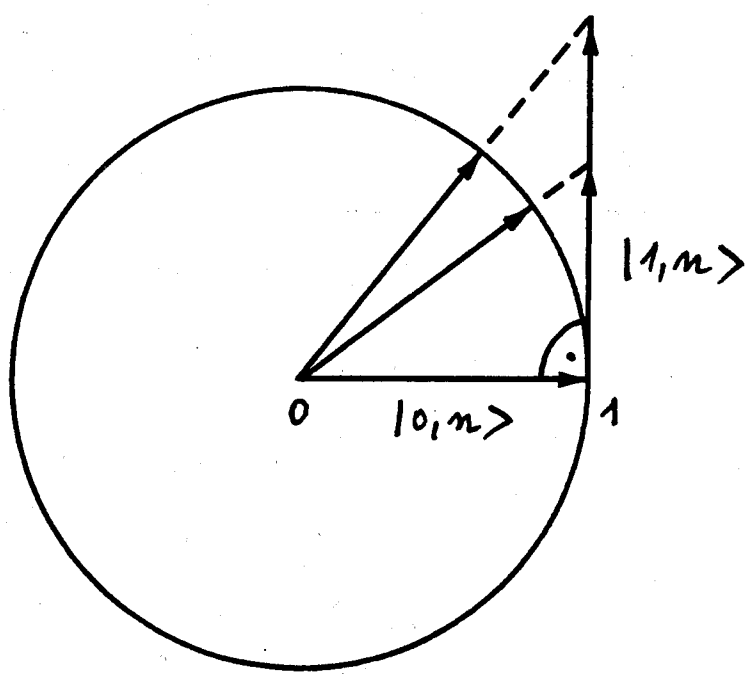
On Rayleigh-Schrödinger perturbations of slightly non-orthogonal states and IPA

ETH Eidgenössische Technische Hochschule Zürich
Ecole polytechnique fédérale de Zurich
Politecnico federale di Zurigo
Swiss Federal Institute of Technology Zurich

IPS Interdisziplinäres Projektzentrum für Supercomputing
Interdisciplinary Project Center for Supercomputing

On Rayleigh-Schrödinger Perturbations of Slightly Non-Orthogonal States and IPA

Gabor G. Groh



November 1993
IPS Research Report No. 93-17

IPS
ETH-Zentrum
CH-8092 Zurich

On Rayleigh-Schrödinger perturbations of slightly non-orthogonal states and IPA

Gabor G. Groh

Swiss Federal Institute of Technology (ETH), Interdisciplinary Project Center for Supercomputing (IPS) and Institute of Energy Technology (IET-LSM), ETH-Zentrum, CH-8092 Zürich, Switzerland, groh@iet.ethz.ch

Abstract

Rayleigh-Schrödinger perturbation theory is reviewed under the practical assumptions that a finite number N of slightly non-orthogonal unperturbed wave functions are used. It is shown that the formalism is internally consistent in first order for arbitrary complex dimensions N (not for odd real dimensions), including the limit $N \to \infty$ in l_2, for Hamiltonians with discrete non-degenerate spectrum. In inverse perturbation analyses (IPA), the sum-over-states representation of the first-order corrections of the wave functions can be used as an indicator of extraneous oscillations which are caused by non-orthogonality.

1 Direct and inverse perturbation analyses

1.1 Review of RS theory

In order to contrast the non-orthogonal case in Section 2, consider first-order perturbation theory (RS) following Schrödinger, Rellich and Kato [5,6,11,12,13]. The problem is to solve the time-independent Schrödinger equation

$$H|n\rangle = E_n|n\rangle \quad (n \in \{0,1,\ldots\}) \tag{1}$$

with almost diagonal Hamiltonian ($0 < \varepsilon \ll 1$):

$$H = H^0 + \varepsilon H^1, \quad \|H^1\| = O(1). \tag{2}$$

The unperturbed problem is assumed to be solved and, the analysis in the following sections will be restricted to the case that the spectrum of H^0 is purely discrete and not (almost) degenerate:

$$H^0|0,n\rangle = E_n^0|0,n\rangle, \quad E_k^0 - E_n^0 = O(1) \quad (k \neq n), \quad (k,n \in \{0,1,\ldots\}). \tag{3}$$

The eigenfunctions $|0,n\rangle$ form an orthonormal basis in the appropriate Hilbert space:

$$\langle 0,k|0,n\rangle = \delta_{kn} \quad (k,n \in \{0,1,\ldots\}). \tag{4}$$

Thus H^0 is diagonal in this basis and the eigenvalues are the Rayleigh quotients: $E_n^0 = \langle 0,n|H^0|0,n\rangle$. The first order correction $\varepsilon|1,n\rangle$ of the nth wave function $|0,n\rangle$ is a solution of the inhomogeneous self-adjoint linear (partial) differential equation:

$$(H^0 - E_n^0)|1,n\rangle = (E_n^1 - H^1)|0,n\rangle, \tag{5}$$

with the definitions (excluding discontinuity for $\varepsilon \to 0$, e.g., the Klauder phenomenon [6]):

$$E_n = E_n^0 + \varepsilon E_n^1 \qquad (|E_n^1| = O(1)), \tag{6}$$

$$|n\rangle = |0,n\rangle + \varepsilon|1,n\rangle \quad (\||1,n\rangle\| = O(1)). \tag{7}$$

By the Fredholm alternative, Eq. (5) admits a solution if and only if the inhomogeneity is orthogonal to the eigenfunction $\langle 0,n|$ of the adjoint operator $(H^0)^\dagger = H^0$ to the eigenvalue E_n^0, leading to Schrödinger's famous formula [13,11,6,12,5]:

$$E_n^1 = \langle 0,n|H^1|0,n\rangle. \tag{8}$$

The solution $|1,n\rangle$ can be obtained by solving Eq. (5) directly, or in variational form, or using the sum-over-states representation:

$$|1,n\rangle = \sum_{k=0}^{\infty} |0,k\rangle x_{n,k}^1, \qquad x_{n,k}^1 := \langle 0,k|1,n\rangle, \tag{9}$$

which yields a verification of the asymptotic error laws in Eqs. (6) and (7), i.e., the integrity of the exponents of ε [13]. Upon multiplying Eq. (5) by the bra $\langle 0,m|$ on the left:

$$\sum_{k=0}^{\infty}(E_k^0 - E_n^0)\delta_{mk}x_{n,k}^1 = (E_m^0 - E_n^0)x_{n,m}^1 = E_n^1\delta_{mn} - \langle 0,m|H^1|0,n\rangle, \qquad (10)$$

the expansion coefficients are given by the well-known expressions [13,11,6,12,5]:

$$x_{n,m}^1 = \frac{\langle 0,m|H^1|0,n\rangle}{E_n^0 - E_m^0} \quad (m \neq n), \qquad x_{n,n}^1 = 0. \qquad (11)$$

Although the series (9) is also used as an asymptotic series in ε (see the review article [6]), the classical theory relies on l_2-convergence of the series (9), as assumed already in Eq. (7):

$$\langle 1,n|1,n\rangle = \sum_{k=0}^{\infty}|x_{n,k}^1|^2 < \infty, \qquad (12)$$

implying the smallness condition for the perturbation operator εH^1 [13,5,12].

The second set of equations (11) expresses the fact that the nth first-order correction $\varepsilon|1,n\rangle$ of the nth state $|0,n\rangle$ is exactly orthogonal to that state before normalization of the resultant and $\||0,n\rangle + \varepsilon|1,n\rangle\| = 1 + O(\varepsilon^2)$ (cf. [11]). Simple vector addition in Hilbert space shows that this is reasonable; the normalization issue can be handled in different ways [4,6,9,12].

1.2 Alternative derivation of the energy correction

The above principle of orthogonal corrections leads directly to Schrödinger's energy correction formula (8). Thus define, instead of Eqs. (2) and (7):

$$H = H^0 + \varepsilon^\alpha h \qquad (\|h\| = O(1), \qquad \alpha > 0), \qquad (13)$$

$$|n\rangle = |0,n\rangle + \varepsilon^\beta|h,n\rangle \quad (\||h,n\rangle\| = O(1)), \quad \beta > 0), \qquad (14)$$

where ε^α and ε^β are first-order infinitesimals. Upon inserting these expressions into Eq. (1), using the self-adjointness of H^0 and the normalization $\langle 0,n|0,n\rangle = 1$, one obtains:

$$E_n - E_n^0 = \langle 0,n|\varepsilon^\alpha h|0,n\rangle + O(\varepsilon^{\alpha+\beta}). \qquad (15)$$

This first-order energy correction is independent of β, precisely because of the essential orthogonality $\langle 0,n|h,n\rangle = 0$. The error law of $|h,n\rangle$ appears only in higher order after normalization (Section 1.1). Thus a formally more general result follows from a simple geometric argument in Hilbert space without assuming exact orthogonality of the kets $|0,n\rangle$ (cf. also Section 3.2). The above features pertain to infinite-dimensional separable unitary spaces.

1.3 Inverse perturbation analysis (IPA)

RS theory can also be used to solve the inverse perturbation problem, as has been demonstrated by Kosman and Hinze [7] and by Vidal et al. (see, e.g. [14,15]). In such an inverse perturbation analysis (IPA), a certain number of parameters α_j is determined in a model perturbation Hamiltonian $H^1(\alpha_j)$ by solving a highly overdetermined system of linear integral or algebraic equations (based upon Eq. (8)) by a least-squares method:

$$\rho_n := E_n^1 - \langle 0,n|H^1(\alpha_0,\alpha_1,\ldots,\alpha_M)|0,n\rangle = O(\varepsilon) \quad (M \leq N, n \in \{0,1,\ldots,N-1\}). \tag{16}$$

The energy corrections εE_n^1 are given from spectroscopic data. Once the values of the parameters which minimize the sum of squares of the residuals ρ_n are determined, the Schrödinger equation (1) is solved with the corresponding H^1:

$$(H^0 + \varepsilon H^1(\alpha_0^{(1)},\ldots,\alpha_{M(1)}^{(1)}))|(1),n\rangle = E_n|(1),n\rangle = (E_n^0 + \varepsilon E_n^1)|(1),n\rangle \tag{17}$$

for the improved eigenfunctions $|(1),n\rangle$. The solution of the Schrödinger equation (17) can represent a considerable task in practice, and it has to be performed more than once (and for all n) in IPA, in order to evaluate the matrix elements in (16) for the next step. The refinement of the values $\alpha_j^{(\nu)}$ of the parameters α_j is thus enhanced iteratively ($\nu = 1,2,\ldots$). Although Eq. (16) could be derived without explicitly assuming orthogonality of the unperturbed wave functions (Section 1.2), this property is obviously dictated by general theory and has been realized with sufficient accuracy in [7,14,15].

2 Consequences of slight non-orthogonality

It appears from Eq. (16) that the *collective* behavior of the unperturbed wave functions is especially important in IPA, since these are tied together in the above linear system (16) by the unknown parameters, whereas in a direct perturbation analysis (RS) each energy correction E_n^1 obviously follows from the knowledge of the *individual* $|0,n\rangle$. This observation leads to the question: ***What happens if Schrödinger's formula*** (8) ***is used with a finite number of slightly non-orthogonal basis functions?***

In practice the expansion (9) has to be truncated and the basis vectors $|0,n\rangle$ may be slightly non-orthogonal, albeit of higher order $q \geq 2$ in ε. This leads to a Hermitian positive-definite metric matrix (Gram matrix) with real eigenvalues near one:

$$0 < S_{mk} = \langle 0,m|0,k\rangle = O(\varepsilon^q)\ (m \neq k), \quad S_{kk} = 1\ (m,k \in \{0,1,\ldots,N-1\}). \tag{18}$$

The orthogonality problem is known to be nontrivial if high accuracy is desired [1,3,4,8,9]. Normalization $S_{kk} = 1$ is relatively easy to realize. As a result of Eq. (18) the components $x_{n,k}^1$ are not given by (11) but are related linearly to the

defining inner products in Eq. (9) as follows:

$$\begin{aligned} S\mathbf{x}_n^1 &= \mathbf{b}_n^1, \quad (n \in \{0,1,\ldots,N-1\}) && (19)\\ \mathbf{x}_n^1 &:= (x_{n,0}^1, x_{n,1}^1, \ldots, x_{n,N-1}^1)^T, && (20)\\ \mathbf{b}_n^1 &:= (b_{n,0}^1, b_{n,1}^1, \ldots, b_{n,N-1}^1)^T, && (21)\\ b_{n,m}^1 &:= \langle 0,m|1,n\rangle, && (22)\\ S &:= (S_{mk}) = \overline{S}^T = S^\dagger. && (23) \end{aligned}$$

Instead of Eq. (10) one obtains in first order:

$$\sum_{k=0}^{N-1} (E_k^0 - E_n^0) S_{mk} x_{n,k}^1 = E_n^1 S_{mn} - \langle 0,m|H^1|0,n\rangle \quad (n \in \{0,1,\ldots,N-1\}). \quad (24)$$

The $x_{n,k}^1$ are determined for $m = n$:

$$\sum_{k=0}^{N-1} (E_k^0 - E_n^0) S_{nk} x_{n,k}^1 = E_n^1 - \langle 0,n|H^1|0,n\rangle \quad (n \in \{0,1,\ldots,N-1\}). \quad (25)$$

These equations can be casted in matrix form:

$$AX^1 = R^1, \quad (26)$$

with the definitions:

$$\begin{aligned} A := (A_{nk}) \quad &, \quad A_{nk} := (E_k^0 - E_n^0) S_{nk}, && (27)\\ X^1 := (X_{kn}^1) \quad &, \quad X_{kn}^1 := x_{n,k}^1, && (28)\\ R^1 := (R_{mn}^1) \quad &, \quad R_{mn}^1 := (E_n^1 - \langle 0,n|H^1|0,n\rangle)\delta_{mn}. && (29) \end{aligned}$$

This is equivalent to N systems of linear equations with the same skew-Hermitian coefficient matrix A ($A = 0$ for $S = I$ =unit matrix):

$$A\mathbf{x}_n^1 = \mathbf{r}_n^1, \quad A^\dagger = \overline{A}^T = -A \quad (30)$$

with the definitions:

$$\begin{aligned} \mathbf{x}_n^1 &:= n\text{th column vector of components in } X^1, && (31)\\ \mathbf{r}_n^1 &:= n\text{th column vector of residuals in } R^1. && (32) \end{aligned}$$

By the assumption of non-degeneracy, the energy differences in A_{nk} are $O(1)$, thus A has off-diagonal elements of order $O(\varepsilon^q)$ and can be scaled as $A = \varepsilon^q A_*$, where the elements of A_* are $O(1)$. As a result, the size $\|\mathbf{x}_n^1\|$ of the solution vector lies in the interval $(\varepsilon^{-q}\|A_*\|^{-1}\|\mathbf{r}_n^1\|, \varepsilon^{-q}\kappa(A_*)\|A_*\|^{-1}\|\mathbf{r}_n^1\|)$ in any consistent pair of vector and matrix norms. The size of the interval depends on the condition number $\kappa(A_*)$ relative to the subordinate matrix norm, which is moderate under the assumptions made on the spectrum of H^0. Therefore, the solution vector is a quantity determined by the residuals R_{nn}^1 of Schrödinger's formula (8) corresponding to the ρ_n in

Eq. (16) (or Eq. (38) in Section 3.2) , amplified by a large factor of ε^{-q}, and there is useful information to be gained from the linear systems (30) for $S \neq I$, even for very small residuals. (No information will be extracted from the fact that A is the first-order term of an almost-identical transformation matrix which is unitary up to order $O(\varepsilon^{2q})$.)

In applications the time-independent wave functions $|0,n\rangle$ are frequently real. (Schrödinger has considered real waves in his seminal article [13].) Suppose then that the kets $|0,n\rangle$ are real-valued functions; the general case with complex kets will be commented where necessary.

The eigenvalues of a real skew-symmetric matrix (like the A above) lie on the imaginary axis, symmetrically with respect to the origin (symmetry being lost for complex skew-Hermitian matrices). Therefore, such matrices of odd order N are always singular. Their characteristic polynomial is of the form:

$$\lambda^N + a_{N-2}\lambda^{N-2} + \cdots + a_2\lambda^2 + a_0 = 0 \quad (N \text{ even}), \tag{33}$$

$$\lambda^N + a_{N-2}\lambda^{N-2} + \cdots + a_3\lambda^3 + a_1\lambda = 0 \quad (N \text{ odd}). \tag{34}$$

If N is odd, the eigenvector corresponding to the vanishing eigenvalue in Eq. (34) is real. Further, the vanishing eigenvalue (or the couple of vanishing eigenvalues for even N) will be simple in the generic case, since $a_2 \neq 0$ in Eq. (33) and $a_1 \neq 0$ in Eq. (34) in the generic case. On the other hand, for even N, $a_0 \neq 0$ in the generic case, and therefore such a matrix A of even order will be nonsingular almost certainly.

These properties of A lead to the following analysis of the systems (30) for real-valued kets $|0,n\rangle$. If Schrödinger's formula (8) is used with slightly non-orthogonal functions (18), then $\mathbf{r}_n^1 := 0$ (all n) and the systems (30) are made homogeneous and become identical. It is natural in applications that Eq. (8) is satisfied approximately, often for a different reason than the slight non-orthogonality (numerical errors, e.g., the residuals ρ_n in the least-squares scheme in IPA; see Section 1.3). Thus it is better to keep the right-hand sides in Eq. (30) nonzero and small, considering the limit of vanishing residuals $\mathbf{r}_n^1 \to 0$.

2.1 N= even integer

For even N, A is nonsingular in the generic case. Barring again ill-conditioning of A, one can infer that the solution $\mathbf{x}_n^1$ of Eq. (30) will exist, will be different for different n, and will tend to the zero vector in the limit $\mathbf{r}_n^1 \to 0$. In particular, because the overlap matrix $S = I + O(\varepsilon^q)$ in Eq. (19) is nonsingular, the reasonable conditions $b_{n,n}^1 = 0$ corresponding to the second set of equations (11) and discussed after Eq. (11) are satisfied: The nth correction of the nth excited state is orthogonal to the nth unperturbed state (principle of orthogonal corrections, cf. also Section 1.2). But $\mathbf{b}_n^1 = \mathbf{0}$ in the limit is a much stronger result: The nth correction of the nth excited state is orthogonal to *all the other* excited states $|0,k\rangle$ ($k \in \{0,1,\ldots,N-1\}$). Therefore, the corrections lie in the orthogonal complement of the N-dimensional approximation space spanned by the unperturbed

states. In the language of Fourier series, one could say that the corrections contain only higher harmonics starting with the $(N+1)$st harmonic. In conclusion, the approximation scheme is internally consistent for even N, but the result is physically reasonable only if the non-orthogonality is very slight or if N is sufficiently small, in which case only a few lower "harmonics" are skipped (or ignored) in the first-order correction. The latter requirement is contrary to the one dictated by the good approximation of $|1,n\rangle$ by the expansion in $|0,k\rangle$ which asks for a large N. This dilemma is created by the fact that *the number of energy values is set equal to the number of basis functions* in the IPA procedure towards which attention is focused. If the kets $|0,n\rangle$ are complex-valued, A is skew-Hermitian, its spectrum is not symmetric about the origin in general, and nothing changes essentially in the above argument (A can be written as a real matrix of even order $2N$).

2.2 $N=$ odd integer

Turning to the case with odd N, the formalism appears to be internally inconsistent for real-valued kets (while for complex-valued ones, there is no difference to the case with even N). Indeed, a real A will always be singular and its rank will be $N-1$ in the generic case by Eq. (34). This invalidates completely the continuity argument which was found reasonable for even N, since the homogeneous systems in the limit $\mathbf{r}_n^1 = \mathbf{0}$ will not only admit the trivial solution but also a parasitic solution. More precisely, they will have a one-parameter family of nontrivial solutions which will become identical for all n after normalization. This is physically unreasonable, signifying that each excited state $|0,n\rangle$ would receive exactly the same correction, independently of the particular quantum mechanical system considered. Moreover, the systems (30) will not admit any solution at all in the generic case, although they have the peculiar property that the rank deficiency is equal to the number of nonzero components of the inhomogeneity vector, namely one. Nonexistence of a solution follows from the general theory of linear systems (Fredholm alternative in finite-dimensional inner-product spaces). The conclusion in this case is that the formalism is internally inconsistent in odd dimensions for real, slightly non-orthogonal kets $|0,n\rangle$.

2.3 $N \to \infty$, $N=$ even integer

The limit $N \to \infty$ (taking convergent subsequences with even N in the real case) is reasonable by virtue of the above features of the linear systems, as long as the matrices are kept square and the set $\{|0,n\rangle : n \in \{0,1,\ldots\}\}$ remains a Riesz basis. This is a very weak condition in the present context with nearly orthogonal basis functions. Further, the weak assumption of equicontinuous wave functions (usual in quantum mechanics) avoids the need to consider the integrals in the sense of Lebesgue. The resulting consistency of the Rayleigh-Schrödinger formalism for even N warrants its use as an indicator of any behavior of the wave function corrections, e.g., in the sense of the discussion in Section 2.1. It should be noted that

this limit is only valid in first order (which is sufficient), since the eigenfunctions of a self-adjoint operator form an orthogonal family (even for degenerate spectra).

3 Alternative formulations

3.1 The energy representation and IPA

There is a rather straightforward way of getting around the above consistency problem (Section 2.2). One starts with Eqs. (1)-(3) but avoids Eqs. (6) and (7), and the main equation can be written *formally* for any metric matrix in Eq. (18) as follows:

$$\sum_{k=0}^{\infty} \left\{ (E_k^0 - E_n) S_{mk} + \langle 0,m|H^1|0,k\rangle \right\} x_{n,k} = 0 \quad (m \in \{0,1,\ldots\}). \tag{35}$$

This is a set of N homogeneous linear equations for the expansion coefficients $x_{n,k}$ of the exact kets $|n\rangle$ in terms of the complete set of unperturbed kets $|0,n\rangle$, sometimes called the energy representation of the Schrödinger equation in the theoretical case $S = I$. This system is just an early stage in the derivation of the Rayleigh-Schrödinger scheme of Section 1.1, before the assumptions on smallness of the energy corrections are made. Thus E_n instead of E_n^1 appears in the system. The condition for existence of a nontrivial solution $\mathbf{x}_n$ is:

$$\det A_n(E_n, H^1) = 0. \tag{36}$$

One may truncate Eq. (35) at a finite size N and use the square system as approximation. (Such systems appear routinely in the eigenspaces of perturbed degenerate eigenvalues where they are solved in first order by an almost-identical pricipal axis transformation [6,11,13,12].) In the direct problem, these secular equations define the exact energies E_n, in the inverse problem the perturbation operator $H^1(\alpha_j^{(\nu)})$. The present formulation is not sensitive to the parity of N or to the non-orthogonality of the basis. Non-orthogonality implies that small $O(\varepsilon^q)$ fractions of the energy-differences are distributed all over the matrix A_n, changing the matrix elements in a higher order $O(\varepsilon^q)$. It has to be conceded that these contaminations could be amplified, if the matrix A_n is ill-conditioned. Notice that the solution is not a correction and does not follow the principle of orthogonal corrections of Sections 1.1 and 1.2.

There is useful information emerging from the nonlinear character of the scheme. The necessity of prescribing a suitable set of starting values $\alpha_j^{(0)}$ for the solution of Eq. (36) ("inner iteration") is a chance to communicate any wishes concerning the physically reasonable behavior of the solution. For example, if $H^1(\alpha_j)$ is a potential correction, then one may have definite ideas concerning its shape [2]. If the model based upon such physical insight is also sufficiently accurate and flexible (in relation with the smallness of ε), then the starter is expected to lie in the basin of attraction of the physical solution. This step corresponds to a damping

(or stabilization) of the IPA scheme by including the approximately desired potential correction in the least-squares fit in Eq. (16) in order to "tame" the (all too flexible) potential correction. Such an extended least-squares scheme is well adapted to the structure of a *multiplicative perturbation operator* (such as a potential function) and can be formulated so as to improve the numerical condition of the design matrix.

3.2 A scaled Schrödinger equation for IPA

Incidentally, the inhomogeneous first-order equation (5) for the correction $\varepsilon|1,n\rangle$ is numerically superior to Eq. (17) because it does not contain quantities of different orders of magnitude. In the first step (1) of IPA one has for $n \in \{0,1,\ldots,N-1\}$:

$$(H^0 - E_n^0)|1(1),n\rangle = (E_n^1 - H^1(\alpha_j^{(1)}))|0,n\rangle \quad (j \in \{0,1,\ldots,M^{(1)}\}). \tag{37}$$

By the Fredholm alternative, this equation is only solvable in leading order since the residual $\rho_n^{(1)}$ in Eq. (16) does never vanish in practice:

$$(H^0 - E_n^0)|1(1),n\rangle = \{\langle 0,n|H^1(\alpha_j^{(1)})|0,n\rangle - H^1(\alpha_j^{(1)})\}|0,n\rangle + \rho_n^{(1)}|0,n\rangle. \tag{38}$$

Therefore, the required accuracy of the solution is automatically determined by the size of the residual which at any rate is $O(\varepsilon)$. At each step (ν) the order constant is reduced by a certain amount, and it is not necessary nor adequate to solve these equations with excessive accuracy ("ignoring the residuals"). The Fredholm alternative specifies further that the solution $|1(1),n\rangle$ is unique up to an additive multiple of the solution $|0,n\rangle$ of the associated homogeneous equation and, the free constant is determined by requiring that the solution is orthogonal to $|0,n\rangle$:

$$|1(1),n\rangle := |1(1),n\rangle - \langle 0,n|1(1),n\rangle|0,n\rangle + O(\varepsilon). \tag{39}$$

Since the solution is a correction, it could be expected that the principle of orthogonal correction (Section 1.2) would reappear, now as a direct result of the Fredholm alternative.

4 Conclusion

The analysis of the truncated Rayleigh-Schrödinger formalism in Section 2 has exhibited a rather peculiar behavior with slightly non-orthogonal basis functions. It was possible to show that the scheme remains internally consistent in first order (not for odd real dimensions).

In the presence of slight non-orthogonality, the truncated sum-over-states representation appears to be governed essentially by the residuals (random errors in IPA) of Schrödinger's first-order energy correction formula. The sum-over-states representation collapses to zero in the limit of vanishing residuals, indicating that the first-order correction has no components anymore in the direction of the original approximation space. Thus slight non-orthogonality causes extraneous oscillations in the corrections, corresponding to the dominant components in the

orthogonal complement in Hilbert space. These in turn may induce oscillations in the unknown least-squares parameters in IPA. Normally this effect, which increases with N, may be damped by the quadratures in the matrix elements and it will be obscured further by other numerical errors in actual computations. Nevertheless, the sum-over-states representation acts as an indicator of that particular effect, much like spectral analysis. It is not aimed at replacing the direct solution by a truncated sum over states, as this would produce a gross truncation error, but rather to acknowledge partial information regarding the first N spectral components. The consistence of the RS formalism in the sense of Section 2 implies that the projection onto this subspace is correct.

Successful results with IPA are reported in [7,14,15] for particular quantum mechanical systems, apparently using "sufficiently orthogonal" wave functions. However, the tradeoff mechanism between the described ocillations and the stabilizing effect of solving the Schrödinger equation again (and iteratively in IPA) is obviously problem-dependent and it lacks theoretical foundation or numerical validation in molecular systems. The least-squares parameters could well drift away from the correct solution towards a parasitic solution, and such a danger is always imminent in inverse problems (non-uniqueness problem).

A study of the Banach algebra structure of the resolvent operator of the whole IPA iteration, not just of the Schrödinger equation (17), may shed light on the general approximation problem, but the role of orthogonality requires consideration of the inner-product (unitary) space structure of the functionals (wave functions). Nonlinearity of the resolvent is made plausible by the observation that Schrödinger's first-order energy correction formula is a sesquilinear form ("quadratically " nonlinear in $|0,n\rangle$) induced by the bounded linear perturbation operator H^1 (Theorem of Lax-Milgram).

The alternative formulation in Section 3.1 has been included in order to contrast its behavior with that of RS theory in the presence of non-orthogonality, to illustrate the importance of a good starter solution in IPA, and to present (not advocate) a possible computational tool.

The scaled first-order version of the Schrödinger equation (Section 3.2) has apparently not yet been used in IPA and is proposed as a numerically promising alternative.

Finally, it should be mentioned that other perturbation techniques with non-orthogonal basis functions are discussed in Morse and Feshbach [10].

5 Acknowledgments

This study is part of the theoretical basis of a damped IPA scheme, within the framework of a project on nonadiabatic effects in the vibrations and rotations of the hydrogen molecule and its isotopes. The author is indebted to the project leader, Professor Kurt Dressler at the Physical Chemistry Laboratory (ETH), for an appointment as guest scientist during the initial phase of this work.

References

[1] M. Cohen and A. Dalgarno, Rev. Mod. Phys. **35**, 506 (1963).

[2] K. Dressler (personal communication).

[3] W. Gautschi, in *Orthogonal Polynomials and their applications*, C. Brezinski et al. (editors), J. C. Baltzer AG, Scientific Publishing Co (IMACS), 61-71 (1991).

[4] Y. J. I'Haya and T. Morikawa, Adv. Quantum Chemistry **12**, 43 (1980).

[5] T. Kato, *Perturbation Theory of Linear Operators*, Springer-Verlag, Berlin, 1966.

[6] J. Killingbeck, Rep. Prog. Phys. **40**, 963 (1977).

[7] W. M. Kosman and J. Hinze, J. Molec. Spectrosc. **56**, 93 (1975).

[8] L. Molinari, Chapter 4 in *Numerische Prozeduren: aus Nachlass und Lehre von Heinz Rutishauser*, W. Gander et al. (editors), Internat. series of numerical mathematics **33**, Birkhäuser, Basel, 1977.

[9] T. Morikawa and Y. J. I'Haya, J. Math. Phys. **22** (6), 1186 (1981).

[10] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics, Part II, p. 1034*, McGraw-Hill, New York, 1953.

[11] W. Pauli, *Die allgemeinen Prinzipien der Wellenmechanik*, Springer-Verlag, Berlin, 1990.

[12] F. Rellich, *Perturbation Theory of Eigenvalue Problems*, Gordon and Breach, New York, 1969.

[13] E. Schrödinger, Annalen der Physik **4**, Bd. 80, 437, (1926).

[14] C. R. Vidal and H. Scheingraber, J. Molec. Spectrosc. **65**, 46 (1977).

[15] C. R. Vidal, Comm. At. Mol. Phys. **17** (4), 173 (1986).

Recent IPS Research Reports

93-01 M. TROYER, H. TSUNETSUGU, T.M. RICE, J. RIERA AND E. DAGOTTO: Spin Gap and superconductivity in the one-dimensional $t - J$ model with Coulomb repulsion.

93-02 F. PYTHOUD: Mass fluctuations of a Sandpile automaton.

93-03 C. DE GROOT: Nonlinear time series analysis with connectionist networks.

93-04 K. AKEMI, P. DE FORCRAND, M. FUJISAKI, T. HASHIMOTO, H.C. HEGE, S. HIOKI, O. MIYAMURA, A. NAKAMURA, M. OKUDA, I.O. STAMATESCU, Y. TAGO AND T. TAKAISHI: Quantum chromodynamics simulations on a nondedicated highly parallel computer.
QCD_TARO collaboration

93-05 C. BERNARD, T. BLUM, T.A. DEGRAND, C. DETAR, S. GOTTLIEB, A. KRASNITZ, R.L. SUGAR AND D. TOUSSAINT: Finite-size and quark mass effects on the QCD spectrum with two flavors.

93-06 R. SCHNIDRIG: Bildrauschunterdrückung und Texturesegmentierung mit autoregressiven Prozessen auf einem massiv parallelen Rechner.

93-07 J. HUTTER, H.P. LÜTHI AND F. DIEDERICH: The structures and vibrational frequencies of the carbon molecules C_2–C_{18} calculated by density functional theory.

93-08 R. RÖHSE, W. KLOPPER AND W. KUTZELNIGG: CISD-R12 calculations of the ground state of H_3^+. A benchmark study.

93-09 K. AKEMI, P. DE FORCRAND, M. FUJISAKI, T. HASHIMOTO, S. HIOKI, O. MIYAMURA, A. NAKAMURA, M. OKUDA, I.O. STAMATESCU, Y. TAGO AND T. TAKAISHI: Scaling study of pure gauge lattice QCD by Monte Carlo renormalization group method.

93-10 A. KRASNITZ AND R. POTTING: Sphaleron transitions in a realistic heat bath.

93-11 M.H. GUTKNECHT AND M. HOCHBRUCK: Look-ahead Levinson and Schur algorithms for non-Hermitian Toeplitz systems.

93-12 S. CABAY, M.H. GUTKNECHT AND R. MELESHKO: Stable rational interpolation?

93-13 M.H. GUTKNECHT AND M. HOCHBRUCK: The stability of inversion formulas for Toeplitz matrices.

93-14 M.H. GUTKNECHT: The multipoint Padé table and general recurrences for rational interpolation.

93-15 K.E. GATES AND W.P. PETERSEN: A technical description of some parallel computers.

93-16 A. KRASNITZ AND R. POTTING: Sphaleron-like processes in a realistic heat bath.

93-17 G.G. GROH: On Rayleigh-Schrödinger perturbations of slightly non-orthogonal states and IPA.