ground, did stimulate interest in the inversion problem (of determining the potential $V(r)$ from the scattering data or phase shifts, δ_l) and in a theoretical justification for the Breit–Wigner resonance formula used to fit nuclear scattering data. In this latter undertaking, the analyticity of the S matrix suggested by Kramers was of central importance. Kronig in 1946 recalled that causality (essentially in the form of a first-signal principle) had been related to the analyticity properties of the index of refraction in optics to obtain dispersion relations. He suggested that the causality principle might also play a central role for Heisenberg's S matrix. Some progress was made in deriving the Breit–Wigner formula from the causality requirement and dispersion relations were used as an aid in quantum field theory calculations of scattering amplitudes. In 1955 Wigner showed quite explicitly how the causality requirement implied observable consequences for the behavior of a scattering phase shift near a resonance. Interest in problems such as these provided a background out of which emerged in the mid 1950s the dispersion-theory program for the strong interactions.

3

Dispersion relations

There exist some excellent technical reviews of relativistic dispersion relations (Goldberger, 1960, 1961; Jackson, 1961), as well as Goldberger's (1970) own informal recollections of the period from about 1954–69. In addition, Cini (1980) and Pickering (1989a) have written about some of the sociological influences on that program. We shall comment on these later in this chapter and in the concluding chapter. However, let us begin with a few general observations about the mood of theoretical physics in the United States just after the Second World War, at least in one Physics Department, namely the University of Chicago. This is relevant for what follows, since that Department became a center of activity for the dispersion theory program. Wentzel and Fermi were on the faculty then and Goldberger and Chew were graduate students there. One frequent attendee[1] at the theory seminars at Chicago during those years recalls that at that time (around 1948) the general spirit of many of the younger, exceptionally gifted theorists was that physics was something to *do* (never mind studying the works of the great masters) and that nothing significant had been done (in their areas of interest) prior to this. In fact, many other people had worked on these problems (e.g., fixed-source field theory) before, but there was little sense of history among the younger generation. Wentzel's comments (say, from the audience at a talk) on the previous literature were typically received with impatience. None of this, of course, detracts from the significant and original contributions made by this talented and numerically large generation of younger physicists. While Goldberger and Chew were excited by Heisenberg's ideas, which they learned about from Wentzel's lectures and Møller's articles, they did not really pursue this. The S matrix simply was not a hot topic of discussion then.[2] Furthermore, in the work that led to their 1954 paper (with Thirring) on

dispersion relations, Goldberger and Gell-Mann were interested in *applications* and in obtaining relations among finite, experimentally measurable quantities. As quantum field theory went downhill with its inability to handle the strong interactions, there were two major reactions. One was the use of dispersion relations to avoid unreliable perturbation calculations and the other was axiomatic quantum field theory. We begin with the dispersion theory program, about which Goldberger (1970, p. 689) recalled:

> ... Wheeler and Wigner [at Princeton], through their students Toll, Schützer and Tiomno, had gotten the whole subject [dispersion theory] started.

One of the first problems that had to be faced was that of finding some justification for using dispersion relations for processes in which particles can be created and destroyed. That is, dispersion relations and analyticity properties had previously been established (either classically or quantum mechanically) only for interactions in which the number of particles remains fixed during a reaction (as is the case for the nonrelativistic Schrödinger equation). A foundation was needed to extend those relations to situations in which massive particles could be created and annihilated. After World War II there was an explosion in the amount of accelerator scattering data available on high-energy, strongly interacting particles. Quantum field theory was of little help in providing guidance to the experimentalists or in correlating these data. Dispersion relations were promising, if they could plausibly be believed for the data at hand. There was pressure to organize the data and give some coherence to the expanding field of high-energy physics. It was not at first a case of 'high' theory dictating to experiment, but the enterprise was a pragmatic and coupled one.

The key to a justification of dispersion relations was again provided by a causality condition, but this time in the form of a microcausality. The basic motivation is simple enough to explain in general terms. In ordinary quantum mechanics, observables that cannot be measured simultaneously (i.e., incompatible observables) are represented in the theory by operators that do not commute. A typical example is the pair q (position) and p (momentum) for which the commutator does not vanish since $[q, p] = i$. If two observables are compatible, then their commutator does vanish. In relativistic quantum field theory, the basic mathematical entities out of which all observables are finally built are quantized field operators $\phi(x)$. These operators are labeled with space-time variables x (i.e., x stands collectively for $\mathbf{x}$ and t). If two field

operators commute, then observables built from (or corresponding to) them should be compatible or *independently* observable. If two space-time points are spacelike separated (i.e., too far apart for a light signal to be able to propagate between them in the elapsed time between the occurrence of the two events: $t_2 - t_1 < |\mathbf{x}_2 - \mathbf{x}_1|/c$), then an event at one point should be independent of an event at the other point. The first-signal principle of special relativity would lead us to expect that. The events would be causally independent. So, a natural way to implement causality would be to require that $\phi(x)$ and $\phi(y)$ commute if x and y are spacelike separated: $[\phi(x), \phi(y)] = 0$. This is termed *microcausality* because, as stated, it is to hold for *all* spacelike separated x and y, even on a microscopic scale (of, say, subnuclear dimensions). We shall discuss how this causality constraint was employed to argue for relativistic dispersion relations.

This initial success did lead to attempts to establish a general and coherent theoretical framework for high-energy phenomena. Much of this later work was purely theoretical and not motivated directly by experiment.

3.1   Goldberger, Gell-Mann, Thirring and microcausality

In 1953–54, Goldberger and Gell-Mann became interested in the dispersion-theory problem of the type discussed by Schützer and Tiomno and by van Kampen. Around December 1952 to January 1953, they saw the paper by Rohrlich and Gluckstern (1952) on forward Delbrück scattering and the extension by Bethe and Rohrlich (1952) on the use of dispersion relations applied to nonforward Delbrück scattering. Goldberger and Gell-Mann had not been aware of the Kramers–Kronig relations and do not recall having been motivated by the work of Wheeler and Toll (although Toll was in Chicago with his thesis around May, 1953).[2] However, Richard Eden does recall spending the spring of 1954 in Princeton and doing some work there with Goldberger on dispersion relations. In Eden's opinion, Goldberger was aware of Eden's (1952) paper on the analytic structure of the S matrix in quantum field theory, a paper that had its origins in Eden's own earlier interest in Heisenberg's S matrix and that explicitly refers to Heisenberg's program.[3] This does indicate a continuity (by way of problem background, at the very least) with Heisenberg's older program, a point also made by other physicists who were active at that time.[4] Goldberger and Gell-Mann were interested in applications and

in finding a calculational tool that involved only finite, measurable quantities. In their attempt to understand, on the basis of first principles, these dispersion relations in situations where particle creation and annihilation are possible, a crucial remark was made to Goldberger (1970, p. 687) by Thirring at Princeton. Thirring observed that the causality condition (Cushing, 1986c) in field theory might be taken care of by the requirement that the commutator of the vector potential $A_\mu(x)$ vanish if x and y are spacelike separated,

$$[A_\mu(x), A_\nu(y)] = 0. \tag{3.1}$$

Gell-Mann, Goldberger and Thirring (GGT) (1954) studied the problem in perturbation theory and derived the Kramers–Kronig dispersion relation for forward photon scattering. There they quite explicitly reference the work of Kronig (1926, 1946), Kramers (1927), Schützer and Tiomno (1951), Toll (1952) and van Kampen (1953a, 1953b). Let us outline the argument used in that paper (GGT) because it is the seminal paper for strong-interaction dispersion relations.

3.2   'Proofs' of dispersion relations

For the interested reader, a few technical details are first given. The general logic of the argument is then discussed.

For a *noninteracting* scalar field of mass m, the (Heisenberg) field operators satisfy the equation of motion[5]

$$(\Box^2 + m^2)\phi(x) = 0. \tag{3.2}$$

Equation (3.2) is essentially a generalization of the wave equation, but now for massive 'particles' (unlike the massless photon that goes with the electromagnetic potential $A_\mu(x)$). This is again an example of a 'reinterpretation' of classical equations (i.e., Maxwell's equations). The canonical commutation relations are

$$[\phi(x), \phi(y)] = i\Delta(x - y). \tag{3.3}$$

For our purposes it is sufficient to note that $\Delta(x)$ vanishes whenever x is spacelike; that is, when

$$x^2 = t^2 - r^2 < 0, \tag{3.4}$$

which means a light signal could not travel the distance r in a time t. So, microcausality is satisfied here. If this field $\phi(x)$ is

Fourier decomposed as

$$\phi(x) = \frac{1}{\sqrt{2k_0}} \sum_k [a_k(x_0) e^{-ikx} + a_k^\dagger(x_0) e^{ikx}], \tag{3.5}$$

then the creation operators $a_k^\dagger$ and the destruction operators a_k satisfy

$$[a_k, a_{k'}^\dagger] = \delta_{kk'}. \tag{3.6}$$

One-particle states $|k\rangle$ are created from the vacuum state $|0\rangle$ as

$$|k\rangle = a_k^\dagger |0\rangle, \tag{3.7a}$$

and a_k annihilates the vacuum as

$$a_k |0\rangle = 0. \tag{3.7b}$$

These operators create and destroy quanta ('particles') of mass m and momentum k.

For an interacting system, the Heisenberg field operators $\phi(x)$ now have as their equation of motion

$$(\Box^2 + m^2)\phi(x) = j(x) \tag{3.8}$$

where $j(x)$ is the interaction current. The commutation relations (3.3) and (3.6) are modified and their form cannot be written down until the $\phi(x)$ themselves have been found. However, the *assumption* of (micro-) causality requires that local field observables be independent at spacelike separated points. Since such observables are built up from expectation values of the field operators, we might *expect* that the microcausality condition can be implemented through the requirement that $\phi(x)$ and $\phi(y)$ commute[6] at spacelike distances,

$$[\phi(x), \phi(y)] = 0, \quad (x - y)^2 < 0. \tag{3.9}$$

Finally, we define the function $\eta(x - y)$ by the condition

$$\eta(x - y) = \begin{cases} 1, & x_0 > y_0 \\ 0, & x_0 < y_0 \end{cases}. \tag{3.10}$$

A *compelling* argument for the connection between causality and the commutator condition of Eq. (3.9) does not seem to exist (Goldberger, 1961, p. 199):

> Has anyone ever made any effort to correlate the commutator condition with any measurable properties?

The example that GGT (1954) began with was the scattering of scalar particles (represented by $\phi(x)$) from a fixed force center. Since Eq. (3.9) is an operator identity that holds for all spacelike separations, it follows that the expectation value

$$G_{\text{ret}}(x,y) \equiv i<0|[\phi(x),\phi(y)]|0>\eta(x-y) \qquad (3.11)$$

is a function that vanishes whenever $(x-y)^2<0$ and/or $x_0<y_0$. We shall indicate a bit more fully below how $G_{\text{ret}}(x,y)$ is related to the Fourier transform of the forward scattering amplitude. However, we already can see that this *might* be reasonable since $\phi(x)$ creates one-particle states from the vacuum so that Eq. (3.11) has something (although it is not exactly clear yet just what) to do with the overlap of such states.

For the case of photons, whose field operator is $A_\mu(x)$, interacting with a quantized matter field, the analogue of (3.9) which GGT (1954) began with was

$$[A_\mu(x), A_\nu(y)] = 0, \quad (x-y)^2 < 0 \qquad (3.12)$$

and from this they formed

$$G_{\mu\nu}(x,y) \equiv i\langle 0, f|[A_\mu(x), A_\nu(y)]|i, 0\rangle\eta(x-y). \quad (3.13)$$

Here $|i, 0\rangle$ is the initial state (i) of the matter field with no photons (0) present and $|f, 0\rangle$ the final state (f) of the matter field with no photons (0) present. What GGT now did was to compute the $A_\mu(x)$ in a perturbation expansion to order e^2 in the electric charge (e) and then by brute force manipulations to relate the result to the Fourier transform of the forward scattering amplitude for the scattering of light.

The logic of the argument then is that this forward scattering amplitude is the Fourier transform of a function that vanishes identically outside some region. As the discussion following Eq. (A.65) or that given in Section 2.6 indicates, such an amplitude is an analytic function satisfying a Kramers–Kronig type dispersion relation. Although we shall discuss some of the details below, the basic point to be made here is that the microcausality condition, Eq. (3.12), and a perturbation expansion in e together implied the Kramers–Kronig relation for the forward scattering amplitude of light by a matter field. A short time later, Goldberger (1955a) was able to obtain the Kramers–Kronig relation without having to resort to a perturbation expansion for $A_\mu(x)$.

He discovered an application of what would later be known as the reduction technique that allowed him to express the forward scattering amplitude in terms of an expectation value of the *retarded* (because of the presence of the function $\eta(x-y)$) commutator of the current operators $j_\mu(x)$

$$\eta(x-y)[j_\mu(x), j_\nu(y)] \qquad (3.14)$$

where these $j_\mu(x)$ are the currents determining the $A_\mu(x)$ from the equation of motion

$$\Box^2 A_\mu(x) = j_\mu(x). \qquad (3.15)$$

Notice that from Eqs. (3.12) and (3.15) it follows that

$$[j_\mu(x), j_\nu(y)] = 0, \qquad (x-y)^2 < 0. \qquad (3.16)$$

That is, the microcausality condition has been transferred from the fields $A_\mu(x)$ (cf. Eq. (3.12)) to the currents $j_\mu(x)$. Goldberger based his derivation on formal solutions for the $A_\mu(x)$ obtained previously by Yang and Feldman (1950) and by Källen (1950, 1952). We now outline the reduction technique he used. We do not mean to imply that our outline represents the way Goldberger (1955a) obtained his results. However, here the reader is given some idea of what goes on in that important paper. It is also true that Low (1954) had used a reduction result similar to Goldberger's and later Low (1955) indicated how to obtain this result, but by beginning with the formal field-theory perturbation expansion for the S matrix (cf. Eq. (1.23)) in terms of the interaction Hamiltonian density $H_I(x)$ (Mandl, 1959, pp. 83–4)[7]

$$S = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int d^4x_1 \ldots \int d^4x_n P\{H_I(x_1), \ldots H_I(x_n)\} \qquad (3.17)$$

and verifying directly, for example, the identity $\langle p'k'|S|pk\rangle = \langle p'k' \text{ out}|pk \text{ in}\rangle$, which will be defined below. In a sense, though, one must know the form of the result in order to verify it in this fashion.

The basic idea (Jackson, 1961, pp. 11–21) behind the reduction technique as Goldberger developed it for the Heisenberg field operator $\phi(x)$ is that in the remote past $(x_0 \to -\infty)$ they approach operators $\phi^{\text{in}}(x)$ representing the incoming fields before the interaction begins. The advantage is that we know more about these asymptotically free operators (cf. Eq. (3.3)) than we do about the $\phi(x)$ them-

selves. That is, for arbitrary states $|i\rangle$ and $|j\rangle$ we have

$$\lim_{x_0 \to \pm\infty} \langle i|a_k^{\overset{\text{out}}{\text{in}}}|j\rangle = \langle i|a_k^{\overset{\text{in}}{\text{in}}}|j\rangle. \tag{3.18}$$

For example, the state $a_{k_1}^{\text{in}\dagger}a_{k_2}^{\text{in}\dagger}|0\rangle$ represents a state of incoming particles of momenta k_1 and k_2, that can collide and produce outgoing particles (or waves) (cf. discussion accompanying Eqs. (2.10)–(2.15) and Note 2.10). Similarly, we have $\phi(x)\xrightarrow[x_0 \to +\infty]{}\phi^{\text{out}}(x)$.These $\phi^{\text{in}}(x)$ and $\phi^{\text{out}}(x)$ satisfy the free-field equation of motion, Eq. (3.2), the free-field commutation relations, Eq. (3.3), and have an expansion like Eq. (3.5) in terms of $a_k^{\text{in}}, a_k^{\text{out}}$ (and their adjoints) with Eq. (3.6) holding for the in and out creation and destruction operators. Because the $\Delta(x-y)$ of Eq. (3.3) satisfies

$$(\Box^2 + m^2)\Delta(x-y) = \delta(x-y), \tag{3.19}$$

we can write a formal solution of Eq. (3.8) as (cf. Eqs. (A.8)–(A.10) for the type of reasoning used)

$$\phi(x) = \phi^{\text{in}}(x) + \int d^4x' \Delta_R(x-x')j(x') \tag{3.20}$$

where $\Delta_R(x) = \eta(x)\Delta(x)$[8]. (That is, one can verify directly that Eq. (3.20) satisfies Eq. (3.8).) This solution is such that $\phi(x)\xrightarrow[x_0 \to -\infty]{}\phi^{\text{in}}(x)$. From the expansion (3.5) and use of the total energy-momentum operator P (actually P_μ) as the translation operator according to

$$\phi(x+a) = e^{iPa}\phi(x)e^{-iPa}, \tag{3.21}$$

it follows that, for single-particle states $|k, \text{in}\rangle$, $|k, \text{out}\rangle$,

$$\langle 0|\phi(x)|k, \text{in}\rangle = \langle 0|\phi(x)|k, \text{out}\rangle. \tag{3.22}$$

The basic idea there is that one can use Eq. (3.21) to 'translate' $\phi(x)$ back into the distant past, where it becomes $\phi^{\text{in}}(x)$. Equation (3.18) implies that

$$\int_{-\infty}^{\infty} dx_0 \frac{\partial}{\partial x_0}\langle j|a_k(x)|i\rangle = \langle j|a_k^{\text{out}}|i\rangle - \langle j|a_k^{\text{in}}|i\rangle. \tag{3.23}$$

These formal results are essential for the reduction technique, which we now illustrate for the scattering of two particles,

$$p + k \to p' + k'. \tag{3.24}$$

The Heisenberg S-matrix element is defined as (cf. Note 2.10)

$$S_{\beta\alpha} = \langle \beta \text{ out}|\alpha \text{ in}\rangle = \langle p'k' \text{ out}|pk \text{ in}\rangle. \tag{3.25}$$

Here we have used α to signify collectively the particles in the initial state and β those in the final state. Use of

$$a_{k'}^{\text{out}\dagger}|0\rangle = |k' \text{ out}\rangle, \tag{3.26}$$

Eq. (3.23), Eq. (3.5), the equations of motion and integration by parts finally yields, after considerable manipulation (cf. Jackson, 1961, p. 16),

$$S_{\beta\alpha} = \langle p'|a_k^{\text{out}}|pk \text{ in}\rangle$$
$$= \delta_{\beta\alpha} + \int d^4x \, e^{ik'x}\langle p'|j(x)|pk \text{ in}\rangle. \tag{3.27}$$

A similar reduction of $|k \text{ in}\rangle$, plus use of Eq. (3.21), leads to an expression for the T-matrix (cf. Eq. (2.17)) of the form[9]

$$T_{\beta\alpha} = i\int d^4x \, e^{ik'x}\langle p|\eta(x)[j(x), j(0)]|p'\rangle. \tag{3.28}$$

with $k' = p + k - p'$.

All the mathematical details aside, the importance of this reduction technique result is that it gives an *exact* expression for the scattering amplitude $T_{\beta\alpha}$ in terms of the retarded current commutator in the integrand. The vanishing of that commutator in a certain region of integration leads to analyticity properties of the $T_{\beta\alpha}$ (just as occurred for the Fourier transform discussed in Chapter 2).

If we now restrict ourselves to the forward direction ($k'=k, p'=p$), go to the laboratory frame in which $\mathbf{p}=0$ and consider massless particles for which $k=(\omega, \mathbf{k})$ and $\omega=|\mathbf{k}|$, Eq. (3.28) becomes

$$T(\omega) = i\int dt \, d\mathbf{r} \, e^{i\omega(t-\hat{\mathbf{k}}\cdot\mathbf{r})}\langle p|\eta(x)[j(x), j(0)]|p\rangle. \tag{3.29}$$

Since $[j(x), j(0)]$ vanishes unless $x^2 = t^2 - r^2 > 0$ and since $\eta(x)$ vanishes unless $t > 0$, we see that the t integral in Eq. (3.29) runs only over values of t such that $t > r$ so that $(t - \hat{\mathbf{k}}\cdot\mathbf{r}) > 0$ in the integrand. By the type of argument we have discussed more than once by now, this implies that $T(\omega)$ is analytic in the upper half ω plane and yields a dispersion relation for $T(\omega)$. From Eq. (3.29) we also see immediately that

$$T^*(\omega) = T(-\omega), \tag{3.30}$$

which is the analogue of Eq. (A.68). This is a special case of an important general property known as crossing (Gell-Mann and Goldberger, 1954) and we shall discuss it at length in the next chapter. Although we have found this property from the general expression

(3.29), GGT (1954) obtained this result from their perturbation calculation of $T(\omega)$ and Gell-Mann and Goldberger (1954) argued for crossing on the basis a general property satisfied by the Feynman diagrams representing the perturbation expansion.

Of course, we have glossed over several important points in this sketch. Goldberger (1955a) was dealing not with a massless scalar field $\phi(x)$ but with the four-vector field $A_\mu(x)$ of quantum electrodynamics. Nevertheless, in spite of considerable technical complications, similar manipulations go through. Also, we have argued for the analyticity of $T(\omega)$ for $\mathrm{Im}\,\omega > 0$, but we have overlooked the behavior of $T(\omega)$ as $\omega \to \pm \infty$ along the real axis. Furthermore, there is, in fact, the possibility that $T(\omega) \xrightarrow[\omega \to \infty]{} \omega^n$, where n is a finite power. We shall return to these problems in a later section. In spite of all these caveats, we must not lose sight of the deep and beautiful connection established between causality and analyticity. The essential ingredient in the proof is the microcausality statement, Eqs. (3.9) or (3.12). Of course, this is strictly an assumption appended to the field-theory framework since there are no known rigorous solutions to any realistic (four-dimensional) quantum field theory satisfying such a causality condition.

At this point (say, after Goldberger's (1955a) perturbation-theory independent proof of the forward dispersion relation for photon scattering), only the scattering of massless particles had been handled with any degree of rigor. Here, as throughout the dispersion-theory and S-matrix theory programs, applications of assumed dispersion relations often far outstripped what could be proven (even in a loose sense of that term). This gap between what could be *proven* and what was pragmatically *assumed* to get on with things would become a recurrent characteristic of the dispersion-theory and S-matrix theory programs. In their original paper, GGT (1954) implied that their dispersion-relation proof was valid for massive particles as well. They were criticized for this by van Kampen (GGT, 1954, p. 1613; Goldberger, 1970, p. 687). The basic difficulty when $\omega = \sqrt{k^2 + m^2}$, or $k = \sqrt{\omega^2 - m^2}$, can be seen from the factor

$$e^{ikx} = e^{i(\omega t - \mathbf{k} \cdot \mathbf{r})} = e^{i[\omega t - \sqrt{\omega^2 - m^2}\,\hat{\mathbf{k}} \cdot \mathbf{r}]} \tag{3.31}$$

which does *not* have oscillatory behavior for $0 < \omega < m$. Goldberger (1955b) indicated how this difficulty might be overcome, but a true proof had not yet been constructed. Goldberger (1970, p. 688) himself states about his (1955a) proof: 'Of course, my derivation was not really correct; since the result was correct, I was sure someone would prove it eventually.'

From the analytic properties they established using causality, Gell-Mann, Goldberger and Thirring (1954) did derive the dispersion relation for the electromagnetic forward scattering amplitude $f(\omega)$ (cf. Eqs. (A.66)–(A.70) for details)

$$\mathrm{Re}\,f(\omega) - \mathrm{Re}\,f(0) = \frac{2\omega^2}{\pi}\,\mathbb{P}\int_0^\infty \frac{\mathrm{d}\omega'\,\mathrm{Im}\,f(\omega')}{\omega'(\omega'^2 - \omega^2)} \tag{3.32a}$$

$$= \frac{\omega^2}{2\pi^2}\,\mathbb{P}\int_0^\infty \frac{\mathrm{d}\omega'\,\sigma_t(\omega')}{(\omega'^2 - \omega^2)}. \tag{3.32b}$$

These are just Eqs. (A.70c) and (A.70d) which are familiar from classical electrodynamics. Here a previously expected result had been justified within the framework of a wider theory. The limit of Eq. (3.32b) as $\omega \to \infty$ is

$$\mathrm{Re}\,f(\infty) - \mathrm{Re}\,f(0) = -\frac{1}{2\pi^2}\int_0^\infty \sigma_t(\omega')\,\mathrm{d}\omega'. \tag{3.33}$$

For scattering of light from a bound electron (Rayleigh scattering), we have $f(0) = 0$, while for scattering from a free electron (Thomson scattering) $f(0) = -e^2$ [10] (Kroll and Ruderman, 1954; Low, 1954; Gell-Mann and Goldberger, 1954). If for the quatum electrodynamics (QED) case, the high-energy limit for scattering from a bound electron *were* $f(\infty) = -e^2$ (as it is classically; cf. Eq. (A.83)), then GGT (1954) would also have obtained the old Thomas-Kuhn sum rule, Eq. (A.85). Such sum rules are important direct relations between experimentally measurable quantities (here e^2 and σ_t). Similar sum rules would prove important in later dispersion-theory and S-matrix theory work. However, the QED value for $f(\infty)$ was unclear then.[11] Nevertheless, they did indicate how to apply the result (3.33) to the scattering of photons by a nucleus and by protons even though no convincing experimental checks of their predictions were yet possible.

3.3   Phenomenological use of dispersion relations

Karplus and Ruderman (1955) assumed that the GGT (1954) dispersion relations were valid for the scattering of massive particles and showed how these could be used to analyze experimental data. They pointed out that Eqs. (3.32) would, in general, have to be supplemented by pole terms (representing bound states) and that, if m is the mass of the

lightest particle scattered, then for incident energy $E = \sqrt{k^2 + m^2}$,

$$\operatorname{Im} f = 0, \qquad 0 < E < m^2. \tag{3.34}$$

This requirement can be related to the optical theorem, Eq. (2.9) (since physical scattering cannot take place below $E = m^2$) or to unitarity, Eq. (2.20) (cf. comment preceding Eq. (2.51)). That is, if the energy available is not even enough to include the mass-energy (mc^2) of the scattered particle, then the reaction cannot go and the cross section must vanish. Karplus and Ruderman (1955) restricted themselves to the case of no bound states (and, hence, no pole terms in the scattering amplitude) and observed that Eq. (3.32a) could (in principle) be compared with experiment through the relation

$$\frac{d\sigma}{d\Omega}(\theta = 0) = |f(k)|^2 = [\operatorname{Re} f(k)]^2 + \frac{k^2}{16\pi^2}[\sigma_t(k)]^2. \tag{3.35}$$

That is, from the observed differential scattering cross section ($d\sigma/d\Omega$) and total scattering cross section, one can find $\operatorname{Re} f(k)$ from (3.35) and then compare it with $\operatorname{Re} f(k)$ from Eq. (3.32b). They also suggested using dispersion relations as a guide to doing a phase shift analysis, as follows. Since Eq. (3.35) shows that $d\sigma/d\Omega$ only determines $|f(k)|$, there is an ambiguity in the sign of $f(k)$ and this is reflected in there being more than one solution for the phase shift $\delta(k)$ of Eq. (A.23a). If we consider only s-waves so that the partial-wave expansion for $f(k)$ becomes approximately

$$f(k) \approx \frac{e^{i\delta} \sin \delta}{k} \tag{3.36}$$

and for simplicity of illustration drop $f(0)$ in Eq. (3.32b), we have

$$\operatorname{Re} f(k) \approx \frac{\sin[2\delta(k)]}{2k} = \frac{k^2}{2\pi^2} \int_0^\infty \frac{dk' \, \sigma_t(k')}{(k'^2 - k^2)}. \tag{3.37}$$

Karplus and Ruderman (1955) assumed isospin conservation (a form of charge independence for the nuclear forces) and applied this to 'data'[12] on $\pi^0 p$ scattering and found numerical agreement with the results obtained previously by determining the phase of $f(k)$ through interference (for $\pi^\pm p$ data) with the Coulomb scattering amplitude at small angles. That is, these charged particles (pions and protons) interact through the electromagnetic force (because they carry electric charge) and through the strong (or nuclear) force. Therefore, the full scattering amplitude can be expressed as the sum of a Coulomb (or purely electromagnetic) amplitude and another term representing the strong

interactions. This sum of two amplitudes can produce interference effects (just as for optical and acoustic phenomena). This interference dip in the observed cross section can be seen near the forward (i.e., $\theta = 0$) direction. Karplus and Ruderman's basic point was that causality (through dispersion relations) could generally resolve the standard phase shift ambiguity, just as Coulomb inference effects had done in the past.

However, Karplus and Ruderman (1955, p. 772) had used the crossing relation in the form

$$f(-E) = f^*(E) \tag{3.38}$$

which, as Goldberger pointed out, was correct only for uncharged particles. Goldberger, Miyazawa and Oehme (GMO) (1955) provided the correct details for the general pion–nucleon case. They showed that a suitable linear combination of the physical $\pi^+ p$ and $\pi^- p$ elastic amplitudes satisfied the dispersion relations previously written down by Goldberger (1955b) for massive particles. These proper linear combinations do satisfy simple crossing relations like Eq. (3.38), even though the individual amplitudes themselves do not. Goldberger, Miyazawa and Oehme (1955) gave the correct generalizations of Eqs. (3.32) and (3.33) for $\pi^\pm p$ scattering. If the subscript $+$ is used to denote quantities for $\pi^+ p$ scattering and $-$ those for the $\pi^- p$ reactions, then the GMO dispersion relations for the real part of the scattering amplitudes are

$$D_+(k) - \frac{1}{2}\left(1 + \frac{\omega}{\mu}\right) D_+(0) - \frac{1}{2}\left(1 - \frac{\omega}{\mu}\right) D_-(0)$$
$$= \frac{2f^2}{\mu^2} \frac{k^2}{\omega^2 - \mu^2/2M} + \frac{k^2}{4\pi^2} \int_\mu^\infty \frac{d\omega'}{k'}\left[\frac{\sigma_+(\omega')}{\omega' - \omega} + \frac{\sigma_-(\omega')}{\omega' + \omega}\right], \tag{3.39a}$$

$$D_-(k) - \frac{1}{2}\left(1 + \frac{\omega}{\mu}\right) D_-(0) - \frac{1}{2}\left(1 - \frac{\omega}{\mu}\right) D_+(0)$$
$$= -\frac{2f^2}{\mu^2} \frac{k^2}{\omega^2 + \mu^2/2M} + \frac{k^2}{4\pi^2} \int_\mu^\infty \frac{d\omega'}{k'}\left[\frac{\sigma_-(\omega')}{\omega' - \omega} + \frac{\sigma_+(\omega')}{\omega' + \omega}\right]. \tag{3.39b}$$

Here $\omega = \sqrt{k^2 + \mu^2}$, μ is the pion mass, M the proton mass, f the pion–nucleon coupling constant (about which more will be said later), and principal-value integrals are understood. The great value of these justly famous dispersion relations was that they were directly applicable to realistic situations represented by the experiments that could actually be performed at the time. Anderson, Davidon and Kruse (1955) applied these relations to the phase shift analysis of πp data between 0 and

1.9 Bev, and Davidon and Goldberger (1956) did further work on $\pi^{\pm}N$ data. The parameter f had to be adjusted to obtain good agreement between Eqs. (3.39) and the experimental data. They used a value of f found by other means. The determination of f, by several independent means, will prove to be an important point in subsequent sections. Goldberger, Nambu and Oehme (1957) eventually published the details of the dispersion relations for the much more complicated case of nucleon–nucleon scattering.

3.4   Pragmatic attitude of many practitioners

It should be evident that much of this work on dispersion relations was characterized by an attitude that placed a premium on getting practical things done, at time in an almost cavalier fashion. For example, in his (1955b) 'proof' of dispersion relations for massive boson fields, Goldberger essentially assumed without any proof or argument part of the answer he wanted (cf. Goldberger, 1955b, note added in proof, p. 985). 'We made up rules as we went along, proceeded on the basis of hope and conjecture, which drove the purists mad'.[13] As Goldberger (1961, p. 196) himself observed: 'It is perhaps of historical interest to relate that almost all of the important applications (Chew, Goldberger, Low and Nambu, 1957a, 1957b) of the nonforward dispersion relations were carried out before even the forward scattering relations were proved rigorously.' Pickering (1989a) has characterized as 'pragmatic' this attitude in the applications of dispersion relations. Cini (1980) has made a similar observation. On the other hand, those interested in a more rigorous approach were not particularly impressed by this style. As Jost states[14]:

> If I remember correctly (and if you allow me a judgment) very few of the so called 'purists' paid much attention or took much interest in Goldberger's conjectures .... [T]hey were absorbed in their own work.

This contrast between the pragmatism of the Americans, who were much concerned with experimental results, and the rigor of the Europeans, who were distant from experiment, is consistent with the view presented by Cini (1980). He also sees this concern with phenomenology as the key to understanding the renewed interest in dispersion relations, an old idea in physics. The locus of high-energy experimental physics was largely in America during this period.

Still, Goldberger usually knew the right answer even if he couldn't always prove it. The nature of the problem was now rather well defined and this provided the basis for a viable research program: to 'prove' dispersion relations for a wider class of cases (e.g., for massive particles; in the nonforward direction; eventually, in two or more variables; etc.) on the basis of general field-theory axioms and to apply these relations to many experimentally accessible situations. For example, when John C. Polkinghorne (who would later become a leader of the 'Cambridge school' of S-matrix theorists) went to the California Institute of Technology as a postdoctoral fellow in 1955 to work with Gell-Mann, he learned from Gell-Mann that dispersion relations might provide an on-mass-shell, alternative formulation of quantum field theory.[15] Polkinghorne (1956) gave a heuristic proof for single-variable dispersion relations for mesons scattered from a nucleon, allowing for the possibility of meson creation or annihilation. He specifically referred to the Heisenberg program (Polkinghorne, 1956, p. 217) and suggested the possibility that it might provide an alternative formulation to quantum field theory. Once again we see a continuity going back to the early motivations for the S-matrix program.

If rigorous proofs were not available, one assumed the dispersion relations needed for a particular application (subject, of course, to their being no recognizable contradictions). It was only in 1957 that Symanzik (1957) provided a rigorous proof for the forward case. Dispersion relations in the energy variable, but now for fixed, nonzero values of the scattering angle, were heuristically derived, for example, by Salam (1956), by Salam and Gilbert (1956) and by Capps and Takeda (1956). More rigorous proofs of nonforward relations were given by Bogoliubov, Medvedev and Polivanov (1958) and then by Bremermann, Oehme and Taylor (1958) using the theory of several complex variables. No proof good for all angles was ever constructed.

3.5   More general proofs

There were several attempts to derive general results that would remain valid independently of the details of any particular quantum field theory or model. Thus, Lehmann, Symanzik and Zimmermann (LSZ) (1955) began by assuming the existence of Heisenberg field operators and then deriving as much as possible about the S-matrix elements. Nambu (1955) derived dispersion representations of quantum field Green's functions using the causality requirement. However, the attempt

probably most in the spirit of Heisenberg's old *S*-matrix program (which LSZ (1955) explicitly cited) was that of Lehmann, Symanzik and Zimmermann (1957) who started with a pure *S*-matrix formalism. They began by assuming the existence of a scattering operator that connected the causal in and out operators and then sought to prove from this the existence of causal field operators that interpolated between the asymptotic in and out operators. Theirs was one attempt to provide a general framework based on first principles. They were careful (1957) to stress that it remained unknown whether a causal *S* matrix actually exists:

> We begin with a pure *S*-matrix formalism [Heisenberg, 1943a, 1943b], (p. 319)

> As is customary we call a field operator $A(x)$ causal – without discussing here the physical interpretation – if $[A(x), A(y)] = 0$ for $(x - y)^2 > 0$. (p. 323)

> It is an entirely open question whether any scattering matrix exists which is causal in the sense of this definition or whether this is too stringent a demand. (p. 324)

This program was in a sense the converse of their earlier work (LSZ, 1955). That is, this new approach reversed the emphasis from fields being primary to the *S* matrix being the entity of central interest. Many impressive mathematical results were obtained within and as an outgrowth of the LSZ program. Lehmann (1958) used the LSZ (1957) formalism (which relies heavily on reduction techniques similar to those discussed earlier in this chapter) and causality to establish the analyticity of the two-particle scattering amplitude in the momentum-transfer variable (i.e., roughly $\cos \theta$ where θ is the center of momentum scattering angle) at *fixed* energy. This included the region of momentum transfer for which single dispersion relations in the energy variable had been established. Lehmann's proof employs Dyson's (1958) integral representation of expectation values of causal commutators of the field operators. This was a remarkable and important paper, since Lehmann was the first to show explicitly how causality implied analyticity in the momentum transfer variable, even when that variable was outside the physical region (i.e.,outside the values of that variable accessible to experiment). Much subsequent work, which was largely, but not exclusively, a European undertaking, attempted to make these discussions more mathematically rigorous and more general. Res Jost was a central figure in much of this work, both by his own technical

contributions and especially by the influence he had upon theorists associated with the ETH (Eidgenössische Technische Hochschule) in Zürich where Jost was a Professor. Steinmann (1960), Ruelle (1961) and Hepp (1965) related the LSZ formalism to Wightman's (1956) mathematically rigorous axiomatic formulation of quantum field theory[16], although the LSZ program does not follow from Wightman's axioms alone. Bros, Epstein and Glaser (1965) established crossing (which we discuss in the next section) for the four-particle (i.e., two in and two out) amplitude within the LSZ framework and they (1972) proved some analyticity properties for the n-particle amplitude. Still, it was not possible to underpin completely the type of dispersion relations and analyticity properties of general scattering amplitudes that were being used in applications.[17] We do not pursue this axiomatic QFT program further here, mainly because it is an involved and highly technical story requiring a separate study of its own and because it is not really central to our purposes in this study. Nevertheless, the motivation for the axiomatic study of scattering amplitudes was nicely summarized by Lehmann (1959, pp. 153–4):

> It is hoped that by obtaining explicit properties of the scattering matrix we may:
> (a) Correlate and thereby partially understand experimental results.
> (b) Test the correctness of the axioms by confronting them with empirical facts.
> Another motivation is of course the lack of anything better. That is, despite many attempts, nobody has succeeded in formulating *specific* interactions ... in such a way that reliable quantitative calculations can be made, in the case of strongly interacting relativistic particles.

This last paragraph gives some indication of the value placed on getting on with a program in the sense of handling the enormous amount of experimental data then available.

For reference later, let us give here a list (Table 3.1) of the prominent workers in the dispersion-theory program (Goldberger, 1961, p. 181), just to indicate both the international character of the undertaking and the fact that the enterprise was dominated by Americans.

3.6   Other applications of dispersion relations

In this section, we indicate that use of dispersion relations was not restricted to scattering amplitudes alone (Chew, Karplus, Gasiorowicz

Table 3.1. *Prominent workers in the dispersion-theory program*

Americans	Europeans	Russians
R. Blankenbecler	R. Eden	N. Bogoliubov
G. Chew	S. Fubini	E. Fainberg
R. Cutkosky	R. Jost	A. Lagunov
S. Drell	H. Lehmann	L. Landau
F. Dyson	R. Omnès	I. Pomeranchuk
W. Frazer	J. Polkinghorne	D. Shirkov
M. Gell-Mann	K. Symanzik	
M. Goldberger	J. C. Taylor	
F. Low	J. G. Taylor	
S. MacDowell	W. Thirring	
S. Mandelstam	W. Zimermann	
Y. Nambu		
R. Oehme		
J. Toll		
S. Treiman		
J. Wheeler		
E. Wigner		
D. Wong		
F. Zachariasen		

and Zachariasen, 1958). This also allows us to introduce some concepts that will be useful in subsequent chapters. The basic problem here is to discuss the electromagnetic structure of strongly interacting particles (i.e., of those particles, such as the proton or the pion, that undergo both strong and electromagnetic interactions). Let us begin with the relatively simple example of electron(e)–pion(π) scattering, say

$$e^+ e^- \to \pi^+ \pi^-, \tag{3.40}$$

illustrated in Figure 3.1(*a*). Figure 3.1(*b*) shows the approximation to order e^2 for the scattering amplitude of the process in (3.40). This amplitude is found by computing the S-matrix element between the initial and final states of Eq. (3.40). It turns out to be essentially the product of the matrix element for the process $e^+ e^- \to \gamma$, the photon propagator (i.e., the function representing the propagation of the photon) and the matrix element for $\gamma \to \pi^+ \pi^-$. This last vertex is given as

$$\langle q' | j_\mu(0) | q \rangle = e(q' + q)_\mu F_\pi(t) \tag{3.41}$$

where $F_\pi(t)$ is the electromagnetic form factor of the pion and t, the four momentum transfer, is defined as

$$t = (q' - q)^2. \tag{3.42}$$

(Equation (3.41) is another example of the great usefulness of the currents j_μ that also entered so crucially into the derivation of dispersion relations as in Eq. (3.16).) The function $F_\pi(t)$ incorporates all of the (unknown and in practice uncalculable) strong-interaction effects of the $\gamma \to \pi\pi$ vertex. In the nonrelativistic limit (which will not be the case generally of interest to us here), $F_\pi(t)$ is simply the Fourier transform of the charge distribution $\rho(\mathbf{r})$ of the electric charge of the pion,

$$F_\pi(t) \stackrel{\text{N.R}}{=} \int \rho(\mathbf{r}) e^{i(\mathbf{q} - \mathbf{q}') \cdot \mathbf{r}} d^3 r, \tag{3.43}$$

where, by definition,

$$F_\pi(0) = 1. \tag{3.44}$$

Reduction techniques similar to those outlined earlier in this chapter allow one to obtain a dispersion relation for $F_\pi(t)$ as

$$F_\pi(t) = 1 + \frac{t}{\pi} \int_{4m^2}^{\infty} \frac{dt' \operatorname{Im} F_\pi(t')}{t'(t' - t)}. \tag{3.45}$$

Either formal field theory manipulations or the unitarity condition of Eq. (2.20) show that the imaginary (or absorptive) part of $F_\pi(t)$ is related to all the intermediate processes that can connect the γ state to the $\pi\pi$ state. Since the two-pion state is the lowest-mass intermediate state allowed by energy-momentum conservation and the discrete conservation laws (such as charge conservation), the lowest order
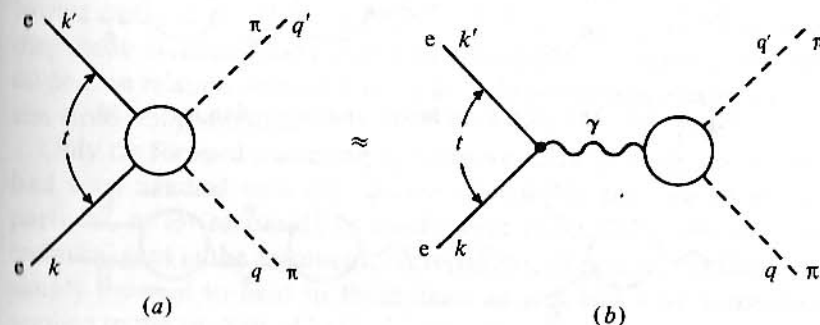

Figure 3.1   The electromagnetic form factor of the pion.

approximation is (see Figure 3.2)

$$\text{Im } F_\pi \approx F_\pi^* \, e^{i\delta} \sin \delta \qquad (3.46)$$

where δ is the p-wave $\pi\pi$ phase shift for $\pi\pi \to \pi\pi$ scattering. Here $e^{i\delta} \sin \delta$ is just the $\pi\pi$ elastic scattering amplitude in the $l = 1$ angular momentum state. Figure 3.2 illustrates the important principle that any reaction proceeds *via* (or 'communicates' with) all possible intermediate states to which it can be connected (subject to the constraints of the operative conservation laws). An important point is that the integral equation resulting when Eq. (3.46) is substituted into Eq. (3.45) can be solved exactly (Omnès, 1958) so that if the p-wave $\pi\pi$ phase shift were known, $F_\pi(t)$ would be (approximately) known.

The problem of more immediate physical interest is the electromagnetic structure of the nucleon. This was the case first discussed by Federbush, Goldberger and Treiman (1958) in terms of dispersion relations. The electron–proton process

$$\text{ep} \to \text{ep} \qquad (3.47)$$

can, much as in Figure 3.1, be approximated by the one-photon intermediate state in terms of the nucleon form factor[18] for

$$\gamma \to N\bar{N}. \qquad (3.48)$$

Again, a dispersion relation can be written for $F_N(t)$ (Drell and Zachariasen, 1961). Now, however, the diagram for $\text{Im } F_N(t)$ becomes Figure 3.3. That is, the equation for $F_N(t)$ involves $F_\pi(t)$ (which, as we have just seen, is known in terms of the $\pi\pi$ phase shifts) and the amplitude for $\pi\pi \to N\bar{N}$. The electromagnetic structure of the pion
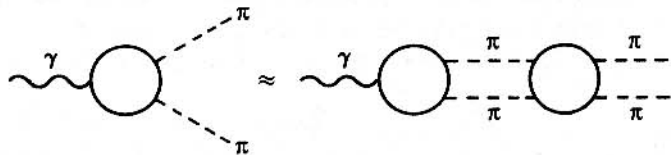

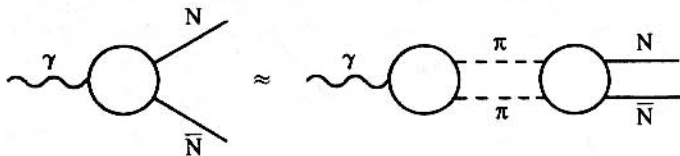Figure 3.2   The pion form factor and pion–pion scattering.

Figure 3.3   The nucleon form factor and pion–nucleon scattering.

plays an essential role in the electromagnic structure of the nucleon. In a later chapter we discuss how the amplitudes for $\pi\pi \to \pi\pi$ and $\pi\pi \to N\bar{N}$ were eventually found and what the experimental consequences of $F_\pi(t)$ and $F_N(t)$ are. A central point to appreciate, though, is that the 'coupling' of various processes illustrated in Figures 3.1–3.3 is essentially a result of unitarity (recall Eqs. (2.19) and (2.20)) in the form

$$\text{Im } T_{ij} = \sum_k T_{ik} T_{jk}^*. \qquad (3.49)$$

That is, speaking loosely but graphically, unitarity ties one process to many others. This insight is essential for the bootstrap concept to be discussed later.

3.7   Summary

In the early 1950s, Goldberger and Gell-Mann at the University of Chicago saw Rohrlich and Gluckstern's paper on the use of dispersion relations to calculate scattering amplitudes in quantum field theory. They were interested in finding a calculational scheme that involved only finite, measurable quantities. Dispersion relations appeared to be a likely tool, but Goldberger and Gell-Mann sought some understanding of them in terms of first principles, such as causality. Thirring pointed out that in quantum field theory the causality requirement can be implemented by demanding that the effects of the field operators $\phi(x)$, $\phi(y)$ acting at spacelike separated points x and y be independent (cf. Eq. (3.1) or Eq. (3.9)). In their first paper on dispersion relations, Gell-Mann, Goldberger and Thirring in 1954 used this commutator causality condition and perturbation theory to derive the Kramers–Kronig dispersion relation for forward photon scattering to lowest order in e^2. The key point was that causality had led to this dispersion relation. Soon afterwards, Goldberger obtained the same dispersion relation without having to use a perturbation expansion for the electromagnetic field.

Only the forward scattering of massless particles (here, the photon) had been handled with any degree of rigor. In the case of massive particles, or for scattering in nonforward directions, there remained technical gaps in the argument. Nevertheless, dispersion relations were simply *assumed* to hold in these cases as well and were successfully applied to the analysis of both elelctromagnetic and strong interaction scattering data. This was indicative of a pragmatic attitude of many

theorists at the time. Success was used as a justification for this approach. Rigorous proofs lagged far behind these applications and many of these assumed results were never proved on the basis of an underlying theory. A research program had been born.

8

The duality program

In this chapter we examine the origins of a program that emerged from the S-matrix formalism but that soon took on a quite independent existence. Its greatest interest for present high-energy physics is that it led to superstring theories. This evolution is also of interest methodologically, as an example of an abandoned research program giving rise to a possibly 'correct' theory that might otherwise never have been formulated.

The concept of duality was an outgrowth of the Regge program. The simple Regge-pole form for the scattering amplitude arises from the exchange (in the crossed channel) of a Regge pole and is valid at *high* energy. When that form is extended (or extrapolated) to low energies, it gives the same result as the average value of an amplitude generated by the exchange of resonances in the direct channel (see the 'pictoral' representation of Eq. (8.1) below). More specifically, when sum rules (an example of which we have seen in Eq. (6.31)) that are integrals over differences of total cross sections are evaluated once numerically using experimental data and once 'theoretically' using the simple Regge form (at *all* energies), the results agree. This type of equivalence between direct-channel resonance and crossed-channel Regge exchange became known as *duality* and was a self-consistency form of the bootstrap. Veneziano offered a specific and simple analytical model that provided a concrete instantiation of the duality conjecture. This example was important both for subsequent theoretical progress and for immediate phenomenological applications. His remarkable model showed how several of the S-matrix principles, along with Regge asymptotic behavior and duality, could be incorporated analytically and consistently to form the basis of a workable bootstrap program. Subsequently, Veneziano suggested a perturbation type of expansion of dual scatter-

ing amplitudes, not in terms of a coupling constant (or smallness parameter) as had been familiar from QFT, but rather in terms of the topological complexity of various dual amplitudes (or diagrams). This gave some understanding of the origin of the Pomeron, which had remained a mystery in previous *S*-matrix theory. Duality supported a successful program of phenomenology that was able to correlate a large body of data.

Topology was incorporated into the *S*-matrix program and gave it new life and promise. A specific topological property of these dual amplitudes – 'order' – was conjectured to hold in nature. This order is a formal mathematical property of the topology of certain dual amplitudes and is not present (in exact, unbroken form) in nature. Nevertheless, the assumption of order did allow explanation of several important and known, but previously mysterious, regularities in the systematics of hadron phenomenology. This is not only an obvious example of hypothetical-deductive reasoning to justify an assumption (here, order). Even better, it is a case of formalism leading physical intuition, since it is unclear what physical property of an elementary particle could correspond to this notion of order. It seems to be a purely formal (even if quite useful) device. Furthermore, in terms of the hierarchy of this topological complexity, some mathematical objects are more fundamental (or 'elementary') than others (much as were Gell-Mann's 'mathematical' quarks in his own early formulation of the quark model; cf. Section 6.7 above). That is, compatibility was now seen as a possibility between the *S*-matrix principles and a quark picture (of elementary entities). Much of the new topological SMT was even formulated in terms of the language of its competitor (quark-gauge-field theory). Again, though, as with earlier versions of the *S*-matrix program, the theory was so complex mathematically that, after some initial and promising success, further progress became difficult.

Nevertheless, the duality program itself did lead to string theory when Nambu recognized that the formal structure of dual amplitudes could also be obtained from a field theory for strings. Again, we have an important example of mathematical formalism leading physical intuition. Not only did SMT and its descendant, duality, provide the origin for string theories, but these string models (in quantized form) require a certain number of space-time dimensions for mathematical consistency (including, possibly, a removal of the old renormalization infinities of QFT). This would be the ultimate bootstrap. As a bonus, which was neither sought nor even contemplated in the early days of duality and string theory, current superstring theory may include gravity (as well as strong and electroweak forces) and be truly the theory of everything (TOE). Such a unique, universal theory was the bootstrapper's dream. There are indications that SMT and present quantum field theories are either equivalent or limits of a larger theory. It is also possible that SMT may bootstrap superstrings. Whether or not any of these conjectured scenarios is realized, it does remain that superstring theories owed their formulation to duality and to the *S*-matrix program.

8.1   *S*-matrix origin of duality

In a previous chapter (6) we saw that Igi (1962) used Regge asymptotic behavior for the πN reaction amplitudes to derive the sum rule of Eq. (6.31). In 1966 de Alfaro, Fubini, Furlan and Rossetti discovered a similar idea of superconvergence relations. They first derived these relations from current commutators in quantum field theory using a technique employed by Gell-Mann for weak interactions. De Alfaro *et al.* (1966), as well as Dolen, Horn and Schmid (1967), also related these superconvergence relations to a limit of ordinary dispersion relations. Schwarz (1967) obtained superconvergence relations from helicity (or spin) amplitudes and an assumed Regge asymptotic behavior.[1] Igi and Matsuda (1967) applied these to the πN charge exchange reaction. Dolen, Horn and Schmid (1967, 1968) developed the use of finite-energy sum rules (FESR) to determine Regge parameters by connecting high-energy Regge behavior with integrals over low-energy data. The dual representation or duality of an amplitude in terms of resonances at low-energy and Regge behavior at high energy provided a bootstrap or consistency condition imposed by analyticity. Here Regge and resonance behaviors are both contained as different approximations to the same exact amplitude and they cannot both be added together independently (Schmid, 1970). Symbolically we can express this as (Mandelstam, 1974b)

$$\tag{8.1}$$

A simple example of how one obtains a FESR is provided by considering (Dolen, Horn and Schmid, 1967, 1968) an amplitude F that is odd in the variable[2]

$$\nu = \frac{(s-u)}{2M} \tag{8.2}$$

and that satisfies an unsubtracted dispersion relation

$$F(v) = \frac{2v}{\pi} \int_0^\infty \frac{\mathrm{Im}\, F(v')\,\mathrm{d}v'}{v'^2 - v^2}. \tag{8.3}$$

FESR'S are based on the assumptions that $F(v)$ is analytic and that the asymptotic behavior is given by the Regge form

$$F(v) \underset{v \to \infty}{\approx} \frac{\beta(t)[1 - \mathrm{e}^{-\mathrm{i}\pi\alpha}]}{\sin(\pi\alpha)\Gamma(\alpha+1)} v^\alpha \equiv R(v). \tag{8.4}$$

If the leading Regge term has $\alpha < -1$, then the $v \to \infty$ limit of Eq. (8.3) requires the superconvergence relation

$$\int_0^\infty \mathrm{Im}\, F(v)\,\mathrm{d}v = 0. \tag{8.5}$$

If *only* the leading Regge trajectory has $\alpha > -1$, then we can subtract R from F and obtain the superconvergence relation

$$\int_0^\infty \mathrm{Im}(F - R)\,\mathrm{d}v = 0. \tag{8.6}$$

Assuming that the Regge approximation of Eq. (8.4) is valid for $v > N$, one obtains from Eq. (8.5) the FESR

$$\int_0^N \mathrm{Im}[F(v)]\,\mathrm{d}v = \sum_\alpha \frac{\beta N^{\alpha+1}}{\Gamma(\alpha+2)}. \tag{8.7}$$

This relation should be valid at large, but *finite*, energy.

An early and important use of FESR'S (Dolen, Horn and Schmid, 1967, 1968; Freund, 1968) was a new type of bootstrap (Chew, 1986b), or self-consistency, condition to determine the parameters of the ρ trajectory from pion–nucleon charge-exchange data (recall Figure 6.9). However, for our purposes here, this πN case led to an important observation (Schmid, 1968a, 1968b, 1970). Only the ρ trajectory is above $\alpha = -1$ (at $t = 0$) so that Eq. (8.6) becomes

$$\int_0^N \mathrm{Im}[F(v) - F_\rho(v)]\,\mathrm{d}v \approx 0 \tag{8.8}$$

where $F_\rho(v)$ is the one-pole ρ Regge amplitude (Eq. (8.4)). Here $\mathrm{Im}\, F(v)$ can be related (through the optical theorem) to the experimentally measurable $[\sigma_{\mathrm{tot}}(\pi^- p) - \sigma_{\mathrm{tot}}(\pi^+ p)] \equiv \Delta\sigma(v)$. Numerical evaluation of the integrals verified Eq. (8.8). This meant that, *on the average*, the high-energy Regge amplitude, extrapolated to low energies, gave the same average behavior as the resonances which dominate the

low-energy region. This is illustrated in Figure 8.1. This equivalence of the average of the resonances and of the Regge term is known as FESR duality. This duality is the meaning of the symbolic Eq. (8.1) above. If no approximations were made, this equality or equivalence should be exact. As applications of the FESR bootstrap, other resonances (denoted as the ρ, f, and g ones) in the direct $\pi\pi$ channel were used to generate the ρ Regge pole in the t channel and to calculate $\alpha_\rho(t)$ (Schmid, 1968a) and the Regge πN charge-exchange amplitude was projected into partial waves to reproduce the low-energy πN resonances observed experimentally. Chew and Pignotti (1968) used duality to justify a Regge multiperipheral model at all energies, as in the production reaction πN $\to$ πNρ.

8.2   The Veneziano model

Veneziano's (1968) model was of central importance for the subsequent development of the duality program (and finally, indirectly, for string theory). Since the mathematics involved is fairly simple, we outline the general features of his model below.

Veneziano (1968) constructed a particularly simple model that exhibited this duality property explicitly. Let us take as an example the (somewhat fictitious) case of the elastic
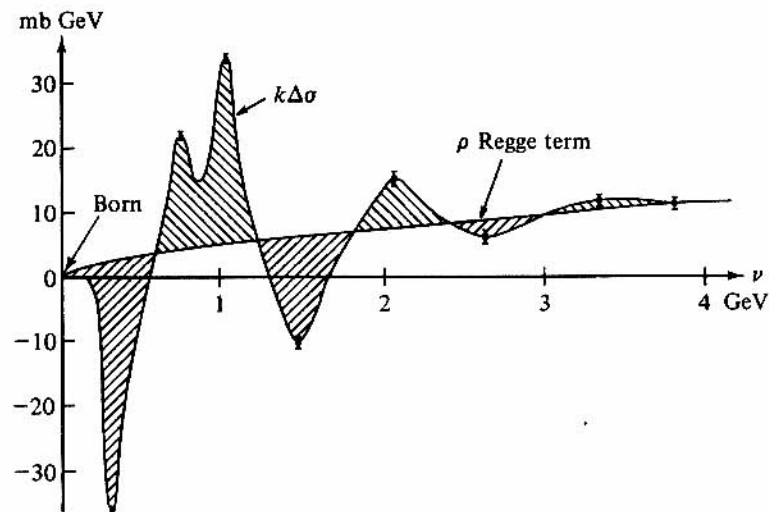

Figure 8.1   Resonance–Regge duality for pion–nucleon data (after Schmid (1970), p. 114).

scattering amplitude for identical scalar (spin zero) particles, $s+s \rightarrow s+s$. The complete amplitude $F(s,t,u)$

$$F(s,t,u) = A(s,t) + A(s,u) + A(t,u) \qquad (8.9)$$

is given in the Veneziano model as the 'narrow' (actually, zero-width) resonance approximation as

$$A(s,t) = \frac{\Gamma[-\alpha(s)]\Gamma[-\alpha(t)]}{\Gamma[-\alpha(s)-\alpha(t)]} = \int_0^1 dx\, x^{-\alpha(s)-1}$$
$$\times (1-x)^{-\alpha(t)-1} \quad (8.10)$$

where $\Gamma(z)$ is the usual gamma (or factorial) function and the Regge trajectory $\alpha(s)$ is assumed to be linear

$$\alpha(s) = \alpha_0 + \alpha' s. \qquad (8.11)$$

(Recall that $s+t+u=4m^2$.) Here α' is positive to represent a rising Regge trajectory (corresponding to physical particles of positive spin) as shown in Figure 8.2. Incidentally, that these *exactly* linear trajectories correspond to the *zero*-width resonance approximation can be seen from the dispersion relation for $\alpha(s)$ (Mandelstam, 1974b)

$$\alpha(t) = a + bt + \frac{(t-t_0)^2}{\pi} \int \frac{dt'\, \mathrm{Im}\,\alpha(t')}{(t'-t)(t'-t_0)^2}, \qquad (8.12)$$

once we recall that $\mathrm{Im}\,\alpha$ is proportional to the width of a resonance (cf. Section 5.4 and Eq. (6.34)).

Since $\Gamma(z)$ has simple poles when $z=-n$ (i.e., at the negative integers), the amplitude $A(s,t)$ of Eq. (8.10) has poles at those values of s for which $\alpha(s)=n$, $n=0,1,2,\ldots$.
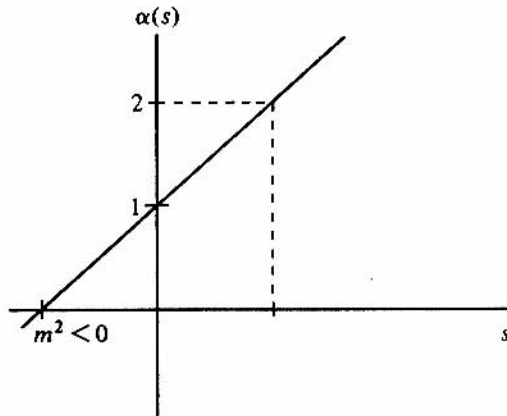

Figure 8.2   A linear Regge trajectory in a Veneziano model.

This is the expected resonance behavior in the direct (here, s) channel. Furthermore, the Stirling approximation for $\Gamma(z)$ implies

$$\frac{\Gamma(z+a)\Gamma(z+b)}{\Gamma(z+c)\Gamma(z+d)} \xrightarrow[|z|\to\infty]{} z^{a+b-c-d} \qquad (8.13)$$

so that[3]

$$A(s,t) \xrightarrow[s\to\infty]{} \Gamma[-\alpha(t)][-\alpha(s)]^{\alpha(t)} \approx s^{\alpha(t)}. \qquad (8.14)$$

which is the desired asymptotic behavior in the s channel. The full amplitude of Eq. (8.9) has the asymptotic form (Schwarz, 1973)

$$F(s,t,u) \xrightarrow[s\to\infty]{} \Gamma[-\alpha(t)][1+e^{-i\pi\alpha(t)}][\alpha(s)]^{\alpha(t)}. \quad (8.15)$$

Notice that, when $\alpha(t)=$ odd integer, the second term in square brackets vanishes. This is the signature factor and it implies that a given trajectory contributes a pole to the physical amplitude only at alternate integers.[4] In the Veneziano model we can avoid a recurrence at $\alpha=1$ as follows. From Eqs. (8.9) and (8.10), as well as the standard recursion relation

$$\Gamma(z+1) = \Gamma(z), \qquad (8.16)$$

we find that the residue of $F(s,t,u)$ at $\alpha(s) \equiv 1$ is proportional to $[-2-\alpha(t)-\alpha(u)]$. We can arrange to have this residue vanish if

$$\alpha(t) + \alpha(u) = -2. \qquad (8.17)$$

Combined with $\alpha(s)=1$, this constraint implies (for the linear trajectories of Eq. (8.11))

$$\alpha(s) + \alpha(t) + \alpha(u) = -1. \qquad (8.18)$$

The 'bootstrap' condition that the Regge trajectory pass through zero (for a spin *zero* particle) at $s=m^2$

$$\alpha(m^2) \equiv \alpha_0 + \alpha' m^2 = 0 \qquad (8.19)$$

then requires that

$$\alpha' m^2 = -1 \qquad (8.20)$$

or that

$$\alpha_0 = 1. \qquad (8.21)$$

This corresponds to $m^2 < 0$ which is a tachyon (i.e., a particle that travels faster than the speed of light).[5] This

unphysical property proved to be one of the persistent problems for dual models and we shall return to it later. In fact, Veneziano (1968) in his original paper considered the case $\pi\pi \to \pi\rho$ and specifically avoided the case $s + s \to s + s$ because of the unphysical result of Eq. (8.21).

Lovelace (1968) made a very nice application of Veneziano's model to the data on decays of the type $p\bar{n} \equiv X \to 3\pi$. In order to fit experiment, Lovelace introduced, *ad hoc*, an imaginary part, $\text{Im}\,\alpha = \text{const}\sqrt{s - 4m^2}$, to the linear Regge trajectory $\alpha(s)$ of the Veneziano model. The phenomenological fit to the data was quite good, although the model was no longer wholly consistent since Veneziano's results obtained only for *linear* trajectories. Veneziano's model had been for a 'four point' scattering amplitude (i.e., two particles, or 'lines' in and two out, for a total of four lines). Bardakci and Ruegg (1968) generalized Veneziano's integral representation of Eq. (8.10) to the five-point amplitude (e.g., two lines in and three out) and Virasoro (1969b) also gave a similar result. Chan Hong-Mo (1969), Goebel and Sakita (1969), and Koba and Nielsen (1969) found a generalization of Veneziano's model for the N-point function. Of course, since such a model violates unitarity, it can at best be considered only a zeroth-order approximation to the true scattering amplitude. That is, from the list of desirable properties for scattering amplitudes (e.g., Lorentz invariance, unitarity, analyticity, crossing, Regge asymptotic behavior), the duality program began by giving up unitarity and then attempting to implement that requirement 'perturbatively'. More complicated generalizations (e.g., Virasoro, 1969a), including integral representations of Veneziano's model, were proposed and extended to include arbitrary values of the isotopic spin for the particles (Paton and Chan, 1969). An entire operator formalism (Mandelstam, 1974a) was generated to guarantee the imposition of subsidiary conditions, such as factorization of residues and the elimination of ghost states (resonances with negative decay probabilities). We shall return later to this and its successor – the quantized string theories – after we have followed a different offshoot of the duality program.

8.3   A topological expansion

Harari (1969) and Rosner (1969) independently suggested a graphical representation to exhibit the duality property of scattering amplitudes. They assumed, in terms of (the group theoretical) SU(3) quark

terminology, that *all* 'legal' particles (external to a diagram) and poles (internal to a diagram) are either baryons (which are fermions) consisting of three quarks ($B \to qqq$) or mesons (which are bosons) consisting of a quark–antiquark pair ($M \to q\bar{q}$).[6] However these duality diagrams are 'cut' in a given channel, only 'legal' states should appear. Thus, Figure 8.3(a) is an allowed duality diagram, whereas Figure 8.3(b) is not. The contribution from the Pomeron is not given by these diagrams, but the rest of the scattering amplitude in any channel is given by the sum of single-particle states. These duality diagrams can equally well be viewed as a sum of resonances in either channel, as indicated in Figure 8.4. All of the allowed duality diagrams are connected and planar, in the sense that the 'quark' lines do not cross one another (i.e., they all lie uncrossed in a plane). When these planar dual diagrams (e.g.,
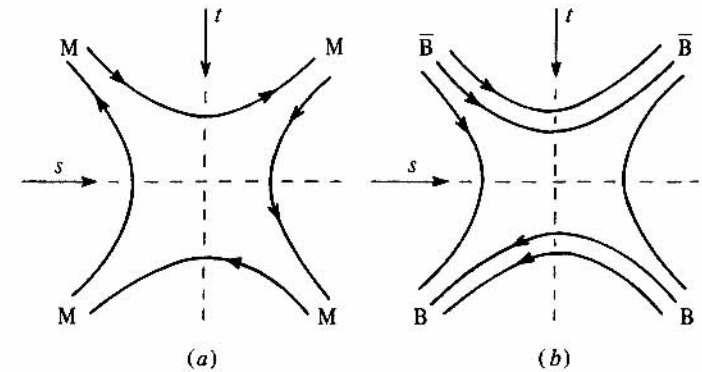

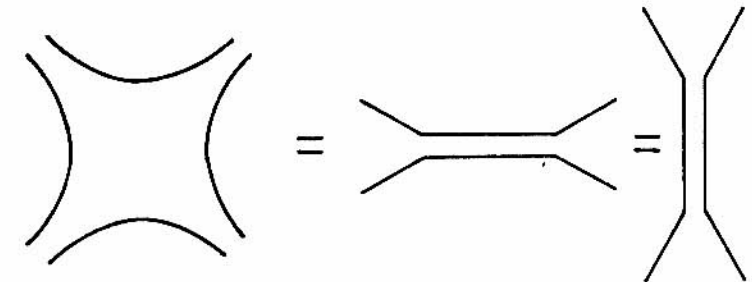Figure 8.3   Harari–Rosner diagrams.

Figure 8.4   Duality diagrams.

Figure 8.3(a)) are combined *via* unitarity, one obtains higher-order diagrams, both planar (Figure 8.5(a)) and nonplanar (Figure 8.5(b)). Duality diagrams could be seen as a first (Born) approximation.

These duality diagrams, which began largely as a phenomenological set of rules, bear some resemblance to the Okubo (1963)–Zweig (1964)–Iizuka (1966) (or OZI) rule, also termed (Okubo, 1977) the quark line-rule (or QLR). Although Okubo (1963) originally derived a phenomenologically successful rule from a 'nonet hypothesis' for SU(3), Zweig (1964) and Iizuka (1966) stated the rule in terms of connected *planar* quark diagrams. Only processes that can be represented by these connected planar diagrams, such as figure 8.6(a), are allowed. A process like Figure 8.6(b) is not allowed, according to the OZI rule. This rule, which had no independent theoretical basis, proved quite successful in that reactions violating it were strongly suppressed. Its origin remained somewhat of a mystery, although its interpretation was straightforward enough in terms of the quark model (Freund, Waltz and Rosner, 1969).

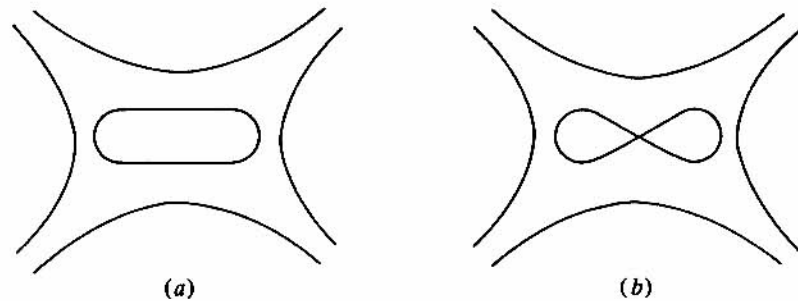The concept of planar diagrams as the lowest-order terms in an

(a)   (b)

Figure 8.5   Planar and nonplanar diagrams.
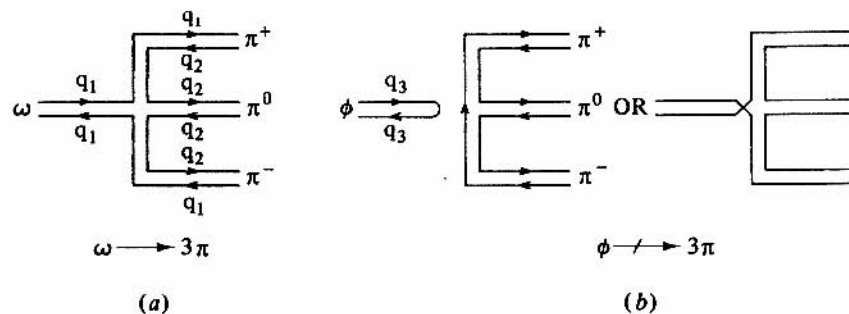

(a)   (b)

Figure 8.6   The Okubo–Zweig–Iizuka (OZI) rule.

expansion of the physical scattering amplitude became important. Lee (1973) gave an indication of how dual resonance amplitudes in a multiperipheral model could bootstrap the Pomeron. Veneziano (1974a) suggested the expansion and unitarization of dual amplitudes in a hierarchy of topological complexity, beginning with the planar at the lowest level.

> If one sums ... 'per topology' ... , a new perturbation expansion emerges (Veneziano, 1974a, p. 365)
>
> which we name the 'topological expansion'. (Veneziano, 1974b, p. 220)

That is, rather than expanding in terms of a coupling constant (or smallness parameter), Veneziano now suggested expanding in terms of topological complexity. The hope was that the topologically simpler diagrams would dominate the expansion.

> [W]e have proposed a new 'order of summation' for implementing unitarity in planar dual models. (Veneziano, 1974b, p. 220)
>
> [We] hope to associate the ... equation
>
> Planar × Planar = Planar + Non-Planar
>
> with another equation
>
> Reggeon × Reggeon = Reggeon + Pomeron.
>
> (Veneziano, 1974a, p. 367)

The unitarity sum generates nonplanar output terms from planar input, just as the Pomeron may be generated through unitarity by Reggeon exchange. This led to calculations, by Rosenzweig and Veneziano (1974) and by Schaap and Veneziano (1975), of Regge couplings (i.e., a vertex joining two particles to a Regge pole as in Figure 6.7) and Regge trajectories with this planar duality and unitarization scheme. An early effort to calculate diffractive scattering (associated with the Pomeron) was made within the framework of a dual Regge picture (Chan, Paton and Tsou, 1975) and the results produced reasonable fits to a wide range of experimental data. A Reggeon bootstrap was attempted (Chan, Paton, Tsou and Ng, 1975) with a dual model.[7] The results were encouraging but not conclusive. In 1974 Chew, while visiting CERN where Veneziano was, became interested in Veneziano's topological expansion. Rosenzweig, who had been working with Veneziano, began a collaboration with Chew and then went to Berkeley. The origin of the Pomeron was clarified as a correction to the planar level in the

topological expansion (Rosenzweig and Chew, 1975; Chew and Rosenzweig, 1975), a bootstrap origin of the Okubo–Zweig–Iizuka (OZI) (Okubo, 1977) rule was found (Chew and Rosenzweig, 1976) and the relation between G parity (another conserved quantum number) and Regge-trajectory exchange degeneracy (for certain pairs of trajectories as, for example, the ρ–A_2 and the f–ω) was studied (Chew and Rosenzweig, 1977). Balázs (1976, 1977a, 1977b, 1977c, 1977d, 1977e) used a dual multiperipheral model to make approximate numerical calculations of the Pomeron parameters, triple Regge couplings (a vertex at which three Regge trajectories meet) and a Regge trajectory giving a linear meson spectrum in agreement with experiment. Capra (1977) and Stapp (1977) made some tentative suggestions for incorporating baryons into this topological expansion scheme.

8.4   The topological *S*-matrix program

In the process of writing a review paper, Chew and Rosenzweig (1978) put forward as an hypothesis the concept of *order* in the topological S matrix. This gave some understanding of several previously observed regularities of high-energy scattering amplitudes. For example, we have already seen that it had long been known that as a matter of principle there must exist Regge cuts (which are moving singularities other than poles), in addition to Regle poles, in a relativistic scattering amplitude. Nevertheless, the effect of these cuts is remarkably absent from much experimental data. That is, good empirical fits can often be obtained using only Regge poles without any Regge cuts. This fact remained a curiosity for which there was no compelling explanation. Furthermore, the empirical OZI rule, which is simple to state and quite accurate, had no theoretical explanation either in QFT or in SMT. So, both the absence of Regge cuts and the OZI rule appeared to be accidents.

It was found that by adding one more postulate to SMT, these facts, plus many predictions that had been characteristic of the quark model in QFT as applied to meson–meson scattering, can be explained (Chew and Rosenzweig, 1978). This additional property is *order* and can best be explained here by means of a specific example. Consider the scattering process (or reaction) in which particles a and b scatter into (or become) particles c and d as

$$a + b \rightarrow c + d \tag{8.22}$$

The analytic scattering amplitude (or S-matrix element) describing the

process (8.22) equally well describes the reaction

$$a + b \rightarrow d + c. \tag{8.23}$$

The order of the particles in the final state (or in the initial state or in both) is immaterial. This is found to be the case in nature. However, we could begin by postulating an ordered S-matrix for which the amplitudes corresponding to the ordered diagrams of the theory have the property

$$A_0(a + b \rightarrow c + d) \neq A_0(a + b \rightarrow d + c). \tag{8.24}$$

Such an S-matrix can be consistently relativistically invariant, unitary, crossing symmetric and analytic. In other words, order is a perfectly acceptable S-matrix principle. By demanding self-consistency (a key to any bootstrap program) when unitarity is imposed, one can show that only a certain (unique) subset of graphs (from among all conceivable ordered connected graphs) is allowed (Capra, 1979). The only difficulty is that the actual physical S-matrix is not ordered. One can attempt to remedy this by defining a physical, unordered scattering amplitude as[8]

$$A_p(a + b \rightarrow c + d) = A_0(a + b \rightarrow c + d) + A_0(a + b \rightarrow d + c). \tag{8.25}$$

Unfortunately, this so-called planar S-matrix now lacks just one of the properties required of an S-matrix – unitarity. If we leave this drawback aside for a moment, we find that the planar S-matrix has several nice properties, among which are the OZI selection rule, exchange degeneracy, and lack of any need for Regge cuts. (That is, there is no inconsistency if Regge cuts are absent).

One can then set up an iterative scheme in terms of the topological structure of these amplitudes and of the unitarity relation to generate corrected amplitudes that do satisfy unitarity. It appears plausible that there may be a dynamical mechanism which suppresses the higher-topology terms to produce convergence of the topological expansion (Chew and Rosenzweig, 1978, p. 293), but such a convergence has not been proved. The hope is that the planar diagrams provide the dominant contribution to the physical amplitudes so that many remnants of order do survive in the physical world. Incidentally, one can in retrospect 'see' in Veneziano's paper on planar dual models (1974a, Fig. 1, Eqs. (3.1) and (3.2)) what is essentially the ring diagram of Chew's later program, even though order does not appear to have been in the air at that earlier date. This dual topological unitarization (DTU) program required the existence of baryonium states (or exotic mesons) which, in terms of the quark model language, would represent

qqq̄q̄ states (Chew and Rosenzweig, 1978, p. 320). Lack of firm experimental evidence in the available data (Nicolescu, Richard and Vinh Mau, 1979; Pickering, 1984, pp. 312–315; Nicolescu and Poénaru, 1985) for these high-mass narrow resonances (with zero baryon number) was *the* central problem for DTU and order. After the 'November Revolution' of the discovery of the J/ψ particle (Aubert *et al.*, 1974; Augustin *et al.*, 1974), there was a loss of interest among the theoretical physics community in the problem of the topological program in favor of the gauge-field-theory program. If the idea of order had come a few years earlier, perhaps there would have been a wider interest in the concept.[9]

It was in his (1979a) paper on the bootstrap theory of quarks that Chew began to give this entire program a highly sophisticated formulation in terms of a topology that could handle baryons as well and he has claimed (1979b) that the weak and electromagnetic interactions may also be able to be accommodated. In fact, electromagnetism now plays a central role in the basic measurement process for Chew (1985). However, as ever in the S-matrix approach, calculations remain difficult. (Some sense of these complexities can be gotten from the papers of Balázs, Gauron and Nicolescu (1984); Balázs and Nicolescu (1980, 1983); Jones and Uschershon (1980a, 1980b); Nicolescu (1981); Gauron, Nicolescu and Ouvry (1981).) In a paper setting forth the axioms of the new topological S-matrix theory (TSMT), Chew and Poénaru (1981) acknowledged the effectiveness of the rival QCD approach and reiterated their belief that the two approaches might indeed be compatible.

> Bootstrap theory developed slowly because of its essential nonlinearity and lost favor when the capabilities of the seemingly opposite quark approach, eventually formalized within quantum chromodynamics (QCD), became recognized. It was, however, never established that conflict exists between quark and bootstrap ideas. This paper describes a bootstrap theory which explains quarks and their properties on the basis of S-matrix consistency. (p. 59)

Some workers in TSMT continue to see signs of this equivalence:[10]

> [W]hen we repeatedly said [in TSMT] that 2-dimensional surfaces are the essence of hadron dynamics, people were more or less laughing. Now they take very seriously into account this idea (for all particles) simply because it comes from superstrings.

An example of what Nicolescu may have in mind here is the importance that Schwarz (1982a, p. 12) assigns to topology in the current superstring theories.

> The two-dimensional world sheets may be classified by their topological complexity – the number of holes and the handles – which is a convenient way of organizing the perturbation expansion.

Nicolescu also appeals to an overarching principle as a motivation for pursuing the TSMT program[11] (see also, Nicolescu, 1985, p. 63):

> For me it is very important that the bootstrap principle succeeds to unite 'world-view' and 'scientific theory' in a single approach.

An exposition of this topological S-matrix theory involves complex concepts quite different from any we have discussed thus far and would add relatively little of new conceptual interest to our case study, aside from the possibility that a bootstrap may be possible even with (topological) hierarchies. (A review can be found in Capra (1984) and in Nicolescu and Poénaru (1985).) This model, like the duality and string models (to which we turn next), began as a mathematical 'toy' model for the mesons but, once it had been extended to include spin and baryons (Chew and Poénaru, 1980, 1981), became for its proponents a candidate for a theory of all particle interactions. Even though first formulated for the hadrons, the TSMT now includes also the electroweak interactions. This theory predicts exactly three 'quarks' in a baryon, four as the maximum number of quarks and antiquarks in an elementary hadron, and just four hadron, generations (for a total of eight flavors) (Finkelstein, 1985). The electroweak interactions may be required in order to produce a consistent topological theory for the hadrons and the fine structure constant ($\alpha \simeq 1/137$) may be fixed in the process (Chew, 1983a). The standard (Salam–Weinberg) electroweak model has been seen as a natural outgrowth of TSMT (Chew and Finkelstein, 1983). Also, *separate* T (time reversal), C (charge conjugation) and P (parity) invariances for strong interactions in TSMT have been deduced (Jones and Finkler, 1985; Finkler and Jones, 1985; Jones, 1985). This brings us up to some of the latest claims of TSMT. Capra's review article (1984) can be consulted for recent further developments in that program. Let us leave this topic for now and turn to the operator formalism and quantized string program that grew out of the duality models.

8.5   The emergence of quantized string theories

Shortly after generalizations of Veneziano's (1968) original model had

been proposed, Fubini and Veneziano (1969) considered the factorization problem for the N-point duality amplitude. If an N-point amplitude A, such as that of Figure 8.7(a), is to exhibit duality in every channel, then it must be representable as a sum of resonance terms in *any* channel. For example, in the s channel it must reduce to a sum of terms (cf. Figure 8.7(b))

$$A \xrightarrow[s \to s_m]{} \frac{R(p,q)}{s - s_m} \qquad (8.26)$$

and the residue $R(p,q)$ must factorize as

$$R(p,q) = F(p)F(q) \qquad (8.27)$$

for a single level, or at least as

$$R(p,q) = \sum_{i=1}^{d_n} F_i(p)F_i(q) \qquad (8.28)$$

for d_n levels. This simultaneous factorization in *all* channels is an extremely stringent requirement. It must be demonstrated explicitly that such is the case, since an arbitrary function with the proper pole (or resonance) structure will not usually possess this property. Furthermore, if ghosts (states of negative norm) are to be avoided, the sum in Eq. (8.28) must be positive-definite.[12] Fubini and Veneziano (1969) were able to show that the duality amplitudes factored, but the residues were not necessarily positive definite. However, they demonstrated that because of a set of invariance transformations of these amplitudes, many (but not necessarily all) of these ghost states were effectively decoupled from the physical states. Using a representation of the N-point function due to Koba and Nielsen (1969), Fubini, Gordon and Veneziano (1969) were able to express these amplitudes in terms of
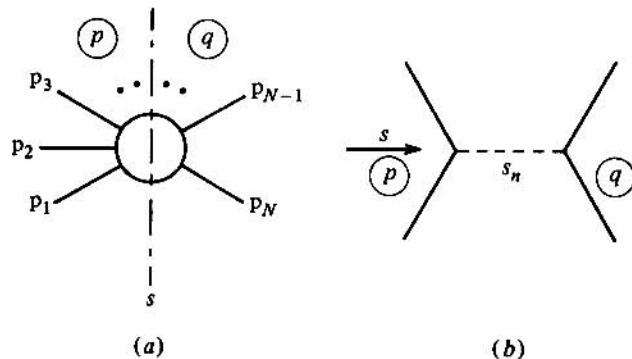

Figure 8.7   Residue factorization at a pole.

expectation values of functions of creation ($a^\dagger$) and annihilation (a) operators in such a way that this factorization property was manifestly guaranteed. This operator formalism was further explored by Fubini and Veneziano (1970) (Alessandrini, Amati, Le Bellac and Olive, 1971). Nambu (1970) also handled factorization by means of an operator formalism in which the basic operators a and $a^\dagger$ satisfy the 'harmonic oscillator' commutation relations[13]

$$[a, a^\dagger] = 1. \qquad (8.29)$$

On the basis of a formal analogy for the energy of an intermediate state, Nambu (1970) observed that the internal energy of a meson can be interpreted in terms of the normal modes of vibration of a string of finite length. It is interesting to note that, although string theories *could* (logically) have been formulated decades ago, they might never in fact have been considered had it not been for SMT *via* the duality models (Mandelstam, 1985). John Schwarz (1975, p. 64), one of the central figures in the construction of superstring theories, has made a similar observation.

> The theory of light strings could have been investigated 40 years ago, but there was little motivation for doing so until the connection with dual models was suspected.

Goddard, Goldstone, Rebbi and Thorn (1973) showed that the Lagrangian in the string theories had a large invariance group which generated the Virasoro (1970) gauges to eliminate the ghosts. Scattering of strings can be pictured as strings breaking and joining at the ends (Mandelstam, 1973a, 1973b).[14] Thus, dual resonance models can be treated either as S-matrix theories (Veneziano, 1974c) or as a system of interacting strings (Rebbi, 1974), although the string formulation can be made manifestly ghost-free (Mandelstam 1974b; Frampton, 1974; Jacob, 1974). We see that the string theory can be looked upon as a natural outgrowth of S-matrix ideas. The existence of ghosts (which involve negative probabilities, the sum of which need not equal unity – a violation of unitarity) can be seen as a conflict between the demands that the theory satisfy the principles of both (special) relativity and quantum mechanics (Schwarz, 1975). In the simplest string model, the space-time dimension must be 26 for these simultaneous demands to be met! (Lovelace, 1971; Cremmer and Scherk, 1972; Olive and Scherk, 1973a). A model allowing both fermions and mesons (Neveu and Schwarz, 1971a, 1971b; Thorn, 1971) can be proven consistent for a space-time of 10 dimensions (Olive and Scherk, 1973b). The tachyon problem (violation of causality) remained.

All of these string models contained a massless spin-2 particle. This was considered undesirable for a theory of *strong* interactions (Schwarz, 1982a), since there is no such strongly-interacting particle known. However, Scherk and Schwarz (1974a, 1974b, 1975) proved that this particle interacted as a graviton would. They suggested using string models as a basis for *all* particles, not just for the hadrons. Gliozzi. Scherk and Olive (1976, 1977) interpreted the string model as a supersymmetry[15] in which fermions and bosons are represented in the same state. Such a supersymmetry must be spontaneously broken since bosons and fermions do not occur in nature with the same mass. Superstring theories have no tachyons (Green and Schwarz, 1981; Schwarz 1982b). At low energies, some of these superstring theories have the possibility of reducing to the conventional gauge field theories.

These 10-dimensional superstring models are candidates for a quantum theory of all interactions in which all calculated quantities are finite (Schwarz, 1982a; Green and Schwarz, 1984a, 1984b, 1984c, 1985). The internal degrees of freedom of the string (as opposed to a structureless *point* particle) allows a *single* string to have an infinite number of states with masses and spins that increase without limit.[16] It appears that conventional (point) field theories cannot remain finite, or even renormalizable,[17] when the gravitational interactions are included. Extended (e.g., string) field theories, in a severely restricted number of cases, appear to have the possibility (likelihood?) of being renormalizable (or even finite) when gravity is included (Green, 1985). These latter theories may, in fact, be (nearly) *unique* (Schwarz, 1984, 1985; Gross, 1985).

> This was a unique state of affairs: the gauge symmetry of a theory being imposed by the requirement of quantum consistency. (Schwarzchild, 1985, p. 19).

A similar observation has been made by Richard Capps, a theoretical physicist who was prominently and actively involved in the dispersion theory and *S*-matrix bootstrap program.[18]

> [R]ecently I have been struck by the similarity of the motivation of the bootstrap idea and the motivation of modern string theory. ... The modern consistency condition is based on anomaly cancellation [a type of inconsistency that must be removed or prevented from occurring in a theory], rather than *S*-matrix analyticity, but the two approaches are more similar than most people realize.

Spontaneous compactification (a type of 'curling up' of certain dimensions), *via* an old idea due to Kaluza (1921) and Klein (1926), has been suggested to reduce the (observable) number of space-time dimensions from 10 down to the 4 of the real world. (Of course, this is an *ad hoc* theoretical move.) These theories even suggest the existence of a 'shadow world' of matter that can interact with ordinary matter *only* through the gravitational field (but not through the strong or electroweak ones so that it is 'invisible' to them) (Kolb, Seckel and Turner, 1985). This could offer a possible resolution of the 'missing mass' problem of cosmology (i.e., the discrepancy between the observed mass of the universe and that required by the observed motion of stars near the galaxies in the universe). Here, too, we have come to some of the latest developments in superstring theory, an area of much activity at the current time (Schwarzschild, 1985; Neveu, 1987).

8.6   Superstrings – the ultimate bootstrap

Let us make two observations about the relationship of string field theories (and gauge field theories like QCD) on one hand and *S*-matrix theory (including TSMT) on the other. Frazer (1985) has reiterated the possibility that the current rage – QCD – may be, like a previous theory of great interest – SMT – not really a *fundamental* theory. Both SMT and QCD have had successes in *different* domains of high energy physics (Pickering, 1984) and SMT is certainly not a theory that has been proven incorrect. Frazer (1985) summarizes Weinberg's (1985) two 'folk theorems': (1) that a quantum field theory with the most general Lagrangian is without content, being simply a way of implementing *S*-matrix theory; and (2) that, at low enough energies, a simple effective Lagrangian field theory can always be found. Frazer (1985) has termed this latter a 'harmonic oscillator theorem' (presumably in analogy with the result from classical mechanics that 'any' small oscillation about a stable equilibrium point is essentially a harmonic oscillation, no matter how complex the mechanical system).

The possibility certainly exists that SMT and QCD could be equivalent and/or 'low energy' limits of some larger theory (such as superstrings or TSMT). When one looks at the topological diagrams and expansions of superstrings and TSMT, one cannot help but be impressed by their (perhaps only formal and superficial) resemblance. In a recent paper, Balázs (1986, p. 1759) attempts to bootstrap superstrings from a self-consistent *S*-matrix formation.

> It is proposed that the superstring and its spectrum may not be fundamental but may themselves correspond to an approximate

solution of a self-consistent dynamical theory built from the basic principles of S-matrix unitarity, analyticity and crossing. In particular, it is shown that a simple approximate dual unitary scheme based on these principles does generate linear Regge trajectories and selects closed (rather than open) strings.

This would imply that superstrings would be *demanded* by the S-matrix principles. On this ambiguously optimistic note we end our technical discussion of the S-matrix program.

8.7   Summary

Within the framework of the S-matrix program of the late 1960s, finite energy sum rules were discovered and a duality was recognized which allowed one to consider a scattering amplitude either in terms of resonances at low energy or in terms of Regge behavior at high energy. Resonances in one channel were used to bootstrap a Regge pole in the crossed channel. Veneziano in 1968 constructed a simple mathematical model which explicitly exhibited this duality behavior and satisfied several S-matrix postulates *except* unitarity. This model was generalized to the N-point function and Veneziano in 1974 suggested a topological expansion to unitarize these amplitudes. This led to an independent duality program for calculations in strong-interaction physics.

In the late 1970s and early 1980s the S-matrix program was resuscitated in the form of a topological S-matrix theory. But, as regards novel empirical content, it was largely a camp follower to the highly successful gauge field theories. Matters of principle continue to motivate a relatively small band of present-day S-matrix theorists.

Nambu in 1969 discovered a connection between duality models, which had emerged from the S-matrix program, and the field theory of a string. Interest in quantized string theories arose because of the duality program, which had been rooted in S-matrix concepts. Although quantized strings were plagued for many years with consistency problems and with undesirable features, Scherk and Schwarz in 1974 recognized that they could be interpreted as representing gravitational as well as strong interactions. Green and Schwarz have recently found indications that all quantities calculated in these theories may be finite and that this requirement may yield a *unique* theory of all fundamental interactions.

On the verge of *Umdeutung* in Minnesota: Van Vleck and the correspondence principle. Part one

Anthony Duncan · Michel Janssen

Abstract   In October 1924, *The Physical Review*, a relatively minor journal at the time, published a remarkable two-part paper by John H. Van Vleck, working in virtual isolation at the University of Minnesota. Using Bohr's correspondence principle and Einstein's quantum theory of radiation along with advanced techniques from classical mechanics, Van Vleck showed that quantum formulae for emission, absorption, and dispersion of radiation merge with their classical counterparts in the limit of high quantum numbers. For modern readers Van Vleck's paper is much easier to follow than the famous paper by Kramers and Heisenberg on dispersion theory, which covers similar terrain and is widely credited to have led directly to Heisenberg's *Umdeutung* paper. This makes Van Vleck's paper extremely valuable for the reconstruction of the genesis of matrix mechanics. It also makes it tempting to ask why Van Vleck did not take the next step and develop matrix mechanics himself.

A. Duncan
Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, USA

M. Janssen
Program in the History of Science, Technology, and Medicine,
University of Minnesota, Minneapolis, USA

M. Janssen (✉)
Tate Laboratory of Physics, 116 Church St. NE, Minneapolis, MN 55455, USA
e-mail: janss011@umn.edu

1 Introduction

Most historians of modern physics agree that the famous *Umdeutung* [reinterpretation] paper with which Werner Heisenberg (1901–1976) laid the basis for matrix mechanics (Heisenberg, 1925c) grew out of a paper he and Hendrik A. (Hans) Kramers (1894–1952) co-authored on dispersion theory (Kramers and Heisenberg, 1925). Although hardly impartial as one of Kramers' students and his biographer, Max Dresden (1987) calls the Kramers–Heisenberg paper "the direct, immediate, and exclusive precursor to the Heisenberg paper on matrix mechanics" (p. 275). Martin J. Klein (1970) is more restrained but agrees that "this work was the immediate predecessor of Heisenberg's new quantum mechanics" (p. 31). To understand the origin of matrix mechanics, one thus has to come to grips with the contents of the Kramers–Heisenberg paper. According to Jagdish Mehra and Helmut Rechenberg, this paper was written "in such a way that every physicist, theoretician or experimentalist, interested in the subject could understand" (Mehra and Rechenberg, 1982–2001, Vol. 2, p. 181).[1] An uniniated modern reader turning to the Kramers–Heisenberg paper after these encouraging words is likely to be disappointed. The authors assume their readers to be thoroughly familiar with techniques, borrowed from celestial mechanics, for dealing with multiply-periodic systems, including canonical transformations, action-angle variables, and related perturbation methods. As far as their contemporaries in theoretical physics were concerned, this was undoubtedly a reasonable assumption. So, Mehra and Rechenberg are probably right to the extent that the intended audience would have had no special difficulties with the paper. The same cannot be said for most modern readers, who no longer have the relevant techniques at their fingertips. Fortunately, there is another paper from the same period covering some of the same terrain that is much easier to follow for such readers.

Immediately preceding the translation of (Kramers and Heisenberg, 1925) in the well-known anthology on the development of matrix mechanics edited by Bartel Leendert van der Waerden (1903–1996) (1968) is a paper by the American theoretical physicist John Hasbrouck Van Vleck (1899–1980) (1924b). Like the Kramers–Heisenberg paper, it combines some sophisticated classical mechanics with the correspondence principle of Niels Bohr (1885–1962) and elements of the quantum radiation theory of Albert Einstein (1879–1955). In the last section of this paper, Van Vleck showed that the Kramers dispersion formula, which Kramers (1924a,b) had only presented in two short notes in *Nature* at that point, merges with the classical formula in the limit of high quantum numbers. Van Vleck's paper is a paragon of clarity. In an interview by Thomas S. Kuhn for the *Archive for History of Quantum Physics* (AHQP) in 1963,[2] Van Vleck acknowledged the influence of his father, the

---

[1] This multi-volume history of quantum physics brings together a wealth of information and we shall frequently refer to it. However, it needs to be used with some caution (see, e.g., notes 5, 79, and 172 below as well as the review of the first few volumes by John L. Heilbron (1985)).

[2] Between February 1962 and May 1964, about 95 people were interviewed for the AHQP project (Kuhn et al., 1967, p. 3). With one exception (see Sect. 2.4) the exact dates of these interviews are unimportant for our purposes and will not be given when we quote from the transcripts. We consulted the copy of the *Archive for History of Quantum Physics* (cited hereafter as AHQP) at Walter Library, University of Minnesota.

mathematician Edward Burr Van Vleck (1863–1943), in developing his exceptionally lucid writing style:

> My father got after me for my very poor style of scientific exposition. I feel I owe a great deal to him for his splitting up my sentences into shorter sentences, avoiding dangling participles—i.e., tightening up my prose style—the same kind of drill I try to give my own graduate students now.[3]

Van der Waerden only included the quantum part (Van Vleck, 1924b), of a two-part paper in his anthology. In the second part, Van Vleck (1924c) clearly laid out the results from classical mechanics needed to understand the first part as well as those parts of (Kramers and Heisenberg, 1925) that are most important for understanding Heisenberg's *Umdeutung* paper. This is true even though Van Vleck only covered *coherent* scattering, in which the frequency of the incident and the scattered radiation is the same, whereas a large part of the Kramers–Heisenberg paper is devoted to *incoherent* scattering, first predicted in (Smekal, 1923) and verified experimentally a few years later (Raman 1928; Landsberg and Mandelstam 1928). In his interview with Kuhn, Heisenberg emphasized the importance of this part of his paper with Kramers for the *Umdeutung* paper.[4] Of course, this is also the part to which Heisenberg materially contributed.[5] Still, the non-commutative multiplication rule introduced in the *Umdeutung* paper may well have been inspired, as Heisenberg suggests, by manipulations in this part of the Kramers–Heisenberg paper. To understand where the arrays of numbers subject to this rule come from, however, it suffices to understand how coherent scattering is treated in Kramers' dispersion theory: indeed, the only explicit use of dispersion theory in the *Umdeutung* paper are results for coherent scattering.

1.1 On the verge of Umdeutung

As in the case of (Kramers and Heisenberg, 1925), one is struck in hindsight by how close (Van Vleck, 1924b,c) comes to anticipating matrix mechanics. During the AHQP interview, Kuhn reminded Van Vleck of a remark he had made 2 years earlier to the effect that, had he been "a little more perceptive," he "might have taken off from that paper to do what Heisenberg did." "That's true," Van Vleck conceded, but added with

---

[3] P. 21 of the transcript of the first of two sessions of the interview, quoted in (Fellows, 1985, p. 57). Van Vleck is talking specifically about the summer of 1925, when he was working on his book-length (Van Vleck, 1926), but his father had probably given him a few pointers before. (Van Vleck, 1924b) definitely belies the author's harsh judgment of his earlier writing style.

[4] P. 18 of the transcript of session 4 of a total of 12 sessions of the AHQP interview with Heisenberg.

[5] According to Dresden (1987, pp. 273–274), Kramers added Heisenberg's name to (Kramers and Heisenberg, 1925) mainly as a courtesy. For Heisenberg's side of the story, see pp. 15–18 of the transcript of session 4 of the AHQP interview with Heisenberg, several passages of which can be found in (Mehra and Rechenberg, 1982–2001, Vol. 2, pp. 178–179), although the authors cite their own conversations with Heisenberg as their source (cf. the foreword to Vol. 2).

characteristic modesty: "Perhaps I should say *considerably* more perceptive."[6] In the biographical information he supplied for the AHQP, Van Vleck noted:

> In the two or three years after my doctorate…my most significant paper was one on the correspondence principle for absorption…It was somewhat related to considerations based on the correspondence principle that led Heisenberg to the discovery of quantum mechanics, but I did not have sufficient insight for this.[7]

This modest assessment is reflected in the discussion of the relation between Van Vleck's work and matrix mechanics by Fred Fellows (1985, pp. 74–81), who wrote a superb dissertation covering the first half of Van Vleck's life and career. In a biographical memoir about his teacher and fellow Nobel laureate, Phil Anderson (1987)[8] is less reserved: "This paper comes tantalizingly close to the kind of considerations that led to Heisenberg's matrix mechanics" (p. 506).

Van Vleck did not pursue his own research any further in 1924 and instead spent months writing—and, as he jokingly put it, being a "galley slave" (Fellows, 1985, p. 100) of—a *Bulletin* for the *National Research Council* (NRC) on the old quantum theory (Van Vleck, 1926). With his masterful survey he would surely have rendered a great service to the American physics community had it not been for the quantum revolution of 1925–1926. Like the better-known *Handbuch* article by Wolfgang Pauli (1900–1958) (1926), the *Bulletin* was, as Van Vleck (1971) recognized, "in a sense…obsolete by the time it was off the press" (p. 6).[9] One is left wondering what would have happened, had the young assistant professor at the University of Minnesota continued to ponder the interaction between radiation and matter and the correspondence principle instead of fulfilling his duties as a newly minted member of the American physics community.

That Kramers and Van Vleck—and, one may add, Max Born (1882–1970) and Pascual Jordan (1902–1980)—came so close to beating Heisenberg to the punch makes the birth of matrix mechanics reminiscent of the birth of special relativity. The comparison seems apt, even though none of these authors anticipated as much of the new theory as H. A. Lorentz (1853–1928) and Henri Poincaré (1854–1912) in the case of relativity.[10] Heisenberg (1971, p. 63) himself actually compared his *Umdeutung* paper

---

[6] See p. 24 of the transcript of the first session of the interview. Kuhn's recollection is that Van Vleck's earlier remark was made during a meeting in Philadelphia in March 1961 to plan for the AHQP project (Kuhn et al., 1967, p. viii). Van Vleck was Kuhn's Ph.D. advisor and the two men co-authored (Kuhn and Van Vleck, 1950) (Anderson, 1987, p. 518). It was Van Vleck who approached Kuhn in February 1961 to offer him the directorship of the AHQP project (Kuhn et al., 1967, p. viii) (see also Baltas et al., 2000, pp. 302–303).

[7] Biographical information prepared for the American Institute of Physics project on the history of recent physics in the United States (included in the folder on Van Vleck in the AHQP), p. 1.

[8] Van Vleck, Anderson, and Sir Nevill Mott shared the 1977 Nobel Prize "for their fundamental theoretical investigations of the electronic structure of magnetic and disordered systems." Van Vleck won for the work begun in the early 1930s that earned him the title of "father of modern magnetism."

[9] For the reception of Van Vleck's *Bulletin*, see (Fellows, 1985, pp. 88–89). Van Vleck's *Bulletin* and Pauli's *Handbuch* article were not the only treatises on the old quantum theory that were out of date before the ink was dry. (Born, 1925) and (Birtwistle, 1926), two books on atomic mechanics, suffered the same fate.

[10] In his autobiography, Born (1978, pp. 216–217) exaggerated how close he came to matrix mechanics before Heisenberg.

to Einstein's relativity paper (Einstein, 1905), arguing that what they had in common was their insistence on allowing only observable quantities into physical theory. The analogy is considerably richer than that.

The breakthroughs of Einstein and Heisenberg consisted, to a large extent, in reinterpreting elements already present in the work of their predecessors, extending the domain of application of these elements, and discarding unnecessary scaffolding. Einstein recognized the importance of Lorentz invariance beyond electromagnetism, reinterpreted it as reflecting a new space–time structure, and discarded the ether (Janssen 2002). In the case of (Heisenberg, 1925c), the element of *Umdeutung* or reinterpretation is emphasized in the title of the paper. Heisenberg reinterpreted elements of the Fourier expansion of the position of an electron entering into the demonstration that the Kramers dispersion formula merges with the classical result in the correspondence limit, discarded the orbits supposedly given by that position, and recognized that the non-commuting arrays of numbers associated with transitions between different states and representing position in his new scheme were meaningful far beyond the dispersion theory from which they originated.

A further point of analogy is that neither Einstein nor Heisenberg presented the new theory in a particularly elegant mathematical form. In the case of relativity, this had to await the four-dimensional geometry of Hermann Minkowski (1864–1909) and the theory's further elaboration in terms of it by Arnold Sommerfeld (1868–1951), Max Laue (1879–1960), and others (Janssen and Mecklenburg, 2006). Even so, a modern reader will have no trouble recognizing special relativity in Einstein's 1905 paper. The same reader, however, will probably only start recognizing matrix mechanics in two follow-up papers to the *Umdeutung* paper, (Born and Jordan, 1925b) and (Born, Heisenberg, and Jordan, 1925), the famous *Dreimännerarbeit*.[11] Born first recognized that Heisenberg's new non-commuting quantities are matrices. Born and Jordan first introduced the familiar commutation relations for position and momentum. In the *Umdeutung* paper Heisenberg had used the Thomas–Kuhn sum rule, a by-product of the Kramers dispersion formula, as his fundamental quantization condition. As we shall see, Van Vleck had actually been the first to find the sum rule, although he only recognized the importance of the result later.

In the collective memory of the physics community, major discoveries understandably tend to get linked to singular events even though they are almost invariably stretched over time. The "discovery" of the electron by J. J. Thomson (1856–1940) in 1897 or the "discovery" of the quantum of action by Max Planck (1858–1947) in 1900 are well-known examples of this phenomenon. Special relativity is another good example of a "discovery" that has come to be associated with a single flash of insight, Einstein's recognition of the relativity of simultaneity, and a single emblematic text, "On the electrodynamics of moving bodies" (Einstein, 1905). Much the same can be said about Heisenberg's famous trip to Helgoland in June 1925 to seek relief from his seasonal allergies and the *Umdeutung* paper resulting from his epiphany on this barren

---

[11] During a lunch break in his AHQP interview, Alfred Landé (1888–1976) told Heilbron and Kuhn: "Heisenberg stammered something. Born made sense of it" (p. 10a of the transcript of sessions 1–4 of the interview; cf. note 174). Kuhn and Heilbron report that they wrote this down right after the conversation took place and call it a "Quasi-Direct Quote."

island. The way in which such stories become part of physics lore can be seen as a manifestation of what Robert K. Merton (1968) has dubbed the "Matthew effect," the disproportional accrual of credit to individuals perceived (sometimes retroactively) as leaders in the field.[12] We do, of course, recognize the singular importance of the contributions of Einstein to special relativity and of Heisenberg to matrix mechanics. But there is no need to exaggerate the extent of their achievements. They may have been the first to enter the promised land, to use another admittedly strained biblical metaphor, but they would never have laid eyes on it without some Moses-figure(s) leading the way.

In his biography of Kramers, Dresden makes a convincing case that his subject deserves more credit for matrix mechanics than he received: "Kramers certainly hoped and probably expected to be the single author of the Kramers–Heisenberg paper. It is probably futile to speculate how the credit for the discovery of matrix mechanics would have been distributed in that case. There would be an indispensable preliminary paper by Kramers alone, followed by a seminal paper by Heisenberg; this might well have altered the balance of recognition" (Dresden, 1987, p. 252). Citing this passage, Dirk ter Haar (1998, p. 23), like Dresden one of Kramers' students, raises the question whether Kramers would have shared Heisenberg's 1932 Nobel Prize in that case. In a curmudgeonly review of Dresden's book, however, Nico van Kampen, another one of Kramers' students, takes issue with the pattern of "near misses" that Dresden (1987, pp. 446–461) sees in Kramers' career, the discovery of matrix mechanics being one of them (Dresden, 1987, pp. 285–288). Van Kampen asks: "Is it necessary to explain that, once you have, with a lot of sweat and tears, constructed a dispersion formula on the basis of the correspondence principle, it is not possible to forget that background and that it takes a fresh mind to take the next step?" (Van Kampen, 1988). Similar claims can be made and similar questions can be raised in the case of Van Vleck, even though his work, unlike that of Kramers, did not directly influence Heisenberg.

Van Vleck's contribution has receded even further into the background in the history of quantum mechanics than Kramers'. (Van Vleck, 1924b,c) is not discussed in any of the currently standard secondary sources on quantum dispersion theory and matrix mechanics, such as (Jammer, 1966), (Dresden, 1987), or (Darrigol, 1992). Nor is it mentioned in Vol. 2 of (Mehra and Rechenberg, 1982–2001) on the discovery of matrix mechanics, although it is discussed briefly in Vol. 1 (pp. 646–647) on the old quantum theory and in Vol. 3 (p. 55) on the elaboration of matrix mechanics.[13] That he worked in faraway Minnesota rather than in Copenhagen or Göttingen, we surmise, is a major factor in this neglect of Van Vleck. Whatever the reason, the neglect is regrettable. For a modern reader, it is much easier to see in (Van Vleck, 1924b,c) than in (Kramers and Heisenberg, 1925) or in (Born, 1924) that matrix mechanics did not come as a bolt out of the blue, but was the natural outgrowth

---

[12] The effect is named for the following passage from the Gospel According to St. Matthew: "For unto everyone that hath shall be given, and he shall have in abundance: but from him that hath not shall be taken away even that which he hath."

[13] It is also mentioned in (Van der Waerden and Rechenberg, 1985, pp. 330–331) and in (Hund, 1984, pp. 131–132). As noted in (Mehra and Rechenberg, 1982–2001, Vol. 6, p. 348, note 407), Van Vleck's work is discussed prominently in a paper by Hiroyuki Konno (1993) on Kramers' dispersion theory.

of earlier applications of the correspondence principle to the interaction of radiation and matter.

Aitchison et al. (2004) have recently given a detailed reconstruction of the notoriously opaque mathematics of (Heisenberg, 1925c). By way of motivating their enterprise, they quote the confession of Steven Weinberg (1992) that he has "never understood Heisenberg's *motivations* for the mathematical steps in his paper" (p. 67; our emphasis). These authors clearly explain the mathematical steps. The motivations for these steps, however, cannot be understood, we submit, without recourse to the dispersion theory leading up to his paper. And if we want to retrace Heisenberg's steps on his sojourn to Helgoland, Van Vleck may well be our best guide.

1.2 Structure of our paper

Like Van Vleck's 1924 paper, our paper comes in two parts, the second providing the technical results needed to understand the first in full detail. To provide some context for Van Vleck's work, undertaken far from the European centers in quantum theory, we begin Part One by addressing the question of America's "coming of age" in theoretical physics in the 1920s (Sect. 2). In Sect. 3, we relate the story of how matrix mechanics grew out of dispersion theory in the old quantum theory, drawing on the extensive secondary literature on this episode as well as on the materials brought together in the AHQP. This story is usually told from a Eurocentric perspective. Following our discussion in Sect. 2, we shall look at it from a more American vantage point. Discussion of the famous BKS theory (Bohr, Kramers, and Slater, 1924a), which is prominently mentioned in many papers on dispersion theory in 1924–1925, is postponed until Sect. 4. We shall pay special attention to the role of Van Vleck's fellow graduate student at Harvard, John C. Slater (1900–1976).[14] The reason for keeping the discussion of BKS separate from the discussion of dispersion theory is that we want to argue that the rise and fall of BKS was largely a sideshow distracting from the main plot line, which runs directly from dispersion theory to matrix mechanics. In hindsight, BKS mainly deserves credit for the broad dissemination of its concept of "virtual oscillators." Contrary to widespread opinion, both among contemporaries and among later historians, these virtual oscillators did not originate in the BKS theory. They were introduced the year before, under a different name and in the context of dispersion theory, by the Breslau (now Wrocław, Poland) physicists Rudolf Ladenburg (1882–1952) and Fritz Reiche (1883–1969), who called them "substitute oscillators" [*Ersatzoszillatoren*[15]] (Ladenburg and Reiche, 1923, p. 588, p. 590). This paper is important in its own right and underscores the key achievement of Van Vleck's two-part paper. Both Van Vleck (1924b,c) and Ladenburg and Reiche (1923) discuss the relation between quantum and classical expressions for emission, absorption, and dispersion in view of Bohr's correspondence principle. Van Vleck's discussion is impeccable in all three cases; Ladenburg and Reiche made serious errors in the case of both dispersion and absorption. The expertise Van Vleck had gained in classical mechanics through

---

[14] On Slater, see, e.g., (Schweber, 1990).

[15] We follow the translation used in (Konno, 1993, e.g., p. 139).

his work on the problem of helium in the old quantum theory (Van Vleck 1922a,b)
put him in an ideal position to correct these errors. We suggest that this is in part what
he wanted to do with (Van Vleck, 1924b,c).[16]

In Sect. 5, the first section of Part Two, we give an elementary and self-contained
presentation, drawing on (Van Vleck, 1924b,c), of the technical results on which
our narrative in Sects. 3 and 4 rests. In particular, we use canonical perturbation
theory in action-angle variables to derive a classical formula for the dispersion of
radiation by a charged harmonic oscillator and apply the correspondence principle to
that formula to obtain the Kramers dispersion formula for this special case.[17] This fills
an important pedagogical gap in the historical literature. Given the central importance
of the Kramers dispersion formula for the development of quantum mechanics, it is
to be lamented that there is no explicit easy-to-follow derivation of this result in the
extensive literature on the subject. In the later parts of Sect. 5 and in Sect. 6, we take
a closer look at Van Vleck's main concerns in his 1924 paper, which was absorption
rather than dispersion and the extension of results for the special case of a charged
harmonic oscillator (which suffices to understand how matrix mechanics grew out of
dispersion theory) to arbitrary non-degenerate multiply-periodic systems. In Sect. 7,
we present a simple modern derivation of the Kramers dispersion formula and related
results, which we hope will throw further light on derivations and results in Sects. 5
and 6 as well as on the narrative in Sects. 3 and 4. Finally, in Sect. 8, we bring together
the main conclusions of our investigation.

2 Americans and quantum theory in the early 1920s

"[A]lthough we did not start the orgy of quantum mechanics, our young theorists
joined it promptly" (Van Vleck, 1964, p. 24).[18] This is how our main protagonist,
known to his colleagues simply as "Van", described the American participation in
the quantum revolution of the mid-1920s for an audience in Cleveland in 1963. Van
Vleck spoke as the first recipient of an award named for America's first Nobel Prize
winner in physics, Albert A. Michelson (1825–1931). Van Vleck was selling himself
and his countrymen short by characterizing the American contribution to the quantum
revolution as simply a matter of joining an orgy started by the Europeans and in full
swing by the time the Americans arrived on the scene.

Eight years later, Van Vleck, in fact, took exception to what sounds like a simi-
lar characterization given by another leading American physicist of his generation,
Isidor I. Rabi (1898–1988). Van Vleck quoted a comment that Rabi made in a TV
documentary about Enrico Fermi (1901–1954):

---

[16] (Ladenburg, 1921) and (Ladenburg and Reiche, 1923) are cited in (Van Vleck, 1924b, p. 339).

[17] Van Vleck did it the other way around: he derived the classical formula and showed that it merges with
Kramers' quantum formula in the correspondence limit. In Sect. 5.2, we shall quote from an exchange
between Born and Van Vleck that makes it clear that Van Vleck felt that it did not really matter whether
one used the correspondence principle to *construct* quantum formulae or to *check* them.

[18] Quoted and discussed in (Coben, 1971, p. 456).

> We had produced a large number of people who had been brought up to a certain level, then needed some help, some leadership to get over the hump. Once they were over the hump they were tremendous. People of my generation brought them over the hump, largely from attitudes, tastes, and developments which we had learned in Europe (Van Vleck, 1971, p. 7).

As Kuhn and others have emphasized, Rabi's point was that American physicists returning from Europe rather than European émigrés were mainly responsible for the coming of age of American physics.[19] This issue has been hotly debated in the history of physics literature.[20] Our study of some early American contributions to quantum theory supports the observation by Sam Schweber (1986) that in the 1930s theoretical physics was "already a thriving enterprise in the United States. The refugee scientists resonated with and reinforced American strength and methods: they did not create them" (p. 58).

Commenting on Rabi's remark, Van Vleck (1971) reiterated the point of his Michelson address that "quantum mechanics was a basically European discovery" (p. 6). In (Van Vleck, 1929), he had likewise characterized it as "the result of the reaction of mind on mind among European talent in theoretical physics" (p. 467). In 1971, however, he added that "there has been too much of an impression that American physicists, even in the application of quantum mechanics, were effective only because they had the aid of European physicists, either by going to Europe, or because of their migration to America" (Van Vleck, 1971, p. 6). Van Vleck, who was proud to be a tenth-generation American,[21] received his entire education in the United States. He hardly had any contact with European physicists before 1925, although he did meet a few on a trip to Europe with his parents in the summer of 1923. In Copenhagen, he called on Bohr, who suggested that he get in touch with Kramers,[22] Bohr's right-hand man throughout the period of interest to us. Kramers was not in Denmark at the time but in his native Holland. Decades later, when he received the prestigious Lorentz medal from the *Koninklijke Akademie van Wetenschappen* in Amsterdam, Van Vleck recalled how he had searched for Kramers high and low. After he had finally tracked him down—it can no longer be established whether this was in Bergen aan Zee or in Schoorl—the two men went for a long walk in the dunes along the North-Sea coast: "This was the beginning of a friendship that lasted until his passing in 1952" (Van Vleck, 1974, p. 9). Unfortunately, Van Vleck does not tell us what he and Kramers talked about.

---

[19] See p. 20 of the transcript of the last of five sessions of Kuhn's AHQP interview with George E. Uhlenbeck (1900–1988).

[20] For a concise summary and detailed references to the older literature, see (Moyer, 1985, pp. 171–173). Whereas our focus will be on American contributions to atomic physics, Alexi Assmus (1992, 1999) has argued that American theoretical physics came of age in molecular physics (cf. note 45 below).

[21] He could trace his ancestry back to the fifteenth century, to a certain Johan van Vleeck of Maastricht. One of the latter's descendants, Tielman van Vleeck (or von Fleck), left Bremen for New Amsterdam in 1658 (Fellows, 1985, pp. 5–6).

[22] See p. 14 of the transcript of session 1 of the AHQP interview with Van Vleck.

2.1 Education

Van Vleck learned the old quantum theory of Bohr and Sommerfeld at Harvard as one of the first students to take the new course on quantum theory offered by Edwin C. Kemble (1889–1984), the first American physicist to write a predominantly theoretical quantum dissertation. Kemble, Van Vleck (1992) wrote in an autobiographical note accompanying the published version of his Nobel lecture, "was the one person in America at that time qualified to direct purely theoretical research in quantum atomic physics" (p. 351). Kemble's course roughly followed (Sommerfeld, 1919), the bible of the old quantum theory. Van Vleck supplemented his studies by reading (Bohr, 1918) and (Kramers, 1919) (Fellows, 1985, p. 17).

Van Vleck was part of a remarkable cohort of young American quantum theorists, which also included Slater, Gregory Breit (1899–1981), Harold C. Urey (1893–1981), Robert S. Mulliken (1896–1987), and David M. Dennison (1900–1976). Just as Van Vleck was the first to write a purely theoretical dissertation at Harvard in 1922, Dennison was the first to do so at the University of Michigan in 1924.[23] Dennison could take advantage of the presence of Oskar Klein (1894–1977), an early associate of Bohr,[24] who was a visiting faculty member in the physics department in Michigan from 1923 to 1925 (Sopka, 1988, p. 321). This is where Klein came up with what is now known as the Klein–Gordon equation; it is also where he made his contribution to what is now known as the Kaluza–Klein theory.[25]

Reminiscences about the early days of quantum physics in the United States can be found in (Van Vleck, 1964, 1971), (Slater, 1968, 1973, 1975), and (Rabi, 2006). It is also an important topic of conversation in the AHQP interviews with Van Vleck, Slater, Dennison, and Kemble. These interviews need to be handled with care. In the case of Slater and Van Vleck, one can say, roughly speaking, that the former had a tendency to exaggerate the importance of American contributions, especially his own, while the latter tended to downplay their importance. In sharp contrast, for instance, to the modest remarks by Van Vleck quoted in Sect. 1.1, Slater boasted that he "was really working toward quantum mechanics before quantum mechanics came out. I'm sure if it was delayed a year or so more, I would have got it before the others did."[26]

The older generation—men such as Michelson and Robert A. Millikan (1868–1953)—recognized that the United States badly needed to catch up with Europe in quantum physics. The Americans were already doing first-rate experimental work. Theory, however, was seriously lagging behind. As the German–American–Dutch physicist Ralph Kronig (1904–1995) described the situation in an important essay in the Pauli memorial volume:

> While in experimental physics a number of investigators like Michelson, Millikan, Langmuir, Compton and R. W. Wood, ranking among the foremost in the world, continued a tradition of pioneer research that went back to Franklin,

[23] See p. 10 of the transcript of the first of three sessions of Kuhn's AHQP interview with Dennison.

[24] See (O. Klein, 1967) for his reminiscences about his early days in Copenhagen.

[25] See p. 13 of the transcript of session 5 of the AHQP interview with Uhlenbeck.

[26] See p. 40 of the transcript of the first session of the AHQP interview with Slater.

Henry, and Rowland, theoretical physics, after the meteoric appearance of Gibbs, could not boast of a similar record…There was, it is true, a somewhat disperse group of younger men in America, endeavouring to come up to scratch in [atomic physics], of which I should mention Kemble, Van Vleck, Breit, Slater, and Mulliken, but their mutual contacts were limited (Kronig, 1960, p. 17).

Kronig, born in Dresden, came to the United States in 1919 and got his Ph.D. at Columbia University in 1924. After an extended trip to Europe on a Columbia traveling fellowship, he taught in Columbia for 2 years before returning to Europe for good in 1927 (see the folder on Kronig in the AHQP). Kronig's impression is confirmed by Van Vleck's teacher, Ted Kemble:

[T]he only theoretical physicists in the country at that time were really men on whom the load of teaching all the mathematical physics courses lay, and they all spent their time teaching. It wasn't, as I remember, a constructive occupation.[27]

The one theorist who, in Kemble's estimation, was active in research in classical theory, Arthur Gordon Webster (1863–1923), was never able to make the transition to quantum theory. Webster, Kemble said,

just couldn't keep up with what was going on when the quantum theory began. I always understood that the reason he killed himself was simply because he discovered that suddenly physics had gone off in a new direction and he was unable to follow, and couldn't bear to take a seat in the back and be silent.[28]

When quantum theory arrived on the scene, some experimentalists tried their hands at teaching it themselves (Coben, 1971, p. 444). In this climate, young American physicists with a knack for theory became a hot commodity. They received fellowships to learn the theory at the feet of the masters in Europe and offers of faculty positions straight out of graduate school.[29]

2.2 Postdocs and faculty positions

The careers of the young theorists listed above amply illustrate the new opportunities in the mid-1920s. Slater went to Europe on a Sheldon fellowship from Harvard and spent the first half of 1924 with Bohr and Kramers in Copenhagen. During this period, Urey and Frank C. Hoyt (1898–1977) were in Copenhagen as well, Urey on a small fellowship from the American–Scandinavian Foundation, Hoyt on a more generous NRC fellowship paid for by the Rockefeller foundation.[30] Among the visitors the Americans got to meet in Bohr's institute were Heisenberg and Pauli. Hoyt, a promising

---

[27] See p. 4 of the transcript of the last two of three sessions of the AHQP interview with Kemble. See also p. 10 of the transcript of the first of session.

[28] See p. 12 of the transcript of the first session of the AHQP interview with Kemble.

[29] For further discussion of quantum physics in America before the mid-1930s, see (Coben, 1971), (Seidel, 1978), (Kevles, 1978, pp. 168–169), (Weart, 1979), (Schweber, 1986), (Holton, 1988), and, especially, (Sopka, 1988).

[30] See (Robertson, 1979, p. 157), (Sopka, 1988, pp. 71, 97), and Slater to Van Vleck, July 27, 1924 (AHQP).

student who never reached the level of distinction of the cohort immediately following him,[31] was in Copenhagen for almost 2 years, from October 1922 to September 1924, Urey for less than one, from September 1923 to June 1924, and Slater only for a few months, from December 1923 to April 1924. Slater did not have a good experience in Copenhagen. This transpires, for instance, in the letter he wrote to Van Vleck on his way back to the United States. Off the coast of Nantucket, a few hours before his ship—The Cunard R.M.S. "Lancastria"—docked in New York, he wrote:

> Don't remember just how much I told you about my stay in Copenhagen. The paper with Bohr and Kramers [proposing the BKS theory] was got out of the way the first six weeks or so—written entirely by Bohr and Kramers. That was very nearly the only paper that came from the institute at all the time I was there; there seemed to be very little doing. Bohr does very little and is chronically overworked by it…Bohr had to go on several vacations in the spring, and came back worse from each one.[32]

In October 1924, Dennison arrived in Copenhagen, on an *International Education Board* (IEB) fellowship, another fellowship paid for by the Rockefeller Foundation.[33] The state of quantum theory in America was already beginning to change at that point. Like Hoyt, Dennison had been awarded a NRC fellowship, but had been told that he could only spend the money at an American institution.[34] In 1923, the NRC had likewise rejected the proposal of Mulliken to go work with Ernest Rutherford (1871–1937) in Cambridge. Mulliken became a NRC research fellow at Harvard instead (Assmus, 1992, p. 23).

Van Vleck and Slater, who both started graduate school at Harvard in 1920 (Van Vleck in February, Slater in September) and lived in the same dormitory,[35] had at one point discussed going to Copenhagen together upon completion of their Ph.D. degrees in 1923. In the end, Van Vleck went to Minneapolis instead. In the biographical note accompanying his Nobel lecture from which we already quoted above, he reflected:

> I was fortunate in being offered an assistant professorship at the University of Minnesota…with purely graduate courses to teach. This was an unusual move by that institution, as at that time, posts with this type of teaching were usually

---

[31] He wrote several papers on applications of Bohr's correspondence principle (Hoyt, 1923, 1924, 1925a,b). The first two are cited in (Van Vleck, 1924b, p. 334) and all but the second are cited in (Van Vleck, 1926, pp. 124, 146). The second paper is cited in (Ladenburg and Reiche, 1924, p. 672). Hoyt also translated Bohr's Nobel lecture into English (Bohr, 1923a). Hoyt ended up making a career in weapons research rather than in academic physics. After the war, he worked at Argonne National Laboratory, Los Alamos, and Lockheed. He was interviewed for the AHQP by Heilbron but did not remember much of the early days of quantum theory.

[32] Slater to Van Vleck, July 27, 1924 (AHQP). The second sentence of this passage is quoted by Dresden (1987, p. 165) in the course of his detailed discussion of Slater's reaction to his experiences in Copenhagen.

[33] Bohr arranged for one of these fellowships to pay for Heisenberg's visit to Copenhagen in the fall of 1924 (Cassidy, 1991, pp. 180, 183). See also the acknowledgment in (Heisenberg, 1925b, p. 860).

[34] See p. 12 of the transcript of session 1 of the AHQP interview with Dennison.

[35] See Van Vleck, *1920–1930. The first ten years of John Slater's scientific career.* Unpublished manuscript, American Institute of Physics (AIP), p. 2.

reserved for older men, and recent Ph.D.'s were traditionally handicapped by heavy loads of undergraduate teaching which left little time to think about research (Van Vleck, 1992, p. 351).

When the university hired Van Vleck it also hired Breit so that its new recruits would not feel isolated.[36] Breit is one of the more eccentric figures of twentieth-century American physics. He was born in Russia and came to the United States in 1915. In a biographical memoir of the *National Academy of Sciences*, we read that

John Wheeler relates a story told to him by Lubov [Gregory's sister] that she and Gregory were vacationing on the sea when the call to leave Russia came, and they "came as they were." For Gregory this meant dressed in a sailor suit with short pants; he was still wearing it when he enrolled in Johns Hopkins (at age sixteen!). Wheeler attributes some of Gregory's subsequent reticence to the ragging he took at the hand of his classmates for his dress (Hull, 1998, pp. 29–30).

True to form, Breit declined to be interviewed for the AHQP. In a memorandum dated April 8, 1964 (included in the folder on Breit in the AHQP), Kuhn describes how they met for lunch, but did not get beyond "casual reminiscences." Kuhn ends on a positively irritated note: "we broke off amicably but with zero achievement to report for the project."

Breit and Van Vleck replaced W. F. G. Swann (1884–1962) who had left Minneapolis for Chicago, taking his star graduate student Ernest O. Lawrence (1901–1958) with him. As Van Vleck (1971) notes wryly: "A common unwitting remark of the lady next to me at a dinner party was "Wasn't it too bad Minnesota lost Swann—it took two men to replace him!"" (p. 6).

Just as Minnesota hired both Breit and Van Vleck in 1923, the University of Michigan hired not one but two students of Paul Ehrenfest (1880–1933) in 1927, Uhlenbeck and Samuel A. Goudsmit (1902–1978) (Coben, 1971, p. 460).[37] In addition Michigan hired Dennison, its own alumnus, upon his return from Copenhagen. Ann Arbor thus became an important center for quantum theory, especially in molecular physics (Assmus, 1992, pp. 4, 26, 30). While Uhlenbeck and Goudsmit essentially remained in Ann Arbor for the rest of their careers, neither Breit nor Van Vleck stayed long in Minneapolis. Breit left for the Carnegie Institution of Washington after only 1 year, Van Vleck for the University of Wisconsin, his *alma mater*, after five.[38] Van Vleck agonized over the decision to leave Minnesota, where he had been promoted to associate professor in June 1926 and, only a year later, to full professor (Fellows, 1985, Ch. VII). Moreover, on June 10, 1927, he had married Abigail Pearson (1900–1989),

[36] See p. 14 and p. 18 of the transcript of session 1 of the AHQP interview with Van Vleck.

[37] See also (Sopka, 1988, p. 149) and the AHQP interview with Dennison. The recruiters were Walter F. Colby (1880–1970) and Harrison M. Randall (1870–1969).

[38] The mathematics building in Madison is named after Van Vleck's father, who was a professor of mathematics at the University of Wisconsin from 1906 until his retirement in 1929.

whom he had met while she was an undergraduate at the University of Minnesota and who had strong ties to Minneapolis.[39]

To replace Van Vleck, Minnesota made the irresistible offer of a full professorship to the young Edward U. Condon (1902–1974). Minnesota had offered Condon an assistant professorship the year before. At that point, Condon had received six such offers and had decided on Princeton (Condon, 1973, p. 321). His laconic response to this embarrassment of riches: "The market conditions for young theoretical physicists continues [sic] to surprise me" (Coben, 1971, p. 463). Before his first Minnesota winter as a full professor, Condon already regretted leaving New Jersey. He returned to Princeton the following year. Condon, Rabi, and J. Robert Oppenheimer (1904–1967)[40] were the leaders of the cohort of American quantum theorists graduating right after the quantum revolution of 1925. The cohort most relevant to our story graduated right before that watershed.

2.3 The physical review

It was during Van Vleck's tenure in Minnesota that his senior colleague John T. (Jack) Tate (1889–1950) took over as editor-in-chief of *The Physical Review* (Sopka, 1988, pp. 142–145, 203, note 11). Tate edited the journal from 1926 to 1950.[41] Van Vleck (1971) described the change of editorship as "another revolution" in the "middle of the quantum revolution" (pp. 7–8). Van Vleck was highly appreciative of Tate's role: "He published my papers very promptly, and also often let me see manuscripts of submitted papers, usually to referee" (ibid.). Thanks in no small measure to Van Vleck and other young whippersnappers in quantum theory, Tate turned what had been a lack-luster publication into the prestigious journal it still is today. Van Vleck recalled the transformation:

> *The Physical Review* was only so-so, especially in theory, and in 1922 I was greatly pleased that my doctor's thesis [Van Vleck, 1922] was accepted for publication by the *Philosophical Magazine* in England…By 1930 or so, the relative standings of *The Physical Review* and *Philosophical Magazine* were interchanged…Prompt publication, beginning in 1929, of "Letters to the Editor" in *The Physical Review*…obviated the necessity of sending notes to *Nature*, a practice previously followed by our more eager colleagues [see, e.g., Breit, 1924b, Slater, 1924, 1925c] (Van Vleck, 1964, pp. 22, 24).

Van Vleck's impression is corroborated by two foreign-born theorists who made their careers in the United States, Rabi and Uhlenbeck (Coben, 1971, p. 456). Rabi was

---

[39] After her husband's death, Abigail made a generous donation to the University of Minnesota to support the Abigail and John van Vleck Lecture Series. Phil Anderson gave the inaugural lecture in 1983 and the series has brought several Nobel Prize winners to Minneapolis since. The main auditorium in the building currently housing the University of Minnesota physics department is also named after the couple.

[40] Oppenheimer enrolled as an undergraduate at Harvard in 1922, 2 years after Van Vleck and Slater started graduate school there.

[41] It is largely in recognition of this achievement that the current Minnesota physics building is named after him.

born in Galicia but moved to New York City as an infant. Rabi liked to tell the story of how, when he returned to Europe to study quantum theory in Germany in 1927, he discovered that *The Physical Review* "was so lowly regarded that the University of Göttingen waited until the end of the year and ordered all twelve monthly issues at once to save postage" (ibid.). On other occasions, Rabi told this story about Hamburg University (Rigden, 1987, p. 4). He told Jeremy Bernstein (2004) that "in Hamburg so little was thought of the journal…that the librarian uncrated the issues only once a year" (p. 28). The following exchange between Kuhn and Heisenberg, talking about the early 1920s, is also revealing:

> Heisenberg: "What was the American paper at that time?"
> Kuhn: "The *Physical Review*?"
> Heisenberg: "No, that didn't exist at that time. I don't think so. Well, in these early times it probably didn't play a very important role."[42]

In a talk about Condon, Rabi elaborated on the mediocrity of *The Physical Review*:

> it was not a very exciting journal even though I published my dissertation in it. And we felt this very keenly. Here was the United States, a vast and rich country but on a rather less than modest level in its contribution to physics, at least per capita. And we resolved that we would change the situation. And I think we did. By 1937 the *Physical Review* was a leading journal in the world (Rabi, 1975, p. 7).

Uhlenbeck remembered how as a student in Leyden he viewed *The Physical Review* as "one of the funny journals just like the Japanese."[43] His initial reaction to the job offer from Michigan suggests that, at least at the time, his disdain for American physics journals extended to the country as a whole: "If it had been Egypt or somewhere like that, I would have gone right away, or China, or even India, I always wanted to go to exotic places [Uhlenbeck was born in Batavia in the Dutch East Indies, now Jakarta, Indonesia]; but America seemed terribly dull and uninteresting" (Coben, 1971, p. 460). In the AHQP interview with Uhlenbeck, one finds no such disparaging remarks. In fact, Uhlenbeck talks about how he had reluctantly agreed to return to the Netherlands in 1935 to replace Kramers, who had left Utrecht for Leyden to become Ehrenfest's successor after the latter's suicide.[44] Uhlenbeck was back in Ann Arbor in 1939.

2.4 The lack of recognition of early American contributions to quantum theory

Given the disadvantage they started out with, American theorists in the early 1920s would have done well had they just absorbed the work of their European counterparts and transmitted it to the next generation. They did considerably better than that. Even before the breakthrough of Heisenberg they started making important contributions themselves. According to Assmus (1992), however, "[a]tomic physics was shark

---

[42] See p. 5 of the transcript of session 3 of the AHQP interview with Heisenberg.

[43] See p. 20 of the transcript of session 5 of the AHQP interview with Uhlenbeck.

[44] See p. 9 of the transcript of session 5 of the AHQP interview with Uhlenbeck.

infested waters and was to be avoided; U.S. physicists would flourish and mature in the calmer and safer tidepools of molecular physics" (p. 8; see also Assmus, 1999, p. 187). She sees the early contributions of Van Vleck and Slater to atomic physics, which will be the focus of our study, as exceptions to this rule:

> Van Vleck and Slater viewed themselves as the younger generation, as central figures in the "coming of age" of U.S. physics. They had been given the knowledge that Kemble and his generation could provide and felt themselves capable of pushing into areas where the physics community in the United States had not dared to venture. Still, after experiences had muted their youthful exuberance, they turned to the by-then traditional problems of American quantum physics[,] problems that addressed the building up of matter rather than its deconstruction (Assmus, 1992, p. 22).

We hope to show that American work in atomic physics was significantly more important—if not in quantity, then at least in quality—than these remarks suggest.[45] Slater was one of the architects of the short-lived but highly influential Bohr–Kramers–Slater (BKS) theory (Bohr, Kramers, and Slater, 1924a) (see Sect. 4). Van Vleck's two-part article in *The Physical Review* (Van Vleck, 1924b,c), which is the focus of our study, is less well-known.

Originally, Van Vleck's paper was to have three parts. A rough draft of the third part has been preserved.[46] Van Vleck did not finish the third part at the time. As he explained in a letter to Born on November 13, 1924 (AHQP): "Part III which is not yet ready relates to classical black body radiation rather than quantum theory." It was only toward the end of his life that he returned to the masterpiece of his youth. Three years before he died he published a paper, co-authored with D. L. Huber, that can be seen as a substitute for part III. As the authors explain:

> Part III was to be concerned with the equilibrium between absorption and emission under the Rayleigh–Jeans law. It was never written up for publication because in 1925 the author was busy writing his book [Van Vleck, 1926a] and of course the advent of quantum mechanics presented innumerable research problems more timely than a purely classical investigation. The idea occurred to him to use the 50th anniversary of Parts I and II as the date for publishing a paper which would start with Part III and might even bear its title. Although he did not succeed in meeting the deadline, it still provided a partial motivation for collaborating on the present article (Van Vleck and Huber, 1977, p. 939).

It was at the suggestion of Jordan that van der Waerden included the first (quantum) part of Van Vleck's 1924 paper in his anthology on matrix mechanics (Van der Waerden, 1968, see the preface).[47] Interviewing Van Vleck for the AHQP in October 1963,

---

[45] Assmus is probably right, however, that the Americans contributed more to molecular than to atomic physics. This would fit with the thesis of (Schweber, 1990) that "Americans contributed most significantly to the development of quantum mechanics in quantum chemistry" (pp. 398–406).

[46] *American Institute of Physics*, Van Vleck papers, Box 17. We are grateful to Fred Fellows for alerting us to this manuscript.

[47] See also (Sopka, 1988, pp. 110–111).

Kuhn claimed that Jordan had told him that Born and Jordan "were working quite hard in an attempt to reformulate it [Van Vleck, 1924b,c] and had been multiplying Fourier coefficients together,[48] just at the time they got the Heisenberg paper that was going to be matrix mechanics."[49] In fact, a paper by Born and Jordan (1925a) building on (Van Vleck, 1924b,c) was submitted to *Zeitschrift für Physik* on June 11, 1925, several weeks *before* Heisenberg's breakthrough (Cassidy, 1991, p. 198). We therefore suspect that Kuhn misremembered or misconstrued what Jordan had told him during an interview for the AHQP in June 1963, a few months before the interview with Van Vleck. Van Vleck's paper is brought up during the second session of the interview (see p. 14 of the transcript). In this exchange Kuhn insisted that (Born and Jordan, 1925a) had come out *before* (Van Vleck, 1924b,c). Jordan corrected Kuhn at the beginning of the third session, which prompted some further discussion of Van Vleck's paper. However, it was Kuhn, not Jordan, who suggested at that point that Born and Jordan continued to pursue the ideas in Van Vleck's paper even after publishing (Born and Jordan, 1925a). Jordan did not confirm this. Still, although Kuhn probably embellished the story, there is no question that Van Vleck's paper had a big impact on the work of Born and Jordan. Jordan emphasized this in the interview with Kuhn, in a letter to van der Waerden of December 1, 1961 (quoted in Van der Waerden, 1968, p. 17), and in (Jordan, 1973). We quote from this last source:

> Van Vleck gave a derivation of Einstein's laws of the relation between the probabilities of spontaneous emission and positive and negative absorption. This result of Einstein's had been looked upon for a long time in a sceptical manner by Niels Bohr; now it was highly interesting to see, just how from Bohr's preferred way of thinking, a derivation of Einstein's law could be given. Born and I performed a simplified mathematical derivation of the results of Van Vleck. Our article on this topic [Born and Jordan, 1925a] *did not contain anything new apart from our simpler form of the calculation*, but by studying this topic we both came to a more intimate understanding of Bohr's leading ideas (Jordan, 1973, p. 294, our emphasis).[50]

Incidentally, Van Vleck (1971, p. 7) pointed to this important pre-1925 contribution of his own as well as to Slater's role in BKS and Kemble's work on helium to demonstrate the inaccuracy of Rabi's characterization of American work in quantum theory quoted earlier. Even at the time, Van Vleck had felt that the Europeans were not giving the Americans their due. He complained about this in a letter to Born:

> I am writing this letter regarding some of the references to my work in your articles. I fully realize that an occasional error in a reference is unavoidable, for I have made such mistakes myself. I would gladly overlook any one error, but inasmuch as there are two or three instances, it is perhaps worth while to call

---

[48] The multiplication of quantum–theoretical quantities corresponding to classical Fourier components is one of the key elements of Heisenberg's *Umdeutung* paper.

[49] See p. 24 of the transcript of session 1 of the AHQP interview with Van Vleck.

[50] See Sects. 5.2, 5.3 and 6.1 for discussion of Van Vleck's correspondence principles for emission and absorption. As in the case of (Kramers and Heisenberg, 1925), we suspect that (Born and Jordan, 1925a) is actually more difficult to follow for most modern readers than (Van Vleck, 1924b,c).

them to your attention. On p. 332 of your treatise on "Atommechanik" [Born, 1925], the reference to my work on the crossed-orbit model of the normal helium atom is given as [Van Vleck, 1923]. This reference is only to the abstract of some work on *excited* helium and the references to my articles on *normal helium* are [Van Vleck, 1922a]…and especially [Van Vleck, 1922b], where the details of the computations are given. This incorrect reference to a paper on another subject published a year later makes it appear as though my computation was published simultaneously or later than that of Kramer[s] [(Kramers, 1923), cited in the same footnote as (Van Vleck, 1923) in (Born, 1925, p. 332)]. The same error is also found in your article [Born, 1924b] on perturbation theory…Also in your book on Atommechanik [(Born, 1925, p. 332), the sentence with the footnote referring to (Kramers, 1923) and (Van Vleck, 1923)] you say "das raumliche [sic] Modell ist ebenfalls von Bohr vorgeschlagen" [the spatial model has also been proposed by Bohr], without any mention of the name Kemble, who proposed the crossed-orbit model in [Kemble, 1921] before [Bohr, 1922].[51]

Van Vleck then comes to the most egregious case, Born's failure to properly acknowledge his two-part paper on the correspondence principle in (Born and Jordan, 1925a). Especially in view of Jordan's comments on the importance of this paper quoted above, the authors were very stingy in giving him credit.

Van Vleck's letter continues:

I was much interested in your recent article on the Quantization of Aperiodic Systems, in which you show that the method of Fourier integrals gives many results obtained by "Niessen and Van Vleck" [Born and Jordan, 1925a, p. 486], placing my name after Niessen's [Kare Frederick Niessen (1895–1967)], even though his paper [Niessen, 1924] did not appear until Dec. 1924 while the details of my computations were given in the Physical Review for Oct. 1924 [Van Vleck, 1924b,c] and a preliminary notice published in the Journal of the Optical Society for July 1924 [Van Vleck, 1924a], before Niessen's article was even submitted for publication. I think you wrote me inquiring about my work shortly after the appearance of this preliminary note, and so you must be aware that it was the first to appear…inasmuch as Niessen's discussion is somewhat less general than my own, it seems to me that it scarcely merits being listed first (Ibid.).

Writing from Cambridge, Massachusetts, where he was visiting MIT, Born apologized.[52] Born had indeed written to Van Vleck concerning (Van Vleck, 1924a), albeit a little later than the latter remembered:

While we already came close to one another in the calculation of the helium atom, I see from your paper "A Correspondence Principle for Absorption" [Van Vleck, 1924a] that we now approach each other very closely with our trains of

[51] Van Vleck to Born, October 19, 1925, draft (AHQP).

[52] Born to Van Vleck, November 25, 1925 (AHQP). Born had been less generous in the case of a similar complaint from America a few years earlier (see Sect. 3.2).

thought…I am sending you my paper "On Quantum Mechanics" [Born, 1924], which pursues a goal similar to yours.[53]

This goes to show—Rabi's anecdotal evidence to the contrary notwithstanding—that at least some European physicists did keep up with theoretical work published in American journals, the *Journal of the Optical Society of America* in this case, even if they were not particularly generous acknowledging its importance in print.

3 Dispersion theory as the bridge between the old quantum theory and matrix mechanics

From the point of view of modern quantum mechanics, the old quantum theory of Bohr and Sommerfeld—especially in the hands of the latter and members of his Munich school—was largely an elaborate attempt at damage control. In classical physics the state of a physical system is represented by a point in the phase space spanned by a system's generalized coordinates and momenta (q_i, p_i). All its properties are represented by functions $f(q_i, p_i)$ defined on this phase space. In quantum mechanics the state of a system is represented by a ray in the Hilbert space associated with the system; its properties are represented by operators acting in this Hilbert space, i.e., by rules for *transitions* from one ray to another. In the old quantum theory, one bent over backward to retain classical phase space. Quantum conditions formulated in various ways in (Sommerfeld, 1915a), (Wilson, 1915), (Ishiwara, 1915), (Schwarzschild, 1916), and (Epstein, 1916) only restricted the allowed orbits of points in phase space. These conditions restricted the value of so-called action integrals for every degree of freedom of some multiply-periodic system to integer multiples of Planck's constant h,

$$\oint p_i dq_i = n_i h, \tag{1}$$

where the integral is extended over one period of the generalized coordinate q_i (there is no summation over i). This condition must be imposed in coordinates in which the so-called Hamilton–Jacobi equation for the system is separable.

Imposing such quantum conditions on classical phase space would not do in the end. As the picture of the interaction of matter and radiation in the old quantum theory already suggests, more drastic steps were required. In Bohr's theory the frequency $v_{i \to f}$ of the radiation emitted when an electron makes the transition from an initial state i to a final state f is given by the energy difference $E_i - E_f$ between the two states divided by h. Except in the limiting case of high quantum numbers, this radiation frequency differs sharply from the frequencies with which the electron traverses its quantized orbits in classical phase space before and after emission. This was widely recognized as the most radical aspect of the Bohr model. Erwin Schrödinger (1887–1961), for instance, opined in 1926 that this discrepancy between radiation frequency and orbital frequency

---

[53] Born to Van Vleck, October 24, 1924 (AHQP).

…seems to me, (and has indeed seemed to me since 1914), to be something so *monstrous*, that I should like to characterize the excitation of light in this way as really almost *inconceivable*.[54]

Imre Lakatos (1970) produces a lengthy quotation from an obituary of Planck by Born (1948), in which the same point is made more forcefully. It even repeats some of the language of Schrödinger's letter:

That within the atom certain quantized orbits…should play a special role, could well be granted; somewhat less easy to accept is the further assumption that the electrons moving on these curvilinear orbits…radiate no energy. But that the sharply defined frequency of an emitted light quantum should be different from the frequency of the emitting electron would be regarded by a theoretician who had grown up in the classical school as *monstrous* and *almost inconceivable* (Lakatos, 1970, pp. 150–151, our emphasis).

Unfortunately, this passage is nowhere to be found in (Born, 1948)!

One area of the old quantum theory in which the "monstrous" element became glaringly and unavoidably apparent was in the treatment of optical dispersion, the differential refraction of light of different colors. It was in this area that physicists most keenly felt the tension between orbital frequencies associated with individual states (the quantized electron orbits of the Bohr–Sommerfeld model) and radiation frequencies associated with *transitions* between such states. One of the key points of Heisenberg's *Umdeutung* paper was to formulate a new theory not in terms of properties of individual quantum states but in terms of quantities associated with transitions between states *without even attempting to specify the states themselves*.[55] What, above all, prepared the ground for this move, as we shall show in this section, was the development of a quantum theory of dispersion by Ladenburg, Reiche, Bohr, Kramers, and others. As Friedrich Hund (1896–1997) put it in his concise but rather cryptic history of quantum theory:

In 1924 the question of the *dispersion of light* came to the foreground. It brought new points of view, and *it paved the way for quantum mechanics* (Hund, 1984, p. 128).

By comparison, many of the other preoccupations of the old quantum theory, such as a detailed understanding of spectral lines, the Zeeman and Stark effects, and the extension of the Bohr–Sommerfeld model to multi-electron atoms (in particular, helium) mostly added to the overall confusion and did little to stimulate the shift to the new mode of thinking exemplified by the *Umdeutung* paper.[56]

The same is true—*pace* Roger Stuewer (1975)—for the broad acceptance of Einstein's 1905 light-quantum hypothesis following the discovery in late 1922 by

---

[54] Schrödinger to Lorentz, June 6, 1926 (M. Klein, 1967, p. 61).

[55] As Klaas Landsman (2007) emphasizes, "Heisenberg…identified the mathematical nature of the observables, whereas Schrödinger…found the description of the states" (p. 428).

[56] For detailed analyses of some of these bewildering developments, see, e.g., (Serwer, 1977) and (Forman, 1968, 1970).

Arthur H. Compton (1892–1962) of the effect soon to be named after him. What *was* crucial for the development of matrix mechanics were the *A* and *B* coefficients for emission and absorption *even though* they had been introduced in the context of a theory involving light quanta (Einstein, 1916a,b, 1917). Physicists working on dispersion theory were happy to use these coefficients but were just as happy to continue thinking of light as consisting of waves rather than particles. John Hendry (1981) makes the provocative claim that "since Sommerfeld was the only known convert to the light-quantum concept as a result of the Compton effect whose opinions were of any real historical importance, this places Stuewer's thesis on the importance of the effect in some doubt" (p. 197). It is our impression that the Compton effect *did* convince many physicists of the reality of light quanta, just as Stuewer says it did, but we agree with Hendry (1981, p. 6) that this made surprisingly little difference for the quantum revolution of 1925–1926.

3.1 The classical dispersion theory of Helmholtz, Lorentz and Drude

Optical dispersion can boast of a venerable history in the annals of science reaching back at least to Descartes' rainbow and Newton's prism. The old quantum theory was certainly not the first theory for which dispersion presented serious difficulties. Both proponents of Newtonian particle theories of light in the eighteenth century and proponents of wave theories of light in the nineteenth century struggled with dispersion (Cantor, 1983).

In a review article on wave optics for the *British Association for the Advancement of Science* published in 1886, Richard Tetley Glazebrook (1854–1935), a student of James Clerk Maxwell (1831–1879), divided the nineteenth century into three periods. During the first period, which lasted well into the 1860s, optical phenomena were explained purely in terms of properties of the luminiferous ether, the medium thought to carry light waves. Refraction and dispersion, for instance, were explained by assuming that some property of the ether inside transparent media is different from what it is outside. The dispersion theories of this period typically also depend on the distance between the molecules of the transparent medium with which the ether was supposed to co-exist, but there was no consideration of any dynamical interaction between the ether and the transparent medium. This changed in the second period, which began in the late 1860s. Theorists now began to account for refraction and dispersion in terms of waves in the ether setting harmonically bound particles inside transparent media oscillating. During the third period, which was just starting when Glazebrook wrote his review article and which would not bear fruit until the 1890s, the models proposed in the second period were reworked to reflect that it had meanwhile become clear that light is an electromagnetic wave and that the particles in matter with which they interact are charged particles, to be identified with electrons by the end of the 1890s.

Some of the better known physicists and mathematicians contributing to the theory of dispersion during the first period distinguished by Glazebrook were Augustin Jean Fresnel (1788–1827), James MacCullagh (1809–1847), and Augustin Louis Cauchy (1789–1857). What distinguished their theories from one another was to a large extent simply which property of the ether was made responsible for the different

behavior of light in different media. MacCullagh and Cauchy, who did their most important work on dispersion in the 1830s, assumed that the rigidity or the elasticity of the ether was the key variable (Glazebrook, 1886, p. 158, pp. 164–165). Many theorists, however, followed Fresnel's original idea that it was its density. Fresnel assumed that the index of refraction is proportional to the square of the ether density inside the transparent medium (ibid., p. 157). This view became popular even though it implied that a transparent medium contains different amounts of ether for different colors. The index of refraction, after all, must depend on frequency to account for dispersion. This also affected the optics of moving bodies. To account for the absence of any signs of motion of the earth with respect to the ether, Fresnel, in 1818, introduced the "drag" coefficient. A transparent medium with index of refraction n would carry along the ether inside of it with a fraction $f = 1 - 1/n^2$ of its velocity with respect to the ether. Although it was widely recognized that the drag coefficient was needed to account for the null results of numerous ether drift experiments, many physicists throughout the nineteenth century expressed strong reservations about the underlying physical mechanism proposed by Fresnel, since it implied that, because of dispersion, matter had to drag along different amounts of ether for different colors (Janssen and Stachel, 2004; Stachel, 2005).

Despite such conceptual difficulties and despite limited agreement with the experimental data, progress was made in the first half of the nineteenth century in understanding such phenomena as dispersion with, to use Glazebrook's terminology, "theories based solely on the elastic solid theory [of the ether]" (Glazebrook, 1886, p. 210), in which all optical phenomena in transparent matter are attributed to some modification of the properties of the ether inside. Concluding his discussion of such theories in his review article, Glazebrook wrote:

> while the elastic solid theory, taken strictly, fails to represent all the facts of experiment, we have learnt an immense amount by its development, and have been taught where to look for modifications and improvements. We may, I think, infer that the optical differences of bodies depend mainly on differences in the density or effective density of the ether in those bodies, and not on differences of rigidity (ibid. p. 211).

Glazebrook then turned to the second period and the second class of theories that he distinguished in his review article: "[t]heories based on the mutual reaction between ether and matter" (Glazebrook, 1886, Part III, pp. 212–251). In such theories, the ether is typically assumed to have the same properties everywhere and refraction and dispersion are explained in terms of momentum transfer between the ether and the molecules of ponderable matter. Before discussing various theories of this kind, Glazebrook explained why it was to be expected that a satisfactory theory for the behavior of light in transparent media calls for such a theory:

> The properties we have been considering depend on the presence of matter, and we have to deal with two systems of mutually interpenetrating particles. It is clearly a very rough approximation to suppose that the effect of the matter is merely to alter the rigidity or the density of the ether. The motion of the ether will be disturbed by the presence of the matter; motion may even be set up in

the matter particles. The forces to which this gives rise may, so far as they affect the ether, enter its equations in such a way as to be equivalent to a change in its density or rigidity, but they may, and probably will, in some cases do more than this. The matter motion will depend in great measure on the ratio which the period of the incident light bears to the free period of the matter particles. If this be nearly unity, most of the energy in the incident vibration will be absorbed in setting the matter into motion, and the solution will be modified accordingly (ibid. p. 212).

More than anything else, it was the phenomenon of anomalous dispersion that necessitated these more sophisticated theories. Anomalous dispersion was first noticed in 1840 by the early photographer William Henry Fox Talbot (1800–1877) but only recognized for what it really was in 1870 by the Danish physicist Christian Christiansen (1843–1917) (Buchwald, 1985, p. 233). Whereas in normal dispersion the angle of refraction increases (if we consider some fixed angle of incidence) with the frequency of the refracted light, in anomalous dispersion there are frequency intervals in which the angle of refraction decreases with increasing frequency. As Glazebrook emphasized, this phenomenon is inexplicable in the older class of theories.[57] Anomalous dispersion calls for a theory "based on the mutual reaction between ether and matter."

The first such theory appears to have been formulated in 1867 by Joseph Valentin Boussinesq (1842–1929) (Glazebrook, 1886, p. 213). Independently of Boussinesq, it seems, Wolfgang Sellmeier, a student of Franz Neumann (1798–1895), developed a similar theory and used it in 1872 to account for dispersion, including anomalous dispersion. Roughly, according to Sellmeier's theory, what happens when a light wave of a certain frequency hits a transparent medium is that it produces (additional) oscillations of harmonically bound particles in the medium. The result of this interaction, as Glazebrook points out in the passage quoted above, will depend on how close the frequency of the incoming light is to the resonance frequencies of these particles. The dispersion formula given by Sellmeier has a pole at the resonance frequencies. The same is true for the dispersion formula derived from electromagnetic theories later in the century (cf. Eqs. (6)–(7)). If these poles are in the ultraviolet, optical dispersion is normal (i.e., the angle of refraction increases with frequency throughout the optical spectrum); if, however, there are poles at optical frequencies, the angle of refraction decreases with frequency in the range immediately above them (Glazebrook, 1886, p. 219).

In 1875, Hermann von Helmholtz (1821–1894) proposed a highly influential dispersion theory in the spirit of Boussinesq and Sellmeier but substantially improving on their work (Buchwald, 1985, Ch. 27). Helmholtz's theory was based on "twin equations" for the coupled oscillations in the ether and in the transparent medium (ibid. p. 235; cf. Glazebrook 1886, p. 222). Helmholtz's theory, like Sellmeier's, is a purely mechanical one. Newtonian mechanics governs both ether and matter. In 1893, 7 years after Glazebrook's review article, Helmholtz adapted his theory to reflect that

---

[57] "The suggestions of Cauchy and [Charles Auguste] Briot [(1817–1882)]…lead to expressions for the relation between the refractive index and wave length which agree well with experiment so long as we steer clear of substances which present the phenomena of anomalous dispersion, but of this they give no account" (Glazebrook, 1886, p. 212; see also p. 217).

light is an electromagnetic wave and that such waves act on and are emitted by charged particles in matter (Buchwald, 1985, Sect. 27.2; Darrigol, 2000, Sect. 8.3). The year before and apparently unbeknownst to Helmholtz, Lorentz had already published such an electromagnetic dispersion theory, building on work he had done in the 1870s, which in turn built on Helmholtz's synthesis of British and continental ideas about electromagnetism.[58] Lorentz's theory of 1892 is similar but superior to Helmholtz's theory of 1893. For one thing, Lorentz immediately derived the Fresnel drag coefficient from his theory, whereas it was left to Richard August Reiff (1855–1927) to do so for Helmholtz's theory later in 1893 (Darrigol, 2000, p. 322). Moreover, while it is perfectly clear that Lorentz derived the drag coefficient without introducing any actual ether drag, this is not so clear in the case of Reiff (Buchwald, 1985, p. 241). Neither Helmholtz's theory nor Lorentz's left room for ether drag. In both theories, the ether is immobile, its properties are the same everywhere, and the index of refraction is related to the polarization of harmonically bound electric charges. These theories thus avoid the absurdity in Fresnel's original picture that matter drags along different amounts of ether for different frequencies.

Lorentz's theory constituted a much more radical move into microphysics than Helmholtz's and partly as a result of that, it seems, held less appeal for German physicists in the 1890s, although Helmholtz's greater authority in the German physics community may also have been a factor (Buchwald, 1985, pp. 238–241). The approach to optics based on an electron theory à la Helmholtz and Lorentz only became popular with the appearance of *Lehrbuch der Optik* (Drude, 1900), in which Paul Drude (1863–1906) presented and extended the theory.[59] The English translation of Drude's book in 1905 made the approach popular in Britain and the United States as well.

This classical electron theory of dispersion was remarkably successful in accounting for the experimental data. Hence, two centuries after Newton, there finally was a reasonably satisfactory theory for dispersion, including anomalous dispersion. Only two decades later, however, the model of matter underlying this theory was called into question again with the rise of the old quantum theory (Jammer, 1966, p. 189). The electrons oscillating inside atoms in the Helmholtz–Lorentz–Drude model were replaced by electrons orbiting the nucleus in the Rutherford–Bohr model. As we shall see, the classical electron theory of dispersion nonetheless played an important role in the development of a quantum theory of dispersion in the early 1920s.

---

[58] This earlier work by Lorentz is not mentioned in the brief discussion of electromagnetic theories in (Glazebrook, 1886, Part IV, pp. 251–261). For a detailed discussion of Lorentz's dispersion theory, see (Buchwald, 1985, Appendix 7).

[59] Olivier Darrigol (1992, p. 331) suggests that Drude converted to Lorentz's theory after the 1898 *Naturforscherversammlung* in Düsseldorf, where Lorentz was the guest of honor for a session on the problem of optics and electrodynamics in moving bodies. Jed Buchwald (1985, p. 250), however, points out that (Drude, 1900) only refers to Lorentz in the discussion of optics in moving bodies and suggests that Drude, like most German physicists, followed Helmholtz rather than Lorentz. Dispersion is covered in Pt. II, Sect. II, Ch. V of Drude's book.

   The basic model of dispersion in this classical theory is very simple.[60] Suppose an electromagnetic wave of frequency ν (we are not concerned with how and where this wave originated) strikes a charged one-dimensional simple harmonic oscillator with characteristic frequency ν_0. We focus on the case where the frequency ν of the electromagnetic wave is far from the resonance frequency ν_0 of the oscillator. We can picture the oscillator as a point particle with mass m and charge $-e$ (where e is the *absolute value* of the electron charge) on a spring with equilibrium position $x = 0$ and spring constant k, resulting in a restoring force $F = -kx$. The characteristic angular frequency $\omega_0 = 2\pi\nu_0$ is then given by $\sqrt{k/m}$. The electric field E of the incident electromagnetic wave[61] will induce an additional component of the motion at the imposed frequency ν. This component will be superimposed on any preexisting oscillations at the characteristic frequency ν_0 of the unperturbed system. It is this additional component of the particle motion, coherent with the incident wave (i.e., oscillating with frequency ν), that is responsible for the secondary radiation that gives rise to dispersion. The time dependence of this component is given by:

$$\Delta x_{\text{coh}}(t) = A \cos \omega t, \tag{2}$$

where $\omega = 2\pi\nu$. To determine the amplitude A, we substitute Eq. (2) into the equation of motion for the system. As long as we are far from resonance, radiation damping can be ignored and the equation of motion is simply:[62]

$$m\ddot{x} = -m\omega_0^2 x - eE \cos \omega t, \tag{3}$$

where dots indicate time derivatives and where we have made the innocuous simplifying assumption that the electric field of the incident wave is in the x-direction. Substituting $\Delta x_{\text{coh}}(t)$ in Eq. (2) for $x(t)$ in Eq. (3), we find:

$$-m\omega^2 A \cos \omega t = (-m\omega_0^2 A - eE) \cos \omega t. \tag{4}$$

It follows that

$$A = \frac{eE}{m(\omega^2 - \omega_0^2)}. \tag{5}$$

---

[60] The theory is covered elegantly in Chap. 31 of Vol. 1 of the Feynman lectures (see also Chap. 32 of Vol. 2). Feynman makes it clear that this classical theory remains relevant in modern physics: "we will assume that the atoms are little oscillators, that is that the electrons are fastened elastically to the atoms…You may think that this is a funny model of an atom if you have heard about electrons whirling around in orbits. But that is just an oversimplified picture. The correct picture of an atom, which is given by the theory of wave mechanics, says that, *so far as problems involving light are concerned*, the electrons behave as though they were held by springs" (Feynman et al., 1964, Vol. 1, Sect. 31-4).

[61] We need not worry about the effects of the magnetic field B. The velocity of electrons in typical atoms is of order αc, where c is the velocity of light and $\alpha \simeq 1/137$ is the fine-structure constant. The effects due to the magnetic field are thus a factor $1/137$ smaller than those due to the electric field and can be ignored in all situations considered in this paper.

[62] In Sect. 5.3, we show how to take into account the effects of radiation damping.

The central quantity in the classical dispersion theory is the dipole moment $p(t) \equiv -e\Delta x_{\text{coh}}(t)$ of the oscillator induced by the electric field of the incident electromagnetic wave. From Eqs. (2) and (5) it follows that:

$$p(t) = -e\Delta x_{\text{coh}}(t) = \frac{e^2 E}{4\pi^2 m (v_0^2 - v^2)} \cos 2\pi vt. \qquad (6)$$

For groups of n_i oscillators of characteristic frequencies v_i per unit volume, this formula for the dipole moment naturally generalizes to the following result for the polarization (i.e., the dipole moment per unit volume):

$$P(t) = \frac{e^2 E}{4\pi^2 m} \sum_i \frac{n_i}{v_i^2 - v^2} \cos 2\pi vt. \qquad (7)$$

The number of oscillators of characteristic frequency v_i will be some fraction f_i of the numbers of atoms in the volume under consideration. This fraction was often called the "oscillator strength" in the literature of the time. The polarization P determines the index of refraction n (see, e.g., Feynman et al., 1964, Vol. 1, 31-5). The agreement of Eq. (7) with the data from experiments on dispersion was not perfect, but dispersion was nonetheless seen as an important success for the classical theory.

3.2 The Sommerfeld–Debye theory and its critics

An early and influential attempt to bring dispersion theory under the umbrella of the old quantum theory was made by Sommerfeld (1915b, 1917) and by his former student Peter Debye (1884–1966) (Debye, 1915).[63] Clinton J. Davisson (1881–1958), then working at the Carnegie Institute of Technology in Pittsburgh, also contributed (Davisson, 1916).[64] The Sommerfeld–Debye theory, as it came to be known, was based on the dubious assumption that the secondary radiation coming from small

---

[63] For other historical discussions of the development of quantum dispersion theory, see, e.g., (Darrigol, 1992, pp. 224–230), (Dresden, 1987, pp. 146–159, pp. 215–222), (Jammer, 1966, p. 165 and Sect. 4.3, especially pp. 188–195), (Mehra and Rechenberg, 1982–2001, Vol. 1, Sect. VI.1; Vol. 2, Sect. III.5, pp. 170–190; Vol. 6, Sect. III.1 (b), pp. 348–353), and (Whittaker, 1953, Vol. 1, p. 401; Vol. 2, pp. 200–206). (Van Vleck, 1926, Sect. 49, pp. 156–159) briefly discusses the early attempts to formulate a quantum theory of dispersion in his review article on the old quantum theory. We focus on the theory of Sommerfeld and Debye of the late 1910s and on the theories developed by Ladenburg and Reiche and by Kramers in the early 1920s. Van Vleck also mentions theories of the latter period by Charles Galton Darwin (1887–1962), Adolf Gustav Smekal (1895–1959), and Karl F. Herzfeld (1892–1978). All three of these theories make use of light quanta. In addition, strict energy conservation is given up in the theory of Darwin (1922, 1923), while in the theories of Smekal (1923) and Herzfeld (1924) orbits other than those picked out by the Bohr–Sommerfeld condition are allowed, a feature known as "diffuse quantization." For other (near) contemporary reviews of dispersion theory, see (Pauli, 1926, pp. 86–96), (Andrade, 1927, pp. 669–682), and (Breit, 1932). (Stolzenburg, 1984, pp. 17–18) briefly discusses Bohr's critical reaction to Darwin's dispersion theory.

[64] In 1927 at Bell Labs, Davisson and his assistant Lester H. Germer (1896–1971) would do their celebrated work on electron diffraction (Davisson and Germer, 1927), another great American contribution to (experimental) quantum physics for which the authors received the 1937 Nobel Prize (Kevles, 1978, pp. 188–189).

perturbations of a Bohr orbit induced by incident radiation could be calculated on the basis of ordinary classical electrodynamics, even though, by the basic tenets of the Bohr model, the classical theory did *not* apply to the original unperturbed orbit. In other words, it was assumed that, while the large accelerations of electrons moving on Bohr orbits would produce no radiation whatsoever, the comparatively small accelerations involved in the slight deviations from these orbits caused by weak incident radiation would produce radiation.[65] Otherwise, the theory stayed close to the classical theory, substituting small deviations in the motion of electrons from their Bohr orbits for small deviations from the vibrations of simple harmonic oscillators at their characteristic frequencies.

Both the Swedish physicist Carl Wilhelm Oseen (1879–1944) and Bohr severely criticized the way in which Sommerfeld and Debye modeled their quantum dispersion theory on the classical theory. Oseen (1915) wrote: "Bohr's atom model can in no way be reconciled with the fundamental assumptions of Lorentz's electron theory. We have to make our choice between these two theories" (p. 405).[66] Bohr agreed. The central problem was that in Bohr's theory the link between radiation frequencies and orbital frequencies had been severed. As Bohr explained to Oseen in a letter of December 20, 1915, if the characteristic frequencies involved in dispersion

…are determined by the laws for quantum emission, the dispersion cannot, whatever its explanation, be calculated from the motion of the electrons and the usual electrodynamics, which does not have the slightest connection with the frequencies considered (Bohr, 1972–1996, Vol. 2, p. 337).

Bohr elaborated on his criticism of the Sommerfeld–Debye theory in a lengthy paper intended for publication in *Philosophical Magazine* in 1916 but withdrawn after it was already typeset.[67] Bohr argued (we leave out the specifics of the experiments on dispersion in various gases that Bohr mentions in this passage):

[E]xperiments…show that the dispersion…can be represented with a high degree of approximation by a simple Sellmeier formula in which the characteristic frequencies coincide with the frequencies of the lines in the…spectra…[T]hese frequencies correspond with transitions between the normal states of the atom…On this view we must consequently assume that the dispersion…depends on the same mechanism as the transition between different stationary states, and that it cannot be calculated by application of ordinary electrodynamics from the configuration and motions of the electrons in these states (Bohr, 1972–1996, Vol. 2, pp. 448–449).

In the next paragraph, Bohr added a prescient comment. Inverting the line of reasoning in the passage above that dispersion should depend on the same mechanism as transitions between states, he suggested that transitions between states, about which

---

[65] Sommerfeld (1915b, p. 502) realized that this assumption was problematic and tried (unconvincingly) to justify it.

[66] Quoted and discussed in (Bohr, 1972–1996, Vol. 2, p. 337)

[67] It can be found in (Bohr, 1972–1996, Vol. 2, pp. 433–461). For further discussion of Bohr's early views on dispersion, see (Heilbron and Kuhn, 1969, pp. 281–283).

the Bohr theory famously says nothing, should depend on the same mechanism as dispersion: "[i]f the above view is correct…we must, on the other hand, assume that this mechanism [of transitions between states] shows a close analogy to an ordinary electrodynamic vibrator" (ibid).

As we shall see, in the quantum dispersion theory of the 1920s, the oscillators of the classical theory were grafted onto the Bohr model. For the time being, however, it was unclear how to arrive at a satisfactory quantum theory of dispersion. The quasi-classical Sommerfeld–Debye theory led to a formula for the induced polarization of the form of Eq. (7) but with resonance poles at the orbital frequencies. As Oseen and Bohr pointed out, this was in blatant contradiction with the experimental data, which clearly indicated that the poles should be at the radiation frequencies, which in Bohr's theory differed sharply from the orbital frequencies.

This criticism is repeated in more sophisticated form in a paper by Paul Sophus Epstein (1883–1966) with the subtitle "Critical comments on dispersion." This paper is the concluding installment of a trilogy on the application of classical perturbation theory to problems in the old quantum theory (Epstein, 1922a,b,c). Epstein, a Russian Jew who studied with Sommerfeld in Munich, was the first European quantum theorist to be lured to America. In 1921 Millikan brought him to the California Institute of Technology in Pasadena, despite prevailing anti-Semitic attitudes (Kevles, 1978, pp. 211–212).[68] In his 1926 review article Van Vleck emphasizes the importance of the work of his colleague at Caltech and notes that it "is rather too often overlooked" (Van Vleck, 1926, p. 164, note 268), to which one might add: "by European physicists." As we saw in Sect. 2.4, Van Vleck felt the same way about his own contributions. Like Van Vleck, Epstein apparently complained about this lack of recognition to Born. This can be inferred from a letter from Born to Sommerfeld of January 5, 1923, shortly before a visit of the latter to the United States:

> When you talk to Epstein in Pasadena and he complains about me, tell him that he should show you the very unfriendly letter he wrote to me because he felt that his right as first-born had been compromised by the paper on perturbation theory by Pauli and me [Born and Pauli, 1922, which appeared shortly after Epstein's trilogy]. Also tell him that I do not answer such letters but that I do not hold a grudge against him because of his impoliteness (to put it mildly)…In terms of perturbative quantization we are ahead of him anyway (Sommerfeld, 2004, p. 137).[69]

To deal with the kind of multiply-periodic systems that represent hydrogenic atoms (i.e., atoms with only one valence electron) in the old quantum theory, Epstein customized techniques developed in celestial mechanics for computing the perturbations of the orbits of the inner planets due to the gravitational pull of the outer ones.[70] The

---

[68] For further discussion of Epstein's position at Caltech, see (Seidel, 1978, pp. 507–520).

[69] This letter is quoted and discussed in (Eckert, 1993, p. 96).

[70] One of the sources cited by Epstein (1922a, p. 216) is (Charlier, 1902–1907). This source is also cited in (Bohr, 1918, p. 114), (Kramers, 1919, p. 8), and (Born and Pauli, 1922, p. 154). In their interviews for the AHQP, both Van Vleck (p. 14 of the transcript of session 1) and Heisenberg (p. 24 of the transcript of session 5) mention that they studied Charlier as well.

perihelion advance of Mercury due to such perturbations, for instance, is more than ten times the well-known 43″ per century due to the gravitational field of the sun as given by general relativity. Such calculations in classical mechanics are also the starting point of the later more successful approach to dispersion theory by Kramers and Van Vleck. Epstein clearly recognized that these calculations by themselves do not lead to a satisfactory theory of dispersion. In the introduction of his paper, Epstein (1922c, p. 92) explains that he discusses dispersion mainly because it nicely illustrates some of the techniques developed in the first two parts of his trilogy. He warns the reader that he is essentially following the Sommerfeld–Debye theory, and emphasizes that "this point of view leads to internal contradictions so strong that I consider the Debye–Davysson [sic] dispersion theory [as Epstein in Pasadena referred to it] to be untenable" (ibid.). The central problem is once again the discrepancy between radiation frequencies and orbital frequencies. As Epstein wrote in the conclusion of his paper:

> the positions of maximal dispersion and absorption [in the formula he derived] do not lie at the position of the emission lines of hydrogen but at the position of the mechanical frequencies of the model…*the conclusion seems unavoidable to us that the foundations of the Debye-Davysson* [sic] *theory are incorrect* (Epstein, 1922c, pp. 107–108).

Epstein recognized that a fundamentally new approach was required: "We believe that…dispersion theory must be put on a whole new basis, in which one takes the Bohr frequency condition into account from the very beginning" (ibid. p. 110).[71]

3.3 Dispersion in Breslau: Ladenburg and Reiche

Unbeknownst to Epstein, quantum dispersion theory had already begun to emerge from the impasse he called attention to in 1922. The year before, Ladenburg had introduced one of two key ingredients needed for a satisfactory treatment of dispersion in the old quantum theory: the emission and absorption coefficients of Einstein's quantum theory of radiation. The other critical ingredient, as we shall see below, was Bohr's correspondence principle.

Ladenburg spent most of his career doing experiments on dispersion in gases. He started in 1908, about 2 years after he joined the physics department, then headed by Otto Lummer (1860–1925), at the University of Breslau, his hometown (Ladenburg, 1908).[72] He stayed in Breslau until 1924, when he accepted a position at the *Kaiser Wilhelm Institut* in Berlin. There he continued his work with the help of students such as Hans Kopfermann (1895–1963), Agathe Carst, S. Levy, and G. Wolfsohn.

---

[71] Epstein had already voiced this criticism before he left for the United States. From Zurich, he had written to Einstein on October 15, 1919: "Meanwhile, I have carried out the calculations for dispersion theory from the point of view of quantum theory that I mentioned at one point in conversation: the result is definitely that the Debye–Sommerfeld theory only has the status of an approximation and that the true theory must take into account the [Bohr] frequency condition. It is not surprising, therefore, that Sommerfeld's results are occasionally off" (Einstein, 1987–2006, Vol. 9, Doc. 136).

[72] See the entry on Ladenburg by A. G. Shenstone (1973) in the *Dictionary of Scientific Biography*.

Ladenburg and his group reported the results of their experiments on dispersion in a series of papers published between 1926 and 1934.[73] Ladenburg's direct involvement ceased with his emigration to the United States in 1931.

Ladenburg and Stanislaw Loria (1883–1958) had established early on that the frequency of the H_α line in the Balmer series in the hydrogen spectrum corresponds to a pole in the classical dispersion formula (Ladenburg and Loria, 1908, p. 866). Given that the Sommerfeld–Debye theory flies in the face of this experimental fact, Ladenburg was never attracted to that theory. He simply kept using a dispersion formula with poles at the observed radiation frequencies. He focused on the numerator rather than the denominator of the dispersion formula. This is made particularly clear in the AHQP interviews with two of his collaborators in the early 1920s—Rudolph Minkowski (1895–1976), a nephew of Hermann Minkowski, who took his doctorate under Ladenburg in 1921 and co-authored (Ladenburg and Minkowski, 1921); and Fritz Reiche, who was appointed in Breslau in 1921.[74] After his doctorate (with Planck) in Berlin in 1907, Reiche had already spent three years in Breslau. He and Ladenburg had become close friends. Reiche had gone back to Berlin in 1911. When he returned to Breslau ten years later, he stayed until he was dismissed in 1933.[75] Reiche's help is prominently acknowledged in (Ladenburg, 1921, p. 140, note). Ladenburg was first and foremost an experimentalist and he welcomed input from his theoretician friend and colleague.[76] The two of them co-authored a pair of follow-up papers (Ladenburg and Reiche, 1923, 1924). Discussing the first of these, Reiche told Kuhn and Uhlenbeck in 1962:

> we did not derive a consistent dispersion theory, in which instead of the revolution numbers the emitted lines came out. We thought it completely self-evident, that one had to change the denominator of the dispersion formula in such a way that the frequencies were the emitted line frequencies, and not something which has to do with (the orbit) [sic].[77]

Reiche made it clear that he and Ladenburg were concerned only with explaining "the N which is on top of the dispersion formula:"

> It never came out correctly equal to the number of atoms, or to the number of atoms multiplied by the number of electrons in an atom. It gave, under certain conditions, even numbers which are less than the whole number of atoms. They were written very often with a German N…This was the main aim of the whole thing [Ladenburg and Reiche, 1923]. There, based on a previous paper by Ladenburg [1921], we found a relation between the German N and the real

---

[73] See (Mehra and Rechenberg, 1982–2001, Vol. 6, Ch. 3(b), pp. 348–353) and (Shenstone, 1973, p. 555) for detailed references and brief discussions.

[74] The following information is based on an autobiographical statement by Reiche published as an appendix to (Bederson, 2005).

[75] It was not until 1941 that he finally managed to emigrate to the United States.

[76] Asked by Kuhn whether Ladenburg was "strictly an experimentalist," Reiche said: "He was, as far as I understand, a very good experimental man, but he was one of the men who could make, let me say, easy theoretical work" (p. 10 of the transcript of the last of three sessions of the interview).

[77] See p. 11 of the transcript of the second of three sessions of the AHQP interview with Reiche.

number of atoms. The f were not 1 or 2 or 3 or something like this, but could be point 5 or the like. And the explanation of this was the aim of this dispersion paper. But it did not come out that we had a correct and consistent theory in which the denominator gave now the emitted frequencies. This, I think, was only done by Kramers [1924a, b], first of all.[78]

Ladenburg's dispersion experiments had indicated all along that the oscillator strength f_i, the number of dispersion electrons with characteristic frequency ν_i per atom, was not on the order of unity, as one would expect on the basis of the classical theory, but much smaller. For the frequency ν_i corresponding to the H_α line in the Balmer series in the hydrogen spectrum, for instance, Ladenburg and Loria (1908, p. 865) found that there was only 1 dispersion electron per 50,000 molecules, and they cited findings of 1 dispersion electron per 200 molecules in sodium vapor. Such low values were quite inexplicable on classical grounds. In the Bohr model the H_α (absorption) line corresponds to a transition from the $n = 2$ to the $n = 3$ state of the hydrogen atom. That Ladenburg found such a low value for what he interpreted classically as the number of dispersion electrons at the frequency of the H_α line is explained in Bohr's theory simply by noting that only a tiny fraction of the atoms will be in the $n = 2$ state (Ladenburg, 1921, p. 156). Ladenburg's key contribution was that he recognized that the oscillator strengths corresponding to various transitions could all be interpreted in terms of transition probabilities, given by Einstein's A and B coefficients. Hence the title of his paper: "The quantum-theoretical interpretation of the number of dispersion electrons" (Ladenburg, 1921).

Ladenburg obtained a relation between the oscillator strengths and the A and B coefficients by equating results derived for what would seem to be two mutually exclusive models of matter, a classical and a quantum model. He calculated the energy absorption rate both for a collection of classical oscillators à la Helmholtz, Lorentz and Drude, resonating at the absorption frequencies, and for a collection of atoms à la Bohr and Einstein with transitions between discrete energy levels corresponding to these same frequencies. Ladenburg set the two absorption rates equal to one another. His paper only gives the resulting expression for the numerator of the dispersion formula. Combining Ladenburg's theoretical relation between classical oscillator strengths and quantum transition probabilities with his experimental evidence that the resonance poles should be at the radiation frequencies, we arrive at the following formula—in our notation, based on (Van Vleck, 1924b)—for the induced polarization of a group of N_r atomic systems in their ground state r

$$P_r(t) = \frac{N_r c^3 E}{32\pi^4} \sum_s \frac{A_{s\to r}}{\nu_{s\to r}^2 (\nu_{s\to r}^2 - \nu^2)} \cos 2\pi \nu t, \tag{8}$$

where $\nu_{s\to r}$ is the frequency for a transition from the excited states s to r and $A_{s\to r}$ is Einstein's emission coefficient for this transition.

---

[78] Ibid. Dispersion is discussed at greater length during the third session of the interview (see pp. 10–14 of the transcript).

Ladenburg's paper initially did not attract much attention. It is not mentioned in Epstein's trilogy the following year, but then Epstein was working in faraway California. More surprisingly, quantum physicists in Göttingen, Munich, and Copenhagen, it seems, also failed to take notice, even though Ladenburg was well-known to his Göttingen colleagues Born and James Franck (1882–1964). Ladenburg had actually prevented that Born, a fellow Breslau native, was sent to the trenches in World War I. Ladenburg had recruited Born for an army unit under his command in Berlin devoted to artillery research (Thorndike Greenspan, 2005, pp. 71–72). Bohr and Ladenburg also knew each other personally: Ladenburg had attended Bohr's colloquium in Berlin in April 1920 and the two men had exchanged a few letters since (Bohr, 1972–1996, Vol. 4, pp. 709–717).

Heisenberg later attributed the neglect of Ladenburg in Göttingen and Munich to the problem of connecting Ladenburg's work, closely tied to Einstein's radiation theory, to the dominant Bohr–Sommerfeld theory.[79] According to Heisenberg, it was only when Kramers (1924a,b) rederived Ladenburg's formula as a special case of his own more general dispersion formula that its significance was widely appreciated.[80] Ladenburg's own derivation had been unconvincing, at least to most physicists.[81] In addition to just assuming the poles in the dispersion formula to be at the radiation frequencies rather than at the orbital frequencies, Ladenburg offered no justification for equating classical and quantum energy absorption rates. Van der Waerden (1968, p. 10) suggests that Ladenburg appealed to Bohr's correspondence principle in his derivation of the relation between oscillator strengths and A and B coefficients, but the correspondence principle is not mentioned anywhere in Ladenburg's paper. The full dispersion formula (8)—admittedly only implicit in Ladenburg's paper but associated with it, not just by later historians but also by his contemporaries—can certainly not be derived with the help of the correspondence principle, since it only holds for atoms in their ground state and *not* for atoms in highly excited states where classical and quantum theory may be expected to merge in the sense of the correspondence principle. Still, if Heisenberg's later recollections are to be trusted, it might have helped the reception of Ladenburg's paper had he made some reference to the correspondence principle.

Unlike his colleagues in Göttingen and Munich, Bohr did take notice of Ladenburg's paper early on. He was just slow, as usual, to express himself about it in print. As noted in (Hendry, 1981, p. 192), Bohr referred to (Ladenburg, 1921) in the very last sentence of a manuscript he did not date but probably started and abandoned in 1921 (Bohr,

---

[79] See p. 8 of the transcript of session 4 of the AHQP interview with Heisenberg, parts of which can be found in (Mehra and Rechenberg, 1982–2001, Vol. 2, pp. 175–176), although the authors cite their own conversations with Heisenberg as their source (cf. note 5).

[80] Jordan had the same impression (see pp. 24–25 of the transcript of the first session of Kuhn's interview with Jordan for the AHQP in June 1963). It also fits with Born's recollections. In his autobiography, Born (1978) notes: "An important step was made by my old friend from Breslau…Ladenburg…A detailed account was given by Ladenburg and Reiche, my other old friend from Breslau…On the basis of these investigations, Kramers…succeeded in developing a complete 'dispersion formula'" (pp. 215–216).

[81] As Kuhn put it in his AHQP interview with Slater: "Of course, there was a good deal that appeared to most physicists as pretty totally ad hoc about the Reiche–Ladenburg work, and the whole question as to why it was the transition frequencies that occurred in the denominator rather than the orbital frequencies." Slater disagreed: "This seemed to me perfectly obvious…" (p. 41 of the transcript of the first session of the interview).

1972–1996, Vol. 3, pp. 397–414). In a paper submitted in November 1922, Bohr (1923b, p. 162) finally discussed Ladenburg's work in print. After repeating some of the observations about dispersion in the passages of his unpublished 1916 paper quoted in Sect. 3.2, Bohr, in his tortuous verbose style, made some highly interesting remarks that anticipate aspects of the BKS theory of 1924 (see Sect. 4):

> the phenomena of dispersion must thus be so conceived that the reaction of the atom on being subjected to radiation is closely connected with the unknown mechanism which is answerable [the German has *verantwortlich*: responsible] for the emission of the radiation on the transition between stationary states. In order to take account of the observations, it must be assumed that this mechanism… becomes active when the atom is illuminated in such a way that the total reaction of a number of atoms is the same as that of a number of harmonic oscillators in the classical theory,[82] the frequencies of which are equal to those of the radiation emitted by the atom in the possible processes of transition, and the relative number of which is determined by the probability of occurrence of such processes of transition under the influence of illumination. A train of thought of this kind was first followed out closely in a work by Ladenburg [1921] in which he has tried, in a very interesting and promising manner, to set up a direct connection between the quantities which are important for a quantitative description of the phenomena of dispersion according to the classical theory and the coefficients of probability appearing in the deduction of the law of temperature radiation by Einstein (Bohr, 1972–1996, Vol. 3, p. 496).

A letter from Bohr to Ladenburg of May 17, 1923 offers further insights into Bohr's developing views on the mechanism of radiation:

> to interpret the actual observations, it…seems necessary to me that the quantum jumps are not the direct cause of the absorption of radiation, but that they represent an effect which accompanies the continuously dispersing (and absorbing) effect of the atom on the radiation, even though we cannot account in detail for the quantitative relation [between these two effects] with the usual concepts of physics (Bohr, 1972–1996, Vol. 5, p. 400).

At the beginning of this letter, Bohr mentioned the vagueness of some of his earlier pronouncements on the topic. After the passage just quoted he acknowledged "that these comments are not far behind the earlier ones in terms of vagueness. I do of course reckon with the possibility that I am on the wrong track but, if my view contains even a kernel of truth, then it lies in the nature of the matter that the demand for clarity in the current state of the theory is not easily met" (ibid.). Bohr need not have been so apologetic. His comments proved to be an inspiration to Ladenburg and Reiche. On June 14, 1923, Ladenburg wrote to Bohr:

> Over the last few months Reiche and I have often discussed [the absorption and scattering of radiation] following up on your comments in [Bohr 1923b] about

---

[82] Note the similarity between Bohr's description here to Feynman's observation (quoted in note 60) that atoms behave like oscillators "so far as problems involving light are concerned."

reflection and dispersion phenomena and on my own considerations [Ladenburg 1921] which you were kind enough to mention there (Bohr, 1972–1996, Vol. 5, pp. 400–401).

In this same letter, Ladenburg announced his forthcoming paper with Reiche in a special issue of *Die Naturwissenschaften* to mark the tenth anniversary of Bohr's atomic theory. In the conclusion of this paper, they wrote:[83]

> Surveying the whole area of scattering and dispersion discussed here, we have to admit that we do not know the real [*eigentlich*] mechanism through which an incident wave acts on the atoms and that we cannot describe the reaction of the atom in detail. This is no different by the way in the case of the real [*eigentlich*] quantum process, be it that an external wave ν_0 lifts electrons into higher quantum states, or be it that a wave ν_0 is sent out upon the return to lower states. We nevertheless believe on the grounds of the observed phenomena that the end result of a process in which a wave of frequency ν acts upon the atom should not be seen as fundamentally different from the effect that such a wave exerts on classical oscillators (Ladenburg and Reiche, 1923, p. 597).

Ladenburg and Reiche (1923, p. 588, p. 590) introduced the term "substitute oscillators" [*Ersatzoszillatoren*] for such classical oscillators representing the atom as far as its interaction with radiation is concerned. They credited Bohr with the basic idea.[84] As we shall see in Sect. 4, these substitute oscillators became the virtual oscillators of BKS. Ladenburg and Reiche (1924, p. 672) themselves noted the following year that substitute oscillators were now called virtual oscillators (Konno, 1993, p. 141). The Berlin physicist Richard Becker (1887–1955) likewise noted in a paper written in the context of BKS the following year: "these virtual oscillators are substantially identical with the 'substitute oscillators' already introduced by Ladenburg and Reiche" (Becker 1924, p. 174, note 2).[85] That same year, Herzfeld (1924, p. 350) still used the term 'substitute oscillators,' citing (Ladenburg and Reiche, 1923). The term can also be found, without attribution, in the famous paper by Born and Jordan (1925b, p. 884) on matrix mechanics.[86]

Unlike Ladenburg in 1921, Ladenburg and Reiche prominently mentioned both Bohr's atomic theory and the correspondence principle in their 1923 paper. The authors' understanding and use of the correspondence principle, however, are still tied strongly to Einstein's quantum theory of radiation. Their "correspondence" arguments apply not to individual quantum systems, for which Bohr's correspondence principle was formulated, but to collections of such systems in thermal equilibrium with the

---

[83] Quoted and discussed in (Hendry, 1981, p. 192).

[84] See also (Ladenburg and Reiche, 1924, p. 672). Van Vleck (1926, p. 159, note 260) reports that Lorentz made a similar suggestion at the third Solvay congress in 1921 (Verschaffelt et al., 1923, p. 24), but does not mention Ladenburg and Reiche in this context, attributing the idea to (Slater, 1924) instead.

[85] Quoted in (Konno, 1993, p. 141).

[86] We are grateful to Jürgen Ehlers for drawing our attention to this passage, which is not in the part of (Born and Jordan, 1925b) included in (Van der Waerden, 1968).

ambient radiation.[87] The authors also do not limit their "correspondence" arguments to the regime of high quantum numbers (Ladenburg and Reiche, 1923, especially Sect. 4–5, pp. 586–589). These problems invalidate many of the results purportedly derived from the correspondence principle in their paper. Drawing on earlier work by Planck, they derived a result for emission consistent with the correspondence principle (i.e., merging with the classical result in the limit of high quantum numbers), but their attempts to derive similar results for absorption and dispersion were unconvincing. In fact, it may well be that these dubious attempts inspired Van Vleck to formulate correspondence principles for emission and absorption himself (see Sect. 5.3 for further discussion).

3.4 The Kramers dispersion formula

Given Bohr's strong interest in the subject, it is not surprising that his first lieutenant Kramers took the next big step in quantum dispersion theory.[88] Formula (8) based on Ladenburg's insights only holds for systems in the ground state. The correspondence principle only applies to highly excited states. Kramers (1924a,b) found that the correspondence principle requires a formula with *two* terms.[89] In our notation—which once again follows (Van Vleck, 1924b, p. 344, Eq. 17)—the induced polarization P_r of N_r atoms in a state labeled by the quantum number r is given by:

$$P_r(t) = \frac{N_r c^3 E}{32\pi^4} \left(\sum_{s>r} \frac{A_{s\to r}}{\nu_{s\to r}^2 (\nu_{s\to r}^2 - \nu^2)} - \sum_{t<r} \frac{A_{r\to t}}{\nu_{r\to t}^2 (\nu_{r\to t}^2 - \nu^2)} \right) \cos 2\pi \nu t, \quad (9)$$

where s and t are the quantum numbers labeling states above and below r, respectively (see Sects. 5.1–5.2 and 6.2 for detailed derivations). For high values of r this formula merges with the classical result. In the spirit of the correspondence principle, Kramers took the leap of faith that it holds all the way down to low quantum numbers. If r is the ground state, the second term vanishes and the Kramers formula (9) reduces to the Ladenburg formula (8). Like Ladenburg and Reiche (1923), Kramers interpreted his formula in terms of oscillators, distinguishing between "absorption oscillators" for the first term and "emission oscillators" for the second term (Kramers, 1924a, pp. 179–180). Kramers introduced the characteristic times $\tau_{i\to f}$ inversely proportional to $(e^2/m)\nu_{i\to f}^2$. So instead of factors $\nu_{i\to f}^2$ in the denominators in the two terms in Eq. (9), the formula given by Kramers (1924a, p. 179, Eq. 5) has factors $(e^2/m)\tau_{i\to f}$ in

---

[87] That Ladenburg and Reiche did not carefully distinguish between individual systems and collections of such systems becomes more understandable if we bear in mind that they were trying to combine Einstein's quantum theory of radiation and Bohr's correspondence principle. These two elements belong to two different strands in the development of quantum physics, characterized as follows in a concise and perceptive overview of the early history of quantum physics: "The first approach, dominated by the Berlin physicists Einstein, Planck, and Nernst, and by…Ehrenfest…involved the thermodynamics properties of matter and the nature of radiation…The other trend, centered socially in Copenhagen, Munich and Göttingen, consisted of the application of the quantum to individual atoms and molecules" (Darrigol, 2002, p. 336).

[88] Hendry (1984, p. 46) goes as far as calling Kramers' theory "the Bohr–Kramers dispersion theory."

[89] In addition to the literature cited in note 63, see (Ter Haar, 1998, pp. 23–30) and, especially, (Konno, 1993) for discussion of Kramers' work on dispersion theory.

the numerators.[90] Because of the minus sign in front of the second term, the emission oscillators appear to have negative mass, which is why Kramers also called them "negative oscillators" (ibid.). Van Vleck (1924a, p. 30, note 2) gave a more satisfactory interpretation of this minus sign, interpreting Kramers' formula for dispersion the same way as a formula for absorption he himself had proposed on the basis of Einstein's quantum radiation theory, as giving the net dispersion in a given quantum state as the difference between contributions from transitions to higher and transitions to lower states.

Kramers initially only published two notes in *Nature* on his new dispersion formula (Kramers, 1924a,b). Since these were submitted *after* (Bohr, Kramers, and Slater, 1924a), he used the new BKS terminology of 'virtual oscillators' in both of them. As we shall see in Sect. 4, this caused considerable confusion, both at the time and in the historical literature, about the relation between BKS and dispersion theory. Kramers' notes, moreover, are short on detail. The first, submitted on March 25, contains only the briefest of hints as to how the new dispersion formula had been found. The second, submitted on July 22 in response to a letter by Minnesota's Gregory Breit (1924b), contains at least an outline of the derivation. Kramers did not get around to publishing the derivation in full until his paper with Heisenberg, completed over the Christmas break of 1924, received by *Zeitschrift für Physik* on January 5, 1925, and published two months later (Mehra and Rechenberg, 1982–2001, Vol. 2, p. 181). According to Slater, however, the basic results had been in place by the time he, Slater, arrived in Copenhagen in December 1923. After dissing Bohr in the letter to Van Vleck quoted in Sect. 2.2, Slater goes on to say that

> Kramers hasn't got much done, either. You perhaps noticed his letter to Nature on dispersion [Kramers, 1924a]; the formulas & that he had before I came, although he didn't see the exact application; and except for that he hasn't done anything, so far as I know. They seem to have too much administrative work to do. Even at that, I don't see what they do all the time. Bohr hasn't been teaching at all, Kramers has been giving one or two courses.[91]

Part of what kept Kramers from his work in early 1924, as can be gathered from correspondence with Ladenburg and Reiche, was that his wife had fallen ill. In 1923, the Breslau physicists had already exchanged a few letters about dispersion with their colleague in Copenhagen.[92] On February 28, 1924, Ladenburg gently reminded Kramers that he had promised in January to give his "opinion on dispersion and its quantum interpretation"[93] within a few days. A little over a month later, on April 2, Ladenburg

---

[90] The polarization given by Kramers' formula is three times the polarization given by Van Vleck (i.e., by our Eq. (9)). This is because Kramers assumed that the vibrations in the atom are lined up with the electric field, whereas Van Vleck assumed the relative orientation of vibrations and fields to be random (Van Vleck, 1924b, p. 344, note 25).

[91] Slater to Van Vleck, July 27, 1924 (AHQP).

[92] See Reiche to Kramers, May 9, 1923 and December 28, 1923, and Ladenburg to Kramers, December 28, 1923 (AHQP). Kramers' responses, it seems, are no longer extant.

[93] Ladenburg to Kramers, February 28, 1924 (AHQP). This fits with Slater's recollection that Kramers already had his new dispersion formula around Christmas 1923.

wrote another letter to Kramers, in which he thanked him for sending what must have been either a manuscript or proofs of (Kramers, 1924a) (which only appeared in the May 10 issue of *Nature*) and, apparently having been informed by Kramers that the delay had been due to his wife's illness, apologized for his impatience.[94]

Understandably, given the importance of their own work for Kramers' break-through, Ladenburg and Reiche were enthusiastic about the new dispersion formula. Immediately after the one sentence devoted to the illness of Kramers' wife, without so much as starting a new paragraph, Ladenburg wrote in his letter of April 2:

> Now your opinion about the dispersion question is of course of the highest interest and I don't want to pass up the opportunity to tell you how much it pleases me that you have managed to give a correspondence derivation of the relation between dispersion and transition probabilities. In this way a solid basis has now been created. Your formula…is undoubtedly preferable to ours because of its greater generality. I also agree with you that one cannot extract contributions of the "negative" oscillators from existing experiments.[95]

Ladenburg thus immediately zeroed in on the key experimental question raised by the new formula. In the late 1920s Ladenburg and his collaborators embarked on a ambitious program to verify the second term in the Kramers dispersion formula experimentally. Reiche, writing to Kramers a week later, focused on the theoretical justification of the new formula:

> I wanted to tell you again how delighted I am with your beautiful correspondence derivation. Following Epstein's paper [Epstein, 1922c] and using the Born-Pauli [1922] method, I easily derived the classical expression for P [the polarization] which you indicate in your letter[96] and have also had no trouble reconstructing the correspondence argument for the transition to the quantum formula.[97]

Fearing that few Germans would have access to *Nature*, Ladenburg and Reiche prepa-red a detailed report on (Kramers, 1924a) for *Die Naturwissenschaften*. In late May, Ladenburg asked Kramers whether he would have any objections if they included a derivation of the new dispersion formula, adding that they were not sure how close it was to Kramers' derivation.[98] Kramers welcomed the idea, telling Ladenburg that their derivation would probably not be all that different from his own. He had every intention of writing a longer paper on dispersion and absorption himself, he added, which would obviously include the derivation of his dispersion formula, but recognized that "it will

---

[94] Ladenburg to Kramers, April 2, 1924 (AHQP). Reiche likewise apologized seven days later (Reiche to Kramers, April 9, 1924 [AHQP]).

[95] Ladenburg to Kramers, April 2, 1924 (AHQP). Ladenburg was not familiar with the BKS paper at this point, neither with the English version which appeared in April 1924, nor with the German translation which only appeared on May 22.

[96] This expression—equivalent to Eq. (14)—is not given in (Kramers, 1924a) but does occur in (Kramers, 1924b, p. 199, Eq. 2*) (reproduced as Eq. (50) in Sect. 5.1).

[97] Reiche to Kramers, April 9, 1924 (AHQP).

[98] Ladenburg to Kramers, 31 May 1924 (AHQP). Ladenburg and Reiche had meanwhile read (Bohr, Kramers, and Slater, 1924b) and, unsurprisingly given the importance of their concept of 'substitute oscillators' for BKS, were instant converts to the theory. For further discussion, see Sect. 4.2.

probably be a while before I have time to write such an article; because of lack of time I have not thought through many details and I consequently would not mind it at all if your note appears first."[99] In the end, the editor of *Die Naturwissenschaften* insisted that Ladenburg and Reiche shorten their article.[100] It eventually appeared without the derivation of Kramers' dispersion formula (Ladenburg and Reiche, 1924).

The first ones to publish a full derivation of this important result were Born and Van Vleck. (Born, 1924) was received by *Zeitschrift für Physik* on June 13, 1924, and was published in August; the two-part paper (Van Vleck, 1924b,c) was signed June 19, 1924 and appeared in *The Physical Review* in October. They thus arrived at their results independently.[101] Van Vleck (1924a) read Kramers' first *Nature* note shortly after he finished his first paper on a correspondence principle for absorption and when he was about to submit (Van Vleck, 1924b,c). In a footnote added to (Van Vleck, 1924a), he wrote:

Since the writing of the present article, Dr. H. A. Kramers has published…a very interesting formula for dispersion, in which the polarization is imagined as coming not from actual orbits, but from "virtual oscillators" such as have been suggested by Slater and advocated by Bohr. Kramers states that his formula merges asymptotically [i.e., in the limit of high quantum numbers] into the classical dispersion. To verify this in the general case, the writer has computed the classical polarization formula for an arbitrary non-degenerate multiply periodic orbit …By pairing together positive and negative terms in the Kramers formula, a differential dispersion may be defined resembling the differential absorption of the present article. It is found that this differential quantum theory dispersion approaches asymptotically the classical dispersion…the behavior being very similar to that in the correspondence principle for absorption. This must be regarded as an important argument for the Kramers formula (Van Vleck, 1924a, p. 30).

It was not clear to Van Vleck on the basis of Kramers' note exactly what Kramers had and had not yet done. Van Vleck thought that his calculations extended Kramers' results. As he wrote to Kramers in September 1924:

I am enclosing under separate cover a reprint [Van Vleck, 1924a] which I think may be of interest to you, especially the footnote at the very end, where I mention some computations I have made relative to your dispersion formula. A longer paper [Van Vleck 1924b, c] is now in proof, and should appear shortly in the Physical Review. This more extensive article was ready to send to the printer about the time we received the copy of Nature containing your dispersion formula. In your note [Kramers, 1924a] I did not understand you to state how generally you had verified the asymptotic connection with the classical dispersion from the actual orbit, and it immediately occurred to me that this question could easily

[99] Kramers to Ladenburg, June 5, 1924 (AHQP).

[100] Ladenburg to Kramers, June 8, 1924 (AHQP).

[101] See Sects. 2.4 and 5.2 for quotations from correspondence between Born and Van Vleck in October–November 1924 pertaining to these papers.

be investigated by the perturbation theory method I had previously developed in connection with what I call the "correspondence principle for absorption". I therefore inserted two sections (# 6 and # 15 …) showing that your formula merged into the classical one.

Inasmuch as the classical dispersion formula had apparently not been developed for the general non-degenerate multiply periodic orbit, and as you did not give this in your note to Nature, I conjectured that you had verified the asymptotic connection only in special cases, such as a linear oscillator, so that my computations on dispersion would not be a duplication of what you had done. However, while visiting at Cambridge, Mass. last week I learned from Dr. Slater that your calculation of the asymptotic connection was almost identical with my own in scope and generality. I have therefore altered the proof of my Physical Review article to include a note [Van Vleck 1924b, p. 345] stating that you have also established the correspondence theorem in the general case. I hope this is satisfactory to you. The concept and introduction of the virtual-oscillator formula is entirely yours, and I refer always to the "Kramers dispersion formula", but I had developed the perturbation theory method for absorption etc. prior to learning of any of your work.

I am sorry that we are again apparently duplicating each other in some of our work. Slater tells me that by extending your computations he has independently derived an absorption formula similar to mine, and also noted the asymptotic connection of the two theories in this case.[102]

As in the case of Ladenburg and Reiche half a year earlier, Kramers did not seem to mind at all that Van Vleck was poaching on his preserves. He generously wrote back to Van Vleck: "Your note on absorption made me much pleasure and I think it very just of Providence that you got it published before hearing of our work."[103]

The construction of the dispersion formula (9) requires as a prelude to the application of the correspondence principle, a derivation of the classical formula for the dipole moment of an arbitrary (non-degenerate) multiply-periodic system. This is where Ladenburg and Reiche (1923) came up short, even though, as we saw above, Reiche was able to reconstruct the derivation once Kramers had outlined it for him. Kramers and Van Vleck, like Epstein before them, used canonical perturbation techniques from celestial mechanics to derive this classical formula. In Part Two of our paper, closely following the classical part of Van Vleck's two-part paper (Van Vleck, 1924c), we shall present a detailed derivation of this crucial classical formula, for the special case of the harmonic oscillator in Sect. 5.1 and for a general non-degenerate multiply-periodic system in Sect. 6.2. Guided by the correspondence principle and introducing the A and B coefficients we then construct a quantum formula that merges with the classical formula for high quantum numbers (see Sects. 5.2 and 6.2).[104] Here we summarize the main steps of this derivation.

---

[102] Van Vleck to Kramers, September 22, 1924 (AHQP).

[103] Kramers to Van Vleck, November 11, 1924 (AHQP).

[104] Recall that Van Vleck actually did it the other way around: he started with the quantum formula and checked that this formula merges with the classical formula in the correspondence limit (see note 17).

In general coordinates (q_i, p_i), Hamilton's equations are:

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}, \tag{10}$$

where dots indicate time derivatives. Given the Hamiltonian H of some multiply-periodic system, one can often find special coordinates (w_i, J_i), so-called *action-angle variables*, in which Hamilton's equations take on a particularly simple form:

$$\dot{w}_i = \frac{\partial H}{\partial J_i} = \nu_i, \quad \dot{J}_i = -\frac{\partial H}{\partial w_i} = 0. \tag{11}$$

The angle variables, $w_i = \nu_i t$, give the characteristic frequencies of the system; the (conserved) action variables are subject to the Bohr–Sommerfeld quantum condition, $J_i = n_i h$. This, of course, is why these variables are of particular interest in this context.

Suppose we have a Hamiltonian H that is the sum of H_0, describing some multiply-periodic system representing an electron orbiting the nucleus of an atom in the Bohr–Sommerfeld theory (or, an inner planet like Mercury orbiting the sun), and $H_{\text{int}} = eEx \cos 2\pi \nu t$, a small perturbation describing the interaction of this system with a weak periodic electric field in the x-direction (or, the periodic weak gravitational interaction with a distant outer planet). To find the induced polarization responsible for dispersion in this system we need to calculate the coherent part Δx_{coh} of the displacement caused by the perturbation (cf. Eqs. (2)–(6) in Sect. 3.1). We assume that the unperturbed system can be solved in action-angle variables, which means that $x(t)$ in the absence of H_{int} can be written as a Fourier series:

$$x(t) = \sum_{i, \tau_i} A_{\tau_i}(J_l) e^{2\pi i \tau_i w_i} \tag{12}$$

(where i runs from 1 to 3 and τ_i runs over all positive and negative integers). The complex amplitudes have to satisfy the conjugacy relations $A_{\tau_i} = A_{-\tau_i}^*$ to ensure that $x(t)$ is real. Assuming the interaction is switched on at $t = 0$, we can use Hamilton's equations in action-angle variables—still those for H_0 rather than those for the full Hamiltonian H[105]—to calculate Δw_i and ΔJ_i due to the perturbation. We insert the results into

$$\Delta x = \sum_k \left(\frac{\partial x}{\partial J_k} \Delta J_k + \frac{\partial x}{\partial w_k} \Delta w_k \right), \tag{13}$$

and collect the coherent terms (i.e., all terms with a factor $e^{2\pi i \nu t}$). The result is:

$$\Delta x_{\text{coh}} = 2eE \sum_{i,k,\tau_i,\tau_k} \tau_k \frac{\partial}{\partial J_k} \left(\frac{\tau_i \nu_i}{\nu^2 - (\tau_i \nu_i)^2} |A_{\tau_i}(J_l)|^2 \right) \cos 2\pi \nu t. \tag{14}$$

---

[105] Here Van Vleck's calculation differs from those of Born (1924) or Kramers and Heisenberg (1925).

For the special case of a charged harmonic oscillator, this expression reduces to the simple expression (6) found earlier (as we shall show in detail at the end of Sect. 5.1).

We now translate this classical formula into a quantum formula. The idea is to construct a quantum formula that merges with the classical formula in the limit of high quantum numbers. This is done in three steps. For high values of the quantum number i, the derivatives $\partial/\partial J_i$ can be replaced by difference quotients,[106] the square of the amplitudes $A_{\tau_i}(J_l)$ by transition probabilities $A_{i \to j}$ (where $|i - j|$ is small compared to i), and orbital frequencies ν_i by transition frequencies $\nu_{i \to j}$. We then take the leap of faith that the resulting formula holds for all quantum numbers. Multiplying by the charge $-e$ and the number of atoms N to get from the coherent part of the displacement of one atom to the polarization of a group of atoms, we arrive at the Kramers dispersion formula (9).

3.5 Heisenberg's *Umdeutung* and dispersion theory

The Kramers dispersion formula was a crucial step in the transition from the old quantum theory to matrix mechanics, and thereby in the transition from functions on classical phase spaces to operators on Hilbert spaces. As Kramers pointed out in his second *Nature* note, the formula

> only contains such quantities as allow of a direct physical interpretation on the basis of the fundamental postulates of the quantum theory…and exhibits no further reminiscence of the mathematical theory of multiple [sic] periodic systems (Kramers, 1924b, p. 311)

This point is amplified in the Kramers–Heisenberg paper:

> we shall obtain, quite naturally, formulae which contain only the frequencies and amplitudes *which are characteristic for the transitions*, while all those symbols which refer to the mathematical theory of periodic systems will have disappeared (Kramers and Heisenberg, 1925, p. 234, our emphasis).

Orbits do not correspond to observable quantities, but transitions do, namely to the frequency $\nu_{i \to f}$ of the emitted radiation, and, through the Einstein coefficients $A_{i \to f}$, to its intensity. In the introduction of his *Umdeutung* paper, Heisenberg (1925c) explained that he wanted "to establish a theoretical quantum mechanics, analogous to classical mechanics, but in which only relationships between observable quantities occur" (p. 262). In the next sentence he identified the Kramers dispersion theory as one of "the most important first steps toward such a quantum-theoretical mechanics" (ibid.).

Rather than using classical mechanics to analyze features of electron orbits and translating the end result into a quantum formula, as Kramers and others had done (cf. Eqs. (10)–(14) above), Heisenberg translated the Fourier series for the position of an electron that forms the starting point of such classical calculations into a

---

[106] This replacement is known as "Born's correspondence rule." In fact, both Kramers and Van Vleck found it independently of Born. We return to this point in Sect. 5.2.

quantum expression. He replaced the amplitudes and frequencies by two-index quantities, referring to the initial and final state of a quantum transition, respectively, and thus replaced classical position by an array of numbers associated with transitions between states. Reinterpreting rather than replacing the old theory, he assumed that these new quantities would satisfy all the familiar relations of Newtonian mechanics. Note that Heisenberg thus formulated a new theory directly in terms of transition quantities without bothering to find a representation for the states connected by the transitions.

The Bohr–Sommerfeld quantization condition (1) has the form of a restriction on orbits in phase space. With the elimination of orbits, it could no longer be used, at least not in its original form. As Heisenberg recalls in his AHQP interview:

> I had, of course, to think about the quantum condition. And that was an important point. But there I knew so much from Copenhagen how important this Thomas–Kuhn sum rule was. That took some time. That I think I had done in Göttingen, [I] had seen how I could translate the Thomas–Kuhn sum rule into what I call a quantum mechanical statement, into a statement in which only differences occurred. I did not see that it was a commutation rule [but with this translation] I can bring this sum rule into my whole scheme and then this sum rule actually fixes everything. I could see that this fixes the quantization.[107]

The Thomas–Kuhn sum rule, a corollary of the Kramers dispersion formula (see Sect. 7.1 for a derivation in modern quantum mechanics), had been found independently by Werner Kuhn (1899–1963) (1925) in Copenhagen[108] and by Willy Thomas (1925) in Breslau.[109] Kuhn (i.e., Thomas Kuhn) pressed Heisenberg a little on how he had settled on this rule as his fundamental quantization condition: "Using the Kuhn–Thomas [sic] rule is a stroke of genius but one supposes that there were a lot of other intermediate attempts." Apparently there were none. Heisenberg insisted:

> No, I would say it was rather trivial for the following reasons: First of all, there was the integral pdq…I felt that perhaps only the difference of integral pdq between one quantum state and the next quantum state is an important thing. So I actually felt, "Well, perhaps I should write down integral pdq in one state minus integral pdq in the neighboring state." Then I saw that if I write down this and try to translate it according to *the scheme of the dispersion theory*, then I get the Thomas–Kuhn sum rule. And that is the point. Then I thought, "Well, that is apparently the way how it is done."[110]

In other words, following the general recipe introduced in the *Umdeutung* paper for the translation of classical formulae into quantum-mechanical ones— "the scheme of

---

[107] See p. 10 of the transcript of session 7 of the AHQP interview with Heisenberg

[108] The publication of Kuhn's paper had been delayed in typical Copenhagen fashion: "A paper on the summation rule had been submitted to Prof. Bohr and Prof. Kramers about half a year before the final one, but it was rejected at that time because it contained besides the main good argument some unsuitable passages" (Werner Kuhn to Thomas Kuhn, May 3, 1962 [included in the folder on Kuhn in the AHQP])

[109] Thomas was a student of Reiche in Breslau who died young of tuberculosis. See p. 14 of the transcript of the third session of the AHQP interview with Reiche.

[110] See p. 10 of the transcript of session 7 of the AHQP interview with Heisenberg (our emphasis).

the dispersion theory"—Heisenberg (1925c, p. 268) was able to convert *a derivative of* the Bohr–Sommerfeld condition into an equation that contains only amplitudes and frequencies. Since Heisenberg's theory only deals with transitions between states, the absolute value of the action J does not matter; only the difference in J-value between two states does.

The sum rule is sometimes called the Thomas–Kuhn–Reiche sum rule because Reiche and Thomas (1925) were the first to publish a detailed derivation of it in a paper *submitted* to the *Zeitschrift für Physik* in early August 1925 about a month before (Heisenberg, 1925c) *appeared* in the same journal. In formulating the goal of their paper, Reiche and Thomas not only used the term 'Umdeutung' in very much the same way as Heisenberg in his *Umdeutung* paper, they also explicitly tied this usage to Kramers' dispersion theory:

> We use…the correspondence principle in the same way in which it was applied by Kramers in the derivation of the dispersion formula by reinterpreting [*umdeuten*] the mechanical orbital frequencies as radiation frequencies, the Fourier coefficients as the "characteristic amplitudes" that determine the quantum radiation, and, finally, in analogy to the Bohr frequency condition,[111] differential quotients as difference quotients. In the realm of high quantum numbers the classical and quantum-theoretical representations become identical. We try to arrive at a general relation, by maintaining the reinterpretation [*Umdeutung*] of classical quantities into quantum-theoretical ones for all quantum numbers (Reiche and Thomas, 1925, pp. 511–512).

In view of the tendency of European theorists to neglect American contributions (see Sect. 2.4), it is also interesting to note that Reiche and Thomas (1925, p. 513) cite (Van Vleck, 1924b).

Although he failed to recognize the importance of the result at the time, Van Vleck had, in fact, been the first to find the sum rule (Sopka, 1988, p. 135, note 184). As he wrote in his NRC *Bulletin*:

> Eq. (62a) [a version of the sum rule] appears to have been first incidentally suggested by the writer [Van Vleck 1924c, pp. 359–360, footnote 43] and then was later and independently much more strongly advanced by Thomas…Kuhn…and Reiche and Thomas (Van Vleck, 1926, p. 152).

Van Vleck is referring to a footnote in the section on dispersion in the classical part of his paper. In this footnote he mentioned two objections that explain why he did not put greater emphasis on the sum rule himself. Van Vleck's idea—which he calls "tempting (but probably futile)" (Van Vleck, 1924c, p. 359, footnote 43)—was that the sum rule would allow him to compute the Einstein A coefficients. He was under the impression, however, that "such a method is hard to reconcile with the [experimental] work of F. C. Hoyt [1923, 1924]" on X-ray absorption and that it "would lead to transitions

---

[111] In the limit of high quantum numbers, the Bohr frequency condition, $v_{i \to j} = (E_i - E_j)/h$, merges with the relation $v_i = \partial H / \partial J_i$ (cf. Eq. (11)). Van Vleck (1924b, p. 333) calls this the correspondence theorem for frequencies.

from positive to negative quantum numbers, which can scarcely correspond to any physical reality" (ibid.).

As Heisenberg (1925c, pp. 269–270) shows briefly in his paper, the sum rule follows from the Kramers dispersion formula (9) if one takes the limit in which the frequency ν of the incident radiation is much greater than any of the absorption frequencies $\nu_{i \to j}$ (see Sect. 7.1). That the quantization condition obtained by massaging the Bohr–Sommerfeld condition also follows from the Kramers dispersion theory, widely recognized as one of the most secure parts of the old quantum theory, must have bolstered Heisenberg's confidence in the translation procedure of his *Umdeutung* paper. It was left to Born and Jordan (1925b) to extract the now standard commutation relations for position and momentum from the Thomas–Kuhn sum rule (in Sect. 7.1 we shall show in detail how this is done). That Heisenberg stopped short of making this move is largely, as we shall argue in Sect. 7.1, because he was thinking in terms of the positions and velocities of the Lagrangian formalism rather than in terms of the positions and momenta of the Hamiltonian formalism.

Although Heisenberg thus relied heavily on dispersion theory in his *Umdeutung* paper, he gave his positivist methodology pride of place. This philosophical outlook probably came from a variety of sources. Pauli, Heisenberg's fellow student and frequent discussion partner (both in person and in writing), was a devoted follower of his godfather Ernst Mach (1838–1916).[112] As Pauli had written to Bohr, for instance, on December 12, 1924:

> We must not…put the atoms in the shackles of our prejudices (of which in my opinion the assumption of the existence of electron orbits in the sense of the ordinary kinematics is an example); on the contrary, we must adapt our concepts to experience (Bohr, 1972–1996, Vol. 5, pp. 35–36).

We already indicated in Sect. 1.1 that Heisenberg himself later claimed that his positivist attitude came in part from his reading of Einstein's 1905 special relativity paper.[113] His biographer David Cassidy (1991, p. 198) makes the suggestive observation that Born and Jordan (1925a, p. 493), in a paper completed by June 11, 1925, not only emphasized the observability principle but also appealed to Einstein's analysis of distant simultaneity in support of it.

As Helge Kragh (1999) notes: "there was no royal road from the observability principle to quantum mechanics" (p. 162). This truism is nicely illustrated by a conversation between Einstein and Heisenberg reported years later by the latter. The following exchange supposedly took place in Berlin in the spring of 1926:

> "But you don't seriously believe," Einstein protested, "that none but observable magnitudes must go into a physical theory?" "Isn't that precisely what you have done with relativity?" I asked in some surprise…"Possibly I did use this kind

---

[112] On Pauli's positivism, see, e.g., (Hendry, 1984, pp. 19–23) and (Gustavson, 2004).

[113] See, e.g., (Holton, 2005, pp. 26–31) for discussion.

of reasoning," Einstein admitted, "but it is nonsense all the same" (Heisenberg, 1971, p. 63).[114]

With his *S*-matrix program in the 1940s,[115] Heisenberg once again tried to force a theoretical breakthrough by restricting himself to observable quantities, this time with the qualification that he had taken to heart Einstein's lesson that, in the end, it is the theory that determines what the observables are. Heisenberg (1971, p. 63) has Einstein make this point a few sentences after the passage quoted above and acknowledges it as a source of inspiration for his 1927 uncertainty principle. Nearly two decades after the *Umdeutung* paper, Heisenberg (1943) wrote: "in this situation it seems useful to raise the question which concepts of the present theory can be retained in the future theory, and this question is roughly equivalent to a different question, namely which quantities of the current theory are "observable"…Of course, it will always only be decided by the completed theory which quantities are truly "observable"" (p. 514).

As Einstein complained in 1917 in a letter to his friend Michele Besso (1873–1955), referring to the excessive Machian positivism of their mutual acquaintance Friedrich Adler (1879–1960): "He is riding the Machian nag [*den Machschen Klepper*] to exhaustion." In a follow-up letter he elaborated: "It cannot give birth to anything living, it can only stamp out harmful vermin."[116] This is true in the case of matrix mechanics as well. Heisenberg's positivism would have been perfectly sterile if it had not been for Kramers' dispersion theory. In that context, positivism was not a blanket injunction against unobservable quantities in general but was directed at a specific set of increasingly problematic unobservables, the electron orbits of the Bohr–Sommerfeld theory.

4 The Bohr–Kramers–Slater (BKS) theory as a detour on the road from dispersion theory to matrix mechanics

4.1 Virtual oscillators and virtual radiation

Kramers presented his work on dispersion theory in the context of the BKS theory, not just in the two preliminary notes to *Nature* discussed in Sect. 3.4, but also in the authoritative exposition of his dispersion theory in the paper with Heisenberg. In the abstract of this paper, the authors announce that

> [t]he arguments are based throughout on the interpretation of the connection of the wave radiation of the atom with the stationary states advocated in a recent paper by Bohr, Kramers and Slater [1924a,b], and the conclusions, should they

---

[114] Quoted and discussed, for instance, in (MacKinnon, 1977, p. 185) and in (Holton, 2005, pp. 30–31). For other versions of the same anecdote, see (Heisenberg, 1983, pp. 113–114) and pp. 18–19 of the transcript of session 5 of the AHQP interview with Heisenberg.

[115] See (Pais, 1986, 497–503), (Dresden, 1987, 453–458), and, especially, (Cushing, 1990) for discussion. See also pp. 20–21 of the transcript cited in note 114.

[116] Einstein to Besso, April 29 and May 13, 1917, respectively (Einstein, 1987–2006, Vol. 8, Docs. 331 and 339). For further discussion, see, e.g., (Holton, 1968).

be confirmed, would form an interesting support for this interpretation (Kramers and Heisenberg, 1925, p. 223).[117]

It should thus come as no surprise that the Kramers dispersion theory has been portrayed as an application of the BKS theory in most older and even in some more recent historical literature.[118] Jammer (1966), for instance, writes that BKS "was the point of departure of Kramers's detailed theory of dispersion" (p. 184). Mara Beller (1999) still characterized (Kramers and Heisenberg, 1925) as a paper that "spelled out, in a rigorous mathematical way, the ideas only roughly outlined in the presentation of Bohr, Kramers, and Slater" (p. 23). More than a decade earlier, Dresden (1987, pp. 144–146, pp. 220–221) had in fact already set the record straight.[119] Darrigol (1992, p. 225) duly emphasizes that *the Kramers dispersion theory was developed before and independently of BKS*. Even before Dresden, Hendry (1981) had already made it clear that *BKS got its virtual oscillators from dispersion theory—the substitute oscillators of (Ladenburg and Reiche, 1923)—and not the other way around*. We briefly review the evidence in support of the italicized claims above.

We know from the passage quoted in Sect. 3.4 from a letter from Slater to Van Vleck that by the time the former arrived in Copenhagen around Christmas 1923 Kramers already had his dispersion formula. Kramers must have used the substitute oscillators of (Ladenburg and Reiche, 1923) at that point even though by the time he finally got around to publishing his formula he called them virtual oscillators (see Sect. 3.4). Slater's arrival in Copenhagen marks the lower limit for the birth of the BKS theory. The theory, after all, grew around an idea that Slater hit upon shortly before he left for Europe late that year.[120] Slater suggested that the wave and particle properties of light might be reconciled by having an electromagnetic field guide corpuscular light quanta.[121] Bohr and Kramers cannibalized Slater's idea and stripped it of all reference to light quanta. Against his better judgment—as he insisted decades later in a letter of November 4, 1964 to van der Waerden (1968, p. 13)—Slater went along and his idea entered the literature via the BKS paper. In a short letter sent to *Nature* a week after this joint paper had been submitted, Slater explained how Bohr and Kramers had convinced him of their point of view. Accordingly, he presented his idea couched in BKS terms:

---

[117] We use the translation of (Stolzenburg, 1984, p. 87) at this point, which is more accurate than the standard translation in (Van der Waerden, 1968, p. 223).

[118] There is an extensive literature on BKS; see, e.g., (Klein, 1970, pp. 23–39), (Stuewer, 1975, pp. 291–305), (Hendry, 1981), the dissertation of Neil Wasserman (1981), (Mehra and Rechenberg, 1982–2001, Vol. 1, Sect. V.2), the essay by Klaus Stolzenburg (1984) in (Bohr, 1972–1996, Vol. 5, pp. 3–96), and (Dresden, 1987, pp. 159–215).

[119] See (Dresden, 1987, p. 221) for a helpful chronology of events in 1923–1925 pertaining to BKS and dispersion theory.

[120] See Slater to his mother, November 8, 1923 (quoted in Dresden, 1987, p. 161); Slater to Kramers, December 8, 1923 (AHQP). For discussions of Slater's idea, see (Klein, 1970, p. 23), (Stuewer, 1975, pp. 291–294), (Hendry, 1981, pp. 213–214), (Stolzenburg, 1984, pp. 6–11), and (Darrigol, 1992, pp. 218–219).

[121] Slater was probably unaware that Einstein and Louis de Broglie (1892–1987) had already made similar suggestions (Hendry, 1981, p. 199; Darrigol, 1992, p. 218).

Any atom may, in fact, be supposed to communicate with other atoms all the time it is in a stationary state, by means of a virtual field of radiation originating from oscillators having the frequencies of possible quantum transitions and the function of which is to provide for the statistical conservation of energy and momentum by determining the probabilities for quantum transitions (Slater, 1924, p. 307).

The final clause about the statistical conservation of energy and momentum was foisted upon Slater by Bohr and Kramers.[122] Bohr had been contemplating such a move for several years, as can be inferred, for instance, from correspondence with Ehrenfest in 1921 in connection with the third Solvay congress held that year (Klein, 1970, p. 19) and with Darwin in 1922 (Stolzenburg, 1984, pp. 13–19). Slater's concept of virtual radiation emitted while an atom is in a stationary state fit nicely with Bohr's tentative ideas concerning the mechanism of emission and absorption of radiation. In Sects. 3.2–3.3, we quoted various comments by Bohr on dispersion from the period 1916–1923 showing how he came to embrace the notion that an atom interacts with radiation like a set of oscillators.

The concept of virtual oscillators is often attributed to Slater, not just by later historians (see, e.g., Stuewer, 1975, p. 291, p. 303) but also by his contemporaries. In the abstract of (Van Vleck, 1924b), for instance, we read that the Kramers dispersion formula "assumes the dispersion to be due not to the actual orbits but to Slater's 'virtual' or 'ghost' oscillators having the spectroscopic rather than orbital frequencies" (p. 330).[123] In the BKS paper itself, however, the concept is unambiguously attributed to Ladenburg:[124]

The correspondence principle has led to comparing the reaction of an atom on a field of radiation with the reaction on such a field which, according to the classical theory of electrodynamics, should be expected from a set of 'virtual' harmonic oscillators with frequencies equal to those determined by $[h\nu = E_1 - E_2]$ for the various possible transitions between stationary states.[125] Such a picture has been used by Ladenburg[126] in an attempt to connect the experimental results on dispersion quantitatively with considerations on the probability of transitions between stationary states (Bohr, Kramers, and Slater, 1924a, pp. 163–164).

As we saw in Sect. 3.3, Ladenburg and Reiche in turn attributed the idea to Bohr. In 1924, for instance, they wrote:

---

[122] See also (Bohr, Kramers, and Slater, 1924a, p. 160).

[123] See also (Van Vleck, 1924a, p. 30), quoted in Sect. 3.4, and (Van Vleck, 1926, p. 163).

[124] The mistakes with the prepositions in the passage below ('reaction on' instead of 'reaction to' and 'considerations on' instead of 'considerations of') would tend to support Slater's claim that the paper was "written entirely by Bohr and Kramers" (Slater to Van Vleck, July 27, 1924, quoted in Sect. 2.2).

[125] At this point, the authors refer to Ch. III, Sect. 3 of (the English translation of) (Bohr, 1923b), the section in which (Ladenburg, 1921) is discussed and which triggered the correspondence between Bohr and Ladenburg discussed in Sect. 3.3.

[126] At this point, the authors append a footnote referring to (Ladenburg, 1921) and (Ladenburg and Reiche, 1923).

Formally, we can describe the relation [between oscillator strengths and transition probabilities] following an assumption introduced by Bohr [1923b, pp. 161–162], by imagining that the atom responds to external radiation like a system of electrical oscillators, whose characteristic frequencies ν agree with the emitted or absorbed frequencies in *possible* quantum transitions (Ladenburg and Reiche, 1924, p. 672).

In the next sentence they use their own term "substitute oscillators" (in quotation marks) and add: "(now called "virtual oscillators")" (ibid.). Likewise, in the introduction of the opening installment of a series of papers on the experimental verification of the Kramers dispersion formula, Ladenburg talks about "the "substitute oscillators"," which were introduced, "at Bohr's suggestion, as the carriers of the scattered radiation needed for dispersion" (Ladenburg, 1928, p. 16).[127]

   Bohr had communicated the idea in a letter to Ladenburg (see Sect. 3.3). This may explain why, when interviewed for the AHQP, Reiche did not remember who originally came up with it:

I do not know whether we or Kramers first used this terminology of virtual oscillators…It might be it is Kramers. If it was Kramers then we certainly at once incorporated it into our thinking.[128]

In his AHQP interview with Slater, Kuhn also asked about virtual oscillators:

to what extent did that come from [the BKS] paper, to what extent does it really go back to the Ladenburg, and Ladenburg–Reiche [papers]? It could have grown out [of the] Ladenburg and Ladenburg–Reiche papers, yet my impression from the literature is that there was little done with that until after the Bohr–Kramers–Slater paper.[129]

Slater concurred, though his comments would have been more valuable had he not been asked such a leading question:

I think that's true. Of course, I was very familiar with the Ladenburg–Reiche things,[130] so was Bohr. I think that we helped popularize it in a sense. Of course, this also came at the same time, approximately, that Kramers was working on his dispersion formula. That again is operating with things very much like the virtual oscillator, so they all seem to hang together, and I think it was a combination of the oscillators from our paper, from the Ladenburg–Reiche, and the Heisenberg–Kramers dispersion that really set them in operation.[131]

Despite the loaded question that elicited this response and even though Slater is wrong to suggest that BKS and Kramers' dispersion theory were developed independently of

---

[127] The passage from which these clauses are taken is quoted in full at the beginning of Sect. 7.

[128] See p. 11 of the transcript of session 3 of the interview with Reiche. It could be, however, that Reiche was only referring to the new *term* for the Bohr–Ladenburg–Reiche concept of substitute oscillators.

[129] See p. 34 of the transcript of the first session of the AHQP interview with Slater.

[130] (Ladenburg, 1921) and (Ladenburg and Reiche, 1923) are cited in (Slater, 1925a, p. 397).

[131] See pp. 34–35 of the transcript of the first session of the AHQP interview with Slater.

the earlier work of Ladenburg and Reiche, the overall characterization of the situation seems to be accurate. BKS officially sanctioned the dual representation of the atom as simultaneously a quantum system à la Einstein and Bohr and a set of oscillators à la Helmholtz, Lorentz and Drude. This dual picture had been implicit in (Ladenburg, 1921). It was made explicit, under Bohr's influence, in (Ladenburg and Reiche, 1923). That it was endorsed by the highest authorities in Copenhagen undoubtedly helped its dissemination. Even so it was typically presented with some trepidation. In his second *Nature* note, Kramers tried to pass it off as merely a matter of words:[132]

> In this connexion it may be emphasized that the notation 'virtual oscillator' used in my former letter [Kramers, 1924a] does not mean the introduction of any additional hypothetical mechanism, but is meant only as a terminology suitable to characterise certain main features of the connexion between the description of optical phenomena and the theoretical interpretation of spectra (Kramers, 1924b, p. 311).

Van Vleck was more upfront:

> The introduction of these virtual resonators is, to be sure, in some ways very artificial, but is nevertheless apparently the most satisfactory way of combining the elements of truth in both the classical and quantum theories. In particular this avoids the otherwise almost insuperable difficulty that it is the spectroscopic rather than the orbital frequencies…which figure in dispersion (Van Vleck, 1924b, p. 344).

Despite such disclaimers, Kramers and Van Vleck—as well as Slater, Born, Breit and others working in the general area of dispersion theory in 1924–1925—used a model of the atom in which the electron orbits of the Bohr–Sommerfeld theory were supplemented by an "orchestra of virtual oscillators"[133] with characteristic frequencies corresponding to each and every transition that an electron in a given orbit can undergo. Thanks to virtual oscillators—to paraphrase Heisenberg's succinct statement to van der Waerden (1968, p. 29) in 1963—at least *something* in the atom was vibrating with the right frequency again.

The dual representation of physical systems (of electrons rather than atoms in this case) was also key to the BKS explanation of the Compton effect. BKS was Bohr's last stand against light quanta after the Compton effect had finally convinced most other physicists that they were unavoidable (Klein, 1970, p. 3).[134] BKS explains the Compton effect without light quanta. It attributes the frequency shift between incoming and scattered X-rays to a Doppler shift in the X-ray wave fronts instead. Compton (1923) thought this option was ruled out because, as he showed in his paper, the recoil velocity needed to get the right Doppler shift is different from the recoil velocity needed to ensure conservation of energy and momentum in the process, and

---

[132] In the work that led to (Kramers and Heisenberg, 1925), however, Kramers, according to Hendry (1981), "ignored their virtual nature altogether and treated the oscillator model as naively as he had the orbital model" (p. 202).

[133] The term "virtual orchestra" comes from (Landé, 1926, p. 456) (Jammer, 1966, p. 187).

[134] Cf. our comments in the introduction to Sect. 3.

one and the same electron cannot recoil with two different velocities. In the BKS theory, however, there is room for two recoil velocities, one for the electron itself, one for the orchestra of virtual oscillators associated with it.[135] The Compton effect can be interpreted as a Doppler shift if the appropriate recoil velocity is assigned to the virtual oscillators. Energy and momentum can be conserved if a different recoil velocity is assigned to the electrons themselves. Bohr and his co-authors wasted few words on the justification of this startling maneuver:

> That in this case the virtual oscillator moves with a velocity different from that of the illuminated electrons themselves is certainly a feature strikingly unfamiliar to the classical conceptions. In view of the fundamental departures from the classical space-time description, involved in the very idea of virtual oscillators, it seems at the present state of science hardly justifiable to reject a formal interpretation as that under consideration as inadequate (Bohr, Kramers, and Slater, 1924a, p. 173).

This is almost as bad as pieces of glass dragging along different amounts of ether for different colors of light in early nineteenth-century ether theory (see Sect. 3.1)!

The problem carries over to the dispersion theory based on the dual representation of atoms in terms of classical orbits and virtual oscillators, as is acknowledged, if only in passing, by Kramers and Heisenberg (1925): "We shall not discuss in any detail the curious fact that the centre of these spherical waves moves relative to the excited atom" (p. 229). This exacerbated the problem of the Bohr–Sommerfeld orbits in the theory. Not only were they responsible for the discrepancy between orbital frequencies and radiation frequencies, they also make it harder to picture an atom in space and time. After all, the system of electron orbits does not even move in concert with its orchestra of virtual oscillators.

Edward MacKinnon (1977, 1982) has suggested that the resulting problem of combining different pictures of the atom into one coherent picture forced Heisenberg to make a choice between them (see also Beller, 1999, p. 23). Since the virtual oscillators carry all the physical information while the electron orbits are completely unobservable, the choice is obvious. MacKinnon (1977, p. 138) has gone as far as describing Heisenberg's *Umdeutung* paper as proposing a theory of virtual oscillators. Of course, there is no explicit reference to virtual oscillators anywhere in the *Umdeutung* paper. MacKinnon (1977, pp. 155–156, 162, 177) speculates that this is because Heisenberg suppressed all talk about virtual oscillators as a response to Pauli's objections to the "virtualization" of physics.[136] We shall return to the relation between BKS and Heisenberg's work in Sect. 4.3.

Pauli had originally promised not to subvert Bohr's efforts to get the physics community to accept the term 'virtual' as used in the context of BKS. Working on the

---

[135] What Compton (1923) actually said in his paper is very suggestive of this option: "It is clear…that so far as the effect on the wave-length is concerned, we may replace the recoiling electron by a scattering electron" with an "effective velocity" different from that of the recoiling electron (p. 487; quoted and discussed in Stuewer, 1975, p. 230).

[136] In a letter of January 8, 1925, Heisenberg told Bohr that Pauli did not believe "in virtual oscillators and is outraged at the 'virtualization' of physics" (MacKinnon, 1977, p. 156).

German translation of the paper (Bohr, Kramers, and Slater, 1924b), Bohr was anxious to ensure that Pauli approved of "the words "communicate" and "virtual", for after lengthy consideration, we have agreed here on these basic pillars of the exposition."[137] In typical Bohr fashion, he first announced that the manuscript would be submitted that same day and that he would enclose a copy, then added a postscript saying that there had been further delays and that it would be sent later.[138] Amused, Pauli wrote back a few days later:

> I laughed a little (you will certainly forgive me for that) about your warm recommendation of the words "communicate" and "virtual" and about your postscript that the manuscript is still not yet completed. On the basis of my knowledge of these two words (*which I definitely promise you not to undermine*), I have tried to guess what your paper may deal with. But I have not succeeded.[139]

The term 'virtual' also puzzled the group of physicists in Ann Arbor studying the BKS paper with Bohr's former associate Klein, who wrote to Bohr on June 30, 1924: "Colby [cf. note 37], who is also most interested in it, asked me about the meaning of the term 'virtual radiation'" (Stolzenburg, 1984, p. 29).

Exactly what does the 'virtual' in virtual oscillator and virtual radiation mean? Virtual oscillators can be thought of in analogy to virtual images in geometric optics. Just as the light reflected from a mirror appears to come from an imaginary point behind the mirror, the light scattered by an atom appears to come from an imaginary oscillator. This analogy, however, is nowhere to be found in the BKS paper. Whatever its exact meaning, the designation 'virtual' does serve as a warning that these oscillators are not just classical oscillators. The authors warn, for instance, that "the absorption and emission of radiation are coupled to different processes of transition, and thereby to different virtual oscillators" (Bohr, Kramers, and Slater, 1924a, p. 171).

Unlike the light coming from virtual images in geometric optics, the radiation coming from virtual oscillators is *also* called virtual in the BKS paper. Again, it is not clear why. As the analogy with geometric optics shows, that a source is virtual does not mean that the radiation must be virtual as well. In Slater's original conception, the radiation might be called virtual in the sense that the light quanta are the primary reality and that the radiation is there only to guide them. In the BKS theory, however, there are no light quanta, only the radiation.

The way Heisenberg later remembered it, the virtual radiation of BKS had a status similar to that of the Schrödinger wave function in Born's statistical interpretation a few years later. As Heisenberg told Kuhn in his AHQP interview:

> What Bohr, Kramers, and Slater did was to establish the probability as a kind of reality…one felt that by making the probability become some kind of reality, you get hold of something which is there. It was at that time of course, very difficult to say what it was that you had gotten hold of. I would say only through

---

[137] Bohr to Pauli, February 16, 1924 (Bohr, 1972–1996, Vol. 5, p. 409).

[138] Contrary to what is suggested by these delays, the German translation simply follows the English original.

[139] Pauli to Bohr, February 21, 1924 (Bohr, 1972–1996, Vol. 5, p. 412; our emphasis).

the paper of Born [1926] did it become quite clear that one should say, "All right, the Schrödinger wave means that probability that an electron should be there." But the main point was that the probability itself was something real. It was not only in the mind of the people, but it was something in nature…Up to that time people had two possibilities. One possibility was that the reality is a wave. There is an electric field, and a magnetic field acting upon an atom, shaking the electron, and then the atom does something, it makes a transition …There is an entirely different picture of reality in which there is a light quantum…hitting the atom, and then something happens. But now the idea is that there is a wave. But this wave is not the reality. This wave is a probability—this wave is a tendency. It means that when this wave is present then the atom gets a tendency to emit light quanta. So this idea of the wave field being a tendency was something just in the middle between reality and non-reality…That was the striking thing about [BKS], you know, this new invention of a possibility which was a reality in some way but not a real reality—a half reality.[140]

Unsurprisingly, Born took exception to Heisenberg's suggestion that the Born interpretation had been anticipated in this way by BKS. As Heisenberg said in a subsequent session of the interview: "I felt once, when I discussed this matter with Born, that he was a bit angry that I had quoted too much the Bohr–Kramers–Slater paper in connection with the probability interpretation of waves."[141] We sympathize with Born. Heisenberg's comments, we feel, have all the flavor of an after-the-fact rationalization.

In subsequent expositions of the BKS theory by both Kramers and Slater, the radiation from virtual oscillators is presented as every bit as real as the external radiation. It is hard to see how this could be otherwise since the two types of radiation are supposed to interfere with one another. Bohr, Kramers, and Slater (1924a) write: "we shall assume that [illuminated atoms] will act as secondary sources of *virtual* wave radiation which interferes with the incident radiation" (p. 167, our emphasis). A few pages later, they talk about the same "secondary wavelets set up by each of the illuminated atoms" (ibid., p. 172) without labeling them virtual. On the following page they suddenly refer to the external radiation as "incident *virtual* radiation" (ibid., p. 173, our emphasis). And the final paragraph of the paper discusses the "(virtual) radiation field" (ibid., p. 175) produced by ordinary antennas. The concluding sentence, which has Bohr written all over it, shows how the authors struggled with their own terminology:

It will in this connexion be observed that the emphasizing of the 'virtual' character of the radiation field, which at the present state of science seems so essential for an adequate description of atomic phenomena, automatically loses its importance in a limiting case like that just considered [i.e., a classical antenna], where the field, as regards its observable interaction with matter, is endowed with all the attributes of an electromagnetic field in classical electrodynamics (Bohr, Kramers, and Slater, 1924a, p. 175).

---

[140] See p. 2 of the transcript of session 4 of the AHQP interview with Heisenberg

[141] See p. 21 of the transcript of session 6 of the AHQP interview with Heisenberg. A very similar discussion of BKS can be found in an essay, "The history of quantum theory" (Heisenberg, 1958, pp. 40–41).

Subsequent expositions of BKS by Slater and Kramers removed much of the tentativeness of this passage.

In a lengthy paper signed December 1, 1924, and published in the April 1925 issue of *The Physical Review*, Slater tried to work out a "consistent detailed theory of optical phenomena" based on BKS (Slater, 1925a, p. 395). Slater presented this work at a meeting of the *American Physical Society* in Washington, D.C., in December 1924 (Slater, 1925b). At this same meeting—which also marked the end of the controversy between Compton and Harvard's William Duane (1872–1935) over the Compton effect (Stuewer, 1975, p. 273)—Van Vleck (1925) talked about (Van Vleck, 1924b,c) and Breit (1925) talked about (Breit, 1924a).[142] Slater sent a copy of his paper to Bohr in December 1924 and defended his elaboration of BKS in a letter to Bohr of January 6, 1925 (Bohr, 1972–1996, Vol. 5, pp. 65–66).

In the introduction of his paper, Slater presents the dilemma that led him to embrace Bohr's statistical conservation laws.[143] The problem, he argues, is that

> in the quantum theory the energy of atoms must change by jumps; and in the electromagnetic theory the energy of a radiation field must change continuously…Two paths of escape from this difficulty have been followed with more or less success. The first is to redefine energy [i.e., to adopt Einstein's light-quantum hypothesis]; the second to discard conservation. Optical theory on [the first interpretation] would be a set of laws telling in what paths the quanta travel…[One way to do this is] to set up a sort of ghost field, similar to the classical field, whose function was in some way to guide the quanta. For example, the quanta might travel in the direction of Poynting's vector in such a field. The author was at one time of the opinion that this method was the most hopeful one for solving the problem…The other direction of escape from the conflict between quantum theory and wave theory has been to retain intact the quantum theory and as much of the wave theory as relates to the field, but to discard conservation of energy in the interaction between them (Slater, 1925a, pp. 396–397).

Slater sketches some difficulties facing this second approach, but makes it clear that this is the approach he now favors:

> An attempt was made by the writer, in a note to Nature [Slater, 1924], enlarged upon in collaboration with Bohr and Kramers, to contribute slightly to the solution of these difficulties. In the present paper, the suggestions made in those papers are developed into a more specific theory (ibid., p. 398).

Slater then describes more carefully how to picture the interaction between matter and radiation in BKS and makes it clear that the proposed mechanism is incompatible with strict energy conservation. According to Slater, the "one…essentially new" suggestion of BKS (note that he does not claim credit for the concept of virtual oscillators) was:

---

[142] The AHQP contains some correspondence between Slater and Van Vleck regarding this meeting and regarding (Slater, 1925a): Slater to Van Vleck, December 8, 1924; Van Vleck to Slater, December 15, 1924.

[143] See also the brief discussion of BKS in (Van Vleck, 1926, pp. 285–286).

that the wavelets sent out by an atom in connection with a given transition were sent out, not as a consequence of the occurrence of the transition, but as a consequence of the existence of the atom in the stationary state from which it could make that transition.[144] On this assumption, the stationary state is the time during which the atom is radiating or absorbing; the transition from one state to another is not accompanied by radiation, but so far as the field is concerned, merely marks the end of the radiation or absorption characteristic of one state, and the beginning of that characteristic of another. The radiation emitted or absorbed during the stationary state is further not merely of the particular frequency connected with the transition which the atom is going to make; it includes all the frequencies connected with all the transitions which the atom could make …Although the atom is radiating or absorbing during the stationary states, its own energy does not vary, but changes only discontinuously at transitions…It is quite obvious that the mechanism becomes possible only by discarding conservation (ibid., pp. 397–398).

On the next page, Slater inserts a disclaimer similar to the one by Van Vleck quoted above:

It must be admitted that a theory of the kind suggested has unattractive features; there is an apparent duplication between the atoms on the one hand, and the mechanism of oscillators producing the field on the other. But this duplication seems to be indicated by the experimental facts, and it is difficult at the present stage to see how it is to be avoided (ibid., p. 399).

Slater's portrayal of BKS agrees with the exposition given by Kramers and Helge Holst (1871–1944) in the German edition (Kramers and Holst, 1925) of a popular book on Bohr's atomic theory originally published in Danish (Kramers and Holst, 1922).[145] In a section, entitled "Bohr's new conception of the fundamental postulates," that was added to the German edition, Kramers explained that BKS breaks with one of the basic tenets of Bohr's original theory, namely that atoms only emit light when one of its electrons makes a transition from, to use his example, the second to the first stationary state. "According to the new conception," Kramers wrote, "radiation with frequency ν_{2-1} is still tied to the *possibility* of a transition to the first state, but it is assumed that the emission takes place during the entire time the atom is in the second state" (Kramers and Holst, 1925, p. 135). Another difference is that "if the atom is in the third state, it will simultaneously emit the frequencies ν_{3-2} and ν_{3-1} until it either jumps to the second or to the first state" (ibid.). Kramers emphasizes that this makes

---

[144] Note the similarity with the comments of Bohr to Ladenburg quoted in Sect. 3.3: "the quantum jumps are not the direct cause of the absorption of radiation, but…represent an effect which accompanies the continuously dispersing (and absorbing) effect of the atom on the radiation" (Bohr, 1972–1996, Vol. 5, p. 400).

[145] The translation was done by Fritz Arndt (1885–1969), a chemist and a colleague of Ladenburg and Reiche in Breslau (see the correspondence between Kramers and Ladenburg of 1923–1925 in the AHQP). The preface of this translation is dated March 1925. Dresden (1987) writes that the treatment of BKS in this book "is without much doubt the most understandable exposition of the BKS ideas" (p. 195).

the new conception preferable to Bohr's original one from the point of view of the correspondence principle:

> This situation shows that the new conception is closer to the classical electron theory than the old one; the simultaneous emission of two frequencies mentioned above has its counterpart in that an electron moving on an ellipse emits both its fundamental tone and its first overtone…while earlier one had to assume that these two frequencies were produced by *different* transitions in *different* atoms. It is a welcome consequence, especially from the point of view of the correspondence principle, that the radiation emitted by a *single* atom contains all the frequencies that correspond to possible transitions; for in the border region of large quantum numbers the radiation demanded by the quantum theory will now merge very smoothly with the radiation demanded by the classical theory (Kramers and Holst, 1925, pp. 135–136).

The final paragraph of the BKS paper itself, from which we quoted above, can be seen as a garbled version of Kramers' argument here. Note that the term 'virtual radiation' is absent from these expositions by Slater and Kramers. In his detailed critique of the physics of BKS, Dresden (1987) struggles mightily to make sense of the "somewhat vague, tenuous relation between the virtual field and the real electromagnetic field" (p. 179). The presentations of BKS by Slater and Kramers suggest that there is no fundamental difference between the two. BKS does not introduce two different kinds of radiation, real and virtual, but a new picture of the *interaction* between radiation and matter, which is different both from the classical picture and from Einstein's light-quantum picture. As Heisenberg put it in his AHQP interview (see the passage quoted above), radiation is a "half reality" in this new picture in that it only determines the *probabilities* of quantum transitions in matter.

4.2 The demise of BKS

The BKS theory was decisively refuted in experiments by Walther Bothe (1891–1957) and Hans Geiger (1882–1945) in Berlin and by Compton and Alfred Walter Simon in Chicago. These experiments showed that energy-momentum is strictly conserved in Compton scattering (i.e., event by event) and not just statistically (Stuewer, 1975, pp. 299–302; Stolzenburg, 1984, pp. 75–80). The detection of a scattered electron almost always coincided with the detection of a light quantum, which went against the BKS picture that light is emitted and absorbed continuously, whereas the electron changes its energy and momentum only at discrete intervals. Of course, radiation is detected via its effect on electrons in some detector and, in the BKS picture, radiation only determines the probability of an electron absorbing energy. The crucial difference between BKS and the light-quantum prediction is that according to the latter there is a perfect correlation between detection of a scattered electron and detection of scattered X-rays whereas the former predicts no such correlation. The experiments that eventually disproved BKS were begun shortly after the BKS paper was published (see Bothe and Geiger, 1924), but the final verdict did not come in until the following year. Bothe and Geiger (1925a,b) published their results in April 1925. The paper

by Compton and Simon (1925) is signed June 23, 1925, and appeared in September 1925.[146] On April 17, 1925, Geiger sent Bohr a letter forewarning him of the results of his experiments with Bothe. When Geiger's letter arrived in Copenhagen four days later, Bohr was in the process of writing to Ralph H. Fowler (1889–1944) in Cambridge. In the postscript to this letter, Bohr conceded that "there is nothing else to do than to give our revolutionary efforts as honourable a funeral as possible" (Stuewer, 1975, p. 301). His co-authors Kramers and Slater took the fall of BKS harder. So did other supporters of the theory, such as Ladenburg, Reiche, and Born. By contrast, Einstein and Pauli, the theory's most vocal critics, rejoiced. As we shall see, Born, Pauli, and Van Vleck all explicitly recognized that the demise of BKS did not affect Kramers' dispersion theory and its virtual oscillators.

Ladenburg and Reiche had first read (the German version of) the BKS paper (Bohr, Kramers, and Slater, 1924b) in May 1924. "We are pleased," Ladenburg wrote to Kramers, "that our considerations harmonize so well with your ideas."[147] In the same letter, Ladenburg invited Kramers to come to Breslau to give a talk and to discuss in person what the two of them and Reiche had been discussing in correspondence (see Sect. 3.4). Kramers accepted the invitation and suggested he talk about the new radiation theory, "which, I hope, will soon meet with approval from most physicists (although I heard that Einstein has expressed a relatively unfavorable opinion)."[148] Less than a week later, Kramers received the following intelligence from Ladenburg, directly addressing his parenthetical remark:

> As far as Einstein's opinion about your new conception of radiation is concerned, I can give you a very precise report, since I attended his talk on May 28 in the Berlin colloquium. His opinion was decidedly not unfavorable. He declared the new conception to be internally fully consistent and not in direct contradiction with any facts. The mechanism of the undulatory theory would have to be preserved in his opinion. He put great emphasis, however, on the conceptual logical difficulties of the new theory, of the "preestablished harmony," which the fundamental introduction of probability instead of causality brings with it. Specific objections that he raised seemed to rest only on a not yet complete knowledge of all your considerations. He pointed to the asymmetry, for instance, that the production of virtual radiation was tied to a specific atomic state. In discussion, I pointed out in response to this that the virtual oscillators have the frequencies of *possible* transitions—at which point he immediately withdrew the objection.[149]

Privately, Einstein was less guarded. A month earlier—in a letter to Born and his wife Hedi (1892–1972) of April 29, 1924—he had already delivered his oft-quoted

---

[146] Stuewer (1975, p. 301) draws attention to a footnote in this paper that makes it clear that the experiment had been discussed even before Slater's arrival in Copenhagen: "The possibility of such a test was suggested by W. F. G. Swann in conversation with Bohr and one of us [Compton] in November 1923" (Compton and Simon, 1925, p. 290, note 6). Swann, the reader may recall, had just started in Chicago that fall, leaving the vacancy in Minnesota that was filled by Breit and Van Vleck (see Sect. 2.2).

[147] Ladenburg to Kramers, May 31, 1924 (AHQP).

[148] Kramers to Ladenburg, June 5, 1924 (AHQP).

[149] Ladenburg to Kramers, June 8, 1924 (AHQP).

put-down that, should BKS turn out to be correct, he "would rather have been a shoe-maker or even an employee in a gambling casino than a physicist" (Klein, 1970, p. 32).[150] Talking to Kramers in late June, Einstein expressed himself more diplomatically again. Kramers stopped in Berlin on his return trip from Breslau, where he had given a well-received talk on BKS on June 24, 1924. As he reported to Ladenburg once he was back in Copenhagen: "It was very interesting to hear Einstein's considerations; as he himself says, they are all arguments based on intuition."[151]

Ladenburg also attended the colloquium in Berlin in May 1925 in which Bothe and Geiger presented their results. Ladenburg had just received a copy of the German edition of Kramers' popular book with Holst from which we quoted above. He clearly had a hard time accepting the refutation of BKS at this point. Referring to the discussion of BKS in chap. 6 of (Kramers and Holst, 1925), he wrote:

> In this connection, I must report to you that yesterday Geiger and Bothe presented their important and beautiful experiments on counting electrons and [light] quanta in the Compton effect. Apparently, as you know, they have shown that the emission of electrons and quanta is simultaneous within one-thousandth of a second or less. Can I ask you to what extent you and Bohr consider this as standing in contradiction to your theory? Does your theory really require the complete independence of these two processes, so that only chance could cause the simultaneous occurrence of the two processes within one-thousandth of a second? You can imagine how these questions also affect us and if you have time to write to me to give your opinion I would be very grateful.[152]

Unfortunately, we do not know whether and, if so, how Kramers replied.

When Slater found out about the experimental refutation of BKS, he dashed off another letter to *Nature* (dated July 25, 1925) announcing that he had once more changed his mind: "The simplest solution to the radiation problem then seems to be to return to the view of a virtual field to guide corpuscular quanta" (Slater, 1925c). Kramers and Bohr concurred: "we think that Slater's original hypothesis contains a good deal of truth."[153] Slater thus reverted to the position that, as he reminds the reader, he had been talked out of by Bohr and Kramers. Slater also noted that Swann had argued for this view during the December 1924 meeting of the *American Association for the Advancement of Science*, unaware that he, Slater, had been thinking along the same lines.[154] The following year, Bohr mentioned in passing in a letter to Slater that he had "a bad conscience in persuading you to our view." Slater told him not to worry about it.[155]

The way in which the BKS paper had come to be written, however, had left Slater with a bitter taste in his mouth (Schweber, 1990, pp. 350–356). We already quoted from

---

[150] For further discussion of Einstein's objections to BKS, see (Klein, 1970, pp. 32–35), (Wasserman, 1981, pp. 255–263), and (Stolzenburg, 1984, pp. 24–28, pp. 31–34).

[151] Kramers to Ladenburg, July 3, 1924 (AHQP).

[152] Ladenburg to Kramers, May 15, 1925 (AHQP).

[153] Kramers to Urey, July 16, 1925, quoted by Stolzenburg (1984, p. 86).

[154] Cf. (Swann, 1925). See (Stuewer, 1975, pp. 321–322) for discussion of Swann's proposal.

[155] Bohr to Slater, January 28, 1926; Slater to Bohr, May 27, 1926 (Bohr, 1972–1996, Vol. 5, pp. 68–69).

his letter to Van Vleck of July 27, 1924, in which his disenchantment with Copenhagen shines through very brightly (see Sects. 2.2 and 3.4). Interestingly, on that very same day, Slater wrote to Bohr, thanking him for his "great kindness and attention to me while I was in Copenhagen. Even if we did have some disagreements, I felt very well repaid for my time there, and I look back to it very pleasantly" (Bohr, 1972–1996, Vol. 5, p. 494). This sounds disingenuous in view of his comments to Van Vleck, but Slater had also been very positive about Bohr writing to his Harvard teacher Percy Bridgman (1882–1961) on February 1, 1924 (Schweber, 1990, p. 354). In his AHQP interview, however, Slater was very negative about Bohr and his institute. In fact, when he found out that Copenhagen would be one of the depositories for the AHQP materials, Slater asked Kuhn to keep the interview out of the copy going to Denmark.[156]

Initially, Slater was angry with both Bohr and Kramers, but his attitude toward the latter later softened (Dresden, 1987, pp. 168–171). His wife, fellow-physicist Rose Mooney (1902–1981), may have had something to do with that (Dresden, 1987, pp. 527–528).[157] Before Ms. Mooney became Mrs. Slater in 1948, she had been close to Kramers, whom she had met at a summer school in Michigan in 1938. The two of them almost certainly had an affair. By the late 1930s Kramers was unhappy in his marriage to Anna 'Storm' Petersen, a Danish singer he had met in artistic circles in Copenhagen and married in 1920 after she got pregnant.[158] In one of the most memorable passages of his book, Dresden (1987, pp. 289–295) reveals that Kramers had told Storm many years after the fact that he himself had on at least one occasion been railroaded by Bohr. Kramers apparently thought of the Compton effect around 1920, well before Compton and Debye did. Bohr, however, detested the notion of light quanta so much that he worked on Kramers until he recanted. According to what Storm told Dresden, Kramers had to be hospitalized after one of these sessions with Bohr! Bohr's victory was complete. Even more strongly than Slater in the case of BKS a few years later, Kramers joined Bohr's crusade against light quanta with "all the passion of a repentant convert" (Dresden, 1987, p. 171).[159] Slater may well have found out about this episode from his wife, Kramers' former mistress. Whether or not he did, in his autobiography, as Dresden (1987, p. 528) points out, Slater (1975) refers to his BKS co-author as "my old friend Kramers" (p. 233).

---

[156] Slater to Kuhn, November 22, 1963, included in the folder on Slater in the AHQP.

[157] A caveat is in order here. As pointed out in a review of (Dresden, 1987), "[t]he wealth of intimate detail about Kramers that Dresden provides relies so heavily on personal interviews (Dresden himself notes the "'soft' character" of this information) that it is difficult for others to assess the evidence until the interviews (which I hope were taped), as well as Kramers's personal papers, are made available to others" (Stachel, 1988, p. 745).

[158] Kramers was on the rebound at the time from the on-again-off-again relationship with his Dutch girlfriend, Waldi van Eck. Dresden's description of Kramers' relationship with van Eck (not to be confused with Van Vleck) conjures up the image of a virtual oscillator: "no commitments were made, no decisions were taken, the relationship was never defined, it was certainly never consummated, nor ever terminated" (Dresden, 1987, p. 525).

[159] Bohr apparently commiserated with Pauli a few years later about Kramers' lingering bitterness over this episode. Pauli later told his colleague Res Jost (1918–1990) at the ETH in Zurich that he had consoled Bohr by arguing that discovering the Compton effect was hardly an impressive feat since Compton and Debye had come up with it independently of one another (Dresden, 1987, p. 294).

Born had also been a supporter of BKS. With only Kramers' *Nature* notes to go on, he assumed that Kramers' dispersion theory was a product of BKS. He had no way of knowing that Kramers had obtained these results before BKS. By the time (Born, 1924) was published, however, Born realized that one did not have to subscribe to all articles of the BKS philosophy to extend the results of Kramers' dispersion theory. At the beginning of the paper, Born still writes as if the two stand or fall together:

> Recently…considerable progress has been made by Bohr, Kramers and Slater on just this matter of the connection between radiation and atomic structure…- How fruitful these ideas are, is also shown by Kramers' success in setting up a dispersion formula…In this situation, one might consider whether it would not be possible to extend Kramers' ideas, which he applied so successfully to the interaction between radiation field and radiating electron, to the case of the interaction between several electrons of an atom…The present paper is an attempt to carry out this idea (Born, 1924, pp. 181–182).

A footnote appended to this passage reads: "By a happy coincidence I was able to discuss the contents of this paper with Mr. Niels Bohr, which contributed greatly to a clarification of the concepts." Bohr had visited Born and Heisenberg in Göttingen in early June 1924 (Cassidy, 1991, pp. 177–179). Heisenberg had already told Born all about BKS and Born had expressed his admiration for the theory in a letter to Bohr of April 16, 1924.[160] Bohr's visit must have further solidified his enthusiasm. A week later, however, Einstein passed through town and trashed BKS.[161] As a result of Einstein's onslaught, Born hedged his bets and did not throw in his fate with the more controversial aspects of BKS (see Mehra and Rechenberg, 1982–2001, Vol. 2, p. 144; Cassidy, 1991, p. 179). At the beginning of Sect. 3 of his paper, he writes:

> it will be profitable to make use of the intuitive ideas, introduced by Bohr, Kramers and Slater…but our line of reasoning will be independent of the critically important and still disputed conceptual framework of that theory, such as the statistical interpretation of energy and momentum transfer (Born, 1924, p. 189).[162]

Born, however, continued to be a true believer in BKS and took its collapse harder than Bohr himself. On April 24, 1925, he wrote to Bohr:

> Today Franck showed me your letter [of April 21, 1925, the day that Bohr had received word from Geiger about the results of the Bothe–Geiger experiment]…which interested me exceedingly and indeed almost shocked me,

---

[160] See (Bohr, 1972–1996, Vol. 5, p. 299), discussed in (Mehra and Rechenberg, 1982–2001, Vol. 2, p. 143).

[161] See Heisenberg to Pauli, June 8, 1924 (Pauli, 1979, Doc. 62). This is the same day that Ladenburg wrote to Kramers that Einstein's opinion of BKS was "decidedly not unfavorable" (see above).

[162] This illustrates the importance of what Beller (1999) has called the "dialogical approach" to the history of quantum mechanics (an approach adopted *avant la lettre* by Hendry [1984]): to resolve the tension between the two quoted passages in Born's paper, it is important to be attuned to the voices of both Bohr and Einstein in his text.

because in it you abandon the radiation theory that obeyed no conservation laws (Bohr, 1972–1996, Vol. 5, p. 84).

In contrast to Born, Pauli called the demise of BKS "a magnificent stroke of luck."[163] Pauli's opposition to BKS was probably fueled by Einstein, who gave him an earful about the theory during the annual meeting of the *Gesellschaft Deutscher Naturforscher und Ärzte* in Innsbruck in September 1924.[164] Pauli clearly recognized that Kramers' dispersion theory was independent of BKS and that the fall of the latter did not affect the former. A footnote in (Pauli, 1925) emphasizes

> that the formulae of [Kramers and Heisenberg, 1925] used here are independent of the special theoretical interpretation concerning the detailed description of the radiation phenomena in the quantum theory taken as a basis by them [i.e., BKS], since these formulas only apply to averages over a large number of elementary phenomena (Pauli, 1925, p. 5).

As he explained to Kramers, Pauli wanted to distance himself from the suggestion in the abstract of (Kramers and Heisenberg, 1925) that "the conclusions, should they be confirmed, would form an interesting support for this [i.e., the BKS] interpretation" (cf. Sect. 4.1). Alerting Kramers to the footnote quoted above, Pauli wrote:

> if I had not added the footnote in question, it would also have been true that the conclusions of *my* paper, if they should be confirmed, 'would form an interesting support for this interpretation.' This impression I had, of course, to counteract![165]

This letter was written after Pauli had read the manuscript of Heisenberg's *Umdeutung* paper, which was much more to his liking. In the same letter, in cruel Pauli fashion, he berated Kramers for pushing BKS. That this did not affect Pauli's appreciation for Kramers' work on dispersion is clear from what he wrote to another correspondent a few months after this scathing letter: "[m]any greetings also to Kramers, whom I am very fond of after all, especially when I think of his beautiful dispersion formula."[166]

In his NRC *Bulletin*, written after the Bothe–Geiger and Compton–Simon experiments, Van Vleck, like Pauli, stressed the independence of the Kramers dispersion theory and BKS. The rejection of BKS and the acceptance of the light-quantum hypothesis, he wrote

> [do] not mean that Slater's concept of virtual oscillators is not a useful one. We may assume that the fields which guide the light-quants come from a hypothetical set of oscillators rather than from the actual electron orbits of the conventional electrodynamics.[167] In this way the appearance of the spectroscopic rather than

---

[163] Pauli to Kramers, July 27, 1925 (Pauli, 1979, pp. 232–234; Bohr, 1972–1996, Vol. 5, p. 87).

[164] See Pauli to Bohr, October 2, 1924 (Pauli, 1979, Doc. 66), quoted and discussed in (Wasserman, 1981, pp. 260–263).

[165] Pauli to Kramers, July 27, 1925 (cf. note 163).

[166] Pauli to Kronig, October 9, 1925, quoted in (Stolzenburg, 1984, p. 91).

[167] At this point, the following footnote is appended: "This viewpoint has been advocated by Slater during the printing of the present Bulletin. See [Slater, 1925a]."

the orbital frequency in dispersion can be explained, and the essential features of the virtual oscillator theory of dispersion…can still be retained. There is an exact conservation of energy between the atoms and the actual corpuscular light-quants, but only a statistical conservation of energy between the atoms and the hypothetical virtual fields (Van Vleck, 1926, pp. 286–287).

Virtual oscillators survived the demise of BKS and happily lived on in the dispersion theory from which they originated.

These observations by Pauli and Van Vleck make it clear that BKS only played a limited role in the developments that led to matrix mechanics. It is important to keep that in mind. As long as we think of the Kramers dispersion theory as part and parcel of BKS, it looks as if matrix mechanics replaced a decisively refuted theory. Once we recognize that the Kramers dispersion theory was developed before and independently of BKS, we see that matrix mechanics grew naturally out of an eminently successful earlier theory. The BKS theory and its refutation by the Bothe–Geiger and Compton–Simon experiments then become a sideshow distracting from the main plot line, which runs directly from dispersion theory to matrix mechanics. A corollary to this last observation is that the acceptance of the light-quantum hypothesis was irrelevant to the development of matrix mechanics. Compton scattering provided convincing evidence for the light-quantum hypothesis and against BKS, but it had no bearing on dispersion theory. The work of Ladenburg, Kramers, Born, and Van Vleck crucially depended on Einstein's A and B coefficients, but not on the theory of light quanta in which these coefficients were originally introduced.

4.3 Heisenberg, BKS, and virtual oscillators

When Heisenberg first read the BKS paper, he was not impressed: "Bohr's paper on radiation is certainly very interesting; but I do not really see any fundamental progress."[168] He subsequently warmed to the theory, writing to Copenhagen on April 6, 1924 that he hoped Bohr had meanwhile convinced Pauli.[169] To Sommerfeld he wrote on November 18, 1924: "Maybe Bohr's radiation theory is a most felicitous [*sehr glücklicke*] description of this dualism [i.e., the wave-particle duality of radiation] after all" (Sommerfeld, 2004, p. 174, quoted in Wasserman, 1981, p. 251). Five years later, Heisenberg was praising BKS effusively:

This investigation represented the real high point in the crisis of quantum theory, and, although it could not overcome the difficulties, it contributed, more than any other work of that time, to the clarification of the situation in quantum theory (Heisenberg, 1929, p. 492; translated and quoted in Stuewer, 1975, p. 291).

And thirty years later, Heisenberg (1955, p. 12) remembered BKS as "the first serious attempt to resolve the paradoxes of radiation into rational physics" (quoted in Klein, 1970, p. 37).

---

[168] Heisenberg to Pauli, March 4, 1924 (Pauli, 1979, Doc. 57); quoted by Dresden (1987, p. 202) and Wasserman (1981, p. 250).

[169] See (Bohr, 1972–1996, Vol. 5, pp. 354–355), cited by Cassidy (1991) to support his claim that "by the end of his March 1924 visit to Copenhagen, Werner was a convert" (p. 176).

Why was Heisenberg so taken with BKS? We already came across part of the answer. As he told Kuhn in his AHQP interview, Heisenberg saw in BKS a precursor to the Born interpretation of the Schrödinger wave function (see Sect. 4.1). This, we feel, mainly helps explain Heisenberg's profuse praise after the fact. In the same interview, however, Heisenberg identified another aspect of BKS that can account for his enthusiasm for BKS before *Umdeutung*—or rather, Kuhn identified it for him. What triggered Heisenberg's ruminations on probability in BKS and in the Born interpretation was the observation by Kuhn that despite the experimental refutation of BKS, "a large part of the basic ideas and the whole use of the Correspondence Principle formulated in terms of virtual oscillators goes on quite unshaken."[170] Heisenberg's response does not address this issue at all, whereupon Kuhn tries again: "In order to do that paper [BKS] one talks not only about…probability…but also transforms one's idea of the atom into a collection of virtual oscillators that operate between states" (ibid., p. 3). This time Heisenberg takes the bait:

> Yes, that was it. This idea, of course, also was there already that an atom was really a collection of virtual oscillators. Now this…was in some way contrary to the idea of an electron moving around a nucleus. The obvious connection, the only possible connection, was that the Fourier components of this motion in some way corresponded, as Bohr said, to the oscillators. But certainly this paper [BKS] then prepared the way for this later idea that the assembly of oscillators is nothing but a matrix. For instance, we can simply say that matrix elements are the collection of oscillators. In this way, you can say that matrix mechanics was already contained in this paper [BKS] (ibid., p. 3).

This supports the thesis in (MacKinnon, 1977) mentioned in Sect. 4.2 that matrix mechanics can be seen as a theory of virtual oscillators. What we want to emphasize is that what initially seems to have attracted Heisenberg to BKS was the notion of virtual oscillators. Given the origin of this concept, Heisenberg's intellectual debt on this point was not to BKS but—once again (see Sect. 3.5)—to dispersion theory. During a subsequent session of the AHQP interview, Heisenberg, in fact, talks about the link between Fourier components and oscillators in the context of Kramers' dispersion theory. "When you say the dispersion formula started from a physical idea," Kuhn asked, "do you have a particular thing in mind?" Heisenberg replied:

> Well, I would say that his [i.e., Kramers'] idea was that there was the Einstein paper [with the *A* and *B* coefficients] and there was the Ladenburg [1921] paper connected with Einstein's. On the other hand there was Bohr's Correspondence Principle and the idea finally that this has to do somehow with Fourier components as oscillators. Kramers had the force to combine these two possibilities in one simple formula—the dispersion formula. And this I think was a very important idea that one should combine the Einstein paper, which was very far from the Bohr model[,] with the Bohr model…Behind this idea was already the idea

---

[170] See p. 2 of the transcript of session 4 of the AHQP interview with Heisenberg.

of connecting the oscillators with the Fourier components, which, as I have said many times, was in the air somehow in these years.[171]

Heisenberg explicitly availed himself of virtual oscillators in (Heisenberg, 1925a), a paper on the polarization of fluorescent light submitted from Copenhagen in November 1924 (i.e., before the Kramers–Heisenberg paper). Talking about this paper in his interview with Kuhn, Heisenberg said:

> I would say that all this is part of the game to make the total table of linear oscillators be the real picture of the atom. One felt that in the Correspondence Principle, one should compare one of these linear oscillators with one Fourier component of a motion …So the whole thing was a program which one had consciously or unconsciously in one's mind. That is, how can we actually replace everywhere the orbits of the electron by the Fourier components and thereby get into better touch with what happens? Well, that was the main idea of quantum mechanics later on. One could see, more and more clearly, that the reality were the Fourier components and not the orbits.[172]

MacKinnon (1977, pp. 148–155) stresses the importance of (Heisenberg, 1925a) for the development of matrix mechanics.[173] Heisenberg agreed. Commenting on a draft of MacKinnon's article, he wrote to the author in July 1974: "I was especially glad to see that you noticed how important the paper on the polarization of fluorescent light has been for my further work on quantum mechanics. Actually, in Copenhagen I felt that this paper contained the first step in which I could go beyond the views of Bohr and Kramers" (MacKinnon, 1977, p. 149, note 29). As he proudly recounts in his AHQP interview (see pp. 13–14 of the transcript of session 4), Heisenberg managed to convince Bohr and Kramers of his approach to this problem, an approach they initially questioned.

MacKinnon (1977, pp. 157–162) also sees (Heisenberg, 1925b) on the anomalous Zeeman effect as an important step on the way to matrix mechanics:

> In the conclusion Heisenberg outlined a new program for quantum theory. One should use the virtual oscillator model to work out all the Fourier components for the electrons in an atom and for the coupling between electrons. In the rest of this article I will attempt to trace through in detail the way Heisenberg implemented this program and developed quantum mechanics (MacKinnon, 1977, pp. 161–162).

Here we part company with MacKinnon. Virtual oscillators are not mentioned at all in (Heisenberg, 1925b) (though Fourier components are). (Heisenberg, 1925b, p. 857) does not even refer to virtual oscillators when discussing results pertaining to incoherent radiation from (Kramers and Heisenberg, 1925). This paper, far from

---

[171] See p. 13 of the transcript of session 6 of the AHQP interview with Heisenberg.

[172] See p. 15 of the transcript of session 4 of the AHQP interview with Heisenberg. Parts of this passage are quoted in (MacKinnon, 1977, p. 155) and in (Mehra and Rechenberg, 1982–2001, Vol. 2, p. 165) (although the latter cite their own conversations with Heisenberg as their source; cf. notes 5 and 79).

[173] For other historical discussions of (Heisenberg, 1925a), see (Cassidy, 1991, pp. 187–188) and (Mehra and Rechenberg, 1982–2001, Vol. 2, pp. 159–169).

being another step toward *Umdeutung*, seems to be mired in the intractable problems of the old quantum theory: the Zeeman effect, multi-electron atoms, and mysterious factors of 2 later to be accounted for in terms of electron spin.

MacKinnon, in our opinion, thus overstates his case. Yet, even if we discard what he has to say about (Heisenberg, 1925b) on the Zeeman effect, ample evidence remains for his claim that "[t]he virtual oscillator model played an essential role in the process of reasoning that led Heisenberg to the development of quantum mechanics" (MacKinnon, 1977, p. 184). In fact, this thesis is not nearly as controversial as MacKinnon makes it sound. In the entry on Kramers for the *Dictionary of Scientific Biography*, the sober-minded Dutch physicist Hendrik B. G. Casimir (1909–2000) states matter-of-factly: "The notion of virtual oscillators was the starting point of Heisenberg's quantum mechanics—the virtual oscillators became the matrix elements of the coordinates" (Casimir, 1973, p. 492). MacKinnon (1977) claims that "after [the *Umdeutung* paper] was written the virtual oscillator model sunk from sight and never resurfaced" (p. 184). We already noted, however, that the term "substitute oscillators" can still be found in the famous post-*Umdeutung* paper of Born and Jordan (1925b) (see Sect. 3.3). What we did not mention so far is that Landé (1926, p. 456) actually introduced the phrase "virtual orchestras" to describe not BKS but matrix mechanics![174] The imagery, if not exactly the language, of an "orchestra of virtual oscillators" was also used in early popular expositions of matrix mechanics. In a popular book of the 1930s that went through many editions and was endorsed by Max Planck in a short preface, Ernst Zimmer wrote:[175]

> The state of an atom should no longer be described by the unobservable position and momentum of its electrons, but by the measurable frequencies and intensities of its spectral lines …Regardless of the nature of the real musicians who play the optical music of the atoms for us, Heisenberg imagines assistant or auxiliary musicians [*Hilfsmusiker*]: every one plays just one note at a certain volume. Every one of these musicians is represented by a mathematical expression, q_{mn}, which contains the volume and the frequency of the spectral line as in expressions in acoustics familiar to physicists. These auxiliary musicians are lined up in an orchestra [*Kapelle*] according to the initial and final states n and m of the transition under consideration. The mathematician calls such an arrangement a "matrix" (Zimmer, 1934, pp. 161–162).

Zimmer's *Kapelle der Hilfsmusiker* was clearly inspired by Landé's *Ersatzorchester der virtuellen Oszillatoren*. Virtual oscillators thus not only survived the demise of BKS but also the transition to matrix mechanics. In fact, as we shall see in Sect. 7.1, the features captured by the notion of virtual oscillators can still readily be identified in

---

[174] Landé had worked with Heisenberg in 1924 (Cassidy, 1991, p. 177), resulting in a joint paper (Landé and Heisenberg, 1924). In his AHQP interview, Landé nonetheless said that Heisenberg (1925c) had been incomprehensible to him and that it had taken (Born, Heisenberg, and Jordan, 1925) for him to understand matrix mechanics (p. 3 of the transcript of session 5 of the interview; cf. note 11). These comments seem to be colored, however, by lingering resentment. Landé felt strongly that Born should have won the Nobel Prize for his contribution to matrix mechanics and that German anti-Semitism was the only reason he had not.

[175] We are grateful to Jürgen Ehlers for drawing our attention to Zimmer's book.

the formalism of modern quantum mechanics. From the point of view of the quantum theory that emerged in the immediate aftermath of Heisenberg's *Umdeutung* paper, in which the atomic system is quantized but not (as yet) the electromagnetic field, virtual oscillators are nothing but the Fourier components of the Schrödinger wave function of the electron. The perturbing electromagnetic field induces additional Fourier components in this wave function, which in turn results in secondary electromagnetic radiation. In terms of the language borrowed from BKS, this radiation is emitted by virtual oscillators.

References

Aitchison, I. J. R., D. A. McManus, and T. M. Snyder (2004). Understanding Heisenberg's "magical" paper of July 1925: A new look at the calculational details. *American Journal of Physics* 72: 1370–1379.

Anderson, P. W. (1987). John Hasbrouck Van Vleck, March 13, 1899–October 27, 1980. *National Academy of Sciences Biographical Memoirs* 56: 501–540.

Andrade, E. N. da C. (1927). *The structure of the atom.* 3rd ed. London: G. Bells and Sons.

Assmus, A. (1992). The Americanization of molecular physics. *Historical Studies in the Physical and Biological Sciences* 23: 1–34.

Assmus, A. (1999). Edwin C. Kemble. *National Academy of Sciences. Biographical Memoirs* 76: 178–197.

Baltas, A., K. Gavroglu, and V. Kindi (2000). A discussion with Thomas S. Kuhn. Pp. 255–323 in: J. Conant and J. Haugeland (eds.), *The road since Structure.* Chicago: University of Chicago Press.

Barut, A. O., H. Odabasi, A. van der Merwe, eds. (1991). *Selected popular writings of E. U. Condon.* New York, Berlin: Springer.

Becker, R. (1924). Über die Absorption und Dispersion in Bohrs Quantentheorie. *Zeitschrift für Physik* 27: 173–188.

Bederson, B. (2005). Fritz Reiche and the emergency committee in aid of displaced foreign scholars. *Physics in Perspective* 7: 453–472.

Beller, M. (1999). *Quantum dialogue. The making of a revolution.* Chicago: University of Chicago Press.

Bernstein, J. (2004). *Oppenheimer. Portrait of an enigma.* Chicago: Ivan R. Dee.

Birtwistle, G. (1926). *The quantum theory of the atom.* Cambridge: The University Press.

Bohr, N. (1913). On the constitution of atoms and molecules (Part I). *Philosophical Magazine* 26: 1–25.

Bohr, N. (1918). On the Quantum Theory of Line-Spectra. Part 1, On the General Theory. *Det Kongelige Danske Videnskabernes Selskab. Skrifter. Naturvidenskabelig og Matematisk Afdeling* 8, no. 4.1: 336. Reprinted in (Van der Waerden, 1968, pp. 95–136)

Bohr, N. (1922). Der Bau der Atome und die physikalischen und chemischen Eigenschaften der Elemente. *Zeitschrift für Physik* 9: 1–67.

Bohr, N. (1923a). The structure of the atom. (Transl. F. C. Hoyt). *Nature* 112: 29–44. Reprinted in (Bohr, 1972–1996, Vol. 467–482).

Bohr, N. (1923b). Über die Anwendung der Quantentheorie auf den Atombau. I. Die Grundpostulate der Quantentheorie. *Zeitschrift für Physik* 13: 117–165. English translation published as a supplement of *Proceedings of the Cambridge Philosophical Society* (1924, pp. 1–42) and reprinted in (Bohr, 1972–1996, Vol. 3, 457–499).

Bohr, N. (1972–1996). *Collected works.* 9 Vols. Edited by L. Rosenfeld *et al.* Amsterdam: North-Holland.

Bohr, N., H. A. Kramers, and J. C. Slater (1924a). The quantum theory of radiation. *Philosophical Magazine* 47: 785–802. Page references to reprint in (Van der Waerden, 1968, pp. 159–176).

Bohr, N., H. A. Kramers, and J. C. Slater (1924b). Über die Quantentheorie der Strahlung. *Zeitschrift für Physik* 24: 69–87.

Born, M. (1924). Über Quantenmechanik. *Zeitschrift für Physik* 26: 379–395. Page references are to the English translation in (Van der Waerden, 1968, pp. 181–198).

Born, M. (1925). *Vorlesungen über Atommechanik.* Berlin: Springer.

Born, M. (1926). Zur Quantenmechanik der Stoßvorgänge. *Zeitschrift für Physik* 37: 863–867.

Born, M. (1948). Max Karl Ernst Ludwig Planck. *Obituary Notices of Fellows of the Royal Society* 6: 161–180.

Born, M. (1978). *My life. Recollections of a Nobel laureate.* New York: Charles Scribner.

Born, M., W. Heisenberg, and P. Jordan (1925). Zur Quantenmechanik II. *Zeitschrift für Physik* 35: 557–615. English translation in (Van der Waerden, 1968, pp. 321–385).

Born, M., and P. Jordan (1925a). Zur Quantentheorie aperiodischer Vorgänge. I. *Zeitschrift für Physik* 33: 479–505.

Born, M., and P. Jordan (1925b). Zur Quantenmechanik. *Zeitschrift für Physik* 34: 858–888. Page references to chs. 1–3 are to the English translation in (Van der Waerden, 1968, pp. 277–306). Ch. 4 is omitted in this translation.

Born, M., and W. Pauli (1922). Über die Quantelung gestörter mechanisher Systeme. *Zeitschrift für Physik* 10: 137–158.

Bothe, W., and H. Geiger (1924). Ein Weg zur experimentellen Nachprüfung der Theorie von Bohr, Kramers, und Slater. *Zeitschrift für Physik* 25: 44.

Bothe, W., and H. Geiger (1925a). Experimentelles zur Theorie von Bohr, Kramers, und Slater. *Naturwissenschaften* 13: 440–441.

Bothe, W., and H. Geiger (1925b). Über das Wesen des Comptoneffekts: ein experimenteller Beitrag zur Theorie der Strahlung. *Zeitschrift für Physik* 32: 639–663.

Breit, G. (1924a). The polarization of resonance radiation. *Philosophical Magazine* 47: 832–842.

Breit, G. (1924b). The quantum theory of dispersion. *Nature* 114: 310.

Breit, G. (1925). Polarization of resonance radiation and the quantum theory of dispersion. *Physical Review* 25: 242.

Breit, G. (1932). Quantum theory of dispersion. *Reviews of Modern Physics* 4: 504–576.

Buchwald, J. Z. (1985). *From Maxwell to microphysics. Aspects of electromagnetic theory in the last quarter of the nineteenth century.* Chicago: University of Chicago Press.

Cantor, G. N. (1983). *Optics after Newton: theories of light in Britain and Ireland, 1704-1840.* Dover, N.H.: Manchester University Press.

Casimir, H. B. G. (1973). Kramers, Hendrik Anthony. Pp. 491–494 in: C. C. Gillispie (ed.), *Dictionary of scientific biography.* Vol. VII. New York: Charles Scribner's Sons.

Cassidy, D. C. (1991). *Uncertainty. The life and science of Werner Heisenberg.* New York: Freeman.

Charlier, C. L. (1902–1907). *Die Mechanik des Himmels.* 2 Vols. (Vol. 1: 1902; Vol. 2: 1907). Leipzig: Veit.

Coben, S. (1971). The scientific establishment and the transmission of quantum mechanics to the United States, 1919–32. *American Historical Review* 76: 442–466.

Compton, A. H. (1923). A quantum theory of the scattering of X-rays by light elements. *Physical Review* 21: 483–502.

Compton, A. H., and A. W. Simon (1925). Directed quanta of scattered X-rays. *Physical Review* 26: 289–299.

Condon, E. U. (1973). Reminiscences of a life in and out of quantum mechanics. Pp. 314–331 in (Barut et al., 1991).

Cushing, J. T. (1990). *Theory construction and selection in modern physics. The S matrix.* Cambridge: Cambridge University Press.

Darrigol, O. (1992). *From c-numbers to q-numbers: the classical analogy in the history of quantum theory.* Berkeley: University of California Press.

Darrigol, O. (2000). *Electrodynamics from Ampère to Einstein.* Oxford: Oxford University Press.

Darrigol, O. (2002). Quantum theory and atomic structure, 1900–1927. Pp. 331–349 in: M. J. Nye (ed.), *The Cambridge history of science*, Vol. 5, *The modern physical and mathematical sciences*. Cambridge: Cambridge University Press, 2002.

Darwin, C. G. (1922). A quantum theory of optical dispersion. *Nature* 110: 841–842.

Darwin, C. G. (1923). The wave theory and the quantum theory. *Nature* 111: 771–773.

Davisson, C. J. (1916). The dispersion of hydrogen and helium on Bohr's theory. *Physical Review* 8: 20–27.

Davisson, C. J., and L. H. Germer (1927). Diffraction of electrons by a crystal of nickle. *Physical Review* 30: 705–740.

Debye, P. (1915). Die Konstitution des Wasserstoff-moleküls. *Sitzungsberichte der mathematisch-physikalischen Klasse der Königlichen Bayerischen Akademie der Wissenschaften zu München.* 1–26.

Dresden, M. (1987). *H. A. Kramers: between tradition and revolution.* New York: Springer.

Drude, P. (1900). *Lehrbuch der Optik.* Leipzig: S. Hirzel. English transl.: *The theory of optics.* transl.: C. R. Mann and R. A. Millikan. New York: Longmans, Green, 1902.

Eckert, M. (1993). *Die Atomphysiker. Eine Geschichte der theoretischen Physik am Beispiel der Sommerfeldschule.* Braunschweig; Wiesbaden: Vieweg.

Einstein, A. (1905). Zur Elektrodynamik bewegter Körper. *Annalen der Physik* 17: 891–921. Reprinted in facsimile as Doc. 23 in (Einstein, 1987–2006, Vol. 2).

Einstein, A. (1916a). Strahlungs-Emission und -Absorption nach der Quantentheorie. *Deutsche Physikalische Gesellschaft. Verhandlungen* 18: 318–323. Reprinted in facsimile as Doc. 34 in (Einstein, 1987–2006, Vol. 6).

Einstein, A. (1916b). Zur Quantentheorie der Strahlung. *Physikalische Gesellschaft Zürich. Mitteilungen* 18: 47–62. Reprinted as (Einstein, 1917) and (in facsimile) as Doc. 38 in (Einstein, 1987–2006, Vol. 6).

Einstein, A. (1917). Zur Quantentheorie der Strahlung. *Physikalische Zeitschrift* 18: 121–128. Reprint of (Einstein, 1916b). English translation in (Van der Waerden, 1968, pp. 63–77).

Einstein, A. (1987–2006). *The collected papers of Albert Einstein.* 9 Vols. Edited by J. Stachel *et al.* Princeton: Princeton University Press.

Epstein, P. S. (1916). Zur Quantentheorie. *Annalen der Physik* 51: 168–188.

Epstein, P. S. (1922a). Die Störungsrechnung im Dienste der Quantentheorie. I. Eine Methode der Störungsrechnung. *Zeitschrift für Physik* 8: 211–228.

Epstein, P. S. (1922b). Die Störungsrechnung im Dienste der Quantentheorie. II. Die numerische Durchführung der Methode. *Zeitschrift für Physik* 8: 305–320.

Epstein, P. S. (1922c). Die Störungsrechnung im Dienste der Quantentheorie. III. Kritische Bemerkungen zur Dispersionstheorie. *Zeitschrift für Physik* 9: 92–110.

Fellows, F. H. (1985). *J. H. Van Vleck: The early life and work of a mathematical physicist.* Ph.D. Thesis, University of Minnesota.

Feynman, R. P., R. B. Leighton, and M. Sands (1964). *The Feynman lectures on physics.* 3 Vols. Reading, MA: Addison-Wesley.

Forman, P. (1968). The doublet riddle and atomic physics circa 1924. *Isis* 59: 156–174.

Forman, P. (1970). Alfred Landé and the anomalous Zeeman effect. *Historical Studies in the Physical Sciences* 2: 153–261.

Glazebrook, R. T. (1886). Report on optical theories. Pp. 157–261 in *British Association for the Advancement of Science. Report—1885.* London: Spottiswoode.

Gustavson, J. R. (2004). *Wolfgang Pauli 1900 to 1930: His early physics in Jungian perspective.* Ph.D. Thesis, University of Minnesota.

Heilbron, J. L. (1985). Artes compilationis. [Review of (Mehra and Rechenberg, 1982–2001, Vols. 1–4)] *Isis* 76: 388–393.

Heilbron, J. L., and T. S. Kuhn (1969). The genesis of the Bohr atom. *Historical Studies in the Physical Sciences* 1: 211–290.

Heisenberg, W. (1925a). Über eine Anwendung des Korrespondenzprinzips auf die Frage nach der Polarisation des Fluoreszenzlichtes. *Zeitschrift für Physik* 31: 617–626.

Heisenberg, W. (1925b). Zur Quantentheorie der Multiplettstruktur und der anomalen Zeemaneffekte. *Zeitschrift für Physik* 32: 841–860.

Heisenberg, W. (1925c). Über die quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen. *Zeitschrift für Physik* 33: 879–893. Page references to English translation in (Van der Waerden, 1968, pp. 261–276).

Heisenberg, W. (1929). Die Entwicklung der Quantentheorie 1918–1928. *Die Naturwissenschaften* 17: 490–496.

Heisenberg, W. (1943). Die 'beobachtbaren Grössen' in der Theorie der Elementarteilchen. *Zeitschrift für Physik* 120: 513–538.

Heisenberg, W. (1955). The development of the interpretation of the quantum theory. Pp. 12–29 in: W. Pauli (ed.), *Niels Bohr and the development of physics*. London: Pergamon Press.

Heisenberg, W. (1958). *Physics and philosophy. The revolution in modern science.* New York: Harper.

Heisenberg, W. (1971). *Physics and beyond. Encounters and conversations.* New York: Harper & Row. Translation of: *Der Teil und das Ganze. Gespräche im Umkreis der Atomphysik.* Munich: Piper Verlag, 1969.

Heisenberg, W. (1983). *Encounters with Einstein. And other essays on people, places, and particles.* Princeton: Princeton University Press.

Hendry, J. (1981). Bohr–Kramers–Slater: A virtual theory of virtual oscillators and its role in the history of quantum mechanics. *Centaurus* 25: 189–221.

Hendry, J. (1984). *The creation of quantum mechanics and the Bohr–Pauli dialogue.* Dordrecht: Reidel.

Herzfeld, K. F. (1924). Versuch einer quantenhaften Deutung der Dispersion. *Zeitschrift für Physik* 23: 341–360.

Holton, G. (1968). Mach, Einstein and the search for reality. *Daedalus* 97: 636–673. Reprinted as Ch. 7 in *Thematic origins of scientific thought.* Rev. ed. Cambridge: Harvard University Press, 1988.

Holton, G. (1988). On the hesitant rise of quantum physics research in the United States. Pp. 147–187 in: *Thematic origins of scientific thought.* Rev. ed. Cambridge: Harvard University Press.

Holton, G. (2005). Werner Heisenberg and Albert Einstein. Pp. 26–35 in *Victory and vexation in science. Einstein, Bohr, Heisenberg, and others.* Cambridge, MA: Harvard University Press.

Hoyt, F. C. (1923). Intensities of spectral lines. *Philosophical Magazine* 46: 135–145.

Hoyt, F. C. (1924). Relative probabilities of the transitions involved in the Balmer series lines of hydrogen. *Philosophical Magazine* 47: 826–831.

Hoyt, F. C. (1925a). The harmonic analysis of electron orbits. *Physical Review* 25: 174–186

Hoyt, F. C. (1925b). Application of the correspondence principle to relative intensities in series spectra. *Physical Review* 26: 749–760.

Hull, M. (1998). Gregory Breit. *National Academy of Sciences. Biographical Memoirs* 74: 26–57.

Hund, F. (1984). *Geschichte der Quantentheorie.* Darmstadt: Wissenschaftliche Buchgesellschaft.

Ishiwara, J. (1915). Die universelle Bedeutung des Wirkungsquantum. *Toyko Sugaku Buturigakkawi Kizi* 8: 106–116.

Jammer, M. (1966). *The conceptual development of quantum mechanics.* New York: McGraw-Hill.

Janssen, M. (2002). Reconsidering a scientific revolution: the case of Lorentz *versus* Einstein. *Physics in Perspective* 4: 421–446.

Janssen, M., and M. Mecklenburg (2006). From Classical to Relativistic Mechanics: Electromagnetic Models of the Electron. Pp. 65–134 in: V. F. Hendricks, K. F. Jørgensen, J. Lützen, and S. A. Pedersen (eds.), *Interactions: mathematics, physics and philosophy, 1860–1930.* Berlin: Springer.

Janssen, M., and J. Stachel (2004). Optics and Electrodynamics in moving bodies. *Max Planck Institute for the History of Science.* Preprint 265. To appear in: John Stachel, *Going critical* (in preparation); and (in Italian) in: Sandro Petruccioli et al. (eds.), *Storia della scienza*, Istituto della Enciclopedia Italiana (in preparation).

Jordan, P. (1973). Early years of quantum mechanics: some reminiscences. Pp. 294–299 in: J. Mehra (ed.), *The physicist's conception of nature.* Dordrecht: Reidel.

Kemble, E. C. (1921). The probable normal state of the helium atom. *Philosophical Magazine* 42: 123–133.

Kevles, D. J. (1978). *The physicists. The history of a scientific community in modern America.* New York: Knopf.

Klein, M. J., ed. (1967). *Letters on wave mechanics.* New York: Philosophical Library.

Klein, M. J. (1970). The first phase of the Bohr-Einstein dialogue. *Historical Studies in the Physical Sciences* 2: 1–39.

Klein, O. (1967). Glimpses of Niels Bohr as scientist and thinker. Pp. 74–93 in: S. Rozental (ed.), *Niels Bohr. His life and work as seen by his friends and colleagues.* London: Interscience Publishers.

Konno, H. (1993). Kramers' negative dispersion, the virtual oscillator model, and the correspondence principle. *Centaurus* 36: 117–166.

Kragh, H. (1999). *Quantum generations. A history of physics in the twentieth century.* Princeton: Princeton University Press.

Kramers, H. A. (1919). Intensities of spectral lines. *Det Kongelige Danske Videnskabernes Selskab. Skrifter. Naturvidenskabelig og Matematisk Afdeling* 8, no. 3.3: 285–386. Reprinted in (Kramers, 1956, pp. 3–108).

Kramers, H. A. (1923). Über das Modell des Heliumatoms, *Zeitschrift für Physik* 13: 312–341.

Kramers, H. A. (1924a). The law of dispersion and Bohr's theory of spectra. *Nature* 113: 673–676. Page references to reprint in (Van der Waerden, 1968, pp. 177–180).

Kramers, H. A. (1924b). The quantum theory of dispersion. *Nature* 114: 310–311. Page references to reprint in (Van der Waerden, 1968, pp. 199–201).

Kramers, H. A. (1956). *Collected scientific papers*. Edited by H. B. G. Casimir et al. Amsterdam: North Holland.

Kramers, H. A., and W. Heisenberg (1925). Über die Streuung von Strahlung durch Atome. *Zeitschrift für Physik* 31: 681–707. Page references to English translation in (Van der Waerden, 1968, pp. 223–252).

Kramers, H. A., and H. Holst (1922). *Bohrs atomteori: almenfatteligt fremstillet*. Copenhagen: Gyldendal Nordisk forlag.

Kramers, H. A., and H. Holst (1925). *Das Atom und die Bohrsche Theorie seines Baues. Gemeinverständlich dargestellt*. F. Arndt, transl. Berlin: Springer.

Kronig, R. (1960). The turning point. Pp. 5–39 in: M. Fierz and V. F. Weisskopf (eds.), *Theoretical physics in the twentieth century. A memorial volume to Wolfgang Pauli*. New York: Interscience Publishers.

Kuhn, T. S., J. L. Heilbron, P. Forman, and L. Allen (1967). *Sources for the history of quantum physics. An inventory and report*. Philadelphia: American Philosophical Society.

Kuhn, T. S., and J. H. Van Vleck (1950). A simplified method of computing the cohesive energies of monovalent metals. *Physical Review* 79: 382–388.

Kuhn, W. (1925). Über die Gesamtstärke der von einem Zustande ausgehenden Absorptionslinien. *Zeitschrift für Physik* 33: 408–412. English translation in (Van der Waerden, 1968, pp. 253–257).

Ladenburg, R. (1908). Über die Dispersion des leuchtenden Wasserstoffs. *Physikalische Zeitschrift* 9: 875–878.

Ladenburg, R. (1921). Die quantentheoretische Deutung der Zahl der Dispersionselektronen. *Zeitschrift für Physik* 4: 451–468. Page references to English translation in (Van der Waerden, 1968, pp. 139–157).

Ladenburg, R. (1926). Die quantentheoretische Dispersionsformel und ihre experimentelle Prüfung. *Die Naturwissenschaften* 14: 1208–1213.

Ladenburg, R. (1928). Untersuchungen über die anomale Dispersion angeregter Gase. I. Teil. Zur Prüfung der quantentheoretischen Dispersionsformel. *Zeitschrift für Physik* 48: 15–25.

Ladenburg, R., and S. Loria (1908). Über die Dispersion des leuchtenden Wasserstoffs. *Deutsche Physikalische Gesellschaft. Verhandlungen* 10: 858–866. Reprinted in *Physikalische Zeitschrift* 9 (1908): 875–878.

Ladenburg, R., and R. Minkowski (1921). Die Verdampfungswärme des Natriums und die Übergangswahrscheinlichkeiten des Na-Atoms aus dem Resonanz- in der Normalzustand auf Grund optischer Messungen. *Zeitschrift für Physik* 6: 153–164

Ladenburg, R., and F. Reiche (1923). Absorption, Zerstreuung und Dispersion in der Bohrschen Atomtheorie. *Die Naturwissenschaften* 11: 584–598.

Ladenburg, R., and F. Reiche (1924). Dispersionsgesetz und Bohrsche Atomtheorie. *Die Naturwissenschaften* 12: 672–673.

Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. Pp. 91–196 in: I. Lakatos and A. Musgrave (eds.), *Criticism and the growth of knowledge*. Cambridge: Cambridge University Press.

Landé, A. (1926). Neue Wege der Quantentheorie. *Die Naturwissenschaften* 14: 455–458.

Landé, A., and W. Heisenberg (1924). Termstruktur der Multipletts höherer Stufe. *Zeitschrift für Physik* 25: 279–286.

Landsberg, G. and L. Mandelstam (1928). Eine neue Erscheinung bei der Lichtzerstreuung in Krystallen. *Die Naturwissenschaften* 16: 557–558.

Landsman, N. P. (2007). Between classical and quantum. Pp. 417–553 in: J. Butterfield and J. Earman (eds.), *Philosophy of physics*. Part A. Amsterdam: Elsevier.

MacKinnon, E. M. (1977). Heisenberg, models, and the rise of matrix mechanics. *Historical Studies in the Physical Sciences* 8: 137–188.

MacKinnon, E. M. (1982). *Scientific explanation and atomic physics*. Chicago: University of Chicago Press.

Mehra, J., and H. Rechenberg (1982–2001). *The historical development of quantum theory*. 6 Vols. New York, Berlin: Springer.

Merton, R. K. (1968). The Matthew effect in science. *Science* 159 (January 5, 1968): 56–63.

Millikan, R. A. (1916). A direct photoelectric determination of Planck's '*h*'. *Physical Review* 7: 355–388.

Moyer, A. E. (1985). History of phyics. Pp. 163–182 in: S. Gregory Kohlstedt and M. W. Rossiter (eds.), *Historical writing on American science.* Baltimore: Johns Hopkins University Press.

Niessen, K. F. (1924). Ableitung des Planckschen Strahlungsgesetzes für Atome mit zwei Freiheitsgraden. *Annalen der Physik* 75: 743–780.

Oseen, C. W. (1915). Das Bohrsche Atommodell und die Maxwellschen Gleichungen. *Physikalische Zeitschrift* 16: 395–405.

Pais, A. (1986). *Inward bound. Of matter and forces in the physical world.* Oxford: Clarendon Press; New York: Oxford University Press.

Pauli, W. (1925). Ueber die Intensitäten der im elektrischen Feld erscheinenden Kombinationslinien. *Matematisk-fysiske Meddelelser udgivet af Det Kongelige Danske Videnskabernes Selskab (København)* 7, No. 3: 3–20. Reprinted in facsimile in (Pauli, 1964, Vol. 2, pp. 233–250).

Pauli, W. (1926). Quantentheorie. Pp. 1–278 in H. Geiger and K. Scheel (eds.), *Handbuch der Physik,* Vol. 23. Berlin: Springer. Reprinted in facsimile in (Pauli, 1964, Vol. 1, pp. 269–548).

Pauli, W. (1964). *Collected Scientific Papers.* 2 Vols. Edited by R. Kronig and V. F. Weisskopf. New York: Interscience Publishers.

Pauli, W. (1979). *Scientific correspondence with Bohr, Einstein, Heisenberg a.o. Volume I: 1919–1929.* Edited by A. Hermann, K. von Meyenn, and V. F. Weisskopf. Berlin: Springer.

Rabi, I. I. (1975). E. U. Condon: the physicist and the individual. Pp. 4–9 in (Barut *et al.*, 1991).

Rabi, I. I. (2006). Stories from the early days of quantum mechanics. (A colloquium given in Toronto, April 5, 1979, transcribed and edited by R. Fraser Code.) *Physics Today* 59 (8): 36–41.

Raman, C. V. (1928). A new radiation. *Indian Journal of physics* 2: 387–398.

Reiche, F., and W. Thomas (1925). Über die Zahl der Dispersionselektronen, die einem stationären Zustand zugeordnet sind. *Zeitschrift für Physik* 34: 510–525.

Rigden, J. S. (1987). *Rabi, scientist and citizen.* New York: Basic books.

Robertson, P. (1979). *The early years. The Niels Bohr institute.* Copenhagen: Akademisk Forlag.

Schwarzschild, K. (1916). Zur Quantenhypothese. *Königlich Preussische Akademie der Wissenschaften* (Berlin). *Sitzungsberichte* 1916: 548–568.

Schweber, S. S. (1986). The empiricist temper regnant: Theoretical physics in the United States 1920–1950. *Historical Studies in the Physical and Biological Sciences* 17: 55–98.

Schweber, S. S. (1990). The young John Clarke Slater and the development of quantum chemistry. *Historical Studies in the Physical and Biological Sciences* 20: 339–406.

Seidel, R. W. (1978). *Physics research in California: The rise of a leading sector in American physics.* Ph.D. Thesis, University of California, Berkeley.

Serwer, D. (1977). *Unmechanischer Zwang*: Pauli, Heisenberg, and the rejection of the mechanical atom, 1923–1925. *Historical Studies in the Physical Sciences* 8: 189–256.

Shenstone, A. G. (1973). Ladenburg, Rudolf Walther. Pp. 552–556 in: C. C. Gillispie (ed.), *Dictionary of scientific biography*. Vol. VII. New York: Charles Scribner's Sons.

Slater, J. C. (1924). Radiation and atoms. *Nature* 113: 307–308.

Slater, J. C. (1925a). A quantum theory of optical phenomena. *Physical Review* 25: 395–428.

Slater, J. C. (1925b). The nature of resonance radiation. *Physical Review* 25: 242.

Slater, J. C. (1925c). The nature of radiation. *Nature* 116: 278.

Slater, J. C. (1968). Quantum physics in America between the wars. *Physics Today* 21 (1): 43–51.

Slater, J. C. (1973). The development of quantum mechanics in the period 1924–1926. Pp. 19–25 in: W. C. Price, S. S. Chissick, and T. Ravensdale (eds.), *Wave mechanics. The first fifty years.* New York, Toronto: Wiley.

Slater, J. C. (1975). *Solid-state and molecular theory: a scientific biography.* New York, London, Sydney, Toronto: John Wiley & Sons.

Smekal, A. (1923). Zur Quantentheorie der Dispersion. *Die Naturwissenschaften* 11: 873–875.

Sommerfeld, A. (1915a). Zur Theorie der Balmerschen Serie. *Königlich Bayerische Akademie der Wissenschaften* (München). *Sitzungsberichte* 1915: 425–458.

Sommerfeld, A. (1915b). Die allgemeine Dispersionsformel nach dem Bohrschen Model. Pp. 549–584: in K. Bergwitz (ed.), *Festschrift Julius Elster und Hans Geitel.* Braunschweig. Reprinted in (Sommerfeld, 1968, Vol. 3, pp. 136–171).

Sommerfeld, A. (1917). Die Drudesche Dispersionstheorie vom Standpunkte des Bohrschen Modelles und die Konstitution von H_2, O_2, and N_2. *Annalen der Physik* 53: 497–550.

Sommerfeld, A. (1919). *Atombau und Spektrallinien*. 1st ed. Braunschweig: Vieweg.

Sommerfeld, A. (1922). *Atombau und Spektrallinien*. 3rd ed. Braunschweig: Vieweg.

Sommerfeld, A. (1968). *Gesammelte Schriften*. 4 Vols. Edited by F. Sauter. Braunschweig: Vieweg.

Sommerfeld, A. (2004). *Wissenschaftliche Briefwechsel. Band 2: 1919–1951*. Edited by M. Eckert and K. Märker. Berlin, Diepholz, Munich: Deutsches Museum, Verlag für Geschichte der Naturwissenschaften und der Technik.

Sopka, K. R. (1988). *Quantum physics in America. The years through 1935*. Tomash Publishers/American Institute of Physics.

Stachel J. (1988). Inside a Physicist [Review of Dresden, 1987]. *Nature* Vol. 332, No. 6166 (21 April 1988): 744–745.

Stachel J. (2005). Fresnel's (dragging) coefficient as a challenge to 19th century optics of moving bodies. Pp. 1–13 in: J. Eisenstaedt and A. J. Kox (eds.), *Einstein studies*, Vol. 11, *The universe of general relativity*. Boston: Birkhäuser.

Stolzenburg, K. (1984). Introduction. Part 1. The theory of Bohr, Kramers, and Slater. Pp. 3–96 in (Bohr, 1972–1996, Vol. 5).

Stuewer, R. H. (1975). *The Compton effect. Turning point in physics*. New York: Science History Publications.

Swann, W. F. G. (1925). The trend of thought in physics. *Science* 61: 425–435.

Ter Haar, D. (1998). *Master of theory. The scientific contributions of H. A. Kramers*. Princeton: Princeton University Press.

Thomas, W. (1925). Über die Zahl der Dispersionselektronen, die einem stationären Zustande zugeordnet sind (Vorläufige Mitteilung). *Die Naturwissenschaften* 13: 627.

Thorndike Greenspan, N. (2005). *The end of the certain world. The life and science of Max Born*. New York: Basic Books.

Van der Waerden, B. L., ed. (1968). *Sources of quantum mechanics*. New York: Dover.

Van der Waerden, B. L., and H. Rechenberg (1985). Quantum mechanics (1925–1927). Pp. 329–343 in: Werner Heisenberg, *Gesammelte Werke/Collected Works*. Series A/Part I. Berlin: Springer, 1985.

Van Kampen, N. G. (1988). Book review (boekbespreking) of (Dresden, 1987). *Nederlands Tijdschrift voor Natuurkunde* A54 (1): 40

Van Vleck, J. H. (1922a). The dilemma of the helium atom. *Physical Review* 19: 419–420.

Van Vleck, J. H. (1922b). The normal helium atom and its relation to the quantum theory. *Philosophical Magazine* 44: 842–869.

Van Vleck, J. H. (1923). Note on the quantum theory of the helium arc spectrum. *Physical Review* 21: 372–373.

Van Vleck, J. H. (1924a). A correspondence principle for absorption. *Journal of the Optical Society of America* 9: 27–30.

Van Vleck, J. H. (1924b). The absorption of radiation by multiply periodic orbits, and its relation to the correspondence principle and the Rayeigh-Jeans law. Part I. Some extensions of the correspondence principle. *Physical Review* 24: 330–346. Reprinted in (Van der Waerden, 1968, pp. 203–222).

Van Vleck, J. H. (1924c). The absorption of radiation by multiply periodic orbits, and its relation to the correspondence principle and the Rayeigh-Jeans law. Part II. Calculation of absorption by multiply periodic orbits. *Physical Review* 24: 347–365.

Van Vleck, J. H. (1925). Virtual oscillators and scattering in the quantum theory. *Physical Review* 25: 242–243.

Van Vleck, J. H. (1926). *Quantum principles and line spectra*. Washington, D. C.: National Research Council (Bulletin of the National Research Council 10, Part 4).

Van Vleck, J. H. (1929). The new quantum mechanics. *Chemical Reviews* 5: 467–507.

Van Vleck, J. H. (1964). American physics comes of age. *Physics Today* 17 (6): 21–26.

Van Vleck, J. H. (1971). Reminiscences of the first decade of quantum mechanics. *International Journal of Quantum Chemistry. Symposium No. 5, 1971* (a symposium held in honor of Van Vleck). Edited by Per-Olov Lödwin. New York, London, Sydney, Toronto: John Wiley & Sons. Pp. 3–20.

Van Vleck, J. H. (1974). Acceptance speech. *Koninklijke Nederlandse Akademie van Wetenschappen. Bijzondere bijeenkomst der afdeling natuurkunde…28 september 1974…voor de plechtige uitreiking van de Lorentz-medaille aan Prof. Dr. J. H. Van Vleck.*

Van Vleck, J. H. (1992). John Hasbrouck Van Vleck. Pp. 351–252 in: S. Lundqvist (ed.), *Nobel lectures in physics (1971–1980)*. Singapore: World Scientific Publishing.

Van Vleck, J. H., and D. L. Huber (1977). Absorption, emission, and line breadths: A semihistorical perspective. *Reviews of Modern Physics* 49: 939–959.

Verschaffelt, J. E., M. de Broglie, W. L. Bragg, and L. Brillouin, eds. (1923). *Atomes et électrons. Rapports et discussions du Conseil de Physique tenu à Bruxelles du 1er au 6 avril 1921 sous les auspices de l'Institut International de Physique Solvay.* Paris: Gauthier-Villars.

Wasserman, N. H. (1981). *The Bohr–Kramers–Slater paper and the development of the quantum theory of radiation in the work of Niels Bohr.* Ph.D. Thesis, Harvard.

Weart, S. R. (1979). The physics business in America, 1919–1940: A statistical reconnaissance. Pp. 295–358 in: N. Reingold (ed.), *The sciences in the American context: New perspectives.* Washington, D.C.: Smithsonian Institution Press.

Weinberg, S. (1992). *Dreams of a final theory.* New York: Pantheon. Page reference is to the edition of Vintage Books (New York) first published in 1994.

Whittaker, E. T. (1953). *A history of the theories of aether and electricity.* 2 Vols. London: Nelson. Reprinted as Vol. 7 of *The history of modern physics, 1800–1950* (Thomas Publishers/American Institute of Physics, 1987).

Wilson, W. (1915). The quantum theory of radiation and line spectra. *Philosophical Magazine* 29: 795–802

Zimmer, E. (1934). *Umsturz im Weltbild der Physik.* Munich: Knorr & Hirth.

On the verge of *Umdeutung* in Minnesota: Van Vleck and the correspondence principle. Part two

Anthony Duncan · **Michel Janssen**

Abstract   This is the second installment of a two-part paper on developments in quantum dispersion theory leading up to Heisenberg's *Umdeutung* paper. In telling this story, we have taken a 1924 paper by John H. Van Vleck in *The Physical Review* as our main guide. In this second part we present the detailed derivations on which our narrative in the first part rests. The central result that we derive is the Kramers dispersion formula, which played a key role in the thinking that led to Heisenberg's *Umdeutung* paper. We derive classical formulae for the dispersion, emission, and absorption of radiation and use Bohr's correspondence principle to construct their quantum counterparts both for the special case of a charged harmonic oscillator (Sect. 5) and for arbitrary non-degenerate multiply-periodic systems (Sect. 6). We then rederive these results in modern quantum mechanics (Sect. 7).

A. Duncan
Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, USA

M. Janssen
Program in the History of Science, Technology, and Medicine, University of Minnesota,
Minneapolis, USA

M. Janssen (✉)
Tate Laboratory of Physics, 116 Church St. NE, Minneapolis, MN 55455, USA
e-mail: janss011@umn.edu

5 Van Vleck and the application of the correspondence principle to the interaction of matter and radiation

In the two-part paper that forms the focal point of our study, Van Vleck (1924b,c) explored in a systematic and physically cogent fashion the implications of the correspondence principle for several aspects of the interaction of matter and radiation. The paper is signed June 19, 1924 and appeared in the October 1924 issue of *The Physical Review*. In this paper, Van Vleck gave a detailed derivation of the correspondence principle for absorption, which he had introduced in a short note in the *Journal of the Optical Society in America*, signed April 7, 1924 (Van Vleck, 1924a). In addition, he thoroughly examined the issues involved in connecting Einstein's *A* and *B* coefficients to features of classical electron orbits. Finally, as we mentioned in Sect. 3.4 in Part One of our paper, he showed that, in the limit of high quantum numbers, Kramers' quantum formula for polarization merges with the classical formula for polarization in arbitrary non-degenerate multiply-periodic systems.

In part I of his paper, reproduced in (Van der Waerden, 1968), Van Vleck (1924b) discusses the transition from quantum-theoretical expressions for emission, absorption, and dispersion to corresponding classical expressions that one expects to hold in the limit of high quantum numbers. It is only in part II, not included in (Van der Waerden, 1968), that Van Vleck (1924c) derives the classical expressions for absorption and dispersion of radiation by a general non-degenerate multiply-periodic system, using standard methods of canonical perturbation theory in action-angle variables. Van Vleck could assume his audience to be thoroughly familiar with these techniques. This is no longer true today. For the sake of clarity of exposition, we therefore invert the order of Van Vleck's own presentation.

In Sect. 5.1, we present the basic elements of the canonical formalism in action-angle variables and use it to rederive the classical formula (6) in Sect. 3.1 for the dipole moment of a charged one-dimensional simple harmonic oscillator. Though much more complicated than the derivation in Sect. 3.1, this new derivation has two distinct advantages. First, it suggests a way of translating the classical formula into a quantum formula with the help of Bohr's correspondence principle and Einstein's *A* and *B* coefficients. Secondly, both the derivation of the classical formula and its translation into a quantum formula can easily be generalized to arbitrary non-degenerate multiply-periodic systems.

In Sect. 5.2, we translate the classical formula for the dipole moment of a simple harmonic oscillator into a quantum formula. In Sect. 5.3, we similarly convert classical formulae for emission and absorption by a simple harmonic oscillator to the corresponding quantum formulae. Both the mathematical manipulations and the physical interpretation are particularly transparent in the case of a simple harmonic oscillator, and Van Vleck himself frequently used this example for illustrative purposes. The generalization of the various results to arbitrary non-degenerate multiply-periodic systems, which is a primary focus of Van Vleck's paper, will be deferred to Sect. 6. In Sect. 7, we present (or outline) modern derivations of various results in Sects. 5 and 6. In Sect. 8, we summarize our conclusions.

5.1 Deriving the classical formula for the dipole moment of a simple harmonic oscillator using canonical perturbation theory

In this subsection we rederive formula (6) in Sect. 3.4 for the dipole moment of a charged one-dimensional simple harmonic oscillator, using canonical perturbation theory in action-angle variables. Like Kramers, Van Vleck was a master of these techniques in classical mechanics. As Van Vleck recalled 50 years after the fact:

> In 1924 I was an assistant professor at the University of Minnesota. On an American trip, Ehrenfest gave a lecture there …[He] said he would like to hear a colloquium by a member of the staff. I was selected to give a talk on my "Correspondence Principle for Absorption" …I remember Ehrenfest being surprised at my being so young a man. The lengthy formulas for perturbed orbits in my publication on the three-body problem of the helium atom [Van Vleck, 1922] had given him the image of a venerable astronomer making calculations in celestial mechanics (Van Vleck, 1974, p. 9).[176]

We begin by reviewing some of the mathematical tools we need.[177] Consider a classical Hamiltonian system with phase space coordinates $(q_i, p_i), i = (1, 2, \ldots, N)$ and Hamiltonian $H(q_i, p_i)$, which does not explicitly depend on time. Hamilton's equations are

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}. \tag{15}$$

Consider a contact transformation $(q_i, p_i) \rightarrow (q_i', p_i')$ preserving the form of Hamilton's equations, in the sense that there exists a new Hamiltonian H' such that

$$\dot{q}_i' = \frac{\partial H'}{\partial p_i'}, \quad \dot{p}_i' = -\frac{\partial H'}{\partial q_i'}. \tag{16}$$

Since Hamilton's equations (15) and (16) must hold simultaneously, the variational principles

$$\delta \int_{t_1}^{t_2} \left(\sum_i p_i \dot{q}_i - H(q_i, p_i) \right) dt = 0, \quad \delta \int_{t_1}^{t_2} \left(\sum_i p_i' \dot{q}_i' - H'(p_i', q_i') \right) dt = 0 \tag{17}$$

---

[176] Van Vleck failed to conform to Ehrenfest's image of a young physicist in another respect. In an interview in 1973, "Van Vleck recalled, "I shocked Ehrenfest …when I told him I liked popular music." Ehrenfest, he said, "thought that was completely irreconcilable with my having written any respectable papers." (Fellows, 1985, p. 54)

[177] This material is covered in standard graduate textbooks on classical mechanics, such as (Goldstein, 1980), heavily influenced by (Born, 1925) and (Goldstein, 1980, pp. 429, 493, 540). We recommend (Matzner and Shepley, 1991).

for arbitrary times t_1 and t_2 must also hold simultaneously. This implies that the difference between the two integrands in Eq. (17) must be a total time derivative

$$\left(\sum_i p_i \dot{q}_i - H(q_i, p_i) - \sum_i p'_i \dot{q}'_i + H'(p'_i, q'_i) \right) dt = dF, \qquad (18)$$

which will not contribute to the variation of the action. The apparent dependence of F on the $4N + 1$ variables $(q_i, p_i, q'_i, p'_i, t)$ can be reduced to $2N + 1$ variables via the equations for the contact transformation $(q_i, p_i) \rightarrow (q'_i, p'_i)$. If we choose to write F as a function of the initial and final coordinates, $F = F(q_i, q'_i, t)$, then the partial derivatives of F can be read off directly from Eq. (18):

$$\frac{\partial F}{\partial t} = H' - H, \quad \frac{\partial F}{\partial q_i} = p_i, \quad \frac{\partial F}{\partial q'_i} = -p'_i. \qquad (19)$$

By solving (at least in principle!) the second of these three equations for q'_i as a function of (q_i, p_i), and then substituting the result in the third to obtain p'_i, we see that the function F encodes the full information of the transformation $(q_i, p_i) \rightarrow (q'_i, p'_i)$. This function is called the *generating function* of the transformation. Given F the form of the new Hamiltonian H' can be obtained (again, in principle!) from the first of Eqs. (19).

A special case of great interest occurs when the generating function F can be chosen so that the resulting Hamiltonian is independent of the new coordinates q'_i (which are then called *ignorable*). Hamilton's equations then immediately imply that the associated momenta p'_i are time independent, and that the new coordinates q'_i are linear in time. In this circumstance the new momenta are usually called *action variables*—the notation J_i is conventional for these—while the new coordinates are dubbed *angle variables*, with the conventional notation w_i.

To illustrate the above with a concrete example, which we shall be using throughout this section, consider a one-dimensional simple harmonic oscillator with Hamiltonian:[178]

$$H = \frac{p^2}{2m} + \frac{1}{2} m \omega_0^2 q^2. \qquad (20)$$

Consider the transformation induced by

$$F = \frac{1}{2} m \omega_0 q^2 \cot q'. \qquad (21)$$

---

[178] A short digression on the (almost inevitable) notational confusions lurking in this subject is in order. We shall continue to use the conventional notation ω to denote angular frequencies, with the ordinary frequency (reciprocal period) denoted by the Greek letter ν. Unfortunately, Van Vleck uses ω to denote ordinary frequency! Moreover, there is the embarrassing similarity of the angle variables w_i to the frequencies ω_i. Also, there is the need to distinguish between the frequencies of the isolated mechanical system ($\omega_0 = 2\pi \nu_0$ for the simple harmonic oscillator) and the frequency of an applied electromagnetic wave, which we shall denote as $\omega = 2\pi \nu$ throughout.

This function does not explicitly depend on time, so $H' = H$ (see Eq. (19)). Eq. (19) also tells us that

$$p = \frac{\partial F}{\partial q} = m\omega_0 q \cot q', \quad p' = -\frac{\partial F}{\partial q'} = \frac{1}{2}m\omega_0 q^2 \csc^2 q'. \tag{22}$$

From the latter equation it follows that $q^2 = (2p'/m\omega_0) \sin^2 q'$ or that

$$q = \sqrt{\frac{2p'}{m\omega_0}} \sin q'. \tag{23}$$

Inserting this expression for q into the expression for p, we find

$$p = \sqrt{2m\omega_0 p'} \cos q'. \tag{24}$$

Substituting Eqs. (23)–(24) for q and p into Eq. (20) we find

$$H = \omega_0 p'. \tag{25}$$

Since $H' = H$, this means that the new coordinate variable q' is *ignorable*, as desired. Hamilton's equations for (q', p') are:

$$\dot{q}' = \frac{\partial H}{\partial p'} = \omega_0, \quad \dot{p}' = -\frac{\partial H}{\partial q'} = 0, \tag{26}$$

from which it follows that $q' = \omega_0 t + \epsilon$ and that $p' = H/\omega_0$ is time independent. Instead of the canonically conjugate variables (p', q') it is customary to employ rescaled action-angle variables

$$J \equiv 2\pi p', \quad w \equiv \frac{1}{2\pi} q'. \tag{27}$$

Hamilton's equations for (J, w) are:

$$\dot{w} = \frac{\partial H}{\partial J} = \nu_0, \quad \dot{J} = -\frac{\partial H}{\partial w} = 0. \tag{28}$$

It follows that $J = H/\nu_0$ and $w = \nu_0 t + \epsilon$ (appropriately redefining the arbitrary phase ϵ) for our one-dimensional oscillator.

The connection to the terminology *action variable* is easily seen in this example. In this simple case, the action is defined as the area enclosed by a single orbit of the periodic system in the two-dimensional phase space spanned by the coordinates (p, q):

$$J = \oint p \, dq. \tag{29}$$

Inserting Eqs. (23) and (24) into the integrand, we find

$$
\oint \left(\sqrt{2m\omega_0 p'} \cos q' \right) \mathrm{d} \left(\sqrt{\frac{2p'}{m\omega_0}} \sin q' \right) = \int\limits_0^{2\pi} 2p' \cos^2 q' \mathrm{d}q' = 2\pi p', \qquad (30)
$$

which is just the expression for J in Eq. (27).

The result (23) represents, of course, the solution of the equation of motion of the oscillator

$$
q(t) = D \cos 2\pi \nu_0 t = D \cos 2\pi w, \qquad (31)
$$

where we have chosen the phase shift ϵ to start the oscillator at maximum displacement at $t = 0$, and where the amplitude is a function of the action variable

$$
D = \sqrt{\frac{J}{m\pi\omega_0}}. \qquad (32)
$$

We now turn to our basic model for dispersion, i.e., a charged one-dimensional simple harmonic oscillator subjected to the periodically varying electric field of an electromagnetic wave. Earlier, we used elementary techniques of classical mechanics to analyze this system (see Eqs. (2)–(6) in Sect. 3.1). Although such methods are physically transparent, they depend on an explicit treatment of the equations of motion of a specific and completely specified Hamiltonian. The same results can be obtained by the methods of canonical perturbation theory, where general formulas can be obtained for the perturbation in the coordinate(s) of the system *completely independently of the specific nature of the dynamics*. As Van Vleck put it:

> If we were to study the perturbations in the motion produced by the incident wave purely with the aid of [Newton's second law] it would be impossible to make further progress without specializing the form of the potential function [such as, e.g., $\frac{1}{2}m\omega_0^2 q^2$ in Eq. (20)] …However, it is quite a different story when we seek to compute the perturbations …in the "angle variables" w_1, w_2, w_3 and their conjugate momenta J_1, J_2, J_3 …In fact by using them rather than x, y, z, which is the essential feature of the present calculation, the periodic properties of the system come to light even without knowing the form of [the potential] (Van Vleck, 1924c, p. 350).

Using canonical perturbation theory in action-angle variables, we rederive Eq. (6) of Sect. 3.1 for the polarization of a one-dimensional charged simple harmonic oscillator. In Sect. 6.2, we turn to the general case of an arbitrary non-degenerate multiply-periodic system.

The Hamiltonian is now the sum of the Hamiltonian H_0 given by Eq. (20) and a perturbative term H_{int} describing the interaction between the harmonic oscillator and

the electromagnetic wave[179]:

$$H = H_0 + H_{\text{int}} = \frac{p^2}{2m} + \frac{1}{2}m\omega_0^2 x^2 + eEx\cos\omega t. \tag{33}$$

The subscript '0' in ν_0 or ω_0 refers to the characteristic frequency of the unperturbed oscillator. Without subscript ν and ω refer to the frequency of the external electric field.

Absent a perturbing field ($E = 0, H = H_0$), we can write $x(t)$ in terms of the action-angle variables J and $w = \nu_0 t$:

$$x(t) = \sum_{\tau=\pm 1} A_\tau(J)e^{2\pi i\tau w}, \tag{34}$$

where A_τ has to satisfy the conjugacy relation $A_\tau = A^*_{-\tau}$ to ensure that $x(t)$ in Eq. (34) is real ($x(t) = x^*(t)$). Note that we have changed notation somewhat compared to Eq. (31). We returned to Cartesian coordinate notation (x instead of q), and the amplitude has been redefined:[180]

$$D = 2|A_\tau|. \tag{35}$$

The action-angle variables $J = H_0/\nu_0$ and $w = \nu_0 t$ satisfy Hamilton's equations (see Eq. (28)):

$$0 = -\dot{J} = \frac{\partial H_0}{\partial w}, \quad \frac{\partial H_0}{\partial J} = \dot{w} = \nu_0. \tag{36}$$

It is a special feature of the simple harmonic oscillator that the frequency ν_0 is independent of the amplitude of motion (and thereby of the action). The generating function for the contact transformation from (x, p) to (w, J) is time independent (cf. Eq. (21)), so Eq. (19) implies that the old and new Hamiltonians coincide in value (i.e., one simply reexpresses the original Hamiltonian in the new variables). Even with the perturbation turned on *we shall continue to use the same contact transformation*, computing the perturbations ($\Delta w, \Delta J$) induced by the applied field in the action-angle variables (w, J) as an expansion in E. These are *not* action-angle variables for the full Hamiltonian $H_0 + H_{\text{int}}$, only for the unperturbed Hamiltonian H_0 (cf. Van Vleck, 1926, pp. 200–201).

Eventually, we are interested in the displacement Δx in the particle coordinate (to first order in E) induced by the applied field. To first order, Δx is given by

$$\Delta x = \frac{\partial x}{\partial J}\Delta J + \frac{\partial x}{\partial w}\Delta w. \tag{37}$$

---

[179] As before, we assume that the electric field is in the direction of motion of the oscillator (cf. Sect. 3.1). It follows from Eq. (33) that the force $F = -\partial V/\partial x$ of the electric field on the charge is $-eE\cos\omega t$, in accordance with Eq. (3) in Sect. 3.1 (recall that we use e to denote the absolute value of the electron charge).

[180] Inserting $A_\tau = |A_\tau|e^{i\varphi}$ into Eq. (34), we find $x(t) = (|A_\tau| + |A_{-\tau}|)\cos(2\pi w + \varphi)$. Since $A_\tau = A^*_{-\tau}$, $|A_\tau|^2 = A_\tau A^*_\tau$ is equal to $|A_{-\tau}|^2 = A_{-\tau}A^*_{-\tau}$. The phase angle φ is immaterial.

Using Eq. (34) to evaluate $\partial x/\partial J$ and $\partial x/\partial w$, we can rewrite this as:

$$\Delta x = \sum_{\tau=\pm 1} \left(\frac{\partial A_\tau}{\partial J} \Delta J + 2\pi i \tau A_\tau \Delta w \right) e^{2\pi i \tau w}. \tag{38}$$

Assuming the external field to be switched on at time zero, the first-order shifts Δw and ΔJ are given by:

$$\Delta J = \int_0^t \Delta \dot{J} dt, \quad \Delta w = \int_0^t \Delta \dot{w} dt, \tag{39}$$

where the integrands $\Delta \dot{J}$ and $\Delta \dot{w}$ are determined by Hamilton's equations.

The perturbation in Eq. (33) will induce a time dependence in the action variable, as Hamilton's equation for the action variable in the presence of the perturbing field now reads

$$\dot{J} = -\frac{\partial H_0}{\partial w} - eE \frac{\partial x}{\partial w} \cos 2\pi \nu t = -eE \frac{\partial x}{\partial w} \cos 2\pi \nu t. \tag{40}$$

Note that we still have $\partial H_0/\partial w = 0$, so $\Delta \dot{J} = \dot{J}$. At this point it is convenient to replace $\cos 2\pi \nu t$ by $\frac{1}{2}(e^{2\pi i \nu t} + e^{-2\pi i \nu t})$. Inserting Eq. (34) into Eq. (40), we find

$$\Delta \dot{J} = -\pi i eE \sum_{\tau=\pm 1} \tau A_\tau \left(e^{2\pi i(\tau w + \nu t)} + e^{2\pi i(\tau w - \nu t)} \right). \tag{41}$$

To obtain the polarization, which is a linear effect in the applied field E, we only need ΔJ and Δw to first order in E. This means that the angle variables w in the exponents in Eq. (41) can be taken to zeroth order, i.e., $w = \nu_0 t$. Integrating $\Delta \dot{J}$ we find:

$$\Delta J = \int_0^t \Delta \dot{J} dt = \frac{eE}{2} \sum_{\tau=\pm 1} \tau A_\tau \left\{ \frac{1 - e^{2\pi i(\tau \nu_0 t + \nu t)}}{\tau \nu_0 + \nu} + \frac{1 - e^{2\pi i(\tau \nu_0 t - \nu t)}}{\tau \nu_0 - \nu} \right\}. \tag{42}$$

Next, we need to compute the first order shift Δw in the angle variable w. Hamilton's equation for the angle variable w in the presence of the perturbation is:[181]

$$\begin{aligned} \dot{w} &= \frac{\partial H_0}{\partial J} + eE \frac{\partial x}{\partial J} \cos 2\pi \nu t \\ &= \nu_0 + \frac{eE}{2} \sum_{\tau=\pm 1} \frac{\partial A_\tau}{\partial J} \left(e^{2\pi i(\tau w + \nu t)} + e^{2\pi i(\tau w - \nu t)} \right). \end{aligned} \tag{43}$$

---

[181] It is a special feature of the simple harmonic oscillator that the characteristic frequency ν_0 is independent of the amplitude and thus of the action variable J (see Eq. (32)). In general, ν_0 will be a function of J. The first term on the right-hand side of Eq. (43) would then become $\partial H_0/\partial J = \nu_0(J) = \nu_0 + (\partial \nu_0/\partial J)\Delta J$.

Once again, w may be replaced by $v_0 t$ in the exponentials in Eq. (43). Integrating the second term in Eq. (43), which gives the shift $\Delta \dot{w}$ due to H_{int}, we find:

$$\Delta w = \int_0^t \Delta \dot{w} \, dt = \frac{ieE}{4\pi} \sum_{\tau = \pm 1} \frac{\partial A_\tau}{\partial J} \left\{ \frac{1 - e^{2\pi i(\tau v_0 t + vt)}}{\tau v_0 + v} + \frac{1 - e^{2\pi i(\tau v_0 t - vt)}}{\tau v_0 - v} \right\}. \quad (44)$$

Substituting expressions (42) and (44) for ΔJ and Δw into Eq. (38), we find

$$\Delta x = \frac{eE}{2} \sum_{\tau' = \pm 1} \sum_{\tau = \pm 1} \left\{ \frac{\partial A_{\tau'}}{\partial J} \tau A_\tau - \tau' A_{\tau'} \frac{\partial A_\tau}{\partial J} \right\} \frac{1 - e^{2\pi i(\tau v_0 t - vt)}}{\tau v_0 - v} e^{2\pi i \tau' v_0 t} \quad (45)$$
$$+ (v \to -v),$$

where "$(v \to -v)$" here and below is shorthand for: "the same term with v replaced by $-v$ everywhere." The coherent contribution to the polarization comes from the terms in Eq. (45) with the same time dependence as the applied field, i.e., from terms in which the time dependence is given by the factor $e^{\pm 2\pi i vt}$. In the terminology of Van Vleck (1924c): "the part of the displacement which is resonant to the impressed wave" (p. 361). These are the terms in which the summation indices, which in the case of the simple harmonic oscillator only take on the values ± 1, have opposite values, i.e., $\tau = -\tau'$. The contribution of such terms to the first-order displacement is

$$\Delta x_{\text{coh}} = \frac{eE}{2} \sum_{\tau = \pm 1} \left\{ \left(\frac{\partial A_{-\tau}}{\partial J} \tau A_\tau + \tau A_{-\tau} \frac{\partial A_\tau}{\partial J} \right) \frac{-e^{-2\pi i vt}}{\tau v_0 - v} + (v \to -v) \right\}$$
$$= \frac{eE}{2} \sum_{\tau = \pm 1} \tau \frac{\partial |A_\tau|^2}{\partial J} \left\{ \frac{e^{-2\pi i vt}}{v - \tau v_0} - \frac{e^{2\pi i vt}}{v + \tau v_0} \right\}. \quad (46)$$

The imaginary part of this expression is a sum over the product of odd and even functions of the index τ,

$$-\frac{eE}{2} \sum_{\tau = \pm 1} \tau \frac{\partial |A_\tau|^2}{\partial J} \left(\frac{1}{v - \tau v_0} + \frac{1}{v + \tau v_0} \right) \sin 2\pi vt, \quad (47)$$

and therefore vanishes, leaving only the real part:

$$\Delta x_{\text{coh}} = \frac{eE}{2} \sum_\tau \tau \frac{\partial |A_\tau|^2}{\partial J} \left(\frac{1}{v - \tau v_0} - \frac{1}{v + \tau v_0} \right) \cos 2\pi vt$$
$$= \frac{eE}{2} \sum_\tau \tau \frac{\partial |A_\tau|^2}{\partial J} \left(\frac{2\tau v_0}{v^2 - \tau^2 v_0^2} \right) \cos 2\pi vt. \quad (48)$$

Since $|A_\tau|^2 = |A_{-\tau}|^2$ (see note 180) and since τ only takes on the values ± 1 in the case of the simple harmonic oscillator, $\tau^2 = 1$ and the two terms in the summation

over τ are identical. Although in this special case the derivative with respect to J only acts on $|A_\tau|^2$, we are free to include the expression $2\nu_0/(\nu^2 - \nu_0^2)$ within the scope of the derivative (recall that ν_0 does not depend on J in this case). Equation (48) then becomes

$$\Delta x_{\text{coh}} = 2eE \frac{\partial}{\partial J} \left(\frac{\nu_0}{\nu^2 - \nu_0^2} |A_\tau|^2 \right) \cos 2\pi \nu t. \tag{49}$$

The resulting expression for the dipole moment, $p(t) = -e\Delta x_{\text{coh}}$, of a one-dimensional charged simple harmonic oscillator is a special case of the expressions for the dipole moment of a general non-degenerate multiply-periodic system with the same charge given by Kramers and Van Vleck. Kramers (1924b, p. 310, Eq. 2*) denotes this quantity by P and gives the following formula:

$$P = \frac{E}{2} \sum \frac{\partial}{\partial I} \left(\frac{C^2 \omega}{\omega^2 - \nu^2} \right) \cos 2\pi \nu t. \tag{50}$$

In the special case of a one-dimensional charged simple harmonic oscillator, ω, I, and C correspond to ν_0, J, and $2|A_\tau|$ in our notation, respectively. There appears to be a factor e^2 missing in Kramers' formula. We shall derive the corresponding formula (41) in (Van Vleck, 1924c, p. 361) in Sect. 6.2.

Equation (49) is equivalent to Eq. (6), the result of our much simpler derivation in Sect. 3.1. Recalling that (cf. Eqs. (31)–(32), (34)–(35) and note 180)

$$x(t) = 2|A_\tau| \cos 2\pi \nu_0 t = \sqrt{\frac{J}{2\pi^2 m \nu_0}} \cos 2\pi \nu_0 t, \tag{51}$$

we have $|A_\tau|^2 = J/(8\pi^2 m \nu_0)$, and Eq. (49) reduces to

$$\Delta x_{\text{coh}} = \frac{eE \cos 2\pi \nu t}{4\pi^2 m (\nu^2 - \nu_0^2)}. \tag{52}$$

The dipole moment is thus given by:

$$p(t) = -e\Delta x_{\text{coh}} = \frac{e^2 E}{4\pi^2 m (\nu_0^2 - \nu^2)} \cos 2\pi \nu t, \tag{53}$$

in agreement with Eq. (6) in Sect. 3.1.

The preceding discussion employs a version of canonical perturbation theory in which a single set of action-angle variables, chosen for the unperturbed Hamiltonian, is used throughout the calculation, even after the time-dependent perturbation is switched on. Accordingly, the new action variables are no longer constant, and the new angle variables are no longer linear in time. The same classical polarization result is derived in a somewhat different manner by Born (1924) and Kramers and Heisenberg (1925). Born performs a contact transformation in which the generating function F (cf. Eq. (18)) is chosen as a function of (q_i, p_i'), the old coordinates and the new

momenta, which is then evaluated systematically order by order in the perturbation to maintain the constancy of the new action variables. In (Kramers and Heisenberg, 1925) the same procedure is followed, but as only the first order result is needed, it suffices to use the infinitesimal form of the contact transformation.[182]

5.2 Converting the classical formula for dispersion to a quantum formula in the special case of a simple harmonic oscillator

Using Bohr's correspondence principle as our guide, we now 'translate' the classical formula (49) for displacement (and thence for polarization) into a quantum formula. Two main ingredients go into this particular application of the correspondence principle: (1) a rule—commonly attributed to Born (1924)[183] but found independently by both Kramers and Van Vleck (see below)—for replacing derivatives with respect to the action variables in classical formulae by difference quotients involving neighboring quantum states; (2) the A and B coefficients of Einstein's quantum theory of radiation. In general, the "translation" of a classical formula into a quantum formula involves a third step. The orbital frequencies need to be replaced by transition frequencies. The case of a simple harmonic oscillator has the special features that the only relevant transitions are between adjacent states and that the transition frequency $\nu_{i \to f}$ coincides with the mechanical frequency ν_0. Another special feature is that the correspondence between quantum and classical results for large quantum numbers continues to hold all the way down to the lowest quantum numbers, due to the extremely simple form of the energy spectrum, with uniformly spaced levels.

Using the rule for replacing derivatives by difference quotients, we obtain the quantum formula for polarization from Eq. (49) by the formal correspondence replacement

$$\left. \frac{\partial F(w, J)}{\partial J} \right|_{J=rh} \to \frac{1}{h}(F(r+1) - F(r)), \tag{54}$$

where F can be any dynamical quantity of the system. Classically, it is a function $F(w, J)$ on phase space. $F(r)$ denotes its value, $F(w, J = rh)$, in the quantum state specified by the integer quantum number r. In the correspondence limit where r gets very large, the difference between the values rh and $(r + 1)h$ for the action variable J becomes so small that the difference quotient to the right of the arrow in Eq. (54) approaches the derivative on the left. With this prescription, the classical formula Eq. (49) turns into a quantum expression for the coherent part of the displacement of the particle in quantum state r:

$$\Delta x_{\text{coh}}^r = \frac{2eE}{h} \left(\frac{\nu_0 |A^{r+1}|^2}{\nu^2 - \nu_0^2} - \frac{\nu_0 |A^r|^2}{\nu^2 - \nu_0^2} \right) \cos 2\pi \nu t. \tag{55}$$

---

[182] For a discussion of infinitesimal canonical transformations, see Chapter 11 of (Matzner and Shepley, 1991).

[183] See, e.g., (Jammer, 1966, p. 193), (MacKinnon, 1977, p. 148), (Cassidy, 1991, pp. 178, 186, 188), or (Aitchison et al., 2004, p. 1372).

The amplitudes A^r correspond to the A_τ (with $\tau = \pm 1$) in Eq. (49), and are related to the amplitudes D_r in Eq. (32) for an oscillator in state r by $D_r = 2|A^r|$ (see Eq. (51)). As we saw in Sect. 3.3, Ladenburg (1921) showed how these amplitudes can be connected to the Einstein A coefficients for spontaneous emission (not to be confused with the amplitudes A^r).

At this point we briefly review Einstein's quantum theory of radiation (Einstein, 1916a,b, 1917), using the notation of (Van Vleck, 1924b). Imagine an ensemble of atoms—or indeed, any conceivable quantized mechanical system, such as one-dimensional quantized oscillators—in interaction and statistical equilibrium with an ambient electromagnetic field of spectral density $\rho(\nu)$. If we label the stationary states of the atoms by indices $r, s, \ldots$, the number of atoms in state r (of energy E_r) by N_r, and recall the Bohr frequency condition $\nu_{rs} = (E_r - E_s)/h$, Einstein's analysis gives the average rate of energy emission of light of frequency ν_{rs} for an atom in state r as

$$\frac{dE_{r \to s}}{dt} = h\nu_{rs}\left(A_{r \to s} + B_{r \to s}\rho(\nu_{rs})\right), \tag{56}$$

and the average rate of energy absorption of light of frequency ν_{rs} by an atom in state s as

$$\frac{dE_{s \to r}}{dt} = h\nu_{rs}B_{s \to r}\rho(\nu_{rs}), \tag{57}$$

where $A_{r \to s}$, $B_{r \to s}$, and $B_{s \to r}$ are the transition probabilities for spontaneous emission, stimulated emission, and absorption, respectively. Einstein's analysis of the requirements for thermodynamic equilibrium and comparison with Planck's law of black-body radiation then yields the critical relations

$$B_{r \to s} = B_{s \to r} = \frac{c^3}{8\pi h\nu_{rs}^3}A_{r \to s}. \tag{58}$$

For a charged simple harmonic oscillator, the only allowed transitions amount to changes in the action by one unit of Planck's constant h, so there is only a single Einstein coefficient for spontaneous emission from the state $r + 1$, namely $A_{r+1 \to r}$. The correspondence principle dictates that we associate the rate of spontaneous energy emission for *high* quantum numbers,

$$\frac{dE_{r+1 \to r}}{dt} = h\nu_0 A_{r+1 \to r} \tag{59}$$

(cf. Eq. (56), in the absence of external radiation) with the classical result for the power emitted by an accelerated (in this case, oscillating) charge, given by the Larmor formula (Jackson, 1975; Feynman et al., 1964, Vol. 1, Ch. 32):

$$P = \frac{2}{3}\frac{e^2}{c^3}\dot{v}^2. \tag{60}$$

For an oscillator in state r, with $x(t) = D_r \cos \omega_0 t$, this becomes, for the instantaneous power emission P_r in state r

$$P_r = \frac{2}{3}\frac{e^2}{c^3}\omega_0^4 D_r^2 \cos^2 \omega_0 t, \tag{61}$$

the time average of which, $\frac{1}{3}(e^2/c^3)\omega_0^4 D_r^2$, then gives the desired connection between the amplitudes $D_r = 2|A^r|$ appearing in Eq. (55) and the Einstein coefficient $A_{r+1\to r}$ in the correspondence limit:

$$h\nu_0 A_{r+1\to r} = \frac{4}{3}\frac{e^2}{c^3}\omega_0^4 |A^{r+1}|^2$$

$$|A^{r+1}|^2 = \frac{3hc^3}{64\pi^4 e^2 \nu_0^3} A_{r+1\to r}. \tag{62}$$

Van Vleck (1924b) refers to this connection as the "correspondence principle for emission" (p. 333). Substituting the expression for $|A^{r+1}|^2$ in Eq. (62) into Eq. (55) for $\Delta x_{\mathrm{coh}}^r$ and multiplying by the charge $-e$ to obtain the dipole moment per oscillator and by the number density of oscillators n_{osc}, we arrive at the following result for the polarization induced by the electric field E:

$$P_r = 3\frac{n_{\mathrm{osc}}c^3}{32\pi^4}E\left(\frac{A_{r+1\to r}}{\nu_0^2(\nu_0^2 - \nu^2)} - \frac{A_{r\to r-1}}{\nu_0^2(\nu_0^2 - \nu^2)}\right)\cos 2\pi\nu t. \tag{63}$$

Of course, for the special case of the ground state of the oscillator, $r = 0$, the second term in Eq. (63) cannot be present. Ladenburg's quantum formula for dispersion accordingly only had the equivalent of the first term in Eq. (63) (see Eq. (8) in Sect. 3.3). The full equation corresponds to Eq. (5) in (Kramers, 1924a), and to Eq. (17) in (Van Vleck, 1924b), except for a factor of 3, as we have not assumed random orientation of the oscillators (Van Vleck, 1924b, footnote 25).

One may easily guess that the corresponding formula for a more general, multiply-periodic system will take the form of (Van Vleck, 1924b, Eq. (17)), in analogy to Eq. (63):

$$P_r = 3\frac{n_{\mathrm{osc}}c^3}{32\pi^4}E\left(\sum_s \frac{A_{s\to r}}{\nu_{sr}^2(\nu_{sr}^2 - \nu^2)} - \sum_t \frac{A_{r\to t}}{\nu_{rt}^2(\nu_{rt}^2 - \nu^2)}\right)\cos 2\pi\nu t, \tag{64}$$

where the sum over s (resp. t) corresponds to states higher (resp. lower) than the state r, and where ν_{ij} is Van Vleck's notation for the transition frequency $\nu_{i\to j}$. In the correspondence limit where r is very large and neither s nor t differ much from r, the transition frequencies ν_{sr} and ν_{rt} become equal to the orbital frequencies in the orbits characterized by the values rh, sh, and th for the action variable J. For the harmonic oscillator, the sums in Eq. (64) degenerate to a single term each (with $s = r + 1$, $t = r - 1$), and the transition frequencies ν_{sr}, ν_{rt} are all equal to the mechanical frequency ν_0. In Sect. 6.2 we shall present Van Vleck's derivation of Eq. (64) in detail.

As we indicated above, there is some disagreement in the historical literature as to who was (or were) responsible for the key move in the construction of the quantum

dispersion formula on the basis of the correspondence principle, viz. the replacement (54) of derivatives with respect to the action variable by difference quotients. Jammer (1966, p. 193) and Mehra and Rechenberg (1982–2001, Vol. 2, p. 173) suggest that Kramers got the idea from Born via Heisenberg. Dresden (1987, p. 222) makes it clear that Kramers found the rule before Born, but allows for the possibility that Born found it independently, as Kramers (1924a) did not state the rule in his first *Nature* note, the only presentation of the Kramers dispersion formula that Born had seen when he wrote (Born, 1924). Van Vleck certainly discovered the replacement (54) of derivatives by difference quotients for himself. Since Van Vleck (1924a) announced the correspondence principle for absorption, which he could not have derived without this rule, in a paper submitted in April 1924, whereas (Born, 1924) was not received by *Zeitschrift für Physik* until June 1924, Van Vleck clearly could not have taken the rule from Born's paper. That Kramers, Van Vleck, and probably Born independently hit upon the same idea, underscores that the rule (54) for replacing derivatives by difference quotients is so natural that it readily comes to mind when one is trying to connect quantum-theoretical expressions to classical ones on the basis of the correspondence principle.

Writing to Born in 1924, Van Vleck sounds slightly annoyed at Born's insinuation that he, Van Vleck, did not realize that one needs to replace derivatives by difference quotients to get from classical to quantum-theoretical expressions. In the letter from which we already quoted in Sect. 2.4, Born had written:

> I am sending you my paper On Quantum Mechanics [Born, 1924], which pursues a goal similar to yours. While you limit yourself to the correspondence with high quantum numbers, I conversely aim for rigorous laws for arbitrary quantum numbers.[184]

To which Van Vleck replied:

> I have read with great interest your important, comprehensive article. There is, as you say, considerable similarity in the subject matter in your article and mine, especially as regards to dispersion[185] …As noted in your letter you mention more explicitly than do I the fact that formulas of the quantum theory result from those of the classical theory by replacing a derivative by a difference quotient. I have stressed the asymptotic connection of the two theories but I think it is clear in the content of my article that in the problems considered the classical and quantum formulas are connected as are derivatives and difference quotients.[186]

As we already mentioned in Sect. 1.2 (see note 17) and Sect. 3.4 (note 104), Van Vleck used the correspondence principle—in particular, the replacement of derivatives by difference quotients—to *check* that quantum formulae merge with classical formulae

---

[184] Born to Van Vleck, 24 October, 1924 (AHQP).

[185] Van Vleck seems to be talking here about (Van Vleck, 1924b,c), whereas Born was talking about (Van Vleck, 1924a). Born asked Van Vleck to send him "an offprint of your extensive calculations." Van Vleck obliged: "As you requested, I am sending you under separate cover a reprint of Parts I and II of my computations," presumably (Van Vleck, 1924b,c).

[186] Van Vleck to Born, November 30, 1924 (AHQP).

in the limit of high quantum numbers, whereas Born wanted to use the principle to *construct* quantum formulae out of their classical counterparts. We sympathize with Van Vleck's point in response to Born that the difference between the two approaches should not be exaggerated. Although Van Vleck could take the quantum formulae for emission and dispersion from (Ladenburg and Reiche, 1923) and (Kramers, 1924a), respectively, he had to construct his own quantum formula for absorption on the basis of Einstein's quantum theory of radiation. The formula for absorption was undoubtedly easier to guess than the one for dispersion, which took the combined efforts of Ladenburg and Kramers (see Sects. 3.3–3.4), but, given that Ladenburg and Reiche (1923) got it wrong, it was not completely trivial either (see Sect. 5.3). Moreover, Van Vleck's crucial insight that what matters is the differential absorption was guided by the analogy between the quantum and the classical case. In fact, as Van Vleck (1924a, p. 30) himself pointed out (in a passage we quoted in Sect. 3.4), his insight that one needs to take into account the effects of "negative absorption" (stimulated emission) to arrive at a quantum absorption formula that merges with the classical absorption formula in the correspondence limit, is precisely analogous to the insight that one needs to add a term describing such effects to the Ladenburg dispersion formula to arrive at a quantum dispersion formula, the Kramers formula, that merges with the classical dispersion formula in the correspondence limit. As we also saw in Sect. 3.4, Van Vleck formulated his correspondence principle of absorption before Kramers (1924a) published his dispersion formula. In short, Van Vleck knew perfectly well how to construct quantum formulae on the basis of correspondence considerations when he had to. And while it is true that Born put more emphasis on the constructive use of the correspondence principle, this did not lead Born to additional results of any consequence for subsequent developments (Mehra and Rechenberg, 1982–2001, Vol. 3, pp. 55–57). It was left to Heisenberg to show how one could use the correspondence principle as a guide not just to a few new formulae but to a whole new theory. In the aftermath of Heisenberg's breakthrough, the Göttingen–Copenhagen attitude seems to have been that the correspondence principle had been the ladder that had allowed physicists to get from the old quantum theory to the new matrix mechanics, a ladder that in the safe possession of the new theory could be discarded. Interestingly, Van Vleck's attitude toward the correspondence principle did not change. In early 1928 he published a paper the aim of which is described as follows in the introduction:

> In studying the very significant statistical interpretation [of quantum mechanics], the writer at first experienced considerable difficulty in understanding how the quantum formulas for averages and probabilities merge into the analogous classical expressions in the region of large quantum numbers and also, of course, in the limit $h = 0$. In the present note we shall aim to trace through the asymptotic connection between the formulas of the two theories, which does not seem to have been adequately elucidated in existing papers (Van Vleck, 1928, p. 178).[187]

---

[187] We are grateful to John Stachel for drawing our attention to this paper, in which Van Vleck introduced what has become known as the "Van Vleck determinant."

Even post-*Umdeutung*, Van Vleck thus continued to think of the correspondence principle in terms of checking rather than constructing quantum formulae.

5.3 Emission and absorption classically and quantum-theoretically in the special case of a simple harmonic oscillator

Before we present Van Vleck's "correspondence principle for absorption" (for the special case of a simple harmonic oscillator), we gather some useful results from the classical theory of a charged oscillator (of natural frequency v_0) coupled to a Maxwellian electromagnetic field. Such an oscillator (1) emits electromagnetic radiation of frequency v_0 in the absence of an external field, (2) absorbs energy from an applied electromagnetic field of frequency v, and (3) undergoes a net displacement coherent with an applied electromagnetic field (or "polarization", analyzed above).

The Larmor formula (60) gives the power loss due to radiation by our charged harmonic oscillator. The energy loss of the oscillating system can be ascribed to a radiative reaction force given by

$$F_{\text{rad}} = \frac{2e^2}{3c^3} \ddot{v} \equiv m\tau_{\text{D}} \ddot{v}, \tag{65}$$

where we shall assume that the characteristic radiation damping time τ_{D} is very short in comparison to the mechanical period: $\omega_0 \tau_{\text{D}} \ll 1$, so that radiation damping is very slow on the time scale of the mechanical oscillations of the system. The equation of motion of the oscillator (in the absence of external applied forces) now becomes

$$\dot{v} - \tau_{\text{D}} \ddot{v} + \omega_0^2 x = 0. \tag{66}$$

To a good approximation, the coordinates and velocities of this system are still behaving as harmonic oscillations of frequency ω_0 so we may assume $\ddot{v} \simeq -\omega_0^2 v$ in Eq. (66) and obtain

$$\ddot{x} + \tau_{\text{D}} \omega_0^2 \dot{x} + \omega_0^2 x = 0. \tag{67}$$

Inserting the *Ansatz* $x(t) = De^{-\alpha t}$ into Eq. (67), we find:

$$(\alpha^2 - \tau_{\text{D}} \omega_0^2 \alpha + \omega_0^2) De^{-\alpha t} = 0. \tag{68}$$

Neglecting a term with $\tau_{\text{D}}^2 \omega_0^4$ (recall that $\omega_0 \tau_{\text{D}} \ll 1$, so that $\tau_{\text{D}}^2 \omega_0^4 \ll \omega_0^2$),[188] we can rewrite the expression in parentheses as:

$$(\alpha - \frac{1}{2} \tau_{\text{D}} \omega_0^2 + i\omega_0)(\alpha - \frac{1}{2} \tau_{\text{D}} \omega_0^2 - i\omega_0). \tag{69}$$

It follows that:

$$\alpha \simeq \frac{1}{2} \tau_{\text{D}} \omega_0^2 \pm i\omega_0 \equiv \Gamma/2 \pm i\omega_0. \tag{70}$$

---

[188] Such terms are treated incorrectly in any event by the approximation leading to Eq. (67).

Thus, we have a solution of the form

$$x(t) = De^{-\Gamma t/2} \cos \omega_0 t, \tag{71}$$

from which the average rate of oscillator energy loss from the Larmor formula (60) at small times (i.e., when damping due to the $e^{-\Gamma t/2}$ factor can be ignored) is easily seen to be

$$-\frac{\mathrm{d}E_{\mathrm{osc}}}{\mathrm{d}t} = \frac{e^2}{3c^3} D^2 \omega_0^4 = \frac{16\pi^4 e^2}{3c^3} D^2 \nu_0^4 \tag{72}$$

(where we used that $\dot{\upsilon} \simeq \omega_0^2 D$). The constant $\Gamma = \tau_{\mathrm{D}}\omega_0^2$ is called the radiative decay constant. We emphasize again that the preceding discussion presupposes the narrow resonance limit, $\Gamma \ll \omega_0$. In terms of Γ, the basic equation of motion (67) can be written as

$$\ddot{x} + \Gamma \dot{x} + \omega_0^2 x = 0. \tag{73}$$

Now suppose that our charged oscillator is immersed in an ambient electromagnetic field, characterized by a spectral function (energy density per unit spectral interval) $\rho(\nu)$. As we are dealing with one-dimensional oscillators we shall simplify the discussion by assuming that only the x-component of the electric field is relevant as all the oscillators are so aligned. Then (using overbars to denote time averages) the average value of the electromagnetic energy density is (in Gaussian units) $(1/4\pi)\bar{\mathbf{E}}^2 = (3/4\pi)\bar{E}_x^2 = \rho(\nu)\Delta\nu$ in the frequency interval $(\nu, \nu + \Delta\nu)$. If $E_x = E \cos 2\pi \nu t$ we have $\bar{E}_x^2 = E^2/2$ so finally we have

$$E^2 = \frac{8\pi}{3} \rho(\nu)\Delta\nu. \tag{74}$$

The equation of motion (73) must be modified to include the coupling to the external field (switching back temporarily to angular frequencies, $\omega = 2\pi\nu$, and using complex notation to encode amplitude and phase information):

$$\ddot{x} + \Gamma \dot{x} + \omega_0^2 x = \frac{eE}{m}\mathrm{e}^{\mathrm{i}\omega t} \equiv F_{\mathrm{app}}/m, \tag{75}$$

and the average rate of energy absorption of the oscillator from the ambient field is simply the time average $\langle F_{\mathrm{app}}\dot{x}\rangle$. This linear second order equation is solved by a sum of transients (i.e. solutions of the homogeneous equation: see Eq. (73))

$$x_{\mathrm{tr}}(t) = De^{-\Gamma t/2} \cos \omega_0 t, \tag{76}$$

plus the following particular solution coherent with the applied perturbation

$$x_{\mathrm{coh}}(t) = Re \, \frac{eE}{m} \frac{\mathrm{e}^{\mathrm{i}\omega t}}{\omega_0^2 - \omega^2 + i\Gamma\omega}, \tag{77}$$

so that the desired time average $\langle F_{\text{app}}\dot{x}\rangle = \langle F_{\text{app}}(\dot{x}_{\text{tr}} + \dot{x}_{\text{coh}})\rangle$ giving the energy absorption rate becomes

$$\langle F_{\text{app}}\dot{x}\rangle = \left\langle eE\cos\omega t \,\frac{eE}{m}\,\text{Re}\left(\frac{i\omega e^{i\omega t}}{\omega_0^2 - \omega^2 + i\Gamma\omega}\right)\right\rangle. \tag{78}$$

Note that the transient part of the particle coordinate $x_{\text{tr}}(t)$ is *not* coherent with the applied field (we assume $\omega \neq \omega_0$), and therefore does not contribute to the time average of the energy absorption. This explains why the amplitude D of the oscillations is absent from the final result, which will instead depend only on the specific energy density of the ambient field. In other words, even though the charged particle may be executing very large amplitude oscillations $x_{\text{tr}}(t)$, the only part of the full coordinate $x(t)$ responsible for a nonvanishing average absorption is the part of the displacement $x_{\text{coh}}(t)$ induced by the applied field, which is proportional to E and does not involve the amplitude D. As we shall see below, the corresponding feature of the quantum calculation in the correspondence limit led Van Vleck to the very important realization that the net energy absorption involves a *difference* in the amount of absorption and stimulated emission as described in Einstein's quantum theory of radiation.

Only the cosine part of the complex exponential in Eq. (78) will contribute to the time average. Using $\langle\cos^2\omega t\rangle = 1/2$ and Eq. (74), we find

$$\begin{aligned}\langle F_{\text{app}}\dot{x}\rangle &= \frac{e^2 E^2 \Gamma}{2m}\frac{\omega^2}{(\omega_0^2 - \omega^2)^2 + \Gamma^2\omega^2}\\ &= \frac{4\pi e^2}{3m}\rho\left(\frac{\omega}{2\pi}\right)\Gamma\frac{\omega^2}{(\omega_0^2 - \omega^2)^2 + \Gamma^2\omega^2}\frac{1}{2\pi}\Delta\omega\end{aligned} \tag{79}$$

for the energy absorption rate due to the ambient field in the frequency interval $(\nu, \nu + \Delta\nu) = (\omega, \omega + \Delta\omega)$. Since Eq. (79) contains the electric field E squared, it is apparent that the generalization of this linear simple harmonic oscillator result to an arbitrary multiply-periodic system will require a second-order canonical perturbation theory calculation, which will necessarily be more involved than the corresponding classical polarization calculation, which only involves the electric field to the first order. In the case of interest, where $\Gamma \ll \omega_0$, the line resonance shape in Eq. (79) is highly peaked around the resonance frequency ω_0, so we may use the distributional limit

$$\frac{\epsilon}{x^2 + \epsilon^2} \to \pi\delta(x), \quad \epsilon \to 0 \tag{80}$$

with $x = \omega^2 - \omega_0^2$ and $\epsilon = \Gamma\omega$ to execute the integration over ω in Eq. (79) and compute the total absorption rate:

$$\begin{aligned}\langle F_{\text{app}}\dot{x}\rangle &\approx \frac{2e^2}{3m}\int\rho\left(\frac{\omega}{2\pi}\right)\Gamma\frac{\pi}{\Gamma\omega}\omega^2\delta(\omega^2 - \omega_0^2)d\omega\\ &= \frac{\pi e^2}{3m}\rho(\nu_0).\end{aligned} \tag{81}$$

This classical result is found in (Planck, 1921) (Van Vleck, 1924b, p. 339, note 12)[189] and gives the rate at which a classical charged oscillator gains energy when immersed in an ambient classical electromagnetic field.

In Eq. (62) we found the connection in the limit of high quantum numbers between the Einstein A coefficients and the amplitudes $D_r = 2|A^r|$ of the mechanical motion in the emitting state r:

$$A_{r \to s} \simeq \frac{16\pi^4 e^2}{3hc^3} D_r^2 v_{rs}^3.$$ (82)

From the Einstein relation (58) this implies a corresponding result for the B-coefficients:

$$B_{r \to s} = B_{s \to r} = \frac{2\pi^3 e^2}{3h^2} D_r^2.$$ (83)

In the r-th quantized state of the oscillator, we have $J = rh$ so from Eq. (32) the corresponding amplitude D_r^{qu} of the *quantized* motion becomes

$$D_r^{\text{qu}} = \sqrt{\frac{rh}{2\pi^2 m v_0}},$$ (84)

and the quantum result for the A coefficients in the present case of a linear simple harmonic oscillator becomes

$$A_{r \to r-1} = \frac{8\pi^2 e^2 v_0^2 r}{3mc^3},$$ (85)

while the quantum result for the B coeffficients takes the form

$$B_{r \to r-1} = B_{r-1 \to r} = \frac{\pi e^2 r}{3hm v_0}.$$ (86)

The Einstein analysis of A and B coefficients makes it clear that at the quantum level we must consider what Van Vleck (1924b, p. 340) calls the "differential absorption rate": the rate of energy absorption of the oscillator in state r going to state $r + 1$ via Eq. (57) *minus* the stimulated emission induced by the ambient field and causing the transition r to $r - 1$ (the B term in Eq. (56)). From Eq. (86) we therefore have for the

---

[189] Van Vleck probably got the references to (Planck, 1921) from (Ladenburg and Reiche, 1923). Both Van Vleck (1924b, p. 339, note 12; p. 340, note 14) and Ladenburg and Reiche (1923, p. 588, note 19; p. 591, note 30) cite "equations (260) and (159)" and "section 158" in (Planck, 1921).

differential absorption rate of an oscillator in state r

$$\begin{aligned}
\frac{\mathrm{d}E_{\text{net}}}{\mathrm{d}t} &= h\nu_0(B_{r\to r+1} - B_{r\to r-1})\rho(\nu_0) \\
&= h\nu_0(B_{r+1\to r} - B_{r\to r-1})\rho(\nu_0) \\
&= h\nu_0(r+1-r)\frac{\pi e^2}{3hm\nu_0}\rho(\nu_0) \\
&= \frac{\pi e^2}{3m}\rho(\nu_0),
\end{aligned} \tag{87}$$

which is precisely the classical result (81). Note that the dependence on the quantum state r (or classically, the amplitude of the motion D_r) has cancelled in the differential absorption rate, corresponding to the lack of coherence discussed previously between the transient and impressed motion.

Van Vleck derived this result in Sect. 4 of his paper. He concluded:

> We thus see that in the limiting case of large quantum numbers, where [Eq. (86)] is valid, the classical value [in Eq. (81)] for the rate of absorption of energy is nothing but the differential rate of absorption in the quantum theory. This connection of the classical and quantum differential absorption we shall term the correspondence principle of absorption (Van Vleck, 1924b, p. 340).[190]

In Sect. 5, he generalized the result to arbitrary non-degenerate multiply-periodic systems. Van Vleck's correspondence principle for 'differential absorption' (i.e., the excess of absorption over stimulated emission) also clarifies the correspondence principle for dispersion. As Kramers (1924a,b) emphasized, the negative terms in the dispersion formula were difficult to account for on the basis of purely classical concepts—they somehow corresponded to a *negative* value for e^2/m for those virtual oscillators corresponding to transitions from the initial atomic level to lower energy levels (see Sect. 3.4). Similar negative contributions in the case of absorption are physically much more transparent: transitions to higher levels result in a positive absorption of energy from the ambient electromagnetic field, whereas transitions to lower levels result in energy being returned to the field. The latter process was therefore known as "negative absorption" at the time, a term used by both Kramers (1924a, p. 676) and Van Vleck (1924b, p. 338). Noticing the greater physical transparency of his correspondence-principle results for absorption, and under the impression that Kramers' correspondence-principle arguments for the dispersion formula rested only on a treatment of harmonic oscillators, Van Vleck added sections on dispersion to his paper. Section 6, "The general correspondence principle basis for Kramers' dispersion formula," was added to the first quantum-theoretical part of the paper; Sect. 15, "Computation of polarization," to the classical part (see the letter from Van Vleck to Kramers of September 1924, quoted in Sect. 3.4).

---

[190] Van Vleck points out that this "is a purely mathematical consequence of the correspondence principle for emission, which was used in deriving [Eq. (86)]" (ibid.). A few pages later, Van Vleck (1924b, p. 343) notes that he could also have done the reverse, deriving the correspondence principle for emission from that for absorption.

When Kuhn in his AHQP interview with Van Vleck brought up the paper on the correspondence principle for absorption, Van Vleck said: "I think that was one of my better papers." "How did you get into that?," Kuhn wanted to know. Van Vleck told him:

> Through a misunderstanding of something Gregory Breit [Van Vleck's colleague in Minnesota at the time] told me. He said that the net absorption was the difference between the fluctuations up and the fluctuations down, referred to some paper of—I think it was (Kretschmann)—but that was an entirely different thing. It was concerned with the fact that under certain phase relations the light did work on the atom and under certain phase relations the atom did work on the light. It was dealing essentially with statistical fluctuations. I misunderstood his remark and proceeded to try and get the differential effect between the absorption up from a given stationary state and a[b]sorption going down.[191]

The paper Breit was referring to is presumably (Kretschmann, 1921). In this paper, Erich Kretschmann (1887–1973), a student of Planck better known for his work in general relativity (Kretschmann, 1917), gave a purely classical discussion of the emission and absorption of radiation. What Van Vleck says here about this paper fits with its contents. Van Vleck's comments, however, are also reminiscent of the following passage in (Ladenburg and Reiche, 1923):

> …according to Einstein's assumptions the effect of external radiation on a quantum atom corresponds to the effect a classical oscillator experiences from an incident wave. When the frequency of such a wave does not differ much or not at all from the characteristic frequency of the oscillator, the reaction of the oscillator consists in an increase or a decrease of its energy, depending on the difference in phase between the external wave and the motion of the oscillator. In analogy to this, Einstein assumes that the atom in state i has a probability characterized by the factor b_{ik} to make a transition to a higher state k under absorption of the energy $h\nu$ of the incident wave ("positive irradiation") and that the atom in state k has another probability (b_{ki}) to return to the state i under the influence of an external wave ("negative irradiation") (Ladenburg and Reiche, 1923, p. 586)

As we mentioned in Sect. 3.3, Ladenburg and Reiche appealed to the correspondence principle to justify their quantum formulae for emission, absorption, and dispersion. Except in the case of emission, however, their arguments were fallacious. We conjecture that this is what inspired Van Vleck to use his expertise in techniques from celestial mechanics to find the correct expressions for emission and absorption merging with

---

[191] See p. 22 of the transcript of the first session of the AHQP interview with Van Vleck. Van Vleck told this story in somewhat greater detail to Katherine Sopka. He also explained to her why he acknowledged Breit in (Van Vleck, 1924a, p. 28) but not in (Van Vleck, 1924b,c): "As he [Van Vleck] remembers it, he wanted to thank Breit in the latter, but Breit objected on the ground that the phase fluctuations he had in mind were quite different from the difference effect employed by Van Vleck and so, overmodestly, felt no acknowledgment was in order" (Sopka, 1988, p. 135, note 184; this note makes no mention of Kretschmann).

classical results in the sense of the correspondence principle.[192] Van Vleck (1924b, p. 339, note 13; p. 344, note 21) cited Ladenburg and Reiche but gave no indication that their paper was an important source of inspiration for his own. It is not implausible, however, that Van Vleck simply preferred to pass over their badly flawed calculations in silence rather than touting his own clearly superior results. As we mentioned in Sect. 3.3, one of the problems with the "correspondence" arguments of Ladenburg and Reiche is that, following (Planck, 1921) and in the spirit of the derivation of the A and B coefficients in (Einstein, 1917), they focus on collections of atoms in thermal equilibrium rather than on individual atoms. What is suggestive of a possible influence of (Ladenburg and Reiche, 1923) on (Van Vleck, 1924b,c) is that the exact same passages of (Planck, 1921) are cited in both papers (see note 189 above) and that Van Vleck (1924b) explicitly comments on the issue of many atoms in thermal equilibrium versus single atoms, noting that in Planck's discussion "no explicit mention is made of the asymptotic connection of the classical absorption and the differential absorption for a single orbit (where thermodynamic equilibrium need not be assumed) which is the primary concern of the present paper" (p. 340, note 14). The topic of the third installment that Van Vleck originally planned to add to his two-part paper also becomes understandable in light of our conjecture about the connection between (Van Vleck, 1924b,c) and (Ladenburg and Reiche, 1923). As Van Vleck explained in 1977 (see Sect. 2.4): "Part III was to be concerned with the equilibrium between absorption and emission under the Rayleigh-Jeans law" (Van Vleck and Huber, 1977, p. 939). If Ladenburg and Reiche did indeed stimulate Van Vleck's work, however, it is somewhat puzzling that he does not seem to have recognized that the virtual oscillators of BKS, which, as we saw in Sects. 3.4, 4.1, and 4.2, he consistently attributed to Slater, were essentially just the substitute oscillators of (Ladenburg and Reiche, 1923). We also saw, however, that Van Vleck was hardly alone in associating virtual oscillators with BKS. We thus conclude that it is plausible that Van Vleck was inspired by (Ladenburg and Reiche, 1923) to formulate correspondence principles for emission and absorption. For one thing, this would explain why Van Vleck, who had not worked on radiation theory before, turned his attention to the interaction between matter and radiation.

6 Generalization to arbitrary non-degenerate multiply-periodic systems

6.1 The correspondence principle for absorption

The primary result of (Van Vleck, 1924b,c) was an extension of Eq. (87) to an arbitrary non-degenerate multiply-periodic system of a single particle in three dimensions, and the demonstration that the quantum-differential absorption coincides with this more general result in the correspondence limit. Before giving Van Vleck's result we recall some basic features of multiply-periodic systems, which we shall in any event need in Sect. 6.2, where we give a completely explicit derivation (following, with minor notational changes, the one laid out by Van Vleck) of the corresponding formula for polarization.

---

[192] As we saw in Sect. 3.4, Van Vleck's calculations for dispersion were inspired by (Kramers, 1924a).

The transition from one-dimensional periodic (and harmonic) systems such as the linear simple harmonic oscillator to three-dimensional multiply-periodic ones is fairly straightforward. Apart from the obvious need to introduce vector quantities, there are only two significant additional features. First, there is the appearance of multiple overtones in the general multiply-periodic expansion (so that the multiplicity variables in the analogue of Eq. (34) take arbitrary positive and negative integral values, not just ± 1). Second, the mechanical frequencies ν_1, ν_2, ν_3 (with $\nu_i = \partial H_0 / \partial J_i$) of the separated coordinates are now in general functions of the amplitude of the classical path, which is to say, of the action variables J_i (with $i = 1, 2, 3$). We assume as before that the imposed electric field is in the X-direction so the x-coordinate of our electron is the relevant one for computing the induced coherent polarization, and in analogy to Eq. (34) we now have

$$x(t) = \sum_{\vec{\tau}} A_{\vec{\tau}} e^{2\pi i \vec{\tau} \cdot \vec{w}}, \tag{88}$$

where in the absence of the external field the angle variables $\vec{w} = (w_1, w_2, w_3) = (\nu_1, \nu_2, \nu_3)t \equiv \vec{\nu} t$ and $\vec{\tau} = (\tau_1, \tau_2, \tau_3)$ with τ_i taking on all (positive and negative) integer values. It will be useful to write Eq. (88) in an alternative purely real form, as a cosine expansion:

$$x(t) = \sum_{\vec{\tau}, \vec{\tau} \cdot \vec{\nu} > 0} X_{\vec{\tau}} \cos\left(2\pi \vec{\tau} \cdot \vec{\nu} t\right). \tag{89}$$

The complex amplitudes $A_{\vec{\tau}}$ satisfy the conjugacy condition $A_{\vec{\tau}} = A_{-\vec{\tau}}^*$ to ensure that $x(t)$ is real and we have the relation $X_{\vec{\tau}}^2 = 4 A_{\vec{\tau}} A_{-\vec{\tau}}$.[193]

As before (cf. Eq. (33)), the full Hamiltonian has the form

$$H = H_0 + e E x(t) \cos 2\pi \nu t. \tag{90}$$

The subscripted mechanical frequencies ν_i with $i = 1, 2, 3$ (comprising the vector $\vec{\nu}$) must be distinguished from the single frequency ν (unsubscripted) corresponding to the applied field.

With these notations, Van Vleck's (1924b, p. 342, Eq. (16)) result for the absorption rate becomes:

$$\frac{dE_{\text{net}}}{dt} = \frac{2}{3}\pi^3 e^2 \left[\rho(\vec{\tau} \cdot \vec{\nu}) \tau_k \frac{\partial G_\tau}{\partial J_k} + \rho'(\vec{\tau} \cdot \vec{\nu}) G_\tau \tau_k \frac{\partial}{\partial J_k}(\vec{\tau} \cdot \vec{\nu}) \right]. \tag{91}$$

where $\rho' \equiv \partial \rho / \partial \nu$ and where summation over $k = (1, 2, 3)$ is implied and where $G_\tau \equiv \vec{\tau} \cdot \vec{\nu} D_{\vec{\tau}}^2$ with $D_{\vec{\tau}}^2 \equiv X_{\vec{\tau}}^2 + Y_{\vec{\tau}}^2 + Z_{\vec{\tau}}^2$. In the special case of the harmonic oscillator, the term with ρ', the derivative of the spectral function, vanishes as there is only a single mechanical frequency $\nu = \nu_0$, which is independent of the action variable J. In the first term, we get simply

$$\frac{dE_{\text{net}}}{dt} = \frac{2}{3}\pi^3 e^2 \rho(\nu_0) \frac{\partial}{\partial J}(\nu_0 D^2). \tag{92}$$

---

[193] Cf. Eqs. (34)–(35) and note 180 in Sect. 6.1.

Using Eq. (32), $D = \sqrt{J/m\pi\omega_0}$, for the amplitude, we recover the previous result, Eq. (87).

Equation (91) is the product of a highly nontrivial application of canonical perturbation techniques, where quantities of *second* order in the applied field need to be properly evaluated (cf. our discussion following Eq. (79)). The polarization calculation presented in full in Sect. 6.2 only involves canonical perturbation theory to first order. For the absorption calculation, the variation in the action variables ΔJ_k in particular is needed to second order, and the integration of the result obtained for a monochromatic incident field needed to pass to the case of continuous radiation specified by an arbitrary spectral function $\rho(\nu)$ requires considerable care.

Slater also tried his hand at this calculation, as can be inferred from a letter from Kramers to Van Vleck, from which we already quoted in Sect. 3.4. Kramers wrote:

> Slater had, on my request, made the same calculation, and he stated that the classical mean-absorption formula gave the right result in the limit of high quantum numbers. I did, however, not see his formula, and am not quite sure that he had not forgotten the term with $\partial\rho/\partial\nu$, without which the thing is not complete of course.[194]

Van Vleck clearly remembered this point almost 40 years later. Talking to Kuhn about his 1924 absorption papers, he mentioned: "I got the term in partial rho with respect to nu. I'm very proud of the fact that I picked that one up ...Slater, at Kramers' suggestion I guess, made a completely parallel calculation in Copenhagen which he never published."[195]

6.2 The correspondence principle for polarization

In this section we retrace the derivation given by Van Vleck (1924c) of the classical polarization formula for a general non-degenerate multiply-periodic system (with a single electron) in three dimensions. We remind the reader that this result is by no means new to Van Vleck, nor, for that matter, to Born or Kramers, who also produced derivations of the same result at around this time, using slightly different versions of canonical perturbation theory (cf. our comments at the end of Sect. 5.1). The formula obtained is basically identical to a formula originally derived in celestial mechanics to compute the perturbation in the orbits of the inner planets due to the outer ones. As we saw in Sect. 3.2, Epstein had been the first to use the relevant techniques from celestial mechanics in the context of the old quantum theory. As Van Vleck reminded Slater: "The classical formula analysis to the Kramer[s] formula appears to be first ca[lc]ulated by Epstein [1922c]."[196]

The derivation is basically a straightforward generalization of the derivation of Sect. 5.1 for the special case of a charged simple harmonic oscillator in an electromagnetic field (see Eqs. (37)–(49)). The first-order perturbation in the coordinate $x(t)$

---

[194] Kramers to Van Vleck, November 11, 1924 (AHQP).

[195] See p. 22 of the transcript of the first session of the AHQP interview with Van Vleck.

[196] Van Vleck to Slater, December 15, 1924 (AHQP).

(the direction of the electric field in the incident electromagnetic wave) corresponding to the shifts $(\Delta J_l, \Delta w_l)$ in the action-angle variables is given by the three-dimensional version of Eq. (37):

$$\Delta x = \sum_l \left(\frac{\partial x}{\partial J_l} \Delta J_l + \frac{\partial x}{\partial w_l} \Delta w_l \right). \tag{93}$$

As in Sect. 5.1, we imagine that the external field is switched on at time zero, so that the shifts $(\Delta J_l, \ \Delta w_l)$ are the integrals of their time derivatives from 0 to t. In analogy with Eq. (42) and using Eq. (88) for $x(t)$, we can immediately write down the equation for ΔJ_l to first order in E:

$$\Delta J_l = \int_0^t \dot{J}_l \mathrm{d}t = \frac{eE}{2} \sum_{\vec{\tau}} \tau_l A_{\vec{\tau}} \left\{ \frac{1 - e^{2\pi \mathrm{i}(\vec{\tau} \cdot \vec{v} + v)t}}{\vec{\tau} \cdot \vec{v} + v} + (v \to -v) \right\}. \tag{94}$$

All the terms inside the summation can be taken to zeroth order in the applied field. The computation of the first-order shifts Δw_l is a little more involved as new terms, not present in the harmonic-oscillator case, enter (cf. note 181). The Hamilton equation for $\dot{w}_l$ for the full Hamiltonian Eq. (90) is (cf. Eq. (43)):

$$\dot{w}_l = v_l + \frac{eE}{2} \sum_{\vec{\tau}} \frac{\partial A_{\vec{\tau}}}{\partial J_l} \left\{ e^{2\pi \mathrm{i}(\vec{\tau} \cdot \vec{v} + v)t} + (v \to -v) \right\}. \tag{95}$$

Both terms in Eq. (95) contribute to the first-order deviation $\Delta \dot{w}_l$ from the value of v_l for the unperturbed system. Since v_l depends on J_k, there will be a term $\sum_k (\partial v_l / \partial J_k) \Delta J_k$ (cf. note 181). The second term is just the generalization of the corresponding term in Eq. (43). Hence, we get:

$$\Delta \dot{w}_l = \sum_k \frac{\partial v_l}{\partial J_k} \Delta J_k + \frac{eE}{2} \sum_{\vec{\tau}} \frac{\partial A_{\vec{\tau}}}{\partial J_l} \left\{ e^{2\pi \mathrm{i}(\vec{\tau} \cdot \vec{v} + v)t} + (v \to -v) \right\}. \tag{96}$$

Upon substitution of Eq. (94) for ΔJ_k this turns into

$$\Delta \dot{w}_l = \frac{eE}{2} \sum_{\vec{\tau},k} \left\{ \frac{\partial A_{\vec{\tau}}}{\partial J_l} e^{2\pi \mathrm{i}(\vec{\tau} \cdot \vec{v} + v)t} + \tau_k \frac{\partial v_l}{\partial J_k} A_{\vec{\tau}} \frac{1 - e^{2\pi \mathrm{i}(\vec{\tau} \cdot \vec{v} + v)t}}{\vec{\tau} \cdot \vec{v} + v} \right\}$$
$$+ (v \to -v). \tag{97}$$

Integrating Eq. (97), we find

$$\Delta w_l = \frac{eE}{4\pi} \sum_{\vec{\tau},k} \left\{ \mathrm{i} \frac{\partial A_{\vec{\tau}}}{\partial J_l} \frac{1 - e^{2\pi \mathrm{i}(\vec{\tau} \cdot \vec{v} + v)t}}{\vec{\tau} \cdot \vec{v} + v} \right.$$
$$\left. + \tau_k \frac{\partial v_l}{\partial J_k} A_{\vec{\tau}} \frac{2\pi(\vec{\tau} \cdot \vec{v} + v)t - \mathrm{i}(1 - e^{2\pi \mathrm{i}(\vec{\tau} \cdot \vec{v} + v)t})}{(\vec{\tau} \cdot \vec{v} + v)^2} \right\} + (v \to -v). \tag{98}$$

Inserting Eq. (88) into Eq. (93), we arrive at

$$\Delta x(t) = \sum_{\vec{\tau}',l} \left(\frac{\partial A_{\vec{\tau}'}}{\partial J_l} \Delta J_l + 2\pi \mathrm{i} A_{\vec{\tau}'} \tau_l' \Delta w_l \right) \mathrm{e}^{2\pi \mathrm{i} \vec{\tau}' \cdot \vec{v} t}. \tag{99}$$

Inserting Eqs. (94) and (98) for ΔJ_l and Δw_l, respectively, into this expression, we obtain

$$
\begin{aligned}
\Delta x(t) = \frac{eE}{2} \sum_{\vec{\tau},\vec{\tau}',k,l} \Bigg\{ &\tau_l \frac{\partial A_{\vec{\tau}'}}{\partial J_l} A_{\vec{\tau}} \frac{1 - \mathrm{e}^{2\pi \mathrm{i}(\vec{\tau}\cdot\vec{v}+v)t}}{\vec{\tau}\cdot\vec{v}+v} - \tau_l' \frac{\partial A_{\vec{\tau}}}{\partial J_l} A_{\vec{\tau}'} \frac{1 - \mathrm{e}^{2\pi \mathrm{i}(\vec{\tau}\cdot\vec{v}+v)t}}{\vec{\tau}\cdot\vec{v}+v} \\
&+ A_{\vec{\tau}} A_{\vec{\tau}'} \tau_k \frac{\partial v_l}{\partial J_k} \tau_l' \frac{2\pi \mathrm{i}(\vec{\tau}\cdot\vec{v}+v)t + 1 - \mathrm{e}^{2\pi \mathrm{i}(\vec{\tau}\cdot\vec{v}+v)t}}{(\vec{\tau}\cdot\vec{v}+v)^2} \\
&+ (v \to -v) \Bigg\} \mathrm{e}^{2\pi \mathrm{i}\vec{\tau}'\cdot\vec{v}t}.
\end{aligned} \tag{100}
$$

As in Sect. 5.1, we are only interested in the coherent contribution to the polarization, so we omit all terms in Eq. (100) whose time dependence is not precisely $\mathrm{e}^{\pm 2\pi \mathrm{i} v t}$ and find, writing for convenience $\sum_k \tau_k (\partial/\partial J_k) \equiv \vec{\tau} \cdot \vec{\nabla}_J$,

$$
\begin{aligned}
\Delta x_{\mathrm{coh}} = \frac{eE}{2} \sum_{\vec{\tau}} \Bigg\{ &-\vec{\tau} \cdot \vec{\nabla}_J (A_{\vec{\tau}} A_{-\vec{\tau}}) \frac{\mathrm{e}^{2\pi \mathrm{i} v t}}{\vec{\tau}\cdot\vec{v}+v} \\
&+ A_{\vec{\tau}} A_{-\vec{\tau}} \vec{\tau} \cdot \vec{\nabla}_J (\vec{\tau}\cdot\vec{v}) \frac{\mathrm{e}^{2\pi \mathrm{i} v t}}{(\vec{\tau}\cdot\vec{v}+v)^2} \Bigg\} + (v \to -v).
\end{aligned} \tag{101}
$$

Note that the coherent contribution derives from terms in which $\vec{\tau}' = -\vec{\tau}$, as otherwise the uncancelled overtones from the mechanical system would shift the spectral line (as in Raman scattering). Essentially the only additional physics of (Kramers and Heisenberg, 1925) in comparison to (Van Vleck, 1924b,c) is a detailed examination of such terms, predicted earlier in (Smekal, 1923). The terms in Eq. (101) involving $\sin 2\pi v t$ vanish, as can be seen with the help of the identities

$$\sum_{\vec{\tau}} \tau_j \left(\frac{1}{\vec{\tau}\cdot\vec{v}+v} - \frac{1}{\vec{\tau}\cdot\vec{v}-v} \right) \cdot (\text{even function of } \vec{\tau}) = 0$$

$$\sum_{\vec{\tau}} \tau_j \tau_k \left(\frac{1}{(\vec{\tau}\cdot\vec{v}+v)^2} - \frac{1}{(\vec{\tau}\cdot\vec{v}-v)^2} \right) \cdot (\text{even function of } \vec{\tau}) = 0.$$

Thus Eq. (101) simplifies to

$$\Delta x_{\text{coh}} = -\frac{eE}{2} \cos 2\pi \nu t \sum_{\vec{\tau}} \left\{ \vec{\tau} \cdot \vec{\nabla}_J \left(\frac{A_{\vec{\tau}} A_{-\vec{\tau}}}{\vec{\tau} \cdot \vec{\nu} + \nu} \right) + (\nu \to -\nu) \right\}$$

$$= -eE \cos 2\pi \nu t \sum_{\vec{\tau}} \vec{\tau} \cdot \vec{\nabla}_J \left(\frac{\vec{\tau} \cdot \vec{\nu} A_{\vec{\tau}} A_{-\vec{\tau}}}{(\vec{\tau} \cdot \vec{\nu})^2 - \nu^2} \right). \tag{102}$$

With the replacement $X_{\vec{\tau}}^2 = 4A_{\vec{\tau}} A_{-\vec{\tau}}$, we may go over to the cosine form of the expansion in Eq. (102) (cf. Eqs. (88)–(89)), summing over only positive values of $\vec{\tau} \cdot \vec{\nu}$ (with a factor of 2):

$$\Delta x_{\text{coh}} = -\frac{eE}{2} \cos 2\pi \nu t \sum_{\vec{\tau}, \vec{\tau} \cdot \vec{\nu} > 0} \vec{\tau} \cdot \vec{\nabla}_J \left(\frac{\vec{\tau} \cdot \vec{\nu} X_{\vec{\tau}}^2}{(\vec{\tau} \cdot \vec{\nu})^2 - \nu^2} \right). \tag{103}$$

This is the generalization of Eq. (49) for the harmonic oscillator.

Finally, we obtain the polarization by multiplying the displacement by N_r, the number of electrons per unit volume (the subscript r refers to the fact that we shall shortly consider only electrons in a particular quantum state r), and by $-e$ for the electron charge

$$P = N_r \frac{e^2}{2} E \cos 2\pi \nu t \sum_{\vec{\tau} \cdot \vec{\nu} > 0} \vec{\tau} \cdot \vec{\nabla}_J \left(\frac{\vec{\tau} \cdot \vec{\nu} X_{\vec{\tau}}^2}{(\vec{\tau} \cdot \vec{\nu})^2 - \nu^2} \right) \tag{104}$$

which is Eq. (41) in (Van Vleck, 1924c; in Van Vleck's notation, $\vec{\tau} \cdot \vec{\nu}$ is written ω_τ) and equivalent to Eq. 2* in (Kramers, 1924b) (see Eq. (50) above).

The equivalence of Eq. (104) to the Kramers dispersion formula (64) in the correspondence limit is sketched in (Kramers, 1924b) and fully explained in Sect. 6 of (Van Vleck, 1924b).[197] Here we follow the latter. So we begin with Eq. (64) for the polarization of a quantized system in state r, without the factor of 3 corresponding to the assumption that all oscillators be aligned with the applied field (rather than randomly in three-dimensional space), and writing N_r instead of n_{osc}:

$$P_r = \frac{N_r c^3}{32\pi^4} E \cos 2\pi \nu t \left(\sum_s \frac{A_{s \to r}}{\nu_{sr}^2 (\nu_{sr}^2 - \nu^2)} - \sum_t \frac{A_{r \to t}}{\nu_{rt}^2 (\nu_{rt}^2 - \nu^2)} \right). \tag{105}$$

The sums over s (resp. t) refer to states higher (resp. lower) in energy than the fixed state r under consideration. In the correspondence limit, we take the state r to correspond to very high quantum numbers (n_1, n_2, n_3). The states s, t are associated to the central

---

[197] Cf. Van Vleck to Kramers, September 22, 1924 (AHQP), quoted in Sect. 3.4.

state r in symmetrical pairs:

$$
\begin{aligned}
s &\rightarrow (n_1 + \tau_1, n_2 + \tau_2, n_3 + \tau_3), \\
r &\rightarrow (n_1, n_2, n_3), \\
t &\rightarrow (n_1 - \tau_1, n_2 - \tau_2, n_3 - \tau_3),
\end{aligned}
\tag{106}
$$

with $\vec{\tau} \cdot \vec{\nu} > 0$ so that the states s (resp. t) do indeed correspond to higher (resp. lower) energy states. Furthermore, we assume that $\vec{\tau} \cdot \vec{\nu} \ll \vec{n} \cdot \vec{\nu}$ so that the transitions $s \rightarrow r \rightarrow t$ correspond to very slight changes in the classical orbitals (and differences approximate well to derivatives). The Bohr–Sommerfeld quantization condition (1) associates action values $J_i = n_i h$ with a given quantized state, so the formal correspondence principle becomes (cf. Eq. (54) in Sect. 6.2):

$$
\delta_{\vec{\tau}} F(\vec{n}) \equiv F(\vec{n}) - F(\vec{n} - \vec{\tau}) \rightarrow h\vec{\tau} \cdot \vec{\nabla}_J F.
\tag{107}
$$

In this notation, formula (105) the polarization can be written as

$$
P_r = \frac{N_r c^3}{32\pi^4} E \cos 2\pi \nu t \sum_{\vec{\tau}} \delta_{\vec{\tau}} \left(\frac{A_{s \rightarrow r}}{\nu_{sr}^2 (\nu_{sr}^2 - \nu^2)} \right),
\tag{108}
$$

with $A_{s \rightarrow r}$ given by Van Vleck's "correspondence principle for emission" (see Eqs. (82) and (62))

$$
A_{s \rightarrow r} = \frac{16\pi^4 e^2}{3hc^3} D_s^2 \nu_{sr}^3,
\tag{109}
$$

where $D_s^2 = (X_{\vec{\tau}}^{(s)})^2 + (Y_{\vec{\tau}}^{(s)})^2 + (Z_{\vec{\tau}}^{(s)})^2$ is the full vector amplitude squared for the Fourier component of the classical path responsible for the transition $\vec{n} + \vec{\tau} \rightarrow \vec{n}$. Substituting Eqs. (107) and (109) into Eq. (108) and replacing the difference frequency ν_{sr} by its classical counterpart $\vec{\tau} \cdot \vec{\nu}$, we obtain,:

$$
\begin{aligned}
P_r &= N_r E \cos 2\pi \nu t \frac{c^3}{32\pi^4} \frac{16\pi^4 e^2}{3hc^3} h \sum_{\vec{\tau} \cdot \vec{\nu} > 0} \vec{\tau} \cdot \vec{\nabla}_J \left(\frac{\vec{\tau} \cdot \vec{\nu} D_s^2}{(\vec{\tau} \cdot \vec{\nu})^2 - \nu^2} \right) \\
&= N_r \frac{e^2}{2} E \cos 2\pi \nu t \sum_{\vec{\tau} \cdot \vec{\nu} > 0} \vec{\tau} \cdot \vec{\nabla}_J \left(\frac{\vec{\tau} \cdot \vec{\nu} \frac{1}{3} D_s^2}{(\vec{\tau} \cdot \vec{\nu})^2 - \nu^2} \right).
\end{aligned}
\tag{110}
$$

With the replacement $\frac{1}{3} D_s^2 \rightarrow X_{\vec{\tau}}^2$ appropriate for randomly oriented atoms, Eq. (110) becomes identical to the classical formula (104). This shows that the Kramers dispersion formula (105) does indeed merge with the classical result in the limit of high quantum numbers, as Van Vleck set out to demonstrate.

7 Derivation of the formulae for dispersion, emission, and absorption in modern quantum mechanics

Describing the impact of the new quantum mechanics on dispersion theory, Van Vleck wrote in 1929:

> Dispersion was particularly bothersome in the old quantum theory, which could never explain why the resonance frequencies in dispersion were experimentally the spectroscopic frequencies given by the Bohr frequency condition rather than the altogether different frequencies of motion in orbits constituting the stationary states [cf. our discussion in the introduction of Sect. 3]. The new mechanics, however, yields the Kramers dispersion formula, previously derived semi-empirically from the correspondence principle …As the result of the masterful treatment by Dirac [1927], a mechanism has at last to a certain extent been found for the previously so mysterious quantum jumps between stationary states (Van Vleck, 1929, pp. 494–495).

That same year, in the first installment of what would turn out to be an eight-part paper entitled "Investigations of anomalous dispersion in excited gases," Ladenburg likewise provided a brief synopsis of recent developments in dispersion theory:

> The first successful treatment of dispersion phenomena on the basis of Bohr's atomic theory implicitly contained the assumption that the orbital frequencies of the Bohr electrons are the special values at which dispersion changes sign.[198] In contrast to this, the point of departure of the *newer* development of dispersion theory is the empirical fact that not the orbital frequencies of the electrons but the frequencies, observable in emission and absorption, of "quantum jumps," i.e., spectral lines, are the singular values of anomalous dispersion. These correspond to the characteristic frequencies of quasi-elastically bound electrons in the classical electron theory [discussed in Sect. 3.1]. Tying together the notions of this theory with Bohr's atomic theory has taught us that the "strength" of the dispersion or of the "substitute oscillators," which at Bohr's suggestion were introduced as carriers of the scattered radiation needed for dispersion, is determined in non-classical fashion by the "strength," i.e., the probability of quantum jumps[,] and by the density of atoms in the "lower" atomic state involved in such quantum jumps.[199] H. A. Kramers then showed,[200] through correspondence considerations, that the dispersion formula obtained by the author [cf. Eq. (8) in Sect. 3.3] only holds exactly in the case of non-excited or meta-stable atoms; in the case of excited non-meta-stable atoms, which can also make spontaneous transitions to states of lower energy, this formula is incomplete and has to be supplemented by terms of "negative dispersion," which correspond to the "negative absorption"

---

[198] At this point, Ladenburg refers to the papers by Sommerfeld, Debye, and Davisson and the criticism of them by Bohr and Epstein that we discussed in Sect. 3.2.

[199] At this point, Ladenburg refers to his own work, Bohr's favorable reaction to it, and his subsequent work with Reiche, all discussed in Sect. 3.3.

[200] At this point, Ladenburg refers to Kramers' two *Nature* notes and to the Kramers-Heisenberg paper discussed in Sect. 3.4.

[i.e., stimulated emission] of the radiation theory of Planck and Einstein. Thus originated the "quantum-theoretical dispersion formula" [cf. Eq. (9) in Sect. 3.4] which has finally been given a fully consistent foundation in quantum mechanics and wave mechanics;[201] this new quantum theory completely avoids concepts like orbital frequencies of electrons in stationary states, and one of its points of departure was precisely the quantum-theoretical interpretation of dispersion phenomena mentioned above (Ladenburg, 1928, pp. 15–16)

Rather than pursuing the history of dispersion post-*Umdeutung*, we shall present our own modern derivations of quantum formulae for dispersion (Sect. 7.1), (spontaneous) emission (Sect. 7.2), and absorption (Sect. 7.3). Seeing how modern quantum mechanics sanctions the formulae found by Kramers, Van Vleck and others in the old quantum theory on the basis of Einstein's quantum theory of radiation and Bohr's correspondence principle will illuminate various aspects of the relation between the old and the new theory.

First, we show how the orchestra of virtual oscillators of pre-*Umdeutung* dispersion theory survives in the guise of a sum over matrix elements of the position operator. Second, we show how the diagonal matrix elements of the fundamental commutation relation for position and momentum, $[X, P] = i\hbar$, are given by the high-frequency limit of the Kramers dispersion formula, a formula known as the Thomas–Kuhn (–Reiche) sum rule (Thomas, 1925, Kuhn, 1925, Reiche and Thomas, 1925). This formula replaces the Bohr–Sommerfeld condition as the fundamental quantization condition in the *Umdeutung* paper (see Sect. 3.5). Heisenberg obtained the sum rule by applying the procedure introduced in the *Umdeutung* paper for translating classical quantities into quantum-theoretical ones to (a derivative of) the Bohr–Sommerfeld quantization condition. He then showed that the sum rule also obtains by comparing the high-frequency limit of the Kramers dispersion formula with the polarization of a charged harmonic oscillator in the limit where $\nu >> \nu_0$ (see our Eq. (53)). In hindsight, we can see clearly in the *Umdeutung* paper how close Heisenberg came to recognizing the presence of the commutation relation between position and momentum in the sum rule serving as his quantization condition. As he told Kuhn:

I had written down, as the quantization rule the Thomas-Kuhn sum rule, but I had not recognized that this was just pq minus qp. That I had not seen.[202]

That he did not take this step is probably due to two important obstacles, one conceptual, the other technical. The conceptual framework of the entire *Umdeutung* paper is Lagrangian (as opposed to Hamiltonian): the essential problem is to find a quantum-theoretical reinterpretation of the classical concepts of position $x(t)$ and *velocity* $\dot{x}(t)$ of a particle. Indeed, the conventional symbol for momentum, p, appears only *once*

---

[201] At this point, Ladenburg refers to the treatments of dispersion in (Born, Heisenberg, and Jordan, 1925, pp. 330–338) [see also (Born and Jordan, 1930, pp. 240–250)], (Schrödinger, 1926), and (Dirac, 1927). For discussion of Schrödinger's wave-mechanical treatment of dispersion, see (Mehra and Rechenberg, 1982–2001, Vol. 5, pp. 789–796).

[202] See p. 12 of the transcript of session 5 of the AHQP interview with Heisenberg. See also p. 9 of the transcript of session 7. Cf. our discussion in Sect. 3.5. Heisenberg obtained his result by computing $(\oint p\,dq)_{n+1} - (\oint p\,dq)_n$.

in the entire paper, in the statement of the Bohr–Sommerfeld quantization condition (Eq. (12) in the paper). From this point on, p is replaced everywhere by $m\dot{x}(t)$. The canonical connection between position and momentum (so central, ironically, to the canonical perturbation theory that led to the dispersion formula in the first place[203]) seems simply to have vanished from Heisenberg's thinking at this point. The other, technical, obstacle was an inconvenient division of the sum over quantum states in the sum rule, which, though very natural from the point of dispersion theory, obscured its connection to a commutator, as we shall see below.

It will also become clear in the course of our modern derivation that the Kramers dispersion formula is an even more general result in modern quantum mechanics than it was in the old quantum theory. In the old quantum theory, it held for any non-degenerate multiply-periodic system with an unperturbed Hamiltonian such that the unperturbed motion can be solved in action-angle variables. In modern quantum mechanics, the result holds for any system with a Hermitian Hamilton operator such that the unperturbed part has a spectrum that is at least partially discrete. This helps to explain why the Kramers dispersion formula carries over completely intact from the old quantum theory to modern quantum mechanics.

7.1 Dispersion

In this subsection, we derive the Kramers dispersion formula in time-dependent perturbation theory. We then examine the high-frequency limit of this formula and discuss the role it played in (Heisenberg, 1925) as the fundamental quantization condition replacing the Bohr–Sommerfeld condition.

We consider a quantized charged system (valence electron) with states labeled by discrete indices $r, s, t, ...$, and with the Hamilton operator

$$H = H_0 + V(t) = H_0 + eEx \cos \omega t. \tag{111}$$

We want to calculate the first-order perturbation (in the electric field E) in the expectation value of the electron position in a particular state $|r, t\rangle$. It is convenient to work in the interaction picture.[204] The state $|r, t\rangle_{\text{int}}$ in the interaction picture is related to the state $|r, t\rangle$ in the Schrödinger picture via:

$$|r, t\rangle_{\text{int}} \equiv e^{i H_0 t/\hbar}|r, t\rangle. \tag{112}$$

An operator $O_{\text{int}}(t)$ in the interaction picture is related to the corresponding operator O in the Schrödinger picture via

$$O_{\text{int}}(t) \equiv e^{i H_0 t/\hbar} O e^{-i H_0 t/\hbar}. \tag{113}$$

---

[203] Of course, it was also central to (Dirac, 1925).

[204] The special role of H_0 in the time dependence of states and operators in the interaction picture is analogous to the choice of action-angle variables for the free rather than the full Hamiltonian in the version of canonical perturbation theory used by Van Vleck. This is what lies behind the close similarities between the calculations in this section and those in Sects. 5.1 and 6.2.

It follows that expectation values are the same in the two pictures:

$$_{\text{int}}\langle r, t | O_{\text{int}}(t) | r, t \rangle_{\text{int}} = \langle r, t | O | r, t \rangle. \tag{114}$$

The evolution of the states in the interaction picture is given by:

$$\frac{\partial}{\partial t} | r, t \rangle_{\text{int}} = \frac{i}{\hbar} e^{i H_0 t / \hbar} H_0 | r, t \rangle + e^{i H_0 t / \hbar} \frac{\partial}{\partial t} | r, t \rangle$$

$$= \frac{i}{\hbar} e^{i H_0 t / \hbar} (H_0 - H) | r, t \rangle, \tag{115}$$

where in the last step, we used the Schrödinger equation

$$\frac{\partial}{\partial t} | r, t \rangle = -\frac{i H}{\hbar} | r, t \rangle. \tag{116}$$

Since $H_0 - H = -V(t)$ (see Eq. (111)), we can write Eq. (115) as:

$$\frac{\partial}{\partial t} | r, t \rangle_{\text{int}} = -\frac{i}{\hbar} e^{i H_0 t / \hbar} V(t) e^{-i H_0 t / \hbar} | r, t \rangle_{\text{int}}$$

$$= -\frac{i}{\hbar} V_{\text{int}}(t) | r, t \rangle_{\text{int}}, \tag{117}$$

where we used Eqs. (112)–(113). To first order in $V_{\text{int}}(t)$ (i.e., to first order in the field E), the solution of Eq. (117) is

$$| r, t \rangle_{\text{int}} = | r, 0 \rangle_{\text{int}} - \frac{i}{\hbar} \int_0^t d\tau \, V_{\text{int}}(\tau) | r, 0 \rangle_{\text{int}}$$

$$= | r, 0 \rangle_{\text{int}} - \frac{ieE}{\hbar} \int_0^t d\tau \, x_{\text{int}}(\tau) \cos \omega \tau | r, 0 \rangle_{\text{int}}. \tag{118}$$

At $t = 0$ the states (and operators) in the interaction picture coincide with those in the Schrödinger picture. From now on we thus simply write $| r \rangle$ for $| r, 0 \rangle_{\text{int}}$. The dual ('bra') of the vector ('ket') in Eq. (118) is:

$$_{\text{int}}\langle r, t | = \langle r | + \frac{ieE}{\hbar} \int_0^t d\tau \cos \omega \tau \langle r | x_{\text{int}}(\tau). \tag{119}$$

To find the dipole moment $P_r(t)$ of the system in state r to first order in E, we calculate the first-order contribution to the expectation value of the displacement $\langle \Delta x \rangle_r$ in the state r induced by the field E:

$$\langle \Delta x \rangle_r \equiv {}_{\text{int}}\langle r, t | x_{\text{int}}(t) | r, t \rangle_{\text{int}} - \langle r | x_{\text{int}}(t) | r \rangle. \tag{120}$$

Inserting Eqs. (118)–(119) into this expression, we find:

$$\langle \Delta x \rangle_r = \frac{ieE}{\hbar} \int_0^t d\tau \, \langle r | \{ x_{\text{int}}(\tau) x_{\text{int}}(t) - x_{\text{int}}(t) x_{\text{int}}(\tau) \} | r \rangle \cos \omega \tau. \qquad (121)$$

Writing $\cos \omega \tau = \frac{1}{2}(e^{i\omega\tau} + e^{-i\omega\tau})$, and inserting a complete set of eigenstates of the unperturbed Hamiltonian H_0 $(1 = \sum_s |s\rangle\langle s|)$ between the two coordinate operators, we obtain

$$
\begin{aligned}
\langle \Delta x \rangle_r = {} & \frac{ieE}{2\hbar} \sum_s \int_0^t d\tau \left(\langle r | e^{iH_0\tau/\hbar} x e^{-iH_0\tau/\hbar} |s\rangle \langle s| e^{iH_0 t/\hbar} x e^{-iH_0 t/\hbar} |r\rangle \right. \\
& \left. - \langle r | e^{iH_0 t/\hbar} x e^{-iH_0 t/\hbar} |s\rangle \langle s| e^{iH_0\tau/\hbar} x e^{-iH_0\tau/\hbar} |r\rangle \right) e^{i\omega\tau} \\
& + (\omega \to -\omega) \\
= {} & \frac{ieE}{2\hbar} \sum_s \int_0^t d\tau \left(e^{i(E_r - E_s + \hbar\omega)\tau/\hbar} e^{i(E_s - E_r)t/\hbar} \right. \\
& \left. - e^{i(E_r - E_s)t/\hbar} e^{i(E_s - E_r + \hbar\omega)\tau/\hbar} \right) \langle r|x|s\rangle \langle s|x|r\rangle \\
& + (\omega \to -\omega).
\end{aligned}
\qquad (122)
$$

We introduce the notation $X_{rs} \equiv \langle r|x|s\rangle$ for the matrix elements of the coordinate operator. Note that these matrix elements in Eq. (122) are accompanied by time-development phases $e^{i(E_r - E_s)t/\hbar}$ of *purely harmonic form*: they are the precise correlates in modern quantum mechanics of the substitute oscillators of Ladenburg and Reiche (1923) or, equivalently, the virtual oscillators of BKS, as was clearly recognized, for instance, by Landé (1926) (cf. the discussion at the end of Sect. 4.3).[205]

Performing the time integral in Eq. (122), we find

$$
\begin{aligned}
\langle \Delta x \rangle_r = {} & \frac{eE}{2} \sum_s \left[\frac{e^{i(E_r - E_s + \hbar\omega)t/\hbar} - 1}{E_r - E_s + \hbar\omega} e^{i(E_s - E_r)t/\hbar} \right. \\
& \left. - \frac{e^{i(E_s - E_r + \hbar\omega)t/\hbar} - 1}{E_s - E_r + \hbar\omega} e^{i(E_r - E_s)t/\hbar} \right] X_{rs} X_{sr} \\
& + (\omega \to -\omega).
\end{aligned}
\qquad (123)
$$

(cf. Eqs. (42) and (44) in Sect. 6.1 and Eqs. (94) and (98) in Sect. 5.2). The coherent terms in $\langle \Delta x \rangle_r$, i.e. the terms with a time-dependence $e^{\pm i\omega t}$ (cf. Eq. (46) in Sect. 6.1

---

[205] Once the electromagnetic field itself is quantized, it becomes more natural to identify the virtual oscillators of BKS with the Fourier components of the quantized electromagnetic field, which correspond to time-dependent operators creating (or destroying) the photons emitted (or absorbed) by the atom.

and Eq. (101) in Sect. 5.2), are:

$$\langle \Delta x_{\text{coh}} \rangle_r = \frac{eE}{2} \sum_s X_{rs} X_{sr} e^{i\omega t} \left[\frac{1}{E_r - E_s + \hbar\omega} - \frac{1}{E_s - E_r + \hbar\omega} \right] \quad (124)$$
$$+ (\omega \to -\omega).$$

Using the Bohr frequency condition $\hbar\omega_{rs} = E_r - E_s$, we can write the expression in square brackets in Eq. (124) as:

$$\frac{1}{\hbar\omega_{rs} + \hbar\omega} - \frac{1}{\hbar\omega_{sr} + \hbar\omega} = \frac{2\omega_{rs}}{\hbar(\omega_{rs}^2 - \omega^2)}. \quad (125)$$

Inserting this result into Eq. (124) and noting that the terms proportional to $\sin \omega t$ vanish, we find the following result for the dipole moment of the system in state r (cf. Eq. (6) or (53))

$$P_r(t) = -e\langle \Delta x_{\text{coh}} \rangle_r = \frac{2e^2 E}{\hbar} \sum_s \frac{\omega_{sr} X_{rs} X_{sr}}{\omega_{sr}^2 - \omega^2} \cos \omega t. \quad (126)$$

The sum over s can naturally be separated into states s of higher energy than r, with $\omega_{sr} > 0$, and states t of lower energy, with $\omega_{rt} > 0$ ($\omega_{rt} = 0$ for $r = t$):

$$P_r = \frac{2e^2 E}{\hbar} \left(\sum_s \frac{\omega_{sr} X_{sr} X_{rs}}{\omega_{sr}^2 - \omega^2} - \sum_t \frac{\omega_{rt} X_{rt} X_{tr}}{\omega_{rt}^2 - \omega^2} \right) \cos \omega t. \quad (127)$$

If we recall the correspondence principle for emission (82), and identify D_s^2 with $3(X_\tau^s)^2 = 12A_\tau A_{-\tau}$ and the Fourier coefficients $A_\tau \to X_{sr}$, $A_{-\tau} \to X_{rs}$ we get

$$A_{s \to r} = \frac{64\pi^4 e^2}{hc^3} \nu_{sr}^3 X_{sr} X_{rs}, \quad (128)$$

whence we recover the original form (64) of the dispersion formula

$$P_r = \frac{c^3}{32\pi^4} E \cos \omega t \left(\sum_s \frac{A_{s \to r}}{\nu_{sr}^2 (\nu_{sr}^2 - \nu^2)} - \sum_t \frac{A_{r \to t}}{\nu_{rt}^2 (\nu_{rt}^2 - \nu^2)} \right). \quad (129)$$

Of course, the above identification of classical Fourier components with matrix elements of the position operator is at the core of Heisenberg's 1925 breakthrough.

Returning for a moment to Eq. (127), we see that in the Thomson limit where the frequency of incident radiation far exceeds the difference frequencies ω_{rs} for the

electron states r, s,[206] the polarization P_r becomes asymptotically

$$P_r \simeq -\frac{2e^2 E}{\hbar \omega^2} \left(\sum_s \omega_{sr} X_{sr} X_{rs} - \sum_t \omega_{rt} X_{rt} X_{tr} \right) \cos \omega t. \qquad (130)$$

The preceding equation is in content identical with the next to last (unnumbered) equation in Sect. 2 in (Heisenberg, 1925), where the Kramers dispersion theory is explicitly invoked. For large frequencies, we expect the polarization to approach our previously derived result (see Eq. (6) or (53)) for the polarization of a charged harmonic oscillator in the limit where $\nu >> \nu_0$:[207]

$$P_r = -\frac{e^2 E}{m \omega^2} \cos \omega t, \qquad (131)$$

Comparing Eq. (130) with Eq. (131) we find Eq. (16) in (Heisenberg, 1925):

$$h = 4\pi m \left(\sum_s \omega_{sr} X_{sr} X_{rs} - \sum_t \omega_{rt} X_{rt} X_{tr} \right). \qquad (132)$$

This result is first obtained by Heisenberg from the Bohr–Sommerfeld quantization condition by applying the quantum-theoretical transcription procedure, which was introduced in Sect. 1 of the *Umdeutung* paper and had been inspired by dispersion theory. It replaces the Bohr–Sommerfeld condition as the fundamental quantization constraint in Heisenberg's new theory. That the same result can be obtained directly from the high-frequency limit of the Kramers dispersion formula is clearly regarded by Heisenberg as strong evidence for the validity of his transcription procedure. Using Eq. (132), together with the formal transcription of the classical equation of motion, $\ddot{x} + f(x) = 0$ (Eq. (11) of the *Umdeutung* paper), Heisenberg (1925) asserts the possibility of "a complete determination not only of frequencies and energy values, but also of quantum-theoretical transition probabilities" (p. 268). As Heisenberg points out, Eq. (132) is completely equivalent to the sum rules for oscillator strengths given by Thomas (1925) and Kuhn (1925).[208]

The realization that Eq. (132) is equivalent to (the diagonal matrix elements of) the fundamental commutator relation $[P, X] = \hbar/i$ of modern quantum theory came shortly after this, in the work of Born and Jordan (1925). The recognition of Eq. (132)

---

[206] Or, alternatively, when the incident photon energy far exceeds the energy needed to ionize the electron, so that the latter can be regarded as essentially a free, unbound particle.

[207] This result is obtained in (Kuhn, 1925) by equating the energy scattered by an electron in the Thomson limit to the radiation emitted by an oscillating dipole according to the Larmor formula.

[208] Heisenberg's logic is slightly different from ours. Instead of pointing out that the high-frequency limit (130) of the Kramers dispersion formula and the well-established classical result (131) imply Heisenberg's quantization condition (132), Heisenberg (1925, pp. 269–270) points out that Eqs. (132) and (130) imply Eq. (131). This is only a cosmetic difference. The point of the exercise is still to show that the new quantization condition, found through *Umdeutung* of the derivative of the Bohr–Sommerfeld condition, follows from well-established results in Kramers' dispersion theory and classical electrodynamics. We are nonetheless grateful to Christoph Lehner for alerting us to this point.

as a commutator is mathematically obscured by the separation of the sum into states higher (s) and lower (t) than the given state r—a separation which is very natural given the history of the Kramers dispersion formula. If Heisenberg had applied his own transcription rules for associating classical variables with quantum two-index quantities to the momentum $P \equiv m\dot{X}$ in the unnumbered equation immediately following Eq. (13) in the *Umdeutung* paper (Heisenberg, 1925, p. 267), he would have found (using modern matrix notation):[209]

$$P_{rs} = im\omega_{rs} X_{rs}. \tag{133}$$

That Heisenberg did not write down this equation is probably, as we suggested above, because he was thinking in terms of the Lagrange rather than the Hamilton formalism. Rewriting Eq. (132) as a *single* sum over all states s, but splitting the sum into two equal pieces via the identity $2\omega_{sr} = \omega_{sr} - \omega_{rs}$, we find

$$h = 4\pi m \sum_s \omega_{sr} X_{rs} X_{sr}$$

$$= 2\pi m \sum_s (X_{rs}\omega_{sr} X_{sr} - \omega_{rs} X_{rs} X_{sr}) \tag{134}$$

$$= -2\pi i \sum_s (X_{rs} P_{sr} - P_{rs} X_{sr}),$$

where in the last step we used Eq. (133). In modern notation, this last expression is immediately recognized as the diagonal matrix element of the fundamental commutator $[X, P] = i\hbar$:

$$i\frac{h}{2\pi} = \langle r|XP - PX|r\rangle$$

$$= \sum_s (\langle r|X|s\rangle\langle s|P|r\rangle - \langle r|P|s\rangle\langle s|P|r\rangle). \tag{135}$$

Although Heisenberg recognized the significance of the noncommutativity of quantum-theoretic quantities in his formalism (see the last three paragraphs of Sect. 1), the simplicity of $x(t)p(t) - p(t)x(t)$ implied by his fundamental quantization relation (132) eluded him. He was thinking in terms of velocity rather than momentum. Moreover, even if he had been thinking in terms of momentum, the origin of his quantization condition in dispersion theory might well have prevented him from rewriting the summations the way we did in Eq. (134).

---

[209] Following Heisenberg's procedure in the *Umdeutung* paper for translating classical equations into quantum-mechanical ones, we would translate his classical equation for momentum, $m\dot{x} = m\sum_\alpha a_\alpha(n)i\alpha\omega_n e^{i\alpha\omega_n t}$, into the following quantum-mechanical equation: $P(n, n + \alpha) = ima(n, n + \alpha)\omega(n, n + \alpha)$. In modern notation, this becomes: $P_{rs} = imX_{rs}\omega_{rs}$ (no summation).

7.2 Spontaneous emission

To begin with, we note that we are dealing throughout with the dipole approximation, which is implicit in the 1924 work, corresponding to the regime where the wavelength of light is much larger than atomic dimensions (or equivalently, where photon momentum is much smaller than electron momentum). Once again, note that the notation of Van Vleck (1924b, Eq. (1)),

$$
\begin{aligned}
x &= \sum_{\tau_1 \tau_2 \tau_3} X(\tau_1, \tau_2, \tau_3) \cos\{2\pi(\tau_1\omega_1 + \tau_2\omega_2 + \tau_3\omega_3)t + \cdots\} \\
&= \sum \left\{ \frac{1}{2} X(\tau_1, \tau_2, \tau_3) e^{+2\pi i(\tau_1\omega_1 + \tau_2\omega_2 + \tau_3\omega_3)t + \cdots} \right. \\
&\quad \left. + \frac{1}{2} X(\tau_1, \tau_2, \tau_3) e^{-2\pi i(\tau_1\omega_1 + \tau_2\omega_2 + \tau_3\omega_3)t + \cdots} \right\},
\end{aligned}
\tag{136}
$$

implies that van Vleck's $D^2 = X^2 + Y^2 + Z^2$ (Van Vleck, 1924b, line following Eq. (8)) corresponds to four times the square of the matrix element of the quantum position operator appearing in the dipole transition formulas of modern quantum mechanics. For the latter we shall follow the treatment of (Baym, 1969, Ch. 13).

In the dipole approximation, the spontaneously emitted power per unit solid angle is given by (Baym, 1969, p. 282, Eq. 13–100), for emitted light of polarization vector $\vec{\lambda}$, in a transition from state r to state s:

$$
\begin{aligned}
\frac{\mathrm{d}P}{\mathrm{d}\Omega} &= \frac{\omega^4 e^2}{2\pi c^3} \langle r|\vec{\lambda} \cdot \vec{x}|s\rangle \langle s|\vec{\lambda} \cdot \vec{x}|r\rangle \\
&= \sum_{i,j=1}^{3} \frac{\omega^4 e^2}{2\pi c^3} \lambda_i \lambda_j \langle r|x_i|s\rangle \langle s|x_j|r\rangle.
\end{aligned}
\tag{137}
$$

Here (unlike Baym) we take real polarization vectors $\vec{\lambda}$ (plane polarized) rather than complex (circularly polarized) ones as our basis. We want the total spontaneously emitted power in any event, summed over the two possible polarizations for any momentum vector $\vec{k}$ of the emitted photon (so the basis of photon states is irrelevant). This requires the polarization sum

$$
\sum_{\lambda=1}^{2} \lambda_i \lambda_j = \delta_{ij} - \hat{k}_i \hat{k}_j, \quad (i, j = 1, 2, 3),
\tag{138}
$$

which follows from the fact that the two polarization vectors are any pair of orthogonal unit vectors perpendicular to the unit vector $\hat{k}$ along the photon direction. Finally, we want the total power emitted in any direction, so the polarization sum (138) must be integrated over all solid angles:

$$
\int \mathrm{d}\Omega_{\hat{k}} (\delta_{ij} - \hat{k}_i \hat{k}_j) = 4\pi \left(\frac{2}{3} \delta_{ij} \right).
\tag{139}
$$

The Einstein coefficient $A_{r \to s}$ in (Van Vleck, 1924b, Eqs. (5) and (9)) refers to a rate of photon emission (not energy emission) so we must divide Eq. (137) by $\hbar \omega$. Putting together the above results, we find:

$$A_{r \to s} = \frac{1}{\hbar \omega} \int d\Omega_{\hat{k}} \frac{dP}{d\Omega_{\hat{k}}} = \frac{\omega^4 e^2}{2\pi \hbar \omega c^3} \frac{8\pi}{3} \sum_i \langle r | x_i | s \rangle \langle s | x_i | r \rangle. \qquad (140)$$

Using the notation $X_{rs} \equiv \langle r | x | s \rangle$, etc. for the matrix elements of position introduced above we can rewrite this as:

$$A_{r \to s} = \frac{\omega^4 e^2}{2\pi \hbar \omega c^3} \frac{8\pi}{3} \left(|X_{rs}|^2 + |Y_{rs}|^2 + |Z_{rs}|^2 \right). \qquad (141)$$

Replacing the matrix elements X_{rs}, Y_{rs}, and Z_{rs} by the amplitude D_r in the correspondence limit as indicated in the preceding section (cf. the remarks preceding Eq. (128)) and substituting $\omega = 2\pi \nu$, we arrive at:

$$A_{r \to s} = \frac{16\pi^4 e^2 \nu^3}{3hc^3} D_r^2. \qquad (142)$$

D_r^2 is the amplitude defined in (Van Vleck, 1924b) immediately following Eq. (8), to be replaced by $D_r(\tau_1, \tau_2, \tau_3)^2$ in Eq. (9), with which Eq. (142) is seen to be identical.

7.3 Absorption

The Einstein formula for absorption (Van Vleck, 1924b, Eq. (6)), when combined with the stimulated emission ("negative absorption") term to yield (ibid., Eq. (15)), leads directly to the correspondence limit result (ibid., Eq. (16)). Here, we check the identity of Eq. (15) in (Van Vleck, 1924b) (more precisely, the unnumbered equation immediately following this one) with the modern absorption calculation given in (Baym, 1969). For the rate of absorption of light leading to a transition from state s to (higher) state r, (Baym, 1969, Eq. 13–40) reads (in the dipole approximation, $\vec{j}_{\vec{k}} \to \vec{p}/m$):

$$\Gamma_{s \to r}^{abs} = \frac{2\pi e^2}{\hbar^2 c^2} \frac{\omega^2}{(2\pi c)^3} \int d\Omega_{\hat{k}} \sum_{\lambda} \langle s | \vec{\lambda} \cdot \frac{\vec{p}}{m} | r \rangle \langle r | \vec{\lambda} \cdot \frac{\vec{p}}{m} | s \rangle |A_{\vec{k}\lambda}|^2. \qquad (143)$$

As usual, in the dipole approximation we can use (Baym, 1969, Eq. 13–98) to replace matrix elements of the momentum operator with those of the coordinate operator (using the equations of motion). For Hamiltonians of the form $H = (\vec{p}^2/2m) + V(\vec{x})$,

$$[H, x_j] = \frac{1}{2m}[p_i p_i, x_j] = \frac{1}{m} p_i[p_i, x_j] = \frac{p_i}{m} \frac{\hbar}{i} \delta_{ij} = \frac{\hbar}{i} \frac{p_j}{m}, \qquad (144)$$

whence

$$
\begin{aligned}
\langle r|\frac{\vec{p}}{m}|s\rangle &= \frac{\mathrm{i}}{\hbar}\langle r|[H,\vec{x}]|s\rangle \\
&= \frac{\mathrm{i}}{\hbar}(E_r - E_s)\langle r|\vec{x}|s\rangle \\
&= \mathrm{i}\omega\langle r|\vec{x}|s\rangle,
\end{aligned}
\tag{145}
$$

where $\hbar\omega = E_r - E_s$. Once again, in Eq. (145), we see the "monstrous" difference frequencies characteristic of quantum theory, which wreaked havoc on classical interpretations of radiation phenomena, making their appearance in the modern formalism. Accordingly, Eq. (143) becomes

$$
\Gamma_{s\to r}^{\mathrm{abs}} = \frac{2\pi e^2}{\hbar^2 c^2}\frac{\omega^4}{(2\pi c)^3}\int \mathrm{d}\Omega_{\hat{k}}\sum_\lambda \langle s|\lambda_i x_i|r\rangle\langle r|\lambda_j x_j|s\rangle|A_{\vec{k}\vec{\lambda}}|^2.
\tag{146}
$$

Now we are going to assume that the ambient light is unpolarized and isotropic so that the squared amplitude $|A_{\vec{k}\vec{\lambda}}|^2$ is in fact independent of λ, $\hat{k}$, and the only angular dependence comes in via the polarization vectors. The angle average of the polarization sum in Eq. (146) can then be performed as in Eq. (139) to yield

$$
\Gamma_{s\to r}^{\mathrm{abs}} = \frac{4\pi e^2}{3\hbar^2 c^2}\frac{\omega^4}{(2\pi c)^3}\langle s|x_i|r\rangle\langle r|x_i|s\rangle\int \mathrm{d}\Omega_{\hat{k}}|A_{\vec{k}\vec{\lambda}}|^2.
\tag{147}
$$

Next, we need to establish the relation between the squared mode amplitudes $|A_{\vec{k}\vec{\lambda}}|^2$ and the specific energy density function $\rho(\nu)$ defined as the energy per unit volume per unit frequency interval. The mode amplitudes $A_{\vec{k}\vec{\lambda}}$ correspond to discrete modes for electromagnetic radiation in a box of volume V, with each mode contributing energy density

$$
\frac{1}{V}|A_{\vec{k}\vec{\lambda}}|^2\frac{\omega}{2\pi c^2}
\tag{148}
$$

(Baym, 1969, Eq. 13–14). As the box volume goes to infinity we have the usual correspondence

$$
\frac{1}{V}\sum_k \to \int \frac{k^2 \mathrm{d}k\,\mathrm{d}\Omega_{\hat{k}}}{(2\pi)^3},
\tag{149}
$$

so that the total energy density *between frequency ν and frequency $\nu + \Delta\nu$* is

$$
\begin{aligned}
\rho(\nu)\Delta\nu &= \frac{1}{V}\sum_{2\pi\nu<kc<2\pi(\nu+\Delta\nu)} 2|A_{\vec{k}\vec{\lambda}}|^2\frac{\omega^2}{2\pi c^2} \\
&\to \frac{1}{(2\pi)^3}\int \mathrm{d}\Omega_{\hat{k}}\int_{2\pi\nu/c}^{2\pi(\nu+\Delta\nu)/c} \mathrm{d}k\;k^2\frac{\omega^2}{2\pi c^2}2|A_{\vec{k}\vec{\lambda}}|^2.
\end{aligned}
\tag{150}
$$

Note that although we continue to write the mode amplitudes $A_{\vec{k}\lambda}$ as depending on polarization and momentum vector of the photon, we are really assuming that there is no dependence on the polarization or photon *direction*. Hence the factor of 2, with no remaining sum over λ. Equation (150) gives

$$\rho(\nu)\Delta\nu = \frac{1}{(2\pi)^3}\frac{2\pi}{c}k^2\frac{\omega^2}{2\pi c^2}2\int d\Omega_{\hat{k}}|A_{\vec{k}\lambda}|^2\Delta\nu, \qquad (151)$$

or, equivalently

$$\int d\Omega_{\hat{k}}|A_{\vec{k}\lambda}|^2 = \frac{4\pi^3c^5}{\omega^4}\rho(\nu). \qquad (152)$$

Inserting Eq. (152) into Eq. (147) and multiplying by $\hbar\omega$ to get the rate of energy absorption (instead of the number rate of photon absorption) we find, using the usual association of squares of matrix elements of the position operator to the classical orbit amplitude $\frac{1}{4}D_r^2$,

$$\begin{aligned}\hbar\omega\Gamma_{s\to r}^{\text{abs}} &= \frac{4\pi e^2\omega}{3\hbar c^2}\frac{\omega^4}{(2\pi c)^3}\frac{4\pi^3c^5}{\omega^4}\rho(\nu)\frac{1}{4}D_r^2 \\ &= \frac{2\pi^3e^2}{3h}\nu\rho(\nu)D_r^2,\end{aligned} \qquad (153)$$

which coincides with the first term in van Vleck's equation (Van Vleck, 1924b, the equation following Eq. (15)) for the part of the total absorption rate due to upward transitions. Of course, the second (negative absorption, or stimulated emission) term is of exactly the same form (with a minus sign) due to the symmetry of the Einstein B coefficients.

8 Conclusion

Our study of Van Vleck's two-part paper on the application of the correspondence principle to the interaction of matter and radiation (Van Vleck, 1924b,c) has led us to consider three clusters of questions. First, there are questions about the paper itself. What made Van Vleck decide to work in this area? He had not published on radiation theory before. And—as one is inevitably tempted to ask—why did Van Vleck not take the next step and arrive at something like matrix mechanics? That gets us to the second cluster of questions, about the developments in quantum theory that provide the natural context for Van Vleck's work, especially the transition from the old quantum theory of Bohr and Sommerfeld to matrix mechanics. What was important for this development and what was not? The third group of questions concerns the relative importance of American contributions to these developments. In this final section we collect the (partial) answers we have found to these biographical, conceptual, and sociological questions.

Let us first dispose of the issue of American contributions to early quantum theory. Since we focused on the work of only two individuals, Van Vleck and Slater, we are in no position to draw strong conclusions. Still, it seems safe to say that our study

supports the thesis of Sam Schweber (1986) and others that, by the early 1920s, the United States had a homegrown tradition in quantum theory, which, to be sure, was reinforced, but certainly not created by the influx of European émigrés in the 1930s. We are less sanguine about the thesis of Alexi Assmus (1992) that American theorists contributed mainly to molecular rather than to atomic physics, although she may be right that Slater and Van Vleck are just exceptions to the rule (see Sect. 2.4). However, we did come across several other contributions (some admittedly minor) to atomic theory by Americans (Breit, Davisson, Hoyt, Kemble) or by Europeans working in America (Epstein, Swann). And we do want to emphasize that the contributions to atomic theory by our main protagonists were absolutely first rate, even if they did not always receive the recognition they deserved from their European colleagues (see the correspondence between Born and Van Vleck cited in Sects. 2.4 and 5.2). The quickly refuted but highly influential Bohr–Kramers–Slater (BKS) theory was built around Slater's idea of a virtual radiation field emitted by an atom while in a stationary state (see Sect. 4.1). The derivation of a correspondence principle of absorption for a general non-degenerate multiply-periodic system, the centerpiece of (Van Vleck, 1924b,c), is a *tour de force* that may well have been the most sophisticated application of the correspondence principle in the old quantum theory. All in all, the Americans had definitely established a presence in atomic theory by the early 1920s. In the period we examined, they were certainly more prominent than the British, not to mention the French. Ultimately, however, the decisive steps were taken in Europe, not in the United States.

This brings us to the question of why Van Vleck stopped short of these decisive steps. Before we offer our best guess as to why Van Vleck did not do what he did not do, we want to say a few words about why he did what he did. His papers on the correspondence principle for absorption (Van Vleck, 1924a,b,c) constitute his first foray into quantum radiation theory. His earlier publications had dealt with such topics as the extension of Bohr's model of the atom to helium and the specific heat of molecular hydrogen. The formulation of a correspondence principle for absorption, Van Vleck told Kuhn in his interview for the AHQP in 1963, had been triggered by a comment of his Minnesota colleague Breit (see also Van Vleck, 1924a, p. 28). Breit's remark, we conjectured (in Sect. 5.3), may have directed Van Vleck to the work of Ladenburg and Reiche (1923), who proposed quantum formulae for emission, absorption, and dispersion, invoking but not always correctly implementing the correspondence principle. Van Vleck likewise proposed quantum formulae for emission and absorption and used his considerable expertise in classical mechanics to show that these formulae as well as the Kramers dispersion formula merged with the classical formulae in the limit of high quantum numbers.

So why did Van Vleck not take the next step? The trivial explanation is that he was too busy working on his *Bulletin* for the *National Research Council* on the old quantum theory (Van Vleck, 1926) to pursue his own research. But even if he had not been burdened by this *Bulletin*, we seriously doubt that Van Vleck would have done what Heisenberg did—as he himself acknowledged both in a biographical statement prepared for the AHQP and in his interview for the project (see Sect. 1.1). Van Vleck, it seems, was too wedded to the orbits of the Bohr–Sommerfeld theory to completely discard them, a prerequisite for Heisenberg's *Umdeutung*. This is clear at several points

in (Van Vleck, 1924b). At the end of Sect. 1, for instance, we find a formula expressing
the Einstein coefficient $A_{r\to s}$ as an average over the frequencies of orbits, not allowed
by the Bohr–Sommerfeld quantization condition, between the initial state r and the
final state s. Section 2 of the paper is devoted to "a correspondence principle for *orbital
distortions*" (Van Vleck, 1924b, p. 334, our emphasis). On the issue of how seriously
one should take the orbits of the Bohr–Sommerfeld theory, Van Vleck might have
benefited from direct contact with the Europeans. He had the distinct disadvantage
of *reading* Sommerfeld instead of *talking* to Bohr and his circle.[210] Bohr and Pauli
certainly prepared Heisenberg for the step of leaving orbits behind.

The emphasis on observable quantities in the *Umdeutung* paper, however, struck a
chord with Van Vleck, who had been primed for such a positivist turn by his Harvard
teacher Bridgman.[211] Explaining the new quantum mechanics in *Chemical Reviews*
in 1929,[212] he wrote:

> Heisenberg's epoch-making development of the matrix theory was spurred by
> Born's repeated emphasis to his colleagues at Göttingen that the reason the old
> quantum theory was then (1925) failing was that we were all too anxious to use
> the same concepts of space and time within the atom as in ordinary measurable
> large-scale events. …the concepts of distance and time have a meaning only when
> we tell how they can be measured. This is very nicely emphasized in Bridgman's
> recent book, "The Logic of Modern Physics" [Bridgman, 1927] …one cannot
> use a meter stick to measure the diameter of an atom, or an alarm clock to record
> when an electron is at the perihelion of its orbit. Consequently we must not
> be surprised if within the atom the correlation of space and time is something
> which cannot be visualized, and that models cannot be constructed with the same
> kind of mechanics as Henry Ford uses in designing an automobile. …The goal
> of theoretical physics and chemistry must ever be to explain observable rather
> than unobservable phenomena …What the physicist observes about an atom is
> primarily its radiations …We may say that we have a sound atomic theory when
> we have a set of a small number of mathematical postulates from which these
> observed things can be calculated correctly, even though it forces us to discard
> the usual space-time models (Van Vleck, 1929, p. 468).

Van Vleck was thus ready enough to give up orbits once Heisenberg had shown the
way. He failed to take this step on his own.

---

[210] According to Assmus (1992, pp. 8, 15), Americans had a tendency to follow Sommerfeld rather than Bohr anyway.

[211] In the biographical note written for the AHQP, Van Vleck wrote: "I suspect that Bridgman's operational philosophy may have subconsciously influenced my approach to theoretical physics." At a ceremony honoring Bridgman's 1946 Nobel prize, Slater went as far as suggesting a genetic link between Bridgman's operationalism and Heisenberg's uncertainty principle! Schweber (1990) quotes Slater as saying on this occasion: "It is very likely that this principle, so much like Bridgman's attitude, is actually derived to a very considerable extent from Bridgman's thinking" (p. 391).

[212] For the benefit of the chemists, Van Vleck (1929) compared a matrix to a baseball schedule: "the entry in row 3 and column 2, for instance, gives information about a transition between a 3 and 2 quantum state, just as the analogous baseball entry does about the meetings between teams 3 and 2" (p. 469).

The study of Van Vleck's paper illuminates various aspects of the transition from the old quantum theory to matrix mechanics that tend to get obscured when one approaches these developments through, say, (Kramers and Heisenberg, 1925). Most importantly perhaps, following (Van Vleck, 1924b,c) rather than (Kramers and Heisenberg, 1925) or (Born, 1924), we were able to give a transparent and explicit version of the derivation needed to show that the crucial Kramers dispersion formula reduces to the classical formula in the limit of high quantum numbers (see Sects. 5.1–5.2 for the special case of a simple harmonic oscillator, Sect. 6.2 for the generalization to arbitrary non-degenerate multiply-periodic systems, and Sect. 7.1 for a closely analogous derivation of the Kramers formula in modern quantum mechanics). That Van Vleck confirmed the Kramers dispersion formula without relying on the Bohr–Kramers–Slater (BKS) theory makes it particularly clear that matrix mechanics grew directly out of dispersion theory and that BKS was mainly a sideshow (see Sect. 4). The only element of the BKS theory used by Van Vleck is the concept of virtual oscillators. We saw that this concept actually predates BKS. 'Virtual oscillators' was Bohr's new name for the substitute oscillators introduced into dispersion theory the year before and at Bohr's suggestion by Ladenburg and Reiche (1923). In addition to popularizing the notion of virtual oscillators, BKS may have contributed to instilling skepticism about the electron orbits of the Bohr–Sommerfeld theory. In that sense, it might have helped Van Vleck had he embraced BKS more wholeheartedly. Overall, however, we argued that BKS played only a limited role in the breakthrough to matrix mechanics. The broad acceptance of Einstein's light-quantum concept following the discovery of the Compton effect played no role in this development. Physicists working in dispersion theory, while accepting the Compton effect as decisive evidence for light quanta, happily continued to treat light as a wave phenomenon.

What was it about dispersion theory that made it so important for the transition from the Bohr–Sommerfeld theory to the theory of Heisenberg's *Umdeutung* paper? As we suggested in the introduction of Sect. 3, the answer is that the discrepancy between orbital frequencies and radiation frequencies—one of the most radical, if not *the* most radical aspect of the Bohr model of the atom—manifested itself glaringly and unavoidably in dispersion theory. The natural approach to adapting the successful classical dispersion theory of Helmholtz, Lorentz and Drude to Bohr's new theory inevitably led to a dispersion formula with resonance poles at the orbital frequencies (Sommerfeld, 1915b, Debye, 1915, Davisson, 1916, Epstein, 1922), whereas experiment clearly indicated that the resonance poles should be at the radiation frequencies, associated in Bohr's theory with transitions between orbits. Employing Einstein's *A* and *B* coefficients and Bohr's correspondence principle (in conjunction with techniques from celestial mechanics customized to the problems at hand) and building on pioneering work by Ladenburg (1921) and Ladenburg and Reiche (1923), Kramers (1924a,b) constructed a quantum formula for dispersion with resonance poles at the transition frequencies rather than at the orbital frequencies and claimed that this formula merged with the classical formula in the limit of high quantum numbers. Van Vleck (1924b,c) and Born (1924) were the first to publish an explicit proof that the Kramers quantum formula does indeed merge with the classical formula for dispersion in a general non-degenerate multiply-periodic system in the correspondence limit. The three key moves in translating the classical formula into a quantum-theoretical one were to (1) replace

orbital frequencies by transition frequencies; (2) relate amplitudes to Einstein's *A* coefficients; and (3) replace derivatives with respect to the action variable by difference quotients. The first move goes back to the embryonic version of the correspondence principle in (Bohr, 1913) (Heilbron and Kuhn, 1969, pp. 274–275). Ladenburg (1921) introduced the second move. It was made more precise by Kramers and Van Vleck (cf. Jordan's remarks quoted in Sect. 2.4). Born (1924) is usually credited with the third move and the rule for replacing derivatives by difference quotients is sometimes even called "Born's correspondence rule" (Jammer, 1966, p. 193) or "Born's discretizing rule" (Cassidy, 1991, p. 181). It was found independently, however, by both Kramers and Van Vleck (see the discussion at the end of Sect. 5.2).

The Kramers dispersion formula no longer contains any reference to the orbits of the Bohr–Sommerfeld theory, but only to transitions between them. This signaled to Heisenberg that orbits could be dispensed with altogether. Dispersion theory further told Heisenberg how to generate quantum formulae from classical formulae in his *Umdeutung* scheme. The procedure consisted of the same three moves listed above: one had to replace (1) classical frequencies (more specifically: the Fourier overtones of the classical mechanical motion) by quantum transition frequencies; (2) classical amplitudes associated with definite orbits by quantum transition amplitudes associated with pairs of stationary states; and (3) derivatives by difference quotients.[213] Dispersion theory also furnished the fundamental quantization condition for Heisenberg's new theory. Heisenberg formulated this condition by applying his *Umdeutung* procedure to the Bohr–Sommerfeld quantum condition, which was no longer acceptable because of its explicit reference to orbits. That Heisenberg's new condition also emerged in the high-frequency limit of the Kramers dispersion formula (see Sect. 7.1) convinced him that he had found a sensible replacement for the Bohr–Sommerfeld condition. The relevant formula had been found in quantum dispersion theory before and was known as the Thomas–Kuhn(–Reiche) sum rule (Thomas, 1925, Kuhn, 1925, Reiche and Thomas, 1925). Van Vleck actually had been the first to find this rule, even though he did not emphasize the result because he thought it was problematic (see Sect. 3.5). In his later years, Van Vleck nonetheless used to mention this achievement with pride to several of his colleagues (Roger Stuewer, private communication). The Kramers dispersion formula and the Thomas–Kuhn sum rule are the critical physical ingredients in the first two sections of (Heisenberg, 1925), in which the *Umdeutung* procedure is motivated. Van Vleck was fully cognizant of these same ingredients by mid-1924. Van Vleck can thus truly be said to have been on the verge of *Umdeutung* in Minnesota in the summer of 1924.

---

[213] Another factor in Van Vleck's failure to take the next step might have been that he used these correspondence-principle replacements to check rather than to construct quantum formulae. However, even though Born had emphasized the constructive use of these replacements (see Sect. 5.2), he did not do what Heisenberg did either.

References

Aitchison, I. J. R., D. A. McManus, and T. M. Snyder (2004). Understanding Heisenberg's "magical" paper of July 1925: A new look at the calculational details. *American Journal of Physics* 72: 1370–1379.

Assmus, A. (1992). The Americanization of molecular physics. *Historical Studies in the Physical and Biological Sciences* 23: 1–34.

Baym, G. (1969). *Lectures on quantum mechanics.* Reading, MA: Addison-Wesley.

Bohr, N. (1913). On the constitution of atoms and molecules (Part I). *Philosophical Magazine* 26: 1–25.

Bohr, N. (1972–1996). *Collected works.* 9 Vols. Edited by L. Rosenfeld *et al.* Amsterdam: North-Holland.

Bohr, N., H. A. Kramers, and J. C. Slater (1924). The quantum theory of radiation. *Philosophical Magazine* 47: 785–802. Page references to reprint in (Van der Waerden, 1968, pp. 159–176).

Born, M. (1924). Über Quantenmechanik. *Zeitschrift für Physik* 26: 379–395. Page references are to the English translation in (Van der Waerden, 1968, pp. 181–198).

Born, M. (1925). *Vorlesungen über Atommechanik.* Berlin: Springer.

Born, M., and P. Jordan (1925). Zur Quantenmechanik. *Zeitschrift für Physik* 34: 858–888. English translation in (Van der Waerden, 1968, pp. 277–306).

Born, M., and P. Jordan (1930). *Elementare Quantenmechanik.* Berlin: Springer.

Born, M., W. Heisenberg, and P. Jordan (1925). Zur Quantenmechanik II. *Zeitschrift für Physik* 35: 557–615. English translation in (Van der Waerden, 1968, pp. 321–385).

Bridgman, P. W. (1927). *The logic of modern physics.* New York: MacMillan.

Cassidy, D. C. (1991). *Uncertainty. The life and science of Werner Heisenberg.* New York: Freeman.

Davisson, C. J. (1916). The dispersion of hydrogen and helium on Bohr's theory. *Physical Review* 8: 20–27.

Debye, P. (1915). Die Konstitution des Wasserstoff-moleküls. *Sitzungsberichte der mathematisch-physikalischen Klasse der Königlichen Bayerischen Akademie der Wissenschaften zu München.* 1–26.

Dirac, P. A. M. (1925). The fundamental equations of quantum mechanics. *Proceedings of the Royal Society of London* A109: 642–653.

Dirac, P. A. M. (1927). The quantum theory of dispersion. *Proceedings of the Royal Society of London* A114: 710–728.

Dresden, M. (1987). *H. A. Kramers: between tradition and revolution.* New York: Springer.

Einstein, A. (1916a). Strahlungs-Emission und -Absorption nach der Quantentheorie. *Deutsche Physikalische Gesellschaft. Verhandlungen* 18: 318–323. Reprinted in facsimile as Doc. 34 in (Einstein, 1987–2004, Vol. 6).

Einstein, A. (1916b). Zur Quantentheorie der Strahlung. *Physikalische Gesellschaft Zürich. Mitteilungen* 18: 47–62. Reprinted as (Einstein, 1917) and (in facsimile) as Doc. 38 in (Einstein, 1987–2004, Vol. 6).

Einstein, A. (1917). Zur Quantentheorie der Strahlung. *Physikalische Zeitschrift* 18: 121–128. Reprint of (Einstein, 1916b). English translation in (Van der Waerden, 1968, pp. 63–77).

Einstein, A. (1987–2004). *The collected papers of Albert Einstein.* 9 Vols. Edited by J. Stachel *et al.* Princeton: Princeton University Press.

Epstein, P. S. (1922). Die Störungsrechnung im Dienste der Quantentheorie. III. Kritische Bemerkungen zur Dispersionstheorie. *Zeitschrift für Physik* 9: 92–110.

Fellows, F. H. (1985). *J. H. Van Vleck: The early life and work of a mathematical physicist.* Ph.D. Thesis, University of Minnesota.

Feynman, R. P., R. B. Leighton, and M. Sands (1964). *The Feynman lectures on physics.* 3 Vols. Reading, MA: Addison-Wesley.

Goldstein, H. (1980). *Classical mechanics.* 2nd ed. Reading, MA: Addison-Wesley.

Heilbron, J. L., and T. S. Kuhn (1969). The genesis of the Bohr atom. *Historical Studies in the Physical Sciences* 1: 211–290.

Heisenberg, W. (1925). Über die quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen. *Zeitschrift für Physik* 33: 879–893. English translation in (Van der Waerden, 1968, pp. 261–276). Page references are to this volume.

Jackson, J. D. (1975). *Classical electrodynamics.* 2nd ed. New York: Wiley.

Jammer, M. (1966). *The conceptual development of quantum mechanics.* New York: McGraw-Hill.

Kramers, H. A. (1924a). The law of dispersion and Bohrs theory of spectra. *Nature* 113: 673–676. Reprinted in (Van der Waerden, 1968, pp. 177–180).

Kramers, H. A. (1924b). The quantum theory of dispersion. *Nature* 114: 310–311. Reprinted in (Van der Waerden, 1968, pp. 199–201).

Kramers, H. A., and W. Heisenberg (1925). Über die Streuung von Strahlung durch Atome. *Zeitschrift für Physik* 31: 681–707. English translation in (Van der Waerden, 1968, pp. 223–252).

Kretschmann, E. (1917). Über den physikalischen Sinn der Relativitätspostulate. A. Einsteins neue und seine ursprüngliche Relativitätstheorie. *Annalen der Physik* 53: 575–614.

Kretschmann, E. (1921). Über die Wirkung des Planckschen Oszillators auf die spektrale Energieverteilung des Strahlungsfeldes. *Annalen der Physik* 65: 310–334.

Kuhn, W. (1925). Über die Gesamtstärke der von einem Zustande ausgehenden Absorptionslinien. *Zeitschrift für Physik* 33: 408–412. English translation in (Van der Waerden, 1968, pp. 253–257).

Ladenburg, R. (1921). Die quantentheoretische Deutung der Zahl der Dispersionselektronen. *Zeitschrift für Physik* 4: 451–468. English translation in (Van der Waerden, 1968, pp. 139–157).

Ladenburg, R. (1928). Untersuchungen über die anomale Dispersion angeregter Gase. I. Teil. Zur Prüfung der quantentheoretischen Dispersionsformel. *Zeitschrift für Physik* 48: 15–25.

Ladenburg, R., and F. Reiche (1923). Absorption, Zerstreuung und Dispersion in der Bohrschen Atomtheorie. *Die Naturwissenschaften* 11: 584–598.

Landé, A. (1926). Neue Wege der Quantentheorie. *Die Naturwissenschaften* 14: 455–458.

Matzner, R. A., and L. C. Shepley (1991). *Classical Mechanics.* Englewood Cliffs, NJ: Prentice Hall.

MacKinnon, E. M. (1977). Heisenberg, models, and the rise of matrix mechanics. *Historical Studies in the Physical Sciences* 8: 137–188.

Mehra, J., and H. Rechenberg (1982–2001). *The historical development of quantum theory.* 6 Vols. New York, Berlin: Springer.

Planck, M. (1921). *Vorlesungen über die Theorie der Wärmestrahlung.* 4th ed. Leipzig: Barth.

Reiche, F., and W. Thomas (1925). Über die Zahl der Dispersionselektronen, die einem stationären Zustand zugeordnet sind. *Zeitschrift für Physik* 34: 510–525.

Schrödinger, E. (1926). Quantisierung als Eigenwertproblem (Vierte Mitteilung). *Annalen der Physik* 81: 109–139. English translation on pp. 102–123 of E. Schrödinger, *Collected papers on wave mechanics* (London: Blackie & Son, 1928).

Schweber, S. S. (1986). The empiricist temper regnant: Theoretical physics in the United States 1920–1950. *Historical Studies in the Physical and Biological Sciences* 17: 55–98.

Schweber, S. S. (1990). The young John Clarke Slater and the development of quantum chemistry. *Historical Studies in the Physical and Biological Sciences* 20: 339–406.

Smekal, A. (1923). Zur Quantentheorie der Dispersion. *Die Naturwissenschaften* 11: 873–875.

Sommerfeld, A. (1915). Die allgemeine Dispersionsformel nach dem Bohrschen Model. pp. 549–584 in: K. Bergwitz (ed.), *Festschrift Julius Elster und Hans Geitel.* Braunschweig. Reprinted on pp. 136–171 of Vol. 3 of: F. Sauter (ed.), *Arnold Sommerfeld: Gesammelte Schriften.* 4 Vols. Edited by F. Sauter. Braunschweig: Vieweg, 1968.

Sopka, K. R. (1988). *Quantum physics in America. The years through 1935.* Tomash Publishers/American Institute of Physics.

Thomas, W. (1925). Über die Zahl der Dispersionselektronen, die einem stationären Zustande zugeordnet sind (Vorläufige Mitteillung). *Die Naturwissenschaften* 13: 627.

Van der Waerden, B. L., ed. (1968). *Sources of quantum mechanics.* New York: Dover.

Van Vleck, J. H. (1922). The normal helium atom and its relation to the quantum theory. *Philosophical Magazine* 44: 842–869.

Van Vleck, J. H. (1924a). A correspondence principle for absorption. *Journal of the Optical Society of America* 9: 27–30.

Van Vleck, J. H. (1924b). The absorption of radiation by multiply periodic orbits, and its relation to the correspondence principle and the Rayeigh-Jeans law. Part I. Some extensions of the correspondence principle. *Physical Review* 24: 330–346. Reprinted in (Van der Waerden, 1968, pp. 203–222).

Van Vleck, J. H. (1924c). The absorption of radiation by multiply periodic orbits, and its relation to the correspondence principle and the Rayeigh-Jeans law. Part II. Calculation of absorption by multiply periodic orbits. *Physical Review* 24: 347–365.

Van Vleck, J. H. (1926). *Quantum principles and line spectra.* Washington, D. C.: National Research Council (Bulletin of the National Research Council 10, Part 4).

Van Vleck, J. H. (1928). The correspondence principle in the statistical interpretation of quantum mechanics. *Proceedings of the National Academy of Sciences* 14: 178–188.

Van Vleck, J. H. (1929). The new quantum mechanics. *Chemical Reviews* 5: 467–507.

Van Vleck, J. H. (1974). Acceptance speech. *Koninklijke Nederlandse Akademie van Wetenschappen. Bijzondere bijeenkomst der afdeling natuurkunde …28 september 1974 …voor de plechtige uitreiking van de Lorentz-medaille aan Prof. Dr. J. H. Van Vleck.*

Van Vleck, J. H., and D. L. Huber (1977). Absorption, emission, and line breadths: A semihistorical perspective. *Reviews of Modern Physics* 49: 939–959.

THE EARLY HISTORY OF
STRING THEORY AND SUPERSYMMETRY

John H. Schwarz

California Institute of Technology

Pasadena, CA 91125, USA

Abstract

This lecture presents a brief overview of the early history of string theory and supersymmetry. It describes how the S-matrix theory program for understanding the strong nuclear force evolved into superstring theory, which is a promising framework for constructing a unified quantum theory of all forces including gravity. The period covered begins with S-matrix theory in the mid 1960s and ends with the widespread acceptance of superstring theory in the mid 1980s. Further details and additional references can be found in Schwarz (2007).

1   S-Matrix Theory

In UC Berkeley, where I was a graduate student in the mid 1960s, Geoffrey Chew (my thesis advisor), Stanley Mandelstam, and others focussed their efforts on constructing a theory of the strong nuclear force, *i.e.,* a theory of hadrons. Chew's approach to understanding the strong nuclear force was based on *S-matrix theory.* He argued that quantum field theory, which was so successful in describing QED, was inappropriate for describing a strongly interacting theory, where a weak-coupling perturbation expansion would not be useful. One reason for holding this view was that none of the hadrons seemed more fundamental than any of the others. Therefore a field theory that singled out some subset of the hadrons did not seem sensible. Also, it seemed impossible to formulate a quantum field theory with a fundamental field for every hadron. Chew spoke of *nuclear democracy* and the *bootstrap principle* to describe this situation. Chew advocated focussing attention on physical quantities, especially the S Matrix, which describes on-mass-shell scattering amplitudes. The goal was to develop a theory that would determine the hadron spectrum and hadronic S matrix.

The quark concept also arose during this period, but the prevailing opinion in the mid 1960s was that quarks are mathematical constructs, rather than physical entities, whose main use is as a mathematical technique for understanding symmetries and quantum numbers. The SLAC deep inelastic scattering experiments in the late 1960s made it clear that quarks and gluons are physical (confined) particles. It was then natural to try to base a quantum field theory on them, and QCD was developed a few years later with the discovery of asymptotic freedom. Thus, with the wisdom of hindsight, it is clear that Chew *et al.* were wrong to reject quantum field theory. Nonetheless, their insights were very influential, perhaps even

crucial, for the discovery of string theory, which can be regarded as the ultimate realization of the S-Matrix program.

Some of the ingredients that went into the S-matrix theory program, such as unitarity and maximal analyticity of the S matrix, were properties (deduced from quantum field theory) that encode the requirements of causality and nonnegative probabilities. Another important ingredient was analyticity in angular momentum. The idea is that partial wave amplitudes $a_l(s)$, which are defined in the first instance for angular momenta $l = 0, 1, \ldots$, can be extended to an analytic function of l, $a(l, s)$. The uniqueness of this extension results from imposing suitable asymptotic behavior in l. The Mandelstam invariant s is the square of the center-of-mass energy of the scattering reaction. The analytic function $a(l, s)$ can have isolated poles called *Regge poles*. (Branch points are also possible, but they are usually ignored.) The position of a Regge pole is given by a *Regge trajectory* $l = \alpha(s)$. A value of s for which $l = \alpha(s)$ takes a physical value corresponds to a physical hadron of spin l.

Theoretical work in this period was strongly influenced by experimental results. Many new hadrons were discovered in experiments at the Bevatron in Berkeley, the AGS in Brookhaven, and the PS at CERN. Plotting masses squared versus angular momentum (for fixed values of other quantum numbers), it was noticed that the Regge trajectories are approximately linear with a common slope

$$\alpha(s) = \alpha(0) + \alpha' s \qquad \alpha' \sim 1.0 \, (\text{GeV})^{-2}\,.$$

Using the crossing-symmetry properties of analytically continued scattering amplitudes, one argued that exchange of Regge poles (in the t channel) controlled the high-energy, fixed momentum transfer, asymptotic behavior of physical amplitudes:

$$A(s, t) \sim \beta(t)(s/s_0)^{\alpha(t)} \qquad s \to \infty, \, t < 0.$$

In this way one deduced from data that the intercept of the ρ trajectory, for example, was $\alpha_\rho(0) \sim .5$. This is consistent with the measured mass $m_\rho = .76 \, \text{GeV}$ and the Regge slope $\alpha' \sim 1.0 \, (\text{GeV})^{-2}$.

The approximation of linear Regge trajectories describes long-lived resonances, whose widths are negligible compared to their masses. This approximation is called the *narrow resonance approximation*. In this approximation branch cuts in scattering amplitudes, whose branch points correspond to multiparticle thresholds, are approximated by a sequence of resonance poles. This is what one would expect in the tree approximation to a quantum field theory in which all the resonances appear as fundamental fields. However, there was also another discovery, called *duality*, which clashed with the usual notions of quantum field theory. In this context duality means that a scattering amplitude can be expanded in an infinite series of s-channel poles, and this gives the same result as its expansion in an infinite series of t-channel poles. To include both sets of poles, as usual Feynman diagram techniques might suggest, would amount to double counting.

2   The Discovery of String Theory

Veneziano (1968) discovered a simple analytic formula that exhibits duality with linear Regge trajectories. It is given by a sum of ratios of Euler gamma functions:

$$T = A(s,t) + A(s,u) + A(t,u), \quad \text{where} \quad A(s,t) = g^2 \frac{\Gamma(-\alpha(s))\Gamma(-\alpha(t))}{\Gamma(-\alpha(s) - \alpha(t))},$$

g is a coupling constant, and α is a linear Regge trajectory

$$\alpha(s) = \alpha(0) + \alpha' s.$$

The Veneziano formula gives an explicit realization of duality and Regge behavior in the narrow resonance approximation. The function $A(s,t)$ can be expanded as an infinite series of s-channel poles or of t-channel poles. The motivation for writing down this formula was largely phenomenological, but it turned out that formulas of this type describe scattering amplitudes in the tree approximation to a consistent quantum theory!

A generalization to incorporate adjoint $SU(N)$ quantum numbers was formulated by Paton and Chan (1969). Chan–Paton symmetry was initially envisaged to be a global (flavor) symmetry, but it was shown later to be a local gauge symmetry.

Very soon after the appearance of the Veneziano amplitude, Virasoro (1969) proposed an alternative formula

$$T = g^2 \frac{\Gamma(-\tfrac{1}{2}\alpha(s))\Gamma(-\tfrac{1}{2}\alpha(t))\Gamma(-\tfrac{1}{2}\alpha(u))}{\Gamma(-\tfrac{1}{2}\alpha(t) - \tfrac{1}{2}\alpha(u))\Gamma(-\tfrac{1}{2}\alpha(s) - \tfrac{1}{2}\alpha(u))\Gamma(-\tfrac{1}{2}\alpha(s) - \tfrac{1}{2}\alpha(t))},$$

which has similar virtues. Since this formula has total stu symmetry, it describes particles that are singlets of the Chan–Paton symmetry group.

Over the course of the next year or so, *dual models*, as the subject was then called, underwent a sudden surge of popularity, marked by several remarkable discoveries. One was the discovery (by several different groups) of an N-particle generalization of the Veneziano formula

$$A_N(k_1, k_2, \ldots, k_N) = g_{\text{open}}^{N-2} \int d\mu_N(y) \prod_{i<j} (y_i - y_j)^{\alpha' k_i \cdot k_j},$$

where $y_1, y_2, \ldots, y_N$ are real coordinates. I will omit the description of the measure $d\mu_N(y)$, which can be found in Schwarz (2007). This formula has cyclic symmetry in the N external lines. Soon thereafter Shapiro (1970) formulated an N-particle generalization of the Virasoro formula:

$$A_N(k_1, k_2, \ldots, k_N) = g_{\text{closed}}^{N-2} \int d\mu_N(z) \prod_{i<j} |z_i - z_j|^{\alpha' k_i \cdot k_j},$$

where $z_1, z_2, \ldots, z_N$ are complex coordinates. This amplitude has total symmetry in the N external lines.

Both of these formulas for multiparticle amplitudes were shown to have poles whose residues factorize in a consistent manner on an infinite spectrum of single-particle states.

This spectrum is described by a Fock space associated to an infinite number of harmonic oscillators

$$\{a_m^\mu\} \qquad \mu = 0, 1, \ldots, d-1 \qquad m = 1, 2, \ldots$$

where d is the dimension of Minkowski spacetime, which was initially assumed to be four. There is one set of such oscillators in the Veneziano case and two sets in the Shapiro–Virasoro case. These spectra were interpreted as describing the normal modes of a relativistic string: an open string (with ends) in the first case and a closed string (loop) in the second case. Amazingly, the formulas were discovered before this interpretation was proposed. In the above formulas, the y coordinates parametrize points on the boundary of a string world sheet, where particles that are open-string states are emitted or absorbed, whereas the z coordinates parametrize points on the interior of a string world sheet, where particles that are closed-string states are emitted or absorbed. (It is also possible to construct amplitudes in which both types of particles participate.)

Having found the factorization, it became possible to compute radiative corrections (loop amplitudes). Gross, Neveu, Scherk, and Schwarz (1970) discovered unanticipated singularities in a particular one-loop diagram for which the world sheet is a cylinder with two external particles attached to each of the two boundaries. The computations showed that this diagram gives branch points that violate unitarity. This was a very disturbing conclusion, since it seemed to imply that the classical theory does not have a consistent quantum extension. However, soon thereafter it was pointed out by Lovelace (1971) that these branch points become poles provided that

$$\alpha(0) = 1 \quad \text{and} \quad d = 26.$$

Prior to this discovery, everyone assumed that the spacetime dimension should be $d = 4$. We had no physical reason to consider extra dimensions. It was the mathematics that forced us in that direction. Later, these poles were interpreted as closed-string states in a one-loop open-string amplitude. Nowadays this is referred to as *open-string/closed-string duality*. This is closely related to *gauge/gravity duality*, which was discovered 27 years later.

The analysis also required there to be an infinite number of decoupling conditions, which turned out to coincide with the constraints proposed by Virasoro (1970) and further elaborated upon by Fubini and Veneziano (1971). Since the string has an infinite spectrum of higher-spin states, there are corresponding gauge invariances that eliminate unphysical degrees of freedom. The operators that describe the constraints that arise for a particular covariant gauge choice satisfy the Virasoro algebra

$$[L_m, L_n] = (m-n)L_{m+n} + \frac{c}{12}(m^3 - m)\delta_{m,-n},$$

where m, n are arbitrary integers. These operators can also be interpreted as generators of conformal symmetry for the two-dimensional string world sheet. The central charge (or conformal anomaly) c is equal to the spacetime dimension d. This anomaly cancels for $d = 26$ when the contribution of Faddeev–Popov ghosts is included.

3 The RNS Model and the Discovery of Supersymmetry

In a very inspired and important development, Ramond (1971) constructed a stringy analog of the Dirac equation, which describes a fermionic string. Just as the string momentum p^μ is the zero mode of a density $P^\mu(\sigma)$, where the coordinate σ parametrizes the string, he proposed that the Dirac matrices γ^μ should be the zero modes of densities $\Gamma^\mu(\sigma)$. Then he considered the Fourier modes of the dot product:

$$F_n = \int_0^{2\pi} e^{-in\sigma}\Gamma(\sigma)\cdot P(\sigma)d\sigma \qquad n \in \mathbb{Z}.$$

In particular,

$$F_0 = \gamma\cdot p + \text{additional terms}.$$

He proposed that physical states of a fermionic string should satisfy the following analog of the Dirac equation

$$(F_0 + M)|\psi\rangle = 0.$$

He also observed that in the case of the fermionic string the Virasoro algebra generalizes to a super-Virasoro algebra

$$\{F_m, F_n\} = 2L_{m+n} + \frac{c}{3}m^2\delta_{m,-n}$$

$$[L_m, F_n] = (\frac{m}{2} - n)F_{m+n}$$

$$[L_m, L_n] = (m - n)L_{m+n} + \frac{c}{12}m^3\delta_{m,-n}.$$

Ramond's paper does not include the central extension, which turns out to be $c = 3d/2$, where d is the spacetime dimension. A little later, it was realized that consistency requires $d = 10$ and $M = 0$. These conditions are the analogs of $d = 26$ and $\alpha(0) = 1$ for the bosonic Veneziano string theory.

A couple of months later Neveu and Schwarz (1971a) constructed a new interacting bosonic string theory, which was called the *dual pion model*. It has a similar structure to the fermionic string, but the periodic density $\Gamma^\mu(\sigma)$ is replaced by an antiperiodic one $H^\mu(\sigma + 2\pi) = -H^\mu(\sigma)$. Then the Fourier modes, which differ from an integer by 1/2,

$$G_r = \int_0^{2\pi} e^{-ir\sigma}H\cdot Pd\sigma \qquad r \in \mathbb{Z} + 1/2$$

satisfy a similar super-Virasoro algebra. Neveu and Schwarz (1971a) refers to this algebra as a *supergauge algebra*, a terminology that was sensible in the context at hand. The Neveu–Schwarz bosons and Ramond fermions were combined in a unified interacting theory of bosons and fermions by Neveu and Schwarz (1971b) and by Thorn (1971). This theory (the RNS model) was an early version of superstring theory. As will be explained shortly, a few crucial issues were not yet understood.

After a few more months, Gervais and Sakita (1971) showed that the the RNS model is described by the string world-sheet action

$$S = T \int d\sigma d\tau \left(\partial_\alpha X^\mu \partial^\alpha X_\mu - i\bar{\psi}^\mu \gamma^\alpha \partial_\alpha \psi_\mu \right),$$

where the coefficient T is the string tension. They also explained that it has *two-dimensional supersymmetry*, though that terminology was not used yet, by showing that it is invariant under the transformations

$$\delta X^\mu = \bar{\varepsilon}\psi^\mu, \qquad \delta\psi^\mu = -i\gamma^\alpha \varepsilon \partial_\alpha X^\mu,$$

where ε is an infinitesimal constant spinor. To the best of my knowledge, this is the first supersymmetric theory identified in the literature! There are two possibilities for the world-sheet fermi fields ψ^μ. When it is antiperiodic $\psi^\mu = H^\mu$, which gives the boson spectrum (Neveu–Schwarz sector), and when it is periodic $\psi^\mu = \Gamma^\mu$, which gives the fermion spectrum (Ramond sector).

Five years later, Brink, Di Vecchia, and Howe (1976) and Deser and Zumino (1976) constructed a more fundamental world-sheet action with local supersymmetry. This formulation of the world-sheet theory has the additional virtue of also accounting for the super-Virasoro constraints. From this point of view, the significance of the super-Virasoro algebra is that the world-sheet theory, when properly gauge fixed and quantized, has *superconformal symmetry*. Again, the anomaly cancels for $d = 10$ when the Faddeev–Popov ghosts are included.

At about the same time as Ramond's paper, the four-dimensional super-Poincaré algebra was introduced in a paper by Golfand and Likhtman (1971), who proposed constructing 4d field theories with this symmetry. This paper went unnoticed in the West for several more years. In fact, the celebrated paper of Wess and Zumino (1974), which formulated a class of 4d supersymmetric theories, was motivated by the search for 4d analogs of the 2d Gervais–Sakita world-sheet action. The Wess-Zumino paper launched the study of supersymmetric field theories, which proceeded in parallel with the development of supersymmetric string theory. Wess and Zumino (1974) used the expression *supergauge*, following the terminology of Neveu and Schwarz (1971), but in their subsequent papers they switched to *supersymmetry*, which was more appropriate for what they were doing.

4   The Temporary Demise of String Theory

String theory is formulated as an on-shell S-matrix theory in keeping with its origins discussed earlier. However, the SLAC deep inelastic scattering experiments in the late 1960s made it clear that the hadronic component of the electromagnetic current is a physical off-shell quantity, and that its asymptotic properties imply that hadrons have hard point-like constituents. Moreover, all indications (at that time) were that strings are too soft to describe hadrons with their pointlike constituents.

By 1973–74 there were many good reasons to stop working on string theory: a successful and convincing theory of hadrons (QCD) was discovered, and string theory had severe problems as a theory of hadrons. These included an unrealistic spacetime dimension ($d = 10$ or $d = 26$), an unrealistic spectrum (including a tachyon and massless particles), and the absence of pointlike constituents. A few years of attempts to do better had been unsuccessful. Moreover, convincing theoretical and experimental evidence for the Standard Model was rapidly falling into place. That was where the action was. Even for those seeking to pursue speculative theoretical ideas there were options other than string theory that most people found more appealing, such as grand unification and supersymmetric field theory. Understandably, string theory fell out of favor. What had been a booming enterprise involving several hundred theorists rapidly came to a grinding halt. Only a few diehards continued to pursue it.

5   Gravity and Unification

Among the problems of the known string theories, as a theory of hadrons, was the fact that the spectrum of open strings contains massless spin 1 particles, and the spectrum of closed strings contains a massless spin 2 particle (as well as other massless particles), but there are no massless hadrons. In 1974, Joël Scherk and I decided to take string theory seriously as it stood, rather than forcing it to conform to our preconceptions. This meant abandoning the original program of describing hadron physics and interpreting the massless spin 2 state in the closed-string spectrum as a graviton. Also, the massless spin 1 states in the open-string spectrum could be interpreted as particles associated to Yang–Mills gauge fields. Specifically, Scherk and Schwarz (1974) proposed trying to interpret string theory as a unified quantum theory of all forces including gravity. Neveu and Scherk (1972) had shown that string theory incorporates the correct gauge invariances to ensure agreement at low energies (compared to the scale given by the string tension) with Yang–Mills theory. Yoneya (1973,1974) and Scherk and Schwarz (1974) showed that it also contains gauge invariances that ensure agreement at low energies with general relativity.

To account for Newton's constant, the most natural choice for the fundamental string length scale was $l_s \sim 10^{-33}$ cm (the Planck length) instead of $l_s \sim 10^{-13}$ cm (the typical size of a hadron). Thus the strings suddenly shrank by 20 orders of magnitude, but the mathematics was essentially unchanged. The string tension is proportional to l_s^{-2}, so it increased by 40 orders of magnitude.

The proposed new interpretation had several advantages:
- Gravity and Yang–Mills forces are required by string theory.
- String theory has no UV divergences.
- Extra spatial dimensions could be a good thing.

Let me say a few words about the last point. In a nongravitational theory, the spacetime geometry is a rigid background on which the dynamics takes place. In that setup, the fact

that we observe four-dimensional Minkowski spacetime is a compelling argument to formulate the theory in that background geometry. As you know very well, this is part of the story of the Standard Model. However, in a gravitational theory that abides by the general principles laid out by Einstein, the spacetime geometry is determined by the dynamical equations. In such a setup extra dimensions can make sense provided that the equations of the theory have a solution for which the geometry is the product of four-dimensional Minkowski spacetime and a compact manifold that is sufficiently small to have eluded detection. It turns out that there are many such solutions. Moreover, the details of the compact manifold play a crucial role in determining the symmetries and particle content of the effective low-energy theory in four dimensions, even when the compact dimensions are much too small to observe directly.

6   Supersymmetry, Supergravity, and Superstrings

In the second half of the 1970s the study of supersymmetric field theories become a major endeavor. A few important supersymmetric theories that were formulated in that era included
- $\mathcal{N} = 1$, $d = 4$ supergravity, discovered by Freedman, Van Nieuwenhuizen, and Ferrara (1976) and Deser and Zumino (1976).
- $\mathcal{N} = 1$, $d = 10$ and $\mathcal{N} = 4$, $d = 4$ supersymmetric Yang–Mills theory discovered by Brink, Scherk, and Schwarz (1977) and Gliozzi, Scherk, and Olive (1977).
- $\mathcal{N} = 1$, $d = 11$ supergravity discovered by Cremmer, Julia and Scherk (1978).

Gliozzi, Scherk, and Olive (1976, 1977) proposed a truncation of the RNS string theory spectrum – *the GSO Projection* – that removes half of the fermion states and the "odd G-parity" bosons. In particular, the latter projection eliminates the tachyon. They showed that after the projection the number of physical bosonic degrees of freedom is equal to the number of physical fermionic degrees of freedom at every mass level. This was compelling evidence for *ten-dimensional spacetime supersymmetry* of the GSO-projected theory. Prior to this, we knew about the supersymmetry of the two-dimensional string world-sheet theory, but we had not considered the possibility of spacetime supersymmetry. In fact, the GSO projection is not just an option; it is required for consistency.

In 1979 Michael Green and I began a collaboration, which had the initial goal of understanding and proving the ten-dimensional spacetime supersymmetry of the GSO-projected version of the RNS theory. The highlights of our work included Green and Schwarz (1981, 1984a), which developed a new formalism in which the spacetime supersymmetry of the GSO-projected RNS string is manifest, and Green and Schwarz (1982), which classified the consistent ten-dimensional superstring theories and giving them the names Type I, Type IIA, and Type IIB. We were excited about these (and other) developments, but they did not arouse much interest in the theory community. String theory was still in the doldrums.

In the early 1980s there was growing interest in supersymmetry and extra dimensions. In particular, a small community became intrigued by Kaluza–Klein reduction of 11-dimensional

supergravity. Only the string ingredient was missing from their considerations. That changed following our next discovery.

7   Anomalies

If a unified theory is to make contact with the Standard Model, and have a chance of being realistic, parity violation is an essential ingredient. However, parity-violating classical theories generically have *gauge anomalies*, which means that they cannot be used to define quantum theories. The gauge symmetry is broken by one-loop quantum corrections, rendering the would-be quantum theory inconsistent. In the case of the Standard Model, if one were to change the theory by removing all of the leptons or all of the quarks, the theory would become inconsistent. When both the quarks and the leptons are included all gauge anomalies beautifully cancel, and so the Standard Model is a well-defined quantum theory. These considerations raise the question whether the potential gauge anomalies in chiral superstring theories also cancel, so that they give consistent quantum theories.

We knew that Type I superstring theory is a well-defined ten-dimensional theory at tree level for any $SO(n)$ or $Sp(n)$ gauge group, and that for every such group it is chiral (*i.e.,* parity violating). However, evaluation of a one-loop hexagon diagram in ten-dimensional super Yang–Mills theory, which describes the massless open-string states, exhibits explicit nonconservation of gauge currents, signalling a gauge anomaly. The only hope for consistency is that inclusion of the closed-string (gravitational) sector cancels this gauge anomaly without introducing new ones.

Type IIB superstring theory, which only has a closed-string gravitational sector, is also chiral and therefore potentially anomalous. It was not known how to analyze such anomalies until Alvarez-Gaumé and Witten (1984) derived general formulas for gauge, gravitational, and mixed anomalies in an arbitrary spacetime dimension. Using their results, they discovered that the gravitational anomalies, which would imply nonconservation of the stress tensor, cancel in Type IIB superstring theory. In their calculation this cancellation appears quite miraculous, though the UV finiteness of the Type IIB loop amplitudes implies that it had to work. Thus, Type IIB is a consistent chiral superstring theory. On the other hand, it did not look promising for describing the real world, since it does not contain any Yang–Mills gauge fields. (Many years later, nonperturbative Type IIB solutions that do contain Yang–Mills fields were discovered.) At that time, the last hope for constructing a realistic model seemed to reside with the Type I superstring theories, which are chiral and do contain Yang–Mills fields.

After a couple years of failed attempts, Green and I finally managed to compute the one-loop hexagon diagrams in Type I superstring theory. We found that both the cylinder and the Möbius-strip world-sheet diagrams contribute to the gauge anomaly and realized that there might be a gauge group for which the two contributions cancel. Green and Schwarz (1985) showed that $SO(32)$ is the unique choice for which the cancellation occurs. Since this

computation only demonstrated the cancellation of the pure gauge part of the anomaly, we decided to explore the low-energy effective field theory to see whether the gravitational and mixed anomalies also cancel. Using the results of Alvarez-Gaumé and Witten (1984), Green and Schwarz (1984b) verified that all gauge, gravitational, and mixed anomalies do in fact cancel for the gauge group $SO(32)$.

The effective field theory analysis showed that $E_8 \times E_8$ is a second (and the only other) gauge group for which the anomalies could cancel for a theory with $\mathcal{N} = 1$ supersymmetry in ten dimensions. In both cases, it is crucial for the result that the coupling to supergravity is included. The $SO(32)$ case could be accommodated by Type I superstring theory, but we didn't know a superstring theory with gauge group $E_8 \times E_8$. We were aware of the article by Goddard and Olive (1983) that pointed out (among other things) that there are exactly two even self-dual Euclidean lattices in 16 dimensions, and these are associated with precisely these two gauge groups. However, we did not figure out how to exploit this fact before the problem was solved by Gross, Harvey, Martinec, and Rohm (1985).

8   Epilogue

Following these discoveries there was a sudden surge of interest in superstring theory. After more than a decade, string theory had emerged from the doldrums. In my view, some of the new converts made a phase transition from being too pessimistic about string theory to being too optimistic about the near-term prospects for finding a realistic model. However, after a few years, almost all practitioners had a much more sober assessment of the challenges that remain. Superstring theory (including M-theory, which is part of the same theoretical framework) has remained a very active subject ever since 1984. Even though the construction of a complete and realistic model of elementary particles still appears to be a distant dream, the study of string theory has been enormously productive. For example, insights derived from these studies have had a profound impact on fundamental mathematics and are beginning to inspire new approaches to understanding topics in other areas of physics.

For many years string theory was considered to be a radical alternative to quantum field theory. However, in recent times – long after the period covered by this lecture – dualities relating string theory and quantum field theory were discovered. In view of these dualities, my current opinion is that string theory is best regarded as the logical completion of quantum field theory, and therefore it is not radical at all. There is still much that remains to be understood, but I am convinced that we are on the right track and making very good progress.

References

Alvarez-Gaumé, L. & Witten, E. (1984). Gravitational anomalies. *Nuclear Physics B*, *234*, 269-330.

Brink, L., Di Vecchia, P., & Howe, P. (1976). A locally supersymmetric and reparametrization invariant action for the spinning string. *Physics Letters B*, *65*, 471-474.

Brink, L., Schwarz, J. H., & Scherk, J. (1977). Supersymmetric Yang-Mills theories. *Nuclear Physics B*, *121*, 77-92.

Cremmer, E., Julia, B., & Scherk, J. (1978). Supergravity theory in eleven-dimensions. *Physics Letters B*, *76*, 409-412.

Deser, S., & Zumino, B. (1976a). Consistent supergravity. *Physics Letters B*, *62*, 335-337.

Deser, S., & Zumino, B. (1976b). A complete action for the spinning string. *Physics Letters B*, *65*, 369-373.

Freedman, D. Z., Van Nieuwenhuizen, P., & Ferrara, S. (1976). Progress toward a theory of supergravity. *Physical Review D*, *13*, 3214-3218.

Fubini, S., & Veneziano, G. (1971). Algebraic treatment of subsidiary conditions in dual resonance models. *Annals of Physics*, *63*, 12-27.

Gervais, J. L., & Sakita, B. (1971). Field theory interpretation of supergauges in dual models. *Nuclear Physics B*, *34*, 632-639.

Gliozzi, F., Scherk, J., & Olive, D. (1976). Supergravity and the spinor dual model. *Physics Letters B*, *65*, 282-286.

Gliozzi, F., Scherk, J., & Olive, D. (1977). Supersymmetry, supergravity theories and the dual spinor model. *Nuclear Physics B*, *122*, 253-290.

Goddard, P., & Olive, D. (1983). Algebras, lattices and strings. DAMTP-83/22. Reprinted in Goddard, P. (Ed.), & Olive, D. (Ed.): *Kac-Moody and Virasoro Algebras* (pp. 210-255). World Scientific 1988.

Golfand, Yu. A., & Likhtman, E. P. (1971). Extension of the algebra of Poincaré group

generators and violation of P invariance. *JETP Lett.*, *13*, 323-326.

Green, M. B., & Schwarz, J. H. (1981). Supersymmetrical dual string theory. *Nuclear Physics B*, *181*, 502-530.

Green, M. B., & Schwarz, J. H. (1982). Supersymmetrical string theories. *Physics Letters B*, *109*, 444-448.

Green, M. B., & Schwarz, J. H. (1984a). Covariant description of superstrings. *Physics Letters B*, *136*, 367-370.

Green, M. B., & Schwarz, J. H. (1984b). Anomaly cancellation in supersymmetric $D = 10$ gauge theory and superstring theory. *Physics Letters B*, *149*, 117-122.

Green, M. B., & Schwarz, J. H. (1985). The hexagon gauge anomaly in Type I superstring theory. *Nuclear Physics B*, *255*, 93-114.

Gross, D. J., Neveu, A., Scherk, J., & Schwarz, J. H. (1970). Renormalization and unitarity in the dual resonance model. *Physical Review D*, *2*, 697-710.

Gross, D. J., Harvey, J. A., Martinec, E. J., & Rohm, R. (1985). The heterotic string. *Physical Review Letters*, *54*, 502-505.

Lovelace, C. (1971). Pomeron form factors and dual Regge cuts. *Physics Letters B*, *34*, 500-506.

Neveu, A., & Schwarz, J. H. (1971a). Factorizable dual model of pions. *Nuclear Physics B*, *31*, 86-112.

Neveu, A., & Schwarz, J. H. (1971b). Quark model of dual pions *Physical Review D*, *4*, 1109-1111.

Neveu, A., Schwarz, J. H., & Thorn, C. B. (1971). Reformulation of the dual pion model. *Physics Letters B*, *35*, 529-533.

Neveu, A., & Scherk, J. (1972). Connection between Yang–Mills fields and dual models. *Nuclear Physics B*, *36*, 155-161.

Paton, J. E., & Chan, H. (1969). Generalized Veneziano model with isospin. *Nuclear Physics B*, *10*, 516-520.

Ramond, P. (1971). Dual theory for free fermions. *Physical Review D*, *3*, 2415-2418.

Scherk, J., & Schwarz, J. H. (1974). Dual models for non-hadrons. *Nuclear Physics B*, *81*, 118-144.

Schwarz, J. H. (2007). The early years of string theory: a personal perspective. arXiv:0708.1917 [hep-th].

Shapiro, J. A. (1970). Electrostatic analog for the Virasoro model. *Physics Letters B*, *33*, 361-362.

Thorn, C. B. (1971). Embryonic dual model for pions and fermions. *Physical Review D*, *4*, 1112-1116.

Veneziano, G. (1968). Construction of a crossing-symmetric Regge-behaved amplitude for linearly rising Regge trajectories. *Nuovo Cimento A*, *57*, 190-197.

Virasoro, M. (1969). Alternative constructions of crossing-symmetric amplitudes with Regge behavior. *Physical Review*, *177*, 2309-2311.

Virasoro, M. (1970). Subsidiary conditions and ghosts in dual resonance models. *Physical Review D*, *1*, 2933-2936.

Wess, J., & Zumino, B. (1974). Supergauge transformations in four dimensions. *Nuclear Physics B*, *70*, 39-50.

Yoneya, T. (1973). Quantum gravity and the zero slope limit of the generalized Virasoro model. *Nuovo Cimento Letters*, *8*, 951-955.

Yoneya, T. (1974). Connection of dual models to electrodynamics and gravidynamics. *Progress in Theoretical Physics*, *51*, 1907-1920.

Reminiscence on the Birth of String Theory[1]

Joel A. Shapiro

Department of Physics and Astronomy
Rutgers University, Piscataway, NJ, 08854-8019

Abstract

These are my personal impressions of the environment in which string theory was born, and what the important developments affecting my work were during the hadronic string era, 1968-1974. I discuss my motivations and concerns at the time, particularly in my work on loop amplitudes and on closed strings.

1   Introduction

It is not unusual in theoretical physics for conceptual frameworks to ride roller-coasters, but few have had as extreme highs and lows as in the history of string theory from its beginnings in 1968 to the present. In fact, string theory was so dead in the mid to late '70's that it is a common assumption of many articles in the popular press, and of many younger string theorists, that the field originated in the '80's, completely ignoring the period we are celebrating here, which is primarily 1968-74.

So it was pleasantly surprising to be invited to reminisce about the early days of string theory. Research results from that era have been extensively presented and reviewed, so I will try to give my impression of the atmosphere at the time, and what questions we were trying to settle, rather than review the actual results.

2   The Placenta

In the mid '60's, the framework for understanding fundamental physics was very different from what it is now. We still talk about the four fundamental interactions, but we know that the weak and electromagnetic interactions are part of a unified gauge field theory, that strong interactions are also described by a gauge field theory which might quite possibly unify with the others at higher energy, and that even general relativity is a form of gauge field theory.

---

[1]Invited Contribution to "The Birth of String Theory" Commemorative Volume

In the 1960's things were very different. Not only were the four interactions considered to be of completely different natures, but for the most part the physicists who worked on them were divided into groups by the interactions on which they worked. Of course, every budding particle theorist learned QFT and how wonderfully successful it was in treating QED. But one also learned how these perturbative methods could not be used for strong interactions because the coupling constant was too large, and that for the weak interactions one could only work at the Born approximation, because all existing field theories for the weak interactions were non-renormalizable. So particle theorists were divided into separate groups: one working on strong interactions, one on weak interaction phenomenology, and one doing high order, esoteric QED calculations. Each group had very different techniques and styles.

Even more removed from the world of a strong interaction physicist was the fourth interaction, gravity, which was studied, if at all, by general relativists[2]. When, as a very naïve graduate student who knew nothing of the fields of physics research (I had just received my ScB in Applied Mathematics), I was asked by my future advisor what I might be interested in, I replied "unified field theory". Nonetheless, it was never suggested that I take a course in general relativity!

So the context into which string theory was born was not so much theoretical fundamental physics or even particle theory, but rather strong interaction theory/phenomenology. The principal recent successes in that field had been in searching for patterns and fitting simple models[3] to scattering data. Scattering cross-sections were dominated by resonance peaks and the high energy asymptotic behavior described by Regge trajectories. A huge number of particles and resonances had been found and were listed in the particle data tables [38]. The organization of these particles into (flavor) SU(3) multiplets was the most impressive thing understood about the strong interactions.

The quasi-stable hadrons and the resonances fell beautifully into patterns which could be understood by treating baryons *as if* they consisted of three quarks and mesons of a quark and an antiquark. Even though this very successfully described the dominant experimental observations, theorists were very reluctant to think of the quarks as real constituents of hadrons.

Fits to the data were done by treating the scattering amplitude as a sum of resonance production and decay, together with an additional contribution due to the exchange of the same

---

[2]How separated general relativity and particle physics were in the '60's is discussed by David Kaiser [20]. He argues that funding cuts in particle theory in the late '60's and '70's played a large role in the subsequent bringing together of particle theorists and general relativists.

[3] I am using "model" and "theory" with a distinction that is perhaps not generally accepted. To me, a theory is a comprehensive approach to explaining part of physics in a way which will at least have features which are fundamentally correct, while a model tries, with less ambition, to fit aspects of the data, but cannot be taken as the fundamental truth, even as an approximation of the truth. Thus QED, QCD, and general relativity are theories, even though the last clearly needs modification to include quantum mechanics, while the interference model, DHS duality[12], and my thesis are models. The Dual Resonance Model might be taken to have evolved into a theory when we started calculating unitary corrections in the form of loop graphs.

particles in the form of Regge poles to describe the high energy behavior. This sum was called the interference model. But the experimentalists kept finding more and more resonances, and they were joined by phase-shift analysts. It soon appeared the sequence of resonances continued indefinitely to higher masses and spins, in what clearly looked like linearly rising Regge trajectories. In fact, my Ph. D. thesis[40] was a very naïve non-relativistic model using PCAC, which rather successfully explained the experimental[38] π-nucleon decay widths of a tower of five Δ resonances with spins ranging from 3/2 to 19/2. Unfortunately the top two of these resonances have subsequently dissolved[54]. This infinite sequence of resonances suggested the idea of duality [12], that the amplitude could be described *either* in terms of a sum of resonances *or* in terms of a series of Regge poles. The possibility that a scattering amplitude $A(s,t)$ could be given as an sum of resonant poles in s or alternatively as a sum of Regge poles in t caused great excitement, but also skepticism that such a function could exist.

3 Conception and the Embryonic Period

Thus it seemed miraculous when Veneziano [51] discovered that Euler had given us just such a function in 1772, to describe the $\pi\pi \to \pi\omega$ scattering amplitude. This paper arrived at the Lawrence Radiation Lab in Berkeley in the summer of 1968 while I was away on a short vacation, and I returned to find the place in a whirlwind of interest. Everyone had stopped what they were doing, and were asking if this idea could be extended to a more accessible interaction, such as $\pi\pi \to \pi\pi$. I suggested the very minor modification necessary to remove the tachyon,

$$\frac{\Gamma(-\alpha(s))\Gamma(-\alpha(t))}{\Gamma(-\alpha(s)-\alpha(t))} \longrightarrow \frac{\Gamma(1-\alpha(s))\Gamma(1-\alpha(t))}{\Gamma(1-\alpha(s)-\alpha(t))},$$

and Joel Yellin and I investigated whether this could be taken as a realistic description [48] for $\pi\pi$ scattering. It had a lot of good qualitative features, including resonance dominance, regge behavior, and full duality. We were forced to have exchange degeneracy between the $I = 0$ and $I = 1$ trajectories, which was well fit by the data. We noticed the problem that such an amplitude can wind up with ghosts, with a negative decay width for the ϵ', the 0^+ partner of the f, but also that this problem disappeared if the ρ trajectory intercept exceeded 0.496, very close to the value of 0.48 which we got from fitting the low energy phase shifts. A much more serious problem was that we predicted a ρ' degenerate with the f, which seemed to be ruled out by experimental data. That the simplest function did not produce a totally acceptable model was discouraging, especially to Yellin, although we realized that there was no compelling reason not to add subsidiary terms to the simple ratio of gamma functions, except that to do so removed all predictive power! This convinced Yellin that he didn't want to coauthor the fuller version[41] of our paper. But Lovelace[28], who independently discovered the same amplitude, managed to do a favorable comparison to experiment.

There were a number of papers attempting to do phenomenology with dual models, mostly describing two-body scattering processes. In general the results had, as did our

paper, nice qualitative features but unsatisfactory fitting of the data. At the same time, the formal model was becoming much more serious, as great progress was made in extending the narrow resonance approximation amplitude, first to the 5 point function [5] and then the n-particle [6, 9, 8, 15] amplitudes. A very elegant formulation of these amplitudes was given by Koba and Nielsen[26, 34, 27], in which the external particles correspond to charges given by their momentum, entering on the boundary of a unit circle, and the amplitude is given by an integral, over relative positions of the particles, of the two-dimensional electrostatic energy. Here the conformal invariance was seen to play a crucial role, and in particular the Möbius invariance explained the cyclic symmetry. From the n-point amplitude for ground state particles one could factor in multiparticle channels to extract the scattering amplitudes for all the particles which occurred in intermediate states, determining all amplitudes in what could be considered the equivalent of the tree approximation in a Lagrangian field theory. We took the attitude that the particles of the theory should be all and only those which arose from n-point scattering amplitudes of the ground state particles, as intermediate states in the n-point tree function. The amplitude for an arbitrary particle X connected to p ground states could be found by factoring the $p + q$ ground state amplitude [4], and amplitudes involving more arbitrary states could come from factoring that. Thus one could determine, in the tree approximation, the arbitrary n-particle amplitude. In a sense, this was a form of bootstrap, as the set of particles generated as intermediate states were added to form a consistent set, with the same particles as intermediate states as were considered external states.

4   Birth of String Theory

Of course a set of tree amplitudes is not a unitary theory. In perturbative field theory, the Feynman rules are guaranteed to implement unitarity by specifying loop graphs whose discontinuities give the required sum over intermediate states, because these all come from a Hermitian lagrangian. The possibility of advancing dual models to a unitary theory became possible once we had the tree amplitudes for arbitrary single particle states, as one could sew together the loop graphs to give a perturbative (in the number of loops) theory satisfying unitarity. In perturbative quantum field theory, loop graphs give the appropriate contributions to the optical theorem, satisfying the unitarity of the S-matrix. Bardakçı, Halpern and I (BHS)[3] defined the one loop graph by the requirement of two-particle unitarity. An earlier attempt (KSV)[25] defined the planar loop graph by extending duality to the internal legs, which gave most of the factors in the loop integrand. But to get the full expression, the one-loop amplitude for n ground-state particles σ should be required to have the correct two particle discontinuity, a sum over all possible two-particle intermediate states. Starting with the tree amplitude for n σ's plus $X(p)$ plus $X(-p)$, and summing over

4

all possible states X and momentum p, as shown by the stitch marks here, one is summing not only over X but also over all particles in the left arm, because those are all included in the tree graph.

Of course we called this process "sewing", which led to an amusing battle with Sy Pasternack, the editor of Physical Review, on a subsequent paper[17]. Pasternack thought he needed to uphold a certain formality, and was responsible for "pomeronchukon" rather than "pomeron". He wrote us a very witty letter[35] arguing that "sewing" would lead inevitably to "weaving", "braiding", "darning", "knitting" and "sew-on". We objected, however, that the actual thread lines were shown in the figures. Redrawing figures in those days was a major undertaking. That won the argument.
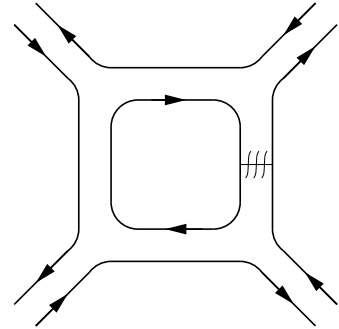
I should point out that at the time, our description of the intermediate states and the amplitudes was quite clumsy, using the rather messy techniques of the Bardakçı-Mandelstam factorization [4]. While we were working on deriving the loop graph, Nambu [31], and Fubini, Gordon and Veneziano [13] were developing the elegant operator formalism, in which the states of the system are described by harmonic oscillator excitation operators $a_n^{\mu\,\dagger}$, $n = 1...\infty$ acting on a ground state $|0\rangle$. Here n corresponds to nodes on a string and we have a Lorentz index μ. The amplitudes can then be written as a matrix element with vertex functions for each external particle, and propagators integrated $\int_0^1 du$ over an internal variable $u \sim e^{-\tau}$, where τ acts like a time describing how long an intermediate state propagates. Thus resonance poles in a tree, or two particle intermediate states in a loop, come from the $\tau \to \infty$, $u \to 0$ limit for the corresponding propagator. This formalism made the calculations much easier. It enabled the authors of KSV to discover independently from us the extra factors that get two particle unitarity correct, except for spurious states. The new formalism was so superior[2] that few people were encouraged to read our paper, and I am still grumpy about that.

The operator formalism made clearer two problems that had already been vaguely seen. In this formalism, the amplitudes appear to lose the Möbius invariance, but the amplitudes do not, due to the existence of Ward identities. That is, there are spurious states, combinations of excitations which decouple from all n ground state amplitudes, and therefore by our philosophy should not be included among the states. Secondly, the time-like creation operators create ghosts, particles with negative widths, which clearly should not be there. The set of these ghosts produced by the lowest node operator, $a_1^{0\,\dagger}$, were precisely those that could be exorcised by those Ward identities known at the time. Much of our effort in BHS involved excluding these spurious states from the loop. Of course the higher time-like modes $a_n^{0\,\dagger}$, $n > 1$ also produce ghost states, but these were also exorcised by the Ward identities found later by Virasoro[53].

One unpleasant feature of the planar loop amplitude we constructed from two-particle unitarity was the presence of additional factors $\prod_r (1-w^r)^{-D}$, where $w = \prod u_j$ is the product of the u factors of all the internal propagators, and D is the dimension of space-time, which

5

at that point we simply wrote as 4. The discontinuities we were building into the loops come from several intermediate state u_i's going to zero, so only the $w \to 0$ endpoint contributes, but the natural integration range for w was from 0 to 1. Of course at $w = 1$ this factor has extremely bad behavior. Eliminating the one set of spurious states known about at the time eliminated just one power of $(1 - w)^{-1}$, which didn't help much. Later, Virasoro [53] discovered that if the Reggeon intercept $\alpha(0) = 1$, there was an infinite set of generators of spurious states, and eliminating those gets rid of all the ghosts (for $D \leq 26$), and one full set of $\prod_r (1 - w^r)^{-1}$. Still, there is a very serious divergence as $w \to 1$, which will turn out to be connected to the Pomeron/closed string. Before that was realized, there was speculation about whether this divergence was removable, and whether the two particle discontinuity had the expected two-Reggeon cut asymptotic form[47] as $s \to \infty$.

It should be mentioned that this effort to raise dual models to the same level of legitimacy as perturbative QFT was a departure which made many uncomfortable. Strong interaction theorists had been divided into field theorists and S-matrix folk, and dual models were generally considered the domain of S-matrix types, but here they were adopting the moral values of a field theorist, even if the context was different. The phenomenologically inclined thought it would be better to simply assign imaginary parts to the regge trajectories in the dual amplitudes to go beyond a narrow resonance fit. As there had been no real data-fitting successes, many had great skepticism about the value of dual model research. One such skeptic asked me why I would work on something so unlikely to be the real physical truth. I recall saying that even though the probability that dual models would be part of the real answer was small, perhaps 10%, at least there was a chance of working towards the truth, while fitting elastic scattering data to Regge poles, to me, seemed not to have any chance of leading to fundamental physical understanding.

I mention this because in 1987, in Aspen where string theory was the superhot theory of everything, I asked some of the younger researchers what their estimate was, of the probability that string theory would be part of the the true theory of physics, and was rather astounded to hear answers upwards of 50%.

Anyway, let's get back to the construction of loop graphs for a complete, unitary dual resonance theory. This was a very active field. Neveu and Scherk [32] used their superior French mathematical education to express the planar loop in terms of elegant Jacobi θ functions, enabling them to extract the divergent behavior. The operator formalism [14, 13, 50] made tractable the calculation of nonplanar loops [24, 16, 21] and multiloop amplitudes[22]. Abelian integrals were used by Lovelace[29, 1, 30], who suggested that experimentalists deprived of funding for higher energy machines could "still construct duality diagrams in tinfoil and measure induced charges" as their contribution to understanding particle physics.

Closed Strings

My second post-doc appointment was at the University of Maryland, which had a pleasant and active high-energy theory group, but no one doing dual models. I felt quite isolated, and while I was able to write some technical papers[42, 44], I missed the stimulating environment

I had had in Berkeley. In particular, the first of these papers was a rather misguided attempt to get rid of the spurious states given by the Ward identity with L_1, before I became aware that Virasoro had found, for the "unrealistic choice of $\alpha(0) = 1$", that there was an infinite set of such Ward identities, enough to get rid of all the ghosts produced by $a_n^{0\dagger}$. Fortunately I was free to visit Berkeley and Aspen during the summers. During my Berkeley stay in 1970 I spent a day at SLAC, where in a discussion with Nussinov and Schwimmer, they asked me a very interesting question. At the time, the n-point Veneziano formula was best described by the Koba-Nielsen picture of external charges (or currents) on the circumference of a disk. The integrand could be interpreted as the electrostatic energy of the charges or as the heat generated by the currents, inside the disk. There was also much interest in this being an approximation of very complex Feynman diagrams called fishnets within the disk. The question Nussinov asked is what would happen if the external particles, instead of residing on the circumference of a disk, where on the surface of a sphere. Nussinov answered his own question with "I bet one would get the Virasoro formula". This is because, with the particles integrated over the surface, there is no cyclic order constraint, and any collection of particles are free to approach each other and produce a singularity in P^2, where P^μ is the sum of their momenta. This is what happens in the Virasoro formula[52].

Should the fishnets, or electric fields, or currents, fill the ball, or should they be confined to the surface with the external particles? In my view, the new Virasoro identities were associated with the local conformal invariance of analytic functions, a much richer group than conformal transformations in Euclidean spaces of higher dimension. Thus filling the three dimensional ball was unlikely to work, but putting fishnets on the surface might be very interesting. So the three of us began to work out electrostatics on the surface of the sphere.

In the Nielsen approach one needs the electrostatic energy of a configuration of point charges at arbitrary locations, and then integrate over the charges' positions. We can solve Poisson's equation for each configuration of charges on a 2-sphere, but we cannot define the electrostatic energy as the integral of $(\vec{\nabla}\phi)^2$, because that includes the infinite self-energy of each charge. Instead we might define $E = \frac{1}{2}\sum_{i\neq j} q_i \phi_j(\vec{r}_i)$, but to do so one needs to be able to find the electrostatic potential of a single charge. We cannot have a source of electric flux without a sink, and we seemed to hit an impasse and let the matter drop. Several weeks later, after we had all gone our separate ways, while I was (I think) in Aspen, I decided to look at this again, and I realized that putting an arbitrary sink for all the fields would do no harm. After all, for two-dimensional electrostatics one takes $\phi \propto \ln|\vec{r} - \vec{r}_0|$ without worrying about the flux which goes off to infinity, and in two dimensions conformal invariance makes infinity no different from any other point. As the sphere is conformally equivalent to the complex plane, the potentials are just logarithms of $|(z_i, z_j, a, b)|$, the absolute value of the cross ratio, including the arbitrary sink point b and a point a at which we can set the potential to zero. The Möbius invariance, now extended to three *complex* parameters, again permits the positions of three charges to be fixed arbitrarily, and the others integrated over the sphere, or complex plane. We have consistency only for $\alpha(0) = 2$, but there we have

a consistent n-point function for closed string[4] scattering[43]. I speculated that this was equivalent to the Pomerons which appeared as a problem in the loop graphs of open strings, and later, with Clavelli[10], I showed that this is indeed the case.

There was at the time not much interest in closed strings, which have no ends. All the semi-successes of dual model phenomenology were based on Harari-Rosner diagrams[18, 39] being incorporated by Chan-Paton factors[36], which required string ends on which to attach quarks. Even I postponed looking at factorization and loop graphs in this model in favor of a paper[44] showing that nonorientable graphs do not enter theories with $SU(n)$ flavors incorporated *à la* Chan-Paton. But the following summer, in Aspen again, with the factorization having been done by others[55, 11], I addressed the one-loop diagram[45]. The propagators now have an integral over the length of the tube and the angle of twist to get to the next particle, and the complex variable w has $|w| = -\ln T$, where T is the combined times of propagation, but w also has a phase given by the angle of twist in sewing the initial end of the tube to the final end. The amplitude involves

$$\int_{|w|\leq 1} d^2w \prod_{r=1}^{\infty} |1 - w^r|^{-2(D-E)},$$

where D is the dimension of space-time and E is the number of factors assumed to mysteriously disappear if one removes the spurious states. E had already been shown to be 1 in general D, but we were hoping, as was found true later, that $E = 2$ in the right D. Still, $D - E$ is a positive number, and $1 - w^r$ vanishes for w any integral power of $e^{2\pi i/r}$, so we have a terrible divergence at every point of the unit circle for which the angle is 2π times a rational number.

Fortunately, by the time I was looking at this, I had read the elegant reformulation[32] of the open-string loop in terms of Jacobi theta functions. This encourages us to look at $\tau = (\ln w)/(2\pi i)$. Of course the integrand is invariant under $\tau \to \tau + 1$, because w is unchanged, but it is also invariant under the Jacobi imaginary transformation, $\tau \to -1/\tau$, provided we have the magic dimensions $D = 26, E = 2$. The world sheet of a loop of closed strings is a donut, conformally equivalent to a parallelogram with sides 1 and $i\tau$, with opposite sides identified. While a Hula hoop[TM] and a Mayflower donut[49] may not look the same, multiplying the parallelogram by $-i/\tau$ maps one into the other. These transformations generate the modular group, so invariance shows that in integrating w over the unit disk we are including an infinite number of copies, while we wanted, for unitarity, only one copy of the region around $w \approx 0$. Thus the right thing to do is restrict our integration to the fundamental region, which is $|\tau| \geq 1$, $-\frac{1}{2} < \mathrm{Re}\,\tau \leq \frac{1}{2}$. In terms of w, this is a subset of $|w| < 0.0044$, so we stay far away from the horrible divergences.

---

[4]For some bizarre reason, I referred to the dual models in terms of what we would now call their world-sheets, as strip and tube models for what we would now call open and closed strings.

If closed strings aroused little interest, loops of them really aroused none. The figure shows the number of citations to [45] each year as listed by Spires. Interest in dual models as models of the strong interactions was fading fast. Firstly, the evidence for partons, pointlike constituents of hadrons, found in deep inelastic scattering starting in 1969, was inconsistent with the soft, extended object picture of strings. Secondly, non-abelian gauge theories were proven renormalizable ('t Hooft, 1971 [19]), explained neutral currents in a unified electroweak theory, and gave quantum chromodynamics as a theory

Citations by year to ref. [45], as listed in the Spires Citation Index.

of the strong interactions. This greatly improved the appeal of conventional field theory at the expense of string theory. And within string theory, the inclusion of fermions by Ramond[37] and Neveu-Schwarz[33] was more exciting than loops of Pomerons.
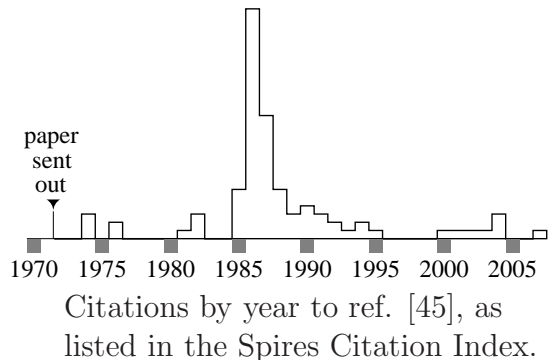
In the fall on 1971 I started an Assistant Professorship at Rutgers in a new high energy theory group headed by Lovelace, and including Clavelli as a postdoc. Lovelace had a very ambitious program for describing arbitrary multiloop diagrams, and Clavelli and I looked at how the closed string intermediate state in the nonplanar loop interacts with the ordinary (open-string) states. Then I struggled with understanding renormalization[46], an effort which would have been totally trivial if I had only realized that $\sum_n n = -1/12$, but unfortunately I had taken too many pure math classes to recognize this fact.

The decision on my tenure was coming up, and dual models did not seem the best way to prove my worth, so I reluctantly got into several other endeavors, and was rather slow to get back into strings when they arose like a phoenix, or perhaps like a fire storm, in 1984. But that is after the period we are considering here.

A Comment on Impact

I want to say a few words about how this field was perceived within the physics community. In recent years there have been numerous attacks from some in the high energy theory community, and from experimentalists, that strings are, like the Pied Piper, leading the bright young theorists astray. String theory was not quite so dominant in the 1969-1974 era, though it did absorb the attention of a very large fraction of the young theorists. It did not get a similar acceptance by most of the more established people, though I think Europe was more receptive than America. In particular Phys. Rev. Lett. published very few articles in this field, but Physics Letters had many of the important papers.

The field did attract quite a bit of attention. In fact, in the early '70's I was interviewed by a sociologist who wanted to do a study of what attracted so many people to working on dual models. Unfortunately she followed a narrow set of preprepared questions which seemed totally off the mark. Her focus was on which experimental data encouraged me to continue working in the field. I don't think anything came of that study, and I haven't heard

of studies done when the field became even more rabid in the mid '80's.

But this is still an important question: Is there any real physics in string theory, and should so many people be working on it. Undoubtedly there will be a shift towards more applied high energy theory as LHC starts giving us more data to work with. It will be very interesting to see where the field goes.

References

[1] V. Alessandrini, "A general approach to dual multiloop diagrams", Nuovo Cimento **2A** (1971) 321.

[2] D. Amati, C. Bouchiat, and J. L. Gervais, Lett. al Nuovo Cimento **2** (1969) 399.

[3] K. Bardakçı, M. B. Halpern and J. A. Shapiro, "Unitary Closed Loops in Reggeized Feynman Theory", Phys. Rev. **185** (1969) 1910-1917.

[4] K. Bardakçı and S. Mandelstam, "Analytic Solution of the Linear-Trajectory Bootstrap", Phys. Rev **184** (1969) 1640.

[5] K. Bardakçı and H. Ruegg, Phys. Lett. **28B** (1968) 342.

[6] K. Bardakçı and H. Ruegg, "Reggeized Resonance Model for Arbitrary Production Processes", Phys. Rev. **181** (1969) 1884-1889.

[7] L. Caneschi, A. Schwimmer, and G. Veneziano, "Twisted propagator in the operatorial duality formalism", Phys. Lett. **30B** (1969) 351.

[8] Chan Hong-Mo and Tsou Sheung Tsun, Phys. Lett. **28B** (1969) 485.

[9] Chan Hong-Mo, Phys. Lett. **28B** (1969) 425.

[10] L. Clavelli and Joel A. Shapiro, "Pomeron Factorization in General Dual Models", Nucl. Phys. **B57** (1973) 490-535.

[11] E. Del Giudice and P. Di Vecchia, "Factorization and Operator Formalism in the Generalized Virasoro Model" Nuovo Cimento **A5** (1971) 90-102.

[12] R. Dolen, D. Horn and C. Schmid, Phys. Rev. **166** (1968) 1768.

[13] S. Fubini, D. Gordon, and G. Veneziano, "A general treatment of factorization in dual resonance models" Phys. Lett. **29B** (1969) 679.

[14] S. Fubini and G. Veneziano, "Level Structure of Dual Resonance Models", Nuovo Cimento **64A** (1969) 811.

[15] C. J. Goebel and B. Sakita, Phys. Rev. Lett. **22**, (1969) 257.

[16] David J. Gross, A. Neveu, J. Scherk, and John H. Schwarz, "Renormalization and Unitarity in the Dual-Resonance Model" Phys. Rev. **D2** (1970) 697.

[17] M. B. Halpern, S. A. Klein and J. A. Shapiro, Phys. Rev. **188** (1969) 2378-2384.

[18] H. Harari, "Duality Diagrams", Phys. Rev. Lett. **22** (1969) 562-565.

[19] G. 't Hooft, "Renormalizable Lagrangian for Massive Yang-Mills Fields", Nucl. Phys. **B35** (1971) 167-188.

[20] David Kaiser, "When Fields Collide", Sci. Am. **296 #6**, (2007) 40-47.

[21] M. Kaku and C. B. Thorn, "Unitary Nonplanar Closed Loops" Phys. Rev. **D1** (1970) 2860.

[22] M. Kaku and L. Yu, "The general multi-loop Veneziano amplitude", Phys. Lett. **33B** (1970) 166.

[23] K. Kikkawa, "Regge Cut from a Nonplanar Duality Amplitude", Phys. Rev. **187** (1969) 2249.

[24] K. Kikkawa, S. A. Klein, B. Sakita, and M. A. Virasoro, "Feynman-Like Diagrams Compatible with Duality. II. General Discussion Including Nonplanar Diagrams" Phys. Rev. **D1** (1970) 3258.

[25] K. Kikkawa, B. Sakita, and M. A. Virasoro, "Feynman-Like Diagrams Compatible with Duality. I. Planar Diagrams" Phys. Rev. **184** (1969) 1701-1713.

[26] Z. Koba and H. B. Nielsen, Nucl. Phys. **B10** (1969) 633.

[27] Z. Koba and H. B. Nielsen, "Manifestly crossing-invariant parameterization of n-meson amplitude", Nucl. Phys. **B12** (1969) 517-536

[28] C. Lovelace, "A Novel Application of Regge Trajectories", Phys. Lett. **28B** (1968) 265.

[29] C. Lovelace, "M-Loop Generalized Veneziano Formula", Phys. Lett. **32B** (1970) 703.

[30] C. Lovelace, "Pomeron Form Factors and Dual Regge Cuts", Phys Lett **34B** (1971) 500-506.

[31] Y. Nambu, "Quark model and the factorization of the Veneziano amplitude", in R. Chand (Ed.) Symmetries and Quark Models, (Gordon and Breach 1970) 269-278.

[32] A. Neveu and J. Scherk, "Parameter-Free Regularization of One-Loop Unitary Dual Diagram", Phys. Rev. **D1**(1970) 2355.

[33] A. Neveu and J. H. Schwarz, "Factorizable dual model of pions", Nucl. Phys. **B31** (1971) 86.

[34] H. B. Nielsen, "An almost physical interpretation of the dual N point function", Nordita preprint (1969) unpublished.

[35] S. Pasternack, private communication, but here it is anyway: http://www.physics.rutgers.edu/∼shapiro/pastnk.pdf

[36] J. E. Paton and Chan H. M, "Generalized Veneziano Model with Isospin", Nucl. Phys. **B10** (1969) 516.

[37] P. Ramond, "Dual Theory of Free Fermions", Phys. Rev. **D3** (1971) 2415.

[38] A. H. Rosenfeld, Rev. Mod. Phys. **39** (1967) 1.

[39] J. L. Rosner, "Graphical Form of Duality", Phys. Rev. Lett. **22** (1969) 689.

[40] Joel A. Shapiro, "A Quarks-on-Springs Model of the Baryons", Ann. Phys. **47** (1968) 439-467. Notice I was just one letter off from a really exciting and prescient idea.

[41] Joel A. Shapiro, "Narrow Resonance Model with Regge Behavior for $\pi\pi$ Scattering", Phys. Rev. **179**, (1969) 1345-1353.

[42] Joel A. Shapiro, "General Factorization of the n-Point Beta Function into Two Parts in a Möbius Invariant Manner", Phys. Rev. **D1** (1970) 3371-3376.

[43] Joel A. Shapiro, "Electrostatic Analogue for the Virasoro Model", Phys. Lett. **33B** (1970) 361-362.

[44] Joel A. Shapiro, "Nonorientable Dual Loop Graphs and Isospin" Phys. Rev. **D4** (1971) 1249-1251.

[45] Joel A. Shapiro, "Loop Graph in the Dual Tube Model", Phys. Rev. **D5** (1972) 1945-1948.

[46] Joel A. Shapiro, "Renormalization of dual models" Phys. Rev. **D11** (1975) 2937.

[47] Joel A. Shapiro and Charles B. Thorn, "Two Particle Discontinuities of One Loop Graphs in the Dual Resonance Model", Phys. Rev. **D5**, (1972) 927-932.

[48] Joel Shapiro and Joel Yellin, "A Model for $\pi\pi$ Scattering", UCRL preprint 18500 (Sept. 1968), Yadernaya Fizika 11, 443-447 (1970); Sov. Journ. Nucl. Phys. **11**, 247-249 (1970).

[49] S. L. Steinberg, "How Doughnuts Won America", New York Times, May 6, 1981, C4.

[50] L. Susskind, "Structure of Hadrons Implied by Duality", Phys. Rev. **D1** (1970) 1182.

[51] G. Veneziano, Nuovo Cimento **57A**, 190 (1968).

[52] M. Virasoro, "Alternate Constructions of Crossing-Symmetric Amplitudes with Regge Behavior", Phys. Rev. **177** (1969) 2309.

[53] M. A. Virasoro, "Subsidiary conditions and ghosts in dual-resonance models", Phys Rev **D1** (1970) 2933-2936.

[54] W.-M. Yao et al., Journal of Physics **G 33** (2006) 1.

[55] M. Yoshimura, "Operatorial Factorization and Symmetry of the Shapiro-Virasoro Model" Phys. Lett. **34B** (1971) 79.