Perturbation Theory in Quantum Mechanics

Luigi E. Picasso, Luciano Bracci and Emilio d'Emilio

Dipartimento di Fisica, Università di Pisa
and
Istituto Nazionale di Fisica Nucleare, Sezione di Pisa, Pisa, Italy

Contents

Glossary

Hilbert space

A Hilbert space $\mathcal{H}$ is a normed complex vector space with a Hermitian scalar product. If φ, $\psi \in \mathcal{H}$ the scalar product between φ and ψ is written as $(\varphi, \psi) \equiv (\psi, \varphi)^*$ and is taken to be linear in ψ and antilinear in φ: if a, $b \in \mathbb{C}$, the scalar product between $a\,\varphi$ and $b\,\psi$ is $a^* b(\varphi, \psi)$. The norm of ψ is defined as $\|\psi\| \equiv \sqrt{(\psi, \psi)}$. With respect to the norm $\|\cdot\|$, $\mathcal{H}$ is a complete metric space. In the following $\mathcal{H}$ will be assumed to be separable, that is any complete orthonormal set of vectors is countable.

States and observables

In quantum mechanics the states of a system are represented as vectors in a Hilbert space $\mathcal{H}$, with the convention that proportional vectors represent the same state. Physicists mostly use Dirac's notation: the elements of $\mathcal{H}$ are represented by $|\cdot\rangle$ ("ket") and the scalar product between $|\varphi\rangle$ and $|\psi\rangle$ is written as $\langle\varphi \mid \psi\rangle$ ("braket"). The observables, i.e. the physical quantities that can be measured, are represented by linear Hermitian (more precisely: self-adjoint) operators on $\mathcal{H}$. The eigenvalues of an observable are the only possible results of the measurement of the observable. The observables of a system are generally the same of the corresponding classical system: energy, angular momentum, etc., i.e. they are of the form $f(q, p)$, with $q \equiv (q_1, \cdots, q_n), p \equiv (p_1, \cdots, p_n)$ the position and momentum canonical variables of the system: q_i and p_i are observables, i.e. operators, which satisfy the commutation relations $[q_i, q_j] \equiv q_i q_j - q_j q_i = 0$, $[p_i, p_j] = 0$, $[q_i, p_j] = i\hbar\,\delta_{ij}$, with $\hbar$ the Planck's constant h divided by 2π.

Representations

Since separable Hilbert spaces are isomorphic, it is always possible to represent the elements of $\mathcal{H}$ as elements of l_2, the space of the sequences $\{u_i\}$, $u_i \in \mathbb{C}$, with the scalar product $(v, u) \equiv \sum_i v_i^* u_i$. This can be done by choosing an orthonormal basis of vectors e_i in $\mathcal{H}$: $(e_i, e_j) = \delta_{ij}$ and defining $u_i = (e_i, u)$; with Dirac's notations $|A\rangle \to \{a_i\}$, $a_i = \langle e_i \mid A\rangle$. Linear

operators ξ are then represented by $\{\xi_{ij}\}$, $\xi_{ij} = (e_i, \xi\, e_j) \equiv \langle\, e_i \mid \xi \mid e_j \,\rangle$. The ξ_{ij} are called "matrix elements" of ξ in the representation e_i. If $\xi^\dagger$ is the hermitian-conjugate of ξ, then $(\xi^\dagger)_{ij} = \xi^*_{ji}$. If the e_i are eigenvectors of ξ then the (infinite) matrix ξ_{ij} is diagonal, the diagonal elements being the eigenvalues of ξ.

Schrödinger Representation

A different possibility is to represent the elements of $\mathcal{H}$ as elements of $L^2[\mathbb{R}^n]$, the space of the square-integrable functions on $\mathbb{R}^n$, where n is the number of degrees of freedom of the system. This can be done by assigning how the operators q_i and p_i act on the functions of $L^2[\mathbb{R}^n]$: in the Schrödinger representation the q_i are taken to act as multiplication by x_i and the p_i as $-i\hbar\partial/\partial x_i$: if $|\, A\,\rangle \to \psi_A(x_1, \cdots x_n)$, then

$$q_i\,|\,A\,\rangle \to x_i\psi_A(x_1, \cdots x_n),\ p_i\,|\,A\,\rangle \to -i\hbar\partial\psi_A(x_1, \cdots x_n)/\partial x_i.$$

Schrödinger Equation

Among the observables, the Hamiltonian H plays a special role. It determines the time evolution of the system through the time dependent Schrödinger equation

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi,$$

and its eigenvalues are the energy levels of the system. The eigenvalue equation $H\psi = E\psi$ is called the Schrödinger equation.

1 Definition of the Subject

In the investigation of natural phenomena a crucial role is played by the comparison between theoretical predictions and experimental data. Those practicing the two arts of the trade continuously put challenges to one another either presenting data which ask for an explanation or proposing new experimental verifications of a theory. Celestial mechanics offers the first historical instance of this interplay: the elliptical planetary orbits discovered by Kepler were explained by Newton; when discrepancies from the elliptical paths definitely emerged it was necessary to add the effects of the heavier planets to the dominant role of the sun, until persistent discrepancies between theory and experiment asked for the drastic revision of the theory of gravitation put forth by Einstein, a revision which in turn offered a lot of new effects to observe, some of which have been verified only recently.

In this dialectic interaction between theory and experiment only the simplest problem, that of a planet moving in the field of the sun within Newton's

theory, can be solved exactly. All the rest was calculated by means of perturbation theory. Generally speaking, perturbation theory is the technique of finding an approximate solution to a problem where to a dominant factor, which allows for an exact solution (zero$^{\text{th}}$ order solution), other "perturbing" factors are added which are outweighed by the dominant factor and are expected to bring small corrections to the zero$^{\text{th}}$ order solution.

Perturbation theory is ever-pervasive in physics, but an area where it plays a major role is quantum mechanics. In the early days of this discipline, the interpretation of atomic spectra was made possible only by a heavy use of perturbation theory, since the only exactly soluble problem was that of the hydrogen atom without external fields. The explanation of the Stark spectra (hydrogen in a constant electric field) and of the Zeeman spectra (atom in a magnetic field) was only possible when a perturbation theory tailored to the Schrödinger equation, which rules the atomic world, was devised. As for heavier atoms, in no case an exact solution for the Schrödinger equation is available: they could only be treated as a perturbation of simpler "hydrogenoid" atoms. Most of the essential aspects of atomic and molecular physics could be explained quantitatively in a few years by recourse to suitable forms of perturbation theory. Not only did it explain the position of the spectral lines, but also their relative intensities, and the absence of some lines which showed the impossibility of the corresponding transitions (selection rules) found a convincing explanation when symmetry considerations were introduced. When later more accurate measurements revealed details in the hydrogen spectrum (the Lamb shift) that only quantum field theory was able to explain, perturbation theory gained a new impetus which sometimes resulted in the anticipation of theory (quantum electrodynamics) over experiment as to the accuracy of the effect to be measured.

An attempt to describe all the forms that perturbation theory assumes in the various fields of physics would be vain. We will limit to illustrate its role and its methods in quantum mechanics, which is perhaps the field where it has reached its most mature development and finds its widest applications.

2 Introduction

An early example of the use of perturbation theory which clearly illustrates its main ideas is offered by the study of the free fall of a body [29]. The equation of motion is

$$\dot{\vec{v}} = \vec{g} + 2\vec{v} \times \vec{\Omega} + \vec{\Omega} \times (\vec{r} \times \vec{\Omega}) \tag{2.1}$$

where $\vec{g}$ is the constant gravity acceleration and $\vec{\Omega}$ the angular velocity of the rotation of the earth about its axis. Ω is the parameter characterizing the perturbation. If we wish to find the eastward deviation of the trajectory to first order in Ω we can neglect the third term in the RHS of (2.1), whose main effect is to cause a southward deviation (in the northern hemisphere). The ratio of the second term to the first one in the RHS of (2.1) (the effective perturbation parameter) is $\Omega\sqrt{h/g} \simeq 10^{-4}$ for the fall from a height $h \sim$ 100 m, so we can find the effect of Ω by writing $\vec{v} = \vec{v}_0 + \vec{v}_1$ in (2.1), where $\vec{v}_0$ is the zero$^{\text{th}}$ order solution ($\vec{v}_0 = \vec{g}t$ if $\vec{v}_0(0) = 0$) and $\vec{v}_1$ obeys

$$\dot{\vec{v}}_1 = 2\vec{v}_0 \times \vec{\Omega} = 2t\vec{g} \times \vec{\Omega}. \tag{2.2}$$

The solution is $\vec{r} = \vec{h} + \frac{1}{2}\vec{g}t^2 + \frac{1}{3}t^3\vec{g} \times \vec{\Omega}$. The eastward deviation is the deviation in the direction of $\vec{g} \times \vec{\Omega}$ and its value is $\delta = \frac{1}{3}\bar{t}^3 g\Omega\cos\theta$, where θ is the latitude and $\bar{t}$ the zero$^{\text{th}}$ order time of fall, $\bar{t} = \sqrt{2h/g}$.

While the above example is a nice illustration of the main features of perturbation theory (identification of a perturbation parameter whose powers classify the contributions to the solution, existence of a zero$^{\text{th}}$ order exact solution) the beginning of modern perturbation theory can be traced back to the work of Rayleigh on the theory of sound [33]. In essence, he wondered how the normal modes of a vibrating string

$$\rho(x)\frac{\partial^2 v}{\partial t^2} = \frac{\partial^2 v}{\partial x^2} \qquad\qquad v(0,t) = v(\pi,t) = 0 \tag{2.3}$$

are modified when passing from a constant density $\rho = 1$ to a perturbed density $\rho + \epsilon\sigma(x)$. To solve this problem he wrote down most of the formulae [33, 10] which are still in use to calculate the first order correction to non-degenerate energy levels in quantum mechanics.

The equation for the normal modes is

$$u''(x) + \lambda\rho(x)u(x) = 0, \qquad\qquad u(0) = u(\pi) = 0. \tag{2.4}$$

Let $u_n^{(0)} \equiv \sqrt{2/\pi}\sin nx$ be the unperturbed solution for the n^{th} mode, $\lambda_n = n^2$, and $u_n^{(0)} + \epsilon u_n^{(1)}$ the perturbed solution through first order, corresponding to a frequency $\lambda_n + \epsilon\mu_n$. By writing the equation for $u_n^{(1)}$

$$\frac{d^2u_n^{(1)}}{dx^2} + \lambda_n u_n^{(1)} + \mu_n u_n^{(0)} + \lambda_n\sigma u_n^{(0)} = 0 \qquad u_n^{(1)}(0) = u_n^{(1)}(\pi) = 0 \tag{2.5}$$

after multiplying by $u_r^{(0)}$ and using Green's theorem he found

$$\mu_n = -\lambda_n \int_0^\pi \sigma(x)(u_n^{(0)})^2\,dx, \tag{2.6}$$

$$a_{rn} \equiv \int_0^\pi u_r^{(0)} u_n^{(1)}\, dx = \frac{\lambda_n}{\lambda_r - \lambda_n} \int_0^\pi \sigma u_r^{(0)} u_n^{(0)}\, dx \;\; (r \neq n) \tag{2.7}$$

$$\int_0^\pi u_n^{(0)} u_n^{(1)}\, dx = 0. \tag{2.8}$$

As an application Rayleigh found the position $\pi/2 + \delta x \equiv \pi/2 + \epsilon\tau$ of the nodal point of the perturbed mode $n = 2$ when the perturbation to the density is $\sigma = \kappa\delta(x - \pi/4)$. The vanishing of $u_2^{(0)} + \epsilon u_2^{(1)}$ determines $2\sqrt{2/\pi}\tau = u_2^{(1)}(\pi/2)$. By (2.7) the function $u_2^{(1)}$ has an expansion $u_2^{(1)} = \sum_{n\neq 2} a_{n2} u_n^{(0)}$, $a_{n2} = \frac{4\kappa}{n^2-4} \sin n\pi/4$. The result for τ is

$$\tau = -\frac{2\kappa}{\pi\sqrt{2}}\big(1 + \frac{1}{3} - \frac{1}{5} - \frac{1}{7} + \frac{1}{9} + \frac{1}{11} - \cdots\big) = -\frac{\kappa}{2}.$$

(The series in brackets is equal to $\int_0^1 \frac{1+x^2}{1+x^4} dx = \frac{1}{2}\int_0^\infty \frac{1+x^2}{1+x^4} dx$, which can be calculated by contour integration.)

Perturbation theory was revived by Schrödinger, who introduced it into quantum mechanics in a pioneering work of 1926 [44]. There, he applied the concepts and methods which Rayleigh had put forth to the case where the zero$^{\text{th}}$ order problem was a partial differential equation with non-constant coefficients, and he wrote down, in the language of wave mechanics, all the relevant formulae which yield the correction to the energy levels and to the wave functions for the case of both non-degenerate and degenerate energy levels. As an application he calculated the shift of the energy levels of the hydrogen atom in a constant electric field by two different methods. First he observed that in parabolic coordinates the wave equation is separable also with a constant electric field, which implies that in the subspace of the states with equal zero$^{\text{th}}$ order energy the perturbation is diagonal in the basis of the parabolic eigenfunctions, thus circumventing the intricacies of the degenerate case. Later, he used the spherical coordinates, which entails a non diagonal perturbation matrix and calls for the full machinery of the perturbation theory for degenerate eigenvalues.

It is of no use to repeat here Schrödinger's calculations, since the methods which they use are at the core of modern perturbation theory, which is referred to as the Rayleigh-Schrödinger (RS) perturbation theory. It rapidly superseded other approaches (as that by Born, Heisenberg and Jordan [7], who worked in the framework of the matrix quantum mechanics), and will be presented in the following sections.

3 Presentation of the Problem and an Example

The most frequent application of perturbation theory in quantum mechanics is the approximate calculation of point spectra. The Hamiltonian H is split into an exactly solvable part H_0 (the unperturbed Hamiltonian) plus a term V (the perturbation) which, in a sense to be specified later, is small with respect to H_0: $H = H_0 + V$. In many cases the perturbation contains an adjustable parameter which depends on the actual physical setting. For example, for a system in an external field this parameter is the field strength. For weak fields one expects the spectrum of H to differ only slightly from the spectrum of H_0. In these cases it is convenient to single out the dependence on a parameter by setting

$$H(\lambda) \equiv H_0 + \lambda V. \tag{3.1}$$

Accordingly we will write the Schrödinger equation as

$$H(\lambda)\psi(\lambda) = E(\lambda)\psi(\lambda). \tag{3.2}$$

We will retain the form (3.1) of the Hamiltonian even when H does not contain a variable parameter, thereby understanding that the actual eigenvalues and eigenvectors are the values at $\lambda = 1$.

The basic idea of the RS perturbation theory is that the eigenvalues and eigenvectors of H can be represented as power series

$$\psi(\lambda) = \sum_0^\infty \lambda^n \psi^{(n)} \qquad E(\lambda) = \sum_0^\infty \lambda^n \epsilon_n, \tag{3.3}$$

whose coefficients are determined by substituting expansions (3.3) into (3.2) and equating terms of equal order in λ. Generally, only the first few terms of the series can be explicitly computed, and the primary task of the RS perturbation theory is their calculation. The practicing scientist who uses perturbation theory never has to tackle the mathematical problem of the convergence of the series. This problem, however, or more generally the connection between the truncated perturbation sums and the actual values of the energy and the wave function, is fundamental for the consistency of perturbation theory and will be touched upon in a later section.

Before expounding the technique of the RS perturbation theory we will consider a simple (two-dimensional) problem which can be solved exactly, since in its discussion several features of perturbation theory will emerge clearly, concerning both the behavior of the energy $E(\lambda)$ and the behavior

of the Taylor expansion of this function. From the physical point of view a system with two-dimensional Hilbert space $\mathbb{C}^2$ can be thought of as a particle with spin 1/2 when the translational degrees of freedom are ignored.

Let us write the Hamiltonian $H = H_0 + \lambda V$ in a representation where H_0 is diagonal:

$$H = \begin{pmatrix} E_1^0 & 0 \\ 0 & E_2^0 \end{pmatrix} + \lambda \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} = \begin{pmatrix} E_1^0 + \lambda V_{11} & \lambda V_{12} \\ \lambda V_{12}^* & E_2^0 + \lambda V_{22} \end{pmatrix}. \quad (3.4)$$

We consider first the case $E_1^0 \neq E_2^0$, $V_{12} \neq 0$. The exact eigenvalues $E_{1,2}(\lambda)$ of H are found by solving the secular equation:

$$E_{1,2}(\lambda) = \frac{1}{2}\left[(E_1^0 + \lambda V_{11}) + (E_2^0 + \lambda V_{22}) \pm \sqrt{\Delta(\lambda)}\right], \quad (3.5)$$

$$\Delta(\lambda) \equiv \big((E_1^0 + \lambda V_{11}) - (E_2^0 + \lambda V_{22})\big)^2 + 4\lambda^2 |V_{12}|^2. \quad (3.6)$$

The corresponding eigenvectors, in the so called intermediate normalization defined by $\big(\psi(0), \psi(\lambda)\big) = 1$, are

$$\psi_1(\lambda) = \quad \Big(1, \ \frac{\sqrt{\Delta(\lambda)} - (E_1^0 - E_2^0) - \lambda(V_{11} - V_{22})}{2\lambda V_{12}}\Big) \quad (3.7)$$

$$\psi_2(\lambda) = \quad \Big(-\frac{\sqrt{\Delta(\lambda)} - (E_1^0 - E_2^0) - \lambda(V_{11} - V_{22})}{2\lambda V_{21}}, \ 1\Big). \quad (3.8)$$

Expanding $E_{1,2}(\lambda)$ through order λ^3 we get:

$$E_1(\lambda) = E_1^0 + \lambda V_{11} + \lambda^2 \frac{|V_{12}|^2}{E_1^0 - E_2^0} - \lambda^3 \frac{|V_{12}|^2 (V_{11} - V_{22})}{(E_1^0 - E_2^0)^2} + O(\lambda^4) \quad (3.9)$$

$$E_2(\lambda) = E_2^0 + \lambda V_{22} - \lambda^2 \frac{|V_{12}|^2}{E_1^0 - E_2^0} + \lambda^3 \frac{|V_{12}|^2 (V_{11} - V_{22})}{(E_1^0 - E_2^0)^2} + O(\lambda^4). \quad (3.10)$$

At order 1 only the diagonal matrix elements of V contribute to $E_{1,2}$. The validity of the approximation requires $\lambda |V_{12}| \ll |E_1^0 - E_2^0|$. If this condition is not satisfied, that is if the eigenvalues E_1^0, E_2^0 are "quasi-degenerate", all terms of the expansion can be numerically of the same order of magnitude and no approximation of finite order makes sense.

Note that, within the first order approximation, "level crossing" ($E_1(\lambda) = E_2(\lambda)$) occurs at

$$\bar{\lambda} = -(E_1^0 - E_2^0)/(V_{11} - V_{22}). \quad (3.11)$$

On the other hand Eq. (3.5) shows that level-crossing is impossible, unless $V_{12} = 0$, in which case the first order approximation yields the exact result.

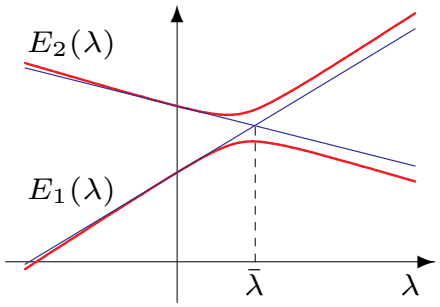

Figure 1: The behavior of the exact eigenvalues $E_{1,2}(\lambda)$ when $V_{12} = 0$ (blue lines) and when $V_{12} \neq 0$ (red lines).

If $V_{12} \neq 0$ the behavior of the levels $E_1(\lambda)$ and $E_2(\lambda)$ near $\bar\lambda$ is shown in Figure 1: the two levels "repel" each other [49].

At first order the eigenvectors $\psi_{1,2}(\lambda)$ are

$$\psi_1^{[1]} = \big(1,\ -\lambda V_{21}/(E_2^0 - E_1^0)\big) \tag{3.12}$$

$$\psi_2^{[1]} = \big(-\lambda V_{12}/(E_1^0 - E_2^0),\ 1\big)\,. \tag{3.13}$$

The expectation value $(\psi_1^{[1]}, H\psi_1^{[1]})/(\psi_1^{[1]}, \psi_1^{[1]})$ of the Hamiltonian over $\psi_1^{[1]}$, for example, is

$$E_1^0 + \lambda V_{11} + \lambda^2 \frac{|V_{12}|^2}{E_1^0 - E_2^0} - \lambda^3 \frac{|V_{12}|^2 (V_{11} - V_{22})}{(E_1^0 - E_2^0)^2} - \lambda^4 \frac{|V_{12}|^4}{(E_1^0 - E_2^0)^3} + O(\lambda^5) \tag{3.14}$$

which agrees with $E_1(\lambda)$ up to the λ^3 terms $\big($the correct fourth order term contains also $|V_{12}|^2 (V_{11} - V_{22})^2/(E_1^0 - E_2^0)^3\big)$. This is an example of Wigner's $(2n+1)$-theorem [52], see Section 4.2.

The power expansions of $E_{1,2}(\lambda)$ and $\psi_{1,2}(\lambda)$ converge in the disk $|\lambda| < |E_1^0 - E_2^0|/\sqrt{(V_{11} - V_{22})^2 + 4|V_{12}|^2}$. The denominator is just twice the infimum over a of the operator norm of $V - aI$. Since adding to V a multiple of the identity does not affect the convergence properties of the Taylor's series of $E(\lambda)$, we see that the convergence domain always contains the disk $|\lambda| < |E_1^0 - E_2^0|/2\|V\|$, a property which holds true for any bounded perturbation in Hilbert space (see Sec. 8).

If H_0 is degenerate, that is $E_1^0 = E_2^0 \equiv E^0$, then the eigenvalues are obtained by diagonalizing V. The degeneracy is removed and the corrections to the eigenvalues are of first order in λ:

$$E_{1,2}(\lambda) = E^0 + \frac{1}{2}\lambda \left(V_{11} + V_{22} \pm \sqrt{(V_{11} - V_{22})^2 + 4|V_{12}|^2} \right) \tag{3.15}$$

while the eigenvectors are λ independent.

The infinite dimensional case is much more involved. In particular, in most cases the perturbation series does not converge at all, that is its radius of convergence vanishes. However, we shall meet again the three situations discussed above: the case of non-degenerate eigenvalues E_n^0 such that $|E_n^0 - E_m^0| \gg |\lambda V_{nm}|$, the case of degenerate eigenvalues and finally the case of "quasi-degenerate" eigenvalues, i.e. groups of eigenvalues $E_{n_i}^0$ such that $|E_{n_i}^0 - E_{n_j}^0| \lesssim |\lambda V_{n_i n_j}|$. As discussed above, in this last case $H_{n_i n_j}$ must be diagonalized exactly prior to applying perturbation theory.

4 Perturbation of Point Spectra: Nondegenerate Case

In this section we consider an eigenvector ψ_0 of H_0 belonging to a non-degenerate eigenvalue E_0 and apply the RS theory to determine the power expansions (3.3) such that Eq. (3.2) is satisfied, the Hamiltonian $H(\lambda)$ being given by (3.1). The case of a degenerate eigenvalue will be considered in Section 5. For both cases the starting point is the substitution of the expansions (3.3) into (3.2), which, upon equating terms with equal powers, yields the following system of equations

$$(H_0 - E_0)\psi^{(n)} + V\psi^{(n-1)} = \sum_{k=0}^{n-1} \psi^{(k)} \epsilon_{n-k}, \qquad n = 1, 2, ... \tag{4.1}$$

A perturbative calculation of the energy and the wave function through order h amounts to calculating ϵ_n and $\psi^{(n)}$ up to $n = h$ and truncating the series in (3.3) at $n = h$.

4.1 Corrections to the energy and the eigenvectors

In the following let ψ_k, E_k be the normalized eigenvectors and the eigenvalues of H_0, and let $\Delta E_{k0} \equiv E_k - E_0$, $V_{hk} \equiv (\psi_h, V\psi_k)$. The correction ϵ_n is recursively defined in terms of the lower order corrections to the energy and the wave function: by left multiplying (4.1) by ψ_0 we find

$$\epsilon_n = (\psi_0, V\psi^{(n-1)}) - \sum_{h=1}^{n-1} (\psi_0, \psi^{(h)}) \epsilon_{n-h}. \tag{4.2}$$

Similarly, the components $(\psi_k, \psi^{(n)})$, $k \neq 0$, are found by left-multiplying by

$\psi_k,\ k \neq 0$:

$$(\psi_k, \psi^{(n)}) = -(\psi_k, V\psi^{(n-1)})\Delta E_{k0}{}^{-1} + \sum_{h=1}^{n-1}(\psi_k, \psi^{(h)})\epsilon_{n-h}\Delta E_{k0}{}^{-1}. \qquad (4.3)$$

Note that, even if the functions $\psi^{(k)}$'s for $k < n$ were known, still $(\psi_0, \psi^{(n)})$ is intrinsically undefined, since to any solution of (4.1) we are allowed to add any multiple of ψ_0. The reason of this indeterminacy is that Eq. (3.2) defines $\psi(\lambda)$ only up to a multiplicative factor $\alpha(\lambda)$. Even the normalization condition $(\psi(\lambda), \psi(\lambda)) = 1$ leaves $\psi(\lambda)$ undetermined by a phase factor $\exp(i\varphi(\lambda)),\ \varphi(\lambda) \in \mathbb{R}$. On the contrary, the corrections ϵ_n as well as all the expectation values (up to order n) are unaffected by these modifications of the wave function $\psi(\lambda)$ [16].

We can turn to our advantage the indeterminacy of $(\psi_0, \psi^{(n)})$ by requiring that in the expression of ϵ_n, Eq. (4.2), the dependence on the values of $(\psi_0, \psi^{(k)})$, $k \leq n-1$, disappears. For example, after writing

$$(\psi_0, V\psi^{(n-1)}) = V_{00}(\psi_0, \psi^{(n-1)}) + \sum_{h\neq 0} V_{0h}(\psi_h, \psi^{(n-1)})$$

the independence of $(\psi_0, \psi^{(n-1)})$ implies $\epsilon_1 = V_{00}$. Next, requiring ϵ_n to be independent of $(\psi_0, \psi^{(n-2)})$ determines ϵ_2 and so on, until finally (4.2) gives ϵ_n. As an example we carry through this procedure for $n = 3$. Starting from

$$\epsilon_3 = V_{00}(\psi_0, \psi^{(2)}) + \sum_{k\neq 0} V_{0k}(\psi_k, \psi^{(2)}) - \epsilon_1(\psi_0, \psi^{(2)}) - \epsilon_2(\psi_0, \psi^{(1)})$$

we first find

$$\epsilon_1 = V_{00}. \qquad (4.4)$$

Next, from (4.3) for $n = 2$, we get

$$\epsilon_3 = -\sum_{k\neq 0}\frac{|V_{0k}|^2}{\Delta E_{k0}}(\psi_0, \psi^{(1)}) - \sum_{h,k\neq 0}\frac{V_{0k}}{\Delta E_{k0}}V_{kh}(\psi_h, \psi^{(1)}) +$$

$$\sum_{k\neq 0}\frac{V_{0k}}{\Delta E_{k0}}(\psi_k, \psi^{(1)})\epsilon_1 - \epsilon_2(\psi_0, \psi^{(1)}).$$

The independence of $(\psi_0, \psi^{(1)})$ implies

$$\epsilon_2 = -\sum_{k\neq 0}\frac{|V_{0k}|^2}{\Delta E_{k0}}. \qquad (4.5)$$

Finally, by using (4.3) for $n = 1$ we find

$$\epsilon_3 = \sum_{h,k\neq 0} \frac{V_{0k}}{\Delta E_{k0}} \frac{V_{kh}}{\Delta E_{h0}} V_{h0} - \epsilon_1 \sum_{k\neq 0} \frac{|V_{0k}|^2}{{\Delta E_{k0}}^2}. \tag{4.6}$$

Note that if ϵ_n is required, the lower order corrections being known, one can use a simplified version of Eqs. (4.2) and (4.3) where the terms $(\psi_0, \psi^{(k)})$ are omitted since the beginning. Once the values of ϵ_k, $k \leq n$, have been determined, Eq. (4.3) yields $(\psi_k, \psi^{(n)})$. For example, for the first order correction to the wave function we have

$$(\psi_k, \psi^{(1)}) = -\frac{V_{k0}}{\Delta E_{k0}}. \tag{4.7}$$

By suitably choosing the arbitrary factor $\alpha(\lambda)$ we referred to after Eq. (4.3) we can impose $(\psi_0, \psi(\lambda)) = 1$. With this choice (known as the "intermediate normalization", since $\psi(\lambda)$ is not normalized) we have $(\psi_0, \psi^{(k)}) = 0$ for any $k > 0$. As a result, for the wave function through order n we find

$$\psi^{[n]} \equiv \psi_0 + \sum_{k=1}^{n} \lambda^k \psi^{(k)} \equiv \psi_0 + \delta_n \psi \tag{4.8}$$

with

$$(\psi_0, \psi^{[n]}) = 1. \tag{4.9}$$

Using the intermediate normalization the expression of ϵ_n is

$$\epsilon_n = (\psi_0, V\psi^{(n-1)}), \tag{4.10}$$

while the value of $(\psi_k, \psi^{(n-1)})$ can be read immediately in the expression of ϵ_n: $(\psi_k, \psi^{(n-1)})$ is obtained from ϵ_n by omitting in each term the factor V_{0k} and the sum over k. For example the wave function $\psi^{[2]} \equiv \psi_0 + \lambda\psi^{(1)} + \lambda^2\psi^{(2)}$ in the intermediate normalization by Equations (4.5) and (4.6) is

$$\psi^{[2]} = \psi_0 - \lambda \sum_{k=1} \psi_k \frac{V_{k0}}{\Delta E_{k0}} +$$
$$\lambda^2 \sum_{h,k=1} \psi_k \frac{V_{kh}}{\Delta E_{k0}} \frac{V_{h0}}{\Delta E_{h0}} - \lambda^2 \epsilon_1 \sum_{k=1} \psi_k \frac{V_{k0}}{{\Delta E_{k0}}^2}. \tag{4.11}$$

In order to calculate expectation values, transition probabilities and so on one needs the normalized wave function

$$\psi_N^{[n]} = N^{1/2}(\psi_0 + \delta_n \psi) \tag{4.12}$$

with $N^{-1} = 1 + (\delta_n\psi, \delta_n\psi)$. N can be chosen real. Note that the wave function $\psi^{[1]}$ is correctly normalized up to first order.

From the above equations one sees in which sense the perturbation V must be small with respect to the unperturbed Hamiltonian H_0: the separation between the unperturbed energy levels must be large with respect to the matrix elements of the perturbation between those levels and the total correction δE to E_0 should be small with respect to $|E_i - E_0|$, E_i standing for any other level of the spectrum of H_0.

4.2 Wigner's theorem

From Eq. (4.1) it follows that

$$H\psi^{[n]} = E^{[n]}\psi^{[n]} + O(\lambda^{n+1}),$$

whence one should infer that, if E is the exact energy, $E - (\psi_N^{[n]}, H\psi_N^{[n]}) = O(\lambda^{n+1})$. It is therefore remarkable Wigner's result that the knowledge of $\psi^{[n]}$ allows the calculation of the energy up to order $2n+1$ (Wigner's $2n+1$ theorem) [52]. Indeed, he proved that, if E is the exact energy,

$$E - \frac{(\psi^{[n]}, H\psi^{[n]})}{(\psi^{[n]}, \psi^{[n]})} = O(\lambda^{2n+2}).$$

To this purpose, let

$$\chi^{(n+1)} = \psi - \frac{\psi^{[n]}}{\sqrt{(\psi^{[n]}, \psi^{[n]})}}$$

where ψ is the normalized exact wave function, $H\psi = E\psi$. Then

$$\chi^{(n+1)} = O(\lambda^{n+1}), \quad (\psi, \chi^{(n+1)}) + (\chi^{(n+1)}, \psi) = -(\chi^{(n+1)}, \chi^{(n+1)}) = O(\lambda^{2n+2}).$$

As a consequence

$$\frac{(\psi^{[n]}, H\psi^{[n]})}{(\psi^{[n]}, \psi^{[n]})} - (\psi, H\psi) = O(\lambda^{2n+2}).$$

We make explicit this point with an example. Since

$$\psi_N^{[1]} = \frac{\psi_0 + \lambda\psi^{(1)}}{\sqrt{1 + \lambda^2(\psi^{(1)}, \psi^{(1)})}},$$

by using (4.7) and recalling (4.5) and (4.6) we have

$$(\psi_N^{[1]}, H\psi_N^{[1]}) = E_0 + \lambda\epsilon_1 + \frac{\lambda^2\epsilon_2 + \lambda^3\epsilon_3}{1 + \lambda^2(\psi^{(1)}, \psi^{(1)})} = E_0 + \lambda\epsilon_1 + \lambda^2\epsilon_2 + \lambda^3\epsilon_3 + O(\lambda^4).$$

4.3 The Feynman–Hellmann theorem

The RS perturbative expansion rests on the hypothesis that both the eigenvalues $E(\lambda)$ and the corresponding eigenvectors $\psi(\lambda)$ admit a power series expansion, in short, that they are analytic functions of λ in a neighborhood of the origin. As we shall see in Sec. 8, as a rule it is not so and the perturbative expansion gives rise only to a formal series. For this reason it is advisable to derive the various terms of the perturbation expansion without assuming analyticity. If we need $E(\lambda)$ and $\psi(\lambda)$ through order n it is sufficient to assume that, as functions of λ, they are C^{n+1}, that is continuously differentiable $(n+1)$ times. The procedure consists in taking the derivatives of (3.2) [15, 28]: at the first step we get

$$H\psi'(\lambda) + V\psi(\lambda) = E'(\lambda)\psi(\lambda) + E(\lambda)\psi'(\lambda) \tag{4.13}$$

and by left multiplication by $\psi(\lambda)$, with $\big(\psi(\lambda), \psi(\lambda)\big) = 1$, we get

$$E'(\lambda) = \big(\psi(\lambda), V\psi(\lambda)\big), \tag{4.14}$$

which is a special case of the Feynman–Hellmann theorem [23, 17]:

$$\frac{\partial E}{\partial \lambda} = (\psi(\lambda), \frac{\partial H}{\partial \lambda}\psi(\lambda)). \tag{4.15}$$

For $\lambda = 0$ we find

$$E'(0) = V_{00}, \tag{4.16}$$

whence $\epsilon_1 = V_{00}$, in agreement with (4.4). Next, after left multiplying (4.13) by ψ_k and taking $\lambda = 0$ we get

$$(\psi_k, \psi') = -\frac{V_{k0}}{\Delta E_{k0}} \tag{4.17}$$

which, again, agrees with (4.7). Taking now the derivative of (2) at $\lambda = 0$ and using (4.16) we obtain

$$E''(0) = 2\sum_{k=1} V_{0k}(\psi_k, \psi') = -2\sum_{k=1} \frac{|V_{0k}|^2}{\Delta E_{k0}} \tag{4.18}$$

whence $\epsilon_2 = \frac{1}{2}E''(0)$, in agreement with (4.5).

It is clear that the procedure can be pursued to any allowed order, and that the results for the energy corrections, as well as for the wave functions,

are the same we obtained earlier by the RS technique. However, the conceptual difference, that no analyticity hypothesis is required, is important since in many cases this hypothesis is not satisfied.

As to the relation of $E^{[n]} \equiv E_0 + \lambda\epsilon_1 + \cdots + \lambda^n\epsilon_n$ with $E(\lambda)$ we recall that, since by assumption $E(\lambda)$ is C^{n+1}, we can write Taylor's formula with a remainder:

$$E(\lambda) = \sum_0^n \frac{E^{(p)}}{p!}\lambda^p \; + \; \frac{E^{(n+1)}(\theta\lambda)}{(n+1)!}\lambda^{n+1}, \qquad 0 < \theta < 1\,. \tag{4.19}$$

As observed in [28], since for small λ the sign of the remainder is the sign of $E^{(n+1)}(0)\lambda^{n+1}$, Eq. (4.19) allows to establish whether the sum in (4.19) underestimates or overestimates $E(\lambda)$. Moreover, if two consecutive terms, say q and $q+1$, have opposite sign, then (for sufficiently small λ) $E(\lambda)$ is bracketed between the partial sums including and excluding the q^{th} term. It is a pity that no one can anticipate how small such a λ should be. (Of course these remarks apply to the RS truncated series as well.)

5 Perturbation of Point Spectra: Degenerate Case

The case when the unperturbed energy E_0 is a degenerate eigenvalue of H_0, i.e. in the Hilbert space there exists a subspace W_0 generated by a set $\{\psi_0^{(i)}\}$, $1 \le i \le n_0$, of orthogonal normalized states, such that each ψ_0 in W_0 obeys $(H_0 - E_0)\psi_0 = 0$, deserves a separate treatment. The main problem is that, if $\psi(\lambda)$ is an eigenstate of the exact Hamiltonian $H = H_0 + \lambda V$, we do not know beforehand which state of W_0 $\psi(0)$ is.

In order to use a more compact notation it is convenient to introduce the projection P_0 onto the subspace W_0 and its complement Q_0

$$P_0\psi = \sum_{i=0}^{n_0} \psi_0^{(i)}(\psi_0^{(i)}, \psi) \qquad Q_0 \equiv I - P_0, \tag{5.1}$$

where $\psi_0^{(i)}$, $1 \le i \le n_0$, is any orthonormal basis of W_0. The Hamiltonian $H = H_0 + \lambda V$ can be written as

$$\begin{gathered} H = (P_0 + Q_0)(H_0 + \lambda V)(P_0 + Q_0) = \\ E_0P_0 + \lambda V_{PP} + \lambda V_{PQ} + \lambda V_{QP} + \lambda V_{QQ} + Q_0H_0Q_0, \end{gathered} \tag{5.2}$$

where

$$V_{PP} = P_0VP_0,\; V_{PQ} = P_0VQ_0,\; V_{QP} = Q_0VP_0,\; V_{QQ} = Q_0VQ_0. \tag{5.3}$$

After projecting the Schrödinger equation onto W_0 and its orthogonal complement $W_0^\perp$, we find

$$(E_0 + \lambda V_{PP})P_0\psi + \lambda V_{PQ}Q_0\psi = EP_0\psi \tag{5.4}$$

$$\lambda V_{QP}P_0\psi + Q_0H_0Q_0\psi + \lambda V_{QQ}Q_0\psi = EQ_0\psi. \tag{5.5}$$

Letting

$$H_{QQ} = Q_0HQ_0 = Q_0H_0Q_0 + \lambda V_{QQ} \tag{5.6}$$

$Q_0\psi$ can be extracted from (5.5):

$$Q_0\psi = \lambda(E - H_{QQ})^{-1}V_{QP}P_0\psi. \tag{5.7}$$

Note that in (5.5) the operator H_{QQ} acts on vectors of $W_0^\perp$ and that $E - H_{QQ}$ does possess an inverse in $W_0^\perp$. Indeed, the existence of a vector ζ in $W_0^\perp$ such that

$$(H_{QQ} - E)\zeta = 0 \tag{5.8}$$

contradicts the assumptions which perturbation theory is grounded in: the separation between $E(\lambda)$ and $E(0)$ should be negligible with respect to the separation between different eigenvalues of H_0. Actually, if ψ_k is such that $H_0\psi_k = E_k\psi_k$, $E_k \neq E_0$, by left multiplying (5.8) by ψ_k we would find

$$(E - E_k)(\psi_k, \zeta) = \lambda(\psi_k, V\zeta)$$

where the LHS is of order 0 in λ, whereas the RHS of order 1.

By substituting (5.7) into (5.4) we have

$$(E_0 + \lambda V_{PP})P_0\psi + \lambda^2 V_{PQ}(E - H_{QQ})^{-1}V_{QP}P_0\psi = EP_0\psi. \tag{5.9}$$

The energy shifts $\Delta E \equiv E - E_0$ appear as eigenvalues of an operator $A(E)$ acting in W_0

$$A(E) \equiv \lambda V_{PP} + \lambda^2 V_{PQ}(E - H_{QQ})^{-1}V_{QP} \tag{5.10}$$

which however still depends on the unknown exact energy E. A calculation of the energy corrections up to a given order is possible, starting from (5.10), provided we expand the term $(E - H_{QQ})^{-1}$ as far as is necessary to include all terms of the requested order.

5.1 Corrections to the energy and the eigenvectors

The contributions ϵ_i are extracted from (5.9) by expanding

$$E = E_0 + \lambda\epsilon_1 + \lambda^2\epsilon_2 + \cdots \qquad P_0\psi = \varphi_0 + \lambda\varphi_1 + \lambda^2\varphi_2 + \cdots$$

and equating terms of equal order. At the first order, since the second term in the LHS of Eq. (5.9) is of order 2 or larger, we have

$$V_{PP}\varphi_0 = \epsilon_1\varphi_0. \tag{5.11}$$

The first order corrections to the energy are the eigenvalues of the matrix V_{PP} and the corresponding zero$^{\text{th}}$ order wave function is the corresponding eigenvector.

In the most favorable case the eigenvalues of V_{PP} are simple, and the degeneracy is completely removed since the first order of perturbation theory. In this case, in order to get the higher order corrections, we can avail ourselves of the arbitrariness in the way of splitting the exact Hamiltonian into a solvable unperturbed Hamiltonian plus a perturbation by putting

$$H = (H_0 + \lambda V_{PP}) + \lambda(V - V_{PP}) \equiv H_0' + \lambda V' \tag{5.12}$$

The eigenvectors of H_0' are the solutions of Eq. (5.11), with eigenvalues $E_0 + \lambda\epsilon_1^{(i)}$, $1 \leq i \leq n_0$, plus the eigenvectors ψ_j of H_0 with eigenvalues $E_j \neq E_0$. Since the eigenvalues $E_0+\lambda\epsilon_1^{(i)}$ are no longer degenerate, the formalism of non-degenerate perturbation theory can be applied, but a warning is in order. When in higher perturbation orders a denominator ΔE_{k0} occurs with the index k referring to another vector of the basis of W_0, this denominator is of order λ and consequently the order of the term containing this denominator is lower than the naive V-counting would imply. In each such term, the effective order is the V-counting order minus the number of these denominators. As shown below, this situation occurs starting from terms of order 4 in the perturbation V. Note that, also in the case of non-complete removal of the degeneracy, the procedure outlined above, with obvious modifications, can be applied to search the higher order corrections to those eigenvalues which at first order turn out to be non-degenerate.

If a residual degeneracy still exists, i.e. an eigenvalue ϵ_1 of Eq. (5.11) is not simple, we must explore the higher order corrections until the degeneracy, if possible, is removed. First of all we must disentangle the contributions of different order in λ from $(E - H_{QQ})^{-1}$. Since

$$E - H_{QQ} = (E - Q_0H_0Q_0)[1 - \lambda(E - Q_0H_0Q_0)^{-1}V_{QQ}],$$

we have

$$(E - H_{QQ})^{-1} = \sum_0^\infty \lambda^n[(E - Q_0H_0Q_0)^{-1}V_{QQ}]^n(E - Q_0H_0Q_0)^{-1}. \tag{5.13}$$

As the energy E still contains contributions of any order, the operator $(E - Q_0 H_0 Q_0)^{-1}$ must in turn be expanded into a series in λ. To make notations more readable, we define

$$\frac{Q_0}{a^n} \equiv (E_0 - Q_0 H_0 Q_0)^{-n}. \tag{5.14}$$

The second order terms from Eqs. (5.9) and (5.13) give

$$V_{PP}\varphi_1 + V_{PQ}\frac{Q_0}{a}V_{QP}\varphi_0 = \epsilon_2\varphi_0 + \epsilon_1\varphi_1. \tag{5.15}$$

Let $P_0^{(i)}$ be the projections onto the subspaces $W_0^{(i)}$ of W_0 corresponding to the eigenvalues $\epsilon_1^{(i)}$:

$$P_0 = \sum_i P_0^{(i)}, \quad V_{PP} = \lambda \sum \epsilon_1^{(i)} P_0^{(i)}, \quad P_1 \equiv P_0^{(1)}, \quad \epsilon_1 \equiv \epsilon_1^{(1)}. \tag{5.16}$$

By projecting onto $W_1 \equiv W_0^{(1)}$ and recalling that φ_0 is in W_1 we get

$$P_1 V_{PQ}\frac{Q_0}{a}V_{QP}\varphi_0 = \epsilon_2\varphi_0, \tag{5.17}$$

whence ϵ_2 is an eigenvalue of the operator

$$V_1 \equiv P_1 V_{PQ}\frac{Q_0}{a}V_{QP}P_1 = P_1 V \frac{Q_0}{a} V P_1. \tag{5.18}$$

Again, if the eigenvalue ϵ_2 is non-degenerate, we can use the previous theory by splitting the Hamiltonian as

$$H = (H_0 + \lambda V_{PP} + \lambda V_1) + \lambda(V - V_{PP} - V_1) \equiv H_0'' + \lambda V''. \tag{5.19}$$

The vectors which make V_1 diagonal belong to non-degenerate eigenvalues of H_0'', hence the non-degenerate theory can be applied. If, on the contrary, the eigenvalue ϵ_2 of V_1 is still degenerate, the above procedure can be carried out one step further, with the aim of removing the residual degeneracy. We work out the calculation for ϵ_3, since a new aspect of degenerate perturbation theory emerges: a truly third order term which is the ratio of a term of order 4 in the potential and a term of first order (see (5.25) below).

From (5.9) and (5.13) we extract the contribution of order 3:

$$\begin{aligned} V_{PP}\varphi_2 + V_{PQ}\frac{Q_0}{a}V_{QP}\varphi_1 - \epsilon_1 V_{PQ}\frac{Q_0}{a^2}V_{QP}\varphi_0 + \\ V_{PQ}\frac{Q_0}{a}V_{QQ}\frac{Q_0}{a}V_{QP}\varphi_0 = \epsilon_1\varphi_2 + \epsilon_2\varphi_1 + \epsilon_3\varphi_0. \end{aligned} \tag{5.20}$$

We want to convert this equation into an eigenvalue problem for ϵ_3. In analogy with (5.16) we have

$$P_1 = \sum_i P_1^{(i)}, \quad V_1 = \sum \epsilon_2^{(i)} P_1^{(i)}, \quad P_2 \equiv P_1^{(1)}, \quad \epsilon_2 \equiv \epsilon_2^{(1)}. \tag{5.21}$$

Since $P_2 V_{PP} = \epsilon_1 P_2$, first we eliminate φ_2 by applying P_2 to (5.20):

$$\begin{aligned} &P_2 V_{PQ} \frac{Q_0}{a} V_{QP} \varphi_1 - \epsilon_1 P_2 V_{PQ} \frac{Q_0}{a^2} V_{QP} \varphi_0 + \\ &P_2 V_{PQ} \frac{Q_0}{a} V_{QQ} \frac{Q_0}{a} V_{QP} \varphi_0 = \epsilon_3 P_2 \varphi_0 + \epsilon_2 P_2 \varphi_1. \end{aligned} \tag{5.22}$$

Writing $\varphi_1 = \sum_i P_0^{(i)} \varphi_1$, since $P_2 V_1 = \epsilon_2 P_2$ the contribution with $i = 1$ of the first term in the LHS of (5.22) is $P_2 V_1 \varphi_1 = \epsilon_2 P_2 \varphi_1$. Hence, Eq. (5.22) reads

$$\begin{aligned} &\sum_{i \neq 1} P_2 V_{PQ} \frac{Q_0}{a} V_{QP} P_0^{(i)} \varphi_1 - \epsilon_1 P_2 V_{PQ} \frac{Q_0}{a^2} V_{QP} \varphi_0 + \\ &P_2 V_{PQ} \frac{Q_0}{a} V_{QQ} \frac{Q_0}{a} V_{QP} \varphi_0 = \epsilon_3 P_2 \varphi_0 + \epsilon_2 \sum_{i \neq 1} P_0^{(i)} \varphi_1. \end{aligned} \tag{5.23}$$

Finally, $P_0^{(i)} \varphi_1$, $i \neq 1$, is extracted from Eq. (5.15) by projecting with $P_0^{(i)}$, $i \neq 1$, and recalling that $P_0^{(i)} \varphi_0 = 0$ if $i \neq 1$:

$$P_0^{(i)} \varphi_1 = P_0^{(i)} V_{PQ} \frac{Q_0}{a} V_{QP} \varphi_0 / (\epsilon_1 - \epsilon_1^{(i)}), \quad i \neq 1. \tag{5.24}$$

Substituting into (5.23) we see that ϵ_3 is defined by the eigenvalue equation for the operator

$$\begin{aligned} V_2 \equiv &P_2 V \frac{Q_0}{a} V \frac{Q_0}{a} V P_2 - \epsilon_1 P_2 V \frac{Q_0}{a^2} V P_2 + \\ &\sum_{i \neq 1} P_2 V \frac{Q_0}{a} V \frac{P_0^{(i)}}{\epsilon_1 - \epsilon_1^{(i)}} V \frac{Q_0}{a} V P_2. \end{aligned} \tag{5.25}$$

Despite the presence of four factors in the potential, the last term is actually a third order term due to the denominators $\epsilon_1 - \epsilon_1^{(i)}$.

The procedure outlined above, which essentially embodies the Rayleigh-Schrödinger approach, can be pursued until the degeneracy is (if possible, see below Sec. 7) completely removed, after which the theory for the non-degenerate case can be used. Rather than detailing the calculations, we present an alternative iterative procedure due to Bloch [4] which allows a more systematic calculation of the corrections to the energy and the wave function.

5.2 Bloch's method

In equations (5.9) and (5.10) we have seen that the energy corrections ΔE and the projections onto W_0 of the vectors $\psi_k(\lambda)$ are eigenvalues and eigenvectors of an operator acting in W_0. This observation is not immediately useful since the operator depends on the unknown exact energy $E(\lambda)$. However, it is possible to produce an operator $B(\lambda)$, which can be calculated in terms of known quantities and has the property that, if $E_k(\lambda)$, $\psi_k(\lambda)$ are eigenvalues and eigenvectors of Eq. (3.2) such that $E_k(0) = E_0$, then

$$B(\lambda)P_0\psi_k(\lambda) = \Delta E_k P_0\psi_k(\lambda). \tag{5.26}$$

First of all, note that the vectors $P_0\psi_k(\lambda)$ are a basis for the subspace W_0. Indeed, it is implicit in the assumption that perturbation theory does work that the perturbing potential should produce only slight modifications of the unperturbed eigenvectors of the Hamiltonian, so that the vectors $P_0\psi_k(\lambda)$ are linearly independent (although not orthogonal). Since their number equals the dimension of W_0, they are a basis for this subspace.

Following [4], we define a λ dependent operator U in this way:

$$UP_0\psi_k(\lambda) = \psi_k(\lambda); \qquad UQ_0 = 0. \tag{5.27}$$

As a consequence we have

$$U = UP_0, \qquad P_0U = P_0, \tag{5.28}$$

$$U\psi_k(\lambda) = \psi_k(\lambda). \tag{5.29}$$

The former of Eqs. (5.28) follows immediately from the definition of U. Hence $P_0U = P_0UP_0$, which implies the latter of (5.28). Equation (5.29) is verified by applying the former of Eqs. (5.28) to $\psi_k(\lambda)$.

Let

$$B(\lambda) \equiv \lambda P_0VU. \tag{5.30}$$

We verify that, if $\Delta E_k \equiv E_k - E_0$, then

$$B(\lambda)P_0\psi_k(\lambda) = \Delta E_k P_0\psi_k(\lambda). \tag{5.31}$$

Indeed, by (5.27) we have $P_0VUP_0\psi_k(\lambda) = P_0V\psi_k(\lambda)$. Writing (3.2) as

$$(H_0 - E_0 + \lambda V)\psi_k(\lambda) = \Delta E_k\psi_k(\lambda)$$

and multiplying by P_0 we find

$$\lambda P_0V\psi_k(\lambda) = \Delta E_k P_0\psi_k(\lambda), \tag{5.32}$$

hence Eq. (5.31) is satisfied.

A practical use of (5.31) requires an iterative definition of U in terms of known quantities. From (5.27) and (5.28) we have

$$U = P_0 U + Q_0 U = P_0 + Q_0 U P_0. \tag{5.33}$$

We calculate the latter term of (5.33) on the vectors $P_0\psi_k(\lambda)$. Since

$$(\lambda V - \Delta E_k)\psi_k = (E_0 - H_0)\psi_k,$$

recalling (5.27) we have

$$Q_0 U P_0 \psi_k(\lambda) = Q_0 \psi_k(\lambda) = \frac{Q_0}{a}(\lambda V - \Delta E_k)\psi_k(\lambda) =$$

$$\lambda \frac{Q_0}{a} V U \psi_k(\lambda) - \Delta E_k \frac{Q_0}{a} U \psi_k(\lambda) = \lambda \frac{Q_0}{a} V U \psi_k(\lambda) - \Delta E_k \frac{Q_0}{a} U P_0 \psi_k(\lambda).$$

By (5.32)

$$Q_0 U P_0 \psi_k(\lambda) = \lambda \frac{Q_0}{a} V U \psi_k(\lambda) - \lambda \frac{Q_0}{a} U P_0 V \psi_k(\lambda) =$$

$$\lambda \frac{Q_0}{a} V U \psi_k(\lambda) - \lambda \frac{Q_0}{a} U P_0 V U \psi_k(\lambda) = \lambda \frac{Q_0}{a} \big(VU - UVU\big) P_0 \psi_k(\lambda).$$

As a consequence the desired iterative equation for U is

$$U = P_0 + \lambda \frac{Q_0}{a}(VU - UVU). \tag{5.34}$$

Equation (5.34) in turn allows an iterative definition of the operator $B(\lambda)$ of (5.30) depending only on quantities which can be computed in terms of the known spectral representation of H_0. Knowing U through order $n-1$ gives $B^{[n]}(\lambda) \equiv \sum_{i=1}^{n} B^{(i)}(\lambda)$, whose eigenvalues are the energy corrections through order n. In fact, if $B = \sum_{i=1}^{\infty} B^{(i)}$ and $P_0\psi_k = \sum_{s=0}^{\infty} \lambda^s \varphi_s$, the order r contribution to (5.31) is

$$\sum_{i=1}^{r} B^{(i)} \varphi_{r-i} = \sum_{i=1}^{r} \epsilon_i \varphi_{r-i}. \tag{5.35}$$

Defining $P_0 \psi_k^{[n]} \equiv \sum_0^n \lambda^r \varphi_r \equiv \varphi^{[n]}$, $\Delta E^{[n]} \equiv \sum_1^n \lambda^r \epsilon_r$, we see that the sum of (5.35) for values of r through n gives

$$B^{[n]} \varphi^{[n]} = \Delta E^{[n]} \varphi^{[n]} + O(\lambda^{n+1}). \tag{5.36}$$

Once $P_0\psi_k^{[n]}(\lambda)$ has been found, Eq. (5.27) gives the component of $\psi_k(\lambda)$ in $W_0^\perp$ through order $n+1$. As an example, for $n=3$ we have

$$U^{[2]} = P_0 + \lambda\frac{Q_0}{a}VP_0 + \lambda^2\frac{Q_0}{a}V\frac{Q_0}{a}VP_0 - \lambda^2\frac{Q_0}{a^2}VP_0VP_0, \tag{5.37}$$

$$B^{[3]} = \lambda P_0VP_0 + \lambda^2 P_0V\frac{Q_0}{a}VP_0 + \lambda^3 P_0V\frac{Q_0}{a}V\frac{Q_0}{a}VP_0 - \lambda^3 P_0V\frac{Q_0}{a^2}VP_0VP_0. \tag{5.38}$$

If W_0 is one dimensional, (5.38) gives for $\lambda\epsilon_1 + \lambda^2\epsilon_2 + \lambda^3\epsilon_3$ the same result as Eqs. (4.4), (4.5) and (4.6).

The main difference between the RS perturbation theory and Bloch's method is that within the former the energy corrections through order n are calculated by means of a sequential computation starting from ϵ_1, with the consequence that at each step the dimension of the matrix to be diagonalized is smaller. Conversely, within Bloch's method one has to diagonalize the matrix $B^{[n]}(\lambda)$, which has the dimension of W_0. However, as noted above, for $n>1$ the eigenvalues of $B^{[n]}(\lambda)$ are different from $\lambda\epsilon_1 + \lambda^2\epsilon_2 + \cdots + \lambda^n\epsilon_n$ by terms of order at least $n+1$. Similarly, the eigenvectors of (5.36) differ from the component in W_0 of $\psi^{[n]} = \psi_0 + \lambda\psi^{(1)} + \cdots + \lambda^n\psi^{(n)}$ by terms of order larger than n.

It is instructive to reconsider the calculation of ϵ_2 and ϵ_3 in the light of Bloch's method. If $P_0 = P_1 + P_1'$, then $P_0VP_0 = \epsilon_1P_1 + P_1'VP_1'$ and

$$B^{[2]}(\lambda) = \lambda\epsilon_1P_1 + \lambda P_1'VP_1' + \lambda^2P_1V\frac{Q_0}{a}VP_1 +$$

$$\lambda^2P_1'V\frac{Q_0}{a}VP_1' + \lambda^2P_1V\frac{Q_0}{a}VP_1' + \lambda^2P_1'V\frac{Q_0}{a}VP_1.$$

The last two terms represent off-diagonal blocks which can be omitted for the calculation of $\lambda\epsilon_1+\lambda^2\epsilon_2$, since the lowest order contribution to the eigenvalues of a matrix X from the off-diagonal terms X_{ij} is $|X_{ij}|^2/(X_{ii}-X_{jj})$. For a second order expansion as $B^{[2]}$ this yields third order contributions of the type

$$\lambda^3P_1V\frac{Q_0}{a}V\frac{P_1'}{\epsilon_1-\epsilon_1'}V\frac{Q_0}{a}VP_1.$$

These are just the contributions to ϵ_3 which we met in the RS approach: the expression of V_2 given in (5.25) combines the block-diagonal term of order 3 with the off-diagonal terms of order 2 giving a third order contribution.

5.3 The quasi–degenerate case

There are cases, in both atomic and molecular physics, where the energy levels of H_0 present a multiplet structure: the energy levels are grouped into "multiplets" whose separation ΔE is large compared to the energy separation δE between the levels belonging to the same multiplet. For instance, in atomic physics this is the case of the fine structure (due to the the so called spin–orbit interaction) or of the hyperfine structure (due to the interaction of the nuclear magnetic moment with the electrons); in molecular physics typically this is the case of the rotational levels associated with the different and widely separated vibrational levels.

If a perturbation V is such that its matrix elements between levels of the same multiplet are comparable to δE, while being small with respect to ΔE, then naive perturbation theory fails because of the small energy denominators pertaining to levels belonging to the same multiplet. To solve this problem, named the problem of quasi-degenerate levels, once again we can exploit the arbitrariness in the way of splitting the Hamiltonian H into an unperturbed Hamiltonian and a perturbation. Let

$$E_0^{(1)} \equiv E_0 + \delta E^{(1)},\ E_0^{(2)} \equiv E_0 + \delta E^{(2)},\ \cdots\ E_0^{(n)} \equiv E_0 + \delta E^{(n)},$$

be the unperturbed energies within a multiplet, with E_0 any value close to the $E_0^{(i)}$'s (for instance their mean value), and $P_0^{(i)}$ the projections onto the corresponding eigenspaces. Let

$$H_0^0 \equiv H_0 - \sum_i \delta E^{(i)} P_0^{(i)}, \qquad \widetilde{V} \equiv \lambda V + \sum_i \delta E^{(i)} P_0^{(i)},$$

so that

$$H = H_0^0 + \widetilde{V}. \tag{5.39}$$

We consider H_0^0 as the unperturbed Hamiltonian and $\widetilde{V}$ as the perturbation. From the physical point of view this procedure, if applied to all multiplets, is just the inclusion into the perturbation of those terms of H_0 that are responsible for the multiplet structure. With the splitting of the Hamiltonian as in (5.39) we can apply the methods of degenerate perturbation theory. The most efficient of these techniques is Bloch's method, which yields a simple prescription for the calculation of the corrections of any order. If for example we are content with the lowest order, we must diagonalize the matrix $P_0 \widetilde{V} P_0$, or equivalently $P_0 H P_0$, that is the energies through first order are the eigenvalues of the equation

$$P_0 H P_0 \psi = E P_0 \psi\,, \tag{5.40}$$

where $P_0 = \sum_i P_0^{(i)}$ is the projection onto W_0, the eigenspace of H_0^0 corresponding to the eigenvalue E_0. These eigenvalues are algebraic functions of λ, and no finite order approximation is meaningful, since all terms can be numerically of the same order, due to the occurrence of small denominators $(\delta E^{(i)} - \delta E^{(j)})^n$.

6 The Brillouin-Wigner Method

Equations (5.10) and (5.13) yield an alternative approach to the calculation of the energy shift ΔE due to a perturbation to a non-degenerate energy level E_0, the so called Brillouin-Wigner method [9, 52, 22]. In this case W_0, the space spanned by the unperturbed eigenvector ψ_0, is one-dimensional. The correction ΔE obeys the equation

$$\Delta E = (\psi_0, A(E)\psi_0), \tag{6.1}$$

where the operator $A(E)$ is defined in (5.10).

Substituting into the expression of A the expansion (5.13) for $(E-H_{QQ})^{-1}$ and noting that, if $\{E_k\}$ is the spectrum of H_0,

$$(\psi_0, V_{PQ}(E - Q_0H_0Q_0)^{-1}V_{QP}\psi_0) = \sum_{k\neq 0} \frac{|V_{0k}|^2}{E - E_k} \ ,$$

$$(\psi_0, V_{PQ}(E - Q_0H_0Q_0)^{-1}V_{QQ}(E - Q_0H_0Q_0)^{-1}V_{QP}\psi_0) = \sum_{k,h\neq 0} V_{0k}(E - E_k)^{-1}V_{kh}(E - E_h)^{-1}V_{h0}$$

and so on, we find the following implicit expression for the exact energy E:

$$E = E_0 + \lambda(\psi_0, V\psi_0) + \lambda^2 \sum_{k\neq 0} \frac{|V_{0k}|^2}{E - E_k} + \lambda^3 \sum_{k,h\neq 0} \frac{V_{0k}}{E - E_k} \frac{V_{kh}}{E - E_h} V_{h0} + \cdots . \tag{6.2}$$

Consistently with the assumption that perturbation theory does work, the denominators in (6.2) are non-vanishing. The equation can be solved by arresting the expansion to a given power n in the potential and searching a solution iteratively starting with $E = E_0$. However, the result differs from the energy $E^{[n]} = E_0 + \lambda\epsilon_1 + \lambda^2\epsilon_2 + \cdots + \lambda^n\epsilon_n$, calculated by means of the RS perturbation theory, by terms of order $n+1$ in the potential. The result of the RS perturbation theory can be recovered from the Brillouin-Wigner approach by substituting in the denominators $E = E_0 + \lambda\epsilon_1 + \lambda^2\epsilon_2 + \cdots + \lambda^n\epsilon_n$,

expanding the denominators in powers of $\lambda^k \epsilon_k / E_0$ and equating terms of equal orders in both sides of (6.2).

As for the perturbed wave function, if the intermediate normalization is used, by (5.7) we have:

$$\psi = \psi_0 + Q_0 \psi = \psi_0 + \lambda (E - H_{QQ})^{-1} V_{QP} \psi_0. \tag{6.3}$$

Again, using the expansion (5.13) we find

$$\psi = \psi_0 + \lambda \sum_{k \neq 0} \psi_k \frac{V_{k0}}{E - E_k} + \lambda^2 \sum_{k,h \neq 0} \psi_k \frac{V_{kh}}{E - E_k} \frac{V_{h0}}{E - E_h} + \cdots \ . \tag{6.4}$$

As for the energy, if we arrest this expression to order n and substitute for E the value calculated by using Eq. (6.2), the result will differ from the one of Rayleigh-Schrödinger perturbation theory by terms of order $n + 1$.

A major drawback of the Brillouin-Wigner method is its lack of size-consistency: for a system consisting of non-interacting subsystems, the perturbative correction to the energy of the total system is not the sum of the perturbative corrections to the energies of the separate subsystems through any finite order. This is best illustrated by the simple case of two systems a, b with unperturbed eigenvectors, energies and interactions ψ_0^a, E_0^a, λV^a and ψ_0^b, E_0^b, λV^b respectively. If for example the expansion (6.2) is arrested at order 2, by noting that the matrix elements $V_{0,ij}$ between the unperturbed state and the states $\psi_i^a \psi_j^b$ are

$$V_{0,ij} \equiv (\psi_0^a \psi_0^b, (V^a + V^b) \psi_i^a \psi_j^b) = (\psi_0^a, V^a \psi_i^a) \delta_{0j} + (\psi_0^b, V^b \psi_j^b) \delta_{0i},$$

for the second order equation defining E we find

$$E = E_0^a + E_0^b + \lambda \epsilon_1^a + \lambda \epsilon_1^b + \lambda^2 \sum_j \frac{|V_{0j}^b|^2}{E - E_0^a - E_j^b} + \lambda^2 \sum_i \frac{|V_{0i}^a|^2}{E - E_0^b - E_i^a} \ . \tag{6.5}$$

On the other hand, for the energy of each system at second order we find

$$E^a = E_0^a + \lambda \epsilon_1^a + \lambda^2 \sum_i \frac{|V_{0i}^a|^2}{E_a - E_i^a}; \qquad E^b = E_0^b + \lambda \epsilon_1^b + \lambda^2 \sum_j \frac{|V_{0j}^b|^2}{E_b - E_j^b} \tag{6.6}$$

It is apparent that the sum of the expression reported in (6.6) does not equal the expression of the energy reported in (6.5). This pathology is absent in the RS perturbation theory, where for non-interacting systems $E(\lambda) = E^a(\lambda) + E^b(\lambda)$, hence, for any j, $\epsilon_j = (1/j!) D^j E(\lambda)|_{\lambda=0} = \epsilon_j^a + \epsilon_j^b$.

7 Symmetry and Degeneracy

In section 5 we applied perturbation theory to the case of degenerate eigenvalues with special emphasis on the problem of the removal of the degeneracy at a suitable order of perturbation theory. The main problem is to know in advance whether the degeneracy can be removed completely, or a residual degeneracy is to be expected. The answer is given by group theory [51, 53, 21].

The very existence of degenerate eigenvalues of a Hamiltonian H is intimately connected with the symmetry properties of this operator. Generally speaking, a group G is a symmetry group for a physical system if there exists an associated set $\{T(g)\}$ of transformations in the Hilbert space of the system such that $|(T(g)\varphi, T(g)\psi)|^2 = |(\varphi, \psi)|^2$, $g \in G$ [53]. It is proven that the operators $T(g)$ must be either unitary or antiunitary [53, 2]. We will consider the most common case that they are unitary and can be chosen in such a way that

$$T(g_1)T(g_2) = T(g_1 g_2), \quad g_1, g_2 \in G, \tag{7.1}$$

so that the operators $\{T(g)\}$ are a representation of G.

A system described by a Hamiltonian H is said to be invariant under the group G if the time evolution operator commutes with $T(g)$. Under fairly wide hypotheses this implies

$$[H, T(g)] = 0, \quad g \in G. \tag{7.2}$$

A consequence is that, for any $g \in G$,

$$H\psi = E\psi \Rightarrow HT(g)\psi = ET(g)\psi, \tag{7.3}$$

that is the restrictions $T(g)|_W$ of the operators $T(g)$ to the space W corresponding to a given energy E are a representation of G. Given an orthonormal basis $\{\psi_i\}$ in W, we have

$$T(g)\psi_i = \sum_j t_{ji}(g)\psi_j \tag{7.4}$$

and the vectors ψ_i are said to transform according to the representation of G described by the matrices t_{ji}.

This representation, apart from the occurrence of the so called accidental degeneracy (which in most cases actually is a consequence of the invariance

of the Hamiltonian under additional transformations) is irreducible: no subspace of W is invariant under all the transformations of the group. As a consequence, knowing the dimensions d_j of the irreducible representations of G allows to predict the possible degree of degeneracy of a given energy level, since the dimension of W must be equal to one of the numbers d_j. If the group of invariance is Abelian, all the irreducible representations are one dimensional, and degeneracy can only be accidental.

Two irreducible representations are equivalent if there are bases which transform with the same matrix $t_{ji}(g)$. Otherwise they are inequivalent. The following orthogonality theorems hold. If a and b are inequivalent representations and $\psi_i^{(a)}$, $\varphi_j^{(b)}$ transform according these representations, then

$$(\psi_i^{(a)}, \varphi_j^{(b)}) = 0 \tag{7.5}$$

while, if a_r and a_s are equivalent, for the basis vectors $\psi_i^{(a_r)}$, $\varphi_j^{(a_s)}$ we have

$$(\psi_i^{(a_r)}, \varphi_j^{(a_s)}) = K_{rs}^{(a)} \delta_{ij}. \tag{7.6}$$

Moreover, if A_{rs} is a matrix which commutes with all the matrices $t_{ij}^{(a)}$ of an irreducible representation b, then $A_{rs} = a\delta_{rs}$ (Schur's Lemma).

7.1 Symmetry and perturbation theory

If $H = H_0 + \lambda V$, let G_0 be the group under which H_0 is invariant. Although it is not the commonest case, we start with assuming that also the perturbation V commutes with $T(g)$ for any element g of G_0. As a rule W_0, the space of eigenvectors of H_0 with energy E_0, hosts an irreducible representation $T(g)$ of G_0. In this case the degeneracy cannot be removed at any order of perturbation theory. While this follows from general principles (for any value of λ, $\psi(\lambda)$ and $T(g)\psi(\lambda)$ are eigenvectors of $H(\lambda)$, and by continuity the eigenspace W_λ will have the same dimension as W_0), it is interesting to understand how the symmetry properties affect the mechanism of perturbation theory.

If $\{\psi_i^0\}$ is a basis of W_0 transforming according to an irreducible representation a of G_0, then the matrix $V_{ij} = (\psi_i^0, V\psi_j^0)$ commutes with all the matrices $t_{ji}^{(a)}(g)$ and, according to Schur's Lemma, $(\psi_i^0, V\psi_j^0) = v\delta_{ij} = \epsilon_1\delta_{ij}$. No splitting occurs at the level of first order perturbation theory, neither can it occur at any higher order. Indeed, when V commutes with the operators $T(g)$, then Bloch's operator U, and consequently the operator $B(\lambda)$ of (5.30), both commute with the $T(g)$'s too. Again by Schur's Lemma, the

operator $B(\lambda)$ is a multiple of the identity. At any order of perturbation the degeneracy of the level is not removed.

In most of the cases, however, the perturbation V does not commute with all the operators $T(g)$. The set

$$G = \{g : \, g \in G_0, \, [T(g), V] = 0\}$$

is a subgroup G of G_0 and the group of invariance for the Hamiltonian H is reduced to G. W_0 generally contains G-irreducible subspaces W_i, $1 \leq i \leq n$: the operators $T(g)|_{W_0}$ are a reducible representation of G. The decomposition into irreducible representations of G is unique up to equivalence.

The crucial information we gain from group theory is the following: the number of energy levels which the energy E_0 is split into is the number of irreducible representations of G which the representation of G_0 in W_0 is split into. The degrees of degeneracy are the dimensions of these representations. What is relevant is that we only need to study the eigenspace W_0 of H_0, which is known by hypothesis.

In fact, let $W(\lambda)$ be the space spanned by the eigenvectors $\psi_k(\lambda)$ of $H(\lambda)$ such that $\psi_k(0) \in W_0$. $W(\lambda)$ is invariant under the operators $T(g)$, $g \in G$, since Bloch's operator U commutes with the operators $T(g)$, $g \in G$. $W(\lambda)$ can be decomposed into G-irreducible subspaces $W_k(\lambda)$, and in each of them by Schur's Lemma the Hamiltonian $H(\lambda)$ is represented by a matrix $E_k(\lambda) I_{W_k(\lambda)}$. The projections $P_0 W_k(\lambda)$ span the space W_0 and transform with the same representation of G as $W_k(\lambda)$, since P_0 commutes with $T(g)$ for any g in G_0, hence for any g in G. Thus, the space W_0 hosts as many irreducible representations of G as $W(\lambda)$. Assuming that the eigenvalues $E_k(\lambda)$ are different from one another, we see that the decomposition of the representation of G_0 in W_0 into irreducible representations of G determines the number and the degeneracy of the eigenvalues of $H(\lambda)$ such that the corresponding eigenvectors $\psi(\lambda)$ are in W_0 for $\lambda = 0$. The possibility that some of the $E_k(\lambda)$ are equal will be touched upon in the next subsection.

Examples where the above mechanism is at work are common in atomic physics. When an atom, whose unperturbed Hamiltonian H_0 is invariant under $O(3)$, is subjected to a constant electric field $\vec{E} = E\hat{z}$ (Stark effect), the invariance group G of its Hamiltonian is reduced to the rotations about the z axis ($SO(2)$) and the reflections with respect to planes containing the z axis. The irreducible representations of this group have dimension at most 2, and the G-irreducible subspaces of W_0 (the space generated by the eigenvectors $\psi_{E_0 lm}$ of H_0 corresponding to the energy E_0) are generated by $\psi_{E_0 l0}$ (one dimensional representation) and $\psi_{E_0 l \pm m}$ (two dimensional representations). Hence, the level E_0 is split into $l + 1$ levels, the states with m and $-m$

remaining degenerate since reflections transform a vector with a given m into the vector with opposite m. Instead, if the atom is subjected to a constant magnetic field $\vec{B} = B\hat{z}$ (Zeeman effect), the surviving invariance group G consists of $SO(2)$ plus the reflections with respect to planes $z = z_0$. G being Abelian, the degeneracy is completely removed and this occurs at the first order of perturbation theory.

In the rather special case that W_0 contains subspaces transforming according to inequivalent representations of G_0, also a G_0-invariant perturbation V can separate in energy the states belonging to inequivalent representations. For example, the spectrum of alkali atoms can be calculated by considering in a first approximation an electron in the field of the unit charged atomic rest, which is treated as pointlike. In this problem the obvious invariance group of the Hamiltonian of the optical electron is $O(3)$, the group of rotations and reflections, and the space W_0 corresponding to the principal quantum number $n > 1$ contains n inequivalent irreducible representations which are labeled by the angular momentum $l \leq n - 1$. When the finite dimension of the atomic rest is taken into account as a perturbation, its invariance under $O(3)$ splits the levels with given n and different l into n sublevels. A more careful consideration, however, shows that also the Lenz vector commutes with the unperturbed Hamiltonian [5], and that the space W_0 is irreducible under a larger group, the group $SO(4)$ [18], which is generated by the angular momentum and the Lenz vector. As a consequence the l degeneracy is by no means accidental: a space irreducible under a given group can turn out to be reducible with respect to one of its subgroups.

Group theory is a valuable tool in degenerate perturbation theory to search the correct vectors $\psi_k(0)$ which make the operator P_0VP_0 diagonal. In fact, let $\psi_i^{(a)}$ be vectors which reduce the representation T of G in W_0 into its irreducible components $T^{(a)}$. The vectors $\psi_i^{(a)}$ and $V\psi_i^{(a)}$ transform according to the same irreducible representation $T^{(a)}$. Hence, by (7.5) and (7.6) we find that the P_0VP_0 is a diagonal block matrix with respect to inequivalent representations:

$$(\psi_i^{(a_r)}, V\psi_j^{(b_s)}) = K_{rs}^{(a)}\delta_{ij}\delta_{ab}, \tag{7.7}$$

with $\delta_{ab} = 1$ if representations a and b are equivalent, $\delta_{ab} = 0$ otherwise. The matrices $K_{rs}^{(a)}$ are generally much smaller than the full matrix of the potential. Thus, the operation of diagonalizing V is made easier by finding the G-irreducible subspaces W_a. Conversely, the reduction of an irreducible representation of a group G_0 in a space W_0 into irreducible representations of a subgroup G can be achieved by the following trick: find an operator V whose symmetry group is just G and interpret W_0 as the degeneracy

eigenspace of a Hamiltonian H_0. The G-irreducible subspaces of W_0 are the eigenspaces of P_0VP_0.

7.2 Level crossing

As shown in the foregoing section, the existence of a non-Abelian group of symmetry for the Hamiltonian entails the existence of degenerate eigenvalues. The problem naturally arises as to whether there are cases when, on the contrary, the degeneracy is truly "accidental", that is it cannot be traced back to symmetry properties. The problem was discussed by J. Von Neumann and E.P. Wigner [49], who showed that for a generic $n \times n$ Hermitian matrix depending on real parameters $\lambda_1, \lambda_2, \cdots$, three real values of the parameters have to be adjusted in order to have the collapse of two eigenvalues (level crossing).

When passing to infinite dimension, arguments valid for finite dimensional matrices might fail. Moreover, often the Hamiltonian is not sufficiently "generic" so that level crossing may occur. As a consequence, we look for necessary conditions in order that, given the Hamiltonian $H(\lambda) = H_0 + \lambda V$, two eigenvalues collapse for some (real) value $\bar{\lambda}$ of the parameter λ: $E_1(\bar{\lambda}) = E_2(\bar{\lambda}) \equiv \bar{E}$. In this case, if $\psi_1(\bar{\lambda})$ and $\psi_2(\bar{\lambda})$ are any two orthonormal eigenvectors of $H(\bar{\lambda}) = H_0 + \bar{\lambda}V$ belonging to the eigenvalue $\bar{E}$, the matrix

$$H_{ij}(\bar{\lambda}) \equiv \big(\psi_i(\bar{\lambda}), (H_0 + \bar{\lambda}V)\psi_j(\bar{\lambda})\big), \quad i, j = 1, 2$$

must be a multiple of the identity:

$$H_{11}(\bar{\lambda}) = H_{22}(\bar{\lambda}) \tag{7.8}$$

$$H_{12}(\bar{\lambda}) = 0. \tag{7.9}$$

Equations (7.8) and (7.9) are three real equations for the unknown $\bar{\lambda}$; hence, except for special cases, level crossing cannot occur.

The condition expressed by Eq. (7.9) is satisfied if the states corresponding to the eigenvalues $E_1(\lambda)$ and $E_2(\lambda)$ possess different symmetry properties, that is if they belong to inequivalent representations of the invariance group of the Hamiltonian or, equivalently, if they are eigenvectors with different eigenvalues of an operator which for any λ commutes with the Hamiltonian $H(\lambda)$ (hence it commutes with both H_0 and V). In this case $H_{12} = 0$ and the occurrence of level crossing depends on whether Eq. (7.8) has a real solution. This explains the statement that level crossing can occur only for states with different symmetry, while states of equal symmetry repel each other. Indeed,

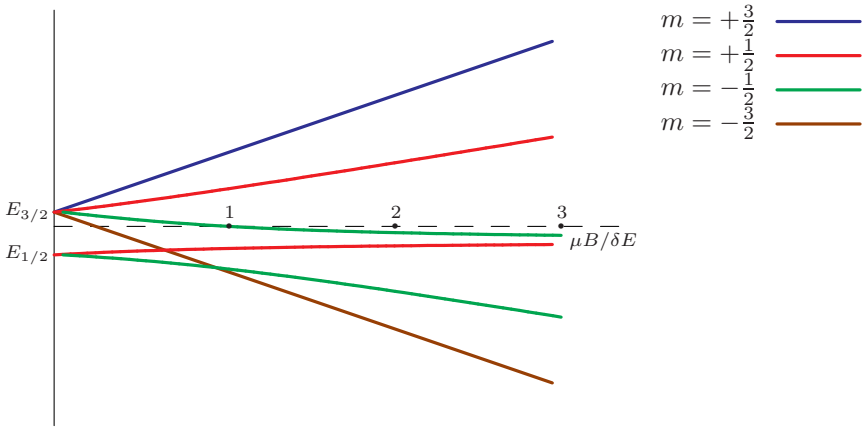

Figure 2: The effect of a magnetic field on the doublet $2p_{1/2}, 2p_{3/2}$ of the lithium whose degeneracies, in the absence of the magnetic field, are respectively 2 and 4. μ is the magnetic moment of the electron and $\mu B/\delta E = 1$ for $B \approx 1.4\,\mathrm{T}$.

if (7.9) is not satisfied, the behavior of two close eigenvalues as functions of λ is illustrated in Figure 1 (Sec. 3).

Figure 2 illustrates the behavior of the quasi-degenerate energy levels $2p_{1/2}, 2p_{3/2}$ of the lithium atom in the presence of an external magnetic field $\vec{B}$. In the absence of the magnetic field they are split by the spin-orbit interaction, with a separation $\delta E \equiv E_{3/2} - E_{1/2} = 0.4 \times 10^{-4}\,\mathrm{eV}$, to be compared with the separation in excess of 1 eV from the adjacent $2s$ and $3s$ levels. This justifies treating the effect of the magnetic field by means of the first order perturbation theory for quasi-degenerate levels.

When the magnetic field is present, the residual symmetry is the (Abelian) group of rotations about the direction of $\vec{B}$. Hence, the Hamiltonian commutes with the component of the angular momentum along the direction of $\vec{B}$, whose eigenvalues are denoted with m. In Figure 2 the energies of states with equal symmetry, that is with the same value of m, are depicted with the same color. No crossing occurs between states with equal m, while the level with $m = -3/2$ does cross both the levels with $m = 1/2$ and with $m = -1/2$ which the $2p_{1/2}$ level is split into.

8 Problems with the Perturbation Series

So far we have assumed that all the power expansions appearing in the

calculations were converging for $|\lambda| \leq 1$, that is we assumed analyticity in λ of $E(\lambda)$. Actually, it is only for rather special cases that analyticity can be proved. For most of the cases of physical interest, even if the terms of the perturbation series can be shown to exist, the series does not converge, or, when it converges, the limit is not $E(\lambda)$. In spite of this, special techniques have been devised to extract a good approximation to $E(\lambda)$ from the (generally few) terms of the perturbation series which can be computed. We will outline the main results existing in the field, without delving into mathematical details, for which we refer the reader to the books of Kato [25] and Reed-Simon [34] and the references therein.

The most favorable case is that of the so called regular perturbations [36, 37, 38, 39], where the perturbation series does converge to $E(\lambda)$. More precisely, if E_0 is a nondegenerate eigenvalue of H_0, for λ in a suitable neighborhood of $\lambda = 0$ the Hamiltonian $H = H_0 + \lambda V$ has a nondegenerate eigenvalue $E(\lambda)$ which is analytic in λ and equals E_0 for $\lambda = 0$. The same property holds for the eigenvector $\psi(\lambda)$. A sufficient condition for this property to hold is expressed by the Kato-Rellich Theorem [36, 37, 38, 39, 26], which essentially states that if the perturbation V is H_0-bounded, in the sense that constants a, b exist such that

$$\|V\psi\| \leq a\|H_0\psi\| + b\|\psi\| \tag{8.1}$$

for any ψ in the domain of V (which must include the domain of H_0) then the perturbation is regular. A lower bound to the radius r such that the perturbation series converges to the eigenvalue $E(\lambda)$ for $|\lambda| < r$ can be given in terms of the parameters a, b appearing in (8.1) and the distance δ of the eigenvalue E_0 from the rest of the spectrum of H_0. We have

$$r = \left[a + \frac{2}{\delta}[b + a(|E_0| + \frac{\delta}{2})]\right]^{-1}. \tag{8.2}$$

It must be stressed, however, that the constants a and b are not uniquely determined by V and H_0. If the perturbation V is bounded ($a = 0$, $b = \|V\|$) condition (8.2) reads $r = \delta/(2\|V\|)$, which implies that the perturbation series for $H = H_0 + V$ with V bounded converges if $\|V\| < \delta/2$ (Kato bound [26]).The analysis of the two-level system (Sec. 3) shows that the figure 1/2 cannot be improved. Still, Kato bound is only a lower bound to r.

A similar statement holds for degenerate eigenvalues [34]: if E_0 has multiplicity m there are m single valued analytic functions $E_k(\lambda)$, $k = 1, ..., m$ such that $E_k(0) = E_0$ and, for λ in a neighborhood of 0, $E_k(\lambda)$ are eigenvalues of $H(\lambda) = H_0 + \lambda V$. Some of the functions $E_k(\lambda)$ may be coincident, and in a neighborhood of E_0 there are no other eigenvalues of $H(\lambda)$.

Regular perturbations are in fact exceedingly rare, a notable case being that of helium-like atoms [47]. Actually, there are cases where, although on physical grounds $H_0 + \lambda V$ does possess bound states, the relationship between $E(\lambda)$ and the RS expansion is far more complicated than for regular perturbations. As pointed out by Kramers [27], with an argument similar to an observation by Dyson [12] for quantum electrodynamics, the quartic anharmonic oscillator with Hamiltonian

$$H = H_0 + \lambda V \equiv \frac{p^2}{2m} + \frac{m\,\omega^2 x^2}{2} + \lambda \frac{m^2\omega^3}{\hbar} x^4 \tag{8.3}$$

is such an example. In fact, on the one hand bound states exist only for $\lambda \geq 0$; on the other hand, if a power series converges for $\lambda > 0$, then the series should converge also for negative values of λ. But for $\lambda < 0$ no bound state exists. Still worse, by estimating the coefficients of the RS expansion it has been proved that the series has vanishing radius of convergence [3].

In spite of this negative result, in this case it has been proved [46] that the perturbation series is an asymptotic series. This means that, for each n, if $\sum_0^n \epsilon_k \lambda^k$ is the sum through order n of the perturbation series, then

$$\lim_{\lambda \to 0} \frac{\sum_0^n \epsilon_k \lambda^k - E(\lambda)}{\lambda^n} = 0. \tag{8.4}$$

We recall the difference between an asymptotic and an absolutely converging series, such as occurs with regular perturbations. For the latter one, given any λ in the convergence range of the series, the distance $|\sum_0^n \epsilon_k \lambda^k - E(\lambda)|$ can be made arbitrarily small provided n is sufficiently large (so that a converging series is also an asymptotic series). On the contrary, for an asymptotic series $|\sum_0^n \epsilon_k \lambda^k - E(\lambda)|$ is arbitrarily small only if λ is sufficiently near 0, but for a definite value of λ the quantity $|\sum_0^n \epsilon_k \lambda^k - E(\lambda)|$ might decrease to a minimum, attained for some value N, and then it could start to oscillate for $n > N$ (this is indeed the case for the anharmonic oscillator). As a consequence, for asymptotic series it is not expedient to push the calculation of the terms of the series beyond the limit where wild oscillations set in.

Any C^∞ function has an asymptotic series, as can be seen by inspection of the Taylor's formula with a remainder (see (4.19)). By this means Krieger [28] argued that, if ϵ_k (or equivalently the k^{th} derivative of $E(\lambda)$) exists for any k, the RS series is asymptotic. However, generally there is not a range where the series converges to $E(\lambda)$, that is $E(\lambda)$ is not analytic. An asymptotic series may fail to converge at all for $\lambda \neq 0$, as noted for the anharmonic oscillator. The asymptotic series of a function, if it exists, is unique, but the converse is not true. For example, for the C^∞ function defined for real x as $f(x) = \exp(-1/x^2)$ if $x \neq 0$, $f(0) = 0$, the asymptotic series vanishes. There

are also cases when the perturbation series is asymptotic for $\arg\lambda$ lying in a range $[\alpha, \beta]$. This occurs for example for the generalized anharmonic oscillator with perturbation $V \propto \lambda x^{2n}$. It has been proved that its perturbation series is asymptotic for $|\arg\lambda| \leq \theta < \pi$ [46] (note that the domain does not include negative values of λ). The result was later extended to multidimensional anharmonic oscillators [19]. General theorems stating sufficient hypotheses for the perturbation series to be asymptotic can be found in the literature. As a rule, however, they do not cover most of the cases of physical interest.

Even in the felicitous case when the perturbation series is asymptotic, it is only known that a partial sum approaches $E(\lambda)$ as much as desired provided λ is sufficiently small. This is not of much help to the practicing scientist, who generally is confronted with a definite value of the parameter λ, which can always be considered $\lambda = 1$ by an appropriate rescaling of the potential V. Recalling that different functions can have the same asymptotic series, it seems hopeless to try to recover the function $E(\lambda)$ from its asymptotic series, but this is possible for the so called strong asymptotic series. A function $E(\lambda)$ analytic in a sectorial region $(0 < |\lambda| < B,\ |\arg\lambda| < \pi/2 + \delta)$ is said to have strong asymptotic series $\sum_0^\infty a_k\lambda^k$ if for all λ in the sector

$$|E(\lambda) - \sum_0^n a_k\lambda^k| \leq C\sigma^{n+1}|\lambda|^{n+1}(n+1)! \tag{8.5}$$

for some constants C, σ. For strong asymptotic series it is proved that the function $E(\lambda)$ is uniquely determined by the series. Conditions that ensure that the RS series is a strong asymptotic series have been given [34].

The problem of actually recovering the function $E(\lambda)$ from its asymptotic series can be tackled by several methods. The most widely used procedure is the Borel summability method [6], which amounts to what follows. Given the strongly asymptotic series $\sum_0^\infty a_k\lambda^k$, one considers the series $F(\lambda) \equiv \sum_0^\infty (a_k/k!)\lambda^k$. This is known as the Borel transform of the initial series, which, by the hypothesis of strong asymptotic convergence, can be proved to have a non-vanishing radius of convergence and to possess an analytic continuation to the positive real axis. Then the function $E(\lambda)$ is given by

$$E(\lambda) = \int_0^\infty F(\lambda x)\exp(-x)\,dx. \tag{8.6}$$

The above statement is Watson's Theorem [50]. Roughly speaking, it yields the function $E(\lambda)$ as if the following exchange of the series with the integral were allowed:

$$E(\lambda) \sim \sum_0^\infty a_k\lambda^k = \sum_0^\infty \frac{a_k}{k!}\int_0^\infty \exp(-x)x^k\,dx\lambda^k =$$

$$\int_0^\infty \exp(-x) \sum_0^\infty \frac{a_k}{k!} (x\lambda)^k \, dx = \int_0^\infty F(\lambda x) \exp(-x) \, dx.$$

A practical problem with perturbation theory is that, apart from a few classroom examples, one is able to calculate only the lower order terms of the perturbation series. Although in principle it is impossible to divine the rest of a series by knowing its terms through a given order, a technique which in some cases turned out to work is that of Padé approximants [32, 1]. A Padé $[M, N]$ approximant to a series is a rational function

$$R_{MN}(z) = \frac{P_M(z)}{Q_N(z)} \tag{8.7}$$

whose power expansion near $z = 0$ is equal to the first $M + N$ terms of the series. It has been proved [31] that the Padé $[N, N]$ approximants converge to the true eigenvalue of the anharmonic oscillator with x^4 or x^6 perturbation. The Padé $[M, N]$ approximant to a function $f(z)$ is unique, but its domain of analyticity is generally larger. Even for asymptotic series whose first terms are known one can write the Padé approximants. One can either use directly the Padé approximant as the value of $E(\lambda)$ for the desired value of λ, or can insert it into the Borel summation method. For the case of the quartic anharmonic oscillator (Eq. (8.3)) both methods have been proved to work (at the cost of calculating some tens of terms of the series).

Another approach to the problem is the method of self-similar approximants [55], whereby approximants to the function $E(\lambda)$ for which the asymptotic series is known are sought by means of products

$$f_{2p}(\lambda) = \prod_{i=1}^{p} (1 + A_i \lambda)^{n_i}. \tag{8.8}$$

The $2p$ parameters A_i, n_i, $1 \leq i \leq p$, are determined by equating the Taylor expansion of $f_{2p}(\lambda)$ with the asymptotic series through order $2p$ ($a_0 = 1$ can be assumed, with no loss of generality, see [55]). Also, odd order approximants f_{2p+1} are possible. For the anharmonic oscillator (Eq. (8.3)) the calculations exhibit a steady convergence to the correct value of the energy of both the even order and the odd order approximants also for $\lambda = 200$.

The problem with the above approaches is that their efficiency seems limited to toy models as the anharmonic oscillator. For realistic problems it is difficult to establish in advance that the method converges to the correct answer.

9 Perturbation of the Continuous Spectrum

In this section we consider the effect of a perturbing potential V on states belonging to the continuous spectrum. Since the problem is interesting mainly for the theory of scattering, we will assume that the unperturbed Hamiltonian H_0 is the free Hamiltonian of a particle of mass m. Also, assuming that the potential $V(\vec{r})$ vanishes at infinity, the spectrum of the free Hamiltonian H_0 and the continuous spectrum of the exact Hamiltonian $H = H_0 + \lambda V$ are equal and consist of the positive real semi-axis. Given an energy $E = \hbar^2 k^2/2m$, the problem is how the potential V affects that particular eigenfunction ψ_0 of H_0 which would represent the state of the system if the interaction potential were absent.

Letting $\psi = \psi_0 + \delta\psi$, the Schrödinger equation reads

$$(E - H_0)\delta\psi = \lambda V(\psi_0 + \delta\psi) \tag{9.1}$$

In the spirit of the perturbation approach, $\delta\psi$ can be calculated by an iterative process provided we are able to find the solution of the inhomogeneous equation

$$(E - H_0)\delta\psi = \zeta \tag{9.2}$$

in the form

$$\delta\psi = \widetilde{G}_0\zeta, \tag{9.3}$$

$\widetilde{G}_0$ being the Green's function of (9.2). Assuming that $\widetilde{G}_0$ is known, we find

$$\begin{aligned} \delta\psi = \quad & \lambda\widetilde{G}_0 V\psi \quad && = \lambda\widetilde{G}_0 V\psi_0 + \lambda\widetilde{G}_0 V\delta\psi = \\ & \lambda\widetilde{G}_0 V\psi_0 && + \lambda\widetilde{G}_0 V(\lambda\widetilde{G}_0 V\psi_0 + \lambda\widetilde{G}_0 V\delta\psi) = \\ & \lambda\widetilde{G}_0 V\psi_0 && + \lambda^2\widetilde{G}_0 V\widetilde{G}_0 V\psi_0 + \lambda^2\widetilde{G}_0 V\widetilde{G}_0 V\delta\psi = \cdots , \end{aligned} \tag{9.4}$$

that is $\delta\psi$ is written as a power expansion in λ in terms of the free wave function ψ_0.

9.1 Scattering solutions and scattering amplitude

One has to decide which eigenfunction of H_0 must be inserted into the above expression, and which Green function $\widetilde{G}_0$ must be used, since, of course, the solution of (9.2) is not unique. The questions are strongly interrelated, and the answers depend on which solution of the exact Schrödinger equation one wishes to find. Since the study of the perturbation of the continuous spectrum is relevant mainly for the theory of potential scattering, we will

focus on this aspect. In the theory of scattering it is shown [24] that, for a potential $V(\vec{r})$ vanishing faster than $1/r$ for $r \to \infty$, a wave function ψ which in the asymptotic region is an eigenfunction of the momentum operator plus an outgoing wave

$$\psi \stackrel{r\to\infty}{\longrightarrow} \exp(i\vec{k}\cdot\vec{r}) + f_{\vec{k}}(\theta,\varphi)\frac{\exp(ikr)}{r} \tag{9.5}$$

(θ, φ being the polar angles with respect to the $\vec{k}$ axis) is suitable for describing the process of diffusion of a beam of free particles with momentum $\vec{k}$ which impinge onto the interaction region and are scattered according the amplitude $f_{\vec{k}}(\theta,\varphi)$. (The character of outgoing wave of the second term in (9.5) is apparent when the time factor $\exp(-iEt)$ is taken into account.) The differential cross section $d\sigma/d\Omega$ is the ratio of the flux of the probability current density due to the outgoing wave to the flux due to the impinging plane wave. One finds

$$\frac{d\sigma}{d\Omega} = |f_{\vec{k}}(\theta,\varphi)|^2. \tag{9.6}$$

In conclusion, we require that ψ_0 is a plane wave, and that the Green function $\widetilde{G}_0$ has to be chosen in such a way as to yield an outgoing wave for large r.

Thus we need to solve the equation

$$(\vec{k}^2 + \Delta)\delta\psi = \lambda\frac{2m}{\hbar^2}V(\exp(i\vec{k}\cdot\vec{r}) + \delta\psi) \equiv U(\vec{r})(\exp(i\vec{k}\cdot\vec{r}) + \delta\psi) \tag{9.7}$$

with the asymptotic condition $\delta\psi \to f_{\vec{k}}\exp(ikr)/r$ for $r \to \infty$. In terms of the Green's function $G_0(\vec{r},\vec{r}\,')$, which satisfies the equation

$$(\Delta + \vec{k}^2)G_0(\vec{r},\vec{r}\,') = \delta(\vec{r}-\vec{r}\,'), \tag{9.8}$$

the solution of (9.7) can be written as

$$\delta\psi(\vec{r}) = \int G_0(\vec{r},\vec{r}\,')U(\vec{r}\,')[(\exp(i\vec{k}\cdot\vec{r}\,') + \delta\psi(\vec{r}\,')]d\vec{r}\,', \tag{9.9}$$

which is a form of the Lippmann-Schwinger equation [30]. The integral operator G_0 with kernel $G_0(\vec{r},\vec{r}\,')$ is connected to the operator $\widetilde{G}_0$ of (9.3) by the equation $\widetilde{G}_0 = 2mG_0/\hbar^2$.

The leading term of $\delta\psi$ for $r \to \infty$ is determined by the leading term of $G_0(\vec{r},\vec{r}\,')$, so we look for a solution of (9.8) with the behavior of outgoing wave for $r \to \infty$. Due to translation and rotation invariance (if both the incoming beam and the scattering potential are translated or rotated by the same amount, the scattering amplitude $f_{\vec{k}}(\theta,\varphi)$ is unchanged), we require for the solution a dependence only on $|\vec{r}-\vec{r}\,'|$.

Recalling that $\Delta\, 1/r = -4\pi\delta(\vec{r})$, we look for a solution of (9.8) with $\vec{r}\,' = 0$ of the form $-F(r)/(4\pi r)$, with $F(0) = 1$. The function $G_0(\vec{r}, \vec{r}\,')$ then will be

$$G_0(\vec{r}, \vec{r}\,') = \frac{F(|\vec{r} - \vec{r}\,'|)}{|\vec{r} - \vec{r}\,'|}. \tag{9.10}$$

The equation for $F(r)$ is

$$F'' + k^2 F = 0, \tag{9.11}$$

whose solutions are $\exp(\pm ikr)$ (outgoing and incoming wave respectively). In conclusion for G_0 we find

$$G_0(\vec{r}, \vec{r}\,') = -\frac{1}{4\pi}\frac{\exp(ik|\vec{r} - \vec{r}\,'|)}{|\vec{r} - \vec{r}\,'|}. \tag{9.12}$$

The solution of the Schrödinger equation with the Green function given in (9.12) is denoted as $\psi^+_{\vec{k}}$ and obeys the integral equation known as the Lippmann-Schwinger equation [30]:

$$\psi^+_{\vec{k}}(\vec{r}) = \exp(i\vec{k}\cdot\vec{r}) - \frac{1}{4\pi}\int \frac{\exp(ik|\vec{r} - \vec{r}\,'|)}{|\vec{r} - \vec{r}\,'|} U(\vec{r}\,')\psi^+_{\vec{k}}(\vec{r}\,')\, d\vec{r}\,'. \tag{9.13}$$

The behavior for $r \to \infty$ can be easily checked to be as in (9.5) by inserting the expansion

$$|\vec{r} - \vec{r}\,'| = r - \frac{\vec{r}\cdot\vec{r}\,'}{r} + O(1/r) \tag{9.14}$$

into the Green function G_0. We find ($\hat{r} \equiv \vec{r}/r$)

$$-\frac{1}{4\pi}\frac{\exp(ik|\vec{r} - \vec{r}\,'|)}{|\vec{r} - \vec{r}\,'|} \overset{r\to\infty}{\longrightarrow} -\frac{1}{4\pi}\frac{\exp[ik(r - \hat{r}\cdot\vec{r}\,')]}{r}\big[1 + \frac{\vec{r}\cdot\vec{r}\,'}{r^2}\big], \tag{9.15}$$

which yields for $\psi^+_{\vec{k}}(\vec{r})$ ($\vec{k}_f \equiv k\hat{r}$)

$$\psi^+_{\vec{k}}(\vec{r}) \overset{r\to\infty}{\longrightarrow} \exp(i\vec{k}\cdot\vec{r}) - \frac{1}{4\pi}\frac{\exp(ikr)}{r}\int \exp\left(-i\vec{k}_f\cdot\vec{r}\,'\right) U(\vec{r}\,')\psi^+_{\vec{k}}(\vec{r}\,')\, d\vec{r}\,'. \tag{9.16}$$

The solutions $\psi^+_{\vec{k}}(\vec{r})$ are normalized as the plane waves $\exp(i\vec{k}\cdot\vec{r})$:

$$(\psi^+_{\vec{k}}, \psi^+_{\vec{k}'}) = (2\pi)^3\delta(\vec{k} - \vec{k}\,'). \tag{9.17}$$

In addition, they are orthogonal to any possible bound state solution of the Schrödinger equation with the Hamiltonian $H = H_0 + \lambda V$. Together with the bound state solutions they constitute a complete set. On a par with

the solutions $\psi^+_{\vec{k}}(\vec{r})$ one can also envisage solutions $\psi^-_{\vec{k}}(\vec{r})$ with asymptotic behavior of incoming wave. They are obtained using for H (see (9.11)) the solution $\exp(-ikr)$. The normalization and orthogonality properties of the functions $\psi^-_{\vec{k}}(\vec{r})$ are the same as for the $\psi^+_{\vec{k}}(\vec{r})$ functions.

From (9.16) we derive an implicit expression for the scattering amplitude $f_{\vec{k}}(\theta,\varphi)$:

$$f_{\vec{k}}(\theta,\varphi) = -\frac{1}{4\pi}\int \exp\left(-i\vec{k}_f\cdot\vec{r}\,'\right)U(\vec{r}\,')\psi^+_{\vec{k}}(\vec{r}\,')\,d\vec{r}\,' \tag{9.18}$$

where the unknown function $\psi^+_{\vec{k}}(\vec{r})$ still appears. Letting $\varphi_f \equiv \exp\left(i\vec{k}_f\cdot\vec{r}\right)$, Eq. (9.18) can also be written as

$$f_{\vec{k}}(\theta,\varphi) = -\frac{1}{4\pi}(\varphi_f, U\psi^+_{\vec{k}}). \tag{9.19}$$

9.2 The Born series and its convergence

Equations (9.13) and (9.19) are the starting point to obtain the expression of the exact wave function $\psi^+_{\vec{k}}(\vec{r})$ and the scattering amplitude $f_{\vec{k}}(\theta,\varphi)$ as a power series in λ, in the spirit of the perturbation approach. Recalling that $U = (2m/\hbar^2)V$ (see (9.7)), if $\varphi_{\vec{k}} \equiv \exp(i\vec{k}\cdot\vec{r})$ for $\psi^+_{\vec{k}}(\vec{r})$ we find

$$\begin{aligned}
\psi^+_{\vec{k}} &= \varphi_{\vec{k}} + \lambda G_0\frac{2m}{\hbar^2}V\varphi_{\vec{k}} + \lambda^2 G_0\frac{2m}{\hbar^2}VG_0\frac{2m}{\hbar^2}V\varphi_{\vec{k}} + \cdots \qquad (9.20)\\
&= \exp(i\vec{k}\cdot\vec{r}) + \lambda\int G_0(\vec{r},\vec{r}\,')\frac{2m}{\hbar^2}V(\vec{r}\,')\exp(i\vec{k}\cdot\vec{r}\,')d\vec{r}\,' +\\
&\lambda^2\int d\vec{r}\,'\int d\vec{r}\,'' G_0(\vec{r},\vec{r}\,')\frac{2m}{\hbar^2}V(\vec{r}\,')G_0(\vec{r}\,',\vec{r}\,'')\frac{2m}{\hbar^2}V(\vec{r}\,'')\exp(i\vec{k}\cdot\vec{r}\,'') + \cdots .
\end{aligned}$$

Inserting the above expansion into (9.19), for the scattering amplitude $f_{\vec{k}}(\theta,\varphi)$ we find

$$f_{\vec{k}}(\theta,\varphi) = \sum_{n=1}^{\infty} f^{(n)}_{\vec{k}}(\theta,\varphi) \tag{9.21}$$

where $f^{(n)}_{\vec{k}}$, the contribution of order n in λ to $f_{\vec{k}}(\theta,\varphi)$, is obtained by substituting $\psi^+_{\vec{k}}$ in (9.19) with the contribution of order $n-1$ of the expansion (9.20). The term of order 1 is called the Born approximation [8], and is given by

$$f^B_{\vec{k}}(\theta,\varphi) \equiv f^{(1)}_{\vec{k}}(\theta,\varphi) = -\frac{1}{4\pi}\left(\frac{2m\lambda}{\hbar^2}\right)\int \exp[i(\vec{k}-\vec{k}_f)\cdot\vec{r}\,']V(\vec{r}\,')\,d\vec{r}\,'. \tag{9.22}$$

The term of order 2 is

$$f_{\vec{k}}^{(2)}(\theta,\varphi) = -\frac{1}{4\pi}\left(\frac{2m\lambda}{\hbar^2}\right)^2 (\varphi_f, VG_0V\varphi_{\vec{k}}) \tag{9.23}$$

and the general term of order n is

$$f_{\vec{k}}^{(n)}(\theta,\varphi) = -\frac{1}{4\pi}\left(\frac{2m\lambda}{\hbar^2}\right)^n (\varphi_f, VG_0V\cdots G_0V\varphi_{\vec{k}}) \quad (n \text{ times } V). \tag{9.24}$$

The scattering amplitude through order n is

$$f_{\vec{k}}^{[n]}(\theta,\varphi) = \sum_{i=1}^{n} f_{\vec{k}}^{(i)}(\theta,\varphi) \tag{9.25}$$

with $f_{\vec{k}}^{(1)}(\theta,\varphi) \equiv f_{\vec{k}}^{B}(\theta,\varphi)$, and the series $\sum_1^\infty f_{\vec{k}}^{(i)}(\theta,\varphi)$ is known as the Born series [8]. Of course, when using (9.25) for calculating the differential cross section $d\sigma/d\Omega$ only terms of order not exceeding n should be consistently retained.

For a discussion of the range of validity and the convergence of the expansions (9.20) and (9.21) it is convenient to pose the problem in the framework of integral equations in the Hilbert space L^2 [40, 54], which provides a natural notion of convergence. To this purpose, since $\psi_{\vec{k}}^+(\vec{r})$ is not square integrable, we start assuming that the potential $V(\vec{r})$ is summable

$$\int |V(\vec{r})| d\vec{r} < \infty \tag{9.26}$$

and multiply Eq. (9.13) by $|V(\vec{r})|^{\frac{1}{2}}$ [41, 20, 45]. Letting $\epsilon_V(\vec{r}) \equiv V(\vec{r})/|V(\vec{r})|$ ($\epsilon_V(\vec{r}) \equiv 0$ if $V(\vec{r}) = 0$) and defining

$$\zeta_{\vec{k}}^+(\vec{r}) \equiv |V(\vec{r})|^{\frac{1}{2}} \psi_{\vec{k}}^+(\vec{r}) \tag{9.27}$$

$$K'(\vec{r},\vec{r}\,') \equiv -\frac{2m}{\hbar^2} G_0(\vec{r},\vec{r}\,')|V(\vec{r})|^{\frac{1}{2}}|V(\vec{r}\,')|^{\frac{1}{2}}\epsilon_V(\vec{r}\,'), \tag{9.28}$$

Eq. (9.13) reads:

$$\zeta_{\vec{k}}^+(\vec{r}) = |V(\vec{r})|^{\frac{1}{2}} \exp(i\vec{k}\cdot\vec{r}) + \lambda \int K'(\vec{r},\vec{r}\,')\zeta_{\vec{k}}^+(\vec{r}\,')\, d\vec{r}\,'. \tag{9.29}$$

Now the function in front of the integral is square integrable and the kernel $K'(\vec{r},\vec{r}\,')$ is square integrable too

$$\int d\vec{r} \int d\vec{r}\,' |K'(\vec{r},\vec{r}\,')|^2 = \int d\vec{r} \int d\vec{r}\,' \frac{|V(\vec{r})||V(\vec{r}\,')|}{|\vec{r}-\vec{r}\,'|^2} < \infty \tag{9.30}$$

provided the potential $V(\vec{r})$ obeys the additional condition

$$\int d\vec{r}\,' \frac{|V(\vec{r}\,')|}{|\vec{r}-\vec{r}\,'|^2} < \infty. \tag{9.31}$$

Equation (9.29) can be formally written as

$$\zeta^+_{\vec{k}} = \zeta_0 + \lambda \hat{K}' \zeta^+_{\vec{k}}, \qquad \zeta_0 \equiv |V(\vec{r})|^{\frac{1}{2}} \exp{(i\vec{k}\cdot\vec{r})}, \tag{9.32}$$

$\hat{K}'$ being the integral operator with kernel K' given in (9.28). The function $\zeta^+_{\vec{k}}$ is formally given as

$$\zeta^+_{\vec{k}} = (I - \lambda \hat{K}')^{-1} \zeta_0 \tag{9.33}$$

where the inverse operator $(I - \lambda \hat{K}')^{-1}$ exists except for those values of λ (singular values) for which $I - \lambda \hat{K}'$ has the eigenvalue 0.

Since by (9.30) $\hat{K}'$ is a compact operator, the singular values are isolated points which obey the inequality $|\lambda| \geq \|\hat{K}'\|^{-1}$, since the spectrum of an operator is contained in the closed disc of radius equal to the norm of the operator. Thus, when $\|\lambda \hat{K}'\| < 1$ the inverse operator $(I - \lambda \hat{K}')^{-1}$ exists and is given by the Neumann series

$$(I - \lambda \hat{K}')^{-1} = I + \lambda \hat{K}' + \lambda^2 \hat{K}'^2 + \cdots \equiv I + R^{K'}_{\lambda}, \tag{9.34}$$

which is clearly norm convergent. By the inequality

$$\|\lambda \hat{K}'\|^2 \leq \lambda^2 \ \mathrm{Tr}(\hat{K}'^{\,\dagger} \hat{K}') = \lambda^2 \int d\vec{r} \int d\vec{r}\,' |K'(\vec{r}, \vec{r}\,')|^2 \tag{9.35}$$

we see that, if

$$\lambda^2 \frac{m^2}{4\pi^2 \hbar^4} \int d\vec{r} \int d\vec{r}\,' \frac{|V(\vec{r})||V(\vec{r}\,')|}{|\vec{r}-\vec{r}\,'|^2} < 1, \tag{9.36}$$

the condition $\|\lambda \hat{K}'\| < 1$ is satisfied and consequently the inverse operator $(I - \lambda \hat{K}')^{-1}$ exists. In conclusion, a sufficient condition for the convergence of the expansion

$$\zeta^+_{\vec{k}} = \zeta_0 + \lambda \hat{K}' \zeta_0 + \lambda^2 K'^2 \zeta_0 + \cdots \tag{9.37}$$

in the L^2 norm is that (9.36) holds [43]. Since $\|\hat{K}'\|^4 \leq \mathrm{Tr}(\hat{K}'^{\,\dagger} \hat{K}' \hat{K}'^{\,\dagger} \hat{K}')$, by the Riemann-Lebesgue Lemma it is possible to prove [56] that for any given λ the condition $\|\lambda \hat{K}'\| < 1$ is satisfied provided the energy $\hbar^2 k^2 / 2m$ is sufficiently large.

The implications for the convergence of the expansion (9.21) of the scattering amplitude are immediate, once Eq. (9.19) is written in the form

$$f_{\vec{k}}(\theta, \varphi) = -\lambda \frac{m}{2\pi\hbar^2} \big(|V(\vec{r})|^{\frac{1}{2}} \varphi_f(\vec{r}) \epsilon_V(\vec{r}), |V(\vec{r})|^{\frac{1}{2}} \psi^+_{\vec{k}}(\vec{r}) \big). \tag{9.38}$$

The Born series converges whenever the iterative solution of (9.32) converges, that is if (9.30) is satisfied. As noted above, for any given λ this happens for sufficiently large energy. An additional useful result is that the Born approximation $f^B_{\vec{k}}(\theta,\varphi)$ (or the expansion $f^{[n]}_{\vec{k}}(\theta,\varphi)$ through any n) converges to the exact scattering amplitude $f_{\vec{k}}(\theta,\varphi)$ when the energy grows to infinity. More precisely [48], if (9.26) holds then

$$|f_{\vec{k}}(\theta,\varphi) - f^{[n]}_{\vec{k}}(\theta,\varphi)| \stackrel{k\to\infty}{\longrightarrow} 0. \tag{9.39}$$

If $\|\lambda\hat{K}'\| \geq 1$ the Neumann series (9.34) does not converge and the perturbation approach is no longer viable. However, if λ is not a singular value Eq. (9.32) can be solved by reducing it to an integral equation with a kernel D of norm less than 1 plus a problem of linear algebra [54]. In fact, for any positive value L the operator $\hat{K}'$ can be approximated by a finite rank operator F

$$F\zeta = \sum_{i=1}^{n} \alpha_i(\vec{r})(\beta_i(\vec{r}\,'), \zeta(\vec{r}\,')) \tag{9.40}$$

such that, if $D \equiv \hat{K}' - F$, $\|D\| < 1/L$. Equation (9.32) then reads

$$(I - \lambda D)\zeta^+_{\vec{k}} = \zeta_0 + \lambda F\zeta^+_{\vec{k}} \tag{9.41}$$

Since for $|\lambda| < L$ we have $\|\lambda D\| < 1$, $\zeta^+_{\vec{k}}$ can be written in terms of the appropriate Neumann series $I + R^D_\lambda$ (see (9.34)):

$$\zeta^+_{\vec{k}} = \zeta_0 + R^D_\lambda\zeta_0 + F\zeta^+_{\vec{k}} + R^D_\lambda F\zeta^+_{\vec{k}}. \tag{9.42}$$

The unknown quantities $(\beta_i, \zeta^+_{\vec{k}})$ which appear in the RHS of (9.42) are determined by solving the linear-algebraic problem obtained by left-multiplying both sides of the equation by β_r, $1 \leq r \leq n$. The values of λ in the range $|\lambda| < L$ for which the algebraic problem is not soluble are the singular values of (9.32) in that range. Thus, Eq. (9.32) can be solved for any non-singular value.

10 Time Dependent Perturbations

A rather different problem is presented by the case that a time independent Hamiltonian H_0, for which the spectrum and the eigenfunctions are known, is perturbed by a time dependent potential $V(t)$. This occurs, for example, when an atom or a molecule interacts with an external electromagnetic field. For the total Hamiltonian $H = H_0 + \lambda V(t)$ stationary states no

longer exist, and the relevant question is how the perturbation affects the time evolution of the system. We assume that the state ψ is known at a given time, which can be chosen as $t = 0$, and search for $\psi(t)$. Obviously, any time t_0 prior to the setting on of the perturbation $\lambda V(t)$ could be chosen instead of $t = 0$.

The time dependent Schrödinger equation reads

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi(t) = H_0\psi(t) + \lambda V(t)\psi(t). \tag{10.1}$$

At any t, $\psi(t)$ can be expanded in the basis of the eigenfunctions $\varphi_n(t)$ of H_0 :

$$\begin{aligned} H_0\varphi_n(t) &= E_n\varphi_n(t) \\ \varphi_n(t) = \varphi_n(0)\exp(-iE_nt/\hbar) &\equiv \zeta_n\exp(-iE_nt/\hbar). \end{aligned} \tag{10.2}$$

For the sake of simplicity we treat H_0 as if it only had discrete spectrum, but the presence of a continuous component of the spectrum does not create any problem.

We can write [11, 42]

$$\psi(t) = \sum_n a_n(t)\varphi_n(t). \tag{10.3}$$

Note that the basis vectors $\varphi_n(t)$ are themselves time dependent (by the phase factor given in(10.2)), whereas the vectors ζ_n are time independent. The isolation of the contribution of H_0 to the time evolution as a time dependent factor allows a simpler system of equations for the unknown coefficients $a_n(t)$.

Substituting expansion (10.3) into (10.1) we find

$$i\hbar\sum_n \dot{a}_n(t)\varphi_n(t) + \sum_n a_n(t)E_n\varphi_n(t) =$$

$$\sum_n a_n(t)E_n\varphi_n(t) + \lambda\sum_n a_n(t)V(t)\varphi_n(t).$$

By left multiplying by $\varphi_k(t)$, for the coefficients $a_k(t)$ we find the system of equations

$$i\hbar\dot{a}_k(t) = \lambda\sum_n a_n(t)(\varphi_k(t), V(t)\varphi_n(t)) \equiv \lambda\sum_n V^I_{kn}(t)a_n(t), \tag{10.4}$$

where we have defined

$$V^I_{kn}(t) = (\varphi_k(t), V(t)\varphi_n(t)). \tag{10.5}$$

The matrix elements $V^I_{kn}(t)$ are the matrix elements of an operator $V^I(t)$ between the time independent vectors ζ_k, ζ_n. Indeed, since

$$\varphi_n(t) = \zeta_n \exp(-iE_n t/\hbar) = \exp(-iH_0 t/\hbar)\zeta_n, \tag{10.6}$$

Eq. (10.5) can be written as

$$V^I_{kn}(t) = \big(\exp(-iH_0 t/\hbar)\zeta_k, V(t)\exp(-iH_0 t/\hbar)\zeta_n\big) \equiv (\zeta_k, V^I(t)\zeta_n), \tag{10.7}$$

where the operator $V^I(t)$ is defined as follows:

$$V^I(t) \equiv \exp(iH_0 t/\hbar)V(t)\exp(-iH_0 t/\hbar). \tag{10.8}$$

System (10.4) must be supplemented with the appropriate initial conditions, which depend on the particular problem. The commonest application of time dependent perturbation theory is the calculation of transition probabilities between eigenstates of H_0. Thus, we assume that at $t = 0$ the system is in an eigenstate of the unperturbed Hamiltonian H_0, say the state φ_1. In this case $a_1(0) = 1$, $a_n(0) = 0$ if $n \neq 1$.

In the spirit of perturbation theory, each a_n is expanded into powers of λ

$$a_1(t) = 1 + \sum_{r=1} \lambda^r a_1^{(r)}(t), \quad a_n(t) = \sum_{r=1} \lambda^r a_n^{(r)}(t) \quad n \neq 1, \tag{10.9}$$

and terms of equal order are equated. For $a_k^{(r)}$ we find

$$i\hbar \dot{a}_k^{(r)} = \sum_n V^I_{kn} a_n^{(r-1)}, \quad r > 0. \tag{10.10}$$

By (10.9), for any $r > 0$, $a_n^{(r)}(0) = 0$. As a consequence, for $r = 1$ we have

$$a_k^{(1)} = \frac{-i}{\hbar}\int_0^t V^I_{k1}(t_1)dt_1 = \frac{-i}{\hbar}\int_0^t (\zeta_k, V(t_1)\zeta_1)\exp(i\Delta E_{k1}t_1/\hbar)\,dt_1, \tag{10.11}$$

where $\Delta E_{k1} \equiv E_k - E_1$. For $r = 2$ we find

$$i\hbar \dot{a}_k^{(2)} = \sum_n V^I_{kn}(t) a_n^{(1)}(t) = \frac{-i}{\hbar}\sum_n V^I_{kn}(t)\int_0^t V^I_{n1}(t_1)dt_1$$

whose solution is

$$a_k^{(2)}(t) = \Big(\frac{-i}{\hbar}\Big)^2 \int_0^t dt_2 \int_0^{t_2} dt_1 \sum_n V^I_{kn}(t_2)V^I_{n1}(t_1) =$$

$$\Big(\frac{-i}{\hbar}\Big)^2 \int_0^t dt_2 \int_0^{t_2} dt_1 \sum_n (\zeta_k, V(t_2)\zeta_n)\exp(i\Delta E_{kn}t_2/\hbar) \times$$

$$(\zeta_n, V(t_1)\zeta_1)\exp(i\Delta E_{n1}t_1/\hbar). \tag{10.12}$$

It is clear how the calculation proceeds for higher values of r. The general expression is

$$a_k^{(r)}(t) = \Big(\frac{-i}{\hbar}\Big)^r \int_0^t dt_r \int_0^{t_r} dt_{r-1} \cdots \int_0^{t_3} dt_2 \int_0^{t_2} dt_1 \times \sum V^I_{kn_r}(t_r) V^I_{n_r n_{r-1}}(t_{r-1}) \cdots V^I_{n_3 n_2}(t_2) V^I_{n_2 1}(t_1). \tag{10.13}$$

By the completeness of the vectors ζ_n, the sums over the intermediate states can be substituted by the identity and the expression of $a_k^{(r)}$ is simplified into

$$a_k^{(r)}(t) = \Big(\frac{-i}{\hbar}\Big)^r \int_0^t dt_r \int_0^{t_r} dt_{r-1} \cdots \int_0^{t_3} dt_2 \int_0^{t_2} dt_1 \big(\zeta_k, V^I(t_r) V^I(t_{r-1}) \cdots V^I(t_2) V^I(t_1) \zeta_1\big). \tag{10.14}$$

It is customary to write (10.14) in a different way. The r-dimensional cube $0 \le t_i \le t,\ 1 \le i \le r$, can be split into $r!$ subdomains

$$0 \le t_{p_1} \le t_{p_2} \cdots \le t_{p_{r-1}} \le t_{p_r} \le t, \tag{10.15}$$

with $\{p_1, p_2, \cdots p_{r-1}, p_r\}$ a permutation of $\{1, 2, \cdots, r-1, r\}$. The time ordered product of r (non-commuting) operators $V^I(t_{p_1})$, $V^I(t_{p_2}), \cdots V^I(t_{p_r})$ is introduced according to the definition

$$T\left[V^I(t_{p_1}) \cdots V^I(t_{p_r})\right] \equiv V^I(t_r) \cdots V^I(t_1), \qquad t_1 \le t_2 \le \cdots \le t_r. \tag{10.16}$$

If $(-i\lambda/\hbar)^r (\zeta_k, T\left[V^I(t_{p_1}) \cdots V^I(t_{p_r})\right] \zeta_1)$ is integrated over the r-cube, then each of the $r!$ subdomains defined by (10.15) yields the same contribution. As a consequence Eq. (10.13) can be written as

$$a_k^{(r)}(t) = \frac{1}{r!}\Big(\frac{-i}{\hbar}\Big)^r \int_0^t dt_r \int_0^t dt_{r-1} \cdots \int_0^t dt_2 \int_0^t dt_1 \big(\zeta_k, T\left[V^I(t_r) V^I(t_{r-1}) \cdots V^I(t_2) V^I(t_1)\right] \zeta_1\big). \tag{10.17}$$

The amplitudes $a_k(t)$ can then be written as

$$a_k(t) = \Big(\zeta_k, T\Big[\exp\Big(\frac{-i\lambda}{\hbar}\Big) \int_0^t V^I(t') dt'\Big] \zeta_1\Big) \tag{10.18}$$

with obvious significance of the T-exponential: each monomial in the V^I operators which appear in the expansion of the exponential is to be time ordered according to the T-prescription. If the initial state is given at time t_0 the integral appearing in the T-exponential should start at t_0. We define

$$U^I(t, t_0) \equiv T\Big[\exp\Big(\frac{-i\lambda}{\hbar}\Big) \int_{t_0}^t V^I(t') dt'\Big]. \tag{10.19}$$

The expansion of the T-exponential into monomials in the V^I operators is called the Dyson series [13, 14]. It is extensively employed in perturbative quantum field theory.

From (10.18) and (10.19) it is easy to derive an expression for the time evolution operator $U(t, t_0)$ such that

$$U(t, t_0)\psi(t_0) = \psi(t). \tag{10.20}$$

Indeed, $\psi(t)$ and $\psi(t_0)$ can be expanded in the basis of the vectors $\varphi_n(t)$ and $\varphi_n(t_0)$ respectively as in (10.3). By the linearity of the Schrödinger equation it suffices to determine $\big(\varphi_k(t), U(t, t_0)\varphi_n(t_0)\big)$, which we already know to be $\big(\zeta_k, U^I(t, t_0)\zeta_n\big)$. We have

$$\big(\varphi_k(t), U(t, t_0)\varphi_n(t_0)\big) = \big(\exp(-iH_0t/\hbar)\zeta_k, U(t, t_0)\exp(-iH_0t_0/\hbar)\zeta_n\big) =$$

$$\big(\zeta_k, \exp(iH_0t/\hbar)U(t, t_0)\exp(-iH_0t_0/\hbar)\zeta_n\big).$$

As a consequence we find the equation

$$U(t, t_0) = \exp(-iH_0t/\hbar)T\Big[\exp\Big(\frac{-i\lambda}{\hbar}\Big)\int_{t_0}^{t} V^I(t')dt'\Big]\exp(iH_0t_0/\hbar), \tag{10.21}$$

which provides the perturbation expansion of the evolution operator $U(t, t_0)$ in powers of λ.

It can be proved that if the operator function $V(t)$ is strongly continuous and the operators $V(t)$ are bounded, then the expansion which defines the T-exponential is norm convergent to a unitary operator, as expected [35]. The restriction to bounded operators $V(t)$ does not detract from the range of applications of Eqs. (10.18) and (10.21), since time dependent perturbation theory is almost exclusively used for treating interactions of a system with external fields, which generate bounded interactions.

11 Future Directions

The long and honorable service of perturbation theory in every sector of quantum mechanics must be properly acknowledged. Its future is perhaps already in our past: the main achievement is its application to quantum field theory where, just to quote an example, the agreement between the measured value and the theoretical prediction of the electron magnetic moment anomaly to ten significant digits has no rivals.

Despite its successes, still perturbation theory is confronted with fundamental questions. In most of realistic problems it is unknown whether the

perturbation series is convergent or at least asymptotic. In non-relativistic quantum mechanics this does not represent a practical problem since only a limited number of terms can be calculated, but in quantum field theory, where higher order terms are in principle calculable, this calls for dedicate investigations. There, in particular, conditions for recovering the exact amplitudes from the first terms of the series by such techniques as the Padé approximants or the self similar approximants, and the estimate on the bound of the error, deserve further investigation.

Somewhat paradoxically, it can be said that the future of perturbation theory is in the non-perturbative results (analyticity domains, large coupling constant behavior, tunneling effect...) - an issue where much work has already been done - since they have proved to be complementary to the use of perturbation theory.

References

[1] Baker GA, Graves-Morris P (1996) Padé approximants, Cambridge Univ. Press, Cambridge

[2] Bargmann V (1964) Note on Wigner's theorem on symmetry operations. Journal Math. Phys. 5:862-868

[3] Bender CM, Wu TT (1969) Anharmonic oscillator. Physical Review 184:1231-1260

[4] Bloch C (1958) Sur la théorie des perturbations des états liées. Nuclear Physics 6:329-347

[5] Böhm A (1993) Quantum mechanics, foundations and applications, Springer Verlag, New York, pp. 208-215

[6] Borel E (1899) Mémoires sur le séries divergentes. Ann. Sci. École Norm. Sup. 16:9-136

[7] Born M, Heisenberg W, Jordan P (1926) Zur Quantenmechanik, II. Z. f. Phys. 35:557-615

[8] Born M (1926) Quantenmechanik der Stossvorgänge. Z. f. Phys, 38:803-827

[9] Brillouin L (1932) Perturbation problem and self consistent field. J. de Phys. et le Radium 3:373-389

[10] Courant R, Hilbert D (1989) Methods of mathematical physics, Vol. I, Wiley, New York, pp 343-350

[11] Dirac PAM (1926) On the theory of quantum mechanics. Proc. Roy. Soc. A112:661-677

[12] Dyson FJ (1952) Divergence of perturbation theory in quantum electrodynamics. Physical Review 85:631-632

[13] Dyson FJ (1949) The radiation theories of Tomonaga, Schwinger and Feynman. Physical Review 75:486-502

[14] Dyson FJ (1949) The S-matrix in quantum electrodynamics. Physical Review 75:1736-1755

[15] Epstein ST (1954) Note on perturbation theory. Amer. Journal Phys. 22:613-614

[16] Epstein ST (1968) Uniqueness of the energy in perturbation theory. Amer. Journal Phys. 36:165-166

[17] Feynman RP (1939) Forces in molecules. Physical Review 56:340-343

[18] Fock VA (1935) Zur Theorie des Wasserstoffatoms. Zeit. f. Physik 98:145-154

[19] Graffi S, Grecchi V, Simon B (1970) Borel summability: application to the anharmonic oscillator. Physics Letters B 32:631-634

[20] Grossman A, (1961) Schrödinger scattering amplitude I. Journal Math. Phys. 3:710-713

[21] Hamermesh M (1989) Group theory and its application to physical problems, Dover, New York, pp. 32-114

[22] Hannabuss K (1997) Introduction to quantum theory, Clarendon, Oxford, pp. 131-136

[23] Hellmann H (1937) Einführung in die Quantenchemie, Deuticke, Leipzig

[24] Joachain J (1983) Quantum collision theory, North-Holland, Amsterdam

[25] Kato T (1966) Perturbation theory for linear operators, Springer Verlag, New York

[26] Kato T (1949) On the convergence of the perturbation method I. Progr. Theor. Phys. 4:514-523

[27] Kramers HA (1957) Quantum mechanics, North-Holland, Amsterdam, pp. 198-202

[28] Krieger JB (1968) Asymptotic properties of perturbation theory. Journal Math. Phys. 9:432-435

[29] Landau LD, Lifshitz EM (1960) Mechanics, Pergamon Press, Oxford, p. 129

[30] Lippmann BA, Schwinger J (1950) Variational principles for scattering processes I. Phys. Rev. 79:469-480

[31] Loeffel J, Martin A, Wightman A, Simon B (1969) Padé approximants and the anharmonic oscillator. Phys. Lett. B30:656-658

[32] Padé H (1899) Sur la représentation approchée d'une fonction pour des fonctions rationelles. Ann. Sci. École Norm. Sup. Suppl. [3] 9:1-93

[33] Rayleigh JW (1894-1896) The theory of sound, Vol. I, Macmillan, London, pp 115-118

[34] Reed M, Simon B (1978) Methods of modern mathematical physics, Vol IV, Academic Press, New York, pp. 10-44

[35] Reed M, Simon B (1975) Methods of modern mathematical physics, Vol II, Academic Press, New York, pp. 282-283

[36] Rellich F (1937) Störungstheorie der Spektralzerlegung I. Math. Ann. 113:600-619

[37] Rellich F (1937) Störungstheorie der Spektralzerlegung II. Math. Ann. 113:677-685

[38] Rellich F (1939) Störungstheorie der Spektralzerlegung III. Math. Ann. 116:555-570

[39] Rellich F (1940) Störungstheorie der Spektralzerlegung IV. Math. Ann. 117:356-382

[40] Riesz F, Sz.-Nagy B (1968) Leçons d'analyse fonctionelle, Gauthier-Villars, Paris, pp. 143-188

[41] Rollnik H (1956) Streumaxima und gebundene Zustände. Z. f. Phys. 145:639-653

[42] Sakurai JJ (1967) Advanced quantum mechanics, Addison-Wesley, Reading, pp. 39-40

[43] Scadron M, Weinberg S, Wright J (1964) Functional analysis and scattering theory. Physical Review 135:B202-B207

[44] Schrödinger E (1926) Quantisierung als Eigenwertproblem. Annalen der Physik 80:437-490

[45] Schwartz J (1960) Some non-self-adjoint operators. Comm. Pure Appl. Math. 13:609-639

[46] Simon B (1970) Coupling constant analyticity for the anharmonic oscillator. Annals of Physics, 58:76-136

[47] Simon B (1991) Fifty years of eigenvalue perturbation theory. Bull. American Math. Soc., 24:303-319

[48] Thirring W (2002) Quantum mathematical physics, Springer Verlag, Berlin, p. 177

[49] von Neumann J, Wigner E (1929) Über das Verhalten von Eigenwerten bei adiabatischen Prozessen. Phys. Zschr. 30: pp. 467-470

[50] Watson G (1912) A theory of asymptotic series. Philos. Trans. Roy. Soc. London Ser. A 211:279-313

[51] Weyl H (1931) The theory of groups and quantum mechanics, Dover, New York

[52] Wigner EP (1935) On a modification of the Rayleigh-Schrödinger perturbation theory. Math. Natur. Anz. (Budapest). 53:477-482

[53] Wigner EP (1959) Group theory and its application to the quantum mechanics of atomic spectra, Academic Press, New York

[54] Yosida K (1991) Lectures on differential and integral equations, Dover, New York, pp. 115-131

[55] Yukalov VI, Yukalova EP (2007) Methods of self similar factor approximants. Physics Letters A 368:341-347

[56] Zemach C, Klein A (1958) The Born expansion in non-relativistic quantum theory I. Nuovo Cimento 10:1078-1087

Books and Reviews

Hirschfelder JO, Byers Brown W, Epstein ST (1964) Recent developments in perturbation theory, in Advances in quantum chemistry, Vol. 1, Academic Press, New York, pp. 255-374

Killingbeck J (1977) Quantum-mechanical perturbation theory. Rep. Progr. Phys. 40:963-1031

Mayer I (2003) Simple theorems, proofs and derivations in quantum chemistry, Kluwer Academic/Plenum Publishers, New York, pp. 69-120

Morse PM, Feshbach H (1953) Methods of theoretical physics, part 2, McGraw-Hill, New York, pp. 1001-1106

Wilcox CH (1966) Perturbation theory and its applications in quantum mechanics, Wiley, New York