Recollections of early solid state physics

BY R. E. PEIERLS, F.R.S.

Nuclear Physics Laboratory, University of Oxford, U.K.


Author of many papers applying quantum mechanics to problems in solid state and in particle physics; author of Quantum theory of solids, 1955. Author with O. R. Frisch in 1940 of 'Confidential memorandum to British Government on possibility and critical size of nuclear bomb'.

My first contact with the electron theory of metals was in Sommerfeld's department at Munich, where I was a student from October 1926 to Easter 1928. These were my third to fifth semesters as a student, and I was only beginning to acquire an understanding of physics, and of the new quantum and wave mechanics.

Sommerfeld was then developing his approach to electrons in metals, with some assistance from two American visitors, W. V. Houston and C. Eckart. I heard him give a series of lectures summarising the problems, and his results, even before the papers were published.\(^1\)

Sommerfeld was very much at home with the classical Lorentz–Drude theory, and familiar with the experimental data, which produced such a mysterious mixture of confirmation (Wiedemann–Franz ratio, 'normal' Hall effect, and many of the thermoelectric and thermomagnetic coefficients) and utter contradiction (paramagnetism, specific heat, temperature dependence of resistivity). Pauli had shown\(^2\) that the application of Fermi statistics to the conduction electrons had resolved the paradox of paramagnetism. Nobody seemed worried at the time about the existence of diamagnetic metals; it was known that the ion cores would be diamagnetic.

Sommerfeld built on Pauli's idea, and showed that Fermi statistics would also account for the absence of an observable electronic specific heat, and this was the greatest success of his approach. He showed that the specific heat of a degenerate Fermi gas would be proportional to \(T\), and should be observable at very low temperatures. It was gratifying that the constancy of the Wiedemann–Franz ratio could still be accounted for, although on the face of it it had appeared to require a classical specific heat, and that the numerical value, which differed a little from the classical value, fitted the data even better.

His theory failed to account for the temperature dependence of the resistivity, and for the magnetoresistance. The latter is now known to depend essentially on variations in the electron mobility. Sommerfeld thought of resistance in terms of a constant mean free path. This made the collision time, and hence the mobility,
Recollections of solid state physics

inversely proportional to the velocity. Since even in a degenerate Fermi gas there is a small spread of velocities of conducting electrons around the Fermi velocity, he could obtain a small spread of collision times, and hence a small magneto-resistance, several orders of magnitude less than the observed effect.

He followed the classical ideas by thinking of the electrons as free, except for collisions with atoms (therefore the magnitude of the mean free path required to explain the observed conductivity, even larger than in the classical picture because of the greater velocity, was another paradox). He also ignored the mutual interaction of the electrons. This had always been done in the classical theory, since electron-electron encounters conserve the total electron momentum, hence the total current, and thus do not contribute to the resistance. I do not recall Sommerfeld mentioning this argument explicitly, but it is always clearly stated in the classical treatments, and there was no reason why he should have taken a different point of view.

As a very junior theoretician I listened to Sommerfeld’s exposition and was duly impressed, but was not yet at the stage of criticizing or questioning the assumptions. It was characteristic of Sommerfeld’s positive attitude that one learnt more about the successful solution of difficulties than about the mysteries that remained. He was completely fair in listing the contradictions – it was just a matter of emphasis.

When at Easter 1928 Sommerfeld left for a sabbatical year, I joined Heisenberg’s group at Leipzig, where Felix Bloch had just completed his treatment of electrons in periodic potentials and his explanation of the order of magnitude and temperature dependence of the resistivity in terms of lattice vibrations. Bloch also did not worry about the electron-electron interaction – I do not know whether he recalled the old arguments that it should not matter, or whether he was simply content to extend the theory by taking in one more factor which had previously been ignored, without necessarily including everything.

My first substantial research assignment in Leipzig was to see whether Bloch’s starting-point of independent electrons (‘electron orbitals’ we would say today) was unavoidable, and how far one could get if one started from a Heitler-London model (which Heisenberg was about this time applying to ferromagnetism). Today it would be obvious to any undergraduate that there could not be any conductivity in the Heitler-London model unless it is supplemented by ionized states, in which some atoms have more, and others fewer, than their normal complement of electrons. But at the time this conclusion was not obvious to me, and evidently not to Heisenberg. It took some struggle with exchange integrals for a many-body system before I concluded that, at least in the linear chain I was using as a model, the only way a current could be obtained was by all electrons making a quantum jump simultaneously, and for a macroscopic dimension this makes the rate astronomically small.

In the summer of 1928 I was fairly sure about this conclusion. I spent the summer vacation in England, mostly as a tourist, but I visited Cambridge and called on Dirac, whom I had met. He introduced me to R. H. Fowler. When Fowler heard I was from Leipzig, he asked me to talk to the Kapitza Club about Bloch’s work. At that time neither my English nor my command of physics was really adequate for
this, but it did not occur to me to refuse, and I did my best. But, more relevant to
the present story, he also introduced me to one of his research students, W. H.
McCrea (now a distinguished astrophysicist), who was also thinking about con­
ductivity using what he and Fowler called the Heitler–London model. Actually
this was a one-electron tight-binding model with two centres of force.

On my return to Leipzig the project was abandoned, and Heisenberg suggested I
looked at the ‘anomalous’, i.e. positive, Hall effect. I tackled this on the basis of
Bloch’s theory of electrons in periodic fields, and first had to convince myself that
the effect of the magnetic field on the wave vector of the electron was the same as
for a free electron of the same velocity, but that the mean velocity of the electron
was given by $\frac{dE}{dk}$, and therefore different from that for a free electron of the
same $k$, if the energy function $E(k)$ was different. It was obvious, in particular,
that in Bloch’s tight-binding model the energy would flatten off near the band edge,
so that the current would there go to zero. Thus for an electron near the band edge
an electric field could cause a decrease, rather than an increase, in the velocity in
the field direction. One’s first shock on seeing this result is the fear that it might
lead to a negative conductivity. One soon realizes, however, that for an ensemble
of electrons in statistical equilibrium the positive acceleration of the electrons near
the bottom of the band outweighs the negative acceleration of those near the top,
until for a full band the current just vanishes.

At this point one was close to an explanation of the positive Hall effect, subject
only to the proof that the rate of change of the wavevector in the magnetic field
is still given by the Lorentz force. At this point I cheated a little by disregarding
inter-band terms, which for the purpose in hand were unimportant, but in other
problems can cause headaches.

So the explanation of the positive Hall effect came out without much difficulty.
I recall a comment by Heisenberg that this was similar to the situation in atomic
spectra (pointed out, I think, by Pauli) where an atom with one, or a few, electrons
missing from a closed shell was dynamically similar to one with just one, or a few,
electrons in that shell, except for some signs. My memory is confused, however, on
the question whether this comment was made when Heisenberg suggested the
problem to me, or when I showed him the answer. In other words, I am not clear
whether Heisenberg had, with his usual powerful intuition, guessed in advance how
the solution would come out. I reported previously that he had, but I am now
rather doubtful whether this was right.

In any event it was gratifying to have solved one of the remaining mysteries.
I wrote a paper on the subject, which was not too clearly written, and also gave
a talk to a conference, of which a summary is published. It contains a sketch of
the Fermi surface for a two-dimensional square lattice for the case of an almost
empty, and an almost full band with tight binding. In the latter case the boundary
consists of circular quadrants inside the four corners of the square which forms the
Brillouin zone for that case. In the longer paper there is also the remark that the
conductivity vanishes for a full band.
This seems relevant to another question on which my memory fails to serve, namely when and how it was first realized that a filled band would give an insulator. In retrospect it seems to be an obvious consequence of the existence of bands, at least in the tight-binding limit, and particularly obvious from the arguments sketched above. It seems almost incredible that this point could have been missed, but I have no clear recollection of when I became aware of it, and it is certainly not mentioned in any paper of that time.

This work was complete by the spring of 1929, and since at that time Heisenberg went on sabbatical leave, I moved to Zurich to work with Pauli. Here I left metals for a while, since Pauli suggested to me the problem of heat conduction in non-metallic crystals, under the influence of the anharmonic forces. This was a problem which, at least at high temperatures, could be treated classically. Pauli had been interested in this problem and had looked at the related question of the absorption of sound waves because of anharmonicity. The abstract of a talk he gave to a meeting is published, and the answer given there is wrong (probably the only error in print under Pauli's name) because it gives a finite damping in a linear chain, for which in fact the three-phonon processes, which he was studying, do not occur. He showed me a few pages of notes on this problem, to start me off. Apart from this guidance I looked at the problem from first principles, and this was probably fortunate, because there were a number of different wrong approaches in the literature, and it was less confusing to find the solution first, and then discover where others had gone wrong.

This led to the concept (and the ugly word) of Umklapp processes, and to the prediction of the exponential rise of the heat conductivity at low temperatures, verified only in 1951 by Berman. In many ways my paper did not dispose of the problem. For example, it failed to point out that a 'pure' substance, to show the exponential rise, had to be also isotopically pure. This omission made the experimental discovery of the effect more difficult. Other, more sophisticated parts of the problem are still not completely sorted out. I submitted a thesis on this topic to Leipzig (my one semester in Zurich not being an adequate residence qualification there) and returned to Zurich as Pauli's assistant.

I then started thinking further about electrons in metals. I felt uncomfortable about having, in my work on the Hall effect, relied on the flattening of the energy surface near the band edge, a result then known only in the tight-binding limit, which was not realistic for conduction electrons. It seemed obvious that in the opposite limit of free electrons this effect was absent, and one therefore did not know what was happening in the intermediate case. It suddenly dawned on me that, if a weak potential was added as a small perturbation, there would be band gaps near the Bragg reflexions, and that the energy surface there had zero slope, though, for a very weak potential this flattening was confined to a very narrow region near the edge, and the slope returned to its free-electron value more rapidly the weaker the potential.

Few pieces of work have given me as much pleasure as this discovery, which
required only a few lines of calculation, both because it satisfied me that the nature of the Bloch bands was now qualitatively the same all the way from tight binding to almost free electrons, and because of the neat method of approximation I had invented. This was, of course, not new, being the standard technique of dealing with the anomalous Zeeman effect, but I was quite ignorant of this. I was satisfied with solving the one-dimensional case, but Brillouin took over the idea and discussed the general three-dimensional problem in depth, which became the theory of Brillouin zones.\(^{(11)}\)

At the same time I started worrying about the conservation of wavevector (pseudomomentum as we say today) in metals. I had found that in non-metallic crystals at low temperatures, when Umklapp processes were rare, it was difficult to get rid of the phonon drift caused by a temperature gradient. In a metal the electric field caused an electron drift, which in electron–phonon collisions tended to be passed on to the phonons. What mechanism restores the phonons to equilibrium?

Bloch had by-passed this problem by assuming that the phonon distribution is always in thermal equilibrium. How far was this justified? I set up a Boltzmann equation for electrons and phonons, allowing for their interaction. For an electron band which is about half full, so that neither the number of electrons nor that of holes is small, or in the presence of several partly filled bands, there is no difficulty in generating Umklapp processes, even at low temperature, and one still finds Bloch’s \(T^5\) law for the low-temperature resistivity. I first found \(T^4\),\(^{(12)}\) because of an invalid approximation, and had to correct this in a later paper.\(^{(13)}\) The factor of the \(T^5\) law could be similar to Bloch’s or different, according to whether phonon–phonon interactions or electron–phonon interactions were dominant in keeping the phonons in equilibrium.

However, in metals in which the Fermi surface did not touch the zone boundary, the theory predicted an exponential rise of the electric conductivity at low temperature. I was worried by the fact that such a behaviour had never been seen. I was aware of the fact that electron–electron collisions could also cause Umklapp processes, and the criterion for this was somewhat less restrictive than for the electron–phonon interaction. Roughly speaking, it would be sufficient if the diameter of the Fermi surface was at least one-half of that of the Brillouin zone. But it was known that the contribution of electron–electron collisions to the resistivity was proportional to \(T^2\); if these collisions were vital to give a finite resistivity it was hard to believe that one would still find a \(T^5\) law.

I was not then aware of the correct explanation: all measurements at these low temperatures were, until recently, made on specimens whose residual (impurity) resistance was much higher than the ‘ideal’ resistivity, and had to be subtracted from the measured value, assuming additivity according to Matthiessen’s rule. This rule lacks any rigorous theoretical foundation, and, while it is empirically quite well satisfied in many circumstances, it is not applicable if the impurities are sufficient to dispose of the excess pseudomomentum, but matter less for the actual
mobility of individual electrons. This suggests that for very pure specimens the resistivity should be substantially less than that computed by Matthiessen's rule from less pure samples. This has now been verified,\(^{(14)}\) so that the ideas of my 1931 paper have at last found some justification.

I have mentioned the electron–electron interaction, and this may be the place at which to refer to the question in Mott's memorandum why this was so generally ignored. I have already referred to the classical argument that collisions between free electrons conserve momentum and hence the current. I realized that this argument was unconvincing, because there was in this respect a complete analogy between electron–electron and electron–phonon interactions. In the absence of Umklapp processes both conserve the wavevector (pseudomomentum). Neither conserves the electric current if the electrons are in Bloch states, but any conservation law causes difficulty in restoring equilibrium, and hence in getting a finite resistivity. Both interactions can involve Umklapp processes, and then do not observe any conservation laws (except for energy).

The original reason for neglecting electron interaction was that one was starting from simple approaches and generalized the model step by step, giving particular weight to those additional factors that one could see would make a major difference. The effect of electron–electron interactions on the resistivity would lead, as I have mentioned, to a \(T^2\) law, which had never been observed, and this suggested that the effect was probably small. This was borne out by a crude order-of-magnitude estimate.

I remember a conversation with Landau in which he explained how to estimate the contributions of various factors to the resistivity by dimensional reasoning. This argument, which was more general than my rough estimate, showed indeed that the electron–electron collisions were negligible, except at temperatures below those in use at the time. He stressed that this was because in a degenerate Fermi gas the only permitted transitions were those in which both electrons were initially, and remained, in the border region of the Fermi distribution, so that the \(T^2\) factor was really \(\left(\frac{kT}{E_R}\right)^2\), whereas for the phonons the characteristic energy was the Debye cut-off \(k\Theta\). Evidently these ideas were the forerunners of his Fermi-liquid theory.

All of these discussions relate only to the effect of the electron–electron interaction on the irreversible processes, as distinct from its effect on the equilibrium state of the electron system. One guess, though not pursued in depth at the time, was that the Bloch wave functions and their energies might retain their meaning, but that their explicit form would have to be modified ('renormalized' in modern terms).

I must confess that I never took too much interest in band calculations or in the calculation of the electron–phonon matrix elements (about which there was a separate, but not unrelated, controversy). I preferred to look at results which were not sensitive to these questions, partly out of laziness but partly out of pessimism about the possibility of making realistic band calculations, which would have to
take electron–electron forces into account. This pessimism has since proved unjustified.

An important episode for my understanding of conduction problems arose from a paper by Kretschmann,(15) who attacked the then accepted theory of conductivity and claimed that the basis of the papers by Bloch and others was quite wrong. He had a number of objections which were mostly not very well conceived, but he claimed, in particular, that in the usual derivation of the Boltzmann equation one had made unjustified use of perturbation theory. In trying to defend the theory I therefore set out to prove that perturbation theory was in order, and to my amazement I found that this was very questionable, if not exactly for the reasons given by Kretschmann. It appeared that the usual application of Fermi’s ‘golden rule’ depended on the inequality

$$\hbar/\tau \ll kT,$$

where \(\tau\) is the collision time. This was not satisfied for many metals. For most the two quantities were of the same order of magnitude. Indeed Landau’s dimensional analysis made them comparable.(16)

This created a dilemma, in which Landau again came to the rescue. He produced an ingenious argument showing that, at least in situations in which the energy transfer in collisions was negligible, i.e. both at high temperatures, above the Debye \(\Theta\), and also where impurity scattering was dominant, it is sufficient that

$$\hbar/\tau \ll E_R,$$

which is much less restrictive, and holds for all normal metals. For intermediate temperatures, and for semiconductors, the point is still obscure.

This argument of Landau’s(17) is still essential today. It is now somewhat easier to visualize in terms of Kubo’s formula, and it is expressed in more highbrow ways in more recent transport theories, unless the problem is simply ignored, which is also not uncommon.

In the 1930s we were strongly influenced by the impressive successes of quantum mechanics in clearing up the basic problems of solid state physics, and this was probably responsible for a tendency to concentrate on general points of principle rather than on specific cases. This was certainly the attitude of Pauli, who had, in a sense, opened up the field with his paper on paramagnetism, and who maintained an interest in the field to the extent of attending conferences devoted to solid state theory (e.g. in Geneva in 1934). But as the work became more detailed, and required more ad hoc assumptions, it seemed to him ‘dirty physics’.

Perhaps it is typical therefore that, when during the Zurich period, I became interested in the optical properties of solids,(18) I looked only at the spectra of pure substances, using the tight-binding model, so that the theory applied strictly only to the rare-earth salts, or to crystals made of large organic molecules in which the optically active atoms were well shielded from each other. I was, of course, aware of the work by Pohl and others on the origin of luminescence and other optical
phenomena, but since these results depended strongly on the state of purity of the substance they seemed too complicated for a basic theoretical approach.

What I looked at was the Frenkel exciton, as it is now called, though I believe I had not read Frenkel's paper. I was particularly puzzled by the origin of the non-radiative decay of such excitations, which involved the transfer of energy to the phonons. One was accustomed to treating such situations by perturbation theory, treating anharmonic effects as small, and on that basis it seemed impossible to understand the decay, since the energy of the electronic excitation was much greater than the maximum phonon energy, so that the decay required the creation of many phonons in a single event. It took some trouble to appreciate that the forces on the excited atom would be appreciably different from those on a normal atom, even as regards the equilibrium configuration, so that a perturbation expansion in terms of displacements from a common equilibrium made no sense. The situation was similar to the familiar Franck-Condon treatment for molecules, except that one was dealing with much more numerous degrees of freedom.

One other worry concerned the magnetoresistance in metals. This was troubled by two difficulties. One was that all the simple theories, such as Sommerfeld's, gave a negligible effect. Similarly, an attempt by Bloch to invoke the paramagnetic spin reversal of some electrons, which alters the distribution of kinetic energies, gave an even smaller result. The other problem was that Kapitza's experiments gave, for many metals, a resistance which, over a long range, was linear in the magnetic field, whereas the usual theories always gave an $H^2$ law. Thinking about the field dependence I noticed that the usual approximation of expanding in powers of $H$ should break down when the Larmor period became shorter than the collision time, and that for very high fields the resistivity should probably tend to a constant limit. This argument therefore predicted a point of inflexion, which was going some way towards Kapitza's linear law. I was proud of this finding, and announced a talk on the subject at a conference in Leipzig in 1930. To my embarrassment, I discovered on the eve of my lecture that, in the model I was using, the coefficient of my pleasing magnetoresistance law was zero. The presentation of my talk was thereby made rather difficult.

I found later that the fault was in using an isotropic model in which, crudely speaking, all electrons had the same mobility, and that anisotropy of the Fermi surface, for example, would cause differences in mobility, and hence a substantial magnetoresistance. This had also been found by Bethe, and I do not remember clearly how far my ideas had been inspired by his paper. One still could not see why the magnetoresistance in the alkalis was of the same order as in other metals, although their Fermi surface was suspected, and has since been proved, to be very accurately spherical. I am not sure who first pointed out that the anisotropy of the phonon spectrum (which is anisotropic even for long waves in a cubic crystal) was responsible.

Landau had visited Zurich when his results on free-electron diamagnetism were new. According to his theory the diamagnetism would compensate one-third of the
Pauli paramagnetism. Some further diamagnetism was due to the ion cores, but it seemed difficult to explain how the total susceptibility could become strongly negative, as in Bi. Clearly the assumption of free electrons was not good enough, and one had to take the periodic potential of the lattice into account. How would one find the Landau diamagnetism for general Bloch states? This seemed a hard problem because Landau’s explanation appeared to depend on the discrete nature of the kinetic energy (at least in the plane perpendicular to the field). One would thus have to find the actual eigenstates in the periodic potential and magnetic field. Moreover, collisions would broaden these levels by more than their spacing if the collision time was shorter than the Larmor period. It was a pleasant surprise to discover that for statistical equilibrium the collision broadening and other effects were unimportant, as long as the energies associated with them were small compared to $kT$, even if they were larger than the level spacing. This made it possible to give a general expression for the Landau diamagnetism for any shape of the energy surface, except for inter-band terms, which were again hopefully regarded as negligible. For Bi this implied a very sharp curvature of the energy function, and later H. Jones put forward his model for Bi, which explained this fact.

This led to speculation what would happen in strong magnetic fields at very low temperatures, and it became evident that one would get an oscillatory behaviour of the susceptibility, reminiscent of the mysterious effect discovered by de Haas and van Alphen. Shoenberg has recently drawn my attention to the fact that this behaviour was already mentioned briefly in Landau’s paper. Landau discounts the effect as unobservable. Evidently Landau had not heard of the de Haas–van Alphen experiments, and they never noticed the remark in his paper. Presumably I never read Landau’s paper carefully, having had its main contents explained by him before publication, or, if I saw the remark, I accepted Landau’s assurance that it was unobservable, and promptly forgot it.

To account for the experiments one had to make what seemed rather exotic assumptions about the shape of the Fermi surface and the occupation numbers in Bi, and even with that I did not manage to get a quantitative fit because I had not allowed for the possibility of several degenerate branches of the Fermi surface. Nevertheless I was convinced that the explanation was basically right. Numerical studies by Blackman later gave a better fit. Then Landau showed how to obtain an approximate answer in closed form. The relation of the de Haas–van Alphen oscillations with the geometry of the Fermi surface was later made transparent by the beautiful argument of Onsager.

My only other major contact with solid state problems in the pre-war days related to the connection between long-range order and the existence of a sharp phase transition. My object was to stress the great qualitative difference between one and three dimensions. The argument also seemed to say that there could not be a sharp melting point in two dimensions, a result only recently upset by the work of Kosterlitz and Thouless.

By the mid-1930s I had become interested in other problems and lost touch with
solid state physics. Work on the Ising model, and on Bethe’s approximation to ordering in alloys, presumably counts as statistical mechanics rather than solid state theory.

My only other excursion into problems of solids was rather accidental. Orowan, who knew much about dislocations, had conceived a model to describe the force on a dislocation as it moves through the lattice. He found the mathematics troublesome, and suggested that I look at it. It was easy enough to formulate the equations for his model, but they led to a nonlinear integral equation. I knew nothing of integral equations, let alone nonlinear ones, and was ready to give up. It was clear physically that the solution to the equation had to be a function that approached a constant value for large positive argument and the same value with opposite sign for large negative argument. In other words the behaviour was similar to an arctan. Just for amusement I inserted an arctan into the equation, and to my great amazement it turned out to be the solution. In the remaining algebra I managed to lose a factor 2, which appears in a large exponent. This error was corrected many years later by Nabarro. I thought this result should be published by Orowan, who had invented the physics, or at most as a joint paper. But he refused, and I did not think the paper important enough to make an issue of it. I had presented it to a meeting in Bristol where, I believe, Orowan was not present. If I had foreseen what rings this small pebble would cause in the pond of dislocation physics I might have insisted, though this would have landed Orowan with a share of the responsibility for the disastrous factor 2, which makes the ‘Peierls–Nabarro force’ quite negligible for most purposes.

References

(1) Sommerfeld, A. 1928 Z. Phys. 47, 1, 43; Eckart, C. 1928 Z. Phys. 47, 38.
(2) Pauli 1926 Z. Phys. 41, 81.
(3) Bloch, F. 1928 Z. Phys. 52, 555.
(10) Peierls, R. 1930 Annln Phys. 4, 121.
(11) Brillouin, L. 1931 Annls Phys. 17, 88; Die Quantenstatistik (Berlin: Springer).
(12) Peierls, E. 1930 Annln Phys. 5, 244.
(18) Peierls, R. 1933 Annln Phys. 13, 905.
(21) Peierls, R. 1933 Z. Phys. 80, 703.
(23) Peierls, R. 1933 Z. Phys. 81, 186.