Theor Chim Acta (1992) 83: 263–312

Theoretica Chimica Acta

Stationary perturbation theory

I. Survey of basic concepts*

Werner Kutzelnigg
Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, W-4630 Bochum, Germany

Received June 14, 1991/Accepted February 7, 1992

Summary. Stationarity of the energy expectation value as required in the variational approach can in most cases of practical interest be formulated with reference to a unitary 'variational group' and its associated Lie algebra. In terms of a basis of this Lie algebra, Brillouin (hypervirial) conditions and a Hessean are defined. The formulation of stationary perturbation theory (as well as of multiple perturbation theory) in the Lie-algebraic framework is straightforward. If the operator Y, which describes the perturbation of the wave function, is expandable in the Lie algebra of the variational group, a Hellmann–Feynman theorem holds (as a special case of Wigner's $[2n+1]$ rule) and the first order operator Y_1 can be obtained from minimization of a Lie-algebraic Hylleraas functional. Under the same condition for two perturbations, Dalgarno's exchange theorem holds. An analysis of the spectrum the Hessean leads to a generalization of the RPA method for any chosen variational group. Any variational group automatically generates a model excitation spectrum. Sometimes one wants to formulate the variational approach in terms of two or more (independent and noncommuting) variational groups. An example is coupled-MC-SCF theory. One must specify the order in which operators of the two groups act, but otherwise there is not much change with respect to the case of a single group. Time-dependent stationary perturbation theory, based on Frenkel's stationarity principle, is possible on similar lines. Singularities related to an indefinite phase, which plague traditional time dependent perturbation theory are automatically avoided. In the framework of stationary perturbation theory the dipole length and dipole velocity formulas for a transition element are equivalent. For a time-dependent Hamiltonian there is no unique definition of the corresponding energy. There are various possibilities to define a 'pseudo-energy'. One of these definitions is consistent with a special form of a time-dependent Hellmann–Feynman theorem. For a perturbation periodic in time so-called Floquet states exist and stationary time-dependent perturbation theory starting from a stationary state of the unperturbed problem automatically leads to these. For Floquet states a genuine stationarity condition can be derived, that is based on the concept of a generalized Hilbert space and that does not suffer from the

* Dedicated in honor of Prof. Klaus Ruedenberg

shortcomings of Frenkel's principle. The perturbation formalism for these is surprisingly close to that of the time-independent theory. For degenerate and near-degenerate states a quasidegenerate generalization of stationary perturbation theory is possible.

Key words: Stationarity – Perturbation theory – Energy expectation value – Lie algebra – Variational approach

1. Introduction

Most methods for an approximate solution of the (time independent) Schrödinger equation:

$$H\Psi = E\Psi \tag{1.1}$$

are based on the variation principle, i.e. on requiring stationarity of the energy functional:

$$\delta\tilde{E} = \delta\{\langle\Psi|H|\Psi\rangle/\langle\Psi|\Psi\rangle\} = 0 \tag{1.2}$$

with respect to a family of variations of Ψ. If Eq. (1.2) holds for all possible variations of Ψ, Ψ is a solution of Eq. (1.1) and $\tilde{E}$ is equal to an eigenvalue of H, otherwise $\tilde{E}$ is an upper bound to the lowest eigenvalue E_0 of H.

One is often faced with the problem to find eigenvalues and eigenfunctions of a Hamiltonian that depends on some parameter λ:

$$H(\lambda)\Psi(\lambda) = E(\lambda)\Psi(\lambda) \tag{1.3}$$

One can then either use the variation principle for some selected values of λ and interpolate, or alternatively try to find $E(\lambda)$ and $\Psi(\lambda)$ as power series in λ:

$$\Psi = \sum_{k=0}^{\infty} \lambda^k \Psi_k; \qquad E = \sum_{k=0}^{\infty} \lambda^k E_k \tag{1.4a}$$

This is possible – within a radius of convergence for λ – if $H(\lambda)$ is analytic in λ, e.g. a polynomial of degree one:

$$H = H_0 + \lambda H_1 \tag{1.4b}$$

and if certain conditions on H are satisfied [1]. Sometimes the expansion of Eq. (1.4a) in λ for H of the form of Eq. (1.4b) is only asymptotic rather than convergent. The formalism used conventionally to find the coefficients of the power series expansions of Ψ and E is that of perturbation theory. In the derivation of perturbation theory, as found in most textbooks, the starting assumption is made that the 'unperturbed problem', i.e. Eq. (1.3) for $\lambda = 0$ can be solved exactly. In this case a few theorems can be derived, e.g. that E_{2n+1} is expressible in terms of the Ψ_k, $k \leqslant n$ only (Wigner's $[2n+1]$-rule), especially that:

$$E_1 = \left.\frac{\partial E(\lambda)}{\partial\lambda}\right|_{\lambda=0} = \langle\Psi_0|H_1|\Psi_0\rangle \tag{1.5}$$

which is known as Hellmann–Feynman theorem, and that Ψ_1 can be obtained by minimizing the Hylleraas functional:

$$E_2 \leqslant F(\Psi_1) = \langle\Psi_1|H_0 - E_0|\Psi_1\rangle + 2Re\langle\Psi|H_1 - E_1|\Psi_1\rangle \tag{1.6}$$

The traditional formalism of perturbation theory is not very helpful if the conditions under which it has been derived are not satisfied, e.g. if the unperturbed problem cannot be solved exactly. To be on a safe ground in actual computations it would be preferable to use a perturbation formalism that has been derived from the variation principle, i.e. from the requirement that the expectation value of $H(\lambda)$ is stationary independently of λ in a neighborhood of $\lambda = 0$, i.e. that:

$$\delta\tilde{E}(\lambda) = 0; \qquad \tilde{E}(\lambda) = \langle\tilde{\Psi}(\lambda)|H(\lambda)|\tilde{\Psi}(\lambda)\rangle/\langle\tilde{\Psi}(\lambda)|\tilde{\Psi}(\lambda)\rangle \tag{1.7}$$

Attempts on these lines to derive a 'stationary perturbation theory' can be found in the literature, e.g. in Epstein's book [2]. Some theorems of 'exact' perturbation theory can be shown to hold in this framework. Noteworthy is Hurley's study [3] of the conditions under which a special Hellmann–Feynman theorem holds for an approximate zeroth order function.

The purpose of the present paper is a unified formulation of the variational approach (for both linear and nonlinear variations) and of stationary perturbation theory (time-independent and time-dependent, for single and multiple perturbations) in a Lie-algebraic framework. The basic idea is that in most cases of interest stationarity of the energy with respect to a family of variations can be formulated in terms of a *variational group*. In terms of this variational group automatically *generalized Brillouin conditions* (in the form of hypervirial relations) and a *Hessean* are defined. A Lie-algebraic version of the Hylleraas functional appears naturally. Moreover, the variational group generates a *model excitation spectrum* of generalized RPA (random phase) type. Variational stability and symmetry breaking or similar features are easily discussed in this framework.

Many well-known as well as a few less-known or unknown results can so be derived in a rather elegant and transparent way. The power of this formalism will become obvious in the second paper of this series, where it is applied to a study of correlation effects on properties [4] and also in a rederivation and a generalization of the IGLO method for magnetic properties [5] in a forthcoming paper [6] (see also [7]).

There are quite a few precursors to this study. The thorough analysis of the variation principle by Epstein [2] has already been mentioned. Relevant in this context is the unitary perturbation theory of Primas [8], much of the literature on hypervirial theorems [9, 10], in particular on the relation of unitary perturbation theory to hypervirial theorems [11–13]. Closely related is the formulation of the quantum mechanical many-body problem in terms of a hierarchy of Lie groups by Fukutome [14]. Concerning perturbative corrections to expectation values various papers by Sadlej deserve interest [15]. Of course, the author's own experience with perturbation theory in Fock space [16] has been helpful. Lie algebras have also been used in time-dependent theory, mainly in studies of the exact solution of a model problem in which the Hamiltonian is a member of a given Lie algebra [17], but also in the context of stationarity conditions like those considered here [18]. The mathematical background on Lie algebras [19] and on differential manifolds [20] can be found in standard textbooks [19, 20].

This paper is organized as follows. In Sect. 2 the variational group is introduced and generalized Brillouin conditions are derived. The Hessean is defined and a Newton–Raphson approach towards a stationary wave function (or density matrix) is formulated. In Sect. 3 various examples of variational groups, e.g. that corresponding to a *scale transformation*, that for *linear varia-*

tions and that relevant for *Hartree–Fock theory*, are presented. The derivation of stationary perturbation theory is given in Sect. 4, the generalization to *multiple perturbations* in Sect. 5. The *Hellmann–Feynman theorem* and *Dalgarno's exchange theorem* arise in a natural way. Essential is that the perturbation is described *within the variational group*. In Sect. 6 we discuss the *eigenvalue spectrum of the Hessean* and the removal of redundancies that would invalidate the solution of the linear systems of equations needed in the variational approach and in stationary perturbation theory. The search for a metric in operator space immediately leads to a generalization of the Random-phase approximation and it is seen that any variational group automatically generates a model excitation and deexcitation spectrum. Special features of real and imaginary perturbations are discussed in Sect. 7. In Sect. 8 stationary perturbation theory is generalized to the practically interesting case that one cares for *stationarity with respect to two (or more) independent and noncommuting groups* $\mathscr{G}_1$ and $\mathscr{G}_2$, but not with respect to the embedding group $\mathscr{G}$ generated from $\mathscr{G}_1$ and $\mathscr{G}_2$. This case is e.g. realized in CAS-SCF theory. Section 9 gives the generalization to *time-dependent* perturbations, where some singularities that plague the traditional Dirac approach are automatically avoided. The concept of the pseudo-energy and its relation to a time-dependent Hellmann–Feynman theorem is introduced and discussed. The perturbation theory of *Floquet states* for a Hamiltonian periodic in time arises in a simple and natural way. In Sect. 10 we finally study the stationary counterpart of *degenerate* or *quasidegenerate perturbation theory*. Some ideas presented here in detail have been briefly outlined previously [7].

The forthcoming second paper of this series [4] will mainly be devoted to the treatment of electron correlation and especially to correlation effects on (first order or second order) properties. It will be shown that many-body perturbation theory (MBPT) can be formulated in the stationary framework, but that for the Møller–Plesset variant (MP) of MBPT some peculiarities arise due to the different treatment of one-particle and many-particle excitations. This has e.g. the consequence that in double perturbation theory one cannot satisfy simultaneously a Brillouin theorem and a Hellmann–Feynman theorem. In a non-perturbative treatment of correlation effects the superiority of methods based on a Brillouin–Brueckner condition will be shown. The present formalism also allows a straightforward access to coupled MC-SCF theory. In another forthcoming paper [6] the present formalism will be used for a reformulation of the IGLO method [5] for magnetic properties of molecules and the MC-SCF generalization of IGLO. To some extent the present work can be regarded as a preparation for these forthcoming papers [4, 6].

2. Lie-algebraic formulation of the variation principle

We shall only admit normalized trial functions. In this case the variation principle of Eq. (1.2) reduces to:

$$\delta\tilde{E} = \delta\langle\tilde{\Psi}|H|\tilde{\Psi}\rangle = 0 \quad \text{for } \langle\tilde{\Psi}|\tilde{\Psi}\rangle = 1 \tag{2.1}$$

Variations of $\tilde{\Psi}$ that conserve the norm can be formulated as unitary transformations [9, 11–13]:

$$\tilde{\Psi} \to \tilde{\Psi}' = U\tilde{\Psi} \tag{2.2}$$

We require stationarity of $\tilde{E}$ with respect to a family of unitary transformations and it is straightforward to require that these transformations form a group. Let us call it the 'variational group' $\mathcal{G}$. This $\mathcal{G}$ is a unitary group, to which a Lie algebra $\mathcal{L}$ is associated. Any $U \in \mathcal{G}$ can be written as:

$$U = \exp X; \quad X = -X^{\dagger}; \qquad X \in \mathcal{L}_r \subset \mathcal{L}_c \tag{2.3}$$

All the X that figure in Eq. (2.3) are elements of the real Lie algebra $\mathcal{L}_r$ of antihermitean operators X, which is the subalgebra of the full (complex) Lie algebra $\mathcal{L}_c$ associated with $\mathcal{G}$ (see appendix A).

Sometimes the requirement of stationarity with respect to all elements of a unitary group $\mathcal{G}$ may be too restrictive and one will, e.g., be satisfied with stationarity with respect to two (or more) noncommuting subgroups of $\mathcal{G}$ independently. We come to this case in Sect. 8.

If we formulate the variation of $\tilde{E}$ in terms of X we get by means of the Lie–Baker–Campbell–Hausdorff expansion (Hausdorff expansion for short):

$$\tilde{E} \to \tilde{E}' = \tilde{E} + \langle \tilde{\Psi} | [H, X] | \tilde{\Psi} \rangle + \tfrac{1}{2} \langle \tilde{\Psi} | [[H, X], X] | \tilde{\Psi} \rangle + \cdots \tag{2.4}$$

and stationarity of $\tilde{E}$ with respect to infinitesimal variations implies the stationarity conditions:

$$\langle \tilde{\Psi} | [H, X] | \tilde{\Psi} \rangle = 0 \quad \forall X \in \mathcal{L}_r \tag{2.5}$$

For the special case that X is a one-particle excitation operator (see Sect. 3) and $\tilde{\Psi}$ a single Slater determinant this is known as *Brillouin theorem* [21–22] or *Brillouin condition* and for $\tilde{\Psi}$ a multiconfiguration wave function as *generalized Brillouin theorem* or condition. The two alternative names (theorem or condition) depend on whether one regards Eq. (2.5) as a consequence of stationarity or as a condition for stationarity. Multiparticle generalizations of the Brillouin theorem have also been discussed [23]. Relations of the form of Eq. (2.5) are generally referred to as hypervirial theorems [2, 9, 10]. There is a subtle conceptual difference between Brillouin conditions and hyperviral relations in spite of their formal indistinguishability. While the former are directly related to a variational approach, the latter may be completely independent from it, e.g. may be satisfied for any appropriately chosen trial function (this is the case for momentum or torque theorems [2]). This distinction is not always clearcut. A typical hyperviral relation such as the classical virial theorem (see Sect. 3) is obviously related to a stationarity condition for the energy [9].

It does not make too much difference whether we require that Eq. (2.5) holds for all $X \in \mathcal{L}_r$ or all $X \in \mathcal{L}_c$. The latter condition implies, of course, the former, but the converse is also true; if Eq. (2.5) holds for all $X_k \in \mathcal{L}_r$ it also holds for all linear combinations of these X_k with complex coefficients, i.e. for all $X \in \mathcal{L}_c$.

We also need a procedure to construct a wave function Ψ which satisfies Eq. (2.5). We can do so by starting from a trial function Φ which is related to the desired $\tilde{\Psi}$ as:

$$\tilde{\Psi} = U\Phi; \qquad U \in \mathcal{G} \tag{2.6}$$

where the unitary transformation:

$$U = \exp \sigma; \qquad \sigma \in \mathcal{L}_r \tag{2.7}$$

or rather its logarithm σ is now our unknown quantity:

$$E(\sigma) = \langle \Phi | e^{-\sigma} H e^{\sigma} | \Phi \rangle \tag{2.8}$$

We expand σ in a basis X_k of $\mathscr{L}_r$:

$$\sigma = \sum_k c_k X_k \tag{2.9}$$

and use the Hausdorff expansion for Eq. (2.8) to get:

$$E(c_1, \ldots, c_n) = E^{(0)} + \sum_k c_k D_k + \tfrac{1}{2} \sum_{k,l} c_k c_l H_{kl} + \cdots \tag{2.10}$$

$$E^{(0)} = \langle \Phi | H | \Phi \rangle \tag{2.11a}$$

$$D_k = \langle \Phi | [H, X_k] | \Phi \rangle \tag{2.11b}$$

$$H_{kl} = \langle \Phi | [[H, X_k], X_l] | \Phi \rangle \tag{2.11c}$$

We call $\underline{\underline{H}} = \{H_{kl}\}$ the Hessean for this variational problem.

Stationarity of E with respect to variation of the c_k leads to the conditions:

$$D_k + \tfrac{1}{2} \sum_l (H_{kl} + H_{lk}) c_l + \cdots = 0 \tag{2.12a}$$

Alternatively one can expand the Brillouin condition of Eq. (2.5):

$$\begin{aligned} 0 &= \langle \phi | \, e^{-\sigma} [H, X_k] \, e^{\sigma} | \phi \rangle = \langle \phi | [H, X_k] + [[H, X_k], \sigma] + \cdots | \phi \rangle \\ &= D_k + \sum_l H_{kl} c_l + \cdots \end{aligned} \tag{2.12b}$$

which agrees with Eq. (2.12a) except that in Eq. (2.12b) the Hessean is not symmetrized. The difference between Eqs. (2.12a) and (2.12b) has to do with the fact that the symmetrized Hessean is the second derivative $\partial^2 E / \partial c_k \, \partial c_l$ at $c_k = c_l = 0$, while H_{kl} is the second derivative at $c_k = c_k(opt)$, $c_l = 0$. One easily sees that at the stationary point H_{kl} becomes symmetric due to the Jacobi identity

$$\langle \tilde{\Psi} | [[H, X_k], X_l] | \tilde{\Psi} \rangle = \langle \tilde{\Psi} | [[H, X_l], X_k] | \tilde{\Psi} \rangle + \langle \tilde{\Psi} | [H, [X_k, X_l]] | \tilde{\Psi} \rangle \tag{2.13}$$

The second term on the r.h.s. of Eq. (2.13) vanishes in view of Eq. (2.5) because $[X_k, X_l] \in \mathscr{L}_c$. Provided that the test function Φ was sufficiently close to the desired Ψ such that the c_k are all rather small, a solution of the system of Eqs. (2.12) is possible by means of a Newton–Raphson algorithm, i.e. by iterative solution of the linearized system:

$$D_k^{(n)} + \sum_l H_{kl}^{(n)} c_l^{(n+1)} = 0 \tag{2.14a}$$

$$\sigma^{(n)} = \sum_k c_k^{(n)} X_k \tag{2.14b}$$

$$D_k^{(n)} = \langle \Phi | \, e^{-\sigma^{(n)}} [H, X_k] \, e^{\sigma^{(n)}} | \Phi \rangle \tag{2.14c}$$

$$H_{kl}^{(n)} = \langle \Phi | \, e^{-\sigma^{(n)}} [[H, X_k], X_l] \, e^{\sigma^{(n)}}] | \Phi \rangle \tag{2.14d}$$

When the iterative procedure converges, the Brillouin condition of Eq. (2.5) is eventually satisfied.

Whether one uses the asymmetric or the symmetrized Hessean in Eq. (2.14) can only affect the rate of convergence of the iteration scheme, not the final result, because eventually the Hessean will become symmetric.

The system (2.14a) has a unique solution only if the matrix $\underline{\underline{H}}^{(n)}$ is non-singular. This implies that one has to remove redundancies before one solves Eq. (2.14a). We come back to this point in Sect. 6.

If all eigenvalues of the converged Hessean are positive, the expectation value E is a *local minimum*, otherwise a saddle point or a higher-order stationary point [24].

Of course, the final result does not only depend on the choice of the 'variational group' $\mathcal{G}$, but also on the trial function Φ. Since all Φ' that are related to a given Φ by means of a unitary transformation contained in $\mathcal{G}$, are equivalent to Φ, they lead to the same approximate results for the same eigenvalue of H. A particular variational approach is specified by $\mathcal{G}$ and by a member of the family of equivalent Φ.

To give an illustration (for more examples and details see Sect. 3) we may choose the variational group as the group of all unitary one-particle transformations on spin-orbital level and Φ as an arbitrary single Slater determinant. This defines unrestricted Hartree–Fock theory. A generalized Hartree–Fock theory on these lines has probably first been derived by B. Levy [22].

The Lie-algebraic formulation of the variation principle is not limited to pure states. It is easily generalizable to ensemble states. In this case the basic equations are:

$$E = \mathrm{Tr}\{H\Gamma\} \tag{2.15a}$$

$$\mathrm{Tr}\{[H, X]\Gamma\} = 0, \quad \forall X \in \mathcal{L}_r \tag{2.15b}$$

Most early papers on the construction of wave functions that satisfy certain hypervirial relations [9, 11] were concerned with variational problems that depend on a single parameter or occasionally on two parameters. In this context it was observed relatively late [13] that simplifications arise if there is a Lie-algebraic structure. In generalized SCF theory this structure has been exploited much earlier [22], at least implicitly.

3. Examples of variational groups

From now on we no longer use a tilde to indicate trial functions or energies.

Our first example is the group of *scale transformations* [2, 9–12]. We consider the family of wave functions (with λ real):

$$\Psi_\lambda = \Psi(\lambda\vec{r}_1, \lambda\vec{r}_2, \ldots \lambda\vec{r}_n) \tag{3.1}$$

Noting that by means of Taylor's theorem:

$$\Psi_\lambda = \exp\left[(\lambda - 1)\left(\frac{\partial}{\partial\lambda}\right)_{\lambda=1}\right]\Psi_\lambda \tag{3.2}$$

and that

$$\frac{\partial}{\partial\lambda}\Psi_\lambda = \sum_{k=1}^{n} \vec{r}_k \cdot \nabla_k \Psi_\lambda \tag{3.3}$$

We get

$$\Psi_\lambda = \exp\left[(\lambda - 1)\sum_{k=1}^{n} \vec{r}_k \cdot \nabla_k\right]\Psi_1 \tag{3.4}$$

The transformation given by Eq. (3.4) is not unitary (norm-conserving), but unitarity can easily be achieved by taking the antihermitean part of the argument

of the exponential. The group of unitary scale transformations hence has the elements:

$$U_\lambda = \exp[(\lambda - 1)X] \tag{3.5a}$$

$$X = \tfrac{1}{2} \sum_{k=1}^{n} (\vec{r}_k \cdot \nabla_k + \nabla_k \cdot \vec{r}_k) \tag{3.5b}$$

and the Lie algebra $\mathscr{L}_r$ consists of the single element X and its multiples.

Let V be homogeneous in the $\vec{r}_k$ of degree v, then the stationarity condition of Eq. (2.5) becomes:

$$0 = \langle \Psi_\lambda | [H, X] | \Psi_\lambda \rangle = \langle \Psi_\lambda | 2T - vV | \Psi_\lambda \rangle \tag{3.6}$$

which is, of course, nothing but the well-known *virial theorem*, a special hypervirial theorem, that holds if the energy is stationary with respect to scale transformations.

The *one-parameter group* of scale transformations characterizes a particular *non-linear variational* problem. Let us now consider *linear variations* for an N-electron system. As usual we start by taking a finite-dimensional N-electron Hilbert space in which an orthonormal basis $\{\phi_k\}$ can be chosen. The variational group is then that of all unitary transformations within the Hilbert space $\mathscr{H}_N$. A basis of the corresponding complex Lie algebra $\mathscr{L}_c$ consists of the shift operators (for ϕ_k normalized to unity):

$$X_{kl} = |\phi_k\rangle\langle\phi_l| \tag{3.7}$$

while the operators:

$$X_{kl}^- = X_{kl} - X_{lk}; \qquad X_{kl}^+ = i\{X_{kl} + X_{lk}\} \quad \text{for } k < l; \qquad X_{kk}^+ = iX_{kk} \tag{3.8}$$

are a basis of the real Lie algebra $\mathscr{L}_r$ (see appendix A).

Of course, one may choose a different basis $\{\phi'_k\}$ related to $\{\phi_k\}$ by means of a unitary transformation within $\mathscr{G}$ and the shift operators in terms of $\{\phi'_k\}$ are an equivalent basis of $\mathscr{L}_c$.

The stationarity conditions of Eq. (2.5) become:

$$\langle \Psi | [H, X_{kl}] | \Psi \rangle = \langle \Psi | H | \phi_k \rangle \langle \phi_l | \Psi \rangle - \langle \Psi | \phi_k \rangle \langle \phi_l | H | \Psi \rangle = 0, \quad \forall k, l \tag{3.9}$$

It is no loss of generality to choose the ϕ_k such that they diagonalize the matrix representation of the Hamiltonian H in $\mathscr{H}_N$. Then one sees easily that Eq. (3.9) is satisfied if Ψ is equal to one of the ϕ_k.

The recommended (and, of course, well-known) way to satisfy Eq. (3.9) consists hence in diagonalizing the matrix representation of the Hamiltonian. Alternatively one can apply the Newton–Raphson approach characterized by Eq. (2.14) or an efficient variant of it, as shown in appendix B.

If in the same variational space we consider stationarity of the energy with respect to an ensemble, Eq. (2.15b) becomes:

$$\mathrm{Tr}\{[\Gamma, H] X_{kl}\} = \mathrm{Tr}\langle \phi_l | [\Gamma, H] | \phi_k \rangle = 0 \tag{3.10}$$

This is satisfied if the projections of Γ and H onto $\mathscr{H}_N$ commute, i.e. if they have common eigenfunctions in $\mathscr{H}_N$, or that Γ is of the form:

$$\Gamma = \sum_k v_k |\phi_k\rangle\langle\phi_k| \tag{3.11}$$

in terms of the ϕ_k that diagonalize the matrix $\langle \phi_k | H | \phi_l \rangle$.

A very important variational group is that of *one-particle transformations*. The generators of this group (hence a basis of $\mathscr{L}_c$) are the one-particle excitation operators a_p^q, if one considers transformations on spin-orbital level, or E_P^Q for the transformations between spinfree orbitals.

$$a_q^q = a^q a_p = a_q^\dagger a_p \tag{3.12a}$$

$$E_Q^P = a_{Q\alpha}^{P\alpha} + a_{Q\beta}^{P\beta} \tag{3.12b}$$

We follow the tensor notation for (particle number conserving) normal ordered products of creation and annihilation operators [14, 16]. The operator $a^p = a_p^\dagger$ creates an electron in spin-orbital χ_p, while a_p is the corresponding annihilation operator. The set $\{\chi_p\}$ of spin-orbitals is, for convenience, chosen orthonormal.

A basis of $\mathscr{L}_r$ consists of:

$$a_q^p - a_p^q, \qquad i(a_q^p + a_p^q) \quad \text{for } p < q; \qquad ia_p^p \tag{3.13}$$

and analogously in the spinfree case.

The group generated by the a_q^p is relevant for unrestricted Hartree–Fock theory, that generated by the E_Q^P for restricted Hartree–Fock theory. Either group defines a non-linear variational problem.

Unfortunately the *two-electron transformations*:

$$U = \exp\left\{\sum_{p,q,r,s} \lambda_{pq}^{rs} a_{rs}^{pq}\right\}; \quad \lambda_{pq}^{rs} = -\lambda_{rs}^{pq*} \tag{3.14}$$

that have been considered occasionally [20], don't constitute a unitary group (except for genuine two-electron systems), since the a_{rs}^{pq} are not a basis of a Lie algebra. In fact, the commutator of two-particle operators a_{rs}^{pq} and a_{vw}^{tu} also contains three-particle operators etc. So with respect to transformations of the type of Eq. (3.14) the present theory is not applicable, unless one includes for an N-electron system excitation operators of all particle rank up to N.

The variational group corresponding to *full CI* is generated by the excitation operators:

$$a_q^p, a_{rs}^{pq}, a_{stu}^{pqr}, \ldots \tag{3.15}$$

up to N-particle excitation rank, in terms of a given one-electron basis $\{\chi_p\}$.

Our main concern is stationarity of E. It may often also be recommended to check whether E is a (local) minimum, i.e. whether all eigenvalues of the Hessean are non-negative [24].

Sometimes one is interested in relating stationarity or minimum properties of an energy expectation values with respect to two variational groups $\mathscr{G}_1$ and $\mathscr{G}_2$ such that one is a subgroup of the other, say $\mathscr{G}_2 \supset \mathscr{G}_1$. If E determined as a minimum with respect to $\mathscr{G}_1$, remains a minimum with respect to $\mathscr{G}_2$, it is said to be 'stable' against extension from $\mathscr{G}_1$ to $\mathscr{G}_2$. This is somehow an indication that $\mathscr{G}_1$ was well chosen. If E changes to first order for first-order variations within $\mathscr{G}_2$, but not within $\mathscr{G}_1$, i.e. if (for $\mathscr{L}_1$ and $\mathscr{L}_2$ the Lie algebras associated with $\mathscr{G}_1$ and $\mathscr{G}_2$):

$$\langle \Psi | [H, X] | \Psi \rangle \neq 0 \quad \text{for } X \in \mathscr{L}_2, \quad X \notin \mathscr{L}_1 \tag{3.16}$$

E is said to be first order unstable. Second order instability is realized if the Brillouin condition holds for all X in $\mathscr{L}_2$, but if E is no longer a minimum, but a saddle point or some other stationary point, i.e. if the Hessean in terms of $\mathscr{L}_2$ has at least one negative eigenvalue (but not that in terms of $\mathscr{L}_1$).

A typical case is that of *Hartree–Fock instability* [24]. Take a closed-shell state with Φ a closed-shell Slater determinant and choose $\mathscr{L}_1$ as the Lie algebra with basis operators defined by Eq. (3.12b) and $\mathscr{L}_2$ with basis operators of Eq. (3.12a). Let the Brillouin condition of Eq. (2.5) be satisfied for all $X \in \mathscr{L}_1$. Then Eq. (2.5) will automatically be satisfied for all $X \in \mathscr{L}_2$, because the extra operators in $\mathscr{L}_2$ not contained in $\mathscr{L}_1$ are of the form:

$$a_{Q\alpha}^{P\alpha} - a_{Q\beta}^{P\beta} \tag{3.17}$$

These are not spin-conserving, but the RHF wave function Φ is a pure-spin state and the Brillouin matrix elements vanish for symmetry reasons. So 1st order Hartree–Fock instability is not possible for a closed-shell state, but 2nd order instability may occur, with the effect that a spin-contaminated Slater-determinant wave function Φ has a lower energy than a pure spin Φ. Since an energy lowering is achieved by a reduction of the symmetry one refers to this type of Hartree–Fock instabilities also as symmetry breaking.

One may further extend the Lie algebra of operators in Eq. (3.12a) by including double creation and double annihilation operators:

$$a^p a^q; \qquad a_p a_q \tag{3.18}$$

One has then still a Lie algebra [14], but the corresponding Lie group is not particle-number-conserving. The Lie group generated by the operators in Eq. (3.12a) is known as $U(n)$, the unitary group of dimension n, the Lie algebra augmented by the operators (3.18) generates $SO(2n)$ the special orthogonal group of dimension $2n$. The Brillouin conditions of Eq. (2.5) with operators of type. (3.18) vanish for particle-number non-conserving wave functions. So again there is no first-order instability. Second-order particle-number instability, i.e. energy lowering for a non-particle-number conserving wave function is supposed to be an essential feature of superconductivity ground states.

4. The perturbation expansion

We now consider a Hamiltonian $H(\lambda)$ that depends parametrically on λ. The expectation value to be made stationary is:

$$E(\lambda) = \langle \Psi(\lambda) | H(\lambda) | \Psi(\lambda) \rangle; \qquad \langle \Psi(\lambda) | \Psi(\lambda) \rangle = 1 \tag{4.1a}$$

$$H(\lambda) = H_0 + \lambda H_1 + \lambda^2 H_2 + \cdots \tag{4.1b}$$

The expansion in powers of λ can be finite (a polynomial) or infinite (a power series), but in the latter case some non-trivial problems concerning the analyticity of H may arise. We assume here that $H(\lambda)$ is a polynomial, without specifying its degree.

We choose the variational group so that the dependence of Ψ on λ can appropriately be described as [11, 12] (this is, in a way, the *key feature of stationary perturbation theory*).

$$\Psi(\lambda) = e^{Y}\Psi_0; \qquad Y(\lambda) \in \mathscr{L}_r; \qquad \Psi_0 = \Psi(0) \tag{4.2}$$

In the philosophy of perturbation theory we make the ansatz:

$$Y(\lambda) = \sum_{k=1}^{\infty} \lambda^k Y_k \tag{4.3}$$

and we expand the expectation value of Eq. (4.1) with Ψ given by Eqs. (4.2, 3) in powers of λ:

$$E(\lambda) = \sum_{k=0}^{\infty} \lambda^k E_k \tag{4.4}$$

$$E_0 = \langle \Psi_0 | H_0 | \Psi_0 \rangle \tag{4.5a}$$

$$E_1 = \langle \Psi_0 | [H_0, Y_1] + H_1 | \Psi_0 \rangle \tag{4.5b}$$

$$E_2 = \langle \Psi_0 | [H_0, Y_2] + \tfrac{1}{2}[[H_0, Y_1], Y_1] + [H_1, Y_1] + H_2 | \Psi_0 \rangle \tag{4.5c}$$

$$E_3 = \langle \Psi_0 | [H_0, Y_3] + \tfrac{1}{2}[[H_0, Y_1], Y_2] + \tfrac{1}{2}[[H_0, Y_2], Y_1] + \tfrac{1}{6}[[[H_0, Y_1], Y_1], Y_1] + [H_1, Y_2] + \tfrac{1}{2}[[H_1, Y_1], Y_1] + [H_2, Y_1] + H_3 | \Psi_0 \rangle \tag{4.5d}$$

We now require that the Brillouin condition of Eq. (2.5) holds for all λ, at least for λ sufficiently close to 0. We hence insert Eqs. (4.2) and (4.3) into Eq. (2.5) and collect powers of λ with the results:

$$C_0 \equiv \langle \Psi_0 | [H_0, X] | \Psi_0 \rangle = 0 \tag{4.6a}$$

$$C_1 \equiv \langle \Psi_0 | [[H_0, X], Y_1] + [H_1, X] | \Psi_0 \rangle = 0 \tag{4.6b}$$

$$C_2 \equiv \langle \Psi_0 | [[H_0, X], Y_2] + \tfrac{1}{2}[[[H_0, X], Y_1], Y_1] + [[H_1, X], Y_1] + [H_2, X] | \Psi_0 \rangle = 0 \tag{4.6c}$$

for all $X \in \mathscr{L}_c$.

Since by assumption $Y \in \mathscr{L}_c$, we can if Eq. (4.6a) holds, i.e. if the stationary condition is satisfied for the uperturbed problem, omit all contributions $\langle \Psi_0 | [H_0, Y_k] | \Psi_0 \rangle$ in Eq. (4.5b,c,d), in particular replace Eq. (4.5b) by:

$$E_1 = \langle \Psi_0 | H_1 | \Psi_0 \rangle \tag{4.7}$$

which is, of course, the *Hellmann–Feynman theorem* in the framework of stationary perturbation theory [9].

If Eq. (4.6a) holds, Eq. (4.6b) can, by means of the Jacobi identity for double commutators and noting that $[X, Y_1] \in \mathscr{L}_r$ be rewritten in the following equivalent forms:

$$\langle \Psi_0 | [H_1, X] + [[H_0, Y_1], X] | \Psi_0 \rangle = 0 \tag{4.8a}$$

$$\langle \Psi_0 | [X_1, X] + \tfrac{1}{2}[[H_0, X], Y_1] + \tfrac{1}{2}[[H_0, Y_1], X] | \Psi_0 \rangle = 0 \tag{4.8b}$$

Equation (4.8b) is easily seen to be the condition that E_2 as given by Eq. (4.5c) is stationary with respect to variation of Y_1, or that the following functional:

$$F(Y_1) = \langle \Psi_0 | [H_1, Y_1] + \tfrac{1}{2}[[H_0, Y_1], Y_1] | \Psi_0 \rangle; \qquad \delta F(Y_1) = 0 \tag{4.9}$$

is stationary with respect to variation of Y_1. We call Eq. (4.9) the *Lie-algebraic Hylleraas functional* (the relation to the conventional form of the latter is pointed out later in this section).

If Eqs. (4.6a) and (4.6b) or Eq. (4.8) is satisfied, the expressions of Eqs. (4.5c,d) for E_2 and E_3 are simplified to:

$$E_2 = \tfrac{1}{2}\langle \Psi_0 | [H_1, Y_1] | \Psi_0 \rangle + \langle \Psi_0 | H_2 | \Psi_0 \rangle \tag{4.10}$$

$$E_3 = \tfrac{1}{6}\langle \Psi_0 | [[[H_0, Y_1], Y_1], Y_1] | \Psi_0 \rangle + \tfrac{1}{2}\langle \Psi_0 | [[H_1, Y_1], Y_1] | \Psi_0 \rangle + \langle \Psi_0 | [H_2, Y_1] + H_3 | \Psi_0 \rangle \tag{4.11}$$

In order to evaluate E_2 and E_3 we only need to know Y_1. (This is a special case of Wigner's $[2n + 1]$-rule.)

If we know the Hessean of Eq. (2.11b) for the unperturbed Hamiltonian H_0, we can construct Y_1 by solution of a linear system of equations. Let us expand Y in $\mathscr{L}_r$ and define the matrix elements V_k:

$$Y_1 = \sum_k c_k X_k; \qquad V_k = \langle \Psi_0 | [H_1, X_k] | \Psi_0 \rangle; \qquad X_k \in \mathscr{L}_r \tag{4.12}$$

then Eq. (4.6b) becomes for $X = X_l$:

$$V_l + \sum_k H^0_{lk} c_k = 0 \tag{4.13a}$$

$$H^0_{lk} = \langle \Psi_0 | [[H_0, X_l], X_k] | \Psi_0 \rangle \tag{4.13b}$$

As already mentioned previously (Sect. 2) a unique solution of a system like Eq. (4.13a) requires that $\underline{\underline{H}}^0$ is non-singular, i.e. that redundancies are removed, as to be indicated in Sect. 6.

To see the relation between Eq. (4.9) and the Hylleraas variation principle, we consider the variational group corresponding to a finite-dimensional Hilbert space with the basis of Eq. (3.7) of the Lie algebra. Choosing the ϕ_k such that they diagonalize $\langle \phi_k | H_0 | \phi_l \rangle$, defining:

$$Y_1 | \Psi_0 \rangle = | \Psi_1 \rangle \tag{4.14}$$

and noting the antihermiticity of Y_1 we get from Eq. (4.9):

$$F(\Psi_1) = 2Re \langle \Psi_0 | V | \Psi_1 \rangle + \langle \Psi_1 | H_0 - E_0 | \Psi_1 \rangle \tag{4.15}$$

which is, in fact, the conventional Hylleraas functional.

If one uses the formalism of this section for the unitary group of one-particle transformations generated by the operators of Eqs. (3.12a) or (3.12b) one is automatically led to coupled-Hartree–Fock theory.

5. Multiple perturbations

We now consider a Hamiltonian that depends on more than one perturbation parameter. For the sake of simplicity we choose:

$$H(\lambda, \mu) = H_0 + \lambda V_{10} + \mu V_{01} \tag{5.1}$$

The counterpart of Eq. (4.1a) is then:

$$E(\lambda, \mu) = \langle \Psi | H_0 + \lambda V_{10} + \mu V_{01} | \Psi \rangle \tag{5.2}$$

and in analogy to Eqs. (4.2, 3) we have:

$$\Psi(\lambda, \mu) = e^Y \Psi_0 \tag{5.3a}$$

$$Y = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \lambda^k \mu^l Y_{kl}; \qquad k = l = 0 \text{ excluded} \tag{5.3b}$$

Instead of Eq. (4.6) we then get (we write 0 for 00):

$$C_0 \equiv \langle \Psi_0 | [H_0, X] | \Psi_0 \rangle = 0 \tag{5.4a}$$

$$C_{10} \equiv \langle \Psi_0 | [[H_0, X], Y_{10}] + [V_{10}, X] | \Psi_0 \rangle = 0 \tag{5.4b}$$

$$C_{01} \equiv \langle \Psi_0 | [[X_0, X], Y_{01}] + [V_{01}, X] | \Psi_0 \rangle = 0 \tag{5.4c}$$

$$C_{11} \equiv \langle \Psi_0 | [[H_0, X], Y_{11}] + \tfrac{1}{2}[[[H_0, X], Y_{01}], Y_{10}] + \tfrac{1}{2}[[[H_0, X], Y_{10}], Y_{01}] + [[V_{10}, X], Y_{01}] + [[V_{01}, X], Y_{10}] | \Psi_0 > = 0 \tag{5.4d}$$

$$C_{20} \equiv \langle \Psi_0 | [[H_0, X], Y_{20}] + \tfrac{1}{2}[[[H_0, X], Y_{10}], Y_{10}] + [[V_{10}, X], Y_{10}] | \Psi_0 \rangle = 0 \tag{5.4e}$$

$$C_{02} \equiv \langle \Psi_0 | [[H_0, X], Y_{02}] + \tfrac{1}{2}[[[H_0, X], Y_{01}], Y_{01}] + [[V_{01}, X], Y_{01}] | \Psi_0 \rangle = 0 \tag{5.4f}$$

The counterpart of Eqs. (4.4, 5) is:

$$E(\lambda, \mu) = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \lambda^k \mu^l E_{kl} \tag{5.5}$$

$$E_0 = \langle \Psi_0 | H_0 | \Psi_0 \rangle \tag{5.6a}$$

$$E_{10} = \langle \Psi_0 | [H_0, Y_{10}] + V_{10} | \Psi_0 \rangle \tag{5.6b}$$

$$E_{01} = \langle \Psi_0 | [H_0, Y_{01}] + V_{01} | \Psi_0 \rangle \tag{5.6c}$$

$$E_{20} = \langle \Psi_0 | [H_0, Y_{20}] + \tfrac{1}{2}[[H_0, Y_{10}], Y_{10}] + [V_{10}, Y_{10}] | \Psi_0 \rangle \tag{5.6d}$$

$$E_{02} = \langle \Psi_0 | [H_0, Y_{02}] + \tfrac{1}{2}[[H_0, Y_{01}], Y_{01}] + [V_{01}, Y_{01}] | \Psi_0 \rangle \tag{5.6e}$$

$$E_{11} = \langle \Psi_0 | [H_0, Y_{11}] + \tfrac{1}{2}[[H_0, Y_{10}], Y_{01}] + \tfrac{1}{2}[[H_0, Y_{01}], Y_{10}] + [V_{10}, Y_{01}] + [V_{01}, Y_{10}] | \Psi_0 \rangle \tag{5.6f}$$

We don't write expressions for E_{21}, E_{12} and E_{22} that are very lengthy. We can simplify Eqs. (5.6), provided that Eqs. (5.4) are satisfied. In analogy to Eq. (4.7) we get:

$$E_{10} = \langle \Psi_0 | V_{10} | \Psi_0 \rangle \tag{5.7a}$$

$$E_{01} = \langle \Psi_0 | V_{01} | \Psi_0 \rangle \tag{5.7b}$$

The simplifications of the type of Eq. (4.10) are:

$$E_{20} = \tfrac{1}{2} \langle \Psi_0 | [V_{10}, Y_{10}] | \Psi_0 \rangle \tag{5.8a}$$

$$E_{02} = \tfrac{1}{2} \langle \Psi_0 | [V_{01}, Y_{01}] | \Psi_0 \rangle \tag{5.8b}$$

$$E_{11} = \tfrac{1}{2} \langle \Psi_0 | [V_{10}, Y_{01}] + [V_{01}, Y_{10}] | \Psi_0 \rangle \tag{5.8c}$$

$$E_{21} = \langle \Psi_0 | \tfrac{1}{2}[[V_{10}, Y_{10}], Y_{01}] + \tfrac{1}{2}[[V_{01}, Y_{10}], Y_{10}] + \tfrac{1}{2}[[V_{10}, Y_{01}], Y_{10}] + \tfrac{1}{6}[[[H_0, Y_{10}], Y_{01}], Y_{10}] + \tfrac{1}{6}[[[H_0, Y_{10}], Y_{10}], Y_{01}] + \tfrac{1}{6}[[[H_0, Y_{01}], Y_{10}], Y_{10}] | \Psi_0 \rangle \tag{5.8d}$$

E_{12} is like E_{21} just with V_{10} exchanged with V_{01} and Y_{10} with Y_{01}.

Using Eqs. (5.4b,c) we can rewrite Eq. (5.8c) as:

$$E_{11} = -\tfrac{1}{2} \langle \Psi_0 | [[H_0, Y_{10}], Y_{01}] + [[H_0, Y_{01}], Y_{10}] | \Psi_0 \rangle \tag{5.9}$$

which can by means of the Jacobi identity, the Lie-algebraic property and Eq. (5.4a) be written as:

$$E_{11} = -\langle \Psi_0 | [[H_0, Y_{10}], Y_{01}] | \Psi_0 \rangle = -\langle \Psi_0 | [[H_0, Y_{01}], Y_{10}] | \Psi_0 \rangle \tag{5.10}$$

Using again Eqs. (5.4b,c) we get finally:

$$E_{11} = \langle \Psi_0 | [V_{10}, Y_{01}] | \Psi_0 \rangle = \langle \Psi_0 | [V_{01}, Y_{10}] | \Psi_0 \rangle \tag{5.11}$$

which is the equivalent of the exchange theorem of traditional perturbation theory [25].

In a similar way, making extensive use of the Jacobi identity of Eq. (5.8d) can be reformulated to:

$$\begin{aligned} E_{21} &= \tfrac{1}{2}\langle\Psi_0|[[V_{10}, Y_{10}], Y_{01}] + [[V_{10}, Y_{01}], Y_{10}] + [[V_{01}, Y_{10}], Y_{10}] \\ &\quad + [[[H_0, Y_{10}], Y_{01}], Y_{10}]|\Psi_0\rangle \\ &= \tfrac{1}{2}\langle\Psi_0|2[[V_{10}, Y_{10}], Y_{01}] + [[V_{01}, Y_{10}], Y_{10}] + [[[H_0, Y_{10}], Y_{10}], Y_{01}]|\Psi_0\rangle \\ &= \tfrac{1}{2}\langle\Psi_0|2[[V_{10}, Y_{01}], Y_{10}] + [[V_{01}, Y_{10}], Y_{10}] + [[[H_0, Y_{01}], Y_{10}], Y_{10}]|\Psi_0\rangle \end{aligned} \tag{5.12}$$

The first of these three expressions is most symmetric, the other two are somewhat simpler. For the evaluation of E_{21} only Y_{10} and Y_{01} are needed.

The expression for E_{12} is again obtained exchanging the two subscripts.

After a somewhat lengthy reformulation one gets the following result for E_{22} in terms of Y_{10}, Y_{01}, Y_{02} and Y_{11} (but not Y_{20}).

$$\begin{aligned} E_{22} &= \tfrac{1}{2}\langle\Psi_0|[[V_{10}, Y_{11}], Y_{01}] + 2[[V_{10}, Y_{10}], Y_{02}]\rangle + [[V_{01}, Y_{10}], Y_{11}] \\ &\quad + [[[V_{10}, Y_{10}], Y_{01}], Y_{01}] + \tfrac{1}{2}[[[V_{01}, Y_{10}], Y_{01}], Y_{10}] + \tfrac{1}{2}[[[V_{01}, Y_{10}], Y_{10}], Y_{01}] \\ &\quad + \tfrac{1}{2}[[[H_0, Y_{10}], Y_{02}], Y_{10}] + \tfrac{1}{2}[[[H_0, Y_{10}], Y_{10}], Y_{02}] + \tfrac{1}{2}[[[H_0, Y_{10}], Y_{11}], Y_{01}] \\ &\quad + \tfrac{1}{2}[[[H_0, Y_{10}], Y_{01}], Y_{11}] + \tfrac{1}{2}[[[[H_0, Y_{10}], Y_{01}], Y_{10}], Y_{01}]|\Psi_0\rangle \end{aligned} \tag{5.13}$$

An alternative expression is obtained if one exchanges everywhere the first with the second subscript. It then involves Y_{20} and Y_{11}, but not Y_{02}. Equation (5.13) may look lengthy, but the corresponding expression is conventional double perturbation theory is hardly more compact.

6. The spectrum of the Hessean and generalized RPA

6.1. The Hessean in a basis of the complex Lie algebra $\mathscr{L}_c$

In principle, the Lie-algebraic formulation of the variation principle as well as of stationary perturbation theory is possible in terms of the real Lie algebra $\mathscr{L}_r$ of antihermitean operators only. It is, nevertheless, often convenient to refer to the full complex Lie algebra $\mathscr{L}_c$ and to admit other than antihermitean operators, because a basis of $\mathscr{L}_c$ is often formally simpler, and also because one gains additional insight into the structure of the theory. We have already seen that it makes no difference for the generalized Brillouin condition of Eq. (2.5) whether we formulate it in terms of $X \in \mathscr{L}_r$ or $X \in \mathscr{L}_c$. However, the original definition of Eq. (2.11b) of the Hessean as well as its specification in Eq. (4.13b) for the unperturbed Hamiltonian is only valid for anithermitean X. A generalized definition is possible for arbitrary X, that reduces to Eq. (4.13b) for anithermitean X, namely:

$$H_{kl} = \langle\Psi_0|[X_k^\dagger, [H_0, X_l]]|\Psi_0\rangle \tag{6.1}$$

We henceforth omit the superscript 0 on H_{kl} to indicate that it refers to the unperturbed Hamiltonian H_0. One is, of course, free how to choose H_0 and the results of this section are independent of this choice.

By means of the Jacobi identity, the Lie-algebraic property and the stationarity condition we can show that an alternative expression equivalent to Eq. (6.1) is:

$$H_{kl} = -\langle \Psi_0 | [H_0, [X_l, X_k^\dagger]] | \Psi_0 \rangle - \langle \Psi_0 | [X_l, [X_k^\dagger, H_0]] | \Psi_0 \rangle$$
$$= \langle \Psi_0 | [X_l, [H_0, X_k^\dagger]] | \Psi_0 \rangle \tag{6.2}$$

From Eqs. (6.1) and (6.2) one sees immediately that $\underline{\underline{H}}$ is a hermitean matrix, i.e. $H_{kl} = H_{lk}^*$. For the special case that $X_k^\dagger = -X_k$, $\underline{\underline{H}}$ is even real symmetric.

If we define the superoperator $\mathscr{H}$ of commutation with H_0:

$$\mathscr{H} X = [H_0, X] \tag{6.3}$$

the matrix elements of the Hessean can also be written as:

$$H_{kl} = \langle \Psi_0 | [X_k^\dagger, \mathscr{H} X_l] | \Psi_0 \rangle = \langle \Psi_0 | [X_l, \mathscr{H} X_k^\dagger] | \Psi_0 \rangle \tag{6.4}$$

As a hermitean matrix $\underline{\underline{H}}$ has only real eigenvalues and an orthonormal set of eigenvectors. The rank of $\underline{\underline{H}}$ is equal to the number of its non-zero eigenvalues.

In perturbation theory in order to construct E_2 and E_3 we have first to evaluate Y_1 from Eq. (4.8a) (or the equivalent expressions of Eqs. (4.6a), (4.8b)). While in Sect. 4 we have expanded Y_1 in a basis of $\mathscr{L}_r$ with the result of Eq. (4.13), we now expand it in a basis $\{X_k\}$ of $\mathscr{L}_c$.

$$Y_1 = \sum_k b_k X_k; \qquad V_k = \langle \Psi_0 | [H_1, X_k] | \Psi_0 \rangle; \qquad X_k \in \mathscr{L}_c \tag{6.5}$$

such that:

$$\sum_k H_{lk} b_k + V_l^* = 0 \tag{6.6}$$

6.2. Eigenvalues, regularization, and inverse of the Hessean

The linear system of Eqs. (4.13a) or (6.6) has unique solutions only if $\underline{\underline{H}}$ is regular, i.e. has no zero eigenvalues. What we need is hence not the matrix $\underline{\underline{H}}$ in terms of a complete basis of $\mathscr{L}_r$ (or $\mathscr{L}_c$), but rather in terms of a *non-redundant* set. Formally this can be achieved by solving the eigenvalue problem:

$$\sum_l H_{kl} c_l^{(\mu)} = \omega_\mu c_k^{(\mu)} \tag{6.7}$$

and discarding all eigenvectors with zero eigenvalues. In practice a non-redundant set of basis operators X_k is often found by inspection, so that a regularization via solution of Eq. (6.7) is not necessary. In the Hartree–Fock case for a closed-shell state a basis of $\mathscr{L}_c$ are the operators a_q^p (see Eq. (3.12a)) in terms of a given orthogonal one-electron basis χ_p. Let the spin-orbitals χ_i be occupied in Ψ_0 and the χ_a unoccupied (virtual), then a non-redundant operator basis (with respect to Ψ_0) consists of the operators a_i^a and a_a^i.

We come back to the regularization of Eq. (6.6) later in this section. We assume now that a basis has been found such that $\underline{\underline{H}}$ is non-singular.

One easily sees that the Y_1 constructed from Eqs. (6.5, 6) is antihermitean. In fact let us alternatively to Eq. (6.5) expand Y_1 in the $X_k^\dagger$. Then from the counterpart of Eq. (6.6) the expansion coefficients turn out to be $-b_k^*$.

In order to construct the first-order perturbation operator Y_1, the solution of the linear system of Eqs. (6.6) is the method of choice if only one perturbation

is considered. However, if one wants to study several perturbations for the same unperturbed problem, other approaches may be preferable. One of these is to first invert $\underline{\underline{H}}$ (of course, after having eliminated all redundancies in the operator basis).

$$\underline{\underline{G}} = \underline{\underline{H}}^{-1} \tag{6.8}$$

One then gets (provided that $H_2 = 0$, cf. Eq. (4.10)):

$$Y_1 = \sum_k b_k X_k = -\sum_{k,l} X_k G_{kl} V_l^* \tag{6.9a}$$

$$E_2 = \tfrac{1}{2} \sum_k b_k V_k = -\tfrac{1}{2} \sum_{k,l} V_k G_{kl} V_l^* \tag{6.9b}$$

The inverse of the Hessean obviously plays a similar role as the resolvent in the conventional formulation of perturbation theory. For a unique definition of $\underline{\underline{G}}$ one must, however, worry about the normalization of operators. Imagine, e.g., that all basis operators are multiplied by a common scalar λ. Then $\underline{\underline{H}}$ is multiplied by λ^2, $\underline{\underline{G}}$ by λ^{-2}, the V_l by λ, while Y_1 and E_2 remain invariant. We come back to this point below.

It is, of course, possible to construct the inverse of $\underline{\underline{H}}$ in the spectral representation. Let us define the operators:

$$\Omega_\mu = \sum_k c_k^{(\mu)} X_k; \qquad X_k \in \mathscr{L}_c \tag{6.10}$$

with the $c_k^{(\mu)}$ satisfying the eigenvalue Eq. (6.7) and the eigenvectors normalized to 1. Then we get:

$$\tilde{H}_{\mu\nu} = \langle \Psi_0 | [\Omega_\mu^\dagger, [H_0, \Omega_\nu]] | \Psi_0 \rangle = \omega_\mu \delta_{\mu\nu} \tag{6.11}$$

and $\underline{\underline{\tilde{G}}}$ is a diagonal operator with diagonal elements ω_μ^{-1}. We can then expand Y_1 in the basis of the Ω_μ (for $\omega_\mu \neq 0$) and get:

$$Y_1 = -\sum_\mu d_\mu \Omega_\mu = -\sum_\mu \omega_\mu^{-1} \tilde{V}_\mu^* \Omega_\mu \tag{6.12a}$$

$$E_2 = \tfrac{1}{2} \sum_\mu d_\mu \tilde{V}_\mu = -\sum_\mu \tilde{V}_\mu^* \omega_\mu^{-1} \tilde{V}_\mu \tag{6.12b}$$

$$\tilde{V}_\mu = \langle \Psi_0 | [H_1, \Omega_\mu] | \Psi_0 \rangle \tag{6.13}$$

This reminds us of the sum-over-state expressions of traditional perturbation theory and it suggests that the eigenvalues ω_μ of the Hessean may be interpreted as transition energies. However, the same reservation must be made as with respect to the inverse $\underline{\underline{G}}$ of $\underline{\underline{H}}$. Imagine that all basis operators are multiplied by λ, then ω_μ is multiplied by λ^2, Ω_μ and V_μ by λ, but Y_1 and E_2 remain unchanged. In order to identify the ω_μ with transition energies we must be able to normalize operators, which requires that we have first to define *a scalar product in operator space.*

Before we study this problem of the normalization of operators, a remark concerning the regularization of the system of Eq. (6.6) by elimination of eigenoperators Ω_μ with $\omega_\mu = 0$ is in order. Two cases are possible. In case (**a**) the matrix elements $\tilde{V}_\mu$ vanish for all Ω_μ with $\omega_\mu = 0$. Then the solutions Y_1 of the regularized and non-regularized system of Eq. (6.6) differ only in an arbitrary solution of the homogeneous system $\underline{\underline{H}}\vec{c} = 0$, which has no effect on E_2. The regularization only serves to make the solution unique, it corresponds to a specific normalization of the wave operator in conventional perturbation theory

[16]. In case (**b**) some $\tilde{V}_\mu$ for $\omega_\mu = 0$ do not vanish. Then the non-regularized system of Eq. (6.6) has no solution at all and the solution of the regularized system may not be meaningful. This case is, e.g., realized if Ψ_0 is degenerate with some $\Omega_\mu \Psi_0$. Then a generalization of stationarity perturbation theory to degenerate states is necessary. We come to this in Sect. 10.

6.3. A metric in operator space. Model transition operators and transition energies

Note that the following considerations are only relevant if we want to define *model excitations* within the framework of the variational group. For the evaluation of Y_1 or E_2 it is perfectly straightforward to diagonalize $\underline{\underline{H}}$ in the naive sense and to use Eqs. (6.12–13) without worrying about normalization.

Assume for a moment that Π_μ is an exact excitation operator. It then satisfies:

$$\mathscr{H}\Pi_\mu = [H_0, \Pi_\mu] = \varepsilon_\mu \Pi_\mu \tag{6.14}$$

where ε_μ is the corresponding exact transition energy. Then:

$$\langle \Psi_0 | [\Pi_\nu^\dagger, [H_0, \Pi_\mu]] | \Psi_0 \rangle = \varepsilon_\mu \langle \Psi_0 | [\Pi_\nu^\dagger, \Pi_\mu] | \Psi_0 \rangle \tag{6.15}$$

The expression on the left side is the matrix element of the Hessean in terms of the Π_μ. This suggests to regard:

$$\Delta_{kl} = \langle \Psi_0 | [X_k^\dagger, X_l] | \Psi_0 \rangle \tag{6.16}$$

as the desired scalar product in operator space. Unfortunately, this Δ_{kl} is not a scalar product in the strict sense since it is not positive-definite. Such products are sometimes referred to as 'binary products'. One might define a genuine scalar product as $\langle \Psi_0 | X_k^\dagger X_l | \Psi_0 \rangle$, but this would involve an ordinary product of operators rather than a Lie product and would hence not be consistent with the Lie-algebraic structure of the theory. With the binary product (6.16) we can now replace the eigenvalue Eq. (6.7) by:

$$\sum_l H_{kl} a_l^{(\mu)} = \varepsilon_\mu \sum_l \Delta_{kl} a_l^{(\mu)} \tag{6.17}$$

The eigenvalues and eigenvectors will no longer be the same as in Eq. (6.7), so we have changed their symbols.

The system of Eq. (6.17) is reminiscent of the random-phase approximation (RPA) and it represents in fact its straightforward generalization to arbitrary variational groups. We note that *any variational group*, which has originally been introduced in view of minimizing an expectation value, *automatically generates a model excitation spectrum* for the respective state.

The eigenvectors $\vec{a}^{(\mu)}$ are orthogonal in the sense:

$$\sum_{k,l} a_k^{(\mu)} \Delta_{kl} a_l^{(\nu)} = 0 \quad \text{for } \varepsilon_\mu \neq \varepsilon_\nu^* \tag{6.18}$$

and we can orthonormalize them so that:

$$\sum_{k,l} a_k^{(\mu)*} \Delta_{kl} a_l^{(\nu)} = \pm \delta_{\mu\nu} \tag{6.19}$$

where some of the vectors are 'normalized' to $+1$, the others to -1. We define the model excitation operators (changing now the meaning of Π_μ with respect to

Eqs. (6.14, 15) from *exact to model* excitation operators):

$$\Pi_\mu = \sum_k a_k^{(\mu)} X_k \tag{6.20}$$

and these are orthogonal with respect to the metric characterized by $\underline{\underline{\Delta}}$ and normalized to either $+1$ or -1. For the matrix elements of the Hessean in terms of the Π_μ we get:

$$\begin{aligned}\tilde{\tilde{H}}_{\mu\nu} &= \langle \Psi_0 | [[\Pi_\mu^\dagger, [H_0, \Pi_\nu]] | \Psi_0 \rangle = \varepsilon_\mu \langle \Psi_0 | [\Pi_\mu^\dagger, \Pi_\nu] | \Psi_0 \rangle \\ &= \varepsilon_\mu \tilde{\tilde{\Delta}}_{\mu\nu} = \tilde{\tilde{\Delta}}_{\mu\mu} \varepsilon_\mu \delta_{\mu\nu} = \pm \varepsilon_\mu \delta_{\mu\nu}\end{aligned} \tag{6.21}$$

We can, in analogy to Eqs. (6.12–13) expand Y_1 in the Π_μ (for $\varepsilon_\mu \neq 0$):

$$Y_1 = \sum_\mu f_\mu \Pi_\mu = -\sum_\mu \varepsilon_\mu^{-1} \tilde{\tilde{\Delta}}_{\mu\mu} \tilde{\tilde{V}}_\mu \Pi_\mu \tag{6.22a}$$

$$E_2 = \tfrac{1}{2} \sum_\mu f_\mu \tilde{\tilde{V}}_\mu = -\tfrac{1}{2} \sum_\mu \tilde{\tilde{\Delta}}_{\mu\mu} \tilde{\tilde{V}}_\mu^* \varepsilon_\mu^{-1} \tilde{\tilde{V}}_\mu \tag{6.22b}$$

$$\tilde{\tilde{V}}_\mu = \langle \Psi_0 | [H_1, \Pi_\mu] | \Psi_0 \rangle \tag{6.22c}$$

A few comments on the eigenvalue system of Eq. (6.17) are in order. While the 'eigenoperators' of the original Hessean are either hermitean or antihermitean (i.e. $\underline{\underline{H}}$ factorizes into noninteracting blocks corresponding to hermitean and antihermitean X_k) this is no longer the case for the eigenvectors in the sense of Eq. (6.17). These are neither hermitean nor antihermitean. Pure hermitean or antihermitean operators have even zero norm with the metric of Eq. (6.16).

On the other hand, if Π_μ is an eigenoperator to the eigenvalue ε_μ, then also $\Pi_\mu^\dagger$ is an eigenoperator, but with eigenvalue $-\varepsilon_\mu^*$. This is, e.g., seen if one writes the eigenvalue Eq. (6.17) in the form:

$$\langle \Psi_0 | [X^\dagger, [H_0, \Pi_\mu]] | \Psi_0 \rangle = \varepsilon_\mu \langle \Psi_0 | [X^\dagger, \Pi_\mu] | \Psi_0 \rangle \tag{6.23a}$$

and takes the complex conjugate:

$$\langle \Psi_0 | [X, [H_0, \Pi_\mu^\dagger]] | \Psi_0 \rangle = -\varepsilon_\mu^* \langle \Psi_0 | [X, \Pi_\mu^\dagger] | \Psi_0 \rangle \tag{6.23b}$$

The Y_1 constructed from Eq. (6.22a) is antihermitean, which means that alternatively to Eq. (6.22a) we can write:

$$Y_1 = -\sum_\mu f_\mu^* \Pi_\mu^\dagger \tag{6.23c}$$

Of course, eigenoperators with excitation energy $\varepsilon_\mu = 0$ in Eq. (6.17) must be discarded. It is not obvious that a non-redundant operator basis in the 'naive' metric and the metric defined by $\underline{\underline{\Delta}}$ should necessarily be identical, i.e. that the operator spaces corresponding to eigenvalues $\omega_\mu = 0$ of Eq. (6.5) or $\varepsilon_\mu = 0$ of Eq. (6.17) will generally agree. In those cases when a non-redundant basis is found by simple inspection this will probably be the case, e.g. for the example mentioned after Eq. (6.7).

Generally one cannot exclude that some eigenvalues ε_μ are complex. This indicates that the chosen variational group does not necessarily generate a physically reasonable model excitation spectrum, but this does not invalidate the construction of Y_1 or E_2.

The model excitation energies ε_μ introduced in a somewhat ad-hoc way as eigenvalues of the system of Eq. (6.17) will get a more direct physical meaning in the context of time-dependent theory (Sect. 9).

Let us now consider a somewhat restricted case. Assume that a non-redundant basis of $\mathscr{L}_c$ can be found which consists of two subsets such that the elements X_μ of one set have the property (which we call the *excitation–deexcitation separation* condition):

$$X_\mu^\dagger|\Psi_0\rangle = 0; \qquad X_\mu|\Psi_0\rangle \neq 0 \tag{6.24}$$

and that the elements of the other set are just the hermitean conjugates $Z_\mu = X_\mu^\dagger$. Then this condition (Eq. (6.24)) guarantees that one can uniquely define a basis of excitation and deexcitation operators with respect to Ψ_0. It does not guarantee – as we shall see soon – that a condition analogous to Eq. (6.24) also holds for the eigenoperators Π_μ. Equation (6.24) does, e.g., hold in coupled-Hartree–Fock theory for Ψ_0 a single Slater determinant, but not necessarily in coupled MC-SCF theory. If Eq. (6.24) holds, $\underline{\underline{H}}$ and $\underline{\underline{\Delta}}$ have the following block structure:

$$\underline{\underline{H}} = \begin{pmatrix} A & B \\ B^\dagger & A^\dagger \end{pmatrix} \qquad \underline{\underline{\Delta}} = \begin{pmatrix} S & 0 \\ 0 & -S^\dagger \end{pmatrix} \tag{6.25}$$

$$\langle\Psi_0|[X_\mu^\dagger, X_\nu]|\Psi_0\rangle = \langle\Psi_0|X_\mu^\dagger X_\nu|\Psi_0\rangle = S_{\mu\nu} \tag{6.26a}$$

$$\langle\Psi_0|[X_\mu^\dagger, Z_\nu]|\Psi_0\rangle = \langle\Psi_0|[Z_\mu^\dagger, X_\nu]|\Psi_0\rangle = 0 \tag{6.26b}$$

$$\langle\Psi_0|[Z_\mu^\dagger, Z_\nu]|\Psi_0\rangle = -\langle\Psi_0|X_\nu^\dagger X_\mu|\Psi_0\rangle = -S_{\nu\mu} \tag{6.26c}$$

$$\langle\Psi_0|[X_\mu^\dagger, [H_0, X_\nu]]|\Psi_0\rangle = A_{\mu\nu} \tag{6.27a}$$

$$\langle\Psi_0|[X_\mu^\dagger, [H_0, Z_\nu]]|\Psi_0\rangle = B_{\mu\nu} \tag{6.27b}$$

In this case the eigenoperators Π_μ are linear combinations of both the basis operators X_k and the $Z_k = X_k^\dagger$. As a consequence, one gets:

$$\Pi_\mu^\dagger|\Psi_0\rangle \neq 0; \qquad \Pi_\mu|\Psi_0\rangle \neq 0 \tag{6.28}$$

The second of these inequalities is as desired, since action of an excitation operator on Ψ_0 should change this to the wave function of an excited state. The first inequality in Eq. (6.28) is less welcome, because an exact deexcitation operator should annihilate Ψ_0. If $\Pi^\dagger$ fails to deexcite Ψ_0, this is an indication of some inconsistency. Either Π is a too poor approximation to the exact excitation operator or Ψ_0 is a too poor approximation to the exact ground state.

This phenomenon is known from traditional RPA theory where the one-particle excitation operators in Eq. (3.12a) are taken as basis of $\mathscr{L}_c$. A remedy has been proposed with the somewhat awkward name 'killer condition' [26], which consists in changing Ψ_0 so that $\Pi_\mu^\dagger|\Psi_0\rangle = 0$. In the present context this would not make too much sense, but rather violate the main philosophy of stationary perturbation theory.

A still more special case is realized if the B block vanishes. Then we only need to solve the eigenvalue problem:

$$A\vec{c}_\mu = \varepsilon_\mu S\vec{c}_\mu \tag{6.29}$$

in order to get the model excitation energies. The ε_μ are all real and to each ε_μ also $-\varepsilon_\mu$ is an eigenvalue of Eq. (6.17), i.e. we get pairs of corresponding excitation and deexcitation operators and energies. The RPA-inconsistency of Eq. (6.28) then no longer arises.

The appearance of both *model excitation* and *deexcitation* operators and energies is characteristic for the Lie-algebraic formulation of the variational approach, and it even arises for linear variations, where the variational group is

generated by the shift operators of Eq. (3.7). Let us consider a basis $\{\phi_\mu\}$ of $\mathscr{H}_N$ such that the matrix representation of the Hamiltonian is diagonal in this basis, i.e.:

$$\langle\phi_\mu|H_0|\phi_\nu\rangle = E_\mu\delta_{\mu\nu}; \qquad \langle\phi_\mu|\phi_\nu\rangle = \delta_{\mu\nu} \tag{6.30}$$

and let $\Psi_0 = \phi_0$, then a non-redundant set of basis operators consists of:

$$X_{\mu 0} = |\phi_\mu\rangle\langle\phi_0|; \qquad X_{0\mu} = |\phi_0\rangle\langle\phi_\mu| = X_{\mu 0}^\dagger, \qquad \mu \neq 0 \tag{6.31}$$

their number is just $2N-2$, if N is the dimension of the basis in $\mathscr{H}_N$.

For the matrix elements (6.26, 27) we get:

$$\langle\Psi_0|[X_{\mu 0}^\dagger, X_{\nu 0}]|\Psi_0\rangle = \delta_{\mu\nu} \tag{6.32a}$$

$$\langle\Psi_0|[X_{0\mu}^\dagger, X_{0\nu}]|\Psi_0\rangle = -\delta_{\mu\nu} \tag{6.32b}$$

$$\langle\Psi_0|[X_{\mu 0}^\dagger, X_{0\nu}]|\Psi_0\rangle = \langle\Psi_0|[X_{0\mu}^\dagger, X_{\nu 0}]|\Psi_0\rangle = 0 \tag{6.32c}$$

$$\langle\Psi_0|[X_{\mu 0}^\dagger, [H_0, X_{\nu 0}]]|\Psi_0\rangle = \delta_{\mu\nu}(E_\mu - E_0) \tag{6.33a}$$

$$\langle\Psi_0|[X_{0\mu}^\dagger, [H_0, X_{0\nu}]]|\Psi_0\rangle = \delta_{\mu\nu}(E_\mu - E_0) \tag{6.33b}$$

$$\langle\Psi_0|[X_{\mu 0}^\dagger, [H_0, X_{0\nu}]]|\Psi_0\rangle = \langle\Psi_0|[X_{0\mu}^\dagger, [H, X_{\nu 0}]]\Psi_0\rangle = 0 \tag{6.33c}$$

The off-diagonal blocks of $\underline{\underline{H}}$ and $\underline{\underline{\Delta}}$ in Eq. (6.25) vanish. The eigenvalues of the first block are just the model excitation energies $E_\mu - E_0$ with eigenoperators $X_{\nu 0}$ and those of the second block are the deexcitation energies $E_0 - E_\mu$ with eigenoperators $X_{0\nu} = X_{\nu 0}^\dagger$.

Let us stress again, that for the evaluation of second-order properties E_2 by perturbation theory the question how good the 'model excitation energies' generated by the variational group, approximate the exact spectral transition energies, is of little relevance. The two requirements that the variational group leads to a good approximation for E_2 and to a good approximation for the excitation spectrum, are often not even compatible. This has to do with the fact that the model spectrum is necessarily discrete, while the physical spectrum contains a discrete and a continuum part. To take care of the role of both parts is only possible via a model spectrum that has little agreement with the physical discrete spectrum. This has also been known in the traditional formulation of perturbation theory in a linear variational space.

The situation changes somewhat if the perturbation is time-dependent and periodic since then resonance phenomena related to spectral excitations arise, and the model excitation energies ε_μ in the sense of Eq. (6.17) get a very concrete meaning (see Sect. 9).

7. Real and complex Hamiltonians

Often (essentially in the absence of magnetic interactions) the Hamiltonian is real. In this case its eigenfunctions can be chosen real and it is sufficient to require stationarity in the sense of Eq. (2.1) for the subgroup $\mathscr{G}_o$ of unitary transformations U that transform a real Ψ to a real Ψ', i.e. for the corresponding subgroup $\mathscr{G}_o$ of real orthogonal transformations imbedded in $\mathscr{G}$. The generators of $\mathscr{G}_o$ are real antihermitean operators (real in the same sense as the operators $U \in \mathscr{G}_o$ are real, see appendix A). The real Lie algebra $\mathscr{L}_r$ of $\mathscr{G}$ is a direct sum of the Lie algebra $\mathscr{L}_o$ of $\mathscr{G}_o$ and the coset $\mathscr{C}_i$ of imaginary antihermitean

operators (note that $\mathscr{C}_i$ is not a Lie algebra since the commutator of two imaginary operators is real).

For a real H the Brillouin condition of Eq. (2.5) is automatically satisfied for all $X \in \mathscr{L}_i$, since the expectation value of an imaginary hermitean operator vanishes. Hence if Eq. (2.5) holds for all $X \in \mathscr{L}_o$ it holds for all $X \in \mathscr{L}_r$ and all $X \in \mathscr{L}_c$.

If $H = H_0 + \lambda H_1$ and if both H_0 and H_1 are real, stationary perturbation theory as outlined in Sect. 4 can be formulated entirely with reference to $\mathscr{L}_o$ instead of $\mathscr{L}_r$. This also holds if H is a higher-order polynomial in λ with all coefficients real.

The next common case is that H_0 is real and H_1 imaginary, or to include the case of magnetic perturbations, that H_k is real for k even and imaginary for k odd. We see from Eq. (4.5) that then the Y_k must be imaginary for k odd and real for k even. If we are only interested in E_2 and Y_1, we have to solve Eqs. (4.6) or (4.8) for Y_1. Obviously, Y_1 is a linear combination (Eq. (4.12)) of the $X_k \in \mathscr{C}_i$. The V_l, H_{lk} and B_k in Eq. (4.13) are then real.

Note that the Hessean matrix H_{kl} needed in Eq. (4.13) is different for real and imaginary perturbations. In the former case it is defined as Eq. (2.11b) in terms of a basis $X_k \in \mathscr{L}_o$, in the latter case for $X_k \in \mathscr{C}_i$.

The formalism of Sect. 6 requires a basis of the full complex Lie algebra $\mathscr{L}_c$, it is especially not possible to formulate the eigenvalue problem of Eq. (6.17) in terms of a basis of $\mathscr{L}_r$ or $\mathscr{L}_o$ only. In the general case there arises no simplification for either real or imaginary perturbations.

These general considerations play a role in coupled-Hartree–Fock theory for real and imaginary perturbations repectively (see paper II of this series [4]).

8. Stationarity with respect to two or more independent variational groups

Often one does not consider stationarity of the energy with respect to one variational group, but with respect to two or more independent groups. Two situations are possible: (**a**) the elements of the two groups commute, (**b**) they don't commute.

In the case (**a**) one can define the direct product $\mathscr{G} = \mathscr{G}_1 \otimes \mathscr{G}_2$ of the two groups and regard this as the new variational group. The new Lie algebra $\mathscr{L}$ is then simply the direct sum of the two Lie algebras $\mathscr{L} = \mathscr{L}_1 \oplus \mathscr{L}_2$. Stationarity with respect to $\mathscr{G}$ is equivalent to stationarity with respect to $\mathscr{G}_1$ and $\mathscr{G}_2$ independently.

Case (**b**) is more complicated. There is, of course, an embedding group $\mathscr{G}$ that consists of all (multiple) products of elements of $\mathscr{G}_1$ and $\mathscr{G}_2$, and the corresponding Lie algebra $\mathscr{L}$ consists of all elements of $\mathscr{L}_1$ and $\mathscr{L}_2$ and their linear combinations plus single and multiple commutators from elements of $\mathscr{L}_1$ and $\mathscr{L}_2$. To require that a Brillouin theorem is satisfied for all $X \in \mathscr{L}$ is a much stronger condition than that it holds for all $X \in \mathscr{L}_1$ and all $X \in \mathscr{L}_2$. Usually we will *not* demand that:

$$\langle \Psi | [H, [X_1, X_2]] | \Psi \rangle = 0; \qquad X_1 \in \mathscr{L}_1, \quad X_2 \in \mathscr{L}_2 \tag{8.1}$$

This has the immediate consequence that, except for special choices of X_1 and X_2:

$$\langle \Psi | [[H, X_1], X_2] | \Psi \rangle \neq \langle \Psi | [[H, X_2], X_1] | \Psi \rangle \tag{8.2}$$

i.e. that the Hessean defined by Eq. (2.11b) is non-symmetric (even if stationarity is achieved). This is, at first glance, puzzling. Let us look, for a moment, at the conventional formulation of the variational approach. One characterizes the admitted wave functions uniquely by a set of parameters, say $\kappa_1, \kappa_2 \ldots \kappa_n$. The energy expectation value E is then a function of these parameters and the stationarity conditions are:

$$\frac{\partial E}{\partial \kappa_k} = 0; \quad k = 1, 2 \ldots n \tag{8.3}$$

In a Newton–Raphson approach to solve Eq. (8.3) starting from some initial values of the κ_k one needs the second derivatives, which are, of course, symmetric:

$$\frac{\partial^2 E}{\partial \kappa_k \, \partial \kappa_l} = \frac{\partial^2 E}{\partial \kappa_l \, \partial \kappa_k} = H_{kl} \tag{8.4}$$

and which constitute the Hessean, while Eq. (8.3) corresponds to the Brillouin condition of Eq. (2.5).

In the present Lie-algebraic approach we characterize the operator U that transforms an initial function Φ to the optimized functions Ψ in terms of a set of parameters c_k:

$$U = \exp \sum_k c_k X_k \tag{8.5}$$

This is always possible and unique if the X_k are a basis of the Lie algebra $\mathscr{L}_r$. It is a fortiori so if the X_k commute. E is, for given Φ and U, a function of the c_k. The Brillouin conditions of Eq. (2.5) are equivalent to Eq. (8.3) and the Hessean of Eq. (8.4) agrees with the symmetrized Hessean of Eq. (2.12a).

If we consider two variational groups that don't commute, we can write a common transformation as:

$$U = U_1 U_2; \qquad U_1 \in \mathscr{G}_1, \quad U_2 \in \mathscr{G}_2 \tag{8.6a}$$

$$U_p = \exp \sigma_p; \qquad \sigma_p = \sum_k c_k^{(p)} X_k^{(p)} \in \mathscr{L}_p \tag{8.6b}$$

However, the result depends on the order of the factors, and a *unique parametrization of the wave function* and hence a unique definition of a variational approach requires that *one specifies the order in which elements of $\mathscr{G}_1$ and $\mathscr{G}_2$ are to be multiplied.* If this specification is made, second derivatives of the energy are again symmetric, although the Hessean in the definition of Eq. (2.11b) is not. Namely:

$$\frac{\partial^2 E}{\partial c_k^{(1)} \, \partial c_l^{(2)}} = \frac{\partial^2 E}{\partial c_l^{(2)} \, \partial c_k^{(1)}} = \langle \Psi \,|\, [[H, X_k^{(1)}], X_l^{(2)}] \,|\, \Psi \rangle \tag{8.7}$$

For the order specified by Eq. (8.6), one must commute H first with an element of $\mathscr{L}_1$, then with one of $\mathscr{L}_2$.

A Newton–Raphson approach is straightforward, based on:

$$\begin{aligned} E = \langle \phi \,|\, U_2^\dagger U_1^\dagger H U_1 U_2 \,|\, \phi \rangle = \langle \phi \,|\, H + [H, \sigma_1] + [H, \sigma_2] \\ + \tfrac{1}{2}[[H, \sigma_1], \sigma_1] + \tfrac{1}{2}[[H, \sigma_2], \sigma_2] + [[H, \sigma_1], \sigma_2] + \cdots \,|\, \phi \rangle \end{aligned} \tag{8.8}$$

$$\frac{\partial E}{\partial c_k^{(1)}} = \langle \phi | [H, X_k^{(1)}] + \tfrac{1}{2}[[H, X_k^{(1)}], \sigma_1] + \tfrac{1}{2}[[H, \sigma_1], X_k^{(1)}] + [[H, X_k^{(1)}], \sigma_2] + \cdots | \phi \rangle = 0 \tag{8.9a}$$

$$\frac{\partial E}{\partial c_k^{(2)}} = \langle \phi | [H, X_k^{(2)}] + \tfrac{1}{2}[[H, X_k^{(2)}], \sigma_2] + \tfrac{1}{2}[[H, \sigma_2], X_k^{(2)}] + [[H, \sigma_1^{(1)}], X_k^{(2)}] + \cdots | \phi \rangle = 0 \tag{8.9b}$$

In stationary perturbation theory in order to be consistent we make the ansatz:

$$\Psi(\lambda) = e^{Y^{(1)}(\lambda)} e^{Y^{(2)}(\lambda)} \Psi_0; \qquad Y^{(1)} \in \mathscr{L}_1; \qquad Y^{(2)} \in \mathscr{L}_2 \tag{8.10}$$

and expand $Y^{(1)}$ and $Y^{(2)}$ as in Eq. (4.3). Subscripts count orders of perturbations theory. The result for the E_k defined by Eqs. (4.4, 5) is then:

$$E_1 = \langle \Psi_0 | [H_0, Y_1^{(1)} + Y_1^{(2)}] + H_1 | \Psi_0 \rangle \tag{8.11a}$$

$$E_2 = \langle \Psi_0 | [H_0, Y_2^{(1)} + Y_2^{(2)}] + \tfrac{1}{2}[[H_0, Y_1^{(1)}], Y_1^{(1)}] + \tfrac{1}{2}[[H_0, Y_1^{(2)}], Y_1^{(2)}], + [[H_0, Y_1^{(1)}], Y_1^{(2)}] + [H_1, Y_1^{(1)} + Y_1^{(2)}] + H_2 | \Psi_0 \rangle \tag{8.11b}$$

The counterparts of the stationarity condition of Eq. (4.6b) are:

$$\langle \Psi_0 | [[H_0, X_k^{(1)}], Y_1^{(1)} + Y_1^{(2)}] + [H_1, X_k^{(1)}] | \Psi_0 \rangle = 0 \tag{8.12a}$$

$$\langle \Psi_0 | [[H_0, X_k^{(2)}], Y_1^{(2)}] + [[H_0, Y_1^{(1)}], X_k^{(2)}] + [H_1, X_k^{(2)}] | \Psi_0 \rangle = 0 \tag{8.12b}$$

The stationary results of Eqs. (4.7) and (4.10) become then:

$$E_1 = \langle \Psi_0 | H_1 | \Psi_0 \rangle$$

$$E_2 = \langle \Psi_0 | \tfrac{1}{2}[H_1, Y_1^{(1)} + Y_2^{(2)}] + H_2 | \Psi_0 \rangle$$

An example for a variational approach with two non-commuting variational groups is MC-SCF. Here $\mathscr{G}_1$ is the group of linear CI type variations within the 'active' space, while $\mathscr{G}_2$ is the group of orbital rotations like in ordinary SCF theory. Consider, e.g., the case of CAS-SCF and label occupied (inactive) orbitals as $i, j, k \ldots$, partially occupied (active) orbitals as $x, y, z, \ldots$ and virtual (unoccupied) orbitals as $a, b, c, \ldots$. Let us start from a Slater determinant, constructed from N_i orbitals of type i and N_x orbitals of type x. We now consider the Lie algebra with basis:

$$\mathscr{L}_1 \colon a_y^x, a_{zu}^{xy}, a_{uvw}^{xyz}, \ldots \tag{8.13a}$$

With a unitary transformation generated by this Lie algebra $\mathscr{L}_1$ we can create any CI *within the active space*. We then apply any one-particle transformation in the entire orbital space generated by:

$$\mathscr{L}_2 \colon a_i^a, a_i^x, a_x^a \text{ and hermitean conjugates} \tag{8.13b}$$

This performs the desired orbital rotations.

The groups obviously don't commute, since, e.g., (for different labels different):

$$[a_{xu}^{xy}, a_x^a] = a_{xu}^{ay} \notin \mathscr{L}_1, \tag{8.14}$$

As long as there are at least two electrons in active orbitals, the embedding group $\mathscr{G}$ of $\mathscr{G}_1$ and $\mathscr{G}_2$ is that of full CI generated by the basis of Eq. (3.15) of $\mathscr{L}_c$. (Labels p, q, r, s in Eq. (3.15) refer to arbitrary orbitals within the given basis.)

In fact, commutation of the result of Eq. (8.14) with a_i^z leads to a_{iu}^{ay}, by two more commutations with a_u^b and a_j^y we arrive at a_{ij}^{ab}, while commutation of a_{iu}^{ay} with a_{yu}^{xz} leads to a_{iuv}^{axz}, from which we can arrive at a_{ijk}^{abc} etc. *So CAS-SCF appears as an approximation to full CI, in which stationarity with respect to two subgroups of the variational group of full CI is achieved.*

Another example would be the combination of a linear variational approach with $\mathscr{G}_2$ generated by the shift operator of Eq. (3.7) and scaling with $\mathscr{G}_1$ generated by Eq. (3.5b).

Since there has been some discussion in the literature [9, 10, 13] on the meaning of stationarity with respect to two non-commuting variational groups (using here the notation of the present paper) the following points must be made clear.

1. Starting from a trial wave function and improving it iteratively in a Newton-Raphson type approach it is always possible to construct a wave function that is simultaneously stationary with respect to both groups. The coupling elements between the two groups in the Hessean have, of course, to be taken care of.

2. Exchanging the factors U_1 and U_2 in the unitary transformation of Eq. (8.6a) defines a different variational problem. Again it is possible to construct a wave function stationary with respect to the two groups, but this need not be (and will in general not be) the same as that obtained with the other order.

3. Often there is a 'natural order' of the two transformations, which leads to a convenient expression for the matrix elements, while the opposite order is difficult to handle. This is, e.g., the case for the example of MC-SCF mentioned above.

In a somewhat different context a decomposition of the relevant Lie algebra into a maximally soluble ideal and a semisimple subalgebra [17], leads to some simplifications. All Lie algebras considered here are simple and hence not decomposable.

9. Time-dependent perturbation theory

9.1. The stationarity principle

An approximative solution of the time-dependent Schrödinger equation:

$$\left\{H - i\hbar \frac{\partial}{\partial t}\right\} \Psi = 0 \tag{9.1a}$$

is often based on the so-called variation principle of Frenkel [27] (which can even be traced back to Dirac [28]):

$$\langle \delta\Psi | H - i\hbar \frac{\partial}{\partial t} | \Psi \rangle = 0 \tag{9.1b}$$

This is certainly not a 'variational principle' in any strict sense, not even a stationarity principle since there is no functional of Ψ for which Eq. (9.1b) is the stationarity condition. This is mainly due to the fact that $i\hbar\,\partial/\partial t$ is not a hermitean operator in the ordinary Hilbert space (see Sect. 9.7). Various discussions and possible improvement of Frenkel's principle can be found in the literature [2, 18, 29–33]. Nevertheless, if Eq. (9.1b) holds for all possible $\delta\Psi$ this

is a necessary and sufficient condition for Ψ to solve Eq. (9.1a). One can rewrite Eq. (9.1b) as:

$$\langle \Psi | X \left(H - i\hbar \frac{\partial}{\partial t} \right) | \Psi \rangle = 0, \tag{9.1c}$$

which is equivalent to Eq. (9.1b) if one considers all possible X. The questions arises what Eq. (9.1c) implies if one restricts it to a set of $X \in \mathscr{L}_c$, that constitute the Lie algebra of a variational group. If we replace X in Eq. (9.1c) by $X^\dagger$ and take the complex conjugate and then the sum as well as the difference of this with Eq. (9.1c) we get:

$$\frac{\partial}{\partial t} \langle \Psi | X | \Psi \rangle = \langle \Psi | \frac{\partial X}{\partial t} | \Psi \rangle + \frac{i}{\hbar} \langle \Psi | [H, X] | \Psi \rangle \tag{9.2a}$$

$$\frac{1}{i\hbar} \langle \Psi | [H, X]_+ | \Psi \rangle = \left\langle \Psi | X \middle| \frac{\partial \Psi}{\partial t} \right\rangle - \left\langle \frac{\partial \Psi}{\partial t} \middle| X | \Psi \right\rangle \tag{9.2b}$$

We recognize Eq. (9.2a) as the 'time-dependent hypervirial relation' or 'generalized Ehrenfest theorem' [2] or as Heisenberg's equation of motion for expectation values. This looks like the searched-for time-dependent generalization of the generalized Brillouin theorem of Eq. (2.5). The meaning of Eq. (9.2b) is less clear.

Let us, for the moment, try to formulate the time-independent case in a similar way. We want to satisfy the time-independent Schrödinger equation:

$$H\Psi = E\Psi \tag{9.3a}$$

by means of the condition:

$$\langle \delta\Psi | H - E | \Psi \rangle = 0 \tag{9.3b}$$

$$\langle \Psi | X(H - E) | \Psi \rangle = 0 \tag{9.3c}$$

in perfect analogy to Eqs. (9.1b) and (9.1c). In analogy to Eq. (9.2) we get:

$$\langle \Psi | [H, X] | \Psi \rangle = 0 \tag{9.4a}$$

$$\tfrac{1}{2} \langle \Psi | [H, X]_+ | \Psi \rangle = E \langle \Psi | X | \Psi \rangle \tag{9.4b}$$

By comparison of the derivation of Eq. (9.4a) – which agrees with Eq. (2.5) – from stationarity of $\langle \Psi | H | \Psi \rangle$ we conclude that Eq. (9.4b) is not needed (see also ref. [18]), except for the special case $X = 1$ or any other constant, i.e. for $\delta\Psi$ proportional to Ψ, where Eq. (9.4b) reads:

$$\langle \Psi | H | \Psi \rangle = E \langle \Psi | \Psi \rangle \tag{9.4c}$$

We hence ignore Eq. (9.2b), i.e. the counterpart of (9.4b) except for $X = 1$, which means:

$$\frac{2}{i\hbar} \langle \Psi | H | \Psi \rangle = \left\langle \Psi \middle| \frac{\partial \Psi}{\partial t} \right\rangle - \left\langle \frac{\partial \Psi}{\partial t} \middle| \Psi \right\rangle \tag{9.5a}$$

and use this together with Eq. (9.2a) as the basis of stationary time-dependent perturbation theory. While Eq. (9.4a) gives no nontrivial result for $X = 1$, we get from Eq. (9.2a):

$$\frac{\partial}{\partial t} \langle \Psi | \Psi \rangle = 0 \tag{9.5b}$$

This obviously expresses the conservation of the norm of the wave function and the unitarity of the time evolution. We satisfy Eq. (9.5b) automatically by choosing Ψ normalized to unity and by considering only unitary variations and unitary time-dependent changes of Ψ.

Starting point for the following derivations is hence that Eq. (9.2a) holds for all $X \in \mathscr{L}_c$, together with Eq. (9.5a).

We shall develop this formalism in Sect. 9.2. There we shall concentrate on the case of a time-independent unperturbed Hamiltonian and a time-dependent perturbation. In this case those states are of special interest that go to stationary states in the limit $\lambda \to 0$. These states conserve some properties of stationary states. For them the energy deserves some interest, although it is strictly not a constant of motion. We shall tentatively define the energy as the expectation value of the Hamiltonian, but we shall see later (Sect. 9.6) that there are other and more appropriate definitions.

After a short comment on the time evolution in Sect. 9.3, we study the special cases of an adiabatically switched static perturbation in Sect. 9.4 and a perturbation periodic in time in Sect. 9.5. There the results of Sect. 6.3 will turn out to be important. We then reconsider the definition of the energy in Sect. 9.6 and define a so-called *pseudoenergy* for which a *Hellmann–Feynman theorem* holds. We then establish the relation to the concept of *Floquet states in Sect.* 9.7.

9.2. Stationary theory for a time-dependent perturbation

We consider the case that the *unperturbed Hamiltonian is time-independent*, and that there is a first order *time-dependent perturbation*, i.e. we choose:

$$H = H_0 + \lambda V(t) \tag{9.6}$$

with H_0 time-independent. For Ψ we make the ansatz (which guarantees that Eq. (9.5b) is automatically satisfied):

$$\Psi(\lambda, t) = e^{Y(\lambda,t)}\Psi_0(t); \qquad Y = -Y^\dagger \in \mathscr{L}_c \tag{9.7a}$$

$$Y(\lambda, t) = \sum_{k=1}^{\infty} \lambda^k Y_k(t) \tag{9.7b}$$

The expectation value of the Hamiltonian expanded in powers of λ is – in analogy to Eq. (4.5):

$$E(t) = \langle \Psi(\lambda, t) | H | \Psi(\lambda, t) \rangle \tag{9.8a}$$

$$E_0(t) = \langle \Psi_0(t) | H_0 | \Psi_0(t) \rangle \tag{9.8b}$$

$$E_1(t) = \langle \Psi_0(t) | V(t) + [H_0, Y_1(t)] | \Psi_0(t) \rangle \tag{9.8c}$$

$$E_2(t) = \langle \Psi_0(t) | [V(t), Y_1(t)] + [H_0, Y_2(t)] + \tfrac{1}{2}[[H_0, Y_1(t)], Y_1(t)] | \Psi_0(t) \rangle \tag{9.8d}$$

$$\begin{aligned} E_3(t) = \langle \Psi_0(t) | & V(t), Y_2(t)] + [H_0, Y_3(t)] + \tfrac{1}{2}[[V(t), Y_1(t)], Y_1(t)] \\ & + \tfrac{1}{2}[[H_0, Y_1(t)], Y_2(t)] + \tfrac{1}{2}[[H_0, Y_2(t)], Y_1(t)] \\ & + \tfrac{1}{6}[[[H_0, Y_1(t)], Y_1(t)], Y_1(t)] | \Psi_0(t) \rangle \end{aligned} \tag{9.8e}$$

We determine $\Psi_0(t)$ and the $Y_k(t)$ from the expansion of the stationary conditions of Eqs. (9.2a) and (9.5a) in powers of λ:

$$C_0(t) \equiv \langle \Psi_0(t) | [H_0, X] - i\hbar \frac{\partial X}{\partial t} | \Psi_0(t) \rangle + i\hbar \frac{\partial}{\partial t} \langle \Phi_0(t) | X | \Psi_0(t) \rangle = 0 \quad (9.9a)$$

$$C_1(t) \equiv \langle \Psi_0(t) | \left[[H_0, X] - i\hbar \frac{\partial X}{\partial t}, Y_1 \right] + [V, X] | \Psi_0(t) \rangle + i\hbar \frac{\partial}{\partial t} \langle \Psi_0(t) | [X, Y_1] | \Psi_0(t) \rangle = 0 \quad (9.9b)$$

$$C_2(t) \equiv \langle \Psi_0(t) | \left[[H_0, X] - i\hbar \frac{\partial X}{\partial t}, Y_2 \right] + [[V, X], Y_1] + \frac{1}{2} \left[\left[[H_0, X] - i\hbar \frac{\partial X}{\partial t}, Y_1 \right], Y_1 \right] | \Psi_0(t) \rangle + i\hbar \frac{\partial}{\partial t} \langle \Psi_0(t) | [X, Y_2] + \tfrac{1}{2}[[X, Y_1], Y_1] | \Psi_0(t) \rangle = 0 \quad (9.9c)$$

$$D_0(t) \equiv Re \langle \Psi_0(t) | H_0 - i\hbar \frac{\partial}{\partial t} | \Psi_0(t) \rangle = 0 \quad (9.10a)$$

$$D_1(t) \equiv Re \langle \Psi_0(t) | [H_0, Y_1] + V - i\hbar \frac{\partial Y_1}{\partial t} | \Psi_0(t) \rangle = 0 \quad (9.10b)$$

$$D_2(t) \equiv Re \langle \Psi_0(t) | [H_0, Y_2] + \tfrac{1}{2}[[H_0, Y_1], Y_1] + [V, Y_1] - i\hbar \frac{\partial Y_2}{\partial t} - \tfrac{1}{2} i\hbar \left[\frac{\partial Y_1}{\partial t}, Y_1 \right] | \Psi_0(t) \rangle \quad (9.10c)$$

The elements of the Lie algebra $\mathscr{L}_c$ are now time-dependent. We realize this by choosing time-independent generators and a time-dependent field of coefficients in the linear combinations of elements.

The condition of Eq. (9.9a) determines the unperturbed wave function $\Psi_0(t)$. So far we have not used that H_0 is time-independent. We now not only make this special choice, but we even specify *that* $\Psi_0(t)$ *describes a stationary state*, i.e. that its time-dependence factors off. In this case Eq. (9.9a) simplifies to the time-independent condition of Eq. (5.4a) and from Eq. (9.10a) we conclude that Ψ_0 is of the form:

$$\Psi_0(t) = \psi_0 \exp \left[\frac{\langle \psi_0 | H_0 | \psi_0 \rangle}{i\hbar} t \right] \quad (9.11)$$

with ψ_0 time-independent.

For stationary time-dependent perturbation theory based on a stationary time-independent unperturbed state we can replace $\Psi_0(t)$ in Eqs. (9.8, 9) and (9.10b,c) by ψ_0, since the operators in these expressions don't contain any differentiations with respect to time. We then get:

$$C_0(t) = \langle \psi_0 | [H_0, X(t)] | \psi_0 \rangle = 0 \quad (9.12a)$$

$$C_1(t) = \langle \psi_0 | [[H_0, X], Y_1] + [V, X] - i\hbar \left[\frac{\partial Y_1}{\partial t}, X \right] | \psi_0 \rangle = 0 \quad (9.12b)$$

$$C_2(t) = \langle\psi_0|[[H_0, X], Y_2] - i\hbar\left[\frac{\partial Y_2}{\partial t}, X\right] + [[V, X], Y_1] + \tfrac{1}{2}[[[H_0, X], Y_1], Y_1] + \frac{i\hbar}{2}\left\{\left[\left[X, \frac{\partial Y_1}{\partial t}\right], Y_1\right] + \left[[X, Y_1], \frac{\partial Y_1}{\partial t}\right]\right\}|\psi_0\rangle = 0 \tag{9.12c}$$

$$D_1(t) = \langle\psi_0|V - i\hbar\frac{\partial Y_1}{\partial t}|\psi_0\rangle = 0 \tag{9.13a}$$

$$D_2(t) = \langle\psi_0|\tfrac{1}{2}[[H_0, Y_1], Y_1] + [V, Y_1] - i\hbar\frac{\partial Y_2}{\partial t} - \tfrac{1}{2}i\hbar\left[\frac{\partial Y_1}{\partial t}, Y_1\right]|\psi_0\rangle = 0 \tag{9.13b}$$

If Eqs. (9.12) are satisfied, Eqs. (9.8) can be simplified to:

$$E_1(t) = \langle\psi_0|V(t)|\psi_0\rangle \tag{9.14a}$$

$$E_2(t) = \tfrac{1}{2}\langle\psi_0|[V(t), Y_1(t)] + i\hbar\left[\frac{\partial Y_1}{\partial t}, Y_1\right]|\psi_0\rangle \tag{9.14b}$$

$$E_3(t) = \langle\psi_0|\tfrac{1}{2}[[V, Y_1], Y_1] + \tfrac{1}{6}[[[H_0, Y_1], Y_1], Y_1] + i\hbar\left[\frac{\partial Y_1}{\partial t}, Y_2\right]|\psi_0\rangle \tag{9.14c}$$

If Eqs. (9.13) hold, alternative expressions for the $E_k(t)$ exist:

$$E_1(t) = i\hbar\langle\psi_0|\frac{\partial Y_1}{\partial t}|\psi_0\rangle \tag{9.15a}$$

$$E_2(t) = i\hbar\langle\psi_0|\frac{\partial Y_2}{\partial t} + \frac{1}{2}\left[\frac{\partial Y_1}{\partial t}, Y_1\right]|\psi_0\rangle \tag{9.15b}$$

$$E_3(t) = i\hbar\langle\psi_0|\frac{\partial Y_3}{\partial t} + \frac{1}{2}\left[\frac{\partial Y_2}{\partial t}, Y_1\right] + \frac{1}{2}\left[\frac{\partial Y_1}{\partial t}, Y_2\right] + \frac{1}{6}\left[\left[\frac{\partial Y_1}{\partial t}, Y_1\right], Y_1\right]|\psi_0\rangle \tag{9.15c}$$

The Eqs. (9.12–15) *are the basis of stationary perturbation theory for a time-dependent perturbation.* Note that in order to evaluate the $E_k(t)$ only Eqs. (9.12) and (9.14) are needed. Equations (9.13) or (9.15) become relevant if one wants to know the *entire* $Y_k(t)$ explicitly, and not only that part of the $Y_k(t)$ which is needed in Eq. (9.14) for the construction of the $E_k(t)$.

To appreciate this subtle detail, note that one can always decompose $Y_1(t)$ in the following way:

$$Y_k(t) = Y_{kD}(t) + Y_{kN}(t) \tag{9.16a}$$

$$Y_{kD} = \langle\psi_0|Y_k(t)|\psi_0\rangle \tag{9.16b}$$

$$Y_{kN} = Y_k - Y_{kD} \tag{9.16c}$$

The subscripts D and N stand for *diagonal* and *nondiagonal* [16]. Obviously only Y_{kD} contributes to Eqs. (9.13a) and (9.15a), while only Y_{kN} contributes to Eqs. (9.14b) or (9.14c). The same decomposition of Eq. (9.16) is also possible in the time-independent theory, but there Y_{kD} is a constant, which must be imaginary, since Y_k has to be antihermitean. Hence the Y_{kD} only give rise to a physically irrelevant phase factor, which can be ignored. In time-dependent theory any Y_{kD} is a function of time and hence induces a *time-dependent phase factor*, and this

does have a physical meaning, though it usually has no effect on expectation values. It is now obvious that Eqs. (9.15) must not be regarded as energy expressions alternative to Eq. (9.14) but rather as conditions for the determination of Y_{kD}.

One should therefore rewrite Eq. (9.15) as:

$$i\hbar\langle\psi_0|\frac{\partial Y_1}{\partial t}|\psi_0\rangle = \langle\psi_0|V(t)|\psi_0\rangle \tag{9.17a}$$

$$i\hbar\langle\psi_0|\frac{\partial Y_2}{\partial t}|\psi_0\rangle = E_2(t) - \tfrac{1}{2}\langle\psi_0|\left[\frac{\partial Y_1}{\partial t}, Y_1\right]|\psi_0\rangle$$

$$= \tfrac{1}{2}\langle\psi_0|[V(t), Y_1(t)]|\psi_0\rangle \tag{9.17b}$$

$$i\hbar\langle\psi_0|\frac{\partial Y_3}{\partial t}|\psi_0\rangle = E_3(t) - \tfrac{1}{2}\langle\psi_0|\left[\frac{\partial Y_2}{\partial t}, Y_1\right] + \left[\frac{\partial Y_1}{\partial t}, Y_2\right]|\psi_0\rangle$$

$$-\tfrac{1}{6}\langle\psi_0|\left[\left[\frac{\partial Y_1}{\partial t}, Y_1\right], Y_1\right]|\psi_0\rangle$$

$$= \langle\psi_0|\tfrac{1}{2}[[V, Y_1], Y_1] + \tfrac{1}{6}[[[H_0, Y_1], Y_1], Y_1] + \tfrac{1}{2}i\hbar\left[\frac{\partial Y_1}{\partial t}, Y_2\right]$$

$$-\tfrac{1}{2}i\hbar\left[\frac{\partial Y_2}{\partial t}, Y_1\right] - \tfrac{1}{6}i\hbar\left[\left[\frac{\partial Y_1}{\partial t}, Y_1\right], Y_1\right]|\psi_0\rangle \tag{6.17c}$$

In order to determine Y_1 (or more precisely Y_{1N}) we expand it in a basis X_l of $\mathscr{L}_c$ and insert this into Eq. (9.12b) and write X_k for X:

$$Y_1 = \sum_l b_l(t)X_l \tag{9.18a}$$

$$\sum_k \left\{\langle\psi_0|[[H_0, X_l], X_k]|\psi_0 > b_k(t) - i\hbar\frac{\partial b_k}{\partial t}\langle\psi_0|[X_l, X_k]|\psi_0\rangle\right\}$$

$$+ \langle\psi_0|[V, X_k]|\psi_0\rangle = 0 \tag{9.18b}$$

This is a linear system of coupled differential equations, that involves the Hessean H_{kl} of Eq. (4.13b) and the metric Δ_{kl} of Eq. (6.16). If we introduce the operators Π_k which diagonalize $\underline{\underline{H}}$ with the metric $\underline{\underline{\Delta}}$, the system of Eq. (9.18) is decoupled. We expand:

$$Y_1 = \sum_l f_l(t)\Pi_l \tag{9.19a}$$

insert this into Eq. (9.12a) with $X = \Pi_k^\dagger$ to arrive at:

$$\sum_l \tilde{\tilde{H}}_{kl}f_l - i\hbar\sum_l \tilde{\tilde{\Delta}}_{kl}\frac{df_l}{dt} + \tilde{\tilde{V}}_k^* = 0 \tag{9.19b}$$

with $\tilde{\tilde{H}}_{kl}$, $\tilde{\tilde{\Delta}}_{kl}$ and $\tilde{\tilde{V}}_k$ defined by Eqs. (6.21) and (6.22c) respectively, with the only difference that now $V = H_1$ as well as V_l are time-dependent. Making use of Eq. (6.21) we can rewrite Eq. (9.19b) as:

$$\left[\varepsilon_k f_k - i\hbar\frac{df_k}{dt}\right]\tilde{\tilde{\Delta}}_{kk} + \tilde{\tilde{V}}_k^* = 0 \tag{9.19c}$$

which is a set of uncoupled differential equations.

We make the transformation (reminiscent of the interaction representation):

$$f_k(t) = g_k(t) \exp\left(\frac{\varepsilon_k}{i\hbar} t\right) \tag{9.20a}$$

$$\frac{df_k}{dt} = \left(\frac{dg_k}{dt} + \frac{\varepsilon_k}{i\hbar} g_k\right) \exp\left(\frac{\varepsilon_k}{i\hbar} t\right) \tag{9.20b}$$

such that Eq. (9.19b) becomes:

$$i\hbar \frac{dg_k}{dt} \tilde{\tilde{\Delta}}_{kk} \exp\left(\frac{\varepsilon_k}{i\hbar} t\right) - \tilde{\tilde{V}}_k^* = 0 \tag{9.20c}$$

Now the g_k can be obtained by direct integration, somewhat like in Dirac's method of the 'variation of constants'. Again the antihermiticity of Y_1 comes out automatically (for V hermitean).

9.3. Perturbation theory of time evolution

The solution of the time-dependent Schrödinger equation is an initial-value problem. Given Ψ at a time $t = t_0$ one wants to know the evolution of Ψ with time. This is determined by the time evolution operator $\exp Y$. Assume again that the unperturbed problem is time-independent and that at $t = t_0$ the system is in a stationary state Ψ_0 of the unperturbed problem, then in view of Eqs. (9.11, 15a, 16):

$$Y_0 = \frac{1}{i\hbar} \langle\psi_0|H_0|\psi_0\rangle t \tag{9.21a}$$

$$Y_1 = \frac{1}{i\hbar} \int_{t_0}^{t} \langle\psi_0|V(t)|\psi_0\rangle \, d\tau + Y_{1N} \text{ etc.} \tag{9.21b}$$

with Y_{1N} determined as solution of Eq. (9.12b) or explicitly Eq. (9.18) to Eq. (9.21).

9.4. Application to a time-independent perturbation

For the special case of an exponentially switched static perturbation:

$$V(t) = e^{\eta t} V_0; \quad \eta > 0; \qquad v_k = \langle\psi_0|[V_0, \Pi_k]|\psi_0\rangle \tag{9.22}$$

which vanishes in the limit $t \to -\infty$, we get from Eq. (9.20):

$$g_k(t)\tilde{\tilde{\Delta}}_{kk} = \frac{1}{i\hbar} v_k^* \int_{-\infty}^{t} \exp\left[\left(\frac{i\varepsilon_k}{\hbar} + \eta\right) t'\right] dt'$$

$$= -v_k^* \frac{1}{\varepsilon_k - i\hbar\eta} \exp\left[\left(\frac{i\varepsilon_k}{\hbar} + \eta\right) t\right] \tag{9.23a}$$

$$f_k(t) \tilde{\tilde{\Delta}}_{kk} = -v_k^* \, e^{\eta t} (\varepsilon_k - i\hbar\eta)^{-1} \tag{9.23b}$$

If we take the limit $\eta \to 0$ we get the same result of Eq. (6.22) as in time-independent stationary perturbation theory for Y_1. We also get trivially the same result

for $E_1(t)$ given by Eq. (9.14a) and for $E_2(t)$ as given by Eq. (9.14b) and for the expressions given by Eq. (9.17), since in the limit $\eta \to 0$, $\partial Y_1/\partial t = 0$. The limit $\eta \to 0$ does apparently not cause any problems, unlike in traditional time-dependent theory [29, 34, 35], where the limit of the wave function for $\eta \to 0$ does not exist.

The singularity of the traditional theory can be traced back even in the present formalism. It namely arises in the 'diagonal part' Y_{1D} of Y_1, as defined by Eq. (9.16b). If we want to construct Y_{1D} from Eq. (9.17a) for the present example:

$$i\hbar\langle\psi_0|\frac{\partial Y_1}{\partial t}|\psi_0\rangle = i\hbar\frac{\partial Y_{1D}}{\partial t} = e^{\eta t}\langle\psi_0|V_0|\psi_0\rangle \tag{9.24}$$

and integrate from $-\infty$ to t we get:

$$Y_{1D} = \frac{1}{i\hbar\eta}e^{\eta t}\langle\psi_0|V_0|\psi_0\rangle \tag{9.25}$$

This obviously diverges in the limit $\eta \to 0$. However if we divide the integration domain into two parts, one from $-\infty$ to 0, the other from 0 to t, we get:

$$Y_{1D} = \frac{1}{i\hbar\eta}\langle\psi_0|V_0|\psi_0\rangle + \frac{1}{i\hbar}(t + \tfrac{1}{2}\eta t^2 + \cdots)\langle\psi_0|V_0|\psi_0\rangle \tag{9.26}$$

The first term in Eq. (9.26), which diverges, is time-independent, it represents an indefinite time-independent phase, while the second term in Eq. (9.26) leads in the limit $\eta \to 0$ to the first order correction of the time-dependent phase factor from $(1/i\hbar)E_0 t$ to $(1/i\hbar)(E_0 + E_1)t$. That the need to correct this phase in a time-dependent theory of time-independent perturbations leads to difficulties, has been observed previously [29, 34, 35]. It is also clear that in the present theory, where one only considers expectation values, an indefinite time-independent phase cannot do any harm.

9.5. A perturbation periodic in time

Of special interest is the case of a perturbation periodic in time:

$$V(t) = \cos(\omega t)V_0 \tag{9.27}$$

We introduce a switching function $\exp(\eta t)$ as in Eq. (9.22) and we define (assuming V_0 to be hermitean):

$$v_k = \langle\psi_0|[V_0, \Pi_k|\psi_0\rangle = -\langle\psi_0|[V_0, \Pi_k^\dagger]|\psi_0\rangle^* \tag{9.28}$$

From Eq. (9.20) we then get:

$$g_k(t)\tilde{\tilde{\Delta}}_{kk} = \frac{1}{i\hbar}v_k^*\int_{-\infty}^{t}\exp\left\{\left(\frac{i\varepsilon_k}{\hbar} + \eta\right)t'\right\}\cos(\omega t')\,dt' \tag{9.29}$$

After integration the limit $\eta \to 0$ can be taken except at the poles where $\hbar\omega = \pm\varepsilon_k$. Applying Eq. (9.20a) we then get:

$$f_k(t)\tilde{\tilde{\Delta}}_{kk} = -\tfrac{1}{2}v_k^*\left\{\frac{\exp(i\omega t)}{\varepsilon_k + \hbar\omega} + \frac{\exp(-i\omega t)}{\varepsilon_k - \hbar\omega}\right\} \tag{9.30}$$

Inserting this into Eq. (9.19a) and then into Eq. (9.14b) we obtain:

$$E_2 = -\tfrac{1}{8}\sum_k |v_k|^2 \tilde{\tilde{A}}_{kk}\varepsilon_k \left\{\frac{2}{\varepsilon_k^2 - \hbar^2\omega^2} + \frac{\exp(2i\omega t)}{(\varepsilon_k + \hbar\omega)^2} + \frac{\exp(-2i\omega t)}{(\varepsilon_k - \hbar\omega)^2}\right\} \tag{9.31}$$

We also give the result for the pseudo-energy, discussed in more detail in Sect. 9.6, namely for that defined by Eq. (9.49b):

$$W_2 = \tfrac{1}{2}\langle\psi_0|[V, Y_1]|\psi_1\rangle = -\tfrac{1}{8}\sum_k |v_k|^2 \tilde{\tilde{A}}_{kk}\left\{\frac{2\varepsilon_k}{\varepsilon_k^2 - \hbar^2\omega^2} + \frac{\exp(2i\omega t)}{\varepsilon_k + \hbar\omega} + \frac{\exp(-2i\omega t)}{\varepsilon_k - \hbar\omega}\right\} \tag{9.32}$$

The singularities of Eqs. (9.30–32) for $\hbar\omega = \pm\varepsilon_k$ indicate that the ε_k must, in fact, be interpreted as model excitation energies, and that the metric which led to their definition (see Sect. 6) does make sense.

The present approach to a regularized time-dependent perturbed theory appears easier and more transparent than others from the literature [29, 24, 35].

Now we show that *in stationary perturbation theory the 'dipole length' and the 'dipole-velocity' formula are equivalent*, provided that:

$$V_0 = e\vec{\mathscr{E}} \cdot \vec{r} \tag{9.33}$$

is exactly expandable in the Lie algebra $\mathscr{L}_c$, with respect to which stationarity is achieved, or rather that $V_0\psi_0$ is expandable in the function basis $X_k\psi_0$, and hence in $\Pi_k\psi_0$ or $\Pi_k^\dagger\psi_0$.

We expand:

$$V_0 = \sum_k a_k \Pi_k^\dagger \tag{9.34}$$

and get, using Eq. (6.15):

$$\begin{aligned} v_k &= \langle\psi_0|[V_0, \Pi_k]|\psi_0\rangle = \sum_l a_l \langle\psi_0|[\Pi_l^\dagger, \Pi_k]|\psi_0\rangle \\ &= \sum_l \frac{a_l}{\varepsilon_k}\langle\psi_0|[\Pi_l^\dagger, [H_0, \Pi_k]]|\psi_0\rangle = \frac{1}{\varepsilon_k}\langle\psi_0|[V_0, [H_0, \Pi_k]]|\psi_0\rangle \\ &= \frac{1}{\varepsilon_k}\langle\psi_0|[\Pi_k, [H_0, V_0]]|\psi_0\rangle \end{aligned} \tag{9.35}$$

we now use:

$$[H_0, V_0] = e\vec{\mathscr{E}} \cdot [H_0, \vec{r}] = \frac{e\hbar^2}{m}\vec{\mathscr{E}} \cdot \nabla \tag{9.36}$$

and get

$$v_k = -\frac{e\hbar^2}{\varepsilon_k m}\langle\psi_0|[\vec{\mathscr{E}} \cdot \nabla, \Pi_k]|\psi_0\rangle \tag{9.37}$$

which is the dipole-velocity formula.

This result is a special case of an off-diagonal Hellmann–Feynman theorem. It has been known that it holds within traditional RPA (with exchange) that is equivalent to linearized time-dependent Hartree–Fock, which is a special case of stationary time-dependent perturbation theory [36, 37].

9.6. *The time-dependent pseudo-energy*

In this section we make the same starting assumptions as in Sect. 9.2, hence the results will be essentially the same, we just use a different definition for the energy. Like in Sect. 9.1 we start with an analogy to *time-independent* Hamiltonians. For these stationary states satisfy the condition:

$$i\hbar \frac{\partial \Psi}{\partial t} = E\Psi \tag{9.38a}$$

with E time-independent. These states are of the form of Eq. (9.11), with ψ_0 making E_0 stationary. One may try to generalize the condition of Eq. (9.38a) for time-dependent Hamiltonians to:

$$i\hbar \frac{\partial \Psi}{\partial t} = W(t)\Psi \tag{9.38b}$$

(with W real), which has the solution:

$$\Psi = \exp\left\{\frac{1}{i\hbar}\int_{t_0}^{t} W(t')\,dt'\right\}\psi \tag{9.39}$$

with ψ time-independent. Let us now regard Eq. (9.39) as an ansatz for Ψ and insert this into the time-dependent Schrödinger equation (9.1a). This leads to the following equation for ψ [30–33]:

$$\left\{H - W(t) - i\hbar \frac{\partial}{\partial t}\right\}\psi = 0 \tag{9.40}$$

Generally the solution ψ of Eq. (9.40) will not be time-independent, i.e. the Ψ given by Eq. (9.39) will not be solution of Eq. (9.38b). However wave functions of the form of Eq. (9.39) with Ψ time-dependent do satisfy the time-dependent Schrödinger equation (9.1a) provided that ψ satisfies Eq. (9.40). In fact, one may choose $W(t)$ rather arbitrarily and always find (at least one) corresponding ψ and Ψ. Even $W(t) \equiv 0$ will do. There is hence so far nothing unique in this procedure. However, we shall later use the freedom left in the definitions of $W(t)$ to make an optimum choice, in the sense of a criterion to be specified. Of course we are interested in states that become stationary in the limit $\lambda \to 0$ and we want to describe them in terms of an energy $W(t)$ and a wave function $\psi(t)$ that are 'as little time-dependent as possible'.

Note that for the case that we consider, which has already been studied in Sect. 9.2, a particular choice of $W(t)$ and $\psi(t)$ does not affect the physics, only possibly the interpretation. We call $W(t)$ the pseudoenergy and $\psi(t)$ the pseudo-eigenfunction. In terms of ψ and W Frenkel's principle of Eq. (9.1b) will be replaced by:

$$\langle \delta\psi \,|\, H - W(t) - i\hbar \frac{\partial}{\partial t} \,|\, \psi \rangle = 0 \tag{9.41}$$

The 'generalized Ehrenfest theorem' of Eq. (9.2a) keeps the same form just with Ψ replaced by ψ. The same holds for Eq. (9.5b) while Eq. (9.5a) has to be replaced by:

$$W(t) = \langle \psi \,|\, H \,|\, \psi \rangle - \frac{i\hbar}{2}\left[\left\langle \psi \,\middle|\, \frac{\partial \psi}{\partial t}\right\rangle - \left\langle \frac{\partial \psi}{\partial t} \,\middle|\, \psi \right\rangle\right] \tag{9.42}$$

We choose again a Hamiltonian in the form of Eq (9.6) and instead of Eq. (6.7a) we have:

$$\Psi = \exp[Y_0(t)]\psi; \qquad \psi = \exp[\tilde{Y}(t)]\psi_0$$

$$\frac{\partial}{\partial t} Y_0(t) = i\hbar W(t); \qquad Y_0 = \sum_{k=0}^{\infty} \lambda^k Y_{0k}; \qquad \tilde{Y} = \sum_{k=0}^{\infty} \lambda^k \tilde{Y}_k \tag{9.43}$$

Since $Y_0(t)$ is a scalar and commutes with $\tilde{Y}(t)$ we can also write Eqs. (9.43) as (9.7a) with:

$$Y = Y_0 + \tilde{Y} - (i\hbar)^{-1} E_0 t \tag{9.44}$$

In Eqs. (9.8) and (9.9) we only need to replace Ψ_0 by ψ_0 and Y_k by $\tilde{Y}_k$ which means that Eq. (9.9) becomes Eq. (9.12) – with Y_k replaced by $\tilde{Y}_k$, while in the counterpart of Eqs. (9.10) or (9.13) now the W_k appear:

$$W_0(t) = \langle \psi_0 | H_0 - i\hbar \frac{\partial}{\partial t} | \psi_0 \rangle \tag{9.45a}$$

$$W_1(t) = \langle \psi_0 | V - i\hbar \frac{\partial \tilde{Y}_1}{\partial t} | \psi_0 \rangle \tag{9.45b}$$

$$W_2(t) = \langle \psi_0 | [V, \tilde{Y}_1] + \tfrac{1}{2}[[H_0, \tilde{Y}_1], \tilde{Y}_1] - i\hbar \frac{\partial \tilde{Y}_2}{\partial t} - i\hbar \left[\frac{\partial \tilde{Y}_1}{\partial t}, \tilde{Y}_1\right] | \psi_0 \rangle \tag{9.45c}$$

$$\begin{aligned} W_3(t) = \langle \psi_0 | & [V, \tilde{Y}_2] + \tfrac{1}{2}[[V, \tilde{Y}_1], \tilde{Y}_1] + \tfrac{1}{2}[[H_0, \tilde{Y}_1], \tilde{Y}_2] \\ & + \tfrac{1}{2}[[H_0, \tilde{Y}_2], \tilde{Y}_1] + \tfrac{1}{6}[[[H_0, \tilde{Y}_1], \tilde{Y}_1], \tilde{Y}_1] - i\hbar \frac{\partial \tilde{Y}_3}{\partial t} \\ & - \frac{i\hbar}{2}\left[\frac{\partial \tilde{Y}_1}{\partial t}, \tilde{Y}_2\right] - \frac{i\hbar}{2}\left[\frac{\partial \tilde{Y}_2}{\partial t}, \tilde{Y}_1\right] \\ & - \frac{i\hbar}{6}\left[\left[\left[\frac{\partial \tilde{Y}_1}{\partial t}, \tilde{Y}_1\right], \tilde{Y}_1\right], \tilde{Y}_1\right] | \psi_0 \rangle \end{aligned} \tag{9.45d}$$

We also get Eqs. (9.14) for the E_k with Y_k replaced by $\tilde{Y}_k$, while instead of Eq. (9.15) we have:

$$W_1(t) = E_1(t) - i\hbar \langle \psi_0 | \frac{\partial \tilde{Y}_1}{\partial t} | \psi_0 \rangle \tag{9.46a}$$

$$W_2(t) = E_2(t) - i\hbar \langle \psi_0 | \frac{\partial \tilde{Y}_2}{\partial t} + \frac{1}{2}\left[\frac{\partial \tilde{Y}_1}{\partial t}, \tilde{Y}_1\right] | \psi_0 \rangle \tag{9.46b}$$

$$\begin{aligned} W_3(t) = E_3(t) - i\hbar \langle \psi_0 | & \frac{\partial \tilde{Y}_3}{\partial t} + \frac{1}{2}\left[\frac{\partial \tilde{Y}_2}{\partial t}, \tilde{Y}_1\right] + \frac{1}{2}\left[\frac{\partial \tilde{Y}_1}{\partial t}, \tilde{Y}_2\right] \\ & + \frac{1}{6}\left[\left[\frac{\partial \tilde{Y}_1}{\partial t}, \tilde{Y}_1\right], \tilde{Y}_1\right] | \psi_0 \rangle \end{aligned} \tag{9.46c}$$

Combining Eq. (9.45) with Eqs. (9.46) and (9.12) (with Y_k replaced by $\tilde{Y}_k$) we get:

$$W_1 = \langle \psi_0 | V | \psi_0 \rangle - i\hbar \langle \psi_0 | \frac{\partial \tilde{Y}_1}{\partial t} | \psi_0 \rangle \tag{9.47a}$$

$$W_2 = \tfrac{1}{2}\langle\psi_0|[V, \tilde{Y}_1]|\psi_0\rangle - i\hbar\langle\psi_0|\frac{\partial \tilde{Y}_2}{\partial t}|\psi_0\rangle \tag{9.47b}$$

$$W_3 = \langle\psi_0|\tfrac{1}{2}[[V, \tilde{Y}_1], \tilde{Y}_1] + \tfrac{1}{6}[[[H_0, \tilde{Y}_1], \tilde{Y}_1], \tilde{Y}_1]|\psi_0\rangle$$
$$- i\hbar\langle\psi_0|\frac{\partial \tilde{Y}_3}{\partial t} + \frac{1}{2}\frac{\partial}{\partial t}[\tilde{Y}_2, \tilde{Y}_1] + \frac{1}{6}\left[\left[\frac{\partial \tilde{Y}_1}{\partial t}, \tilde{Y}_1\right], \tilde{Y}_1\right]|\psi_0\rangle \tag{9.47c}$$

At this point we remember that there was some freedom left in the definition of $W(t)$. In view of Eq. (9.43) $W(t)$ is specified (except for an arbitrary constant) if $Y_0(t)$ is specified; and since Y is determined, the decomposition of Y into Y_0 and $\tilde{Y}$ specifies $W(t)$. Obviously Y_0 is a diagonal operator, the nondiagonal part of $\tilde{Y}$ is hence equal to the nondiagonal part of Y.

$$\tilde{Y}_N = Y_N \tag{9.48}$$

There is hence only a freedom in the choice of $\tilde{Y}_D$. The first possibility is to choose:

$$\tilde{Y}_D = 0; \qquad Y_D = Y_0 = i\hbar\int_{t_0}^{t} W(t')\,dt' \tag{9.48a}$$

For this choice we get:

$$W_1(t) = \langle\psi_0|V|\psi_0\rangle \tag{9.49a}$$

$$W_2(t) = \tfrac{1}{2}\langle\psi_0|[V, \tilde{Y}_1]|\psi_0\rangle \tag{9.49b}$$

$$W_3(t) = \langle\psi_0|\tfrac{1}{2}[[V, \tilde{Y}_1], \tilde{Y}_1] + \tfrac{1}{6}[[[H_0, \tilde{Y}_1], \tilde{Y}_1], \tilde{Y}_1]|\psi_0\rangle$$
$$- i\hbar\langle\psi_0|\frac{1}{2}\frac{\partial}{\partial t}[\tilde{Y}_2, \tilde{Y}_1] + \frac{1}{6}\left[\left[\frac{\partial Y_1}{\partial t}, \tilde{Y}_1\right], \tilde{Y}_1\right]|\psi_0\rangle \tag{9.49c}$$

This is an ad-hoc choice with no clear physical meaning. Nevertheless we have achieved that the expressions for W_1 and W_2 are simpler than those for E_1 and E_2, while W_3 is rather more complicated than E_3.

A second possible choice is to fix the $\tilde{Y}_{kD}$ so that $W(t) = E(t)$. This means that we satisfy:

$$0 = \left\langle\psi\middle|\frac{\partial\psi}{\partial t}\right\rangle - \left\langle\frac{\partial\psi}{\partial t}\middle|\psi\right\rangle = 2\langle\psi_0|\frac{\partial \tilde{Y}}{\partial t} + \frac{1}{2}\left[\frac{\partial \tilde{Y}}{\partial t}, \tilde{Y}\right]$$
$$+ \frac{1}{6}\left[\left[\frac{\partial \tilde{Y}}{\partial t}, \tilde{Y}\right], Y\right] + \cdots|\psi_0\rangle \tag{9.50}$$

to all powers of λ. Then one need not worry at all about $W(t)$ and consider $E(t)$ as the relevant energy.

Other choices of $W(t)$ based on more physical arguments will be made in the following section.

9.7. The time-dependent Hellmann–Feynman theorem

It is easy to see that for a wave function Ψ which is a solution of the time-dependent Schrödinger equation (9.1a) the following relation holds

(for an arbitrary parameter μ):

$$\frac{\partial}{\partial\mu}\langle\Psi|H|\Psi\rangle=\langle\Psi|\frac{\partial H}{\partial\mu}|\Psi\rangle+i\hbar\left\{\left\langle\frac{\partial\Psi}{\partial\mu}\middle|\frac{\partial\Psi}{\partial t}\right\rangle-\left\langle\frac{\partial\Psi}{\partial t}\middle|\frac{\partial\Psi}{\partial\mu}\right\rangle\right\} \tag{9.51a}$$

One may refer to this as a time-dependent Hellmann–Feynman theorem [38]. However, it has little to do with the Hellmann–Feynman theorem for a Ψ which is a solution of the time-independent Schrödinger equation:

$$\frac{\partial E}{\partial\mu}=\frac{\partial}{\partial\mu}\langle\Psi|H|\Psi\rangle=\langle\Psi|\frac{\partial H}{\partial\mu}|\Psi\rangle \tag{9.51b}$$

In fact the second term on the right-hand side of Eq. (9.51a) can be much larger than the first one, as it is, e.g., possible for H time-independent, but Ψ a non-stationary state. In order to have something analogous to Eq. (9.51b) we should refer to a wave function ψ that is as close as possible to the eigenfunction of a stationary state. In fact, consider the pseudoeigenfunction ψ which is a solution of Eq. (9.40) with an appropriately chosen $W(t)$. Then in view of Eqs. (9.42) and (9.40) for $\partial\psi/\partial\lambda$ orthogonal to Ψ:

$$\begin{aligned}\frac{\partial W(t)}{\partial\mu}-\langle\psi|\frac{\partial H}{\partial\mu}|\psi\rangle&=\left\langle\frac{\partial\psi}{\partial\mu}\middle|H|\psi\right\rangle+\langle\psi|H\left|\frac{\partial\psi}{\partial\mu}\right\rangle\\&-\frac{i\hbar}{2}\frac{\partial}{\partial\mu}\left\{\left\langle\psi\middle|\frac{\partial\psi}{\partial t}\right\rangle-\left\langle\frac{\partial\psi}{\partial t}\middle|\psi\right\rangle\right\}=\frac{i\hbar}{2}\frac{\partial}{\partial t}\left\{\left\langle\frac{\partial\psi}{\partial\mu}\middle|\psi\right\rangle-\left\langle\psi\middle|\frac{\partial\psi}{\partial\mu}\right\rangle\right\}\end{aligned} \tag{9.51c}$$

In order to achieve that the Hellmann–Feynman theorem holds in the same form (Eq. (9.51b)) as in the time-independent case (just with E replaced by $W(t)$ we must choose W so that the r.h.s. of Eq. (9.51c) vanishes. It is somewhat tedious but elementary to show that the same condition of Eq. (9.51c) also guarantees the Hellmann–Feynman theorem if ψ is not solution of Eq. (9.40) but obtained from the stationary condition of Eq. (9.41). Of course, we must then express the λ, μ, and t dependence of ψ in terms of $\tilde{Y}$. If we assume again that $H=H_0+\lambda V$ with H_0 time-independent and ψ_0 the time-independent wave function which makes $\langle\psi_0|H|\psi_0\rangle$ stationary, then vanishing of the r.h.s. of Eq. (9.51c) means:

$$\frac{\partial}{\partial t}\langle\psi_0|\frac{\partial\tilde{Y}}{\partial\mu}+\frac{1}{2}\left[\frac{\partial\tilde{Y}}{\partial\mu},\tilde{Y}\right]+\frac{1}{6}\left[\left[\frac{\partial\tilde{Y}}{\partial\mu},\tilde{Y}\right],\tilde{Y}\right]+\cdots|\psi_0\rangle=0 \tag{9.52}$$

Let us first consider the simplest case, namely that μ is identical with the perturbation parameter λ, and that we want to satisfy Eq. (9.52) at $\lambda=0$. We then have:

$$0=\left\{\frac{\partial}{\partial t}\langle\psi_0|\frac{\partial\tilde{Y}}{\partial\lambda}|\psi_0\rangle\right\}_{\lambda=0}=\frac{\partial}{\partial t}\langle\psi_0|\tilde{Y}_1|\psi_0\rangle;\qquad W_1(t)=\langle\psi_0|V|\psi_0\rangle=E_1(t) \tag{9.53}$$

i.e. at $\lambda=0$ the Hellmann–Feynman theorem holds even for W replaced by E. Let us now identify again λ with μ but require that Eq. (9.52) holds at an arbitrary value of λ. This means:

$$0=\frac{\partial}{\partial t}\langle\psi_0|Y_1+2\lambda Y_2+\lambda^2\{3Y_3+[Y_2,Y_1]+\tfrac{1}{2}[Y_1,Y_2]\}+\cdots|\psi_0\rangle \tag{9.54a}$$

Equating all powers of λ to 0 results in addition to Eq. (9.53) in:

$$0 = \frac{\partial}{\partial t}\langle\psi_0|\tilde{Y}_2|\psi_0\rangle; \qquad W_2 = \tfrac{1}{2}\langle\psi_0|[V, \tilde{Y}_1]|\psi_0\rangle \tag{9.54b}$$

$$0 = \frac{\partial}{\partial t}\langle\psi_0|3\tilde{Y}_3 + \tfrac{1}{2}[\tilde{Y}_2, \tilde{Y}_1]|\psi_0\rangle$$

$$W_3(t) = \langle\psi_0|\tfrac{1}{2}[[V, \tilde{Y}_1], \tilde{Y}_1] + \tfrac{1}{6}[[[H_0, \tilde{Y}_1], \tilde{Y}_1], \tilde{Y}_1]|\psi_0\rangle$$

$$- i\hbar\langle\psi_0|\frac{1}{3}\frac{\partial}{\partial t}[\tilde{Y}_2, \tilde{Y}_1] + \frac{1}{6}\left[\left[\frac{\partial\tilde{Y}_1}{\partial t}, \tilde{Y}_1\right], \tilde{Y}_1\right]|\psi_0\rangle \tag{9.54c}$$

The results agree with Eq. (9.49) for W_1 and W_2, while there is a slight difference for W_3.

Let us finally regard λ and μ as independent. This means that we have to apply double perturbation theory for the Hamiltonian:

$$H = H_0 + \lambda V + \mu\Omega \tag{9.55}$$

Using the expansion:

$$\tilde{Y}(t) = \sum_{k,l}\lambda^k\mu^l\tilde{Y}_{kl}(t) \tag{9.56}$$

like in Sect. 5 – and a similar expansion for $W(t)$ – we get from Eq. (9.52):

$$\frac{\partial}{\partial t}\langle\psi_0|\tilde{Y}_{01}|\psi_0\rangle = 0 \tag{9.57a}$$

$$\frac{\partial}{\partial t}\langle\psi_0|\tilde{Y}_{11} + \tfrac{1}{2}[\tilde{Y}_{01}, \tilde{Y}_{10}]|\psi_0\rangle = 0 \tag{9.57b}$$

$$\frac{\partial}{\partial t}\langle\psi_0|\tilde{Y}_{21} + \tfrac{1}{2}[\tilde{Y}_{01}, \tilde{Y}_{20}] + \tfrac{1}{2}[\tilde{Y}_{11}, \tilde{Y}_{10}] + \tfrac{1}{6}[[\tilde{Y}_{01}, \tilde{Y}_{10}], \tilde{Y}_{10}]|\psi_0\rangle = 0 \tag{9.57c}$$

$$W_{01} = \langle\psi_0|\Omega|\psi_0\rangle \tag{9.58a}$$

$$W_{11} = \langle\psi_0|[\Omega, \tilde{Y}_{10}]|\psi_0\rangle$$

$$= \{\psi_0|[V, \tilde{Y}_{01}]|\psi_0\rangle - i\hbar\frac{\partial}{\partial t}\langle\psi_0|[\tilde{Y}_{10}, \tilde{Y}_{01}]|\psi_0\rangle \tag{9.58b}$$

$$W_{21} = \tfrac{1}{2}\langle\psi_0|[[V, \tilde{Y}_{10}], \tilde{Y}_{01}] + [V, \tilde{Y}_{11}] + [[V, \tilde{Y}_{01}], \tilde{Y}_{10}] + [[\Omega, \tilde{Y}_{10}], \tilde{Y}_{10}]|\psi_0\rangle$$

$$+ \tfrac{1}{6}\langle\psi_0|[[[H_0, \tilde{Y}_{10}], \tilde{Y}_{10}], \tilde{Y}_{01}] + [[[H_0, \tilde{Y}_{10}], \tilde{Y}_{01}], \tilde{Y}_{10}]$$

$$+ [[[H_0, \tilde{Y}_{01}], \tilde{Y}_{10}], \tilde{Y}_{10}]|\psi_0\rangle - \frac{i\hbar}{2}\langle\psi_0|\left[\left[\frac{\partial\tilde{Y}_{10}}{\partial t}, \tilde{Y}_{01}\right], \tilde{Y}_{10}\right]|\psi_0\rangle \tag{9.58c}$$

Obviously the Hellmann–Feynman theorem is not symmetric in λ and μ. In fact, in deriving Eqs. (9.57, 58) we have required that Eq. (9.52) is satisfied for $\mu = 0$ but λ arbitrary. If we had required $\lambda = 0$ and μ arbitrary we would have got Eq. (9.58b) with Ω exchanged with V and $\tilde{Y}_{10}$ with $\tilde{Y}_{01}$.

This observation gives a partial answer to the question whether there is a quantity which becomes stationary if Frenkel's principle of Eq. (9.1b) is satisfied. The pseudoenergy $W(t)$ is, in fact, a candidate. If it were possible to satisfy Eq. (9.52) for all possible parametrizations of $\tilde{Y}$, the pseudoenergy $W(t)$ would in

fact be stationary. However, since this is not possible, $W(t)$ is not the desired quantity, and it does not seem possible to find a better one.

We shall see in the next section that for Hamiltonians periodic in time a genuine stationarity principle can be formulated. It involves time-integration over a full period. It is rather obvious that this time integral of Eq. (9.52) vanishes if $\tilde{Y}$ is periodic in time with the same periodicity as the Hamiltonian (and this independently of μ).

9.8. Floquet states

If a Hamiltonian H is *periodic in time* with period τ, then the time-dependent Schrödinger equation has solutions of the form [39–46]:

$$\Psi = \exp\left\{\frac{\tilde{E}}{i\hbar}t\right\}\psi \tag{9.59}$$

with the 'quasi energy' $\tilde{E}$ [40] time-independent, while ψ is periodic in time with the same periodicity as H. States like Eq. (9.59) are referred to as Floquet states [39], since their existence is closely related to Floquet's theorem [47] well known in the theory of differential equations [48]. Other proposed names are 'quasi-periodic' states [42] or 'steady' states [43]. Floquet states play a central role in the theory of the interaction of atoms or molecules with strong laser fields. If H is of the form of Eq. (9.6) it can be shown that in the limit $\lambda \to 0$ the Floquet states converge to the stationary states [43]. This implies that the perturbation theory used here and outlined in Sects. 9.2 and 9.5 leads automatically to Floquet states for *periodic perturbation*, although this name has not been referred to, and although we have not explicitly used wave functions of the form of Eq. (9.59).

Let us limit our attention to Hamiltonians periodic in time. The wave function of Eq. (9.59) is, of course, a special case of Eq. (9.39), just with time-independent $\tilde{E} = W(t)$. Since different choices of $W(t)$ mean a different decomposition of Y_D into Y_0 and $\tilde{Y}_D$ (see Eq. (9.43)), the Floquet form of Eq. (9.59) of the wave function corresponds to a particular definition of $\tilde{Y}_D$ that is determined by the condition

$$0 = \frac{\partial \tilde{E}}{\partial t} = \frac{\partial}{\partial t}\left\{\langle\psi|H|\psi\rangle - \frac{i\hbar}{2}\left[\left\langle\psi\middle|\frac{\partial\psi}{\partial t}\right\rangle - \left\langle\frac{\partial\psi}{\partial t}\middle|\psi\right\rangle\right]\right\} \tag{9.60}$$

We want to study this order by order in perturbation theory and start from the general expressions (9.47a). In order that Eq. (9.47a) is independent of time, the time dependencies of the two terms on the r.h.s. of Eq. (9.47a) must cancel. For perturbations periodic in time this means that the Fourier components for $k \neq 0$ should compensate each other.

We introduce the Fourier basis:

$$f_k(t) = e^{ik\omega t}; \quad \omega = 2\pi/\tau \tag{9.61}$$

and require

$$\langle\!\langle\psi_0|Vf_k|\psi_0\rangle\!\rangle = i\hbar\langle\!\langle\psi_0|\frac{\partial\tilde{Y}_1}{\partial t}f_k|\psi_0\rangle\!\rangle; \quad k \neq 0 \tag{9.62}$$

where the double bracket notation, following Sambe [43], means a scalar product

in an extended Hilbert space which implies additional time averaging:

$$\langle\!\langle \psi | \Omega | \psi \rangle\!\rangle = \frac{1}{\tau} \int_{t_0}^{t_0+\tau} \langle \psi | \Omega | \psi \rangle \, dt \tag{9.63}$$

The freedom left in the definition of $\tilde{Y}_{1D}$ is used so that:

$$\langle\!\langle \psi_0 | \frac{\partial \tilde{Y}_1}{\partial t} | \psi_0 \rangle\!\rangle = \frac{1}{\tau} \langle \psi_0 | Y_1 | \psi_0 \rangle \Big|_{t_0}^{t_0+\tau} = 0; \qquad \tilde{E}_1 = \langle\!\langle \psi_0 | V | \psi_0 \rangle\!\rangle \tag{9.64}$$

From Eq. (9.47b) we take in the same way:

$$\tfrac{1}{2} \langle\!\langle \psi_0 | [V, \tilde{Y}_1] f_k | \psi_0 \rangle\!\rangle = i\hbar \langle\!\langle \psi_0 | \frac{\partial \tilde{Y}_2}{\partial t} f_k | \psi_0 \rangle\!\rangle; \quad k \neq 0 \tag{9.65a}$$

$$\langle\!\langle \psi_0 | \frac{\partial Y_2}{\partial t} | \psi_0 \rangle\!\rangle = 0; \qquad \tilde{E}_2 = \langle\!\langle \psi_0 | [V, \tilde{Y}_1] | \psi_0 \rangle\!\rangle \tag{9.65b}$$

and we similarly obtain from Eq. (9.47c):

$$\tilde{E}_3 = \langle\!\langle \psi_0 | \tfrac{1}{2} [[V, \tilde{Y}_1], \tilde{Y}_1] + \tfrac{1}{6} [[[H_0, \tilde{Y}_1], \tilde{Y}_1], Y_1] - \frac{i\hbar}{6} \left[\left[\frac{\partial \tilde{Y}_1}{\partial t}, \tilde{Y}_1 \right], \tilde{Y}_1 \right] | \psi_0 \rangle\!\rangle \tag{9.65c}$$

One can achieve that $\tilde{E}$ is time-independent to any order in λ, by imposing conditions on $\tilde{Y}_D$. Having guaranteed this, we can, of course, replace $\tilde{E}$ by its time average, remembering that $\tilde{E}$ is a special case of $W(t)$, i.e. using Eq. (9.42):

$$\begin{aligned} \tilde{E} &= \frac{1}{\tau} \int_{t_0}^{t_0+\tau} \tilde{E} \, dt = \frac{1}{\tau} \int_{t_0}^{t_0+\tau} \left\{ \langle \psi | H | \psi \rangle - \frac{i\hbar}{2} \left[\left\langle \psi \middle| \frac{\partial \psi}{\partial t} \right\rangle - \left\langle \frac{\partial \psi}{\partial t} \middle| \psi \right\rangle \right] \right\} dt \\ &= \langle\!\langle \psi | H | \psi \rangle\!\rangle - \langle\!\langle \psi | i\hbar \frac{\partial}{\partial t} | \psi \rangle\!\rangle \end{aligned} \tag{9.66}$$

The nice thing with the extended Hilbert space [43] is that $i\hbar \, \partial/\partial t$ now becomes a hermitean operator. The same is true for the operators:

$$\tilde{H} = H - i\hbar \frac{\partial}{\partial t}; \qquad \tilde{H}_0 = H_0 - i\hbar \frac{\partial}{\partial t} \tag{9.67}$$

and we are able to formulate the time average of the generalized Ehrenfest theorem of Eq. (9.2a):

$$\langle\!\langle \psi | [\tilde{H}, X] | \psi \rangle\!\rangle = 0 \tag{9.68a}$$

as the condition that the time-averaged expection value:

$$\tilde{E} = \langle\!\langle \psi | \tilde{H} | \psi \rangle\!\rangle \quad \text{for } \langle\!\langle \psi | \psi \rangle\!\rangle = 1 \tag{9.68b}$$

is stationary with respect to norm-conserving variations.

It is now straightforward to take over the entire formalism of time-independent stationary perturbation theory just replacing the ordinary brackets by double brackets. In particular we get:

$$\tilde{E}_1 = \langle\!\langle \psi_0 | V | \psi_0 \rangle\!\rangle \tag{9.69a}$$

$$\tilde{E}_2 = \tfrac{1}{2} \langle\!\langle \psi_0 | [V, \tilde{Y}_1] | \psi_0 \rangle\!\rangle \tag{9.69b}$$

$$\tilde{E}_3 = \langle\!\langle \psi_0 | \tfrac{1}{2} [[V, \tilde{Y}_1], \tilde{Y}_1] + \tfrac{1}{6} [[[\tilde{H}_0, \tilde{Y}_1], \tilde{Y}_1], \tilde{Y}_1] | \psi_0 \rangle\!\rangle \tag{9.69c}$$

There is also a generalized Hylleraas functional:

$$\tilde{F}(\tilde{Y}_1) = \langle\!\langle \psi_0 | [V, \tilde{Y}_1] + \tfrac{1}{2}[[\tilde{H}_0, \tilde{Y}_1], \tilde{Y}_1] | \psi_0 \rangle\!\rangle \tag{9.70a}$$

with the stationary condition:

$$\langle\!\langle \psi_0 | [V, X] + \tfrac{1}{2}[[\tilde{H}_0, X], \tilde{Y}_1] + \tfrac{1}{2}[[\tilde{H}_0, \tilde{Y}_1], X] | \psi_0 \rangle\!\rangle = 0 \tag{9.70b}$$

that determines the first-order operator $\tilde{Y}_1$ (or rather its non-diagonal part).

Although the theory of this section is – for Hamiltonians periodic in time – equivalent to that outlined in Sects. 9.2 to 9.7, there is nevertheless a fundamental conceptual difference. While in previous subsections of Sect. 9 time was a simple parameter, it is now an integration variable. This implies that the variational group refers to unitary transformations in the 4-dimensional space including time. A complete set of functions of the time coordinate is given by the Fourier basis of Eq. (9.61) and a basis of the Lie algebra with respect to time transformations are the shift operators:

$$P_{kl} = |f_k\rangle\langle f_l| = e^{i(k-l)\omega t} \tag{9.71a}$$

The operators X in this section are of the form:

$$X = x \cdot P \tag{9.71b}$$

with $x \in \mathscr{L}_c$ time-independent and with P of the form of Eq. (9.71a). Let us now study the stationarity condition of Eq. (9.70b) for $\tilde{Y}_1$. Since ψ_0 is time-independent, only those P_{kl} give a non-vanishing condition for which either $k = 0$ or $l = 0$. We hence get from Eq. (9.70b):

$$\langle\!\langle \psi_0 | [V, x] + \tfrac{1}{2}[[H_0, x], \tilde{Y}_1] + \tfrac{1}{2}[[H_0, \tilde{Y}_1], x] - \tfrac{1}{2}\hbar l\omega[x, \tilde{Y}_1] - \frac{i\hbar}{2}\left[\frac{\partial \tilde{Y}_1}{\partial t}, x\right] \Big\} f_l | \psi_0 \rangle\!\rangle = 0 \tag{9.72a}$$

Here x is time-independent. We expand V and Y_1 into their Fourier components:

$$V(t) = \sum_k v_k f_k(t); \qquad \tilde{Y}_1 = \sum_k y_k f_k(t) \tag{9.72b}$$

and insert this into Eq. (9.72a) with the result:

$$\langle \psi_0 | [v_l, x] + [[H_0, x], y_l] + \hbar l\omega[x, y_l] | \psi_0 \rangle = 0 \tag{9.72c}$$

We expand y_l in a basis x_k:

$$y_l = \sum_k b_k^{(l)} x_k \tag{9.73a}$$

such that Eq. (6.46c) becomes (with x in Eq. (6.46) replaced by x_i):

$$-\langle \psi_0 | [v_l, x_i] | \psi_0 \rangle = \sum_k b_k^{(l)} \langle \psi_0 | [[H_0, x_i], x_k] + \hbar l\omega[x_i, x_k] | \psi_0 \rangle = \sum_k (H_{ik} + \hbar l\omega \Delta_{ik}) b_k^{(l)} \tag{9.73b}$$

If, like in Eq. (9.19a), we expand in the Π_k we get:

$$-\langle \psi_0 | [v_l, \Pi_k] | \psi_0 \rangle = [\varepsilon_k f_k + \hbar l\omega] \tilde{\Delta}_{kk} \tag{9.73c}$$

For the special case of Eq. (9.27) we reproduce Eq. (9.30), while $\tilde{E}_2$ is the time average of the E_2 given by Eq. (9.31), i.e.:

$$\tilde{E}_2 = -\tfrac{1}{4} \sum_k |V_k|^2 \tilde{\tilde{\Delta}}_{kk} \varepsilon_k (\varepsilon_k^2 - \hbar^2\omega^2)^{-1} \tag{9.74}$$

For the quasienergy a Hellmann–Feynman theorem holds in the same form as for time-independent states;

$$\frac{\partial \tilde{E}}{\partial \lambda} = \langle\langle \psi | \frac{\partial \tilde{H}}{\partial \lambda} | \psi \rangle\rangle = \langle\langle \psi | \frac{\partial V}{\partial \lambda} | \psi \rangle\rangle \tag{9.75}$$

and this holds for all λ. One might conclude that therefore:

$$\tilde{E}_2 = \frac{1}{\tau} \int_{t_0}^{t_0+\tau} W(t)\, dt \tag{9.76}$$

with the $W(t)$ derived in the last section that satisfies Eq. (9.51a). This is, in fact, true. However, as one sees from the time-average of Eq. (9.47) this holds for all possible $W(t)$. A Hellmann–Feynman theorem in the time-average is a much weaker condition than this theorem for all t.

9.9. Time-dependent properties, in particular polarizabilities and hyperpolarizabilities

One of the most important applications of time-dependent perturbation theory in quantum chemistry is the calculation of frequency-dependent polarizabilities. We have now a look at them in the framework of stationary time-dependent perturbation theory.

Let the operator Ω describe a property and $\psi = \exp(\tilde{Y})\psi_0$ be a time-dependent pseudoeigenfunction in the sense of Eq. (9.40) (with ψ_0 time-independent) corresponding to a Hamiltonian $H = H_0 + \lambda V(t)$, then:

$$\begin{aligned} \langle \psi | \Omega | \psi \rangle &= \langle \psi_0 | \Omega + [\Omega, \tilde{Y}] + \tfrac{1}{2}[[\Omega, \tilde{Y}], \tilde{Y}] + \cdots | \psi_0 \rangle \\ &= \langle \psi_0 | \Omega + \lambda[\Omega, \tilde{Y}_{10}] + \lambda^2\{[\Omega, \tilde{Y}_{20}] + \tfrac{1}{2}[[\Omega, \tilde{Y}_{10}], \tilde{Y}_{10}]\} + \cdots | \psi_0 \rangle \end{aligned} \tag{9.77}$$

where we have used the notation of the double perturbation expansion of Eq. (5.3), although so far we only consider the single perturbation parameter λ. If we have cared to satisfy the conditions of Eq. (9.55) that a time-dependent Hellmann–Feynman theorem holds, we get alternatively:

$$\langle \psi | \Omega | \psi \rangle = \left\{ \frac{\partial}{\partial \mu} W(t) \right\}_{\mu = 0} \tag{9.78}$$

i.e., as suggested by Rice and Handy [33], we can evaluate time-dependent properties as pseudoenergy derivatives.

On the other hand for periodic perturbations (and we have these in mind) the Floquet-state formalism outlined in Sect. 9.8 is much simpler and more powerful than that in terms of pseudoenergies as explained in Sects. 9.6 and 9.7, and it appears recommended to use this formalism. Of course, $\tilde{E}$ is now time-independent, and time-dependent properties cannot be obtained as derivatives of $\tilde{E}$ even if a Hellmann–Feynman theorem holds. Fortunately a discussion of the time-dependence can now be replaced by a discussion of the various Fourier components.

In the pseudoeigenstate formalism the dipole moment (of a molecule in a time-dependent external field) is the expectation value of the (time-independent) dipole operator. We choose H in the form of Eq. (9.33a) with V the external

oscillating field and Ω as a static field and we get:

$$\langle \Omega \rangle = \frac{\partial W}{\partial \mu} = W_{01} + \lambda W_{11} + \lambda^2 W_{21} + \cdots \tag{9.79}$$

with W_{01} the permanent dipole moment (component) and W_{11} the respective component of the frequency-dependent polarizability.

If V has the frequency ω, it contains the Fourier components $k = +1$ and $k = -1$, then W_{11} will have the same components ± 1, while W_{21} will contain $0, \pm 2$ etc. To obtain all these components we can as well take the time-average of W, i.e. $\tilde{E}$ and replace Ω by the operator of a dipole field oscillating with the frequency $k\omega$, i.e. we consider the Hamiltonian:

$$H + \lambda V + \mu_k \Omega \, e^{ik\omega}$$

and get then for the kth Fourier component of the induced dipole moment:

$$\langle \Omega \rangle_k = \frac{\partial \tilde{E}}{\partial \mu_k}$$

10. Quasidegenerate stationary perturbation theory

The conventional perturbation theory breaks down when the state to be perturbed is degenerate and if this degeneracy is removed by the perturbation. In the sum-over-states formulation zero energy denominators would appear. In our stationary perturbation theory the linear system from which the perturbing operator is evaluated, becomes singular and has no solution. In the case of near-degeneracy, the conventional perturbation theory is still applicable, but small energy denominators appear and convergence of the perturbation series becomes unlikely.

What we basically want to do in perturbation theory for operators in a matrix representation is (**a**) to find a transformation $U(\lambda)$ that transforms a Hamiltonian $H(\lambda)$ to a diagonal operator $L(\lambda)$:

$$L(\lambda) = U^{-1}(\lambda) H(\lambda) U(\lambda) \tag{10.1}$$

and (**b**) to construct $U(\lambda)$ as well as $L(\lambda)$ as power series in λ. Step (**b**) fails in the degenerate or quasidegenerate case. What one can do, however, and this is the essential feature of quasidegenerate perturbation theory, is to renounce on full diagonalization in Eq. (10.1) and to be satisfied with block diagonalization. Say H has the dimension n, and we achieve that L consists of two diagonal blocks of dimension d and $n - d$ respectively, with d sufficiently small and such that our desired eigenvalue $E(\lambda)$ is contained in the $d \times d$ block. Then it is trivial to get $E(\lambda)$ from $L(\lambda)$ after one has constructed $L(\lambda)$ and $U(\lambda)$ as expansions in powers of λ. The subspaces must be selected so that all unperturbed states with energies close to the one to be studied are in the small space of dimension d, called the *model space*, such that in the perturbation expansion no small denominators arise. The diagonal block of $L(\lambda)$ in the model space is referred to as the effective Hamiltonian.

We follow here the formalism preferred by the present author [16] that is equivalent to more traditional presentations of quasidegenerate perturbation

theory. One rewrites Eq. (10.1) as:

$$UL = HU; \qquad L = H_0 + \Delta L; \qquad \Delta L = \sum_{k=1}^{\infty} \lambda^k L_k \tag{10.2a}$$

$$[H_0, U] = U\,\Delta L - \lambda V U; \qquad U = 1 + \sum_{k=1}^{\infty} \lambda^k U_k \tag{10.2b}$$

and expands to powers of λ:

$$[H_0, U_1] = L_1 - V \tag{10.3a}$$

$$[H_0, U_2] = L_2 + U_1 L_1 - V U_1 \text{ etc.} \tag{10.3b}$$

The condition that W should be a 'nondiagonal' and L a 'diagonal' operator (where 'diagonal' here means 'block-diagonal', and 'nondiagonal' that the diagonal blocks vanish) leads to:

$$L_1 = V_D; \qquad [H_0, U_1] = -V_N \tag{10.4a}$$

$$L_2 = (VU_1)_D - (U_1 L_1)_D; \qquad [H_0, U_2] = (U_1 L_1)_N - (VU_1)_N \tag{10.4b}$$

We now want to combine *quasidegenerate* perturbation theory with the concept of *stationary* perturbation theory. Rather than to solve the Schrödinger equation (or to diagonalize a Hamiltonian in the sense of Eq. (10.1)) we want to satisfy a stationarity condition. In the non-degenerate case we care to make the expectation value $\langle \Psi | H | \Psi \rangle$ stationary (for Ψ normalized). The straightforward generalization is to make all matrix elements of the effective Hamiltonian $L(\lambda)$ in the model space stationary, i.e. to require:

$$\delta \langle \Psi_\mu | H | \Psi_\nu \rangle = \delta \langle \phi_\mu | U^\dagger H U | \phi_\nu \rangle = 0, \quad \text{for all } \phi_\mu \text{ in the 'model space'}$$

$$U^\dagger U = 1; \qquad \langle \phi_\mu | H_0 | \phi_\nu \rangle = \delta_{\mu\nu} E_{0\mu} \tag{10.5}$$

The transformation U only removes the coupling between model states and other states, not within the model space. While in traditional quasidegenerate perturbation theory U need not be unitary, the unitary choice is more consistent in the present context. With the exponential formulation $W = \exp \sigma$ we get a set of Brillouin (hypervirial) relations:

$$\langle \Psi_\mu | [H, X] | \Psi_\nu \rangle = 0 \tag{10.6}$$

There are many more conditions than in the non-degenerate case, and one may wonder whether it is possible to satisfy all of them. In fact, let N be the dimension of the Lie algebra of the variational group and d the dimension of the model space, then there are $N \cdot d^2$ conditions, but the number of parameters to express σ and hence W in terms of the X_p is only N. So there are $N \cdot d^2$ equations for only N unknowns. A solution is only possible if at least $N \cdot (d^2 - 1)$ conditions of the type of Eq. (10.6) are satisfied trivially.

We expand Eq. (10.6) first in terms of the Hausdorff expansion for $U = \exp Y$ and then in powers of the perturbation parameter λ for $Y = \sum_k \lambda^k Y_k$:

$$\langle \phi_\mu | [H, X] + [[H, X], Y] + \tfrac{1}{2}[[[H, X], Y], Y] + \cdots | \phi_\nu \rangle = 0 \tag{10.7}$$

$$\langle \phi_\mu | [H_0, X] | \phi_\nu \rangle = 0 \tag{10.8a}$$

$$\langle \phi_\mu | [V, X] + [[H_0, X], Y_1] | \phi_\mu \rangle = 0 \tag{10.8b}$$

exactly like in Sect. 4, just with general matrix elements in the model space rather than simple expectation values.

Take as an example the linear variational group with the basis operators:

$$X_{kl} = |\phi_k\rangle\langle\phi_l| \tag{10.9a}$$

such that the model functions ϕ_μ are included in the basis. Operators of the type $X_{\mu\nu}$ have to be excluded in order to satisfy Eq. (10.3). On the other hand, an X_{kl} in which neither k nor l refers to a function in the model space, gives trivially the value 0 for the Brillouin condition of Eq. (10.8a). A non-redundant basis of antihermitean shift operators hence consists of:

$$\tilde{X}_{\mu k} = |\phi_\mu\rangle\langle\phi_k| - |\phi_k\rangle\langle\phi_\mu|, \quad \phi_k \text{ not in the model space} \tag{10.9b}$$

These are $d \cdot (n-d)$ operators where n is the dimension of the basis $\{\phi_k\}$. One expands Y just in these and the number of non-trivial Brillouin conditions is the same. In fact:

$$\langle\phi_\mu|[H_0, \tilde{X}_{\varrho k}]|\phi_\nu\rangle = h_{\mu\varrho}\delta_{k\nu} - \delta_{\mu\varrho}h_{k\nu} - h_{\mu k}\delta_{\varrho\nu} + \delta_{\mu k}h_{\varrho\nu} \tag{10.10a}$$

$$h_{\mu\varrho} = \langle\phi_\mu|H_0|\phi_\varrho\rangle \tag{10.10b}$$

Since $k = \nu$ or $k = \mu$ is excluded, Eq. (10.10) is trivially non-zero only if $\varrho = \mu$ or $\varrho = \nu$:

$$\langle\phi_\mu|[H, \tilde{X}_{\mu k}]|\phi_\nu\rangle = -h_{k\nu} - h_{\nu k}\delta_{\mu\nu} \tag{10.11a}$$

$$\langle\phi_\mu|[H, \tilde{X}_{\nu k}]|\phi_\nu\rangle = -\delta_{\mu\nu}h_{k\mu} - h_{\mu k} \tag{10.11b}$$

These conditions are satisfied if all $h_{\mu k} = 0$ (k not in the model space), i.e. if the matrix representation of H_0 in the basis $\{\phi_k\}$ is block-diagonal.

We expand Y_1 in terms of the $\tilde{X}_{\mu k}$:

$$Y_1 = \sum_{\mu,k} c_{\mu k}\tilde{X}_{\mu k} \tag{10.12}$$

and get from Eq. (10.8b):

$$\langle\phi_\mu|[V, \tilde{X}_{\varrho l}]|\phi_\nu\rangle + \sum_{\sigma k} c_{\sigma k}\langle\phi_\mu|[[H_0, \tilde{X}_{\varrho l}], \tilde{X}_{\sigma k}]|\phi_\nu\rangle = 0 \tag{10.13}$$

This is a linear system of equations much like in the non-degenerate case. For the Hessean we get:

$$\langle\phi_\mu|[[H_0, \tilde{X}_{\varrho l}], \tilde{X}_{\sigma k}]|\phi_\nu\rangle = -\delta_{\sigma\nu}\,\delta_{kl}h_{\mu\varrho} + \delta_{\varrho\mu}\,\delta_{\sigma\nu}h_{lk} + \delta_{\sigma\mu}\,\delta_{\varrho\nu}h_{kl} - \delta_{\sigma\mu}\,\delta_{kl}h_{\varrho\nu} \tag{10.14}$$

If we choose the ϕ_p such that they diagonalize the matrix $\underline{h}$, the matrix elements of Eq. (10.14) only contain energy differences $\varepsilon_\mu - \varepsilon_k$ between unperturbed states within and outside the model space.

Unfortunately it does not appear possible to define a quasidegenerate Hartree–Fock theory as the first step (for $\lambda = 0$) of a quasidegenerate coupled Hartree–Fock theory. Let us consider a set of quasidegenerate single-determinant states ϕ_μ, let us choose the variational group as usual in Hartree–Fock theory and let us try to satisfy the Brillouin conditions:

$$\langle\phi_\mu|[H_0, a_q^p]|\phi_\nu\rangle = 0 \tag{10.15}$$

but only for those p, q for which:

$$\langle\phi_\mu|a_q^p|\phi_\nu\rangle = 0 \tag{10.16}$$

Let φ_i be a spin-orbital occupied in both ϕ_μ and ϕ_ν and φ_a spin-orbitals neither

occupied in ϕ_μ nor in ϕ_ν, and let φ_x be occupied in ϕ_μ, but φ_y occupied in ϕ_ν. Then a_y^x should be excluded, because it does not satisfy Eq. (10.16). We must hence restrict the a_q^p to a_i^a or a_a^i. Then, however, we have many more conditions to satisfy than coefficients available in the expansion of Y_1.

It is easy to see that stationarity of a single Hartree–Fock state gives just as many Brillouin conditions as there are operators a_i^a. So we cannot make various states stationary with one set of a_i^a. If we want to do so we need excitation operators which discriminate between the model states on which they act. This can be achieved by including 'conditional' excitation operators like a_{ix}^{ax} (also called operators with spectator lines). The notation is used that i, j refer to spin orbital occupied in all model states, a, b in no model state, x, y in some model states. The problem then arises that the Lie-algebraic property of the a_i^a operators is lost, if one includes a_{ix}^{ax} etc. In particular we no longer have a variational group of one-particle transformations. Another way to phrase this is that quasidegenerate perturbation theory in terms of a set of stationary Slater determinants is not possible.

One has to be less demanding and one may, e.g., require that the sum of the energies of the considered states is stationary (average-of-configuration SCF), or that a particular linear combination of Slater determinants is stationary (MC-SCF).

There is, nevertheless, the possibility for stationary quasidegenerate perturbation theory of many-electron systems, namely if one chooses H_0 as the bare-nuclear Hamiltonian without electron interaction. Then just as many Brillouin conditions of the type of Eq. (10.15) are trially satisfied and one has as many unknown parameters as non-trivial conditions. On these lines a stationary formulation of quasidegenerate many-body perturbation theories is possible (see paper II of this series).

11. Concluding remarks

Although a few results of this paper are probably either new or appear in a new context, the main interest of this study is on one side the unified and consistent formulation of stationary perturbation theory in terms of the variational group (or some of its subgroups) and on the other side the power of this formalism for applications to be described in forthcoming papers [4, 6], or already briefly pointed out [7]. Two methods known for the evaluation of molecular properties of molecules, called IGLO [5] and LORG [49] respectively, look at first glance very similar. However, only IGLO satisfies a Hylleraas-type stationary condition for the second-order energy [6, 7].

We have not worried here about mathematical problems of perturbation theory [1], e.g. the conditions for its convergence and the existence (square integrability) of the perturbed wave function [7]. Somehow the present formalism is simpler than the exact theory, because the original problem is reduced to one in a finite number of dimensions (though an originally linear problem may become nonlinear) and some difficulties are avoided in the chosen model, but they may play a role when it comes to comparing approximate with exact solutions.

We have mostly been concerned with ground states, but the formalism is applicable to excited states as well, and even – as mentioned – to ensembles.

It has sometimes been claimed that the unitary formulation of perturbation theory is more complicated than that in intermediate normalization and that

tedious manipulations are required in the unitary formulation to arrive at the final results. (For a comparison of the two possibilities in Fock space perturbation theory see e.g. Ref. [16].) In the present context this is certainly not the case. Of course, lengthy expressions arise if one goes to high orders, but high-order perturbation theory is tedious anyhow and is not our present concern.

Acknowledgements. The author thanks Ch. van Wüllen and J. Noga for discussions on this topic and for critical reading of this manuscript. He is grateful to three unknown referees (especially one of these) for useful and constructive comments. Financial support by Deutsche Forschungsgemeinschaft and Fond der Chemie is gratefully acknowledged.

References

1. Kato T (1966) Perturbation theory of linear operators. Springer, Berlin; see also the appendix by Palmer TW to the review by Hirschfelder JO, Byers-Brown W, Epstein STe (1964) Adv Quant Chem 1:256
2. Epstein ST (1974) The variation method in quantum chemistry. Academic Press, NY
3. Hurley AC (1954) Proc Roy Soc A 226:179
4. Kutzelnigg W, paper II of this series, to be published
5. Kutzelnigg W (1980) Israel J Chem 19:193; Schindler M, Kutzelnigg W (1982) J Chem Phys 76:1919
6. Wüllen Ch van, Kutzelnigg W, to be published
7. Kutzelnigg W (1989) J Mol Struct Theochem 202:11
8. Primas H (1961) Helv Phys Acta 34:331; Primas H (1963) Rev Mod Phys 35:710; see also Goodman GL (1965) J Chem Phys 43:S268
9. Hirschfelder JO (1960) J Chem Phys 33:1462; Epstein ST, Hirschfelder JO (1961) Phys Rev 123:1495; Hirschfelder JO, Coulson CA (1962) J Chem Phys 36:941
10. Fernandez FM, Castro EA (1987) Hypervirial theorems. Lecture Notes in Chemistry, Vol 43, Springer, Berlin
11. Chen JCY (1963) J Chem Phys 39:3167
12. Robinson PD (1964) Proc Roy Soc A283:229
13. Katriel J, Domany E (1973) Chem Phys Lett 21:448
14. Fukutome H, Yamamura M, Nishiyama S (1977) Progr Theor Phys 57:1554; Fukutome H (1981) Progr Theor Phys 65:809
15. Nerbrant PO, Roos B, Sadlej AJ (1979) Int J Quant Chem 15:135; Epstein ST, Sadlej A (1979) Int J Quant Chem 15:147; Sadlej AJ (1983) Int J Quant Chem 23:147, (1982) J Chem Phys 75:320; Diercksen GHF, Roos B, Sadlej A (1981) J Chem Phys 73:1253
16. Kutzelnigg W (1982) J Chem Phys 77:3081, (1984) 80:822, (1985) 82:4166; Kutzelnigg W, Koch S (1983) J Chem Phys 79:4315; Kutzelnigg W (1989) in: Mukherjee D (ed) (1989) Aspects of many-body effects in molecules and extended systems. Lecture Notes in Chemistry, Vol 50, Springer, Berlin, p 35
17. Magnus W (1954) Commun Pure Appl Math 7:649; Wei J, Norman E (1964) Proc Amer Math Soc 15:327, (1963) J Math Phys 4:575; Wolf F (1987) J Phys A 20:L421
18. Kramer P, Saraceno M (1981) Lecture Notes in Physics, Vol 140, Springer, Berlin; Meyer HD, Kučar J, Cederbaum LS (1988) J Math Phys 29:1417
19. Bourbaki N (1961) Elements de mathématique. Algebres de Lie. Hermann, Paris; Jacobson N (1962) Lie Algebras, Interscience, NY; Gilmore R (1974) Lie groups, Lie algebras and some of their applications. Prentice Hall, NY
20. Westenholtz v C (1981) Differential forms in mathematical physics. North-Holland, Amsterdam; Curtis WD, Miller FR (1985) Differential manifolds and theoretical physics. Academic Press, NY
21. Brillouin L (1934) Actualités Sci Ind 159
22. Levy B, Berthier G (1968) Int J Quant Chem 2:307; Levy B (1969) Chem Phys Letters 4:17, (1970) Int J Quant Chem 4:297
23. Kutzelnigg W (1979) Chem Phys Letters 64:383; Kutzelnigg W (1980) Int J Quant Chem 18:3

24. Thouless DJ (1961) The quantum mechanics in many-body systems. Academic Press, NY; Čižek J, Paldus J (1967) J Chem Phys 47:3976; Paldus J, Čižek J (1970) J Chem Phys 52:2919; Fukutome H (1968) Progr Theor Phys 40:998; 1227; Paldus J (1990) in: Carbo R, Klobukowski M (eds) Self-consistent field, theory and applications. Elsevier, Amsterdam, p, 1
25. Dalgarno A, Stewart AL (1956) Proc Roy Soc A 238:269, 276, (1958) 247:245
26. Weiner B, Goszinsky O (1980) Int J Quant Chem 18:1109, (1982) 21:369
27. Frenkel J (1934) Wave mechanics, advanced general theory. Oxford Univ Press, London, p 436
28. Dirac PAM (1926) Proc Roy Soc A 112:673. (1930) Proc Cambridge Phil Soc A 26:376
29. Langhoff PW, Epstein ST, Karplus M (1972) Rev Mod Phys 44:602
30. McLachlan AD (1964) Mol Phys 8:39
31. Heinrichs J (1968) Phys Rev 172:1315, (1968) Chem Phys Lett 2:315
32. Löwdin PO, Mukherjee PK (1972) Chem Phys Lett 14:1
33. Rice JE, Handy NC (1991) J Chem Phys 94:4959; Rice JE, Handy NC, to be published
34. Kutzelnigg W (1991) in: Mukherjee D (ed) Applied many-body methods in spectroscopy and electronic structure. Plenum, NY
35. Bhattacharyya K, Mukherjee D (1986) J Phys A 19:67
36. McLachlan AD, Bell MA (1964) Rev Mod Phys 36:844; Dalgarno A, Victor GA (1966) Proc Roy Soc A17:291
37. Jøgensen P, Simons J (1981) Second quantization-based methods in quantum chemistry. Acad Press, NY
38. Hayes EF, Parr RG (1965) J Chem Phys 43:1831
39. Shirley JH (1965) Phys Rev 138:B979
40. Zel'dovich YaB (1967) Soviet Physics JETP 24:1006
41. Ritus VI (1967) Soviet Physics JETP 24:1041
42. Young RH, Deal WJ, Kestner NR (1969) Mol Phys 17:369
43 Sambe H (1973) Phys Rev A7:2203
44. Salzmann WR (1974) Phys Rev A10:461
45. Howland JH (1989) Ann Inst Henri Poincaré 49:309
46. Breuer HP, Dietz K, Holthaus M (1988) Z. Phys D10:13, (1991) J Phys B 24:1343; Breuer NP, Holthaus M (1989) Z Phys D 11:1, (1991) Annals of Physcs (NY) 211:249
47. Floquet AMG (1883) Ann Ecole Norm Sup 12:47
48. Coddington EQ, Levinson N (1955) Theory of ordinary differential equations. McGraw-Hill, NY, p 78
49. Hansen AE, Boumnan TD (1985) J Chem Phys 82:5035
50. Kutzelnigg W (1990) Z Phys D 15:27
51. Kutzelnigg W (1991) Theoret Chim Acta 80:349

Appendix A

On the Lie algebra of the variational group and its subalgebra

Let $\mathscr{L}$ be a Lie algebra the elements of which are operators that act in a Hilbert space or a Fock space. This algebra is called *real*, if a basis $\{X_k\}$ of $\mathscr{L}$ can be chosen such that all elements of $\mathscr{L}$ are expressible as linear combinations of the X_k with *real* coefficients, otherwise as *complex*. So 'real' or 'complex' refers to the field of numbers upon which the algebra is constructed, not to the elements (operators) themselves. We call an operator X *real* if it transforms a real wave function ψ to a real one and an imaginary ψ to an imaginary one. Conversely an *imaginary* operator transforms a real ψ to an imaginary wave function and vice versa.

In a Lie algebra of operators, the subset of antihermitean operators forms a *real subalgebra*. In fact the commutator of two antihermitean operators is antihermitean. On the other hand, a Lie algebra that consists of hermitean and

antihermitean operators (and linear combinations of these) is necessarily complex, because i times a hermitean operator is antihermitean and vice versa.

Starting from a complex Lie algebra $\mathscr{L}_c$ of operators, we can always construct the real subalgebra $\mathscr{L}_r$ of antihermitean operators.

One can go one step further and consider within $\mathscr{L}_r$ only real operators. Since the commutator of two real operators is again real, the subset $\mathscr{L}_o \subset \mathscr{L}_r$ of real operators is again a Lie algebra. The complementary set of imaginary operators $\mathscr{C}_i \subset \mathscr{L}_r$ is not a Lie algebra, since the commutator of two imaginary operators is real.

If both $\mathscr{L}_c$ and $\mathscr{L}_r$ are associated with a unitary group $\mathscr{G}$ (are 'generating' $\mathscr{G}$), the Lie algebra $\mathscr{L}_o$ generates the real orthogonal group $\mathscr{O} \subset \mathscr{G}$.

Appendix B

The Newton–Raphson procedure for linear variations

We want to solve the iterative linear system of Eq. (2.14a) for the basis of operators of Eq. (3.8). It is recommended to use Eq. (6.1) of the Hessean, rather than Eq. (2.11b). We hence want to solve:

$$D_p^{(n)} + \sum_q H_{pq}^{(n)} c_q^{(n+1)} = 0 \tag{B.1}$$

$$D_p^{(n)} = \langle \Phi^{(n)} | [H, X_p] | \Phi^{(n)} \rangle \tag{B.2a}$$

$$H_{pq}^{(n)} = \langle \Phi^{(n)} | [X_p^\dagger, [H, X_q]] | \Phi^{(n)} \rangle \tag{B.2b}$$

$$\Phi^{(n)} = e^{\sigma^{(n)}} \Phi : \Phi^{(0)} = \Phi \tag{B.3a}$$

$$\sigma^{(n)} = \sum_p c_p^{(n)} X_p ; \quad n \geqslant 1 \tag{B.3b}$$

by iteration. The operator basis consists now of the shift operators $X_{kl} = |\phi_k\rangle\langle\phi_l|$ given by Eq. (3.7). Since each X_{kl} has two labels, each label p, q in Eqs. (B.1–3) has to be replaced by a pair of labels kl or mn. It is recommended to choose the function basis $\{\phi_k\}$ of $\mathscr{H}_N$ such that Φ is a member, say $\Phi = \phi_0$, of this basis and that the basis is orthonormal. Then we have in the first iteration cycle:

$$D_{kl}^{(0)} = D_{kl} = \langle \Phi | [H, X_{kl}] | \Phi \rangle = h_{0k}\delta_{l0} - \delta_{0k} h_{l0} \tag{B.4a}$$

$$\begin{aligned} H_{kl,mn}^{(0)} = H_{kl,mn} = \langle \Phi | [X_{lk}, [H, X_{mn}]] | \Phi \rangle &= \delta_{l0} h_{km} \delta_{n0} \\ - \delta_{l0}\delta_{km} h_{n0} - h_{0m}\delta_{nl}\delta_{k0} &+ \delta_{0m} h_{nl} \delta_{k0} \end{aligned} \tag{B.4b}$$

$$h_{pq} = \langle \phi_p | H | \phi_q \rangle \tag{B.5}$$

The only non-vanishing elements are:

$$D_{0l} = -h_{l0}; \quad l \neq 0; \qquad D_{k0} = h_{0k}; \quad k \neq 0 \tag{B.6a}$$

$$\begin{aligned} H_{k0,m0} &= (h_{km} - \delta_{km} h_{00})(1 - \delta_{k0}) \\ H_{0l,0n} &= (h_{nl} - \delta_{nl} h_{00})(1 - \delta_{l0}) \end{aligned} \tag{B.6b}$$

A non-redundant operator basis hence consists of the operators X_{k0} and X_{0k} $(k \neq 0)$ and we can expand σ as:

$$\sigma = \sum_{k(\neq 0)} c_{k0} X_{k0} + \sum_{k(\neq 0)} c_{0k} X_{0k} \tag{B.7}$$

the coefficients c_{k0} and c_{0k} are obtained from the linear systems of equations:

$$D^*_{k0} + \sum_{l(\neq 0)} H_{k0,l0} c_{l0} = 0 \tag{B.8a}$$

$$D^*_{0k} + \sum_{l(\neq 0)} H_{0k,0l} c_{0l} = 0 \tag{B.8b}$$

Explicitly the system of Eqs. (B.8) reads:

$$h_{k0} + \sum_{l(\neq 0)} h_{kl} c_{l0} - h_{00} c_{k0} = 0 \tag{B.9a}$$

$$-h_{0k} + \sum_{l(\neq 0)} h_{lk} c_{0l} - h_{00} c_{0k} = 0 \tag{B.9b}$$

It is sufficient to solve Eq. (B.9a), because:

$$c_{01} = -c^*_{l0} \tag{B.10}$$

The c_{0l} $(l \neq 0)$ are obtained from a linear system of equations of dimension $N-1$. We choose $c_{00} = 0$ to simplify the summations. We must next construct:

$$\phi^{(n)} = e^{\sigma^{(n)}} \phi_0 \tag{B.11}$$

We note that

$$\sigma^2 \phi_0 = -t^2 \phi_0; \qquad t^2 = \sum_k |c_{k0}|^2 \tag{B.12}$$

$$e^{\sigma} \phi_0 = \cosh \sigma \cdot \phi_0 + \sinh \sigma \cdot \phi_0 = \cos t \cdot \phi_0 + \frac{\sin t}{t} \sigma \phi_0$$

$$= \cos t \cdot \phi_0 + \frac{\sin t}{t} \sum_k c_{k0} \phi_k \tag{B.13}$$

This gives us the ϕ of the next iteration as a linear combination of the basis functions. We can then orthogonalize the basis to the new ϕ and proceed as before.

The transformation to a new basis is a rather-time-consuming step and therefore not recommended in practice. Fortunately this can – in this special case – be completely avoided by a closed summation of the infinite Hausdorff expansion of the Brillouin condition. This is related to an 'even-odd-rule' [16, 50, 51] according to which even powers of σ are diagonal operators (see Eq. (9.16)) and odd powers of σ non-diagonal operators. This manifests itself in Eq. (B.13). One can in fact rewrite the Brillouin condition as (for $\Phi \equiv \phi_0$):

$$0 = \langle \Phi | e^{-\sigma} [H, X] e^{\sigma} | \Phi \rangle = \cos^2 t \Big\{ \langle \Phi | [H, X] | \Phi \rangle$$

$$+ \frac{\tan t}{t} \langle \Phi | [[H, X], \sigma] | \Phi \rangle - \frac{\tan^2 t}{t^2} \langle \Phi | \sigma [H, X] \sigma | \Phi \rangle \Big\} \tag{B.14}$$

If we define:

$$\varrho = \tanh \sigma = \sum_k d_{k0} X_{k0} + \sum_k d_{0k} X_{0k} \tag{B.15a}$$

$$d_{k0} = -d^*_{0k} = (\tan t/t) \cdot c_{k0} \tag{B.15b}$$

the Brillouin condition of Eq. (B.14) can be written as:

$$0 = \langle \Phi | [H, X] | \Phi \rangle + \langle \Phi | [[H, X], \varrho] | \Phi \rangle - \langle \Phi | \varrho [H, X] \varrho | \Phi \rangle \tag{B.16}$$

This expansion in ϱ is finite, while that in powers of σ is infinite. Either expansion is nonlinear and needs to be solved by iteration. But this is much simpler for Eq. (B.16).

If we insert the expansion of Eq. (B.15a) into Eq. (B.16) and use Eq. (B.6) as well as:

$$\langle \Phi | X_{0k}[H, X_{mn}]X_{l0} | \Phi \rangle = h_{km}\delta_{nl} - \delta_{km}h_{nl} \tag{B.17}$$

$$\langle \Phi | \varrho[H, X_{0k}]\varrho | \Phi \rangle = d_{k0} \sum_n d_{0n}h_{n0} \tag{B.18a}$$

$$\langle \Phi | \varrho[H, X_{k0}]\varrho | \Phi \rangle = -d_{0k} \sum_l d_{l0}h_{0l} \tag{B.18b}$$

Eq. (B.16) becomes:

$$h_{k0} + \sum_l h_{kl}d_{l0} - \left(h_{00} + \sum_m d_{0m}h_{m0}\right) d_{k0} = 0 \tag{B.19a}$$

$$h_{0k} - \sum_l h_{kl}d_{0l} + \left(h_{00} + \sum_m d_{k0}h_{0m}\right) d_{0k} = 0 \tag{B.19b}$$

If we ignore the terms bilinear in the d_{k0} or d_{0k}, Eq. (B.19a) becomes identical with Eq. (B.9), i.e. in the first iteration cycle the c_{k0} agree with the d_{k0} and hence ϱ agrees with σ. Of course, the counterpart of Eq. (B.10) holds:

$$d_{0l} = -d_{l0}^* \tag{B.20}$$

The iterative solution of Eq. (B.19) is straightforward via:

$$h_{k0} + \sum_l h_{kl}d_{l0}^{(n+1)} - \left(h_{00} + \sum_m d_{0m}^{(n)}h_{m0}\right) d_{k0}^{(n+1)} = 0 \tag{B.21}$$

This is obviously an iterative approach to the eigenvalue problem:

$$h_{k0} + \sum_l h_{kl}d_{l0} = Ed_{k0} \tag{B.22}$$

with the eigenvector in intermediate normalization. This is easily seen if we consider the expectation value:

$$E = \langle \Phi | e^{-\sigma} H e^{\sigma} | \Phi \rangle = \cos^2 t \langle \Phi | (1 - \varrho)H(1 + \varrho) | \Phi \rangle \tag{B.23}$$

and note that:

$$\cos^2 t = [1 - \langle \Phi | \varrho^2 | \Phi \rangle]^{-1} \tag{B.24}$$

Eq. (B.22) is hence the eigenvalue system that makes E stationary.

If E as given by Eq. (B.23) is stationary with respect to variation of ϱ, an alternative energy expression is:

$$E = \langle \Phi | H | \Phi \rangle + \tfrac{1}{2}\langle \Phi | [H, \varrho] | \Phi \rangle = h_{00} + \sum_m d_{0m}h_{m0} \tag{B.25}$$

which establishes the relation between the non-linear system of Eq. (B.20a) and the eigenvalue system of Eq. (B.22).

There are usually better ways to solve the eigenvalue system of Eq. (B.22) than that via the iteration scheme indicated by Eq. (B.21), but Eq. (B.21) is interesting as it follows directly from the stationarity condition of the expectation value as alternative, but equivalent to the eigenvalue system.