An essay on condensed matter physics in the twentieth century

W. Kohn

Department of Physics, University of California, Santa Barbara, California 93106

DEDICATED TO THE MEMORY OF J. M. LUTTINGER (1923–1997)

[S0034-6861(99)05002-3]

CONTENTS

I. INTRODUCTION

When the 20th century opened, the fields of crystallography, metallurgy, elasticity, magnetism, etc., dealing with diverse aspects of solid matter, were largely autonomous areas of science. Only in the 1940s were these and other fields consolidated into the newly named discipline of "solid state physics" which, two decades later, was enlarged to include the study of the physical properties of liquids and given the name "condensed matter physics" (CMP). At Harvard, for example, J. H. Van Vleck had several times taught a graduate course on magnetism in the 1930s and 1940s. However, the first time a course called "solid state physics" was taught there was in 1949 (by the writer, at Van Vleck's suggestion); it was based largely on the influential, comprehensive monograph *Modern Theory of Solids* by F. Seitz, which had appeared in 1940. In the early 1950s only a handful of universities offered general courses on solids and only a small number of industrial and government laboratories, most notably Bell Telephone Laboratories, conducted broad programs of research in CMP.

Today condensed matter physics is by far the largest subfield of physics. The writer estimates that at least a third of all American physicists identify themselves with CMP and with the closely related field of materials science. A look at the 1998 Bulletin of the March Meeting of the American Physical Society shows about 4500 papers in these fields.

Over the course of this century condensed matter physics has had a spectacular evolution, often by revolutionary steps, in three intertwined respects: new experimental discoveries and techniques of measurement; control of the compositions and atomic configurations of materials; and new theoretical concepts and techniques. To give a brief and readable account of this evolution is immensely difficult due to CMP's extraordinary diversity and many interconnections. Nevertheless, in the following pages the reader will find one theorist's broad-brush—and necessarily very incomplete—attempt at this task. The writer (not a historian of science) had to make many difficult, often rather arbitrary, choices: how to organize this very short essay on a very broad subject and—most painful—what to include and what important material to omit. He begs the reader's indulgence.

II. THE LAST YEARS OF THE CLASSICAL ERA

The very first Nobel Prize was awarded in 1901 to W. C. Roentgen for the discovery of penetrating, so-called x rays. A few years later in 1912, M. von Laue and collaborators demonstrated that these rays were electromagnetic waves of very short wavelengths, which could be diffracted by the atoms of crystals. This discovery dramatically proved beyond a doubt the reality of atomic lattices underlying crystalline solids and at the same time yielded quantitative geometric information about the relative positions of the atoms in solids. It constituted the beginning of CMP on a microscopic scale of 10^{-8} cm.

Building on von Laue's work, M. Born and coworkers (in the 1910s) developed a simple, classical, partially predictive theory of the cohesive energy of alkali halide crystals. Their chemical composition was known to be of the form A^+B^- (where A = alkali and B

=halogen), and the geometric arrangements were known from x-ray experiments. The theory postulated the existence of pairwise interactions consisting of the known long-range Coulomb interactions between the charged ions and short-range repulsions between nearest neighbors, phenomenologically characterized by two parameters: strength and range.

When the two parameters were fitted to the known lattice parameters and elastic bulk moduli, the calculated cohesive energies were in quantitative agreement with experiment on the ~3% level.

Another success based on von Laue's demonstration of the existence of atomic crystal lattices was Born's theory of classical lattice vibrations (1910s). It was based on a Hamiltonian of the form

$$H=\sum_l \frac{M_l}{2}\dot{u}_l^2+\Phi(u_1,u_2,\ldots,u_N), \tag{2.1}$$

where M_l and u_l are the masses and displacements of the atoms, labeled by l, the first sum is the kinetic energy, and Φ is the potential energy, which was expanded up to second order in u, making use of periodicity and other symmetries. Together with convenient periodic boundary conditions, this led to propagating normal modes of vibration with wave vectors q and frequencies $\omega_j(q)$, where j is an additional label. (For monatomic crystals $j=1,2,3$, corresponding to the three so-called acoustic modes; for polyatomic crystals there are also so-called optical modes.) This theory successfully unified the theory of the static elasticity and of long-wavelength sound waves in crystals, and further yielded, in terms of the expansion coefficients of Φ, the normal-mode frequencies $\omega_j(q)$ for arbitrarily large wave vectors q, which were not directly observed until half a century later by inelastic neutron scattering.

Attempts by P. Drude and H. A. Lorentz in the first decade of the century to understand the salient properties of metals in classical terms could not avoid major inconsistencies and had only very limited success. A crucial feature was the (correct) postulate that in a metal some atomic electrons are not attached to specific atoms but roam throughout the entire system. Their scattering by the atomic nuclei was regarded as the cause of electrical resistance. However, this theory could not explain why the resistance of metals generally dropped towards zero linearly as a function of the temperature, or why the expected substantial classical heat capacity, $\frac{3}{2}k$ for each free electron, was never observed.

On the question of ferromagnetism, the observed dependence of the magnetization density M on the applied field H and temperature T could be explained in terms of classical statistical mechanics by assuming a phenomenological effective field acting on an atomic dipole, given by $H_{\rm eff}=H+\alpha M$, where $\alpha\approx 0$ (10^3-10^4). The form of the so-called Weiss field αM (1907), could be roughly understood as due to classical dipole interactions, but the required high magnitude of α was some three orders of magnitude larger than the classical value, of order unity.

The experimental achievement of lower and lower temperatures, culminating in the liquefaction of He at 4.2 K by K. Onnes in 1908, dramatically brought to light the insufficiency of classical concepts. Thus while the classical law of Dulong and Petit, which assigned a heat capacity of 3 k to each atom in a solid, was generally in rather good agreement with experiment at sufficiently high temperatures (typically room temperature or above), all measured heat capacities were found to approach zero as the temperature was reduced towards zero. This was recognized as a critical failure of the concepts of classical statistical mechanics.

The dramatic discovery by Onnes in 1911 of superconductivity, a strictly vanishing resistivity below a critical temperature of a few degrees K, remained a major puzzle for more than four decades.

Thermionic emission from hot metal surfaces or filaments, an important subject during the infancy of electric light bulbs, was partially understood. The velocity distribution of the emerging electrons was, as expected, the classical Maxwell-Boltzmann distribution, $A\exp(-mv^2/2kT)$, where m is the electron mass and k is Boltzmann's constant, but the magnitude of A was not understood.

The photoelectric-effect, the emission of electrons from solid surfaces in response to incident light, was also very puzzling. Light of low frequency ω caused *no* emission of electrons, no matter how high the intensity; however, when ω exceeded a threshold frequency ω_0, electrons were emitted in proportion to the intensity of the incident light.

Thus we see that while the classical theory of CMP at the beginning of this century had some impressive successes, it also had two major, general deficiencies:

(1) When classical theory was successful in providing a satisfactory phenomenological description, it usually had no tools to calculate, even in principle, the system-specific parameters from first principles.

(2) Some phenomena, such as the vanishing of heat capacities at low temperatures and the behavior of the photoemission current as a function of the frequency and intensity of the incident light, could not be understood at all.

Both deficiencies were to be addressed by quantum theory with dramatic success.

III. EARLY IMPACTS OF THE QUANTUM OF ACTION ON CONDENSED-MATTER PHYSICS

As is widely known, Max Planck (1900) ushered in the new century with the introduction of the quantum of action, h, into the theory of blackbody radiation. Much less known is the fact that Einstein received the Nobel Prize not specifically for his work on relativity theory but for "his services to theoretical physics and especially for his discovery of the law of the photoelectric effect." In fact, it was in his considerations of the photoelectric effect in 1905 that Einstein developed the concept of the photon, the quantum of excitation of the radiation field with energy $\hbar\omega$, where $\hbar$ is Planck's constant divided by

2π and ω is the circular frequency. This concept was one of the most important ideas in the early history of quantum theory. It also led naturally to the resolution of the photoelectric effect conundrum: for any particular emitting metal surface the photon must have a surface-specific minimum energy, the so-called work function W, to lift an electron out of the metal into the vacuum. Thus a minimum light frequency is required.

Shortly after this great insight Einstein (1907), not surprisingly, also understood the reason why the lattice heat capacity of a solid approached zero at low temperatures. He modeled each atom as a three-dimensional (3D) harmonic oscillator of frequency $\overline{\omega}$. Again he quantized the excitation energies of each vibrational mode in units of $\hbar\overline{\omega}$, which directly yielded $3\hbar\overline{\omega}/(e^{\hbar\overline{\omega}/kT}-1)$ for the mean energy per mode at temperature T. At high temperatures, $kT \gg \hbar\overline{\omega}$, this yielded the classical result $3\,k$ for the heat capacity per atom, in agreement with the empirical high-temperature law of Dulong and Petit. But at low temperatures the Einstein heat capacity correctly approached zero.

By choosing an appropriate mean frequency $\overline{\omega}$ for a given solid, one could fit experimental results very well, except at the lowest temperatures, where the experimental lattice heat capacity behaved as T^3, while Einstein's theory gave an exponential behavior. This deficiency was repaired by P. Debye (1912), who quantized Born's lattice modes and realized that at low temperatures T only long-wavelength modes with frequencies $\hbar\omega \lesssim kT$ would be appreciably excited. The number of these modes behaves as T^3 and their typical excitation energy is of the order kT. This immediately gave the empirical T^3 law at low temperatures for the lattice heat capacity. The excitation quanta of the normal modes, characterized by a wave vector q, a frequency ω, and an energy $\hbar\omega$, were called phonons and became an indispensable component of CMP.

In these developments we observe (1) the decisive role played by the quantum of action $\hbar$; (2) the importance (in Debye's work) of long-wavelength/low-energy collective modes; and (3) the mutually fruitful interplay between CMP and other fields of science. (For example, in Einstein's work on the photoelectric effect, with quantum electrodynamics.) These features have marked much of CMP for the rest of the century.

IV. THE QUANTUM-MECHANICAL REVOLUTION

The advent of quantum mechanics, particularly in the form of the Schrödinger equation (1926), coupled with the discovery of the electron spin and the Pauli exclusion principle (1925), totally transformed CMP, as it did all of chemistry. While the Bohr theory of the hydrogen atom had brilliantly and accurately described this one-electron system, it proved to be quantitatively powerless even in the face of the two-electron systems He and H_2 let alone condensed matter systems consisting of $\sim 10^{23}$ interacting nuclei and electron. The Schrödinger equation changed all this. The ground-state energy of He was soon calculated by E. Hylleraas (1929) with a fractional accuracy of 10^{-4}, the binding energy and internuclear separation of H_2 was calculated first by W. Heitler and F. London (1927), and then by others, with accuracies of about 10^{-2} to 10^{-3}. This left no reasonable doubt that the Schrödinger equation, applied to both electrons and nuclei, *in principle* was the correct theory for CMP systems.

A very useful organizing principle, the Born-Oppenheimer approximation (1927), was soon articulated: because of the small mass ratio of electrons and nuclei, usually $m/M \sim 10^{-5}$, typical electronic time scales in molecules and presumably also in solids were much shorter than those of nuclei, in proportion to $(m/M)^{1/2}$. This led to the conclusion that the dynamics of electrons and nuclei could, to a good approximation, be decoupled. In the first stage the nuclei are considered fixed in positions $R_1, R_2, \ldots$ and the ground-state electronic energy $E_{\mathrm{el}}(R_1, R_2, \ldots)$ is determined. In the second stage the electrons no longer appear explicitly and the dynamics of the nuclei are determined by the sum of their kinetic energy and an effective potential energy given by $E_{\mathrm{el}}(R_1, R_2, \ldots) + E_{\mathrm{nuc}}(R_1, R_2, \ldots)$, where the last term describes the internuclear Coulomb repulsion.

Since all of condensed matter consists of nuclei and electrons, the field henceforth could, for most purposes, be divided into two parts: one dealing with electron dynamics for fixed nuclear positions (e.g., total energies, magnetism, optical properties, etc.), the other dealing with nuclear dynamics (e.g., lattice vibrations, atomic diffusion, etc.). Important exceptions were phenomena that critically involved the electron-phonon interaction, such as the temperature-dependent part of electrical resistance (F. Bloch, 1930) and, as discovered much later, the phonon-dependent so-called Bardeen-Cooper-Schrieffer (BCS) superconductivity (1957).

Another consequence of the small value of m/M was that, whereas typical electronic energies in solids were of the order of 1–10 eV,[1] those related to the nuclear dynamics were of the order of 10^{-2}–10^{-1} eV. Thus room temperature with $kT \approx 0.025$ eV was generally very cold for electrons but quite warm for nuclear dynamics.

Several of the major failures of classical theory when applied to metals, as described in Sec. II, were soon remedied by the combination of the new quantum mechanics with the Pauli exclusion principle. Of course, a straightforward solution of the Schrödinger equation for $\sim 10^{23}$ strongly interacting electrons was out of the question. But by boldly proposing that, at least roughly, the forces on a given electron due to the other electrons canceled those due to the nuclei, W. Pauli (1927) and, very extensively, A. Sommerfeld (1928) were led to the quantum-mechanical free-electron model of metals: Each electron was described by a plane wave $\varphi_q(r) \equiv \exp(iqr)$ and an up- or down-spin function $\chi_\sigma(\sigma = \pm 1)$. Coupled with the Pauli exclusion principle

[1]However, for metals, electronic *excitation* energies begin at zero.

and the resulting Fermi-Dirac statistics, this model naturally explained the following experimental facts: that in many simple metals, e.g., the alkalis, the magnetic susceptibility due to electronic spins was weak and nearly temperature independent (instead of, classically, large and proportional to T^{-1}); and that the electronic specific heat at low temperatures was small and proportional to T (instead of, classically, 3/2 k per electron and independent of T).

The Pauli-Sommerfeld theory represented major, fundamental progress for metals, but at the same time it left a host of observed phenomena still unexplained: For example, the fact that metallic resistance decreases linearly with temperature and that in some materials the Hall coefficient has the counterintuitive, "wrong" sign. Many of these puzzles were soon greatly clarified by replacing the *uniform* effective potential of the Sommerfeld model by a *periodic* potential reflecting the periodic arrangements of the ions, as will be discussed in the next section. A deeper understanding of the effects of the electron-electron interaction evolved much more slowly.

Another early, spectacular success of the new quantum mechanics was the unexpected explanation by W. Heisenberg (1928) of the "enormous" magnitude of the Weiss effective magnetic field mentioned in the Introduction. Heisenberg realized that the Pauli principle, which prevents two electrons of the same spin from occupying the same state, generates an effective interaction between the spin magnetic moments, quite unrelated to the classical magnetic dipole interaction and typically several orders of magnitude larger.

V. THE BAND-STRUCTURE PARADIGM

Two very significant physical effects were omitted from the Sommerfeld model of metals: the effects of the periodicity and other symmetries of the lattice, and the effects of the electron-electron interaction beyond the Hartree approximation. This section deals with the remarkable consequences of lattice periodicity.

In 1928 F. Bloch posited that electrons could be treated as independent particles moving in some effective potential $v(r)$, which of course had to reflect the periodicity and other symmetries of the lattice. This led to the important concepts of Bloch waves and energy bands, the eigenfunctions and eigenvalues of the single-particle Schrödinger equation,

$$\left(-\frac{\hbar^2}{2m}\nabla^2+v_{\rm per}(r)\right)\psi_{n,k}(r)=\epsilon_{n,k}\psi_{n,k}(r), \tag{5.1}$$

where $v_{\rm per}(r)$ satisfies $v_{\rm per}(r)=v_{\rm per}(r+\tau)$($\tau$ = lattice translation vector); $\psi_{n,k}$ is a quasiperiodic Bloch wave of the form $\psi_{n,k}(r)=u_{n,k}(r)e^{ik\cdot r}$, with $u_{n,k}$ periodic; k is the wave vector, a continuous quantum number describing the phase change from one unit cell to another, and n is an additional discrete quantum number, the so-called band index; the eigenvalues $\epsilon_{n,k}$ as a function of k reflect the periodicity and other symmetries of the lattice. They are periodic functions of k. In terms of these so-called energy bands the essence of most metallic and, as a bonus, insulating and semiconducting behavior (A. H. Wilson, 1930s) could be understood. This is illustrated for a one-dimensional crystal of periodicity a in Fig. 1. One observes that metallic electrons have excitation energies starting from zero; those of insulators and semiconductors have finite gaps. This simple categorization provided a powerful orientation for most simple solids.

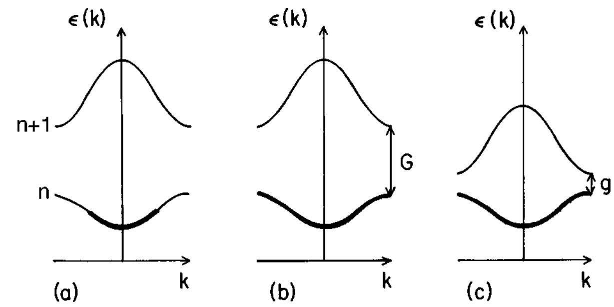

FIG. 1. Schematic energy bands: (a) metal; (b) insulator; (c) semiconductor. Heavy lines denote occupation and light lines nonoccupation by electrons at $T=0$ K. G is the insulating energy gap [$\sim O(5\text{ eV})$]; g is the semiconducting gap ($\lesssim 1$ eV).

The Bloch theory also gave a beautiful explanation of why metallic resistance approached zero at low temperatures, in spite of the presence of *individually* strongly scattering ion cores: it was quantum-mechanical coherence that caused the eigenfunctions in a periodic array of scatterers to remain unscattered. At higher temperatures, ionic positions deviated more and more from perfect periodicity, giving rise to increasing resistance.

This picture also allowed an elegant explanation by Peierls (1929) of the paradoxical Hall effects with the "wrong" sign. For example, if some traps capture electrons from the valence band of a semiconductor [Fig. 1(c)] into localized bound states, this introduces holes into the top of the previously filled valence band, which behave precisely like particles with positive charge $+e$.

Even in the absence of quantitative knowledge of $v_{\rm per}(r)$ a whole host of phenomena could now be studied and, with the help of some experimental input, understood qualitatively or better. A major tool was the quantum-mechanical transport equation (modeled after the classical Boltzmann equation):

$$\frac{\partial f(r,v,t)}{\partial t}=\left.\frac{\partial f(r,v,t)}{\partial t}\right|_{\rm drift}+b-a. \tag{5.2}$$

Here $f(r,v)d\tau$ is the number of electrons in the phase-space element $d\tau=dr dv$, $(\partial f/\partial t)_{\rm drift}d\tau$ is their net drift into a fixed $d\tau$ due to their velocity v and to their acceleration $\dot{v}$, produced by external fields; $b-a$ describes changes in f due to collisions with lattice vibrations or defects which take electrons into and out of $d\tau$.

This equation gave considerable microscopic insight into electrical and thermal conductivities, σ and K. For the venerable universal Wiedemann-Franz constant, first experimentally discovered in 1853, it led to the result

$$\frac{K}{\sigma T}=\frac{\pi^2}{3}\left(\frac{k}{e}\right)^2, \tag{5.3}$$

in good agreement with classical theory and experiment. While σ and K individually depend strongly on specifics, including the collision processes, which are roughly describable by a mean free path l between collisions, the ratio depends only on classical fundamental constants. Electrothermal effects, named after Thomson and Seebeck, could also be successfully described.

Finally, optical properties of solids, including the origin of color, could be understood as due to transitions of electrons between occupied and unoccupied states of the same k but different band quantum number n (see Fig. 1).

A. Total energies

While, as we have seen, the Bloch picture was extremely useful for many purposes, it did not seriously address the extremely important issue of total electronic energies $\mathcal{E}(R_1,R_2,\ldots)$ as a function of the nuclear configuration. For insulators and semiconductors there existed some alternative strategies, e.g., the approach by Born for ionic crystals like Na^+C^- (see Sec. II) and, since the 1930s, L. Pauling's concept of the chemical bond for covalent crystals like Si. This left as a major challenge the work of understanding the total energies of metals.

Here a great advance was achieved by E. Wigner and F. Seitz in their work on the alkali metals beginning in 1933. By a bold, physical argument they proposed that the periodic potential for the valence electrons in any crystal cell l be taken as $v(r-R_l)$, equal to the effective potential for the valence electron in an isolated atom located at R_l. The latter had been accurately determined by comparison with the observed energy spectra of isolated atoms. This was a major step beyond the formal theory of Bloch electrons: The abstract $v_{\text{per}}(r)$ was replaced by a specific, independently determined potential.

In order to obtain the total energy as a function of the lattice parameter, they first calculated the sum of the noninteracting Bloch energies in the periodic potential $\Sigma_l v(r-R_l)$ and argued that a Hartree-like intracell Coulomb interaction energy E_H was approximately canceled by the so-called exchange and correlation energy E_{xc}. (By definition E_{xc} is the difference between the exact physical energy and the energy calculated in the Hartree approximation.) Their results were generally in semiquantitative agreement with experiment for cohesive energies, lattice parameters, and compressibility.

Subsequently they actually estimated the neglected energies E_H and E_{xc} for a uniform electron gas of the appropriate density, confirmed the near-cancellation, and obtained similar results. This involved the first serious many-body study of an infinite system, a uniform interacting electron gas, by E. Wigner in 1938. He arrived at the estimate of $\epsilon_c=-0.288/(r_s+5.1a_0)$ atomic units for the correlation energy per particle, which needs to be added to the Hartree-Fock energy. [r_s is the so-called Wigner-Seitz radius given by $(4\pi/3)r_s^3=(\text{density})^{-1}$, and a_0 is the Bohr radius.] This result has withstood the test of time extremely well. At this time the best available results have been obtained by numerical, so-called Monte Carlo methods (D. Ceperley and others, 1980s) with an accuracy of $\sim 1\times 10^{-2}$.

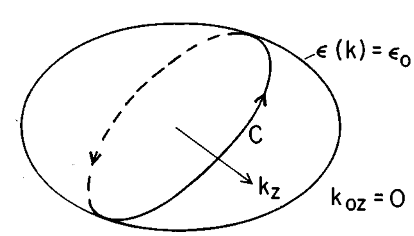

FIG. 2. The Onsager orbit C on the plane $k_x=k_{0z}$ and on the constant-energy surface $\epsilon(k)=\epsilon_0$.

The Wigner-Seitz approach was, of course, very soon tried out on other metals, e.g., the noble metals and Be, which were not so similar to uniform electron gases, but generally with much less success. Not until the advent of density-functional theory 30 years later, in the form of the Kohn-Sham theory were the Bloch and Wigner-Seitz approaches unified in great generality (see Sec. VII).

From the time of Bloch's original paper in 1928 up until the 1950s the band-structure paradigm provided an invaluable conceptual framework for understanding the electronic structure of solids, but very little was known quantitatively about the band structures of specific materials, except the very simplest, like the alkali metals. This now changed dramatically.

B. Fermi surfaces of metals

In a beautiful short note L. Onsager (1952) considered the dynamics of a crystal electron in a (sufficiently weak) magnetic field $B=(0,0,B_z)$. In momentum space it is governed by the semiclassical equations

$$\hbar\dot{k}=\frac{e}{c}[v(k)\times B], \tag{5.4}$$

where $v(k)$ is the velocity,

$$v(k)=\hbar^{-1}\nabla_k\epsilon_k. \tag{5.5}$$

Combining this with purely geometric considerations, Onsager showed that an electron starting at a point k_0 with energy $\epsilon(k_0)$ will return to k_0 cyclically with a so-called cyclotron period

$$T_c=\left(\frac{c}{eB_2}\right)\hbar^2\frac{dS}{d\epsilon}, \tag{5.6}$$

where S is the area in k space enclosed by the curve C, which is generated by the intersection of the plane $k_z=k_{0z}$ and the surface $\epsilon(k)=\epsilon$. (See Fig. 2.) Using this result he showed further that for *any* band structure, no matter how complex, the magnetization is an oscillatory

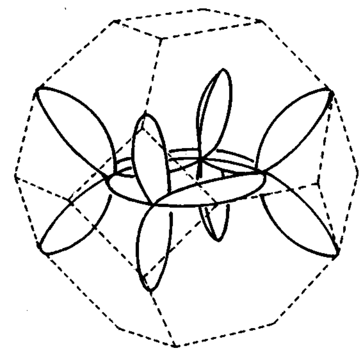

FIG. 3. ''Weird'' topologies of Fermi surfaces: a portion of the Fermi surface of aluminum.

function of B_z^{-1} (the so-called de Haas-van Alphen-Shubnikoff oscillation) with a period (and this was new) given by

$$\Delta\left(\frac{1}{B_z}\right)=\frac{2\pi e}{c}\frac{1}{S}; \tag{5.7}$$

here S is a maximum, minimum, or other stationary cross-sectional area, perpendicular to B, of the so-called *Fermi surface*, in k space, which, by definition, encloses all occupied k vectors. Thus, by tilting the direction of the magnetic field, one could measure geometrically cross-sectional areas with different normals! This was impressively accomplished by D. Schoenberg and his group in Cambridge in the 1950s.

These cross-sectional areas, combined with known symmetries, some rough guidance from approximate band calculations, and the general Luttinger theorem (see Sec. VII), which fixed the volume enclosed by the Fermi surface, generally permitted unique and accurate determination of the entire shape of the Fermi surfaces (see Fig. 3) and ushered in a geometric/topological era of CMP. Since, because of the Pauli exclusion principle, low-energy/low-temperature[2] electronic excitations of metals involve electrons and holes near the Fermi surface, its empirical determination represented a major advance in metal physics.

C. Angle-resolved photoemission and inverse photoemission

Another experimental technique that has shed great light on band structures, both of metals and of nonmetals, is angle-resolved photoemission, which began with the work of W. E. Spicer (1958). This was followed, in 1983, by inverse photoemission. The former explores occupied states, the latter unoccupied states. Photoemission and inverse photoemission have been used to study bulk bands and surface bands. (See also Sec. VI.)

[2]As explained in Sec. IV, ''low'' means typically $\ll 1$ eV or 10^4 K, in fact rather ''high.''

In bulk photoemission a photon of energy $\hbar\omega$ is absorbed by an electron in an occupied state (n,k), which makes a transition to an unoccupied state (n',k'). Here $k'=k$ because the photon momentum can be neglected. Energy conservation requires that

$$\epsilon_{n',k}=\epsilon_{nk}+\hbar\omega. \tag{5.8}$$

$\hbar\omega$ is chosen large enough so that $\epsilon_{n'/k}$ can be taken as free-electron-like and (apart from an additive constant) as known. The external momenta of those final electrons that reach the surface and surmount the dipole barrier give direct information about the initial momenta k and energies $\epsilon_{n,k}$. In this way occupied energy bands of many materials have been directly determined by photoemission.

In inverse photoemission external electrons of known k_e and $\epsilon_e=\hbar^2k_e^2/2m$ may penetrate the surface, occupy unoccupied Bloch states (n',k), and then emit a photon whose frequency ω is again given by Eq. (5.8). From such measurements direct information about the unoccupied $\epsilon_{n',k}$ can be obtained.

A most helpful theoretical tool, at least for so-called simple metals whose valence electrons have at least some resemblance to free electrons, was the concept of the weak, effective pseudopotential $v_{ps}(r)$, due to H. Hellmann (1936) and especially to J. C. Phillips (1958), and widely used by the group of V. Heine (1960s). For the valence electrons, the weak v_{ps} had an effect equivalent to the actual Bloch potential, which is very strong near the nuclei. v_{ps} could be characterized by two or three independent Fourier coefficients and, as in the case of aluminum (see Fig. 3), even a very complicated Fermi surface could be accurately fitted everywhere.

The band-structure paradigm has remained the most important basis for understanding the electronic structure of solids. Even when there are significant, but not radical, effects of interaction and/or nonperiodicity it is usually an indispensable starting point. Theoretical materials science since the 1930s, and especially since about 1970, has increasingly made quantitative use of it. The great silicon revolution of the second half of this century might not have been possible without it.

VI. SURFACES AND INTERFACES

The 20th century has seen a transformation of surface science and, more generally, of two-dimensional (2D) science, made possible by major advances in vacuum technology combined with various techniques like atomic and molecular deposition, beam writing, and etching.

In parallel with the dramatic advances in surface science, the physics of interfaces between two bulk phases has also made major progress. Perhaps its most important practical applications are to highly controlled artificial layer structures in semiconductors and magnetic materials.

At the beginning of the century surfaces were already of great practical interest for the mitigation of corrosion, heterogeneous catalysis, thermionic emission in light bulbs and vacuum tubes, electrochemistry at solid-liquid interfaces, friction, etc. However, the best available vacua were only $\sim 10^{-4}$ torr, and most surfaces at room temperature and below were covered by unknown layers of adsorbed atoms and/or molecules. Atomically clean surfaces were generally achievable only at the highest temperatures when, in favorable circumstances, adsorbates would evaporate.

In 1927 C. J. Davisson, and L. H. Germer demonstrated diffraction patterns in the scattering of electrons by crystal surfaces, an experiment of double significance: it was a direct demonstration of the reality of the wave nature of electrons and the beginning of surface diagnostics on an Angstrom scale—analogous to von Laue scattering of x rays by bulk crystals.

Today, at the end of the century, vacua of 10^{-10} torr can be routinely generated. This has made possible the preparation of atomically clean surfaces or of surfaces covered in a controlled way with an accuracy of $\sim$0.01 monolayers. Since the 1980s we have acquired the ability to check the structural and chemical condition of surfaces point by point with an accuracy of order 10^{-1} Å(!), using electron tunneling and force microscopies (G. Binnig and H. Rohrer, 1980s). A host of other, less local, but powerful diagnostic techniques such as Auger electron spectroscopy, x-ray and neutron reflectometry, low- and high-energy electron diffraction, x-ray and ultraviolet photoelectron spectroscopy (XPS, UPS), have also been highly developed; so have a wide variety of surface treatments.

A major factor driving surface science since the middle of this century has been the dramatic rush to greater and greater miniaturization. While in a structural steel beam surface atoms constitute a fraction $\sim 10^{-9}$ of all the atoms, in a miniaturized semiconducting structure used in contemporary devices the ratio of interface atoms to bulk atoms is on the order of $\sim 10^{-2}$, and surface properties often dominate device performance. For example, in the late 1940s, surface and interface physics was a major aspect of the invention of the first point-contact transistors as well as of later versions, and it has remained a critical element of the subsequent quantum electro-optical revolution, which continues in full swing.

The band-structure paradigm for bulk solids, when appropriately modified, also became a valuable guide for understanding the physics of surfaces and interfaces. A crystal terminated in a plane parallel to $z=0$ remains periodic in the x-y plane so that the electronic states have the 2D Bloch property

$$\psi_{n,k}(r+\tau_j)=e^{ik\cdot\tau_j}\psi_{n,k}(r), \quad j=1,2, \tag{6.1}$$

where the τ_j are 2D lattice vectors in the x-y plane.

Two important qualitatively new features were observed. First, the atoms near the surface may reconstruct so that the symmetry is lowered not only in the z direction (no surprise) but also in the x-y plane. Dimerization is the simplest example, but important cases with much larger supercells have been observed, for example, on the surfaces of silicon. Such reconstruction can radically affect surface electronic structure and interactions with adsorbed atoms or molecules.

Secondly, some electronic wave functions near a surface penetrate into the bulk and some are localized near the surface. The localized surface states, first theoretically proposed by I. Tamm in 1932, are occupied or empty depending on whether they are below or above the electronic chemical potential μ. They have three quantum numbers, k_x, k_y, and m, which describe Bloch-type propagation in the x-y plane and localization along the z direction. If, for example, only states with $m=0$ have energies below μ, we can have a 2D metal. (The motion in the z direction is "frozen out.")

There are two important, microscopically averaged surface properties, the surface energy per unit area ($\frac{1}{2}$ of the cleavage energy) and the electric dipole barrier. They play important roles for thermodynamic and electronic considerations. An early, rather rough, theory for these quantities constitutes the thesis of John Bardeen (1936). Much later, beginning with work by N. Lang and W. Kohn (1970), good quantitative results were obtained by the use of density-functional theory (see Sec. VII).

Another major area of surface science is the joint domain of chemistry and physics: the study of atoms and molecules in various interactive relationships with surfaces—collisions, adsorption, desorption, and, for adsorbed atoms and molecules, diffusion, physical and chemical interactions, and chemical reactions. Again, controlled ultrahigh vacua and deposition methods, combined with the battery of mostly post-1950 diagnostic techniques, have led to spectacular advances. Improved catalytic conversion of noxious automobile emissions and cracking of crude oil are examples of important applications.

Systems involving both surfaces and molecules present a difficult challenge to theory because of the absence of symmetry simplifications and the significant involvement of many (10–10^3) atoms. The first detailed quantitative results using density-functional theory for H_2 on solid surfaces have just begun to appear (A. Gross, M. Scheffler, and others, 1995). These computations would not be possible without the use of the highest state-of-the-art computing power.

The foregoing paragraphs have treated surfaces essentially as two-dimensional versions of bulk crystals. However, we shall see shortly (Sec. X) that, in the presence of disorder and/or electron-electron interaction, entirely new phenomena, such as the quantum Hall effect, exist in two dimensions. In fact, the recognition of the highly nontrivial role of *dimensionality* is one of the hallmarks of CMP in this century. Probably the first example of this recognition is the surprising observation by R. Peierls (1935) that while a harmonic 3D crystal has positional correlation over an infinite range at any finite temperature, this is true of a (free-floating) 2D crystal *only* at $T=0$.

A final remark about surface science and surface technology. The former, both in the laboratory and in the theorist's thinking, deals mostly with as perfect systems as possible. On the other hand, in technological applications surfaces are generally very imperfect, both structurally and chemically. Nevertheless, concepts developed by idealized surface science have been very important guides for practical applications.

VII. MODERATE AND RADICAL EFFECTS OF THE ELECTRON-ELECTRON INTERACTION

The many great successes of the band-structure paradigm in accounting, at least qualitatively, for electronic properties and phenomena in solids strongly implied that interactions often only alter the Bloch picture quantitatively without changing its qualitative features. This turned out to be partly true and partly false. We begin with an account of some moderate interaction effects.

A. Landau Fermi-liquid theory

In 1956, a few years after the first reliable experiments on He^3, the rare *fermionic* isotope of helium, L. Landau published his famous, largely heuristic, theory on the low-energy properties of a uniform gas of mutually repelling neutral fermions. He concluded that there are low-lying excited states that can be described as arising from the ground state by the addition of quasiparticles and quasiholes with momenta $k(-k)$ and energies $\epsilon_k = (\hbar^2/2m^*)|k^2-k_0^2|$, where k_0 is the Fermi momentum of the noninteracting gas, $|k \gtrless k_0$ for electrons and holes, respectively, and $|k-k_0| \ll k_0$; m^* is an effective mass, which for He^3 turned out to be 3.0 times the mass of the bare He^3 atom. Although the quantitative renormalization of the mass was large, these excitations were in 1-to-1 correspondence with those of noninteracting fermions.

Much more interestingly, Landau also introduced an effective spin-dependent interaction $f(\theta)$ between low-energy quasiparticles with momenta k and k', where θ is the angle between k and k'. $f(\theta)$ is usually parametrized in terms of a few spin-dependent angular expansion coefficients $F_0, F_1, \ldots$. While these coefficients are not needed for the low-temperature specific heat, they do enter significantly into the spin susceptibility and compressibility, which, for He^3, are also strongly renormalized by factors of 9.1 and 3.7, respectively. But the most interesting result of Landau's theory was that these interactions lead to a new dynamical collective mode of coherent, interacting quasiparticle-quasihole pairs, the so-called zeroth sound mode, with a linear dispersion relation, $\omega = sq$. The velocity s was also expressible in terms of the F_l. The experimental confirmation of this mode by J. Wheatley and co-workers (1966), and the consistency of the experimentally overdetermined parameters F_0 and F_1 (F_2, etc. are very small) was a great triumph for this theory.

Implications for an interacting electron gas were immediately recognized, the important differences being the presence of the periodic potential due to the nuclei and the long range of the Coulomb interactions. The effect of the latter on a degenerate uniform electron gas had, in fact, been previously shown by D. Pines and D. Bohm (1952) to lead to the collective plasma mode, with dispersion approximately given by $\omega^2 = \omega_p^2 + \frac{3}{5} v_F^2 q^2$; here ω_p is the classical plasma frequency and v_F is the Fermi velocity. Landau's theory provided a unification of the theories of neutral and charged uniform Fermi systems.

Electrons under the influence of both a periodic potential and the Coulomb interaction were soon studied perturbatively using the newly developed machinery of many-body theory. A key result was obtained by J. M. Luttinger (1960), who showed formally that, to all orders in perturbation theory, a sharply defined Fermi surface $k(\theta,\varphi)$ persisted in k space and, though its shape was altered by the interactions, the k-space volume enclosed by it remained unchanged, determined entirely by the mean density $\bar{n}$ of the electrons. This so-called "Luttinger theorem" has been very helpful in studies of metals with complex Fermi surfaces.

B. Strong magnetism

We have already mentioned Heisenberg's qualitative realization that the Pauli exclusion principle combined with the electron-electron interaction can bring about a strong effective interaction between spins. By a dimensional argument it is of the form $\gamma e^2/a$, where a is an effective interelectronic distance and γ is a dimensionless constant. Heisenberg's approach was well suited for insulators describable by localized orbitals, but the traditional ferromagnetic materials, Fe, Ni, and Co are metals.

For these F. Bloch contributed an early insight (1929). He compared the energies of two possible Sommerfeld ground states: (1) A paramagnetic state with each plane wave $k \leq k_0$ (k_0=Fermi wave number), occupied by both a spin-up and a spin-down electron. (2) A ferromagnetic state, in which all spins are pointing in the same direction, say z, and hence each plane wave is occupied by at most one electron. Thus the maximum occupied k is now increased by a factor of $2^{1/3}$, and the kinetic energy by a factor of $2^{2/3}$. However, the exchange energy due to the Coulomb repulsion of the electrons favors the ferromagnetic state, since the Pauli exclusion principle keeps *all* electrons apart from each other, whereas in the paramagnetic state electrons with opposite spin are not kept apart. At sufficiently low density this effect prevails and the ferromagnetic state has the lower energy. (Bloch ignored correlation effects, which, in fact, change the conclusion for the uniform electron gas.) This was the beginning of the concept of itinerant magnetism in metals, soon considerably developed by E. C. Stoner, J. C. Slater, and others.

Bloch soon returned to another aspect of magnetism (1934), this time Heisenberg's localized type. Starting from N atoms, in a perfectly spin-aligned ground state Ψ_0 with a total z spin equal to $\frac{1}{2}N$, he observed that there were gapless, propagating excited states, spin waves of the form $\Phi_k = \Sigma_l \exp(ikR_l) S_l^- \Phi_0$, where S_l^- is the spin operator, which turns over the spin at the site R_l. Their energy spectrum had the form $\epsilon_k \propto k^2$ and their contribution to the low-temperature heat capacity was $\propto T^{3/2}$. The most compelling confirmation of Bloch's spin waves came in the 1960s by means of inelastic neutron scattering, which directly measured the dispersion relation ϵ_k and found remnants of spin waves even above the critical temperature where the average magnetization vanishes.

Since its quantum-mechanical beginning in the late 1920s, the field of strong magnetism has had an explosive growth. One of the most interesting events was the prediction by L. Néel, and the subsequent experimental confirmation (1930s), of a new kind of magnetism, later called antiferromagnetism: the lattice consists of two equivalent sublattices A and B, with all A atoms carrying a magnetic moment m_A and all B atoms carrying the moment $m_B = -m_A$. Thus the total magnetization, in contrast to ferromagnetism, vanishes. However, both the magnetic susceptibility and low-temperature specific heat reveal the "hidden" strong sublattice magnetizations. Again, the incontrovertible proof was provided by direct observation of two magnetic sublattices in elastic neutron scattering (C. G. Shull and J. S. Smart, 1949). Following Néel's work many complex magnetic structures were discovered, especially among heavy metals and metal compounds.

In the first half of the century the most important practical applications of magnetism were electromagnets: generators, motors, electromagnetic relays, etc. Today these are joined by magnetic memory devices, read-in and read-out devices, magnetic layer structures with "giant" magnetoresistance, etc. The field has entered a new, very active phase.

C. Density-functional theory

By the 1960s quantum-chemical methods had been very successful in calculating properties of N-electron systems, with N up to $O(10)$. However, in condensed matter physics $N = O(10^{23})$, and even in the smallest representative clusters $N = O(10^2 - 10^3)$. Density-functional theory (DFT), introduced by P. Hohenberg, W. Kohn, and L. J. Sham in 1964–1965, provided a practical, new approach to electronic structure, applicable also to large-N systems. Density functional theory is couched in terms of the electron density $n(r)$ [or, for magnetized systems, spin densities $n_\sigma(r)$, $\sigma = \pm 1$] instead of the many-electron wave function Ψ. It leads to the Kohn-Sham self-consistent equations, similar to the Hartree equations, in which, however, exchange and correlation effects are included (in principle, exactly) by the addition of the exchange-correlation potential of $v_{\text{xc}}(r)$.

The theory allows parameter-free calculations of densities and spin densities, ground-state energies, as well as related quantities such as lattice structures and constants; elastic coefficients; work functions, surface energies, and atom-surface interaction energies; phonon dispersion relations; magnetic moments, etc. Accuracies typically range from 1–20 % depending on the context; geometries emerge very accurately, typically $\pm 1\%$. Density functional theory is, in principle, exact, but in practice requires an approximation for the exchange-correlation energy E_{xc}. The simplest, the "local-density approximation" (LDA), rests on accurate Monte Carlo calculations of a uniform, interacting electron gas. The scaling of the computation time with N is a relatively very favorable N^α, where $1 \leqslant \alpha \leqslant 3$, so that calculations for finite systems with $N \approx 100-1000$ have been quite feasible.

Density functional theory has become the method of choice for calculating electron densities and energies of most condensed-matter systems. It also leads to nominal energy bands, which are usually a very useful approximation to the physical bands. In the 1980s and 1990s the LDA was greatly improved by density-gradient corrections (A. Becke, J. P. Perdew, and others). Since about 1990, density-functional theory has also been widely used by theoretical chemists, particularly for large, complex molecules and clusters.

D. Collective excitations

Collective excitations are an important hallmark of many-body systems. They depend for their very existence on particle-particle interactions and are delocalized excitations of the entire system. Familiar examples are the vibrations of molecules or of crystal lattices, whose nature has been well understood since about 1910.

We have already mentioned a few other condensed-matter examples: zeroth sound in He^3 and plasmons in a uniform electron gas, as well as spin waves in magnetic systems. The latter represent a separation of electronic spin and charge, already well understood by F. Bloch, who about 1930 is said to have remarked: "If electrons can hop from one atom to another why not spins?" (the condensed matter version of Lewis Carroll's Cheshire cat and his grin).

A collective excitation in insulators was proposed in 1931 by J. Frenkel, now called the Frenkel exciton. It is most easily visualized for a lattice of distinct neutral atoms, say Ar, and is formally analogous to spin waves. Let Ψ_0 be the ground state of the system, with all atoms in their ground state, and let Ψ_l be the state in which the atom at site R_l is in the first excited state. The states $\Psi_1, \Psi_2, \ldots$ are degenerate and, because of the proximity of the atoms, they interact. The correct linear combinations reflecting the lattice periodicity are the excitation waves, or *excitons*, $\Psi_k \equiv A e^{ik \cdot R_l} \Psi_l$ with wave vector k, whose energies ϵ_k are k dependent. Excitons are the lowest excited states of insulators. Being neutral, they carry no electric current. They were first clearly identified in optical spectra in the 1940s.

A different view of excitons is due to G. Wannier (1937). If one ignores interactions between electrons, the lowest-lying excited states, Ψ_{k_c, k_v}, have an electron (k_c) near the bottom of the conduction band and a hole (k_v), with a positive charge, near the top of the valence band. Because of their Coulomb attraction the electron and hole can form lower-lying, traveling bound states with total wave number k, a kind of condensed matter positronium. These are again the excitons. The Frenkel picture is appropriate for small excitons, where the electron and hole are tightly bound to each other (e.g., in Na^+Cl^-); the Wannier exciton is more appropriate in the opposite limit (e.g., in *Si*).

Excitons, consisting of two fermions, are bosons and in principle should exhibit Bose condensation (L. V. Keldysh, 1960s); however, so far this has escaped clear identification.

E. Radical effects

The foregoing paragraphs dealt with what I call moderate effects of the electron-electron (e-e) interaction, when a model of noninteracting effective electrons and/or holes is a good starting point. There are, however, many condensed-matter systems in which this is not the case, whose history will now be briefly addressed. E. Wigner (1938), considering the ground state of a dilute gas of electrons moving in a neutralizing positive charge background, observed that the free-electron kinetic energy per electron behaved as r_s^{-2} while the e-e repulsive energy behaved as r_s^{-1}, and thus in the dilute limit the latter would prevail [r_s, previously defined, is proportional to $(\text{density})^{-1/3}$]. He concluded that the electrons would form an ordered lattice and perform small zero-point vibrations around their equilibrium positions. This so-called *Wigner lattice* was an early indication that there may be condensed matter systems or regimes for which the band paradigm is overwhelmed by the effects of the e-e interaction. (Much later a 2D Wigner crystal was observed for electrons trapped on the surface of liquid He^4.)

In 1949 N. Mott noted that the compound NiO_2 was an insulator, although, based on the number of electrons per unit cell and the Bloch band paradigm, it should be a metal with a half-filled band. This led him to consider a model consisting of H atoms forming a simple cubic lattice with adjustible lattice parameter a. He adopted a tight-binding point of view, in which the many-body wave function is entirely described in terms of atomic 1s orbitals ω_l, centered on the nuclei R_l. He then estimated the effects on the total energy due to electrons' hopping onto neighboring sites. By giving electrons more room, one would cause their kinetic energy to be reduced, while the Coulomb repulsion energy of the electrons would increase by an energy U for each double occupancy. He concluded that, when a exceeded a critical value a_c, this system, which in band language has a half-full band, would nevertheless become an insulator, now called a *Mott insulator*. The internal structure of this insulator is quite different from that of the filled-band Bloch insulator. (There is an obvious relationship between Wigner's and Mott's considerations.)

These ideas were further developed by J. Hubbard (1963) in the so-called Hubbard model of interacting electrons with the Hamiltonian

$$H=\sum_{l\sigma} \epsilon_{l\sigma} n_{l\sigma} + t \sum_{\substack{l,l' \\ (nn)}} c^*_{l\sigma} c_{l'\sigma} + U \sum_i n_{l\uparrow} n_{l\downarrow} . \qquad (7.1)$$

Here l and σ denote sites and spins; the sum over l and l' is over nearest neighbors; $\epsilon_{l\sigma}$ and t are the site-diagonal and hopping energies; U describes the additional energy due to double occupation; and $n_{l\uparrow}, n_{l\downarrow}$ denote the numbers of spin-up (-down) electrons on site l. This Hamiltonian interpolates between isolated atoms ($t=0$) and noninteracting, itinerant electrons ($U=0$). Approximate solutions for U/t finite do indeed yield the Mott metal-insulator transition for a critical value of U/t. But the model has allowed many extensions to more complex systems (e.g., high-T_c superconductors), excitations, defects, effective spin Hamiltonians, magnetic phenomena, longer-range interactions, etc. It has been a valuable guide for understanding systems such as oxides, sulfides, and many other compounds which, under the band paradigm, would be described as narrow-band materials.

In a similar spirit P. W. Anderson (1961) had earlier proposed that an isolated impurity atom, immersed in and hybridized with a sea of conduction electrons, could, due to an intra-atomic e-e repulsion U, develop a finite magnetization.

In 1964 T. Kondo considered the effect of such a localized impurity spin on the scattering of conduction electrons and surprisingly found (in low-order perturbation theory) very unusual behavior (paralleling earlier experimental findings) below what is now called the Kondo temperature T_K. In fact, for $T \ll T_K$ the impurity spin forms a singlet state with the conduction electrons, and its magnetic susceptibility vanishes.

Since the 1980s so-called "heavy-fermion" materials have attracted much attention. They are associated with incompletely filled 4f and 5f shells such as in the Ce and U compounds $CeAl_3$ and UPt_3. At very low temperatures $[T=O(1\,\text{K})]$ this class of materials displays a linear specific heat, γT, with γ values corresponding to an enormously enhanced effective mass, $m^* =(10^2-10^3)m$! They exhibit a great variety of electronic behavior, including paramagnetism (with a huge magnetic susceptibility), various forms of cooperative magnetism with very small magnetic moments, insulating behavior, and superconductivity. It is generally believed that their properties reflect the opposing tendencies of the Kondo mechanism, which tends to suppress localized f moments, and an indirect, so-called RKKY (Ruderman-Kittel-Kasuya-Yosida), interaction between f moments on different sites, via the conduction electrons. Attempts to understand their behavior usually employ a generalization of the Anderson/Hubbard Hamiltonians, including a repulsive energy U for f electrons on the same site and an f-electron/conduction-band hybridization term. In addition to more standard techniques, μSR (positive muon spin rotation and relaxation) has provided a wealth of information about local magnetic fields in these compounds.

At century's end, heavy-fermion systems, together with high-T_c superconductors, represent major challenges to condensed-matter theorists.

VIII. MODERATE AND RADICAL BREAKDOWNS OF LATTICE PERIODICITY

Condensed matter consists, by its nature, of very many significantly interacting atoms. The *periodicity* of crystal lattices was the simplifying feature which, beginning in the decades 1910–1930, gave physicists and chemists the courage to undertake experiments and construct theories that ultimately led to an impressive understanding of crystalline solids on an atomic scale. The periodicity paradigm has remained invaluable ever since. At the same time perfect periodicity for an extensive system is a thermodynamic fiction. For, as is easily shown, at any finite temperature T, the introduction of a small concentration of nonperiodic defects, while raising the internal energy U, also increases the entropy S in such a way that the free energy, $F \equiv U - TS$, is reduced. Thus it is not surprising that, over the course of the century, nonperiodic systems have also been intensively studied, starting with dilute point defects in periodic lattices (1920s) and later including systems for which the periodicity paradigm has little if any relevance, for example, liquids, amorphous solids, and fractals.

A crucial opposite development also took place. Poorly controlled, high levels of structural and/or chemical disorder can prevent meaningful scientific studies or dependable applications. The semiconductors Si and Ge are a case in point. Their electrical and low-frequency optical properties are largely due to very small concentrations of chemical impurities, which could not be adequately controlled until the middle of the century. As a result their applications, e.g., to crystal radios, were highly unreliable. With the advent of zone refining (1940s) structurally excellent crystals of Si and Ge were grown in which the concentration of critical impurities, like B and P, could be controlled at the unprecedented fractional level of $\sim 10^{-8}$. This dramatic accomplishment was indispensable for the semiconductor revolution of this century. Also, the writer recalls seeing (in about 1955) a Si whisker with a *single* structural defect, a so-called screw dislocation, around which it had grown, whose resistance to fracture was many orders of magnitude higher than that of "normally" grown Si.

A. Point defects

Ionic crystals are among the easiest to grow with high structural and chemical perfection and therefore were early subjects of study. The most common point defects are so-called vacancies (missing atoms or ions), followed by interstitials (additional atoms or ions, located in interstices of the periodic lattices). Local electronic neutrality is energetically very strongly favored. Thus in Na^+Cl^- a Cl^- vacancy is typically either paired with a nearby neutralizing Na^+ vacancy or it traps a neutralizing electron. The latter defect was identified as the previously empirically discovered F-center (F=Farbe, i.e., color), which lent a distinctive color to Na^+Cl^- crystals containing them. This and other similar centers became the subject of detailed optical studies and concomitant theoretical work—perhaps the first quantitative application of quantum mechanics to a complex condensed matter system (experiments by R. W. Pohl and coworkers, 1920s; theory by J. H. De Boer and others, 1930s). In roughest approximation the F-center may be regarded as a condensed-matter version of a hydrogen atom, with the net positive charge $+e$, due to the removal of Cl^-, playing the role of the proton in hydrogen. Of course the positive charge is effectively spread out over the volume of the vacancy, and beyond it the electron moves not through a vacuum but through the dynamical ions of the Na^+Cl^- lattice. The primary absorption at 2.7 eV, in the visible spectrum is the greatly modified analog of the 10.2 eV $1s \rightarrow 2p$ line in H. The temperature dependence of the linewidth could be quite well explained as due to the vibrations of the nearby ions. Detailed optical studies of this and other so-called color centers, associated with a single structural point defect or with complexes of several structural defects, were later very effectively complemented by the invention of nuclear and electron spin-resonance techniques and remained an important, highly quantitative field of CMP into the 1960s.

Beginning in the 1950s, analogous but even more precise studies were undertaken of the so-called donor and acceptor defects in covalent semiconductors. These studies were greatly stimulated by the invention of the transistor and decisively aided by the independent measurement of the effective masses of low-energy electrons and holes using so-called cyclotron resonance in external magnetic fields (B. Lax and others, 1950s). Donor and acceptor defects in Si are created by replacement of a four-valent Si atom by, say, a five-valent P atom or a three-valent B atom. Four electrons from the P atom become part of the bonding structure (or filled "band") of the Si matrix. The extra charge $+e$ on the phosphorous nucleus can weakly trap the extra electron of the P atom in one of several hydrogenlike donor states, or "donate" it to the continuum of conduction-band states. An analogous situation obtains for a B acceptor and positive holes. At room temperature the donors and acceptors provide conducting electrons and holes, while in their absence most semiconductors are effectively insulating. The solid-state spectroscopy of trapped donor electrons and acceptor holes has now an astonishing accuracy of $\sim 10^{-3}$. Excited states can be calculated with similarly high accuracy by parameter-free so-called "effective mass theory," in good agreement with experiment. The theory for the more tightly bound donor or acceptor ground states is not so precise.

B. Substitutional alloys

Another class of moderately nonperiodic systems is made up of the substitutional alloys A_xB_{1-x}, where A and B are elements with the same valency, similar

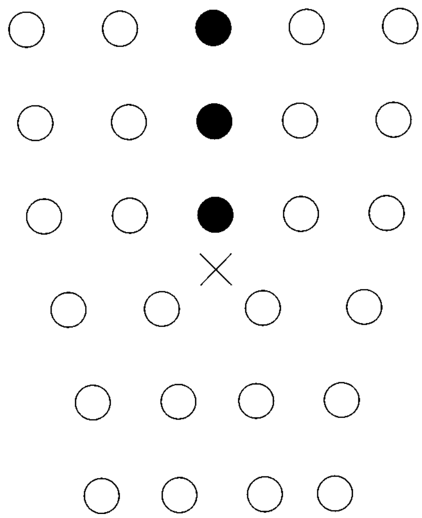

FIG. 4. Edge dislocation line (perpendicular to paper). The solid circles are part of the additional half plane.

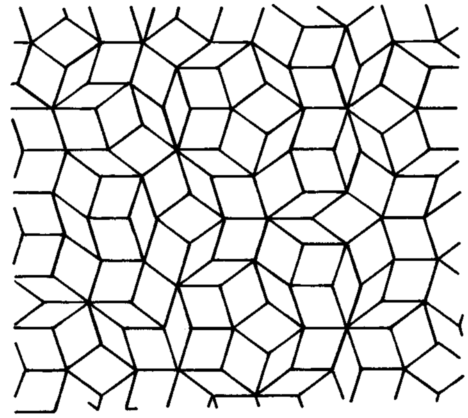

FIG. 5. Quasiperiodic two-dimensional Penrose tiling.

atomic radii, and, in their pure forms, the same crystal structure. Alloys of Cu and Au are an example. Except for special values of x, such as $x=1/4$, when ordered superlattices can form, such alloys display some degree of disorder. Early in this century the theory of such alloys was based on a simple but successful phenomenological mean-field model for the energy (W. L. Bragg and E. J. Williams, 1934): $E=N_{AA}v_{AA}+N_{BB}v_{BB}+N_{AB}v_{AB}$, where N_{AA} are the number of AA-"bonds" and v_{AA} is the corresponding bond energy, etc. Today the energies of many alloy systems have been successfully calculated by parameter-free density-functional theory (Sec. VII) with accuracies of a few percent.

The thermodynamics of the Bragg-Williams model, which is mathematically isomorphic with that of the so-called Ising model for magnetism (1925), has been the subject of intensive theoretical study ever since the 1920s. A major theoretical breakthrough was the exact analytical solution of this model in 2D by L. Onsager in 1944, showing a logarithmic singularity in the specific heat at a critical temperature T_c where long-range magnetic order disappears. For the 3D Ising model, although very precise numerical results are available today, the intensive quest for exact analytical results has so far not succeeded.

C. Dislocations and grain boundaries

In all the examples above, while there are local distortions of the lattice structure, the topology of the underlying periodic lattice remains intact. The major new concept of *dislocation*, a topological defect, was put forward in 1934, independently by G. I. Taylor, E. Orowan, and M. Polanyi, to explain the fact that permanent deformation occurs in metals (e.g., in a Cu wire) under stresses about three orders of magnitude smaller than estimated for a perfect crystal. Figure 4 shows a so-called edge dislocation, which can be thought of as arising from the insertion of an extra half plane of atoms into an otherwise perfect periodic lattice. The *local* properties of the lattice are significantly disturbed only near the terminating edge of the half plane, the so-called dislocation line. However, even far away from this line the topology of the lattice is altered. Thus any circuit enclosing this line, consisting of N steps to the right, N upwards, N to the left, and N down, fails to close by 1 step, no matter how large the number N.

Edge dislocations, just like vacancies and interstitials, are naturally present at finite temperatures. They can be made to slip sideways under much smaller stresses than are needed for the simultaneous slippage of the entire upper half-crystal $z>0$ over the lower half-crystal $z<0$. Slippages of many dislocation lines under stress result in a so-called plastic deformation, which, unlike an elastic deformation, remains when the stress is removed. The concept of dislocations completely revolutionized our understanding of the strength of materials. Single-crystal materials of macroscopic dimensions result only under conditions of extremely slow growth. Otherwise even structurally "good" materials are usually polycrystalline, consisting of small microcrystals (or grains) with different orientations. Grain boundaries can be conceptualized as accumulations of dislocation lines. Their properties are of critical importance in metallurgy.

The year 1984 brought a big surprise in the field of crystallography. Mathematical crystallography had been regarded as a closed subject since the work of Schoenfliess in the 19th century. All possible point groups consistent with periodicity had been listed. In particular the icosahedral point group was not allowed. Yet D. Schechtman and co-workers reported a beautiful x-ray pattern with unequivocal icosahedral symmetry for rapidly quenched AlMn compounds. The appropriate theory was independently developed by D. Levine and P. Steinhardt, who coined the words *quasicrystal* and *quasiperiodic*. Even more curious was the fact that R. Penrose (1984) had anticipated these concepts in purely geometric, so-called Penrose tilings (Fig. 5).

Quasiperiodicity has raised new questions about vibrational lattice modes and electronic structure and growth mechanisms, which continue to attract interest.

D. Totally nonperiodic systems

In all of the preceding examples, the perfect periodic lattice has been the appropriate background against which to understand the nature of the defects. There are other condensed-matter systems in which periodicity is totally or largely absent.

Classical liquids, of course, are nonperiodic. Their macroscopic properties have been and are the subject of mature, specialized fields such as hydrodynamics. On a microscopic scale they are described by classical mechanics and thermodynamics, using a potential-energy function $V(R_1,R_2,\ldots,R_N)$, where the R_l are the positions of the nuclei. For some simple liquids, e.g., Ar, V is simply the sum of molecular pair potentials, accurately known from chemistry. But for metallic liquids pair potentials are generally inadequate. In the 1960s quantum theory began to provide good potential functions also for metals. We mention here further that molecular dynamics, i.e., computer simulation of the classical motions of the constituent atoms, has played an important clarifying role for finite-temperature properties and phenomena since the 1960s (B. Alder, A. Rahman, and others).

Quantum liquids consist of light elements, especially the isotopes of H and He. For these liquids the nuclear dynamics at low temperatures must be described by quantum mechanics and exhibit measurable, striking quantum effects: Rotational phase transitions for liquid H_2, Fermi and Bose statistics for He^3 and He^4; diffusion by quantum tunneling even as $T\rightarrow 0$; superfluidity of ultracold He^3.

Critical phenomena refer to the behavior of a thermodynamic system near its critical point. They constitute one of the most important parts of the discipline of statistical mechanics, with major applications to CMP. Since this essay cannot attempt to deal adequately with the general principles of statistical mechanics we must limit ourselves to very cursory remarks.

The 1910 Nobel Prize in physics was awarded to J. D. van der Waals "for his work on the equation of state for liquids," $P=F(\rho,T)$, where P, ρ, and T denote pressure, density, and temperature. The form of the function F, when expressed in system-dependent, dimensionless variables, was universal for all liquid-gas systems. In particular the theory accounted for the universal existence of the so-called critical point (P_c,ρ_c,T_c), where the line in the (T,P) plane, which separates the liquid and gas phases, terminates.

A magisterial generalization of van der Waals' concepts to a general theory of phase transitions was put forward by L. Landau in the 1930s, couched in terms of the key concept of the order parameter S (spatially uniform or slowly varying), such as the difference between liquid and gas densities at given pressure, or the magnetization per unit volume M of a ferromagnet for a given external magnetic field B. Near a critical point—where the order parameter vanishes—Landau expanded the free energy in powers of S and $(T-T_c)$. The theory predicted a simple analytic behavior [e.g., $(T_c-T)\propto(\rho-\rho_c)^{1/2}$] for the coexistence of liquid and gas. However, beginning in the 1940s, and especially in the 1960s, the exponent was found to be not the rational number $\frac{1}{2}$ but, universally, 0.34—very near the critical point. Similarly, other so-called nonclassical universal exponents were found for other classes of thermodynamic systems. These developments eventually led to the major, radically new concept of the *renormalization group* (K. Wilson, 1971) in which spatial correlation functions of the order parameter at or near the critical point $G(r,r')=\langle S(r)S(r')\rangle$ play a central role. Renormalization-group theory, with its new concepts of scaling, universality, stable and unstable fixed points, and basins of attraction, has led to entirely new thinking about the physics of phase transitions, especially near the critical point, and has had far-reaching impacts not only elsewhere in physics, but beyond.

Glasses, known since antiquity and of enormous practical importance, are still not fully understood. A key advance was made by W. H. Zachariason (1932), who proposed that glassy silica, SiO_2, has the same local bonding structure as crystalline quartz (each Si atom being bonded to four oxygen atoms and each oxygen being bonded to two Si atoms), but that the overall topology of the 3D bonding network has random elements and no periodicity. The nature of the so-called glass transition, in which the viscosity changes by many orders of magnitude over a narrow temperature interval, remains subject to study and controversy.

Spin glasses have received a great deal of attention since the 1960s. They are dilute alloys in which spin-carrying ions, e.g., Mn^{++}, are randomly distributed in a nonmagnetic metallic matrix and interact with each other by long-range oscillatory forces. Spin glasses exhibit novel kinds of dynamic magnetic susceptibilities. They have been explored as possible models of neural networks.

The physics of radically nonperiodic systems is infinitely diversified and the foregoing remarks of necessity could touch on only a very small fraction of the important developments during this century. In particular, the interesting electronic transport process in such systems have not been addressed.

IX. BOSE-EINSTEIN CONDENSATION: SUPERFLUIDITY AND SUPERCONDUCTIVITY

The history of superfluidity and superconductivity in this century is, in the view of many, the most remarkable chapter in the 20th-century history of condensed matter physics. These systems not only behave in radically nonclassical, counterintuitive ways, but also for a long time could not be understood in terms of those quantum-mechanical concepts that had been highly successful in explaining the properties of atoms, molecules, and "normal," i.e., nonsuperconducting/fluid matter. This history had its beginning in 1908, when Kammerlingh Onnes had reached a sufficiently low temperature to liquify He^4 at 4.2 K, thereby creating the important new field of "low-temperature physics" including superconductivity/fluidity. Today, 90 years later, we believe that we understand most of the essential characteristics of these systems, except that the mechanism underlying so-called high-temperature (or high-T_c) superconductivity, which was discovered in 1986, still remains a mystery.

The story of superconductivity/fluidity is a wonderful example of the complexity of scientific progress, involving a mixture of serendipity and planned research, of phenomenology and *a priori* microscopic calculations, of well-designed experiments (in line with Onnes' motto "through measurement to knowledge") and of simple models. This in spite of the fact that in the Schrödinger equation we have, but only in principle, the "theory of everything" in CMP. The history of superfluidity is a good example of the need to be alert to the significance of only rough agreement between theory and experiment, as in the case of the experimental transition temperature of superfluid He^4 (2.2 K) and the theoretical condensation temperature (3.3 K) of a model of noninteracting bosons with the mass of He^4 atoms, as well as the need to be alert to the possible significance of tiny unexpected "blips," as in the experimental discovery of the superfluidity of He^3.

The close relationship between superconductivity of metals, first discovered in 1911, and the superfluidity of He^4, fully established in 1938, was first clearly grasped by F. London (partly in collaboration with his brother Heinz), who saw in both phenomena an underlying long-range order in momentum space. His two books *Superfluids I—Macroscopic Theory of Superconductivity* (1950) and *Superfluids II—Macroscopic Theory of Superfluid Helium* (1954) are marvels of what can be achieved in physics by the application of general fundamental physical principles—in this case, thermodynamics, classical electrodynamics, and general quantum concepts—in a thoughtful analysis of often quite "bizarre" experimental results. I have just re-read London's introduction to the first of these volumes (which actually deals in a coherent way with both superconductivity and He^4 superfluidity), written just a few years before the microscopic theories of Bardeen, Cooper, and Schrieffer (1957) of phonon-mediated superconductivity and the Bogoliubov theory of an interacting Bose gas for He^4. I found it exhilarating to see how very much of the most essential physics London had grasped without any microscopic knowledge of the underlying mechanism for superconductivity or of the effects of interactions on Bose-Einstein condensation.

Both phenomena are now understood to reflect the occurrence of a quasi-Bose-Einstein (BE) condensation, of the bosonic atoms in superfluid He^4, and of the bosonic electron pairs in superconducting metals. (He^3 superfluidity, though much more complex, is fundamentally analogous to superconductivity.) I write "quasi" because, although in superconductors/superfluids the original BE condensation for noninteracting bosons is very strongly modified, the most essential aspects of the BE condensation survive.

The interwoven histories of these systems, both experimental and theoretical, extending over almost a century and still evolving, is extraordinarily complex. Disrupted by two World Wars, it captivated the minds and hands of many of the world's best physicists. The 1957 Bardeen-Cooper-Schrieffer (BCS) microscopic theory of superconductivity, taking advantage of a half century of often inspired experimental and theoretical advances, in one blow provided a coherent quantitative explanation of the key properties and phenomena of simple superconductors. It may be regarded as the dramatic high point of this history, which was, however, followed by many additional major and unanticipated advances.

In this essay I must be content to give a very cursory account of the most crucial milestones before about 1950 and then only to list the subsequent high points. I refer the reader to the article in this volume by J. R. Schrieffer and M. Tinkham and that of A. Leggett for more extended accounts.

The beginning of it all was the totally unexpected discovery in 1911 by K. Onnes and G. Hulst of the sudden vanishing of the electrical resistance, or so-called superconductivity, of Hg when the temperature was reduced to below 4.2 K. Soon some other metals were found to be superconducting at a few degrees K, while others remained normal.

Liquid He^4, the cooling liquid in these experiments, was itself found to have the most unusual properties. In 1930 W. Keesom and J. N. van der Emde discovered accidentally that at very low temperatures liquid He passed through extremely small cracks which, at higher temperatures, were quite impervious to liquid or gaseous He. This suggested a vanishing or extremely small viscosity. At 2.2 K, He^4 exhibited a mysterious liquid/liquid phase transition, with no latent heat but a singularity in the specific heat, most clearly established by W. Keesom and A. P. Keesom in 1932. In 1936 they observed an extraordinarily high apparent heat conductivity. In 1938 P. Kapitza, who conducted closely related experiments, coined the term "superfluid."

On the superconducting front W. Meissner showed in 1933 that superconductors, in addition to the remarkable transport property of vanishing resistance, also had the remarkable thermodynamic property of perfect diamagnetism, i.e., the complete expulsion of a weak magnetic field from the interior.

Beginning in 1935 Fritz London, partly with his brother Heinz, using Meissner's discovery of perfect diamagnetism and general physical principles, proposed the new "London equation"

$$\lambda_L^2 \, \mathrm{curl}\, j + H = 0,$$

as the appropriate constitutive equation for the current density j in superconductors. Here λ_L is the so-called London penetration depth. He also brilliantly put forward the notion (later fully confirmed) that the electrons in a superconductor display a long-range order in momentum space.

Returning to superfluid He^4, F. London's insight (1938) that its extraordinary properties are the reflection of a BE condensation proved to be most fruitful. It became the basis of a highly successful two-fluid phenomenological model put forward by L. Tisza (1938) to describe the available experiments: One fluid was the superfluid, the other a normal fluid, each fluid having its own thermodynamic and dynamical variables like density ρ_s,ρ_n, velocity v_s,v_n, and specific entropy $s_s=0,s_n$. (The vanishing of s_s is a consequence of the occupation of a single quantum state by the macroscopic condensate.) The two fluids were viewed as completely interpenetrating each other. Landau, who at first appears not to have accepted the concept of BE condensation, developed his own two-fluid model (1941 and later) in which the normal fluid consisted of the gas of elementary excitations, phonons, and—a new concept—rotons with vorticity. The two-fluid model led to a new collective mode, called second sound, which eventually was experimentally confirmed.

In 1946 Andronikashvili conducted an experiment which beautifully supported the two-fluid model. The moment of inertia of a slowly rotating stack of discs immersed in superfluid He^4 agreed with the picture that only the normal fluid is dragged along.

The first low-temperature experiments (1949) on the then exceedingly rare isotope He^3, which obeys Fermi statistics, showed no sign of superfluidity down to 0.5 K. This strongly supported F. London's contention that the superfluidity of He^4 depended critically on its Bose statistics.

Following is a mere listing of some of the most important developments since about 1950.

Around 1950, A. B. Pippard, in his microwave experiments on the London penetration depth λ_L, was led to the notion of a second length parameter, the coherence length ξ entering a nonlocal generalized London equation in which the current $j(r)$ is proportional to an average of $A(r')$ over a range $|r'-r|\lesssim\xi$. ξ was found to have a strong dependence on the mean free path l. At about the same time L. D. Landau and V. L. Ginzburg put forward a phenomenological theory of superconductivity, which also included the coherence length, in term of a complex wave function $\psi(r)$, playing the role of a space-dependent order parameter. This theory grew out of Landau's general theory of phase transitions, coupled to general principles of electrodynamics. A crucial parameter was the ratio $\kappa\equiv\lambda_L/\xi$; values of $\kappa<(2)^{-1/2}$ resulted in the "usual" kind of superconductivity with a complete Meissner effect. In 1957 A. A. Abrikosov showed that when $\kappa>(2)^{-1/2}$, in so-called type-II superconductors, magnetic-field tubes can penetrate the superconductor, forming a vortex lattice.

In 1950, H. Fröhlich put forward a (nonpairing) theory of superconductivity, depending on electron-phonon interactions, and several experimental groups discovered independently an isotopic mass dependence of $T_c\propto M^{-1/2}$, consistent with Fröhlich's electron/phonon coupling concept. Independent of the specifics of Fröhlich's theory, the empirical isotope effect showed persuasively that lattice vibrations played an essential role in superconductivity.

The discovery of the isotope effect greatly fired up John Bardeen's old interest in superconductors and in the middle 1950s, he embarked on an intensive research program with two young collaborators, L. Cooper and J. R. Schrieffer. In 1956, Cooper showed that a normal electron gas with attractive electron-electron interaction is unstable with respect to the formation of electronic bound pairs (bosonic, so-called Cooper pairs). In 1957 this led to the microscopic BCS theory of phonon-mediated superconductivity, which gave a coherent and wonderfully successful description of a wide variety of properties and phenomena and has become the main paradigm for superconductivity.

The year 1961 saw the prediction and confirmation of what is now called the Josephson effect: A dc voltage V across a superconducting tunnel junction (superconductor/normal metal/superconductor) gives rise to an ac (!) current of frequency 2 eV/$\hbar$. (The charge $2e$ reflects the electron pairing; the frequency is independent of material properties.)

During more than three decades of painstaking materials research by B. Matthias and many others the highest known superconducting transition temperature rose by about 8 K to ~25 K. Suddenly in 1986 A. Mueller and G. Bednorz, studying a new class of materials containing stacks of hole-doped CuO_2 planes, discovered superconductivity at 30–40 K. Further studies of related compounds, now called high-T_c materials, have taken T_c up to about 160 K! While there is no doubt that the carriers are again electron pairs, there is a wide consensus, consistent with generally small isotope effects, that one or more mechanisms beyond electron-phonon coupling are at work; but there is no consensus about their nature. At century's end, this is a major challenge, as is the experimentalists' dream of reaching room-temperature T_c's.

Returning to the He isotopes: the Bose-Einstein condensation of He^4 was confirmed by painstaking analyses of neutron-scattering experiments, with a small macroscopic occupation of ~6% in the zero-momentum state at $T=0$, according to the best recent estimates. However, in 1995 BE condensates of over 90% of bosonic atoms of certain *dilute gases* with very weak interactions were produced by laser cooling and selective evaporation down to below microdegrees Kelvin and exquisitely studied.

The long-searched-for superfluidity of He^3, analogous to the superconductivity of electrons, but much more complex, was found below 3 millidegrees K by D. M. Lee, D. Osheroff, and R. C. Richardson in 1971.

Finally we mention that the concepts of pairing and of superfluidity/conductivity have had important applications in the theory of nuclear structure and of neutron stars.

All in all a heroic chapter, still unfinished, in the history of science.

X. MESOSCOPICS AND NANOSCIENCE

Sometime in the 1960s Richard Feynman is said to have given a talk entitled "There is always room at the bottom," in which he articulated a then new frontier of science, miniaturization of man-made structures down to dimensions of a few atoms, i.e., 1 nanometer = 10 Å. Today, we have in some respects reached this frontier, in others we are close; even single atoms have been successfully observed and manipulated. Far from being simply a matter of setting new records, this journey has led to some of the most exciting physics and most important technological advances of the last several decades. There can be no question that the journey will continue well into the next century.

The conception, successful fabrication (1948), and dramatic applications of the transistor (see the article by Riordan *et al.* in this issue) highlighted the possibility of controlling the dynamics of electrons in very thin surface and interface layers. This no doubt was a major impulse for what today is called mesoscopics and nanoscience.

A few words about the terminology: mesoscopic systems are "in the middle" between microscopic and macroscopic systems, i.e., they contain between about 10^3 and 10^6 atoms. These limits are very rough and depend on the context. Nanoscience, often overlapping with mesoscopics, emphasizes small dimensions, typically 1–100 nanometers, in one, two, or three dimensions.

Strict control of the chemical, structural, and geometric perfection of the samples has been of the essence. Among the numerous techniques used we mention, in particular, molecular-beam epitaxy (MBE) going back to the 1970s, which has allowed the fabrication of layer structures that are atomically flat and compositionally controllable to an accuracy of about 1%. Combined with lithographic techniques it has permitted fabrication of complex semiconductor/metal structures on a nanoscale.

Mesoscopic and nanosystems highlight the qualitative difference between 1D, 2D, and 3D systems. Of course, literally speaking, all physical systems are three dimensional. However, if, for example, electrons are trapped in a layer of sufficiently small thickness, the motion normal to the layer is quantum-mechanically "frozen out" and the system behaves like a 2D gas. Similarly in a "quantum wire" the electrons are confined to a small cross-sectional area in the x-y plane and constitute a 1D electron gas.

We list here some of the interesting results associated with lower dimensions.

(1) Long-range order: In 1935 Peierls had noticed, by direct exact calculations for harmonic lattices, that while in 3D there is long-range (really infinite) order, even at finite T, in 2D it exists only at $T=0$, and in 1D not even then. Later this observation was generalized by N. D. Mermin and H. Wagner (1966) to include other order parameters like magnetization and the superconducting gap function.

(2) Localization: Calling l the nominal elastic mean free path, an arbitrarily weak static disorder localizes electrons over a distance l in 1D (N. F. Mott and W. T. Twose, 1960) and over generally longer distances in 2D (D. Thouless, 1980s). In 3D, for strong disorder all states are localized, for weak disorder only those near the band edges, so-called Anderson localization (P. W. Anderson, 1958).

(3) Luttinger liquid: As Landau first realized, the effect of electron-electron interaction in 3D is rather mild. In particular the Fermi surface in k space remains sharp at $T=0$. However, in 1D, as shown by J. M. Luttinger (1966), interaction effects smooth out the momentum distribution at the Fermi energy. They also have dramatic effects on low-temperature transport processes. (It has been suggested that a 2D electron gas may also be a Luttinger liquid.)

(4) The 2D quantum Hall effect: The ordinary 3D Hall effect, discovered at the end of the last century, occurs when a dc electronic current, flowing in the x direction, is subjected to a magnetic field B_z in the z direction. The resulting Lorentz force in the y direction is balanced by an electric field E_y (due to induced surface charges). It is given by $E_y=\rho_{yx}J_x$, where J_x is the x-current density and ρ_{yx}, the so-called Hall resistivity, is given by $\rho_{yx}=-B_z/nec$, where n is the electron density. This result is robust under the action of moderate periodic potentials, electron-electron interactions, and electron scattering by impurities or phonons. The 3D Hall effect is very useful for determining the electron density n which, when combined with the conductivity, also yields the electron mobility—a critical figure of merit.

The Hall effect has also been observed for 2D electron gases moving freely in a confining surface layer of a semiconductor but being "frozen out" in the perpendicular, z direction. The expected result was $V_H=-B_zI_x/n_2ec$, where n_2 is the electron number/unit area and I_x the electron current in the x direction. By means of a perpendicular gate voltage V_G the chemical potential μ and hence n_2 could be changed. In 1986, Von Klitzing and co-workers, working with a 2D electron gas of very high perfection, made the startling discovery that the Hall voltage V_h, as a function of V_G, had a series of steps for discrete values of n_2 given by $n_2=\nu(eB_z/h)$ where ν was an integer. This is called the integral quantum Hall effect. Furthermore, associated with the Hall steps, the voltage V_x across the sample *parallel* to the current drops to 0, i.e., the diagonal conductivity in the x direction becomes infinite! Shortly afterwards Hall steps were also found for $\nu=1/3$ and other fractional values $\nu=p/q$: the fractional quantum Hall effect. What is the physical meaning of these numbers? It was first shown by L. Landau in 1933 that a magnetic field B_z bunches the eigenvalues of 2D free electrons into discrete, highly degenerate levels given by $\epsilon_n=(n+\frac{1}{2})\hbar\omega$, where $\omega=(eB_z/mc)$. A value of $\nu=2$ indicates that the electrons exactly fill the two lowest Landau levels, and $\nu=1/3$ corresponds to 1/3 filling of the lowest level.

A beautiful explanation of these experiments was provided by R. B. Laughlin (1981,1983) in which gauge invariance, localization of electrons by disorder, and so-called fractional statistics (a generalization of the "integral" Fermi and Bose statistics) play critical roles.

(5) Universal conductance fluctuations: The conductance of a macroscopic disordered metal will, of course, vary very slightly from sample to sample because of differences in the precise configuration of the atoms. Provided that the inelastic scattering length is much greater than the length of the sample, these conductance fluctuations are of the order of $(e^2/\pi\hbar)$ irrespective of the sample resistance, which can be easily measured (R.

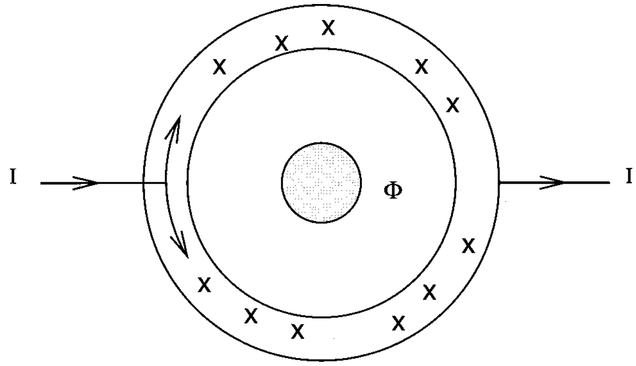

FIG. 6. The Aharonov-Bohm geometry (schematic). Φ is a magnetic flux threading a metallic ring with impurities.

A. Webb and others, 1985). Diffusion of a single impurity, from one equilibrium position to another, produces such a fluctuation. Underlying this and many other mesoscopic phenomena is the fact that, while inelastic electron scattering destroys phase coherence, elastic scattering does not.

(6) The Aharonov-Bohm effect and related effects: The so-called Aharonov-Bohm effect (1959) has found interesting applications in nanoscience. The basic geometry is shown in Fig. 6. Even when the magnetic flux Φ is entirely inside the ring and there is thus no magnetic field acting on the electrons, the conductance dI/dV depends periodically on Φ/Φ_0, where $\Phi_0(\equiv hc/e)$ is the flux quantum. This is due to the fact that the flux introduces *differential phase shifts*, between the parts of an electron wave function propagating in the upper and lower halves of the ring, which affect their interference at the exit point. When Φ/Φ_0 is an integer, the differential phase shift is a multiple of 2π, equivalent to 0. For the observation of this effect, inelastic scattering, which destroys phase coherence, must be negligible. Thus ultralow temperatures and highly miniaturized systems are required ($T \approx 1$ K, dimensions $\approx 1\ \mu$). There is another novel periodicity of $\Phi_0/2$ associated with electron paths and their time-reversed partners. For these path pairs the condition for interference is strictly independent of the particular positions of the impurities [B. Altshuler and others (theory), D. and Yu. Sharvin (experiment), 1981.]

(7) Quantum dots: By a variety of experimental techniques it has become possible (1980s) to fabricate so-called quantum dots, singly or in arrays, of lateral dimensions of ~10–1000 nanometers. The number of mobile electrons in such a dot is typically $O(1-10^4)$. They are sometimes called "artificial atoms" because, due to the small dimensions of the dot, level spacings are correspondingly large. The conductance of such a dot shows enormous fluctuations as a function of gate voltage. These reflect successive resonances of electronic energy levels with the Fermi energy of the attached leads. The positions of these levels are strongly affected by the Coulomb repulsions between electrons, which is the origin of the so-called Coulomb blockade. As miniaturization progresses, it leads to greater spacings Δ between the electronic energy levels. Thus the necessary temperatures, $kT \approx \Delta$, for strong mesoscopic effects is rising. It seems not unreasonable to expect that nanoscience, in the next few decades, will have practical applications at liquid-nitrogen or even room temperature.

XI. SOFT MATTER

The traditional conceptual paradigm of condensed-matter physics, going back to the early part of this century, has been the picture of a dense periodic lattice of ions with valence electrons described by a band structure. This paradigm, with its elaborations and modifications, has been spectacularly successful and continues to underlie much of the ongoing work in CMP, especially for systems with conducting electrons. Let us call it the Bloch paradigm. For insulators a paradigm viewing condensed matter as a collection of interacting atoms or ions (which has been shown to be equivalent to the Bloch paradigm) is often more natural. Both standard metals and insulators have in common periodicity and dense packing. They are very stable and resistant to small perturbations. Let us call them "hard."

Simple liquids of course do not have an underlying periodic lattice. Yet local atomic configurations, densities, cohesive energies, and static compressibilities are very similar above and below the melting temperature. But, unlike crystalline solids, liquids offer no resistance to a static shear stress. In the present context we call them hard/supersoft.

There is, however, another major class of materials, in recent years called "soft matter," whose properties and behavior are covered by different paradigms. Their study has been shared between chemists and physicists. They are characterized by the fact that, unlike gases and ordinary liquids, they do have some shape (or other) stability, but unlike "hard" materials, they respond very strongly to small external disturbances—mechanical, electrical, etc. For example, most edibles, like Jell-O, or fibers, like wood or wool, fall into this category. The bonds between the relevant constituents are usually weak (van der Waals, hydrogen) or easily swiveled or both.

A. Polymers

Polymers are the best known and most extensively studied subclass of soft matter. The simplest ones consist of a chain or "necklace" of identical units, called monomers, strongly and rigidly bonded internally, but flexibly bonded to each other. Monomers on different chains interact by strong short-range repulsions and weak long-range attractions.

Polyethylene,

$$CH_2-CH_2-CH_2-\cdots \text{ or } [CH_2]_N \ (N \approx 10^{3\pm1}), \tag{11.1}$$

is one of the simplest and most thoroughly analyzed. While examples of polymers, e.g., rubber, were known in the 19th century, their nature was clarified only in this century. In 1920 H. Staudinger proved chemically the linear character of polymers. Their flexibility was first

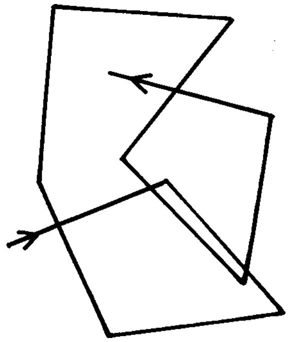

FIG. 7. A polymer or a random walk (in 2D).

demonstrated in 1934 by W. Kuhn for independent polymer chains in dilute solution. He recognized that at finite temperatures a correspondence could be made between the instantaneous shape of a polymer and the trace of a random walk, with the constituent monomers corresponding to the successive steps. This pointed to the very important role of entropy, much higher for a configuration with random orientations of the constituent monomers than for a completely aligned configuration (Fig. 7). From this emerged the picture of a long polymer chain like a tangled ball of yarn after a long period of being "cat-tangled." The elementary theory of random walks led Kuhn to the famous scaling law that the end-to-end distance of a long polymer in solution, and also the effective radius of the tangled 3D ball, obey the scaling law, $R=aN^{1/2}$. Here a is of the order of the length of a monomer, whose precise value depends on the flexibility of the bonds between monomers, etc., but the exponent 1/2 is universal. (A far cry from the paradigms of crystalline solids!)

Kuhn's random-walk analysis was substantially deepened by P. J. Flory (1949), who included the effect of the strong intermonomer repulsions, leading to the analogy with a self-avoiding random walk. A simplified analysis led him to the modified result $R\propto N^{3/5}$, refined by subsequent numerical work to $N^{0.588}$. (On account of these fractional exponents, polymers are examples of so-called fractals, which attracted much attention in the 1980s.)

Because of their enormous and growing scientific and practical importance (as fibers, structural and biological materials, packaging, adhesives, etc.), research and development of polymers has become a rapidly progressing subfield of CMP. The theories of single linear polymer chains have been extended in many directions: polymers consisting of more than one monomer (copolymers); nonlinear, branched chains; mutually entangled chains; polymeric crystals, melts, and glasses; gels; diffusion of polymer chains through the tangle of other chains in a melt (reptation, which means snaking); polymers at interfaces and adhesion.

Polymer physics has enormously benefited from the development of elastic and especially inelastic neutron-scattering techniques led by the work of C. Shull and B. Brockhouse in the middle decades of this century, and these techniques have been applied to polymers since the 1970s by G. Jannink and others. This is because polymers consist largely of light elements (H,C,O, . . .), which interact weakly with x rays but strongly with neutrons. Furthermore, typical frequencies of collective modes of polymers are of the order of $10-10^2$ K, comparable to the energies of cold or thermal neutrons. Finally isotopic replacement of H by D, with very different masses and neutron-scattering properties, has provided a useful research "handle." Thus we owe much of our knowledge of both the structure and the dynamics of polymers to neutron experiments.

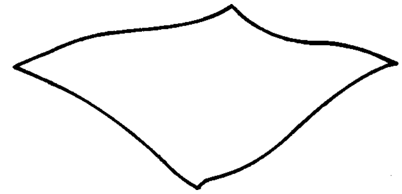

FIG. 8. A free-floating flexible membrane.

B. Membranes

Since the 1970s membranes, another important category of soft matter, have received increasing attention from physicists. Their building blocks are molecules with a hydrophilic (water-loving) head and a hydrophobic (water-fearing) tail.

When dissolved in water these molecules tend to aggregate in flexible two-dimensional membrane structures that "protect" the hydrophobic tails from water. (Fig. 8).

For a given flexed shape the elastic free energy of a single free membrane is given by

$$F=\frac{K}{2}\int dS\left(\frac{1}{R_1}+\frac{1}{R_2}\right)^2, \tag{11.2}$$

where K is the bending stiffness and the R_j are the radii of curvature. Because of thermal fluctuations, memory of the direction of the normal is lost at a characteristic decorrelation distance L_p proportional to $\exp(2\pi K/kT)$, beyond which the membrane will be crumpled.

Membranes are, of course, very important in biology. In fact, the entire broad field of soft matter has become an important bridge between physics and biology.

C. Liquid crystals

The term "liquid crystals" seems to be self-contradictory, but in fact these fascinating materials in some ways strongly resemble conventional, ordered crystals and in other ways conventional liquids. Liquid crystals consist of highly anisotropic weakly coupled molecules. They were first discovered and their essence understood by the French chemist G. Friedel at the end of the 19th century. After a long period of relative neglect their study was actively resumed in the 1960s. There are two main classes, nematics, and smectics. Figure 9 shows a liquid crystal of the "smectic A" type. It

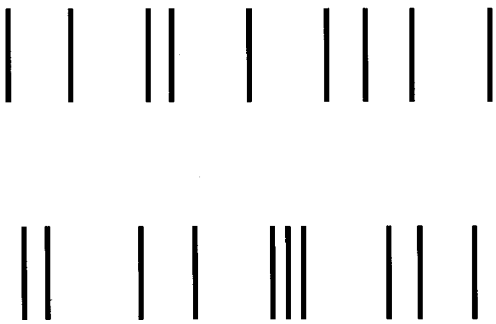

FIG. 9. A "smectic *A*" liquid crystal.

clearly exhibits substantial orientational order, in the orientation both of the molecules (along the z direction) and of the layers (in the x-y plane). Furthermore, the spacing of the layers in the z direction exhibits translational order in the z direction extending over many planes. All these characteristics are reminiscent of conventional crystals. However, the positions and motions of the molecules within any one layer in the x-y plane are highly disordered and resemble those of a two-dimensional liquid.

In some liquid crystals the constituent anisotropic molecules can be realigned by very weak electric fields, which in turn strongly affects optical properties. Such liquid crystals have found extensive use in the displays of electronic watches and other devices (see the article by T. Witten in this volume).

XII. GENERAL COMMENTS

It is perhaps interesting to look at the history of condensed matter physics from the viewpoint of T. S. Kuhn, as expressed in *The Structure of Scientific Revolutions* (Chicago, 1962). He sees scientific history as a succession of (1) periods of "normal" science, governed by serviceable scientific paradigms, followed by (2) transitional, troubled periods in which existing paradigms are found to be seriously wanting, which in turn lead to (3) "scientific revolutions," i.e., the establishment of new paradigms, which may or may not be accompanied by the rejection of the old ones.

Such a linear view seems applicable to the whole field of CMP for some of the broadest revolutions, which directly or indirectly affected a large fraction of the field: x-ray diagnostics yielding crystal structures (1910s); achievement of low temperatures allowing the observation of calmed condensed matter (1900s); quantum mechanics, (1920s); the band-structure paradigm (1920s, 1930s); nuclear and electron spin magnetic-resonance diagnostics (1940s and 50s); neutron elastic and inelastic diagnostics (1950s); many-body electron theories (beginning in the 1930s, with major revolutionary steps in the 1950s and 60s); electronic computer-assisted theory and experiments (1960s-); soft matter (1960s-); and nanoscience (1980s-).

The indicated decades are for rough orientation only. I have included not only conceptual revolutions in the sense of Kuhn but also experimental and technical ones that transformed existing areas of inquiry or opened up important new ones.

Within subfields of CMP there are many additional revolutions. For example: Heisenberg's theory of strong magnetic interactions for magnetism (1920s); scaling and renormalization group for critical phenomena (1960s-); Bose-Einstein condensation and the BCS pairing theory for He^4 and superconductors (1930s and 1950s); masers/lasers for high-intensity, coherent radiation studies (1960s); dislocations for the strength of materials (1930s); high vacua for studies of clean surfaces (1950s).

Others would no doubt choose differently, but few would disagree that the 20th century has been revolutionary for CMP. The fertilizing influence of CMP, conceptual and technical, for other subfields of physics and other sciences has been repeatedly mentioned. Further, condensed matter physics has been a major factor in reshaping technology so that the human experience today is, for most of mankind, very different from what it was 100 years ago.

Looking back over the last century, we see major shifts to the use of more and more sophisticated, man-designed and -fabricated materials, more and more miniaturization, and radically more sensitive diagnostic techniques. Prognostications are, fortunately, beyond the scope of this essay. But it is obvious that the future holds many promises.

ACKNOWLEDGMENTS

I am grateful to Drs. D. Clougherty, A. Gossard, C. Herring, L. Hoddeson, J. Langer, Y. Meir, and P. Pincus for helpful suggestions and comments.

REFERENCES

Brown Laurie, M. Abraham Pais, and Sir Brian Pippard, 1995, Eds., *Twentieth Century Physics* (Institute of Physics, Philadelphia).

Hoddeson, Lillian, Ernst Braun, Jurgen Teichmann, and Spencer Weart, Eds., 1992, *The history of solid state physics* (Oxford University Press, New York).

London, F., 1950, *Superfluids I—Macroscopic Theory of Superconductivity* (Wiley, New York); revised 2nd Ed. (Dover, New York, 1961).

London, F., 1954, *Superfluids II—Macroscopic Theory of Superfluid Helium* (Wiiley, New York); revised 2nd Ed. (Dover, New York, 1964).

Riordan, M., *et al.*, 1999 Rev. Mod. Phys., **71**, this issue.

Warnow-Blewett, Joan, and Jürgen Teichmann, 1992, *Guide to sources for history of solid state physics* (AIP, New York).