Brillouin-Wigner perturbation theory: An introduction

中国科大近代物理系 尤一宁

2017 年 5 月 24 日

1 Introduction

In this paper, we first recall the formula of Rayleigh-Schrödinger perturbation theory (RSPT), then introduce a neater perturbation theory: Brillouin-Wigner perturbation theory (BWPT). A short discussion will focus on their difference and advantage. Afterward, we will talk about generalized BWPT (GBWPT) mainly on the quantitative requirement of GBWPT's convergence, as well as its application on Quantum Chaos theories.

2 Rayleigh-Schrodinger perturbation theory

Different perturbation theories of Quantum Mechanics are widely used in computing and analysing forms and behaviors of perturbed systems' eigenfunctions. In this paper, we denote unperturbed Hamiltonian as H_0 and perturbed Hamiltonian as $H = H_0 + \lambda V$ with perimeter λ depicting the strength of perturbation. Meanwhile, we express their eigenvalues and eigenfunctions as:

$$H_0|n\rangle = E_n^0|n\rangle \tag{1}$$

$$H|\psi_n\rangle = E_n|\psi_n\rangle \tag{2}$$

In the simplest theory *Rayleigh−Schrödinger perturbation theory* (*RSPT*) (non-degenerate), perturbed wave function and energy are expanded upon perimeter. We can present perturbed states as linear overlay of unperturbed eigenfunctions:

$$|\psi_n\rangle = \sum_{i=0}^{\infty} \lambda^i |\xi_i\rangle = |n\rangle + \sum_{i=1}^{\infty} \lambda^i |\xi_i\rangle \tag{3}$$

with $|\xi_i\rangle = \sum\limits_{m\neq n} C_m^{(i)}|m\rangle$, where $C_m^{(i)}$ are coefficients in iterative sequence.

This leads to perturbed states and energy with very complex high order forms (setting $\sum_{m\neq n}$ as $\sum_m{}'$):

$$|\psi_n\rangle = |n\rangle + \lambda \sum_m{}' \frac{V_{mn}}{E_n^0 - E_m^0}|m\rangle + \lambda^2(\sum_m{}' \frac{V_{nn}}{E_m^0 - E_n^0}\frac{V_{mn}}{E_n^0 - E_m^0} - \sum_m{}' \sum_k{}' \frac{V_{mk}}{E_m^0 - E_n^0}\frac{V_{kn}}{E_n^0 - E_k^0})|m\rangle + \cdots \tag{4}$$

$$E_n = E_n^0 + \lambda V_{nn} + \lambda^2 \sum_m{}' \frac{|V_{nm}|}{E_n^0 - E_m^0} \tag{5}$$

As shown above, in RSPT, perturbed states gets irregular and complex from second order term, further calculation gives high order states and energy with much more terms. Meanwhile, in certain models, RSPT perturbed terms converge slower than the following *Brillouin – Wigner perturbation theory*.

3 Brillouin-Wigner perturbation theory

The *Partitioning Technique* in many-body systems divides the system Hamiltonian into an effective space and its orthogonal space[1]. Brillouin-Wigner perturbation theory (BWPT) is more convenient to expand these two space, in which the perturbed states are expanded upon part of the unperturbed eigenfunctions in one term and the other part upon another term. Furthermore, numerical results from Generalized Brillouin-Wigner perturbation theory (GBWPT) can explain the behavior of different part of the eigenfunctions given by Wigner-band random-matrix (WBRM) model[2].

In specialized BWPT, for each unperturbed eigenfunction $|n\rangle$ we define two projection operators, dividing unperturbed space into two parts:

$$P_n = |n\rangle\langle n| \tag{6}$$

$$Q_n = I - P_n = \sum_m{}' P_m \tag{7}$$

P_n projects state vectors into a 1D space, while Q_n projects them into its orthogonal space. Since $H_0|n\rangle = E_n^0|n\rangle$, we can present H_0 as $H_0 = \sum_n E_n^0|n\rangle\langle n|$. It's then easy to verify that $[H_0, Q_n] = 0$.

Now we rewrite $(H_0 + \lambda V)|\psi_n\rangle = E_n|\psi_n\rangle$ as $(E_n - H_0)|\psi_n\rangle = \lambda V|\psi_n\rangle$. Put Q_n to the left of this equation, since H_0 and Q_n commute, we can project $|\psi_n\rangle$ into:

$$Q_n|\psi_n\rangle = \lambda \frac{1}{E_n - H_0} Q_n V|\psi_n\rangle \tag{8}$$

For further use, its important to know the spectral representation of $\frac{Q_n}{E_n - H_0}$, which we use R_n to denote:

$$R_n = \frac{Q_n}{E_n - H_0} = \sum_m{}' \frac{P_m}{E_n - E_m^0} \tag{9}$$

Thus $Q_n|\psi_n\rangle = \lambda R_n V|\psi_n\rangle$.

So the perturbed states can be divided into two parts:

$$\begin{aligned}|\psi_n\rangle = I|\psi_n\rangle &= (P_n + Q_n)|\psi_n\rangle \\ &= P_n|\psi_n\rangle + Q_n|\psi_n\rangle \\ &= |n\rangle\langle n|\psi_n\rangle + Q_n|\psi_n\rangle \end{aligned} \tag{10}$$

It's workable to normalize $\langle n|\psi_n\rangle$ to be 1, because $|n\rangle$ and $|\psi_n\rangle$ correspond to each other. Thus:

$$|\psi_n\rangle = |n\rangle + \lambda R_n V|\psi_n\rangle \tag{11}$$

This equation can iterate into a sequence of different orders:

$$|\psi_n\rangle = |n\rangle + \lambda R_n V|n\rangle + (\lambda R_n V)^2|n\rangle + \cdots \tag{12}$$

In sum, it's $|\psi_n\rangle = (1 - \lambda R_n V)^{-1}|n\rangle$. Meanwhile, in ordinary spectral representation:

$$\begin{aligned}|\psi_n\rangle =& |n\rangle + \lambda \sum_{m_1}{}' \frac{V_{m_1 n}}{E_n - E^0_{m_1}}|m_1\rangle + \lambda^2 \sum_{m_1}{}' \sum_{m_2}{}' \frac{V_{m_1 m_2} V_{m_2 n}}{(E_n - E^0_{m_1})(E_n - E^0_{m_2})}|m_1\rangle + \cdots \\ &+ \lambda^j \sum_{m_1}{}' \sum_{m_2}{}' \sum_{m_3}{}' \cdots \sum_{m_j}{}' \frac{V_{m_1 m_2} V_{m_2 m_3} \cdots V_{m_j n}}{(E_n - E^0_{m_1})(E_n - E^0_{m_2}) \cdots (E_n - E^0_{m_j})}|m_1\rangle + \cdots \end{aligned} \tag{13}$$

Using $\langle n|E_n|\psi_n\rangle = \langle n|H_0|\psi_n\rangle + \lambda\langle n|V|\psi_n\rangle$, we get $E_n = E^0_n + \lambda\langle n|V|\psi_n\rangle$, thus:

$$\begin{aligned}E_n =& E^0_n + \lambda V_{nn} + \lambda^2 \sum_{m_1}{}' \frac{V_{n m_1} V_{m_1 n}}{E_n - E^0_{m_1}} + \cdots \\ &+ \lambda^{j+1} \sum_{m_1}{}' \sum_{m_2}{}' \sum_{m_3}{}' \cdots \sum_{m_j}{}' \frac{V_{n m_1} V_{m_1 m_2} V_{m_2 m_3} \cdots V_{m_j n}}{(E_n - E^0_{m_1})(E_n - E^0_{m_2}) \cdots (E_n - E^0_{m_j})} + \cdots \end{aligned} \tag{14}$$

We can see immediately that BWPT's form is regular in any order, which makes high order perturbation much easier. Besides this, there are also advantages in BWPT[3]:

- BWPT encounters no divergence in degenerate cases. All denominators in BWPT are $E_n - E^0_{m_j}$, as long as perturbed $E_n \neq E^0_{m_j}$, there's no divergence. However, in RSPT, all denominators are in the form $E^0_{m_1} - E^0_{m_2}$, in degenerated cases where $E^0_{m_1} = E^0_{m_2}$, we have to diagonalize the system Hamiltonian instead.

- BWPT converge faster than RSPT in certain problems. For example, when:

$$H = \begin{pmatrix} 0 & 0 \\ 0 & \alpha \end{pmatrix} + \beta \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{15}$$

  BWPT gives $E = \frac{\beta^2}{E - \alpha}$ in second order, which is exact the analytical answer ($E^2 - \alpha E - \beta^2 = 0$) of this perturbed Hamiltonian, while RSPT not.

However, BWPT has also some shortcomings:

- Since E lies in both sides of equation [13] and [14], unlike RSPT, it's impossible to compute E_n analytically. Numerical calculation for perturbed energy in BWPT should start from a testing E_n and iterate until it converges.
- When using BWPT for two non-interactive bodies, sum of separate perturbation of one body gives different answer to perturbation of the whole system under the same order. This lacking of unity may lead to confusion, while RSPT is unified in this issue.

4 Generalized Brillouin-Wigner perturbation theory

The objective of refining perturbation theory is to converge to the exact answer in lower terms. Generally, we can divide the unperturbed state space into not just $1+(N-1)$ combination, for faster convergence, we can divide the P space into more than one state. Later we will discuss how this works.

First of all, we should generalize all the procedures in last section. For each ψ_n, we split the unperturbed eigenfunctions into subspace S_n and $\overline{S}_n$. The two projection operator is thus $P_n = \sum_{|k\rangle\in S_n} |k\rangle\langle k|$ and $Q_n = 1 - P_n$. Projected states are $|\psi_{n_S}\rangle = P_n|\psi_n\rangle$, $|\psi_{n_{\overline{S}}}\rangle = Q_n|\psi_n\rangle$. And again, H_0 commutes with Q_n, so equation [8] is still right:

$$Q_n|\psi_n\rangle = \lambda\frac{1}{E_n - H_0}Q_nV|\psi_n\rangle = \lambda R_nV|\psi_n\rangle \tag{16}$$

$$R_n = \frac{Q_n}{E_n - H_0} = \sum_{|m\rangle\in\overline{S}_n}\frac{|m\rangle\langle m|}{E_n - E_m^0} \tag{17}$$

We will get the same answer, just replace $|n\rangle$ with projected $|\psi_{n_S}\rangle$: $|\psi_n\rangle = (P_n + Q_n)|\psi_n\rangle = |\psi_{n_S}\rangle + \lambda R_nV|\psi_n\rangle$. Thus $|\psi_n\rangle = |\psi_{n_S}\rangle + \lambda R_nV|\psi_{n_S}\rangle + (\lambda R_nV)^2|\psi_{n_S}\rangle + \cdots$, setting $T_n = \lambda R_nV$, we get:

$$|\psi_n\rangle = |\psi_{n_S}\rangle + T_n|\psi_{n_S}\rangle + T_n^2|\psi_{n_S}\rangle + \cdots \tag{18}$$

Now we should talk about convergence in generalized form[4]. In general, when we get a perturbed result $|\psi_n\rangle$, we need it's mode $\langle\psi_n|\psi_n\rangle$ to be finite. This requires it's high orders' inner product, typically like $\langle\psi_n|T_n^{l\dagger}T_n^l|\psi_n\rangle \to 0$ when $l \to 0$. But of course, it will be better if $\langle\psi_n|\psi_n\rangle$ converges fast and stops at finite l. Now, write explicitly:

$$\langle\psi_n|T_n^{l\dagger}T_n^l|\psi_n\rangle = |\lambda|^{2l}\langle\psi_n|V\frac{1}{E_n - H_0}Q_n\cdots V\frac{1}{E_n - H_0}\underbrace{Q_nV\frac{1}{E_n - H_0}Q_n}$$
$$\underbrace{\frac{1}{E_n - H_0}}_{omit}\underbrace{Q_nV\frac{1}{E_n - H_0}Q_n}V\frac{1}{E_n - H_0}Q_nV\cdots\frac{1}{E_n - H_0}Q_nV|\psi_n\rangle$$

To check it more clearly, we put more $Q_nQ_n = Q_n$ (projection operators) into this term and underbrace important sequence:

$$\langle\psi_n|T_n^{l\dagger}T_n^l|\psi_n\rangle = |\lambda|^{2l}\langle\psi_n|V\frac{1}{E_n-H_0}Q_n\cdots\underbrace{Q_nV\frac{1}{E_n-H_0}Q_n}\ \underbrace{Q_nV\frac{1}{E_n-H_0}Q_n}$$
$$\underbrace{\frac{1}{E_n-H_0}}_{omit}\ \underbrace{Q_nV\frac{1}{E_n-H_0}Q_n}\ \underbrace{Q_nV\frac{1}{E_n-H_0}Q_n}\ Q_nV\cdots\frac{1}{E_n-H_0}Q_nV|\psi_n\rangle$$

Since there's only two subspace, the problem is how to split $\overline{S}_n$ out (or split how many states to form $\overline{S}_n$) to make the equation above converge. And this can be simplified to consider $Q_nV\frac{1}{E_n-H_0}Q_n$, which repeats in this term. And it's natural to regard that there are so many sequence of $Q_nV\frac{1}{E_n-H_0}Q_n$, so other terms' contribution: the rest left term $\langle\psi_n|V\frac{1}{E_n-H_0}Q_n$, the middle term $\frac{1}{E_n-H_0}$ and right term $Q_nV|\psi_n\rangle$ has been eliminated.

Thus, the only thing we should care is how can all the matrix elements product to become convergent within the chain of $\cdots Q_nV\frac{1}{E_n-H_0}Q_n\cdots Q_nV\frac{1}{E_n-H_0}Q_n\cdots$. Note that $Q_nV\frac{1}{E_n-H_0}Q_n$ is a sub-matrix in $\overline{S}_n$, because Q_n in left and right have projected V matrix into subspace $\overline{S}$. As discussed above, we believe the largest contribution lies in $Q_nV\frac{1}{E_n-H_0}Q_n$, and this operator is already in subspace $\overline{S}_n$.

Now using operator U_n to denote $Q_nV\frac{1}{E_n-H_0}Q_n$, and note that U_n is not an hermite operator. It's eigenfunctions and eigenvalues are:

$$U_n|\nu_n\rangle = u_n|\nu_n\rangle \tag{19}$$

We can assert that this term: $|\lambda|^{2l}\langle\nu_n|U_n^{2l}|\nu_n\rangle = |\lambda|^{2l}(u_n)^{2l}$ is very close to $\langle\psi_n|T_n^{l\dagger}T_n^l|\psi_n\rangle$, although the ket and bra are totally different, since those two term's related matrix elements are omitted. Meanwhile, though U_n is not an hermite operator, $|\lambda|^{2l}(u_n)^{2l}$ should nearly be real and equal to zero, because it nearly gives $\langle\psi_n|T_n^{l\dagger}T_n^l|\psi_n\rangle$. So $\langle\psi_n|T_n^{l\dagger}T_n^l|\psi_n\rangle \to 0$ is equivalent to $|\lambda|^{2l}(u_n)^{2l} \to 0$ and thus: $|\lambda|^{4l}|u_n|^{4l} \ll 1$. The requirement finally comes to be:

$$|\lambda u_n| \ll 1 \tag{20}$$

We should split out subspace $\overline{S}_n$ to meet this requirement. And naturally, with fewer states in this subspace, fewer matrix elements of V will appear in the inner product. However, when $\overline{S}_n$ is small, then S_n will be big. But as shown in equation [18], $|\psi_n\rangle$ is expanded upon $|\psi_{n_S}\rangle$, a big subspace S_n will leaves $|\psi_{n_S}\rangle$ more complex, leaving us fewer analytical information.

Above we have mentioned it will be better if $\langle\psi_n|\psi_n\rangle$ converges fast and stops at finite order. This requirement will help us find the best eigenfunctions of $\overline{S}_n$. Let's consider inner product

between ideal $|j\rangle \in \overline{S}_n$ with perturbed state $|\psi_n\rangle$:

$$\begin{aligned} C_{\alpha j} &= \langle j|\psi_n\rangle \\ &= \langle j|1 + T_n + T_n^2 + \cdots |\psi_{n_S}\rangle \end{aligned} \tag{21}$$

We need T_n's high order contribution to converge fast, and $\langle j|T_n^m|\psi_{n_S}\rangle$ can represent this. If we choose proper eigenfunctions $|j\rangle$, then $\langle j|T_n^m|\psi_{n_S}\rangle = 0$ for a fairly small m.

Besides, taking the expansion:

$$Q_n \lambda V|\psi_{n_S}\rangle = \sum_\nu h_\nu|\nu_n\rangle \tag{22}$$

Note that this is workable, because $V = \sum_m \sum_n V_{mn}|m\rangle\langle n|$, and its left vectors are projected by Q_n to $\overline{S}_n$ space, its right vectors combines with $|\psi_{n_S}\rangle$ to become numbers. Thus we will get:

$$\begin{aligned} C_{\alpha j} &= \langle j|\psi_n\rangle \\ &= \langle j|1 + T_n + T_n^2 + \cdots + T_n^{m-1}|\psi_{n_S}\rangle \\ &= \langle j|(1 + T_n + T_n^2 + \cdots + T_n^{m-1})\frac{1}{\lambda Q_n V}\sum_\nu h_\nu|\nu_n\rangle \end{aligned} \tag{23}$$

where $T_n = \lambda\frac{Q_n}{E_n - H^0}V$ and $U_n = Q_n V\frac{Q_n}{E_n - H^0}$. Then:

$$\begin{aligned} C_{\alpha j} &= \langle j|(1 + T_n + T_n^2 + \cdots + T_n^{m-1})\frac{1}{(E_n - H^0)T_n}\sum_\nu h_\nu|\nu_n\rangle \\ &= \langle j|\frac{1}{E_n - H^0}[1 + \lambda U_n + (\lambda U_n)^2 + \cdots + (\lambda U_n)^{m-1}]\frac{1}{T_n}\sum_\nu h_\nu|\nu_n\rangle \end{aligned} \tag{24}$$

Now taking the approximation that U_n's eigenvalue is close to real, and U_n should be chosen to be close to an hermite operator. Then T_n=λU_n:

$$\begin{aligned} C_{\alpha j} &= \frac{1}{E_n - H_j^0}\langle j|(1 + \lambda U_n + (\lambda U_n)^2 + \cdots + (\lambda U_n)^{m-1})\frac{1}{\lambda U_n}\sum_\nu h_\nu|\nu_n\rangle \\ &= \frac{1}{E_n - H_j^0}\langle j|\frac{1}{1 - (\lambda U_n)^{-1}}\frac{1}{\lambda U_n}\sum_\nu h_\nu(\lambda U_n)^{m-1}|\nu_n\rangle \\ &= \frac{1}{E_n - H_j^0}\langle j|\frac{1}{\lambda U_n - 1}\sum_\nu h_\nu|\nu_n\rangle(\lambda u_n)^{m-1} \\ &= \frac{1}{E_n - H_j^0}\frac{\sum_\nu h_\nu}{\lambda u_n - 1}\langle j|\nu_n\rangle(\lambda u_n)^{m-1} \end{aligned} \tag{25}$$

Here leaves us a natural way to fix m[1]: the smallest m to make $(\lambda u_n)^m = 0$. If we find proper $|j\rangle$ who meet m's requirement, then $|\psi_n\rangle$'s enpension coefficient on $\overline{S}_n$ is $C_{\alpha j}$.

[1]Still differ from the paper with -1

In Ref.[2], nonperturbative part of perturbed eigenfunctions (the zero order part $|\psi_{n_S}\rangle$) is used to check the relationship with prediction of RMT. In order to avoid degeneration to RSPT, more states from $\overline{S}_n$ should be taken into S_n, this forms the largest part of S_n, whose statistics agrees with RMT under chaotic Hamiltonian.

---

[1]Hubač I, Wilson S. Brillouin-Wigner Methods for Many-Body Systems[M]. Springer Netherlands, 2010.

[2]Wang W. Nonperturbative and perturbative parts of energy eigenfunctions: A three-orbital schematic shell model[J]. Physical Review E Statistical Nonlinear Soft Matter Physics, 2002, 65(3 Pt 2A):036219.

[3]Ingersent K. Brillouin-Wigner Perturbation Theory. (www.phys.ufl.edu/∼kevin/teaching)

[4]Wang W. Perturbative and nonperturbative parts of eigenstates and local spectral density of states: the wigner-band random-matrix model[J]. Physical Review E Statistical Physics Plasmas Fluids Related Interdisciplinary Topics, 1998, 61(1):952.