in (VII.2.22) thus corresponds to the stronger binding of the reference electron in the field of many atom cores as compared with the binding in the field of a single atom core.

The most important part of (VII.2.22) is, however, the third term $2A\cos ka$. This exchange energy arises from the fact that the reference electron does not remain at one atom core but tunnels through the potential hill to the neighboring atom from time to time and spends, on the average, an equal amount of time at every atom core.[1]

In the section Free and Reduced k Values (page 170) we have seen that there are G different eigenfunctions

$$\psi(x) = \psi(x;k) = \frac{1}{\sqrt{G}} \sum_{l=-\infty}^{l=+\infty} e^{jkla}\psi_{at}(x-la) \qquad \text{(VII.2.03)}$$

which differ in the value of the wave number k,

$$k = \frac{2\pi}{a}\cdot\frac{n}{G} \qquad n = \pm\ \text{integer} \qquad \text{(VII.2.05)}$$

In order to cover all the eigenfunctions, n had to traverse only the interval

$$-\frac{G}{2} < n \leq +\frac{G}{2}$$

---

[1] See, e.g., S. Flügge and H. Marschall, "Rechenmethoden der Quantentheorie," pp. 162–164, Springer-Verlag OHG, Berlin, 1947; or H. A. Bethe in Geiger and Scheel, *op. cit.*, vol. XXIV, part 1, p. 335. Here further details are also given concerning the relation between exchange energy and frequency of change of location.

k traversing at the same time the interval of the reduced wave numbers

$$-\frac{\pi}{a} < k \leq +\frac{\pi}{a} \qquad \text{(VII.2.08)}$$

Outside this interval, the same eigenfunctions recurred periodically. Correspondingly, (VII.2.22) shows that also no new energy eigenvalues appear outside the periodicity interval (VII.2.08). Thus the atomic eigenvalue E_{at} splits up, through the simultaneous action of the G atoms of the fundamental domain, into $G/2$ twofold eigenvalues (VII.2.22), since the same eigenvalue E belongs to $k = +2\pi n/aG$ and to $k = -2\pi n/aG$. Since G is supposed to be a very large number, the bringing together of G separate atoms changes the atomic eigenvalue E_{at} into a quasi-continuous band of $G/2$ twofold eigenvalues or G "states." [See Fig. VII.2.7. If we omit the limitation on k by (VII.2.08), we obtain Figs. VII.3.11 and VII.3.12 in §3. This representation will prove advantageous in §6 and §7.] According to the Pauli principle, every one of these states may be occupied by two electrons with opposite spins so that the band can accommodate $2(2G/2) = 2G$ electrons or two electrons per unit cell or per atom.[1] However, this remark oversteps the boundaries of this §2, in which we are not yet concerned with the many-electron problem,[2] but only with an approximate solution of the preliminary one-electron problem.

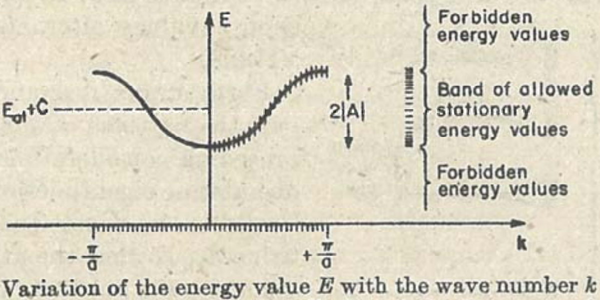

FIG. VII.2.7. Variation of the energy value E with the wave number k (for negative exchange integral, $A < 0$). It is a peculiarity of the linear atomic lattice that the density of states is least in the middle of the band and increases toward both band boundaries. The converse applies in general for actual three-dimensional lattices.

The boundaries of the band are formed by $ka = 0$ and $ka = \pm\pi$. These two values of the wave number have played a special role before, in so far as they led to the standing waves of Fig. VII.2.5. We

---

[1] In this form the statement becomes independent of the arbitrary number of cells G in the fundamental domain and becomes thus physically meaningful.

[2] This will first be taken up in §10.

conclude, hence, that standing waves correspond to the boundaries of the band for the linear atomic lattice here considered. At the edges of the band we have, furthermore,

$$\Delta E \sim (\Delta k)^2 \qquad \text{(VII.2.23)}$$

as follows from (VII.2.22) or Fig. VII.2.7.

i. Band Spectrum of a Crystal, Degenerate Atomic Eigenvalues, Application to Three-dimensional Lattices

So far we have been concerned with a single atomic eigenvalue E_{at}. However, the splitting up into a band of eigenvalues, as isolated atoms are brought together in a crystal, applies equally for all eigenvalues of the atom. Thus we obtain from the discrete spectrum of the atom a band spectrum of the crystal, in which bands of allowed and forbidden energy values alternate (see Fig. VII.2.8).

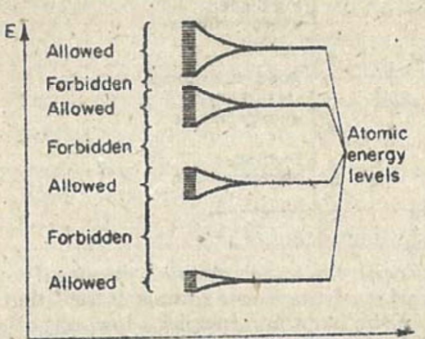

FIG. VII.2.8. The splitting of the discrete atomic spectrum into the band spectrum of the crystal.

Furthermore, it should be pointed out that it was assumed in the preceding considerations that only one atomic eigenfunction $\psi_{at}(x)$ belonged to the selected atomic eigenvalue E_{at}, so that the atomic eigenvalue E_{at} was not degenerate. The rule that a band accommodates two electrons per unit cell was derived with this assumption. For degenerate atomic eigenvalues—and these are in fact the rule—the situation becomes different. However, we shall not discuss further at this point the number of available places in the band since for a three-dimensional lattice the circumstances can become quite complex (see, however, §11). We use this instead as a first indication that caution is advised in applying the several results to the three-dimensional lattice.

Thus we have, e.g., for the linear atomic lattice just treated two different eigenfunctions for one and the same eigenvalue E, namely, a wave running toward the right and one running toward the left. Already in the two-dimensional and particularly in the three-dimensional lattice, many electron waves running in all possible directions belong, in general, to a single eigenvalue. Furthermore, not all the eigenfunctions belonging to the lowest and highest energies, i.e., to the band boundaries, need be standing waves. For example, for a face-centered cubic lattice this is not so for the upper band boundary (for a negative exchange integral). Furthermore, the relation (VII.2.23) does not apply at this boundary.

In the three-dimensional case, the periodicity interval $-\pi/a < k \leq +\pi/a$ on the k scale found for the linear lattice is replaced by a periodicity polyhedron in **k**-vector space. Outside this polyhedron, eigenvalues and eigenfunctions repeat themselves periodically. For example, the Bloch approximation yields for a quadratic point lattice an E surface above the $(\mathbf{k}_x, \mathbf{k}_y)$ plane subdivided into periodicity squares, as shown in Fig. VII.2.9. As already mentioned, this view of the $E(k)$ dependence will prove advantageous in §6 and §7. This periodicity polyhedron is frequently, and wrongly, identified with the so-called first Brillouin zone in **k** space, which will be discussed in §3. In §11, the diamond lattice will provide an example in which the Bloch periodicity polyhedron and the first Brillouin zone differ (see pages 273ff.).

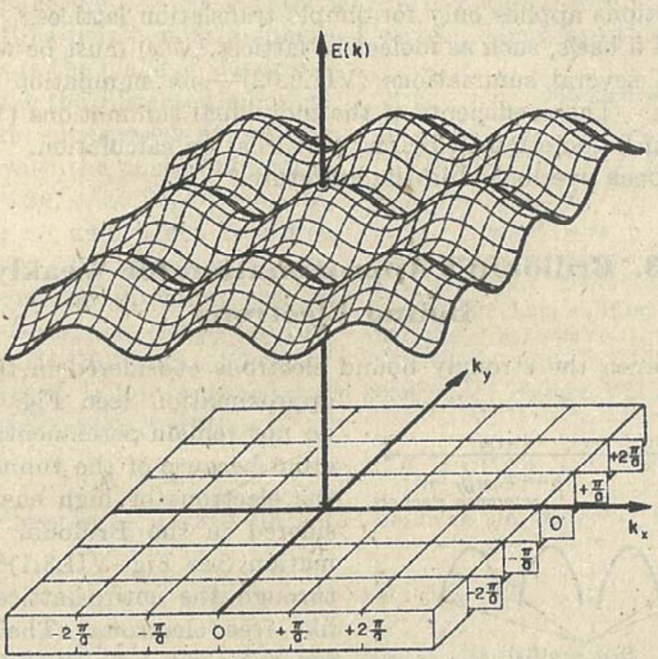

FIG. VII.2.9. Variation of the energy E with the wave number vector **k** for the two-dimensional case (quadratic point lattice).

In the one-dimensional case a quantum state claims a segment $2\pi/(aG)$ of the k scale, as we found on page 170. In the three-dimensional case, correspondingly, a volume $[2\pi/aG]^3$ of **k** space is ascribed to a quantum state. If we introduce the volume $V_{\text{fund}} = (Ga)^3$ of the fundamental domain considered, the volume in **k** space ascribed to one quantum state becomes simply $(2\pi)^3/V_{\text{fund.}}$[1]

In the three-dimensional lattice we must note, furthermore, whether the atomic eigenvalue pertains to an s, a p, or a d term. Thus, in a simple cubic lattice, an atomic p eigenvalue with its three p eigenfunctions splits up into three bands.[2] Energetically these three bands coincide completely, however, an extreme example of band overlap, which will be mentioned repeatedly in §3.

Finally it should be pointed out that the form (VII.2.03) of the eigenfunctions applies only for simple translation lattices.[3] For lattices with a basis, such as molecular lattices,[3] $\psi(x)$ must be written as a sum of several summations (VII.2.02)—one summation for each sublattice. The coefficients of the individual summations (VII.2.02) must then be determined by the perturbation calculation. This has at times been overlooked in the literature.[4]

§3. Brillouin's Approximation for Weakly Bound Electrons

While even the strongly bound electrons considered in the Bloch approximation (see Fig. VII.2.1) do not remain permanently at one atom because of the tunnel effect, the electrons of high energy considered in the Brillouin approximation (see Fig. VII.3.1)[5] can fly through the entire lattice almost like free electrons. Their eigenfunctions will be accordingly very nearly plane waves $e^{j(\mathbf{k}\cdot\mathbf{r})}$. We thus face a problem very similar to that of the bombardment of a crystal with X-rays or cathode rays. The

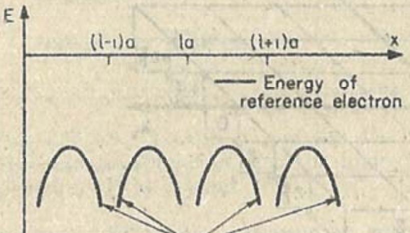

FIG. VII.3.1. Position of a quasi-free electron in the energy diagram.

---

[1] To see how, in the application of this statement in Chap. VIII, §1, the volume of the arbitrary fundamental domain drops out, see the derivation of Eq. (VIII.1.07) on pp. 283 to 286.

[2] See, e.g., H. A. Bethe in Geiger and Scheel, *op. cit.*, vol. XXIV, part 2, pp. 401–404.

[3] For a further discussion of this concept, see pp. 269 to 271.

[4] H. A. Bethe gives in Geiger and Scheel, *op. cit.*, vol. XXIV, part 2, p. 397, a formula (12.17) for the energy spectrum of a lattice with a basis in which apparently account has not been taken of the above-mentioned circumstance.

[5] See footnote 2, p. 192.

phenomena occurring in this instance are well known,[1] and we shall report on them first. The results of Brillouin's perturbation calculation for weakly bound electrons then become highly plausible physically, so that we can omit the perturbation calculation itself.

a. Diffraction of a Three-dimensional Wave by a Linear Series of Points

We begin with the diffraction phenomenon resulting from the incidence of a plane wave on a linear series of points in a direction forming a direction cosine $\alpha = \cos\varphi$ with the row of points (see Fig. VII.3.2). Every lattice point emits a spherical wave. There are then several directions in the plane of the drawing in which the diffracted spherical waves are superposed at infinite distance with the phase differences $\cdots -2\cdot 2\pi,\ -1\cdot 2\pi,\ 0,\ +1\cdot 2\pi,\ +2\cdot 2\pi \cdots$ and hence reinforce each other. In these directions, which form the direction cosines $\alpha_{-2}, \alpha_{-1}, \alpha_0, \alpha_{+1}, \alpha_{+2}$ with the row of points, the diffracted beams of -2d, -1st, 0, $+1$st, $+2$d order are observed. Figure VII.3.2 yields Laue's interference condition

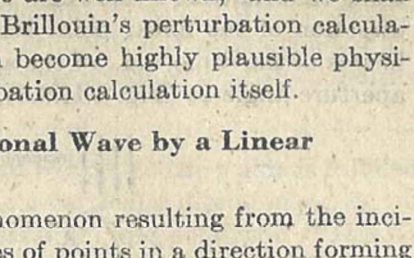

FIG. VII.3.2. Laue's diffraction condition, $a(\alpha - \alpha_0) = h\lambda$.

$$a(\alpha_{h_1} - \alpha) = h_1\lambda \qquad h_1 = 0, \pm 1, \pm 2, \cdots \qquad \text{(VII.3.01)}$$

For the direction α_{h_1} of the diffracted beam of the order h_1, we obtain therefore

$$\alpha_{h_1} = \alpha + h_1\frac{\lambda}{a} \qquad h_1 = 0, \pm 1, \pm 2, \ldots \qquad \text{(VII.3.02)}$$

Since, however, $|\alpha_{h_1}|$, as a direction cosine, cannot exceed unity, not all solutions (VII.3.02) are physically meaningful, but only a limited number of them, up to a certain positive and a certain negative limit of h_1, depending on the smallness of λ/a. We see thus that for an arbitrary, but sufficiently small, wavelength there is a finite number of diffracted beams if we limit ourselves to the plane of the drawing of Fig. VII.3.2, i.e., to the plane of incidence. This is, however, not at all necessary. The interference condition (VII.3.01) for the in-phase superposition of the diffracted spherical waves can be fulfilled equally well for directions which do not lie in the plane of incidence. The only

---

[1] We refer here to the excellent presentations of P. P. Ewald, "Kristalle und Röntgenstrahlen," Springer-Verlag OHG, Berlin, 1923, and in Geiger and Scheel, *op. cit.*, vol. XXIII, part 2.

thing that matters is the angle φ_{h_1} between the diffracted beam and the direction of the linear point lattice. We thus see that all diffracted beams of the same, e.g., of the second order,[1] lie on a cone with a half-aperture angle φ_2 (Fig. VII.3.3).

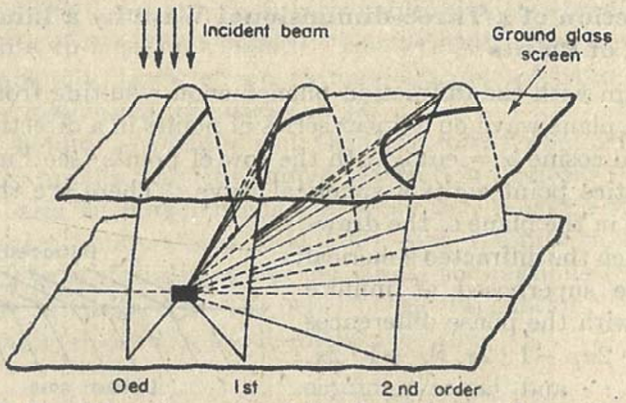

FIG. VII.3.3. The diffraction cones of the linear lattice, for an incident beam perpendicular to the lattice axis. (From P. P. Ewald, "Kristalle und Röntgenstrahlen," p. 43, Springer-Verlag OHG, Berlin, 1923.)

In summary, a one-dimensional point lattice in three-dimensional space will diffract an infinite number of beams from an incident plane wave of arbitrary direction and arbitrary (though sufficiently small) wavelength. All beams of any one order lie on an interference cone whose axis coincides with the direction of the point lattice.

b. Diffraction of a Three-dimensional Plane Wave by a Two-dimensional Point Lattice

We now pass over to a two-dimensional, e.g., quadratic, point lattice in three-dimensional space. Again, in the direction of a diffracted beam, the spherical waves diffracted by all lattice points must be superposed in phase. If this is checked first for the lattice points which form a linear point lattice in the direction of the x axis, we arrive again at Laue's condition (VII.3.01), where α_{h_1} and α are now the direction cosines with respect to the x axis. Hence, once again, cones of order 0, ± 1, ± 2, . . . would be diffracted from a wave of arbitrary direction of incidence and arbitrary (though sufficiently small) wavelength. However, it is here not at all certain that the effects proceeding from parallel series of points 1, 2, 3 (see Fig. VII.3.4) will also be in phase and hence reinforce each other.

We can also state that a gathering of the points of the quadratic

---

[1] That is, $h_1 = +2$.

lattice in the point series 1′, 2′, 3′ parallel to the y axis is equally as valid as their previous grouping in the point series 1, 2, 3 . . . parallel to the x axis and that with the former grouping diffraction cones of the order 0, ±1, ±2, . . . about the y axis are obtained if a Laue condition

$$a(\beta_{h_2} - \beta) = h_2\lambda \qquad \text{(VII.3.03)}$$

for the direction cosines β_{h_2} and β with respect to the y axis is fulfilled.

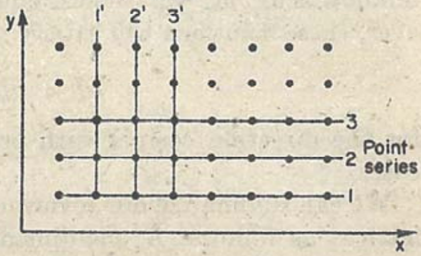

FIG. VII.3.4. The diffraction problem at a quadratic point lattice.

A cooperation of all lattice points in a particular direction α_{h_1}, β_{h_2} will take place only if both condition (VII.3.01) and condition (VII.3.03) are fulfilled. Geometrically, this means that the direction of diffraction of the order h_1, h_2 must lie both on the h_1th diffraction cone about the x axis and the h_2th diffraction cone about the y axis. Thus the only directions of diffraction remaining for the two-dimensional point lattice in three-dimensional space for arbitrary direction and wavelength of the incident waves are given by the two straight lines of intersection of the h_1th cone about the x axis and the h_2th cone about the y axis. Algebraically this is expressed by the fact that (VII.3.01) and (VII.3.03) are two equations for the two unknowns α_{h_1} and β_{h_2}, which are herewith completely determined. It is true that, to fix a direction in space completely, three direction cosines are required; the third direction cosine γ is obtained from the equation

$$\alpha^2 + \beta^2 + \gamma^2 = 1 \qquad \text{(VII.3.04)}$$

which applies for the direction cosines for orthogonal axes.[1]

Since this prescribes a value only to γ^2, not to γ itself, the sign of γ is still undetermined, so that the two Laue conditions (VII.3.01) and (VII.3.03) select two directions in space for the diffracted beam.

c. Diffraction of a Three-dimensional Plane Wave by a Three-dimensional Point Lattice

In similar fashion we must demand of the diffracted beams of a three-dimensional point lattice in three-dimensional space that they be

---

[1] For nonorthogonal axes we would obtain a general condition

$$a\alpha^2 + b\beta^2 + c\gamma^2 + 2d\alpha\beta + 2e\beta\gamma + 2f\gamma\alpha = 1$$

instead of the condition (VII.3.04), which may be regarded as "transformed to principal axes."

the intersections of three cones, e.g., about the x, the y, and the z axis. Since, however, in general, three cones do not intersect in a common straight line, we can conclude that for a three-dimensional point lattice in three-dimensional space an incident wave of arbitrary direction and arbitrary wavelength will not give rise to any diffracted beams.

Algebraically this circumstance is expressed by the fact that a third Laue condition

$$a(\gamma_{h_3} - \gamma) = h_3\lambda \qquad \text{(VII.3.05)}$$

is added to (VII.3.01) and (VII.3.03) and that, then, (VII.3.01), (VII.3.03), and (VII.3.05) represent three equations for the three unknowns $\alpha_{h_1}, \beta_{h_2}, \gamma_{h_3}$. These equations always have solutions. However, these solutions will satisfy the necessary side condition

$$\alpha_{h_1}^2 + \beta_{h_2}^2 + \gamma_{h_3}^2 = 1 \qquad \text{(VII.3.06)}$$

for the direction cosines with orthogonal axes[1] only in rare discrete cases.

We can summarize our review of the diffraction phenomena of point lattices as follows: A one-dimensional point lattice in three-dimensional space diffracts an infinite number of beams of a particular order for arbitrary direction and arbitrary (though sufficiently small) wavelength of the incident wave; a two-dimensional point lattice in three-dimensional space produces under the same circumstances only two diffracted beams; and a three-dimensional point lattice in three-dimensional space produces in general, i.e., for arbitrary direction and arbitrary wavelength of the incident wave, no diffracted beams whatever.

d. The Subdivision of k Space into the Brillouin Zones

The so-called Brillouin zone construction is a means of seeing quickly under what circumstances, i.e., for what directions of incidence and what wavelengths, a three-dimensional point lattice produces a diffracted beam. To arrive at this construction we first write the three Laue conditions (VII.3.01), (VII.3.03), and (VII.3.05) in the following form:

$$\frac{1}{2\pi}\left(\frac{2\pi}{\lambda}\alpha' - \frac{2\pi}{\lambda}\alpha\right) = h_1 \cdot \frac{1}{a_1} + h_2 \cdot 0 + h_3 \cdot 0 \qquad \text{(VII.3.07)}$$

$$\frac{1}{2\pi}\left(\frac{2\pi}{\lambda}\beta' - \frac{2\pi}{\lambda}\beta\right) = h_1 \cdot 0 + h_2 \cdot \frac{1}{a_2} + h_3 \cdot 0 \qquad \text{(VII.3.08)}$$

$$\frac{1}{2\pi}\left(\frac{2\pi}{\lambda}\gamma' - \frac{2\pi}{\lambda}\gamma\right) = h_1 \cdot 0 + h_2 \cdot 0 + h_3 \cdot \frac{1}{a_3} \qquad \text{(VII.3.09)}$$

---

[1] See footnote on page 185.

Here we have designated the direction cosines α_{h_1}, β_{h_2}, γ_{h_3} of the diffracted beam simply with α', β', γ'. Furthermore, we shall base the following treatment right away on the general translation lattice with three oblique axes $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$.

In (VII.3.07), (VII.3.08), and (VII.3.09), $2\pi\alpha/\lambda$, $2\pi\beta/\lambda$, $2\pi\gamma/\lambda$ are the perpendicular projections of the wave vector $\mathbf{k}$ of the incident wave on the three translation axes $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$. The same applies for $2\pi\alpha'/\lambda$ . . . with reference to the wave vector $\mathbf{k}'$ of the diffracted wave. Furthermore, we introduce, in view of the right side of Eqs. (VII.3.07) to (VII.3.09), the three vectors $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ by the requirement that, e.g., the perpendicular projections of the vector $\mathbf{b}_1$ with reference to the three translation axes $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ be equal to $1/a_1$, 0, 0.[1] Thus $\mathbf{b}_1$ must satisfy the three equations

$$\mathbf{b}_1 \cdot \frac{\mathbf{a}_1}{a_1} = \frac{1}{a_1} \qquad \mathbf{b}_1 \cdot \frac{\mathbf{a}_2}{a_2} = 0 \qquad \mathbf{b}_1 \cdot \frac{\mathbf{a}_3}{a_3} = 0$$

or

$$\mathbf{b}_1 \cdot \mathbf{a}_1 = 1 \qquad \mathbf{b}_1 \cdot \mathbf{a}_2 = 0 \qquad \mathbf{b}_1 \cdot \mathbf{a}_3 = 0 \qquad \text{(VII.3.10)}$$

Correspondingly, $\mathbf{b}_2$ and $\mathbf{b}_3$ must fulfill

$$\mathbf{b}_2 \cdot \mathbf{a}_1 = 0 \qquad \mathbf{b}_2 \cdot \mathbf{a}_2 = 1 \qquad \mathbf{b}_2 \cdot \mathbf{a}_3 = 0 \qquad \text{(VII.3.11)}$$

$$\mathbf{b}_3 \cdot \mathbf{a}_1 = 0 \qquad \mathbf{b}_3 \cdot \mathbf{a}_2 = 0 \qquad \mathbf{b}_3 \cdot \mathbf{a}_3 = 1 \qquad \text{(VII.3.12)}$$

The factors of h_1 in (VII.3.07), (VII.3.08), and (VII.3.09) are also the perpendicular projections of $\mathbf{b}_1$ on the three axes $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$. The same applies for the factors of h_2 and h_3 with respect to the vectors $\mathbf{b}_2$ and $\mathbf{b}_3$. Hence we can combine the three Laue conditions in the vector equation

$$\frac{1}{2\pi}(\mathbf{k}' - \mathbf{k}) = h_1\mathbf{b}_1 + h_2\mathbf{b}_2 + h_3\mathbf{b}_3 = \mathbf{h} \qquad \begin{array}{l} h_1 = 0, \pm 1, \pm 2, \ldots \\ h_2 = 0, \pm 1, \pm 2, \ldots \\ h_3 = 0, \pm 1, \pm 2, \ldots \end{array} \qquad \text{(VII.3.13)}$$

It is true that this vectorial Laue condition has for arbitrary $\mathbf{k}$ a solution

$$\frac{1}{2\pi}\mathbf{k}' = \frac{1}{2\pi}\mathbf{k} + \mathbf{h} \qquad \text{(VII.3.14)}$$

However, in order that $\mathbf{k}'$ may be the wave vector of a diffracted wave, its absolute value $|\mathbf{k}'| = 2\pi/\lambda$ must be equal to the absolute value $|\mathbf{k}| = 2\pi/\lambda$ since the wavelength of the diffracted wave must be equal

---

[1] $\mathbf{b}_1$, $\mathbf{b}_2$, and $\mathbf{b}_3$ form the so-called reciprocal lattice.

to that of the incident wave. Thus (VII.3.14) has the side condition

$$|\mathbf{k}'| = |\mathbf{k}| \tag{VII.3.15}$$

which corresponds to the earlier side condition (VII.3.04). A diffracted wave of order h_1, h_2, h_3 exists only when both Eqs. (VII.3.14) and (VII.3.15) are satisfied by the one unknown $\mathbf{k}'$.

We shall now rephrase the question of page 186, namely, for what directions of incidence and what wavelengths a three-dimensional point lattice forms a diffracted beam, in the more restricted form, under what circumstances a beam of a particular order h_1, h_2, h_3 is produced. Then the vector $\mathbf{h}$ in the Laue condition (VII.3.13) is fixed. The addition of the side condition (VII.3.15) requires now that $\mathbf{k}'/(2\pi)$ and $\mathbf{k}/(2\pi)$ form an isosceles triangle with the vector $\mathbf{h}$ (see Fig. VII.3.5*a* and *b*). It follows from this (see Fig. VII.3.5*c*) that for changing direction of incidence the wavelength determined by $|\mathbf{k}| = 2\pi/\lambda$ must vary in such a manner that the end point of $\mathbf{k}$ always lies on the perpendicular plane through the end point of the vector $-\frac{1}{2}\mathbf{h}$.

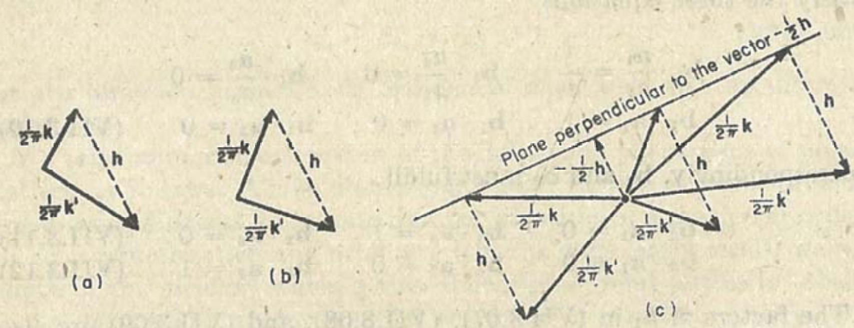

FIG. VII.3.5. Derivation of the Brillouin zone construction. (*a*) Consideration of Laue's diffraction condition alone, $\frac{1}{2\pi}\mathbf{k}' = \frac{1}{2\pi}\mathbf{k} + \mathbf{h}$. (*b*) Added consideration of the side condition $|\mathbf{k}'| = |\mathbf{k}|$. (*c*) The end point of $\mathbf{k}/(2\pi)$ must lie in the plane perpendicular to $-\mathbf{h}/2$ if diffraction is to take place.

Herewith we have already answered the restricted question: Which plane waves $e^{j\mathbf{k}\cdot\mathbf{r}}$ lead to a diffracted beam of the order h_1, h_2, h_3? They are all those waves whose wave vectors divided by 2π end on the perpendicular plane through the end point of the vector $-\frac{1}{2}\mathbf{h}$.

As we return to the more general question of page 186, under what circumstances a plane wave $e^{j\mathbf{k}\cdot\mathbf{r}}$ produces a diffracted beam of any order, we must recall that both the vector $-\mathbf{h}$ and the vector $+\mathbf{h}$ are vectors of the reciprocal lattice, since (VII.3.13) defined the vector $\mathbf{h}$ as follows:

$$\mathbf{h} = h_1\mathbf{b}_1 + h_2\mathbf{b}_2 + h_3\mathbf{b}_3 \qquad \left.\begin{matrix} h_1 \\ h_2 \\ h_3 \end{matrix}\right\} = 0, \pm 1, \pm 2, \pm 3, \ldots \quad \text{(VII.3.16)}$$

Thus, if we draw the reciprocal lattice for the translation lattice under consideration and construct the bisecting perpendicular plane for every lattice vector of the reciprocal lattice, then a diffracted wave $e^{j\mathbf{k}\cdot\mathbf{r}}$ occurs if and only if the end point of $\mathbf{k}/2\pi$ lies on one of the bisecting perpendicular planes.

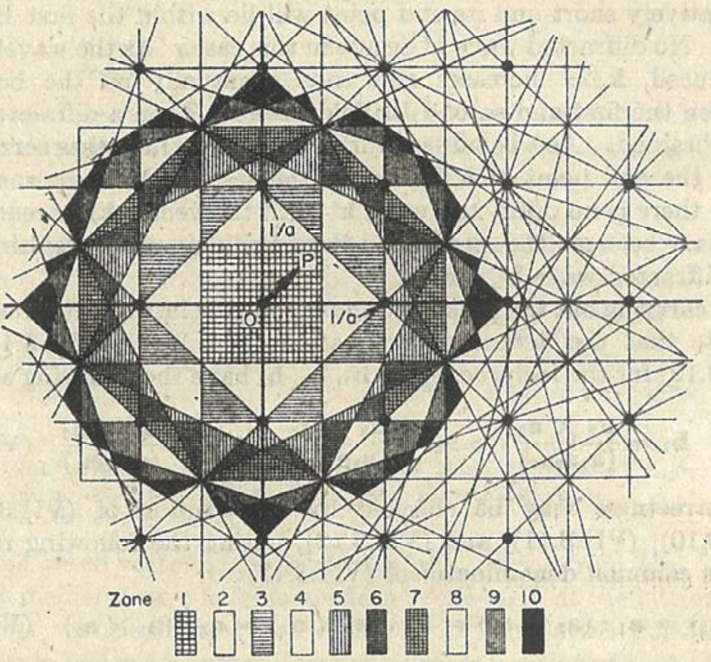

Fig. VII.3.6. Brillouin zones for the quadratic two-dimensional lattice. Strong Bragg reflection occurs if the end point of $\mathbf{k}/(2\pi)$ lies on one of the continuous lines. (From L. Brillouin, "Quantenstatistik," Springer-Verlag OHG, Berlin, 1931.)

The construction of the bisecting perpendicular planes in the reciprocal lattice is designated as Brillouin zone construction. The cubic lattice offers, of course, the simplest example for thecar rying out of this construction. Here the three vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ are mutually orthogonal and all of equal length, namely, equal to the lattice constant a. By (VII.3.10) the vector $\mathbf{b}_1$ must be perpendicular to $\mathbf{a}_2$ and $\mathbf{a}_3$ and coincides hence in direction with $\mathbf{a}_1$. Also by (VII.3.10), its length is equal to $1/a$. The same applies for $\mathbf{b}_2$ and $\mathbf{b}_3$, so that the reciprocal lattice is in the present instance a cubic lattice with the lattice constant $1/a$. Figure VII.3.6 shows the $(\mathbf{b}_1, \mathbf{b}_2)$ plane of this reciprocal lattice and the carrying out of the Brillouin zone construction within this plane. The bisecting perpendicular planes of the reciprocal lattice vectors $h_1\mathbf{b}_1 + h_2\mathbf{b}_2 + 0 \cdot \mathbf{b}_3$ in this $(\mathbf{b}_1, \mathbf{b}_2)$ plane intersect this plane in the lightly drawn straight lines. We see, therefore, how the reciprocal lattice is subdivided into the Brillouin zones.

If a plane wave with the wave vector $\mathbf{k}$ is incident from any direction on the cubic point lattice with the lattice constant a, $\mathbf{k}/2\pi$ must be entered on the reciprocal lattice. For large wavelength λ, $\mathbf{k}/2\pi$ will be relatively short and its end point will lie within the first Brillouin zone. No diffracted wave $\mathbf{k}'$ occurs in this case. As the wavelength λ is reduced, $\mathbf{k}/2\pi$ increases and ends eventually on the boundary between the first and second Brillouin zone. Then a diffracted wave $\mathbf{k}'$ is obtained. As λ is reduced further and $\mathbf{k}/2\pi$ increases correspondingly, the end point of $\mathbf{k}/2\pi$ lies in the second Brillouin zone[1] and, again, there is no diffracted wave $\mathbf{k}'$ until the vector $\mathbf{k}/2\pi$ reaches the boundary between the second and third Brillouin zone, resulting again in a diffracted wave $\mathbf{k}'$, etc.

For carrying out the construction for complex lattices, it is important to note that the determining equations (VII.3.10), (VII.3.11), and (VII.3.12) for the reciprocal axes $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ have the following solution:

$$\mathbf{b}_1 = \frac{\mathbf{a}_2 \times \mathbf{a}_3}{(\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3)} \qquad \mathbf{b}_2 = \frac{\mathbf{a}_3 \times \mathbf{a}_1}{(\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3)} \qquad \mathbf{b}_3 = \frac{\mathbf{a}_1 \times \mathbf{a}_2}{(\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3)} \qquad \text{(VII.3.17)}$$

Its correctness may be checked by substitution of (VII.3.17) in (VII.3.10), (VII.3.11), and (VII.3.12), noting the following relations for the common denominator of (VII.3.17):

$$(\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3) = \mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3) = \mathbf{a}_2 \cdot (\mathbf{a}_3 \times \mathbf{a}_1) = \mathbf{a}_3 \cdot (\mathbf{a}_1 \times \mathbf{a}_2) \qquad \text{(VII.3.18)}$$

e. Interpretation of the Diffraction Phenomenon as Bragg Reflection at a Family of Lattice Planes

For a physical interpretation it may be worth while to note that the planes perpendicular to $\mathbf{h}$ (and hence of course also to $-\frac{1}{2}\mathbf{h}$) are the families of lattice planes with the Miller indices[2] h_1, h_2, h_3. This may be seen in the following manner. According to the definition of the Miller indices, one of the lattice planes has the intercepts on the axes $\mathbf{a}_1/h_1$, $\mathbf{a}_2/h_2$, $\mathbf{a}_3/h_3$. The vectors $(\mathbf{a}_1/h_1 - \mathbf{a}_2/h_2)$ and $(\mathbf{a}_1/h_1 - \mathbf{a}_3/h_3)$

---

[1] Or, in special cases, the fourth zone.

[2] A very clear introduction to the Miller indices is given by Ewald, *op. cit.*, pp. 20 and 26.

thus lie in this lattice plane, and their vector product is perpendicular to it. For this vector product, we find

$$\begin{aligned}
\left(\frac{1}{h_1}\mathbf{a}_1 - \frac{1}{h_2}\mathbf{a}_2\right) \times \left(\frac{1}{h_1}\mathbf{a}_1 - \frac{1}{h_3}\mathbf{a}_3\right) \\
&= \frac{1}{h_1^2}\mathbf{a}_1 \times \mathbf{a}_1 - \frac{1}{h_2 h_1}\mathbf{a}_2 \times \mathbf{a}_1 - \frac{1}{h_1 h_3}\mathbf{a}_1 \times \mathbf{a}_3 + \frac{1}{h_2 h_3}\mathbf{a}_2 \times \mathbf{a}_3 \\
&= 0 + \frac{(\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3)}{h_1 h_2 h_3}\left\{ + h_3 \frac{\mathbf{a}_1 \times \mathbf{a}_2}{(\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3)} + h_2 \frac{\mathbf{a}_3 \times \mathbf{a}_1}{(\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3)} + h_1 \frac{\mathbf{a}_2 \times \mathbf{a}_3}{(\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3)}\right\} \\
&= \frac{(\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3)}{h_1 h_2 h_3}\{h_1\mathbf{b}_1 + h_2\mathbf{b}_2 + h_3\mathbf{b}_3\} \\
&= \frac{(\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3)}{h_1 h_2 h_3}\,\mathbf{h}
\end{aligned}$$

Thus the vector $(\mathbf{a}_1/h_1 - \mathbf{a}_2/h_2) \times (\mathbf{a}_1/h_1 - \mathbf{a}_3/h_3)$ which is perpendicular to the lattice plane (h_1, h_2, h_3) is also parallel to $\mathbf{h}$. Hence the planes perpendicular to $\mathbf{h}$ are the lattice planes with the Miller indices h_1, h_2, h_3.

We recall here Bragg's interpretation of the diffraction process at the point lattice. If, in Fig. VII.3.5, a parallel displacement of the diffracted wave vector $\mathbf{k}'/2\pi$ is carried out, we see from Fig. VII.3.7 that the incident wave $\mathbf{k}/2\pi$ is in effect "reflected" at the lattice plane (h_1, h_2, h_3). The spherical waves diffracted by the lattice points of one plane are in phase because of the symmetrical position of $\mathbf{k}$ and $\mathbf{k}'$ and, consequently, reinforce each other (see Fig. VII.3.8). The fact that the waves reflected by two adjoining lattice planes do not annul each other by destructive interference but, instead, reinforce each other is assured by the relation

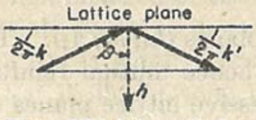

FIG. VII.3.7. "Reflection" at the lattice plane with the Miller indices (h_1, h_2, h_3).

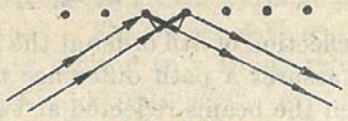

FIG. VII.3.8. For reflection at a lattice plane "adjoining" rays have no path difference.

$$\cos\beta = \frac{\frac{1}{2}|\mathbf{h}|}{\frac{1}{2\pi}|\mathbf{k}|}$$

which may be read off from Fig. VII.3.7. Since $|\mathbf{k}| = 2\pi/\lambda$, we may deduce from this

$$2\frac{1}{|\mathbf{h}|}\cos\beta = \lambda \qquad \text{(VII.3.19)}$$

Furthermore,

$$\frac{1}{|\mathbf{h}|} = \frac{1}{n|\mathbf{h}^*|} = \frac{1}{n} d_{h_1^* h_2^* h_3^*}$$

where n is the largest common denominator of the Miller indices $h_1 = nh_1^*$, $h_2 = nh_2^*$, $h_3 = nh_3^*$ and where the absolute value of the vector $\mathbf{h}^* = h_1^*\mathbf{b}_1 + h_2^*\mathbf{b}_2 + h_3^*\mathbf{b}_3$ formed with the reduced Miller indices h_1^*, h_2^*, h_3^* is equal to the reciprocal of the separation $d_{h_1^* h_2^* h_3^*}$ of two adjoining lattice planes.[1] Thus (VII.3.19) becomes the well-known Bragg reflection condition

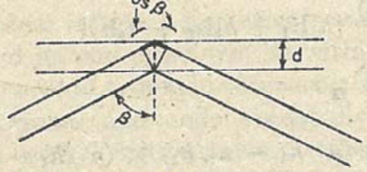

FIG. VII.3.9. For reflection at two adjoining lattice planes the path difference is $2d \cos \beta$.

$$2 \cdot d_{h_1^* h_2^* h_3^*} \cdot \cos \beta = n\lambda \qquad \text{(VII.3.20)}$$

for a reflection of nth order at the lattice planes (h_1, h_2, h_3). This condition assures a path difference $n\lambda$ and hence mutual reinforcement between the beams reflected at two successive lattice planes (see Fig. VII.3.9).

f. The Results of the Brillouin Approximation for Weakly Bound Electrons

We have discussed in some detail the diffraction phenomena at a three-dimensional point lattice and will now compare our results with those of Brillouin's approximation for weakly bound electrons.[2] Brillouin finds:

---

[1] For a proof see, e.g., Ewald, *op. cit.*, pp. 249–250. For a theorem of number theory used by Ewald without proof, see, e.g., B. L. van der Waerden, "Moderne Algebra," part 1, p. 61, Springer-Verlag OHG, Berlin, 1937, or Arnold Scholz, "Einführung in die Zahlentheorie," vol. 1131, p. 22, Sammlung Göschen, 1939.

[2] At the beginning of §3 and in Fig. VII.3.1, we have indicated that Brillouin's approximation is concerned with electrons of high total energy. Since, however, an electron in a region of potential energy which is constant in space behaves like a free electron, the characteristic properties of bound electrons result from spatial variations of potential energy. Accordingly, weakly bound electrons are electrons in a nearly constant potential field. In fact, the decisive assumption in Brillouin's approximation is that the variable part of the potential energy may be regarded as a small perturbation.

1. The eigenfunction of a weakly bound electron is, in general, practically a plane wave $e^{j\mathbf{k}\cdot\mathbf{r}}$.

2. The energy $E(\mathbf{k})$ depends on the wave vector $\mathbf{k}$ practically in the same manner as for a free electron, i.e., according to the law[1]

$$E(\mathbf{k}) = \frac{\hbar^2}{2m}|\mathbf{k}|^2 = \frac{\hbar^2}{2m}(k_x^2 + k_y^2 + k_z^2)$$

See Fig. VII.3.10*a*.

3. The perturbation by the variable part of the potential energy provides the plane wave $e^{j\mathbf{k}\cdot\mathbf{r}}$ with a modulation factor $u(\mathbf{r}; \mathbf{k})$ with the periodicity of the lattice. If the wave vector $\mathbf{k}$ of the free electron[2] is used, this modulation factor has small amplitudes unless $\mathbf{k}$ has a value for which considerations of wave optics indicate a Bragg reflection. For such $\mathbf{k}$ values the eigenfunction is represented no longer simply by a plane wave $e^{j\mathbf{k}\cdot\mathbf{r}}$, but by a superposition of the incident and the diffracted or "reflected" wave. Thus the eigenfunction assumes the character of a standing wave in a direction perpendicular to the reflecting lattice plane. In case $\mathbf{k}$ is perpendicular to the reflecting lattice plane, the eigenfunction becomes a standing wave entirely.

4. The modification of the variation of $E(\mathbf{k})$ by the variable part of the potential energy is, in general, slight. The change with respect to Fig. VII.3.10*a* is small. For $\mathbf{k}$ values leading to a Bragg reflection (i.e., for the Brillouin zone boundaries), discontinuities occur in the otherwise continuous variation $E(\mathbf{k})$ (see Fig. VII.3.10*b*). This arises from the fact that the incident and the reflected waves $e^{j\mathbf{k}\cdot\mathbf{r}}$ and $e^{j\mathbf{k}'\cdot\mathbf{r}}$ have the same wavelength, so that $|\mathbf{k}| = |\mathbf{k}'|$ and that hence the unper-

---

[1] This dependence is found immediately if the formula $\psi(x) = Ae^{j\mathbf{k}\cdot\mathbf{r}}$ is substituted in the Schrödinger equation

$$-\frac{\hbar^2}{2m}\Delta\psi - E\psi = 0$$

of the free electron:

$$-\frac{\hbar^2}{2m}\cdot j^2\cdot(k_x^2 + k_y^2 + k_z^2)A\,e^{j\mathbf{k}\cdot\mathbf{r}} - EA\,e^{j\mathbf{k}\cdot\mathbf{r}} = 0$$

$$E = +\frac{\hbar^2}{2m}|\mathbf{k}|^2 = \frac{\hbar^2}{2m}(k_x^2 + k_y^2 + k_z^2)$$

[2] It is of course also permissible to employ instead of $\mathbf{k}$ a wave vector $\mathbf{k} + 2\pi\mathbf{h}$, where $\mathbf{h}$ is a vector of the reciprocal lattice [see p. 187, in particular footnote 1, and Eq. (VII,3.13)]. In the one-dimensional case, this corresponds to a transition from k to $k + 2\pi h/a$ with $h = \pm 1, \pm 2, \ldots$ The modulation factor then changes over into a modulation factor

$$u(\mathbf{r}; \mathbf{k} + 2\pi\mathbf{h}) = u(\mathbf{r}; \mathbf{k})\,e^{-j2\pi\mathbf{h}\cdot\mathbf{r}}$$

(see also p. 173).

turbed energy eigenvalues $\hbar^2|\mathbf{k}|^2/(2m)$ and $\hbar^2|\mathbf{k}'|^2/(2m)$ are equal to each other. The incident and reflected waves are hence degenerate, and the common unperturbed eigenvalue is split when the perturbation by the variable part of the potential energy is considered. The perturbation produces a forbidden band in the gap-free continuum of the unperturbed energy values. The calculation teaches the following regarding the magnitude of the splitting. In view of its periodicity with the lattice constant a, the lattice potential $U(\mathbf{r})$ may be expanded in a Fourier series

$$U(x) = \sum_{h=-\infty}^{h=+\infty} U_h \, e^{j2\pi \frac{h}{a} x} \qquad \text{(VII.3.21)}$$

[For a three-dimensional lattice, (VII.3.21) is replaced by a threefold

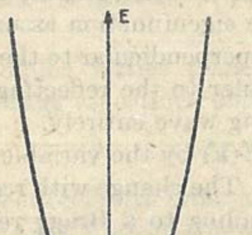

FIG. VII.3.10*a*. The variation of the energy with the wave number *k* for the free electron.

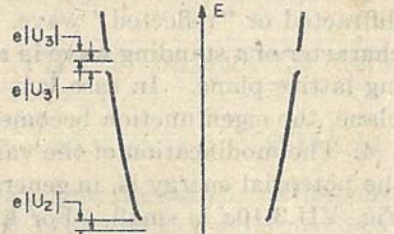

FIG. VII.3.10*b*. The development of forbidden energy regions as result of Bragg reflections.

Fourier expansion.] The split-up energy values are grouped symmetrically about the unperturbed energy value at a distance which is equal to $e|U_h|$ (see Fig. VII.3.10*b*). This makes it particularly plain that the deviation of the electron in a crystal from a free electron results from the variable components in the lattice potential.

5. If the Brillouin approximation is carried out one-dimensionally, i.e., for a linear point lattice with lattice constant a in one-dimensional space, the forbidden energy regions occur at the points $k = \pm\pi/a$, $\pm 2\pi/a$ . . . (see Fig. VII.3.10*b*). We may refer in this one-dimensional case to Fig. VII.3.6, provided that we limit ourselves to the horizontal axis of this figure and to the vertical zone boundaries. Since the zone boundaries occur at $\pm 1/2a$, $\pm 2/2a$, $\pm 3/2a$ . . . ,

Bragg reflections occur if

$$\frac{1}{2\pi}k = \pm h \cdot \frac{1}{2}\frac{1}{a}$$

or

$$k = \pm h\frac{\pi}{a}$$

We see particularly clearly from this simple example that the Brillouin zones constitute a complete subdivision of the range of variation of the independent variable k, whereas the expression "energy bands" directs attention to the existence of forbidden bands or gaps in the range of variation of the dependent variable E.

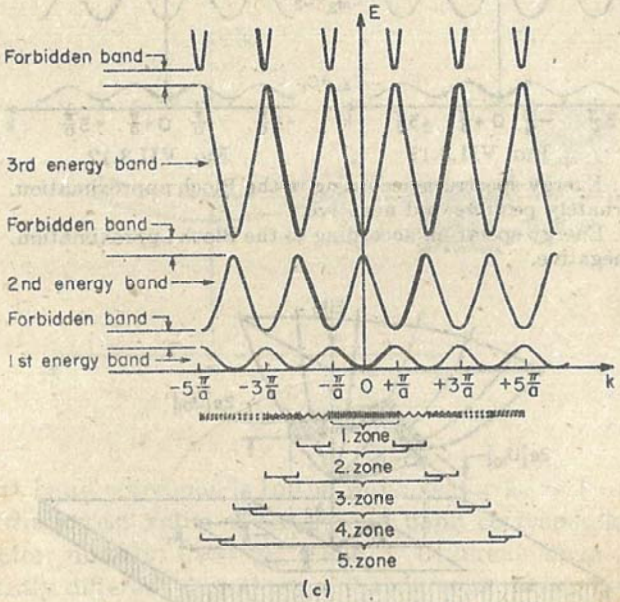

FIG. VII.3.10c. Energy spectrum according to the Brillouin approximation.

6. We have already mentioned, in footnote 2, page 193, that we may employ, in place of the wave number k of the free electron, one of the equivalent wave numbers $k + 2\pi h/a$ with $h = \pm 1, \pm 2, \ldots$. This leaves the eigenfunction $u(x; k)\, e^{jkx}$ and the eigenvalue E unaltered, although the modulation factor with lattice periodicity, $u(x; k)$, changes over into $u(x; k)\, e^{-2\pi jhx/a}$. Hence the variation of $E(k)$ may also be plotted as in Fig. VII.3.10c. This even becomes the natural course when carrying out consistently the degeneracy perturbation calculation required in the neighborhood of a Brillouin zone boundary.[1]

---

[1] To arrive at the representation in Fig. VII.3.10b, it is necessary to suppress one of the eigenvalues supplied by the perturbation calculation at a time, which is really not at all justified.

The comparison with the Bloch approximation to be given at a later point (page 198) is also facilitated hereby.

We turn once more to the Brillouin approximation for a quadratic point lattice in order to describe the phenomenon of band overlap, which is so important for the theory of the metals with a valence of 2 (see §11). We choose this two-dimensional example, because **k** space may then be represented in Fig. VII.3.13 as a horizontal (k_x, k_y) plane above which, in a perpendicular direction, the energy $E(\mathbf{k}) = E(k_x, k_y)$ may be plotted. The first Brillouin zone

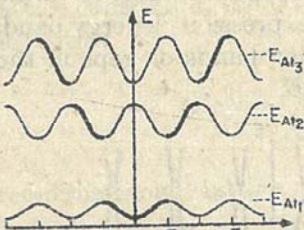

FIG. VII.3.11

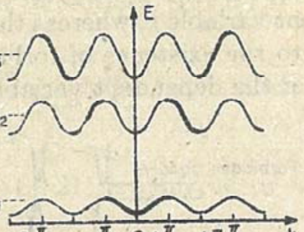

FIG. VII.3.12

FIG. VII.3.11. Energy spectrum according to the Bloch approximation. Exchange integrals alternately positive and negative.

FIG. VII.3.12. Energy spectrum according to the Bloch approximation. Exchange integrals all negative.

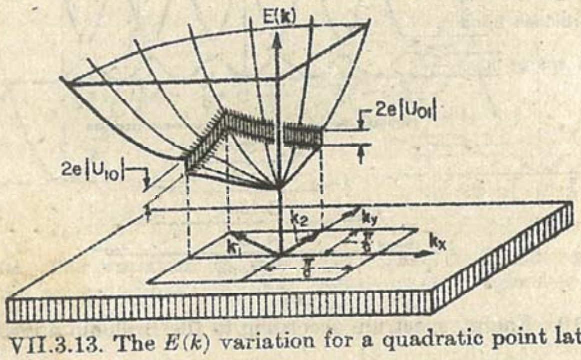

FIG. VII.3.13. The $E(k)$ variation for a quadratic point lattice.

$$-\frac{\pi}{a} \le k_x \le +\frac{\pi}{a} \qquad -\frac{\pi}{a} \le k_y \le +\frac{\pi}{a}$$

is marked out particularly in the horizontal **k** plane. The $E(\mathbf{k})$ surface of a free electron would be a paraboloid of revolution:

$$E(\mathbf{k}) = \frac{\hbar^2}{2m}|\mathbf{k}|^2 = \frac{\hbar^2}{2m}(k_x^2 + k_y^2)$$

The left rear quarter of this paraboloid has been drawn,[1] and it is shown how the paraboloid is cut apart along the two boundaries $k_x = -\pi/a$ and $k_y = +\pi/a$ of the first Brillouin zone. The edges are bent horizontally, so that at the cut there is a gap twice as large as the corresponding Fourier coefficient of the lattice potential (see page 193, item 4).

The two edges of the cut in the energy surface $E(\mathbf{k})$ are represented in Fig. VII.3.14 by drawing the (k_x, E) plane $k_y = +\pi/a$. It is seen here that the highest value of the first band lies, for not too large $|U_{01}|$, higher than the lowest energy value of the second band. The bands overlap. It is true that, according to Fig. VII.3.13, the highest value of the first band corresponds to the wave vector $\mathbf{k}_1 = (-\pi/a, +\pi/a)$, whereas the lowest value of the second band corresponds to another wave vector, namely, $\mathbf{k}_2 = (0, +\pi/a)$. Different wave vectors, and consequently different directions of the electron waves, belong in general to the same energy values in the first and the second band. Thus, in the case of band overlap, the electron can pass from one energy band to the next higher one without energy uptake. At the same time, it must change its direction. This, however, happens at any rate very frequently as the result of thermal collisions (see Chap. VII, §9).

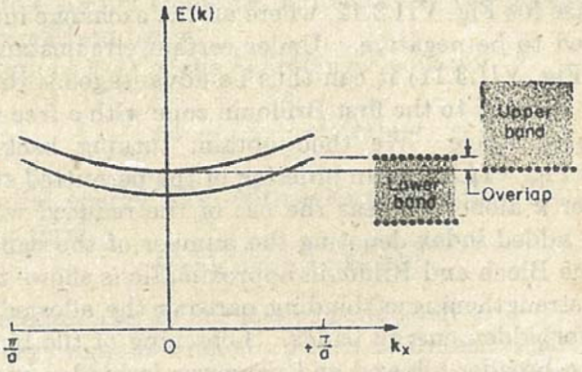

FIG. VII.3.14. Band overlap.

---

[1] Our spatial representation of this two-dimensional example thus corresponds to Fig. VII.3.10b of the one-dimensional case. However, this should not lead to the erroneous conclusion that the possibility of the representation in Fig. VII.3.10c is a peculiarity of the one-dimensional case. It was pointed out already on p. 181 and in Fig. VII.2.9 that the transition from the wave vector $\mathbf{k}$ to a wave vector $\mathbf{k} + 2\pi\mathbf{h}$, without change in the eigenfunction or eigenvalue, is possible also in multidimensional cases.

g. Comparison of the Bloch and Brillouin Approximations. Reduced and Free Wave Vector

We return to Fig. VII.2.7 and retrace it in completed form in Figs. VII.3.11 and VII.3.12. In doing this, we do not reduce the wave number k to the interval $-\pi/a < k \leq +\pi/a$ and draw, furthermore, the variation $E(k)$ for several successive atomic eigenvalues E_{at}. In Fig. VII.3.11 it is assumed that the exchange integrals belonging to the eigenfunctions of successive atomic states are alternately negative and positive. With these—rather artificial—assumptions, the variation $E(k)$ according to Bloch (Fig. VII.3.11) attains a certain similarity with the results of the Brillouin approximation (Fig. VII.3.10*c*). This is not the case for Fig. VII.3.12, where all the exchange integrals have been assumed to be negative. Under certain circumstances (e.g., in the case of Fig. VII.3.11) it can thus be advantageous to replace the wave vector reduced to the first Brillouin zone with a free wave vector with unlimited range. We then obtain, turning back from Fig. VII.3.10*c* to Fig. VII.3.10*b*, an ordering of the permitted states by the wave number k alone, whereas the use of the reduced wave number demands an added index denoting the number of the band. A comparison of the Bloch and Brillouin approximations shows that in both instances a strengthening of binding narrows the allowed bands and widens the forbidden energy bands. Loosening of the binding leads, vice versa, to broader allowed and narrower forbidden energy bands. In the Bloch approximation, the width of the allowed bands is determined by the exchange integral, which becomes small in absolute value when, as a result of strong binding, the atomic eigenfunctions are concentrated closely about the atomic cores. In the Brillouin approximation, the widths of the forbidden energy bands are closely related to the Fourier amplitudes of the potential (see page 193, item 4). As we have pointed out in a footnote on page 192, the binding of the electrons is determined by the variable component of the potential. Thus, strong binding demands large alternating amplitudes of the periodic potential, and these large Fourier amplitudes lead to wide forbidden bands.

A comparison of Figs. VII.3.10*c* and VII.3.11 shows that the Bloch and Brillouin approximations lead to the same subdivision of the k axis in the one-dimensional case. This is true also for many more complex three-dimensional lattices. Thus, for the face-centered cubic and the body-centered cubic lattices, Bloch's periodicity polyhedron and the first Brillouin zone are identical. We will see, however, at a later point (§11) from a lattice with basis, the diamond lattice, that this identity is by no means universal.

§4. General Statements about the Eigenfunctions and the Energy Spectrum of an Electron in the Periodic Potential Field

a. Agreements in the Results of Bloch and Brillouin

In the last two paragraphs we have treated the special cases of strongly bound electrons and nearly free electrons by approximation methods. In spite of quantitative differences, we arrived at a number of qualitative results common to both cases. In fact, the most important qualitative properties of the eigenfunction and of the energy spectrum follow simply from the periodicity of the potential energy. They are summarized below.

1. The solutions $\psi(\mathbf{r})$ of the Schrödinger equation

$$-\frac{\hbar^2}{2m}\Delta\psi - (E + eU(\mathbf{r}))\psi(\mathbf{r}) = 0 \qquad \text{(VII.4.01)}$$

with a potential energy $U(\mathbf{r})$ with lattice periodicity can always be written in the form

$$\psi(\mathbf{r};\mathbf{k}) = u(\mathbf{r};\mathbf{k})\, e^{j\mathbf{k}\cdot\mathbf{r}} \qquad \text{(VII.4.02)}$$

where the function $u(\mathbf{r};\mathbf{k})$ has lattice periodicity.[1]

2. The energy scale is broken up into a number of allowed bands, which are separated by forbidden bands. If E lies in an allowed band, the corresponding ψ has a real $\mathbf{k}$ value. ψ then represents a plane wave with an amplitude modulated with lattice periodicity. $\mathbf{k}$ attains the meaning of a wave number and indicates the number of wavelengths λ in a distance 2π:

$$|\mathbf{k}| = \frac{2\pi}{\lambda} \qquad \text{(VII.4.03)}$$

If E lies in a forbidden band, the corresponding ψ has a complex $\mathbf{k}$ value. ψ then increases or decreases exponentially. The fulfillment of periodic boundary conditions is not possible.[2]

3. A wave vector $\mathbf{k} + 2\pi\mathbf{h}$ leads to the same eigenfunction and hence the same energy E as the wave vector $\mathbf{k}$. $\mathbf{h}$ is here a vector of the reciprocal lattice (see page 187, in particular footnote 1). On the basis of this fact we obtain a subdivision of $\mathbf{k}$ space (or the $\mathbf{k}$ plane or $\mathbf{k}$ axis for two- and one-dimensional problems) into zones within which the eigenfunctions ψ and the eigenvalues E repeat themselves periodically (see, e.g., Figs. VII.2.9, VII.3.10*c*, VII.3.11, VII.3.12).

---

[1] G. Floquet, *Ann. école norm.*, **12:** 47 (1883).

[2] H. A. Kramers, *Physica*, **2:** 483 (1935).

4. Frequently, but not always, the energy shows a variation of the type

$$E = E_{\text{bound}} \mp \text{const} \cdot |\mathbf{k} - \mathbf{k}_{\text{bound}}|^2 \qquad \text{(VII.4.04)}$$

at the upper and the lower boundaries of the energy band.

5. An energy value E, which is forbidden for a particular direction of propagation $\mathbf{k}$, may be allowed for another direction of propagation $\mathbf{k}'$. Thus, if we pass from a consideration of a particular direction of propagation to a consideration of all directions of propagation, the width of the forbidden bands is reduced. This effect can go so far that an energy band forbidden for a particular direction of propagation vanishes when all directions of propagation are considered. We then speak of band overlap.

6. Increasing the binding of the electrons to the atomic cores broadens the forbidden and narrows the allowed energy bands. Weakening of the binding, vice versa, broadens the allowed and narrows the forbidden energy bands.

These are statements which follow quite generally simply from the lattice periodicity of the potential energy $-eU(\mathbf{r})$. It would of course be valuable if, for given $U(\mathbf{r})$, these general statements could be extended, e.g., by indicating in (VII.4.02) the amplitude modulation $u(\mathbf{r}; \mathbf{k})$ or, in (VII.4.04), the values of E_{bound} and of const. Neither the Bloch nor the Brillouin approximation has proved suitable for this in cases of physical interest. The electrons responsible for conduction phenomena in actual solid bodies are neither strongly bound nor almost free, but lie just midway between these two cases. For the attainment of quantitative statements, the cellular method of Wigner and Seitz[1] has proved most successful so far.

b. The Cellular Method of Wigner and Seitz

Wigner and Seitz treat the potential fields about the individual atomic cores as spherically symmetric. They can also adduce a series of reasons why this assumption does not depart greatly from reality. They then subdivide space within the lattice into polyhedral cells about each atomic core by erecting bisecting perpendicular planes on the lines joining an atom core to its nearest neighbors and eventually also to its next-to-nearest neighbors. Within one such cell the spherically symmetrical potential variation is obtained from experimental data or from Hartree tabulations, the Schrödinger equation is separated in spherical coordinates, and the radial part of the eigenfunction is determined by numerical methods.

---

[1] E. Wigner and F. Seitz, *Phys. Rev.*, **43**: 804 (1933).

For the isolated atom the boundary conditions follow from the requirement that the eigenfunction must not become infinitely large either at the origin or at infinite distance. For the atom in the lattice, only the boundary condition at the origin remains unaltered. The boundary condition at infinity is replaced by the requirement that the eigenfunction must pass over continuously into the next cell at the cell boundary. To permit the practical realization of the procedure, this boundary condition is fulfilled only at the centers of the polyhedral faces bounding the cell.

The cellular method, particularly as perfected by Slater,[1] has had remarkable success with monovalent metals. The difficulties increase rapidly if there are several valence electrons per atom. Further efforts have been made to extend the method for this purpose.[2] However, no results which are reliable quantitatively as well as qualitatively have become available up to the present for the valence lattices of the diamond type with four valence electrons per atom (C, Si, Ge, Sn), which are of particular interest for semiconductor physics.[3] Also for lattices with two kinds of atoms the quantitative results are not comparable with those obtained for monovalent and divalent metals.[4]

§5. Mean Momentum, Mean Velocity, and Mean Current of an Electron in a Crystal

The Schrödinger functions

$$\psi(\mathbf{r};\,\mathbf{k}) = u(\mathbf{r};\,\mathbf{k})\,e^{j\mathbf{k}\cdot\mathbf{r}} \tag{VII.5.01}$$

are eigenfunctions of the operator $-\hbar^2\Delta/(2m) + E_{\text{pot}}$ of the total energy. Since the momentum operator

$$\mathbf{p}_{\text{op}} = \frac{\hbar}{j}\,\text{grad} = \frac{\hbar}{j}\left\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right\}$$

does not, in general, commute with the operator of the total energy, they cannot at the same time be eigenfunctions of the momentum.[5]

---

[1] J. C. Slater, *Phys. Rev.*, **45**: 794 (1934).

[2] C. Herring and A. G. Hill, *Phys. Rev.*, **58**: 132 (1940).

[3] For a summary and bibliography of more recent work in this field, see F. Herman, *Proc. IRE*, **43**: 1703 (1955).

[4] See footnote 2, p. 163.

[5] As such, they would have to be pure exponentials $e^{j\mathbf{k}\cdot\mathbf{r}}$. The factor $u(\mathbf{r};\,\mathbf{k})$ with lattice periodicity is thus here in the way. Only for a free electron we have $u(\mathbf{r};\,\mathbf{k}) = \text{const} = A$, so that $\psi(\mathbf{r}) = A\,e^{j\mathbf{k}\cdot\mathbf{r}}$ is an eigenfunction of both the total

Thus, an electron represented by (VII.5.01) possesses a sharply defined total energy $E(\mathbf{k})$, but no sharply defined momentum. Accordingly, only a quantum-mechanical average of the momentum can be given, which must be calculated by the rule

$$\mathbf{p} = \int\limits_{V_{\text{fund}}} \psi^*(\mathbf{r};\, \mathbf{k})\mathbf{p}_{\text{op}}\psi(\mathbf{r};\, \mathbf{k})\, dV = \frac{\hbar}{j} \int\limits_{V_{\text{fund}}} \psi^*(\mathbf{r};\, \mathbf{k}) \text{ grad } \psi(\mathbf{r};\, \mathbf{k})\, dV \tag{VII.5.02}$$

The evaluation of (VII.5.02) requires an explicit knowledge of the eigenfunctions (VII.5.01). Fortunately, however, it is possible to transform (VII.5.02) into

$$\mathbf{p} = \frac{m}{\hbar} \operatorname{grad}_{\mathbf{k}} E(\mathbf{k}) \tag{VII.5.03}$$

or

$$p_x = \frac{m}{\hbar} \frac{\partial}{\partial k_x} E(k_x, k_y, k_z) \qquad p_y = \frac{m}{\hbar} \frac{\partial}{\partial k_y} E(k_x, k_y, k_z)$$

$$p_z = \frac{m}{\hbar} \frac{\partial}{\partial k_z} E(k_x, k_y, k_z)$$

Herewith the mean momentum can be computed simply from the variation of the energy E with the wave number $\mathbf{k}$. This variation $E(\mathbf{k})$ is at least qualitatively known (see, e.g., Figs. VII.3.10*c*, VII.3.11, and VII.3.12), so that qualitative statements regarding the mean momentum belonging to a state (VII.5.01) are possible.

The derivation of (VII.5.03) from (VII.5.02) demands, unfortunately, a certain amount of calculation. We shall make it somewhat more general than is required for the present purpose.[1] The point of departure is the Schrödinger equation (VII.1.01)

$$\left[-\frac{\hbar^2}{2m} \Delta - eU(\mathbf{r}) - E(\mathbf{k}) \right] \psi(\mathbf{r};\, \mathbf{k}) = 0 \tag{VII.5.04}$$

---

energy and the momentum. This is possible in this special case because for the free electron, with $E_{\text{pot}} = \text{const}$, the operators of the total energy and of the momentum commute. For a free electron a state of sharply defined total energy also happens to be a state of sharply defined momentum. See in this connection, e.g., E. Fues in Wien and Harms, "Handbuch der Experimentalphysik, Ergänzungswerk," p. 212, Akademische Verlagsgesellschaft, Leipzig, 1935, and H. A. Kramers, "Die Grundlagen der Quantentheorie" in Eucken and Wolf, "Hand- und Jahrbuch der Chemischen Physik," pp. 138 and 162, Akademische Verlagsgesellschaft, Leipzig, 1938.

[1] The general formula (VII.5.08) is employed, e.g., at the very top of p. 186 in W. V. Houston, *Phys. Rev.*, 57: 184 (1940).

which is differentiated with respect to k_x:

$$\left[-0 + 0 - \frac{\partial}{\partial k_x} E\right] \psi(\mathbf{r}; \mathbf{k}) + \left[-\frac{\hbar^2}{2m} \Delta - eU(\mathbf{r}) - E(\mathbf{k})\right] \frac{\partial}{\partial k_x} \psi(\mathbf{r}; \mathbf{k}) = 0$$

Application of (VII.5.01) leads to

$$-\frac{\partial E}{\partial k_x} \psi(\mathbf{r}; \mathbf{k}) + \left[-\frac{\hbar^2}{2m} \Delta - eU(\mathbf{r}) - E(\mathbf{k})\right] \cdot \left[jxu(\mathbf{r}; \mathbf{k})\, e^{j\mathbf{k}\cdot\mathbf{r}} + \frac{\partial u(\mathbf{r}; \mathbf{k})}{\partial k_x} e^{j\mathbf{k}\cdot\mathbf{r}}\right] = 0$$

We now multiply from the left with $\psi^*(\mathbf{r}; \mathbf{k}')$ and integrate over the fundamental domain[1]

$$-\frac{\partial E}{\partial k_x} \int\limits_{V_{\text{fund}}} \psi^*(\mathbf{r}; \mathbf{k}')\psi(\mathbf{r}; \mathbf{k})\, dV + \int\limits_{V_{\text{fund}}} \psi^*(\mathbf{r}; \mathbf{k}') \left[-\frac{\hbar^2}{2m} \Delta - eU(\mathbf{r}) - E(\mathbf{k})\right] \cdot \left[jx\psi(\mathbf{r}; \mathbf{k}) + \frac{\partial u(\mathbf{r}; \mathbf{k})}{\partial k_x} \cdot e^{j\mathbf{k}\cdot\mathbf{r}}\right] dV = 0$$

In the first integral we utilize the orthogonality of the eigenfunctions. The second integral is split into two integrals:

$$\left.\begin{array}{l} -\dfrac{\partial E}{\partial k_x} \cdot \delta_{\mathbf{k}\mathbf{k}'} \\ + \displaystyle\int\limits_{V_{\text{fund}}} \psi^*(\mathbf{r}; \mathbf{k}') \left[-\frac{\hbar^2}{2m} \Delta - eU(\mathbf{r}) - E(\mathbf{k})\right] jx\psi(\mathbf{r}; \mathbf{k})\, dV \\ + \displaystyle\int\limits_{V_{\text{fund}}} \psi^*(\mathbf{r}; \mathbf{k}') \left[-\frac{\hbar^2}{2m} \Delta - eU(\mathbf{r}) - E(\mathbf{k})\right] \frac{\partial u(\mathbf{r}; \mathbf{k})}{\partial k_x} e^{j\mathbf{k}\cdot\mathbf{r}}\, dV \end{array}\right\} = \left\{\begin{array}{c} \text{I} \\ +\text{II} \\ +\text{III} \end{array}\right\} = 0 \quad \text{(VII.5.05)}$$

We first treat the third term of the sum, in which we can

---

[1] Of course, $\mathbf{k}$ and $\mathbf{k}'$ do not pertain to an incident and diffracted wave, as in §3, but are arbitrary wave vectors. After all, there is no diffracted wave for arbitrary $\mathbf{k}$.

make use of the Hermitian character of the operator of the total energy $[-\hbar^2\Delta/(2m) - eU(\mathbf{r}) - E(\mathbf{k})]$, since both $\psi^*(\mathbf{r};\ \mathbf{k})$ and $[\partial u(\mathbf{r};\ \mathbf{k})/\partial k_x]\ e^{j\mathbf{k}\cdot\mathbf{r}}$ are periodic in the fundamental domain:[1]

$$\mathrm{III} = \int\limits_{V_{\text{fund}}} \frac{\partial u(\mathbf{r};\ \mathbf{k})}{\partial k_x}\, e^{j\mathbf{k}\cdot\mathbf{r}} \left[-\frac{\hbar^2}{2m}\,\Delta - eU(\mathbf{r}) - E(\mathbf{k}) \right] \psi^*(\mathbf{r};\ \mathbf{k}')\, dV$$

With the Schrödinger equation (VII.5.04), with $\mathbf{k}'$ replacing $\mathbf{k}$, we obtain

$$\mathrm{III} = \int\limits_{V_{\text{fund}}} \frac{\partial u(\mathbf{r};\ \mathbf{k})}{\partial k_x}\, e^{j\mathbf{k}\cdot\mathbf{r}}[E(\mathbf{k}') - E(\mathbf{k})]\psi^*(\mathbf{r};\ \mathbf{k}')\, dV$$

and, furthermore, with Eq. (VII.5.01) in which, again, $\mathbf{k}$ has been replaced by $\mathbf{k}'$,

$$\mathrm{III} = \int\limits_{V_{\text{fund}}} \frac{\partial u(\mathbf{r};\ \mathbf{k})}{\partial k_x}\, e^{j\mathbf{k}\cdot\mathbf{r}}[E(\mathbf{k}') - E(\mathbf{k})]u^*(\mathbf{r};\ \mathbf{k}')\, e^{-j\mathbf{k}'\cdot\mathbf{r}}\, dV \qquad \text{(VII.5.06)}$$

We now turn to the second term in (VII.5.05). Since $jx\psi(\mathbf{r};\ \mathbf{k})$ is not periodic in the fundamental domain, we cannot make use of the Hermitian character of the operator $[-\hbar^2\Delta/(2m) - eU(\mathbf{r}) - E(\mathbf{k})]$. We note instead that

$$\Delta(x\cdot\psi(\mathbf{r};\ \mathbf{k})) = x\cdot\Delta\psi(\mathbf{r};\ \mathbf{k}) + 2\cdot 1\cdot\frac{\partial}{\partial x}\,\psi(\mathbf{r};\ \mathbf{k}) + 0$$

---

[1] It is not always noted that the validity of the commutation rule for a Hermitian operator depends on certain properties of the commuted functions. Thus the momentum operator $p_{\text{op}} = (\hbar/j)d/dx$ is Hermitian. However, the relation

$$\int f^* p_{\text{op}} g\, dx = \int g p_{\text{op}}^* f^*\, dx$$

rests on an integration by parts

$$\int\limits_{x=x_1}^{x=x_2} f^* \frac{\hbar}{j}\frac{d}{dx}\, g\, dx = \frac{\hbar}{j}\,[f^*g]_{x=x_1}^{x=x_2} + \int\limits_{x=x_1}^{x=x_2} g\left(\frac{\hbar}{-j}\frac{d}{dx}f^*\right) dx$$

It applies in the simple form here employed only if $[f^*(x_2)g(x_2) - f^*(x_1)g(x_1)]$ vanishes. This is true in most quantum-mechanical problems, since we consider here only functions f and g which vanish at the limits x_1 and x_2. In problems with a periodic potential field, this difference vanishes for a different reason, namely, the periodicity in the fundamental domain $x_1 = -Ga/2$, $x_2 = +Ga/2$:

$$f(x_2) = f\left(+\frac{G}{2}a\right) = f\left(-\frac{G}{2}a\right) = f(x_1)$$

$$g(x_2) = g\left(+\frac{G}{2}a\right) = g\left(-\frac{G}{2}a\right) = g(x_1)$$

$$f^*(x_2)g(x_2) - f^*(x_1)g(x_1) = 0$$

and hence

$$\begin{aligned}\left[-\frac{\hbar^2}{2m}\Delta - eU(\mathbf{r}) - E(\mathbf{k})\right] x\psi(\mathbf{r};\mathbf{k})& \\ &= x\cdot\left[-\frac{\hbar^2}{2m}\Delta - eU(\mathbf{r}) - E(\mathbf{k})\right]\psi(\mathbf{r};\mathbf{k}) - \frac{\hbar^2}{2m}\cdot 2\frac{\partial}{\partial x}\psi(\mathbf{r};\mathbf{k}) \\ &= x\cdot \qquad\qquad 0 \qquad\qquad - \frac{\hbar^2}{m}\frac{\partial}{\partial x}\psi(\mathbf{r};\mathbf{k})\end{aligned}$$

The last is obtained with the aid of (VII.5.04). If we use this simple expression in the integrand of II, we find

$$\begin{aligned}\mathrm{II} &= \int\limits_{V_{\mathrm{fund}}} \psi^*(\mathbf{r};\mathbf{k}')j\left(-\frac{\hbar^2}{m}\frac{\partial}{\partial x}\psi(\mathbf{r};\mathbf{k})\right)dV \\ &= \frac{\hbar}{m}\int\limits_{V_{\mathrm{fund}}} \psi^*(\mathbf{r};\mathbf{k}')\,\frac{\hbar}{j}\frac{\partial}{\partial x}\psi(\mathbf{r};\mathbf{k})\,dV \\ &= \frac{\hbar}{m}\int\limits_{V_{\mathrm{fund}}} \psi^*(\mathbf{r};\mathbf{k}')p_{x_{\mathrm{op}}}\psi(\mathbf{r};\mathbf{k})\,dV = \frac{\hbar}{m}p_{x\mathbf{k}'\mathbf{k}} \qquad \text{(VII.5.07)}\end{aligned}$$

Here $p_{x\mathbf{k}'\mathbf{k}}$ is the so-called matrix element of the x component of the momentum operator (VII.5.02).

Equations (VII.5.05), (VII.5.06), and (VII.5.07) together yield

$$\begin{aligned}-\frac{\partial E}{\partial k_x}\cdot\delta_{\mathbf{k}\mathbf{k}'} + \frac{\hbar}{m}p_{x\mathbf{k}'\mathbf{k}} + [E(\mathbf{k}') - E(\mathbf{k})]\cdot& \\ \int\limits_{V_{\mathrm{fund}}} u^*(\mathbf{r};\mathbf{k}')\,\frac{\partial u(\mathbf{r};\mathbf{k})}{\partial k_x}\,e^{j(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}}\,dV &= 0 \qquad \text{(VII.5.08)}\end{aligned}$$

For the present occasion, we need only this equation for $\mathbf{k} = \mathbf{k}'$. It then yields immediately

$$-\frac{\partial E}{\partial k_x} + \frac{\hbar}{m}p_{x\mathbf{k}\mathbf{k}} + 0 = 0$$

i.e., the x component of the desired relation (VII.5.03)

$$p_x = p_{x\mathbf{k}\mathbf{k}} = \frac{m}{\hbar}\frac{\partial E}{\partial k_x}$$

Herewith the valuable relation (VII.5.03) has been proved. We shall now proceed to its physical evaluation.

In general, the total energy E will depend on a wave vector component, such as k_x, in the manner shown by Figs. VII.3.10c to VII.3.12. Hence, we deduce from (VII.5.03) that the mean value of the corresponding momentum component vanishes at the band edges.[1] The fact, represented in Fig. VII.2.5, that the wave function becomes here a standing wave—at least in the x direction—is in agreement with this.

The curve of the energy shows its steepest inclination near the centers of the bands. These states have, therefore, by (VII.5.03), the largest momentum values.

The mean momentum $\mathbf{p}$ is related to the mean velocity $\mathbf{v}$ by

$$\mathbf{v} = \frac{1}{m}\mathbf{p} \tag{VII.5.09}$$

since the velocity operator $\hbar/jm$ grad and the momentum operator $\hbar/j$ grad differ only by the factor $1/m$. Thus we find from (VII.5.03) for the mean velocity

$$\mathbf{v} = \frac{1}{\hbar}\,\mathrm{grad}_{\mathbf{k}}\,E(\mathbf{k}) \tag{VII.5.10}$$

This rather abstract discussion attains a certain graphic significance from the identity of the mean velocity $\mathbf{v}$ for a state (VII.5.01) with the group velocity of a wave packet formed by neighboring states centered about the state (VII.5.01). This identity may be demonstrated with the aid of (VII.5.10). The time factor for a state (VII.5.01) is $e^{-(j/\hbar)E(\mathbf{k})\cdot t}$, so that the frequency $(1/\hbar)E(\mathbf{k})$ is no longer simply proportional to the wave number $\mathbf{k}$. There is, hence, dispersion.[2] In a one-dimensional example the group velocity is then

$$\frac{\partial\omega}{\partial k} = \frac{1}{\hbar}\frac{d}{dk}E$$

By (VII.5.10) this is identical with the quantum-mechanical mean value

$$\frac{1}{m}\mathbf{p} = \frac{1}{m}\int\limits_{V_{\text{fund}}} \psi^*(\mathbf{r};\mathbf{k})\,\frac{\hbar}{j}\,\mathrm{grad}\,\psi(\mathbf{r};\mathbf{k})\,dV$$

---

[1] Since we are dealing with a quantum-mechanical mean value, any particular measurement can yield a value for the momentum component in question which differs from zero even for a state at the band edge. On the average, for many measurements, however, equally large positive and negative values must occur with equal frequency. The quantum-mechanical mean value of the kinetic energy for a state at the band edge is, on the other hand, by no means zero. Here measured values of p_x^2 are averaged, whereas for the momentum measurements p_x itself is averaged so that positive and negative values can compensate each other. These circumstances are related to the fact that the ψ function which becomes a standing wave at the band edge may be represented by a superposition of equal numbers of plane waves running to the right and to the left.

[2] A. Sommerfeld, "Atombau und Spektrallinien," vol. II, p. 8, Eq. (14), Vieweg-Verlag, Brunswick, Germany, 1944. See also W. Shockley, "Electrons and Holes in Semiconductors," p. 160, Fig. 6.2, D. Van Nostrand Company, Inc., Princeton, N.J., 1950.

All this is familiar for the plane waves of a free electron; however, the validity, here demonstrated, for the lattice-modulated waves of an electron in a crystal is by no means obvious.

To obtain, finally, an expression for the current density corresponding to an electron in a crystal with the eigenfunction

$$\psi(\mathbf{r};\mathbf{k}) = u(\mathbf{r};\mathbf{k})\,e^{j\mathbf{k}\cdot\mathbf{r}}$$

we proceed from the familiar expression[1]

$$\frac{\hbar}{2jm}\,(\psi^*\operatorname{grad}\psi - \psi\operatorname{grad}\psi^*)$$

for the flux of probability density of an electron with normalized eigenfunction ψ. Multiplication with the charge $-e$ leads to the current density

$$-\frac{\hbar}{2j}\frac{e}{m}\,(\psi^*\operatorname{grad}\psi - \psi\operatorname{grad}\psi^*)$$

For the stationary solution $\psi(\mathbf{r};\mathbf{k})$ with a probability density $\psi^*(\mathbf{r};\mathbf{k})\psi(\mathbf{r};\mathbf{k})$ independent of time which we shall substitute in this expression, the divergence of the current density vanishes according to the continuity equation.[2] In the one-dimensional case, the current density must hence be constant in space. In two- and three-dimensional cases a solenoidal component may be added, which, however, must have lattice periodicity since the wave factors $e^{j\mathbf{k}\cdot\mathbf{r}}$ and $e^{-j\mathbf{k}\cdot\mathbf{r}}$, which are periodic in the fundamental domain but do not have lattice periodicity, cancel each other in the foregoing expression for the current density. Such solenoidal portions of the current density, which are repeated from unit cell to unit cell, obviously arise from electron orbits about the atom cores if they exist at all. They are of no interest for the macroscopic current density produced by an electron in a crystal. If we cause them to vanish by taking a space average, we get for the mean current density $\mathbf{i}_{\text{single}}$ of a single electron in a crystal

$$\mathbf{i}_{\text{single}} = -\frac{\hbar}{2j}\frac{e}{m}\frac{1}{V_{\text{fund}}}\int\limits_{V_{\text{fund}}}(\psi^*\operatorname{grad}\psi - \psi\operatorname{grad}\psi^*)\cdot dV \qquad \text{(VII.5.11)}$$

By a transformation we obtain

$$\mathbf{i}_{\text{single}} = -\frac{e}{m}\frac{1}{V_{\text{fund}}}\cdot\frac{1}{2}\cdot\int\limits_{V_{\text{fund}}}\left(\psi^*\cdot\frac{\hbar}{j}\operatorname{grad}\psi + \psi\,\frac{\hbar}{-j}\operatorname{grad}\psi^*\right)\cdot dV$$

---

[1] See, e.g., E. Fues in Wien and Harms, "Handbuch der Experimentalphysik, Ergänzungswerk," p. 150, Eq. (5.6).

[2] *Ibid.*, p. 149, Eq. (5.5).

and, furthermore, if we utilize the momentum operator and its Hermitian character,

$$\mathbf{i}_{\text{single}} = -\frac{e}{m}\cdot\frac{1}{V_{\text{fund}}}\cdot\frac{1}{2}\left[\int\limits_{V_{\text{fund}}} \psi^*\cdot\mathbf{p}_{\text{op}}\psi\,dV + \int\limits_{V_{\text{fund}}} \psi\mathbf{p}^*_{\text{op}}\psi^*\,dV\right]$$

$$\mathbf{i}_{\text{single}} = -\frac{e}{m}\cdot\frac{1}{V_{\text{fund}}}\cdot\int\limits_{V_{\text{fund}}} \psi^*\mathbf{p}_{\text{op}}\psi\,dV\,.$$

According to Eq. (VII.5.02), the integral is equal to the quantum-mechanical mean value $\mathbf{p}$ of the momentum. We obtain, therefore,

$$\mathbf{i}_{\text{single}} = -\frac{e}{m}\cdot\frac{1}{V_{\text{fund}}}\cdot\mathbf{p} \qquad \text{(VII.5.12)}$$

If, finally, we make use of (VII.5.09) and (VII.5.10), we obtain

$$\mathbf{i}_{\text{single}} = -\frac{e}{V_{\text{fund}}}\cdot\mathbf{v} = -\frac{e}{V_{\text{fund}}}\frac{1}{\hbar}\,\text{grad}_{\mathbf{k}}\,E(\mathbf{k}) \qquad \text{(VII.5.13)}$$

If we recall that

$$\rho = -\frac{e}{V_{\text{fund}}}$$

is the charge density for uniform distribution of the electron charge $-e$ over the volume V_{fund} of the fundamental domain, we find complete analogy with the classical formula

Current density = charge density · velocity of convection

The same formula could have been derived with the aid of wave packets, for which the velocity of the center of gravity equals the group velocity $\mathbf{v}$ of the $\psi(\mathbf{r};\mathbf{k})$ waves. Such a derivation would, however, fit in with corpuscular notions. We prefer to give, in §9, pages 246 to 247, a derivation of the equation

$$\mathbf{i}_{\text{single}} = -\frac{e}{V}\mathbf{v}$$

from a completely corpuscular standpoint.

§6. The Effect of an External Field on an Electron in a Crystal and the Effective Mass of an Electron in a Crystal

Conductivity questions require a clarification of the action of an added external force on an electron in a crystal. We shall treat this matter first for a free electron.

a. The Free Electron under the Influence of an External Force F

If an electron is accelerated by a force **F**, its energy E increases with time. Hence, the determination of the law of acceleration is not facilitated by a derivation of the stationary states from the time-free Schrödinger equation with the aid of some boundary conditions, a course which is appropriate in a great majority of quantum-mechanical problems. Instead, we must proceed from the time-dependent Schrödinger equation

$$-\frac{\hbar^2}{2m}\Delta\psi + E_{\text{pot}}\,\psi = j\hbar\frac{\partial\psi}{\partial t} \qquad \text{(VII.6.01)}$$

and follow the evolution in time $\psi(\mathbf{r}, t)$ of a given initial state $\psi(\mathbf{r}, 0)$.

If the free electron is subject to no external force, its potential energy is constant in space:

$$E_{\text{pot}} = -eU \qquad \text{(VII.6.02)}$$

and (VII.6.01) is solved by the plane wave

$$\psi(\mathbf{r}, t) = \psi(\mathbf{r}; \mathbf{k}) \cdot e^{-\frac{j}{\hbar}Et} = A\,e^{j\left[\mathbf{k}\cdot\mathbf{r} - \frac{1}{\hbar}Et\right]} \qquad \text{(VII.6.03)}$$

We may check this by substitution of (VII.6.03) in (VII.6.01) and will then find for the dependence of the energy E on the wave number

$$E(\mathbf{k}) = -eU_0 + \frac{\hbar^2}{2m}|\mathbf{k}|^2 \qquad \text{(VII.6.04)}$$

Let, now, the electron be subject to an external force **F** which, for the sake of convenience, we shall assume to be constant in space. It is hence derived as a negative space gradient from a potential $-\mathbf{F}\cdot\mathbf{r}$[1] so that the potential energy E_{pot} is given by

$$E_{\text{pot}} = -eU_0 - \mathbf{F}\cdot\mathbf{r} \qquad \text{(VII.6.05)}$$

instead of by (VII.6.02). Equation (VII.6.01) therefore assumes the

---

[1] Thus, in component notation we have

$$F_x = -\frac{\partial}{\partial x}(-F_x \cdot x - F_y \cdot y - F_z \cdot z)$$

$$F_y = -\frac{\partial}{\partial y}(-F_x \cdot x - F_y \cdot y - F_z \cdot z)$$

$$F_z = -\frac{\partial}{\partial z}(-F_x \cdot x - F_y \cdot y - F_z \cdot z)$$

or, in a single vector equation,

$$\mathbf{F} = -\operatorname{grad}(-\mathbf{F}\cdot\mathbf{r})$$

form

$$-\frac{\hbar^2}{2m}\Delta\psi - eU_0\psi - \mathbf{F}\cdot\mathbf{r}\psi = j\hbar\frac{\partial\psi}{\partial t} \qquad \text{(VII.6.06)}$$

We shall see that this equation is solved by the following formula:

$$\psi(\mathbf{r}, t) = \psi(\mathbf{r}; \mathbf{k}(t)) \cdot \exp\left[-\frac{j}{\hbar}\int_{\tau=0}^{\tau=t} E(\mathbf{k}(\tau))\,d\tau\right]$$

$$= A \cdot \exp j\left[\mathbf{k}(t)\cdot\mathbf{r} - \frac{1}{\hbar}\int_{\tau=0}^{\tau=t} E(\mathbf{k}(\tau))\cdot d\tau\right] \qquad \text{(VII.6.07)}$$

Here the wave number **k** is assumed to have the following time dependence:

$$\mathbf{k}(t) = \mathbf{k}(0) + \frac{1}{\hbar}\mathbf{F}t \qquad \text{or} \qquad \dot{\mathbf{k}}(t) = \frac{1}{\hbar}\mathbf{F} \qquad \text{(VII.6.08)}$$

By definition, we understand by $E(\mathbf{k}(t))$ the same functional dependence of the energy on the wave vector **k** as for the absence of an external force. Hence, also by definition, Eq. (VII.6.04)

$$E(\mathbf{k}(t)) = -eU_0 + \frac{\hbar^2}{2m}|\mathbf{k}(t)|^2 \qquad \text{(VII.6.04)}$$

continues to be valid for a $\mathbf{k}(t)$ varying with time. We shall verify the formula (VII.6.07) by substitution in (VII.6.06).

Before this, some of our readers may welcome a somewhat more concrete description of the physical situation expressed by the formulas (VII.6.05), (VII.6.06) and the solution (VII.6.07), (VII.6.08), which up to now has been presented in purely mathematical and hence rather abstract terms.

We have assumed the force to be constant in space, so that there exists a constant force field throughout infinite space. The picture contains neither the origin of the lines of force at some surface charge nor their ending at some other surface charge. The solenoidal electrical field in the accelerating tube of a betatron may serve as an example for such a field. The field is produced by the time variation of a central magnetic flux. Its lines of force are therefore circles. They do not originate at or end on electrical charges, but form closed loops. If the accelerating tube has a very large radius, we can forget about its curvature, and we have an unbounded, though not infinite, space in which the same force **F** prevails everywhere—at least in the direction of the field—as soon as the central magnetic flux has begun to change. If, at the instant $t = 0$, even before the "switching on" of the force $\mathbf{F}$, electrons are traveling in the accelerating tube, they will be represented by a wave with an appropriate wave vector $\mathbf{k}$ (see Fig. VII.6.1). Solution (VII.6.07), in combination with (VII.6.08), tells us that the electrons in the accelerating tube are represented also after the switching on of the force by an unlimited "plane" wave whose

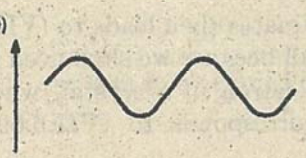

FIG. VII.6.1. Electron wave at the time $t = 0$.

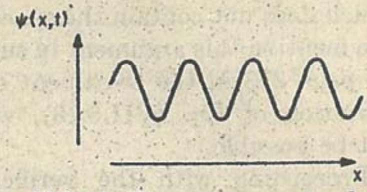

FIG. VII.6.2. Electron wave at a later time $t > 0$, a force F having acted since $t = 0$.

wavelength remains constant in space, but becomes uniformly smaller with time (see Fig. VII.6.2).

We now proceed to the verification of the solution (VII.6.07), (VII.6.08) and obtain first from (VII.6.07) by differentiation

$$\frac{\partial}{\partial x}\psi(\mathbf{r}, t) = jk_x(t)\cdot\psi(\mathbf{r}, t)$$

$$\frac{\partial^2}{\partial x^2}\psi(\mathbf{r}, t) = -k_x^2(t)\cdot\psi(\mathbf{r}, t)$$

$$-\frac{\hbar^2}{2m}\Delta\psi(\mathbf{r}, t) = +\frac{\hbar^2}{2m}\cdot|\mathbf{k}(t)|^2\cdot\psi(\mathbf{r}, t)$$

From this we obtain with the aid of (VII.6.04)

$$-\frac{\hbar^2}{2m}\Delta\psi(\mathbf{r}, t) = [E(\mathbf{k}(t)) + eU_0]\cdot\psi(\mathbf{r}, t)$$

or
$$-\frac{\hbar^2}{2m}\Delta\psi(\mathbf{r}, t) - eU_0\psi(\mathbf{r}, t) = E(\mathbf{k}(t))\cdot\psi(\mathbf{r}, t) \qquad \text{(VII.6.09)}$$

We could have written down Eq. (VII.6.09) even without calculation, since, by definition, $\psi(\mathbf{r}; \mathbf{k}) = A\, e^{j\mathbf{k}\cdot\mathbf{r}}$ satisfies the stationary Schrödinger equation in the absence of an external force:

$$-\frac{\hbar^2}{2m}\Delta\psi(\mathbf{r}; \mathbf{k}) - eU_0\psi(\mathbf{r}; \mathbf{k}) = E(\mathbf{k})\psi(\mathbf{r}; \mathbf{k})$$

If $\mathbf{k}$, instead of being constant, varies with a parameter which does not occur in the stationary equation, namely, time, this does not detract from the fulfillment of the stationary Schrödinger equation provided only that the value $E(\mathbf{k}(t))$ pertaining to the moment of time in question is substituted for the energy $E(\mathbf{k})$. The addition of the time factor

$$\exp\left[-\frac{j}{\hbar}\int_{\tau=0}^{\tau=t} E(\mathbf{k}(\tau))\,d\tau\right]$$

which does not contain the space coordinates then leads to (VII.6.09). We mention this argument in such detail because we shall need it later on page 213 in the treatment of an electron in a crystal, where the deviation of Eq. (VII.6.13), which corresponds to (VII.6.09), will not be possible.

Proceeding with the verification of (VII.6.07), we form from (VII.6.07), making use of (VII.6.08),

$$\begin{aligned} j\hbar\frac{\partial}{\partial t}\psi(\mathbf{r},t) &= j\hbar\cdot j\cdot\left[\dot{\mathbf{k}}(t)\cdot\mathbf{r}-\frac{1}{\hbar}E(\mathbf{k}(t))\right]\psi(\mathbf{r},t)\\ &= -\hbar\cdot\left[\frac{1}{\hbar}\mathbf{F}\cdot\mathbf{r}-\frac{1}{\hbar}E(\mathbf{k}(t))\right]\psi(\mathbf{r},t)\\ &= -[\mathbf{F}\cdot\mathbf{r}-E(\mathbf{k}(t))]\psi(\mathbf{r},t) \end{aligned}$$

which is substituted on the right side of (VII.6.06). On the left side we substitute the first two terms from (VII.6.09). We then obtain an identity, namely,

$$E(\mathbf{k}(t))\psi(\mathbf{r},t)-\mathbf{F}\cdot\mathbf{r}\cdot\psi(\mathbf{r},t) = -[\mathbf{F}\cdot\mathbf{r}-E(\mathbf{k}(t))]\psi(\mathbf{r},t)$$

which demonstrates the fulfillment of (VII.6.06) by formulas (VII.6.07) and (VII.6.08).

b. The Electron in a Crystal under the Influence of an External Force F

We now claim that the behavior of an electron in a crystal under the influence of an external force may be described, just as that of a free electron, by letting the wave number $\mathbf{k}$ in the wave function

$$\psi(\mathbf{r};\mathbf{k}) = u(\mathbf{r};\mathbf{k})\,e^{j\mathbf{k}\cdot\mathbf{r}}$$

of the unperturbed problem depend on time in accordance with the law (VII.6.08). A simple attempt at verification will show that this does not apply with complete rigor.

The statement just made signifies that the time-dependent Schrödinger equation of an electron in a crystal under the influence of an external force

$$-\frac{\hbar^2}{2m}\,\Delta\psi(\mathbf{r}, t) - (eU(\mathbf{r}) + \mathbf{F}\cdot\mathbf{r})\psi(\mathbf{r}, t) = j\hbar\,\frac{\partial}{\partial t}\,\psi(\mathbf{r}, t) \qquad \text{(VII.6.10)}$$

has the solution

$$\psi(\mathbf{r}, t) = \psi(\mathbf{r}; \mathbf{k}(t))\cdot\exp\left[-\frac{j}{\hbar}\int_{\tau=0}^{\tau=t} E(\mathbf{k}(\tau))\,d\tau\right]$$

$$= u(\mathbf{r}; \mathbf{k}(t))\,\exp j\left[\mathbf{k}(t)\cdot\mathbf{r} - \frac{1}{\hbar}\int_{\tau=0}^{\tau=t} E(\mathbf{k}(\tau))\,dt\right] \qquad \text{(VII.6.11)}$$

with

$$\mathbf{k}(t) = \mathbf{k} + \frac{1}{\hbar}\,\mathbf{F}t \qquad \text{or} \qquad \dot{\mathbf{k}}(t) = \frac{1}{\hbar}\,\mathbf{F} \qquad \text{(VII.6.12)}$$

$\psi(\mathbf{r};\ \mathbf{k}(t)) = u(\mathbf{r};\ \mathbf{k}(t))\ e^{j\mathbf{k}\cdot\mathbf{r}}$ is here any solution of the stationary Schrödinger equation

$$-\frac{\hbar^2}{2m}\,\Delta\psi - eU(\mathbf{r})\psi = E\psi$$

Correspondingly, the energy E in the exponent of the time-dependent factor of (VII.6.11) is assumed to depend on the time-dependent wave vector $\mathbf{k}(t)$ in the same manner as on the constant wave number $\mathbf{k}$ for an electron in a crystal with lattice potential $U(\mathbf{r})$ without external force. Examples are provided by Figs. VII.2.9, VII.3.10*c*, and VII.3.12, as well as by Eq. (VII.2.22) for the specific case of the Bloch approximation.

This remark immediately leads to[1]

$$-\frac{\hbar^2}{2m}\,\Delta\psi(\mathbf{r}; \mathbf{k}(t)) - eU(\mathbf{r})\psi(\mathbf{r}; \mathbf{k}(t)) = E(\mathbf{k}(t))\cdot\psi(\mathbf{r}; \mathbf{k}(t))$$

so that, with the addition of the time factor

$$\exp\left[-\frac{j}{\hbar}\int_{\tau=0}^{\tau=t} E(\mathbf{k}(\tau))\,d\tau\right]$$

we have

$$-\frac{\hbar^2}{2m}\,\Delta\psi(\mathbf{r}, t) - eU(\mathbf{r})\psi(\mathbf{r}, t) = E(\mathbf{k}(t))\cdot\psi(\mathbf{r}, t) \qquad \text{(VII.6.13)}$$

This is the first step in a verification of (VII.6.11). To carry it further,

---

[1] The reasoning is the same as that employed for Eq. (VII.6.09).

we must form

$$\begin{aligned}\frac{\partial}{\partial t}\psi(\mathbf{r};\,t) &= \frac{\partial}{\partial t}\left\{u(\mathbf{r};\,\mathbf{k}(t))\cdot\exp j\left[\mathbf{k}(t)\cdot\mathbf{r}-\frac{1}{\hbar}\int_{\tau=0}^{\tau=t}E(\mathbf{k}(\tau))\,d\tau\right]\right\}\\ &= j\left[\dot{\mathbf{k}}(t)\cdot\mathbf{r}-\frac{1}{\hbar}E(\mathbf{k}(t))\right]\psi(\mathbf{r},\,t)+\left(\frac{\partial u(\mathbf{r};\,\mathbf{k}(t))}{\partial k_x}\cdot\dot{k}_x(t)\right.\\ &\qquad\left.+\frac{\partial u(\mathbf{r};\,\mathbf{k}(t))}{\partial k_y}\,\dot{k}_y(t)+\frac{\partial u(\mathbf{r};\,\mathbf{k}(t))}{\partial k_z}\,\dot{k}_z(t)\right)\exp j[\;\cdot\;\cdot\;\;]\end{aligned}$$

$$\begin{aligned}\frac{\partial}{\partial t}\psi(\mathbf{r},\,t) = j\left[\dot{\mathbf{k}}(t)\cdot\mathbf{r}-\frac{1}{\hbar}E(\mathbf{k}(t))\right]\psi(\mathbf{r},\,t)+\operatorname{grad}_{\mathbf{k}}u(\mathbf{r};\,\mathbf{k}(t))\cdot\dot{\mathbf{k}}(t)\\ \cdot\exp j\left[\mathbf{k}(t)\cdot\mathbf{r}-\frac{1}{\hbar}\int_{\tau=0}^{\tau=t}E(\mathbf{k}(\tau))\,d\tau\right]\end{aligned}$$

With (VII.6.12) we then find

$$\begin{aligned}j\hbar\frac{\partial}{\partial t}\psi(\mathbf{r},\,t) = -(\mathbf{F}\cdot\mathbf{r})\psi(\mathbf{r},\,t)+E(\mathbf{k}(t))\psi(\mathbf{r},\,t)+j\mathbf{F}\cdot\operatorname{grad}_{\mathbf{k}}u(\mathbf{r};\,\mathbf{k}(t))\\ \cdot\exp j\left[\mathbf{k}(t)\cdot\mathbf{r}-\frac{1}{\hbar}\int_{\tau=0}^{\tau=t}E(\mathbf{k}(\tau))\,d\tau\right]\qquad\text{(VII.6.14)}\end{aligned}$$

A combination of (VII.6.13) and (VII.6.14) then shows that (VII.6.11) almost satisfies the Schrödinger equation (VII.6.10). The term $j\mathbf{F}\cdot\operatorname{grad}_{\mathbf{k}}u(\mathbf{r};\,\mathbf{k}(t))\cdot\exp j[\;\cdot\;\cdot\;\cdot\;]$ on the right side of (VII.6.10) remains uncompensated, however.

We shall now show that this failure of the assumption (VII.6.12) arises from the fact that it does not take account of transitions of the electron into higher bands which may take place under the influence of the added force **F**. To this end, we shall try to improve on formula (VII.6.11). It appears reasonable to replace the simple expression (VII.6.11) by a sum of such expressions with initially unknown amplitudes A:

$$\begin{aligned}\psi(x,\,t) = \sum_{N'=1}^{\infty}\;\sum_{n'=-\frac{G}{2}}^{n'=+\frac{G}{2}-1}A_{N'n'}\psi_{N'}\left(x;\,k_{n'}+\frac{1}{\hbar}Ft\right)\\ \cdot\exp\left[-\frac{j}{\hbar}\int_{\tau=0}^{\tau=t}E_{N'}\left(k_{n'}+\frac{1}{\hbar}F\tau\right)d\tau\right]\qquad\text{(VII.6.15)}\end{aligned}$$

For the sake of simplicity we have passed over to the one-dimensional case. Hence Eq. (VII.2.05) gives for the individual wave-number values

$$k_n = \frac{2\pi}{a} \cdot \frac{n}{G}$$

The summation over n' in (VII.6.15) thus encompasses all G eigenfunctions of a band, the summation over N', the several bands. Even this extended formula, however, does not lead closer to the goal, since now we have on the right side of Eq. (VII.6.10) a sum of uncompensated terms

$$jA_{N'n'}F \frac{\partial}{\partial k} u_{N'}\left(x; k_{n'} + \frac{1}{\hbar} Ft\right) \exp j[\ \cdot\ \cdot\ \cdot\ ]$$

Compensation becomes possible only if we let the amplitudes $A_{N'n'}$ depend on the time

$$A_{N'n'} = A_{N'n'}(t)$$

Substitution of (VII.6.15) now leads to additional terms on the right side of (VII.6.10):

$$j\hbar \dot{A}_{N'n'}(t)\psi_{N'}\left(x; k_{n'} + \frac{1}{\hbar} Ft\right) \exp\left[-\frac{j}{\hbar}\int_{\tau=0}^{\tau=t} E_{N'}\left(k_{n'} + \frac{1}{\hbar}F\tau\right) d\tau\right]$$

and the fulfillment of (VII.6.10) leads to the requirement

$$0 = \left\{\begin{aligned} &\sum_{N'=1}^{\infty} \sum_{n'=-\frac{G}{2}}^{n'=+\frac{G}{2}-1} j\hbar \dot{A}_{N'n'}(t)\psi_{N'}\left(x; k_{n'} + \frac{1}{\hbar}Ft\right) \\ &\qquad\qquad \cdot \exp\left[-\frac{j}{\hbar}\int_{\tau=0}^{\tau=t} E_{N'}\left(k_{n'} + \frac{1}{\hbar}F\tau\right) d\tau\right] \\ &+ \sum_{N'=1}^{\infty} \sum_{n'=-\frac{G}{2}}^{n'=+\frac{G}{2}-1} jA_{N'n'}(t) \cdot F \cdot \frac{\partial}{\partial k} u_{N'}\left(x; k_{n'} + \frac{1}{\hbar}Ft\right) \\ &\qquad\qquad \cdot \exp j\left[\left(k_{n'} + \frac{1}{\hbar}Ft\right) \cdot x - \frac{1}{\hbar}\int_{\tau=0}^{\tau=t} E_{N'}\left(k_{n'} + \frac{1}{\hbar}F\tau\right) d\tau\right] \end{aligned}\right. \qquad \text{(VII.6.16)}$$

To determine what are in principle an infinite number of unknowns $\dot{A}_{N'n'}$, we multiply (VII.6.16) by

$$\psi_N^*\left(x; k_n + \frac{1}{\hbar}Ft\right) \exp\left[+\frac{j}{\hbar}\int_{\tau=0}^{\tau=t} E_N\left(k_n + \frac{1}{\hbar}F\tau\right) d\tau\right]$$

and integrate over the fundamental domain of the crystal.

As solutions of the stationary problem, the $\psi_N(x; k_n + Ft/\hbar)$ are orthogonal and normalized. We therefore obtain

$$0 = \left\{ \begin{aligned} &+\sum_{N'=1}^{\infty} \sum_{n'=-\frac{G}{2}}^{n'=+\frac{G}{2}-1} j\hbar \dot{A}_{N'n'}(t) \cdot \delta_{NN'}\delta_{nn'} \\ &\qquad \cdot \exp\left[-\frac{j}{\hbar}\int_{\tau=0}^{\tau=t} \Big[E_{N'}\left(k_{n'} + \frac{1}{\hbar}F\tau\right) - E_N\left(k_n + \frac{1}{\hbar}F\tau\right) d\tau\right] \\ &+\sum_{N'=1}^{\infty} \sum_{n'=-\frac{G}{2}}^{n'=+\frac{G}{2}-1} jA_{N'n'}(t) \int_{V_{\text{fund}}} u_N^*\left(x; k_n + \frac{1}{\hbar}Ft\right) \\ &\qquad \cdot F \cdot \frac{\partial}{\partial k} u_{N'}\left(x; k_{n'} + \frac{1}{\hbar}Ft\right) \\ &\qquad \cdot \exp j\Big[\left(k_{n'} + \frac{1}{\hbar}Ft\right)x - \frac{1}{\hbar}\int_{\tau=0}^{\tau=t} E_{N'}\left(k_{n'} + \frac{1}{\hbar}F\tau\right) d\tau \\ &\qquad - \left(k_n + \frac{1}{\hbar}Ft\right)x + \frac{1}{\hbar}\int_{\tau=0}^{\tau=t} E_N\left(k_n + \frac{1}{\hbar}F\tau\right) d\tau\Big] \cdot dV \end{aligned} \right.$$

or

$$\begin{aligned} \dot{A}_{Nn}(t) = -\sum_{N'=1}^{\infty} \sum_{n'=-\frac{G}{2}}^{n'=+\frac{G}{2}-1} A_{N'n'} \frac{1}{\hbar}F \int_{V_{\text{fund}}} u_N^*\left(x; k_n + \frac{1}{\hbar}Ft\right) e^{j(k_{n'}-k_n)x} \\ \cdot \frac{\partial}{\partial k} u_{N'}\left(x; k_{n'} + \frac{1}{\hbar}Ft\right) dV \\ \cdot \exp\left\{+\frac{j}{\hbar}\int_{\tau=0}^{\tau=t} \left[E_N\left(k_n + \frac{1}{\hbar}F\tau\right) - E_{N'}\left(k_{n'} + \frac{1}{\hbar}F\tau\right)\right] d\tau\right\} \end{aligned}$$

We now find that the summation over n' in (VII.6.15) is superfluous.

For every arbitrary but fixed point of time

$$u_N^*\left(x; k_n + \frac{1}{\hbar}Ft\right) \cdot \frac{\partial}{\partial k} u_{N'}\left(x; k_{n'} + \frac{1}{\hbar}Ft\right)$$

is a function with lattice periodicity. We can hence apply the theorem derived in Appendix I, page 381, to the effect that the integral of the product of a function $f(x)$ with lattice periodicity and a factor $e^{j(k-k')x}$ over the fundamental domain vanishes unless $k = k'$. Thus the summation over n' in the last equation for $\dot{A}_{Nn}(t)$ drops out and we obtain

$$\begin{aligned}\dot{A}_{Nn}(t) = -\frac{1}{\hbar}F \sum_{N'=1}^{\infty} A_{N'n}(t) \int\limits_{V_{\text{fund}}} & u_N^*\left(x; k_n + \frac{1}{\hbar}Ft\right) \\ & \cdot \frac{\partial}{\partial k} u_{N'}\left(x; k_n + \frac{1}{\hbar}Ft\right) \cdot dV \qquad \text{(VII.6.17)} \\ \cdot \exp\left\{+\frac{j}{\hbar}\int_{\tau=0}^{\tau=t} \left[E_N\left(k_n + \frac{1}{\hbar}F\tau\right) - E_{N'}\left(k_n + \frac{1}{\hbar}F\tau\right)\right] d\tau\right\}\end{aligned}$$

The dropping out of the summation over n' indicates in particular that for an electron which at the time $t = 0$ is represented by a function $\psi_N(x; k_n)$ no other wave numbers $k_{n'}$ play a role. More precisely, the action of the added force F changes the wave number k in accord with the time variation (VII.6.12), and the simple assumption (VII.6.11) would describe the behavior of the electron completely if states with equal $k(t) = k_n + Ft/\hbar$ were not excited in other bands $N' \neq N$ in the course of time. Equation (VII.6.17) teaches in effect that, even if at $t = 0$ only a single coefficient A_{N_0n} differed from zero and the electron occupied accordingly only the state k_n in the band N_0, all other coefficients $A_{N'n}$ will, in principle, be different from zero after a time t, so that the added force F effects transitions into every other band $N' \neq N_0$ with a certain probability. This probability is measured by the square of the absolute value of $A_{N'n}$. Carrying the treatment of the system of differential equations (VII.6.17) for the infinite number of coefficients $A_{N'n}$ further should indicate with what frequency transitions of the electron into another band occur for a given magnitude of the force F.[1] However, we shall be able to answer this

---

[1] We would use Eq. (VII.6.17) for this purpose. The preceding presentation follows W. V. Houston, *Phys. Rev.*, **57**: 184 (1940). We should also mention in this connection F. Bloch, *Z. Physik*, **52**: 555 (1928); R. Peierls, *Z. Physik*, **53**: 255 (1929); H. A. Bethe in Geiger and Scheel, *op. cit.*, vol. XXIV, part 2, p. 507; H. Jones and C. Zener, *Proc. Roy. Soc. (London)*, **144**: 101–117 (1934); J. C. Slater, *Revs. Mod. Phys.*, **6**: 209 (1934), particularly p. 259; and A. H. Wilson, "The Theory of

question more simply in another manner in §7. We shall see here also that only extraordinarily large forces can effect the transition of electrons into other bands with appreciable frequency. For forces of ordinary magnitude we can state simply that the electron is represented continuously by one[1] solution $\psi(\mathbf{r};\,\mathbf{k})$ of the stationary Schrödinger equation in the absence of an external force, the variation of the wave vector with time being given by

$$\dot{\mathbf{k}} = \frac{1}{\hbar}\mathbf{F} \qquad \text{(VII.6.12)}$$

Equation (VII.6.12) has very important physical consequences, which we shall discuss next.

c. The Effective Mass of an Electron in a Crystal

Equation (VII.6.12) together with (VII.5.10) lead to an analogy for the equation for a free electron

$$\dot{\mathbf{v}} = \frac{1}{m}\mathbf{F} \qquad \text{(VII.6.18)}$$

For the sake of simplicity and to obtain a general view of the situation, we proceed from the variation $E(\mathbf{k})$ (VII.4.04) which is commonly assumed for the neighborhood of the band edges and which we shall now write in the form[2]

$$E = E_{\text{bound}} + \tfrac{1}{2}E''(|\mathbf{k}_{\text{bound}}|)[(k_x - k_{x\text{bound}})^2 + (k_y - k_{y\text{bound}})^2 + (k_z - k_{z\text{bound}})^2] \qquad \text{(VII.6.19)}$$

In this particular variation of $E(\mathbf{k})$ with the wave vector $\mathbf{k}$, the second derivative E'' $(\mathbf{k}_{\text{bound}})$ with respect to the absolute value of the wave

---

Metals," Cambridge University Press, London, 1935. The books of Fröhlich, Seitz, and Mott and Jones give a very simple proof for (VII.6.12) with the aid of an energy theorem. We fear, however, that in this very simple proof a very important part of the matter to be proved—namely, the continuous representation of an electron in the crystal by the solution $\psi(\mathbf{r};\,\mathbf{k})$ of the stationary Schrödinger equation in the absence of an external force—is implicit in the formulation of the energy theorem so that the further deductions only give a more precise picture of the nature of the time dependence of $\mathbf{k}$. For a further objection to this argument, see Shockley, *op. cit.*, pp. 424–425. With regard to this whole series of questions, see also D. Pfirsch and E. Spenke, *Z. Physik*, **137**: 309 (1954).

[1] More recently it has been found, however, that inc ertain problems (related to Ehrenfest's theorem) transitions to higher bands must be considered even for weak forces [see D. Pfirsch and E. Spenke, *Z. Physik*, **137**: 309 (1954)].

[2] The replacement of "const" in Eq. (VII.4.04) by $\tfrac{1}{2}E''$ $(|\mathbf{k}_{\text{bound}}|)$ is obtained from a Taylor expansion of the function $E(\mathbf{k})$ about the point $\mathbf{k} = \mathbf{k}_{\text{bound}}$.