ORCA – Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:https://orca.cardiff.ac.uk/id/eprint/146755/

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Knowles, Peter J. 2022. Perturbation-adapted perturbation theory. The Journal of Chemical Physics 156 , 011101. 10.1063/5.0079853

Publishers page: http://dx.doi.org/10.1063/5.0079853

Please note:
Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See http://orca.cf.ac.uk/policies.html for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.

Perturbation-Adapted Perturbation Theory

Peter J. Knowles[1]
School of Chemistry, Cardiff University, Main Building, Park Place, Cardiff CF10 3AT, United Kingdom

(*Electronic mail: KnowlesPJ@Cardiff.ac.uk)

(Dated: 6 December 2021)

A new general approach is introduced for defining an optimum zero-order Hamiltonian for Rayleigh-Schrödinger perturbation theory. Instead of taking the operator directly from a model problem, it is constructed to be a best fit to the exact Hamiltonian within any desired functional form. When applied to many-body perturbation theory for electrons, strongly improved convergence is observed in cases where the conventional Fock Hamiltonian leads to divergence or slow convergence.

Rayleigh-Schrödinger perturbation theory (RSPT)[1] is a ubiquitous tool for the simplification and analysis of many problems in quantum mechanics. We here consider its application to the ground-state eigenvalue of the time-independent Schrödinger equation. In its usual form, the first step is to identify a model problem that is in some sense close to the target problem, but for which exact solutions to Schrödinger's equation can be obtained. The energy and wavefunction for the target problem are then obtained recursively as a power series in the perturbing Hamiltonian, defined as the difference between the target- and model-problem Hamiltonians. The process is usually formulated under the assumption that the complete set of eigenfunctions of the model problem is available. Success is defined by realizing sufficient accuracy from the series for the energy truncated at low order.

It has long been recognized that the meaningfulness of low-order truncation of perturbation series is dependent at least to some extent on the existence and rapidity of convergence of the series. Although even when a perturbation series is formally divergent, the first terms of an asymptotic expansion can still be useful, it is generally the case that divergence or slow convergence mean that the low-order energies may not be accurate enough to be useful approximations to the true Hamiltonian eigenvalue. In some instances, convergence can be improved by resummation[2–11] or scaling[4,12] techniques, but in general these are not straightforward to apply in a way that preserves extensivity of the energy in many-body theory.

Although the model problem may have a physical reality that aids interpretation and which defines the model Hamiltonian operator $\hat{H}_0$ naturally, there is no intrinsic requirement for either of these for RSPT to be effective. What, however, is essential is that one has an operator whose ground state is a zero-order wavefunction that is a good approximation to the ground state of the target problem; there is complete flexibility about the remainder of its spectrum. We explore here this flexibility, and propose a general procedure for defining an optimum $\hat{H}_0$ given a particular zero-order wavefunction $|0\rangle$ and exact Hamiltonian $\hat{H}$, together with any constraints on the form of $\hat{H}_0$ necessary for sufficiently simple computation.

RSPT starts by choosing an initial zero-order hamiltonian $\hat{H}_0$, which then defines the initial perturbation operator $\hat{H}_1 = \hat{H} - \hat{H}_0$. Schrödinger's equation is then partitioned for all values of the perturbation strength λ,

$$\sum_{n=0} \lambda^n \Big(\left(\hat{H}_0 - E_0\right)|n\rangle + \left(\hat{H} - \hat{H}_0 - E_1\right)|n-1\rangle - \sum_{m=0}^{n-2} E_{n-m}|m\rangle \Big) = 0 \qquad (1)$$

which can be solved order by order in the usual way to obtain the energy and wavefunction.

We now seek an improved zero-order Hamiltonian, $\hat{\Lambda}$, by maximizing the degree to which it resembles the full Hamiltonian, as measured by its action on the first-order wavefunction projected onto a suitable space. The rationale for this approach is that it is known that sometimes the convergence of RSPT can be spoiled if the eigenvalues of $\hat{H}_0$ are markedly different from those of $\hat{H}$; it can then happen that at some point inside the unit circle in the complex plane, $\hat{H}_0 + \lambda(\hat{H} - \hat{H}_0)$, acquires a degenerate ground state, and then at $\lambda = 1$ the perturbation series is divergent[13–19]. Slow convergence can usually also be associated with near degeneracy somewhere in the unit circle. The similarity conditions can be expressed as

$$\langle 0|\hat{\Theta}_p^\dagger(\hat{\Lambda} - \hat{H})|1\rangle = 0 \qquad (2)$$

where $\{\hat{\Theta}_p\}$ is an appropriately-chosen set of operators. It is anticipated that $\hat{\Lambda}$ will be defined through a number of adjustable parameters to be determined from the solution of (2). We denote the power-series expansion of energy and wavefunction based on this partitioning Perturbation-Adapted Perturbation Theory (PAPT), since the zero-order hamiltonian is no longer universal, but depends on the nature and strength of the perturbation.

An obvious choice for the improved zero-order hamiltonian is $\hat{\Lambda} = \hat{H}$. For most problems, this will not offer a practical advantage to simply solving Schrödinger's equation exactly; the linear eigenvalue problem is replaced by a sequence of inhomogeneous linear equations of the same dimension. But we will explore the properties of the approach by applying it to the perturbed 1-dimensional harmonic oscillator. We also note that this choice of the exact Hamiltonian is related to Epstein-Nesbet perturbation theory[20,21], which incorporates the further approximation of ignoring the off-diagonal elements of $\hat{\Lambda}$ for computational simplicity.

More usually, we will construct $\hat{\Lambda}$ to have the same functional form as $\hat{H}_0$, with linear adjustable parameters. We consider below the case of electronic structure theory with the Fock $\hat{H}_0$; in that case, to retain computational efficiency, $\hat{\Lambda}$ should also be a one-body operator whose spectrum mimics that of the two-body $\hat{H}$.

I. APPLICATION TO THE PERTURBED HARMONIC OSCILLATOR

We consider a zero-order hamiltonian for the harmonic oscillator with unit mass and force constant, and taking $\hbar = 1$:

$$\hat{H}_0 = -\tfrac{1}{2}\frac{d^2}{dx^2} + \tfrac{1}{2}x^2 \tag{3}$$

which has eigenvalues $(n+\frac{1}{2}), n = 0,1,2,\ldots$ and known eigenfunctions. We then add a damped quartic perturbation,

$$\hat{H} = \hat{H}_0 + \lambda\, x^4 e^{-x^2/8} \tag{4}$$

It is well known that without the gaussian damping factor, perturbation theory diverges for all $\lambda \neq 0$, associated with the fact that for any finite negative λ there is a negative-energy continuum of eigenstates.

In order to develop PAPT for the ground state of this system, we require that the harmonic-oscillator ground state is an eigenfunction of $\hat{\Lambda}$, but that otherwise $\hat{\Lambda}$ is as close as possible to the actual Hamiltonian $\hat{H}_0 + \lambda(\hat{H} - \hat{H}_0)$. This can be achieved by simple projection,

$$\hat{\Lambda} = \hat{P}\hat{H}_0\hat{P} + (1-\hat{P})(\hat{H}_0 + \lambda(\hat{H} - \hat{H}_0))(1-\hat{P}) \tag{5}$$

where $\hat{P} = |0\rangle\langle 0|$. Table I shows the convergence of RSPT and PAPT as measured by the remaining difference from the exact ground-state eigenvalue after truncating the perturbation series at different orders. The calculations have been carried out using a basis consisting of the first 30 harmonic-oscillator eigenfunctions. RSPT is very slowly convergent, whereas PAPT converges rapidly. It is also possible to obtain similarly rapid convergence of PAPT for the undamped perturbation λx^4, for which RSPT diverges for all λ. We note that in this application of PAPT using the full hamiltonian, the third-order energy (and systematically other higher-order contributions) is zero because the representation of $\hat{H} - \Lambda$ in the first-order interacting space is zero.

Order	RSPT	PAPT
1	7.53×10^{-2}	7.53×10^{-2}
2	-2.30×10^{-3}	-2.19×10^{-5}
3	8.91×10^{-4}	-2.19×10^{-5}
4	-4.63×10^{-4}	1.303×10^{-7}
5	1.74×10^{-4}	1.303×10^{-7}
6	-1.43×10^{-4}	1.303×10^{-7}
7	2.01×10^{-5}	-9.73×10^{-10}

TABLE I. Deviation from exact of truncated perturbation series for the ground-state energy of the damped-quartic-perturbed harmonic oscillator with $\lambda = 0.1$ at different truncation orders.

II. MANY-BODY PERTURBATION THEORY FOR ELECTRONS

Conventional RSPT for the electron correlation problem proceeds using the Møller-Plesset zero-order Hamiltonian[22,23]

$$\hat{H}_0 = \hat{f} = f_{ab}\, a^\dagger b - f_{ij}\, j i^\dagger + f_{ii} + E_{\text{core}} \tag{6}$$

where $i, i^\dagger, a, a^\dagger$ are the usual annihilation and creation operators for, respectively, occupied and virtual orbitals, $\hat{f}$ is the Fock operator of the reference Hartree-Fock calculation, E_{core} is the nuclear-nuclear Coulomb energy, and the Einstein summation convention is used.

We now seek

$$\hat{\Lambda} = \Lambda_{ab}\, a^\dagger b - \Lambda_{ij}\, j i^\dagger + \Lambda_{ii} + E_{\text{core}} \tag{7}$$

such that

$$\hat{\Lambda}|1\rangle \approx \hat{H}|1\rangle \tag{8}$$

where $|1\rangle$ is the first-order RSPT wavefunction,

$$|1\rangle = \tfrac{1}{4} c^{ij}_{ab}\, a^\dagger b j i |0\rangle \tag{9}$$

To find the parameters $\{\Lambda_{ab}\}, \{\Lambda_{ij}\}$, we project (8) onto a suitable space. The vectors

$$|\Theta_{ij}\rangle = \tfrac{1}{4} c^{ik}_{ab}\, a^\dagger b^\dagger k j |0\rangle \tag{10}$$

$$|\Theta_{ab}\rangle = \tfrac{1}{4} c^{ij}_{ac}\, b^\dagger c^\dagger j i |0\rangle \tag{11}$$

are equal in number to the unknowns, and sample the first-order interacting space. The idea is that the reference function plus first-order interacting space contain the dominant part of the wavefunction, and that many higher excited states will arise as unlinked clusters from this space, and so if the similarity condition is obeyed in the first-order space (or even a subspace), it is likely to be reasonable in the full Hilbert space too. The projection vectors contain one linear dependency, $|\Theta_{ii}\rangle = |\Theta_{aa}\rangle = |10\rangle$, and so projection of (8) onto $\{|\Theta_{ij}\rangle\} \oplus \{|\Theta_{ab}\rangle\}$ results in a linear equation system with one redundant equation, which matches a redundant parameter in the parameters: the perturbation equations depend only on differences of eigenvalues of Λ, and are therefore unaffected by a constant shift in all of its diagonal elements.

We further require that $\hat{\Lambda}$ is a physically reasonable, and therefore Hermitian, operator. This can be achieved by enforcing $\hat{\Lambda} = \hat{\Lambda}^\dagger$ and projecting against the space $\{|\Theta_{ij} + \Theta_{ji}\rangle, i \geq j\} \oplus \{|\Theta_{ab} + \Theta_{ba}\rangle, a \geq b\}$.

Figure 1 compares the performance of PAPT and conventional Møller-Plesset (MP) RSPT for the ground singlet state of the CH_2 molecule (bond length 1.102 Å, angle 104.6°). This is an example in which a low-lying electronically-excited state causes convergence of RSPT to be slow. Using the cc-pVDZ basis set[24], and excluding excitations from the lowest (C 1s) orbital, full configuration interaction calculations are possible, and define the exact problem. With the full Slater-determinant basis, perturbation theory is then propagated to any desired order[25,26]. The difference between the perturbation series for the energy, truncated at successive orders, and

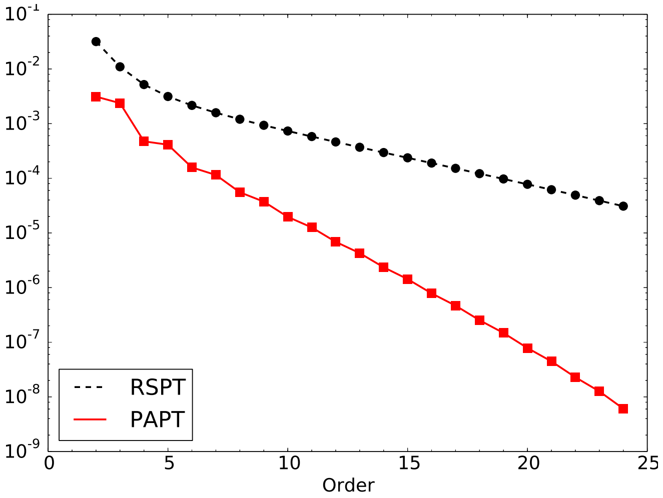

FIG. 1. Deviation from full configuration interaction of truncated perturbation series for the ground-state energy / Hartree of $^1A_1$ CH_2 at different truncation orders.

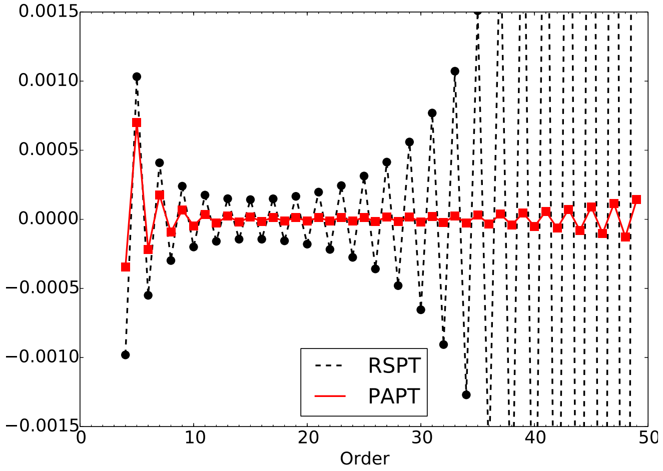

FIG. 2. Deviation from full configuration interaction of truncated perturbation series for the ground-state energy / Hartree of Ne at different truncation orders.

exact is plotted as measure of degree of convergence. The use of a logarithmic scale reflects the fact that both converge monotonically and exponentially with order, but it is seen that the convergence of PAPT is significantly better than that of RSPT, with errors that are more than an order of magnitude smaller at all orders, and a much faster asymptotic rate of convergence.

Figure 2 shows the same measures for the Ne atom in the aug-cc-pVDZ[24,27] basis, and with excitations from the 1s orbital omitted. This has already been established[13] as an example of divergence arising from "back-door intruders", ie highly-excited states which for certain perturbation strengths within the unit circle in the complex plane, and with negative real part, become degenerate with the ground state[13–18]. Not shown on the figure are the third-order errors, which are similar for PAPT (0.00417) and RSPT (0.00470). The RSPT series is oscillatory, and the energies begin to diverge from around 16th order. The PAPT series is also oscillatory, and formally divergent, but divergence begins at higher order, and when the errors are already much smaller.

These examples exhibit a third-order energy that is small but non-zero, and generally smaller contributions from odd orders than from even. This is a consequence of $\hat{\Lambda}$ reproducing $\hat{H}$ in the first-order space generated by $\hat{H} - \hat{H}_0$, which is near, but not identical, to that arising from $\hat{H} - \hat{\Lambda}$. This illustrates a connection between PAPT and Feenberg scaling[12], in which $\hat{H}_0$ is scaled to make the third order energy zero. The Feenberg approach defines a single global scaling parameter, which unfortunately leads to a lack of extensivity, whereas in PAPT, which is defined entirely through linked tensor contractions, extensivity is preserved. One could envisage proceeding further by making an improved zero-order Hamiltonian by using the PAPT first-order wavefunction instead of $|1\rangle$ in equation (8), and even iterating until self-consistent, giving a zero third-order energy, and reduced error at second order. The computational cost for each determination of an improved zero-order Hamiltonian is essentially the same as a third-order calculation; therefore PAPT3 is roughly twice the cost of MP3, and it is probably not worthwhile to iterate further. At higher order, the overhead for PAPT is essentially negligible.

III. CONCLUSION

It is possible to significantly improve the convergence characteristics of Rayleigh-Schrödinger perturation series by adopting a zero-order Hamiltonian that is parameterized to mimic the full Hamiltonian in the first-order interacting space. The method is aspirationally similar to resummation schemes, as well as Similarity Renormalization Group and canonical transformation approaches[28–30], in which singularities arising from actual or near degeneracy are removed by appropriate transformations, but is defined entirely through linear equations. In general, the approach may not offer a significant practical advantage if the resulting zero-order Hamiltonian does not have a simple form. In the special case of electronic structure theory, we have shown that it is possible to find a one-electron operator that represents the full two-electron Hamiltonian, at the cost of an additional third-order perturbation treatment. In the examples considered, the convergence of the perturbation series is substantially improved.

[1] E. Schrödinger, Annalen der Physik **385**, 437 (1926).
[2] S. Wilson and D. M. Silver, Physical Review A **14**, 1949 (1976).
[3] D. M. Silver, S. Wilson, and R. J. Bartlett, Physical Review A **16**, 477 (1977).
[4] A. J. Sadlej and S. Wilson, Molecular Physics **44**, 299 (1981).
[5] D. Z. Goodson, Journal of Chemical Physics **112**, 4901 (2000).
[6] D. Z. Goodson, The Journal of Chemical Physics **113**, 6461 (2000).
[7] R. Roth and J. Langhammer, Physics Letters, Section B: Nuclear, Elementary Particle and High-Energy Physics **683**, 272 (2010).
[8] D. Z. Goodson, Wiley Interdisciplinary Reviews: Computational Molecular Science **2**, 743 (2012).
[9] Z. Mihálka, Szabados, and P. R. Surján, Journal of Chemical Physics **146** (2017), 10.1063/1.4978898.
[10] Z. Mihálka and P. R. Surján, Physical Review A **96**, 1 (2017).

[11] Z. Mihálka, Szabados, and P. R. Surján, Journal of Chemical Physics **150** (2019), 10.1063/1.5083191.

[12] C. Schmidt, M. Warken, and N. C. Handy, Chemical Physics Letters **211**, 272 (1993).

[13] O. Christiansen, J. Olsen, P. Jørgensen, H. Koch, and P.-Å. Malmqvist, Chemical Physics Letters **261**, 369 (1996).

[14] J. Olsen, P. Jørgensen, T. Helgaker, and O. Christiansen, The Journal of Chemical Physics **112**, 9736 (2000).

[15] D. Z. Goodson and A. V. Sergeev, in *Advances in Quantum Chemistry*, Vol. 47 (2004) pp. 193–208.

[16] D. Z. Goodson and A. V. Sergeev, Physics Letters, Section A: General, Atomic and Solid State Physics **359**, 481 (2006).

[17] A. V. Sergeev and D. Z. Goodson, The Journal of Chemical Physics **124**, 094111 (2006).

[18] J. Olsen and P. Jørgensen, Journal of Chemical Physics **151** (2019), 10.1063/1.5110554.

[19] A. Marie, H. G. A. Burton, and P.-F. Loos, Journal of Physics: Condensed Matter **33**, 283001 (2021).

[20] P. S. Epstein, Physical Review **28**, 695 (1926).

[21] R. K. Nesbet, Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences **230**, 312 (1955).

[22] C. Møller and M. S. Plesset, Physical Review **46**, 618 (1934).

[23] J. S. Binkley and J. A. Pople, International Journal of Quantum Chemistry **9**, 229 (1975).

[24] T. H. Dunning Jr, The Journal of Chemical Physics **90**, 1007 (1989).

[25] P. J. Knowles, K. Somasundram, N. C. Handy, and K. Hirao, Chemical Physics Letters **113**, 8 (1985).

[26] N. C. Handy, P. J. Knowles, and K. Somasundram, Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta) **68**, 87 (1985).

[27] R. A. Kendall, T. H. Dunning Jr, and R. J. Harrison, The Journal of Chemical Physics **96**, 6796 (1992).

[28] S. D. Glazek and K. G. Wilson, Physical Review D **49**, 4214 (1994).

[29] S. R. White, Journal of Chemical Physics **117**, 7472 (2002).

[30] F. A. Evangelista, The Journal of Chemical Physics **141**, 054109 (2014).