Regular Perturbation Theory and the Born-Oppenheimer Approximation

Sebastian Gherghe

March 5 2021

Abstract

Understanding the spectrum of a Schrödinger operator is key to understanding the dynamics of the corresponding quantum system. The study of this spectrum can be challenging, and often a more complicated operator is broken down into two pieces,

$$H(\kappa) = H_0 + \kappa V$$

where the spectrum of H_0 is known. Exactly how the spectrum of $H(\kappa)$ varies from $\kappa = 0$ (where one has a lot of information) to $\kappa = \kappa_0$ (where one desires information) is one of the topics of *perturbation theory*.

This talk will consist of two parts. In the first I will present the beautiful Kato-Rellich theory of "regular" perturbations, which gives simple criteria in the form of analytic families in the sense of Kato. I will show that it is sufficient in the case of the Schrödinger operator above for the perturbation V to be H_0-bounded. However more generally this method fails when considering unbounded perturbations. In the second part of this talk I will present the Feshbach projection method (or the Feshbach-Schur method) which is a powerful technique that can allow one to estimate the eigenvalues of $H(\kappa)$. I will finally describe an application to the scenario of the time-independent problem of the Born-Oppenheimer approximation for molecular dynamics.

1 Motivation: Quantum Mechanical Systems

1.1 Basic Theoretical Foundations

1.1.1 The Time-Dependent problem: Linear Evolution

To begin, let us recall the formulation of quantum mechanics via the Schrödinger equation. A basic tenet of quantum theory is that quantum systems, such as atoms, molecules, solids, and to some extent nuclei and even stars, are described by linear differential operators called *Schrödinger operators* or *Hamiltonians*. The Schrödinger operator for a quantum system is the linear partial differential operator

$$H = -\frac{h}{2m}\Delta + V \tag{1}$$

acting on the Hilbert space $L^2(\mathbb{R}^n)$. The constant m is the reduced mass of the system, and the constant $h = 2\pi\hbar$ is called Planck's constant. The real-valued function V is called the *potential*. This family of linear operators, for various potentials V, describes the different quantum systems mentioned above.

For fixed time t, elements of the Hilbert space $L^2(\mathbb{R}^n)$, called *wave functions*, represent the state of the system. The time-evolution of a wave function $\psi(t, x)$ for a quantum system with Schrödinger operator H is controlled by *Schrödinger's equation*,

$$i\hbar\frac{\partial}{\partial t}\psi = H\psi \tag{2}$$

For any reasonable initial state of the system ψ_0, the time-evolution is described by the one-parameter semi-group generated by H. A sufficient condition for this time-evolution operator to be well-defined is the self-adjointness of H. Hence we say that dynamics exist iff H is *self-adjoint*. The various scenarios of time-evolution under (2) in quantum mechanics is referred to as the *time-dependent problem*.

Recall that an operator A acting on a Hilbert space $\mathcal{H}$ is said to be self-adjoint if $A = A^*$, ie. A is a symmetric operator (ie. $\langle\psi, A\phi\rangle = \langle A\psi, \phi\rangle$ for any $\psi, \phi \in \mathcal{H}$) and $\mathcal{D}(A) = \mathcal{D}(A^*)$. An equivalent condition to self-adjointness that is easier to check is if A is closed, symmetric, and if $\mathrm{Ran}(A \pm i) = \mathcal{H}$. (Recall that an operator A with domain $\mathcal{D}(A)$ is said to be *closed* if its graph, $\Gamma(A)$ is a closed subset of $\mathcal{H} \times \mathcal{H}$).

We will not concern ourselves with problems of self-adjointness in this report. Whenever relevant, we shall assume that the operator/Hamiltonian in question is self-adjoint. Work by Kato (and the Kato-Rellich theorem for self-adjointness) has provided accessible conditions to verify in order to conclude self-adjointess. (See [2]).

1.1.2 The Time-Independent Problem: Spectral Theory

Recall that observables in quantum mechanics (ie. self-adjoint operators on $L^2(\mathbb{R}^n)$) correspond to the quantities that one can observe (ie. energy corresponds to the energy operator H, momentum to the momentum operator $p = -i\hbar\nabla$, etc). The postulates of quantum mechanics imply that, for a given observable A,

$$\sigma(A) = \{\text{values of the physical quantity represented by } A\} \tag{3}$$

where $\sigma(A)$ denotes the spectrum of A. The most important quantities are the *energy levels*, i.e. the spectrum of the Schrödinger operator H. The problem of finding $\sigma(H)$ for a given quantum system is referred to as the *time-independent problem*.

Let us recall some definitions again. Let A be a linear operator on X with domain $\mathcal{D}(A)$. We say A is invertible if there exists a bounded operator $A^{-1} : X \to \mathcal{D}(A)$ such that $A^{-1}A = 1_{\mathcal{D}(A)}$ and $AA^{-1} = 1_X$.

The *spectrum* of A, denoted by $\sigma(A)$, is the set of all points $\lambda \in \mathbb{C}$ for which $A - \lambda$ is not invertible. Consequently, the complement of $\sigma(A)$, denoted by $\rho(A)$, is call the *resolvent set* of A and is the set of all $\lambda \in \mathbb{C}$ such that $A - \lambda$ is invertible. If $\lambda \in \rho(A)$, then the inverse $R_A(\lambda) = (A - \lambda)^{-1}$ is called the *resolvent* of A at λ.

Theorem 1. *The resolvent set $\rho(A)$ is an open subset of $\mathbb{C}$ (and hence $\sigma(A)$ is closed) and $R_A(\lambda)$ is an analytic operator-valued function of λ on $\rho(A)$.*

A reference for this (and for the rest of this section) is [3]. Recall that a map $\lambda \in \rho \subset \mathbb{C} \to A(\lambda) \in \mathcal{L}(X)$ is said to be (norm) *analytic* at $\lambda_0 \in \rho$ if $A(\lambda)$ has a power series expansion in $(\lambda - \lambda_0)$ (with coefficients in $\mathcal{L}(X)$) that converges in norm with nonzero radius of convergence, i.e. there exist bounded operators A_n such that

$$A(\lambda) = \sum_{n=0}^{\infty} (\lambda - \lambda_0)^n A_n \tag{4}$$

There are some identities for the resolvent, which while fairly trivial are nonetheless useful. They are called the 1^{st} and 2^{nd} *resolvent identities*, respectively.

Lemma 1. *Let A, B be linear operators on X.*

1) Then for $\mu, \lambda \in \rho(A)$, the 1^{st} resolvent identity:

$$R_A(\lambda) - R_A(\mu) = (\lambda - \mu) R_A(\lambda) R_A(\mu) \tag{5}$$

2) If furthermore A and B are closed operators, with $z \in \rho(A) \cap \rho(B)$, the 2^{nd} resolvent identity:

$$R_A(z) - R_B(z) = R_A(z)(A - B)R_B(z) = R_B(z)(B - A)R_A(z) \tag{6}$$

Let us now recall the structure $\sigma(A)$ and the classification of its different parts. There are basically three ways in which $A - \lambda$ fails to be invertible:

(1) $\ker(A - \lambda) \neq 0$,

(2) $\ker(A - \lambda) = 0$, and $\mathrm{Ran}(A - \lambda)$ is dense so that $(A - \lambda)$ has a densely defined inverse but is unbounded,

(3) $\ker(A - \lambda) = 0$, but $\mathrm{Ran}(A - \lambda)$ is not dense; in this case $(A - \lambda)^{-1}$ exists and may be bounded on $\mathrm{Ran}(A - \lambda)$ but is not densely defined; therefore it cannot be *uniquely* extended to a bounded operator on X.

To account for these situations we classify $\sigma(A)$. If $\lambda \in \sigma(A)$ is such that $\ker(A-\lambda) \neq 0$, then λ is an *eigenvalue* of A and any $u \in \ker(A - \lambda), u \neq 0$ is an eigenvector of A for λ, satisfying $Au = \lambda u$. The lowest eigenvalue of a quantum Hamiltonian is referred to as the *ground state energy* and its eigenvectors are colloquially referred to as *ground states*.

We define the *discrete spectrum* of A, denoted $\sigma_d(A)$, as the set of all eigenvalues with finite algebraic multiplicity and which are isolated points of $\sigma(A)$. Similarly, the *essential spectrum* of A is defined as the complement of $\sigma_d(A)$ in $\sigma(A)$: $\sigma_{ess}(A) := \sigma(A) \setminus \sigma_d(A)$.

1.2 The Stationary Born-Oppenheimer Approximation

We will now describe the main physical scenario to be considered in this set of notes: the *Born-Oppenheimer approximation*, which lies at the foundation of quantum chemistry. It was first proposed by Born and Oppenheimer in their seminal paper [8]. Consider a molecule with N electrons and M nuclei of masses $m_1, \ldots, m_M$. Its state space is the symmetry subspace $L^2_{sym}(\mathbb{R}^{3(N+M)})$ of $L^2(\mathbb{R}^{3(N+M)})$ with Schrödinger operator (in units of $h = 1$ and $c = 1$):

$$H_{mol} = -\sum_1^N \frac{1}{2m}\Delta_{x_j} - \sum_1^M \frac{1}{2m_j}\Delta_{y_j} + V(x, y) \tag{7}$$

Where $x = (x_1, \ldots, x_N)$ and $y = (y_1, \ldots, y_M)$ are the electron and nuclear coordinates, and $V(x, y)$ is the sum of Coulomb interaction potentials between the electrons, between the electrons and the nuclei, and between the nuclei:

$$V(x, y) = \frac{1}{2}\sum_{i \neq j} \frac{e^2}{|x_i - x_j|} - \sum_{i,j} \frac{e^2 Z_j}{|x_i - y_j|} + \frac{1}{2}\sum_{i \neq j} \frac{e^2 Z_i Z_j}{|y_i - y_j|}$$

where the electrons all have charge $-e$ and the nuclei $Z_1 e, \ldots, Z_M e$ (for a neutral molecule, $\sum_j Z_j = N$.)

The molecule will evolve in time subject to the Schrödinger equation (SE):

$$i\hbar\partial_t \Psi = H_{mol}\Psi \tag{8}$$

with $\Psi(x, y, t) \in H^2(\mathbb{R}^{3(N+M)}), \forall t \in \mathbb{R}$.

Let us now remark on the heuristics of the time-dependent Born-Oppenheimer approximation. It is justified by the following physical fact: nuclei are much heavier than electrons and consequently move much slower. When a heavy and slow nucleus moves, that movement disrupts the electrostatic equilibrium of the electrons in the material, and as the electrons respond quickly, they settle into the most optimal configuration and apply a force acting back upon the nucleus (this is the source of "electronic friction". Also, since we assume the electron motion is almost instantaneous, we are in the so-called *adiabatic* regime.)

In the first step one assumes that the electrons instantaneously settle into their ground state, so the configuration of nuclei is fixed and one solves for the ground state of the electron configuration, i.e. for the ground state of the operator

$$H_e(y) = -\sum_1^N \frac{1}{2m}\Delta_{x_j} + V(x, y) \tag{9}$$

which describes the electrons only. Notice that this operator depends on the coordinates y of the nuclei as parameters. It is referred to in the literature as the *Born-Oppenheimer Hamiltonian* and we denote its ground state by $E(y)$ (the lowest eigenvalue in the problem $H_e(y)\psi(y, \cdot) = E(y)\psi(y, \cdot)$). We denote the ground state eigenfunction by $\psi_y(x)$). Actually computing eigenvalues and eigenfunctions of the electronic Schrödinger equation is the primary concern of computational quantum chemistry; see [11] and from a more mathematical viewpoint [12]. Let us just suppose that this problem is solved in some satisfactory way.

In the second step dynamics of nuclei are then modelled by using the ground state energy of the electron configuration as the potential (interaction) energy of the nuclear motion. We can define the nuclear Hamiltonian

$$H_{\text{nucl}} := -\sum_1^M \frac{1}{2m_j}\Delta_{y_j} + E \tag{10}$$

in which we've replaced (9) in (7) with $E = E(y)$, which is the operator of multiplication by $E(y)$.

Expectation: One expects that the eigenvalues of H_e give a good approximation to the eigenvalues of H_{mol} (the *time-independent* or *stationary problem*) and that the dynamics

generated by (10) approximate the true dynamics of the nuclei generated under the exact evolution (8) (the *time-dependent problem*).

In this document we will not discuss the time-dependent Born-Oppenheimer problem. (See for example [5] and [7]). Instead we will focus on the time-independent problem. We can consider

$$T_N = -\sum_1^M \frac{1}{2m_j}\Delta_{y_j} \tag{11}$$

so that we can write $H_{mol} = H_e + T_N$. The stationary problem looks like a perturbation problem $H = H_0 + \kappa W$ where $H_0 := H_e(y)$ and $\kappa W := T_N$ and we define the small parameter κ:

$$\kappa := 1/\min_j m_j$$

However, this is not a typical perturbation problem for two reasons: first the perturbation κW is not H_0-bounded, and second the operators H_0 and κW are operators in different coordinates and hence act on different coordinate spaces.

In what follows, we will first address the "typical" perturbation case of analytic families in the sense of Kato. While this won't apply directly to the above time-independent problem, we shall make use of it in a rather technical way in order to gain some smoothness assumptions on $\psi_y(x)$ in the parameter y. We will then use the isospectrality of the Feshbach projection map to derive a Hamiltonian

$$H_{eff} = H_e + v + w \tag{12}$$

where the extra two terms v, w contain derivatives of the electronic wavefunction $\psi_y(x)$ with respect to y. They are usually neglected in computations, first because they are expensive to compute or simply not available and second by the formal argument – to be taken with caution – that they carry the large nuclear masses $\{m_j\}_{j=1}^M$ in the denominator and are of *lower differentiation order* than the kinetic energy term. We will suggest this *only formally* by arguing that $v = O(\kappa)$ and $w = O(\kappa^2)$.

2 Regular Perturbation Theory (Point Spectra)

In this section we shall address the following situation. An operator H_0 has an eigenvalue E_0, which we will assume is in the discrete spectrum (say, an eigenvalue). Suppose that H_0 is perturbed "a little", ie. consider $H_0 + \kappa V$ where V is some other operator and $|\kappa|$ is small. What eigenvalues of $H_0 + \kappa V$ lie near E_0 and how are they related to V? What are their properties as functions of κ?

In quantum mechanics and other subjects, one may compute *formal* series for the perturbed eigenvalues, called *Rayleigh-Schrödinger series*. This is not only limited to Schrödinger

operators! The main source of this section is [1]. Please consult this source for proofs of all theorems.

2.1 Finite-dimensional perturbation theory

Let us first consider finite-dimensional matrices in order to present explicit formulas in the simplest cases. We shall use the results in this section as motivation for the study of the infinite-dimensional case (Furthermore, degenerate perturbation theory can be treated by reducing it to an essentially finite-dimensional problem.) Recall that E_0 is a *degenerate eigenvalue* when the characteristic equation $\det(H_0 - \lambda) = 0$ has a multiple root at $\lambda = E_0$.

Example 1. *Consider an elementary example:*

$$T(\kappa) = \begin{pmatrix} 1 & \kappa \\ \kappa & -1 \end{pmatrix} \tag{13}$$

$T(\kappa)$ is a matrix-valued analytic function. Its eigenvalues are

$$\lambda_\pm = \pm\sqrt{\kappa^2 + 1} \tag{14}$$

Let us observe a few points about this example.

- Clearly $T(\kappa)$ is entire in κ, but its eigenvalues as functions of κ have branch points at $\kappa = \pm i$.
- The branch points occur on the imaginary axis, not on the real one where $T(\kappa)$ is self-adjoint. So even if there are no branch points at "physical" values, the *perturbation series* (the Taylor series for $\lambda_\pm(\kappa)$ at $\kappa = 0$) has a finite radius of convergence.
- "Level crossing" takes place at the branch points, is at $\kappa = \pm i$ there are fewer distinct eigenvalues (one) than at the other points (where there are two)
- At the singular values of κ the matrix $T(\kappa)$ is not even diagonalizable. For example, we can explicitly compute

$$T(i)\begin{pmatrix} 2 \\ 2i \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad T(i)\begin{pmatrix} 1 \\ -i \end{pmatrix} = \begin{pmatrix} 2 \\ 2i \end{pmatrix}$$

Hence in the basis $\{\begin{pmatrix} 2 \\ 2i \end{pmatrix}, \begin{pmatrix} 1 \\ -i \end{pmatrix}\}$,

$$T(i) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Let us consider a more general case. Suppose $T(\kappa)$ is a matrix-value analytic function in a connected region R of the complex plane (not necessarily linear!) This finite-dimensional case is very useful to study, in that one can reduce an infinite-dimensional, linear, finitely degenerate perturbation problem to a finite-dimensional problem that may no longer be linear. (An example of this is the Feshbach-Schur projection method.)

To find the eigenvalues of this general $T(\kappa)$, we must solve the equation

$$\det(T(\kappa) - \lambda) = (-1)^n[\lambda^n + a_1(\kappa)\lambda^{n-1} + \cdots + a_n(\kappa)] \tag{15}$$

This leads us to the following theorem studying such functions.

Theorem 2. *Let $F(\kappa, \lambda) = \lambda^n + a_1(\kappa)\lambda^{n-1} + \cdots + a_n(\kappa)$ be a polynomial of degree n with leading coefficient 1 and whose coefficients are all analytic functions of κ. Suppose that $\lambda = \lambda_0$ is a simple root of $F(\kappa_0, \lambda)$. Then for $\kappa = \kappa_0$ there is exactly one root $\lambda(\kappa)$ of $F(\kappa, \lambda)$ near λ_0, and $\lambda(\kappa)$ is analytic in κ near $\kappa = \kappa_0$.*

The proof of this is a special case of the implicit function theorem. It immediately implies the following corollary:

Corollary 1. *Suppose $T(\kappa)$ is a matrix-value analytic function near κ_0 and suppose λ_0 is a simple eigenvalue of $T(\kappa_0)$. Then:*

a) For κ near κ_0, $T(\kappa)$ has exactly one eigenvalue, $\lambda_0(\kappa)$, near κ_0.

b) $\lambda_0(\kappa)$ is a simple eigenvalue if κ is near κ_0.

c) $\lambda_0(\kappa)$ is analytic near $\kappa = \kappa_0$.

In the case of eigenvalues with multiplicity greater than one, a more complicated but still straightforward analysis yields:

Corollary 2. *If $T(\kappa)$ is a matrix-analytic function near κ_0 and if λ_0 is an eigenvalue of $T(\kappa_0)$ with multiplicity m, then for κ near κ_0, $T(\kappa)$ has exactly m eigenvalues (counting multiplicity) near κ_0. These eigenvalues are all branches of one or more multi-valued functions analytic near κ_0 with at worst algebraic singularities at κ_0.*

The main take-away is that analytic perturbations preserve (locally) the dimension of the eigenspace. Now, if A and B are self-adjoint, then the perturbed eigenvalues of $A + \kappa B$ are analytic at $\kappa = 0$ even if A has degenerate eigenvalues. The following theorem of Rellich implies that the branch points in the previous corollary cannot occur in this case. The example at the beginning of this subsection shows that branch points can occur for *non-real* κ even in the "self-adjoint case", $T(\kappa)^* = T(\overline{\kappa})$.

Theorem 3. *(Rellich's Theorem) Suppose that $T(\kappa)$ is a matrix-valued analytic function in a region R containing a section of the real axis, and that $T(\kappa)$ is self-adjoint for κ on the real axis. Let λ_0 be an eigenvalue of $T(\kappa_0)$ of multiplicity m. If κ_0 is real, there are $p \leq m$ distinct functions $\lambda_1(\kappa), \ldots, \lambda_p(\kappa)$, single-valued and analytic in κ in a neighbourhood of κ_0, which are all eigenvalues.*

Let us now consider the special case $H(\kappa) = H_0 + \kappa V$. Suppose that E_0 is a non-degenerate eigenvalue of H_0. From Theorem 2, we know that for κ small $H(\kappa)$ has a unique eigenvalue $E(\kappa)$ near E_0 and that $E(\kappa)$ is analytic near $\kappa = 0$. The coefficients of its Taylor series are called *Rayleigh-Schrödinger coefficients* and the Taylor series is called the *Rayleigh-Schrödinger series*.

We shall explicitly write down some formulas to get a sense of what computations may look like. The formulas are simpler when H_0 is self-adjoint, so we restrict ourselves to that case. Since $E(\kappa)$ is the only eigenvalue of $H_0 + \kappa V$ near E_0, if $|E - E_0| < \epsilon$ for small ϵ, then $E(\kappa)$ is the only eigenvalue of the circle $\Gamma = \{E : |E - E_0| = \epsilon\}$. We use the spectral projection,

$$P(\kappa) = -\frac{1}{2\pi i} \oint_\Gamma (H_0 + \kappa V - E)^{-1} dE \tag{16}$$

ie. the projection onto the eigenvector with eigenvalue $E(\kappa)$. We shall justify the analyticity of the resolvent term in κ near $\kappa = 0$ later (see Proposition 1). Hence, $P(\kappa)$ itself is analytic in κ at $\kappa = 0$. So if Ω_0 is the unperturbed eigenvector, then $P(\kappa)\Omega \neq 0$ for κ small since $P(\kappa)\Omega_0 \to \Omega_0$ as $\kappa \to 0$. Since $P(\kappa)\Omega_0$ is an unnormalized eigenvector for $H(\kappa)$,

$$E(\kappa) = \frac{\langle \Omega_0, H(\kappa)P(\kappa)\Omega_0 \rangle}{\langle \Omega_0, P(\kappa)\Omega_0 \rangle} = E_0 + \kappa \frac{\langle \Omega_0, VP(\kappa)\Omega_0 \rangle}{\langle \Omega_0, P(\kappa)\Omega_0 \rangle} \tag{17}$$

This formula is very important for perturbation theory and will play a critical role in what follows! Put into words, it says

> To find the Taylor series for $E(\kappa)$, one only needs to find the Taylor series for $P(\kappa)$.

We do this based on the analyticity of the resolvent term $(H_0 + \kappa V - E)^{-1}$, ie. find its Taylor series and integrate it. But its Taylor series is just a geometric series:

$$\begin{aligned}(H_0 + \kappa V - E)^{-1} =& (H_0 - E)^{-1} - \kappa (H_0 - E)^{-1} V (H_0 - E)^{-1} \\ &+ \cdots + (-1)^n \kappa^n (H_0 - E)^{-1} [V (H_0 - E)^{-1}]^n \ldots\end{aligned}$$

(This is just the Neumann series expansion). Not only is this simple, but there is a nice form for the error term when the series is truncated. However, the actual Rayleigh-Schrödinger

series for $E(\kappa)$ is considerably more complicated because of the contour integration and the division of power series:

$$E(\kappa) = E_0 + \kappa \frac{\sum_{n=0}^{\infty} a_n \kappa^n}{\sum_{n=0}^{\infty} b_n \kappa^n} \tag{18}$$

where

$$a_n = \frac{(-1)^{n+1}}{2\pi i} \oint_\Gamma \langle \Omega_0, [V(H_0 - E)^{-1}]^{n+1} \Omega_0 \rangle dE \tag{19}$$

$$b_n = \frac{(-1)^{n+1}}{2\pi i} \oint_\Gamma \langle \Omega_0, (H_0 - E)^{-1} [V(H_0 - E)^{-1}]^{n+1} \Omega_0 \rangle dE \tag{20}$$

To illustrate, let us explicitly compute $E(\kappa)$ up to order κ^3. Since H_0 is self-adjoint, we can choose a basis of eigenvectors for the space $\{\Omega_0, \ldots, \Omega_{n-1}\}$, such that $H_0\Omega_i = E_i\Omega_i$. Denote $V_{ij} = \langle \Omega_i, V\Omega_j \rangle$. Then, making use of the spectral mapping theorem and Cauchy integral formula:

$$\begin{aligned} b_0 &= -\frac{1}{2\pi i} \oint_\Gamma \langle \Omega_0, (H_0 - E)^{-1} \Omega_0 \rangle dE \\ &= -\frac{1}{2\pi i} \oint_\Gamma (E_0 - E)^{-1} dE = 1 \end{aligned} \tag{21}$$

$$b_1 = \frac{1}{2\pi i} \oint_\Gamma V_{00} (E_0 - E)^{-2} dE = 0 \tag{22}$$

$$b_2 = -\frac{1}{2\pi i} \oint_\Gamma (E_0 - E)^{-2} \sum_{i=0}^{n} (E_i - E)^{-1} V_{0i} V_{i0} dE \tag{23}$$

Notice that in the last sum, the $i = 0$ term behaves very differently:

$$i = 0 : \qquad \frac{1}{2\pi i} \oint_\Gamma (E_0 - E)^{-3} dE = 0 \tag{24}$$

$$i \neq 0 : \qquad \frac{1}{2\pi i} \oint_\Gamma (E_0 - E)^{-2} (E_i - E)^{-1} dE = \frac{1}{2} (E_i - E_0)^{-2} \tag{25}$$

Hence

$$b_2 = -\frac{1}{2} \sum_{i \neq 0} (E_i - E_0)^{-2} V_{0i} V_{i0} \tag{26}$$

And similarly,

$$a_0 = V_{00} \tag{27}$$

$$a_1 = -\sum_{i \neq 0} (E_i - E_0)^{-1} V_{0i} V_{i0} \tag{28}$$

$$a_2 = \sum_{i \neq 0 \neq j} (E_i - E_0)^{-1} (E_j - E_0)^{-1} V_{0i} V_{ij} V_{j0} - 2 \sum_{i \neq 0} (E_i - E_0)^{-2} V_{0i} V_{i0} V_{00} \tag{29}$$

So finally, writing $E(\kappa) = E_0 + \kappa \sum_{n=1} \alpha_n \kappa^n$, we have the first three coefficients:

$$\alpha_1 = a_0 = V_{00} \tag{30}$$

$$\alpha_2 = a_1 = -\sum_{i \neq 0} (E_i - E_0)^{-1} V_{0i} V_{i0} \tag{31}$$

$$\alpha_3 = a_2 - b_2 a_0 \tag{32}$$

$$= \sum_{i \neq 0 \neq j} (E_i - E_0)^{-1} (E_j - E_0)^{-1} V_{0i} V_{ij} V_{j0} - \sum_{i \neq 0} (E_i - E_0)^{-2} V_{0i} V_{i0} V_{00} \tag{33}$$

and so on. The closed-form for the n^{th} Rayleigh-Schrödinger is more complicated than the leading order n^{th} term presented in most quantum mechanics textbooks:

$$\alpha_n = (-1)^{n+1} \sum_{i_1 \neq 0, i_2 \neq 0, \ldots, i_{n-1} \neq 0} \prod_{j=1}^{n-1} (E_{i_j} - E_0)^{-1} V_{0 i_1} V_{i_1 i_2} \ldots V_{i_{n-1} 0} + \ldots \tag{34}$$

We conclude this section with the following remark. Although the terms in the Taylor series are complicated, they arise effectively from a simple geometric series for the resolvent $(H_0 + \kappa V - E_0)^{-1}$. This suggests that to deduce results about $E(\kappa)$ in the infinite-dimensional case, one needs results on the resolvent. Information about the eigenvalues is then obtained via formulas that give the eigenvalue as a ratio of contour integrals of matrix elements of the resolvent.

One more result, whose equivalent in the infinite-dimensional will be extremely important, is about the existence and regularity of an eigenvector $\Omega(\kappa)$ for the perturbed eigenvalue $E(\kappa)$.

Theorem 4. *Let Ω_0 be a non-degenerate eigenvector for T_0 such that $T_0 \Omega_0 = E_0 \Omega_0$. Let $T(\kappa)$ be a matrix-valued analytic function with $T(0) = T_0$. Then for κ small there is a vector-valued analytic function $\Omega(\kappa)$ that obeys*

$$T(\kappa)\Omega(\kappa) = E(\kappa)\Omega(\kappa) \tag{35}$$

where $E(\kappa)$ is the eigenvalue of $T(\kappa)$ near E_0. Moreover if $T(\kappa)$ is self-adjoint for κ real, then $\Omega(\kappa)$ can be chosen so that $\|\Omega(\kappa)\| = 1$ for κ real.

2.2 Regular Perturbation Theory

We now finally deal with the infinite-dimensional case. We will argue that under very general circumstances, the Rayleigh-Schrödinger series has a nonzero radius of convergence for perturbations of unbounded operators in infinite-dimensional Hilbert spaces.

2.2.1 Discrete Spectra

In what follows we will make heavy use of what is called the *Riesz projection*, which we denote by P_λ. We define it in the next theorem. Before stating it we will define an equivalent definition of the discrete spectrum. A point $\lambda \in \sigma(A)$ is called *discrete* if λ is isolated and P_λ is finite-dimensional. If P_λ is one-dimensional, we say λ is a *non-degenerate* eigenvalue.

Theorem 5. *Let A be a closed operator with $\{\mu : |\mu - \lambda| = r\} \subset \rho(A)$. Then*

$$P_\lambda = -\frac{1}{2\pi i} \oint_{|\mu-\lambda|=r} (A-\mu)^{-1} d\mu \tag{36}$$

is a projection. If P has dimension $n < \infty$, then A has at most n points of its spectrum in $\{\mu : |\mu - \lambda| < r\}$ and each is discrete. If $n = 1$, there is exactly one spectral point in $\{\mu : |\mu - \lambda| < r\}$ and it is non-degenerate.

The Riesz projection (36) is also called a *Riesz integral* for A and λ (in the case that it has dimension 1.) The fact that it is a projection is proved below, but in general it is not orthogonal. In the non-degenerate case it follows that $\mathrm{Ran} P_\lambda \supset \ker(A-\lambda)$, while P_λ is only the orthogonal projection onto $\ker(A-\lambda)$ if A is a self-adjoint operator on a Hilbert space.

Proof. Let us first argue that P is a projection. Let $r < R < \epsilon$. Using the 1^{st} resolvent identity (5),

$$\begin{aligned}
P_\lambda^2 &= \frac{1}{(2\pi i)^2} \oint_{|\mu-\lambda|=r} \oint_{|\nu-\lambda|=R} (A-\mu)^{-1}(A-\nu)^{-1} d\nu d\mu \\
&= \frac{1}{(2\pi i)^2} \oint_{|\mu-\lambda|=r} \oint_{|\nu-\lambda|=R} (\nu-\mu)^{-1}[(A-\nu)^{-1} - (A-\mu)^{-1}] d\mu d\nu \\
&= \frac{1}{(2\pi i)^2} [\oint_{|\nu-\lambda|=R} (A-\nu)^{-1} d\nu \oint_{|\mu-\lambda|=r} (\nu-\mu)^{-1} d\mu \\
&\qquad - \oint_{|\mu-\lambda|=r} (A-\mu)^{-1} d\mu \oint_{|\nu-\lambda|=R} (\nu-\lambda)^{-1} d\nu] \\
&= \frac{1}{(2\pi i)^2} [0 - \oint_{|\mu-\lambda|=r} (2\pi i)(A-\mu)^{-1} d\mu] \\
&= P_\lambda
\end{aligned} \tag{37}$$

We note that it can also be argued using Cauchy's theorem that P_λ is independent of the contour chosen, as long as the area inside the contour intersects the spectrum only at the point λ.

Let us now argue that the spaces $G = \mathrm{Ran}\ P$ and $F = \mathrm{Ker}\ P$ are closed, complementary subspaces. In particular, this means that $G + F = \mathcal{H}$ and $G \cap F = \{0\}$. These follow from

elementary arguments given equation (37), ie. that $P_\lambda^2 = P_\lambda$. Note that P_λ is not necessarily orthogonal (this occurs only if A is self-adjoint). Next let us argue that G and F are invariant under A. This means the following: $G \subset \mathcal{D}(A)$, $AG \subset G$ and $F \cap \mathcal{D}(A)$ is dense in F, $A[F \cap \mathcal{D}(A)] \subset F$.

Note that by definition $\mathcal{D}(A)$ is dense in $\mathcal{H}$, hence $F \cap \mathcal{D}(A)$ is dense in F as F is a closed subspace. Now let us argue $G \subset \mathcal{D}(A)$. From the relation $A(A-z)^{-1} = 1 + z(A-z)$, it follows that A commutes with its resolvent, ie. $A(A-z)^{-1} = (A-z)^{-1}A$. Now, write P_λ as a limit of Riemann sums, i.e.

$$P_\lambda^n = \frac{1}{2\pi i} \sum_{i=1}^{2^n} \Delta_i^n (A - z_i)^{-1}$$

Recalling that Γ is a circle, we choose a dyadic partition of the length of the circle into $|\Gamma|/2^N$ and connect those points by straight line segments of length $\Delta_i^n \leqslant \dfrac{|\Gamma|}{2^n}$. We choose z_i^n to be the endpoint of the i^{th} segment. It then follows that P_λ^N are a norm-Cauchy sequence, and hence A commutes with each P_λ^N and thus also with P_λ. Hence any $\psi \in G$ (ie. $\psi = P_\lambda \psi$) can be written as a limit of $\psi_n := P_\lambda^n \psi$. It follows by a similar argument to the above that the sequence $\{A\psi_n\}$ is also Cauchy, i.e.

$$A\psi_n = \frac{1}{2\pi i} \sum_{i=1}^{2^n} \Delta_i^n + \Delta_i^n z_i^n (A - z_i^n)^{-1}$$

is also norm-Cauchy, and hence since A is closed it follows that $\psi \in \mathcal{D}(A)$. Finally, $AG \subset G$ and $A[F \cap \mathcal{D}(A)] \subset F$ follow from the fact that $AP_\lambda = P_\lambda A$.

Now, consider

$$R_\lambda = -\frac{1}{2\pi i} \oint_{|\mu - \lambda| = r} (\lambda - \mu)^{-1} (A - \mu)^{-1} d\mu$$

By using the 1^{st} resolvent identity and computing similar to equation (37), we obtain

$$R_\lambda P_\lambda = P_\lambda R_\lambda (A - \lambda) R_\lambda = R_\lambda (A - \lambda) = 1 - P_\lambda \tag{38}$$

The last equality in equation (2.2.1) applies to vectors in $\mathcal{D}(A)$. We conclude that R_λ takes Ker P_λ into itself, and that $(A_2 - \lambda)R_\lambda = R_\lambda(A_2 - \lambda) = 1$ on Ker P_λ. From this we conclude that $\nu \notin \sigma(A_2)$ if $|\nu - \lambda| < r$, and hence $(A - \nu)^{-1}$ exists if and only if $(A_1 - \nu)^{-1}$ exists.

If G is finite dimensional, A_1 has eigenvalues $\nu_1, \ldots \nu_k$ with $k \leqslant n$, so that $\sigma(A) \cap \{\nu : |\nu - \lambda| < r\}$ is a finite set. Note that Ran $P_\lambda \subset$ Ran P, and hence each spectral point in the circle is discrete.

□

2.2.2 Analytic Families and the Kato-Rellich Theorem

Let us now define our desired families of operators. A (possibly unbounded) operator-valued function $T(\kappa)$ on a complex domain R is called an *analytic family* or an *analytic family in the sense of Kato* if and only if:

(i) For each $\kappa \in R$, $T(\kappa)$ is closed and has a nonempty resolvent set.

(ii) For every $\kappa_0 \in R$, there is a $\lambda_0 \in \rho(T(\kappa_0))$ such that $\lambda_0 \in \rho(T(\kappa))$ for κ near κ_0 and $(T(\kappa) - \lambda_0)^{-1}$ s an analytic operator-valued function of κ near κ_0.

We shall give sufficient conditions for Schrödinger operators to be analytic in the sense of Kato in the next section. Next, we note that the number λ_0 above does not play a special role.

Proposition 1. *Let $\{H(\kappa)\}$ be an analytic family on a domain R. Then*

$$\Gamma = \{(\kappa, \lambda) : \kappa \in \mathbb{R}^{3N},\ \lambda \in \rho(H(\kappa))\} \tag{39}$$

is open and the resolvent $(H(\kappa) - \lambda)^{-1}$ defined on Γ is analytic in both variables.

Proof. First let us argue that Γ is open. Let $(kappa_0, \lambda_1) \in \Gamma$. Suppose that $(H(\kappa) - \lambda_0)^{-1}$ exists and is analytic in κ near κ_0. By the 1^{st} resolvent identity (5),

$$[1 - (\lambda_1 - \lambda_0)(H(\kappa_0) - \lambda_0)]^{-1} = (H(\kappa_0) - \lambda_0)(H(\kappa_0) - \lambda_1)^{-1}$$

Using that the set of invertible linear operators is open, for some (κ, λ) sufficiently close to (κ_0, λ_0), the operator $1 - (\lambda - \lambda_0)(H(\kappa) - \lambda_0)^{-1}$ is invertible. For such (κ, λ) we have

$$(H(\kappa) - \lambda)^{-1} = (H(\kappa) - \lambda_0)^{-1}[1 - (\lambda - \lambda_0)(H(\kappa) - \lambda_0)^{-1}]^{-1}$$

and hence $\lambda \in \rho(H(\kappa))$, thus $(\kappa, \lambda) \in \Gamma$. Thus Γ is open.

Now, note that $1 - (\lambda - \lambda_0)(H(\kappa) - \lambda_0)^{-1}$ is analytic for λ near λ_0 and for κ near κ_0. It then follows that $(H(\kappa) - \lambda)^{-1}$ is analytic in κ and analytic in λ by a general theorem. □

Next we shall need a short and basic lemma. Then we will be ready to state the main theorem.

Lemma 2. *If P and Q are two (not necessarily orthogonal) projections and dim(Ran P) $\neq$ dim(Ran Q), then $\|P - Q\| \geqslant 1$. In particular if $P(x)$ is a continuous projection-valued function of x on a connected topological space, then dim(Ran $P(x)$) is a constant.*

Proof. Without loss of generality suppose $\dim(\text{Ran } P) < \dim(\text{Ran } Q)$. Let $F = \text{Ker } P$ and $E = \text{Ran } Q$. Then $\dim(F^\perp) = \dim(\text{Ran } P) < \dim E$.

This implies that $E \cap F \neq \{0\}$. Consider first that if $E \cap F^\perp = \{0\}$, this is automatically true. Otherwise, the closed subspace $E \cap F^\perp$ will be non-empty and of dimension less than that of E. Choosing an orthonormal basis for $E \cap F^\perp$ and extending it to one for E, we obtain an orthonormal basis for $E \cap F$, and hence that space is non-empty. Then there exists an orthonormal $\psi \neq 0$ with $\psi \in E \cap F$. Hence $P\psi = 0$, $Q\psi = \psi$, so $\|(P-Q)\psi\| = \|\psi\| = 1$, and hence $\|P - Q\| \geqslant 1$.

If $P(x)$ is a continuous projection-valued function of x, then for any $\epsilon > 0$ there exists a $\delta > 0$ such that $\|x_1 - x_2\| < \delta$ implies $\|P(x_1) - P(x_2)\| < \epsilon$. However if $\dim(\text{Ran } P(x_1)) \neq \dim(\text{Ran } P(x_2))$ then $\|P(x_1) - P(x_2)\| \geqslant 1$. This is an obvious contradiction. By connectedness we can extend this argument over the entire topological space, so $\dim(\text{Ran } P(x))$ must be a constant. □

Now we are ready to prove the Kato-Rellich theorem for analytic families! It is the infinite-dimensional analogue of our main results in the last section. What is most important is the analyticity of the eigenvalue $E(\kappa)$ in κ and of the eigenvector $\Omega(\kappa)$ in κ.

Theorem 6. *(Kato-Rellich theorem) Let $T(\kappa)$ be an analytic family in the sense of Kato. Let E_0 be a non-degenerate discrete eigenvalue of $T(\kappa_0)$. Then, for κ near κ_0, there is exactly one point $E(\kappa)$ of $\sigma(E(\kappa))$ near E_0 and this point is isolated and non-degenerate. Furthermore $E(\kappa)$ has an eigenvector $\Omega(\kappa)$, and for κ near κ_0 both are analytic. If $H(\kappa)$ is self-adjoint for $\kappa - \kappa_0$ real, then $\Omega(\kappa)$ can be normalized for $\kappa - \kappa_0$ real.*

Proof. Let $\epsilon > 0$ such that $\sigma(T(\kappa_0)) \cap \{E : |E - E_0| \leqslant \epsilon\} = \{E_0\}$. Observe that the circle $\{E : |E - E_0| = \epsilon\}$ is compact and recall the set Γ (39) is open. Hence, there exists $\delta > 0$ such that $E \in \rho(T(\kappa))$ if $|E - E_0| = \epsilon$ and $|\kappa - \kappa_0| \leqslant \delta$. Then (via repeated applications of the 2^{nd} resolvent identity), the projection

$$P(\kappa) = -\frac{1}{2\pi i} \oint_{|E-E_0|=\epsilon} (T(\kappa) - E)^{-1} dE$$

exists and is analytic for $\kappa \in N := \{\kappa : |\kappa - \kappa_0| \leqslant \delta\}$.

The non-degeneracy of E_0 as an eigenvalue of $T(\kappa_0)$ implies that $P(\kappa_0)$ is one-dimensional and hence by Lemma (2) $P(\kappa)$ is one-dimensional for all $\kappa \in N$. Hence by Theorem 5 there is exactly one eigenvalue $E(\kappa)$ of $T(\kappa)$ with $|E(\kappa) - E_0| < \epsilon$ when $\kappa \in N$ and it is non-degenerate. The analyticity of $E(\kappa)$ follows from the identity

$$(E(\kappa) - E_0 - \epsilon)^{-1} = \frac{\langle \Omega_0, (T(\kappa) - E_0 - \epsilon)^{-1} P(\kappa)\Omega_0 \rangle}{\langle \Omega_0, P(\kappa)\Omega_0 \rangle} \tag{40}$$

and we obtain an analytic eigenvector by choosing $\Omega(\kappa) = P(\kappa)\Omega_0$, or

$$\Omega(\kappa) = (\langle \Omega_0, P(\kappa)\Omega_0 \rangle)^{-1/2} P(\kappa)\Omega_0 \tag{41}$$

in the real case, when Ω_0 is the unperturbed eigenvector. □

2.3 Application to Schrödinger Operators

So, we've seen how easy it is to prove that energy levels are analytic in the coupling constant once we know that $T(\kappa)$ is an analytic family. What makes this very useful is having a convenient criteria for $T(\kappa)$ to be analytic. There are two simple criteria, reflecting the operator/form duality, we will only focus on the operator version.

Let R be a connected domain in the complex plane and let $T(\kappa)$, a closed operator with non-empty resolvent set, be given for each $\kappa \in \mathbb{R}$. Then, we say that $T(\kappa)$ is an *analytic family of type (A)* if and only if

(i) The operator domain of $T(\kappa)$ is some set D independent of κ.

(ii) For each $\psi \in D$, $T(\kappa)\psi$ is a vector-valued analytic function of κ.

We will only mention that is a more general result that every family of type (A) is an analytic family in the sense of Kato. This also holds for Schrödinger operators, but we will not provide the proof here.

Theorem 7. *Let $H_0 + \kappa V$ be an analytic family of type (A) in the region R. Then $H_0 + \kappa V$ is an analytic family in the sense of Kato. In particular, if $0 \in R$ and E_0 is an isolated non-degenerate eigenvalue of H_0, then there is a unique point $E(\kappa)$ of $\sigma(H_0 + \kappa V)$ near E_0 when $|\kappa|$ is small which is an isolated non-degenerate eigenvalue. Moreover, $E(\kappa)$ is analytic near $\kappa = 0$.*

This theorem is essentially the only case we will use. However, when are Schrödinger operators of the form $T(\kappa) = H_0 + \kappa V$ analytic families of type (A)? We will provide and provide a sufficient criterion now:

Lemma 3. *Let H_0 be a closed operator with nonempty resolvent set. Define $H_0 + \kappa V$ on $\mathcal{D}(H_0) \cap \mathcal{D}(V)$. Then $H_0 + \kappa V$ is an analytic family of type (A) near $\kappa = 0$ if and only if*

(a) $\mathcal{D}(V) \supset \mathcal{D}(H_0)$

(b) For some a and b and for all $\psi \in \mathcal{D}(H_0)$,

$$\|V\psi\| \leqslant a\|H_0\psi\| + b\|\psi\| \tag{42}$$

That is, $H_0 + \kappa V$ is type (A) if and only if V *is* H_0*-bounded.* The infimum of such an a is called the *relative bound* of V with respect to A. If the relative bound is zero, we say B is *infinitesimally small* with respect to A and write $B << A$. Usually b must be chosen larger as a is chosen smaller.

This is an important concept in the *perturbations of self-adjoint operators*, which give conditions on B for $A + B$ to be self-adjoint when A already is. Surprisingly (or not!) the fundamental result is also of Kato-Rellich:

Theorem 8. *(Kato-Rellich) Suppose that A is self-adjoint, B is symmetric, and B is A-bounded with relative bound $a < 1$. Then $A + B$ is self-adjoint on $\mathcal{D}(A)$. Furthermore, if A is bounded below by M, then $A + B$ is bounded below by $M - max\{b/(1 - a), a|M| + b\}$, where a and b are given by (42).*

This is the method by which Kato (see [2] for discussion and more historical context, and [13] for his original paper) first proved the self-adjointness of the molecular Hamiltonian (7). He showed the Coulomb potentials in (7) are $-\Delta$-bounded, by showing that $V = L^2 + L^\infty$. As we see from the lemma, operators of the form $-\Delta + \kappa V$ where V is Coulomb are analytic families of type (A).

Let us now prove our lemma.

Proof of Lemma 3. $\rightarrow$: First suppose that $H_0 + \kappa V$ is an analytic family of type (A). Then, $\mathcal{D}(H_0) = \mathcal{D}(H_0 + \kappa V) = \mathcal{D}(H_0) \cap \mathcal{D}(V)$, so condition (a) holds. Next, since H_0 is closed, $D(H_0)$ with the norm $|||\psi||| = \|H_0\psi\| + \|\psi\|$ is a Banach space $\hat{\mathcal{D}}$. Choose $\kappa > 0$ small enough so that κ and $-\kappa$ are both in the domain of analyticity. $H_0 + \kappa V : \hat{\mathcal{D}} \rightarrow \mathcal{H}$ is everywhere defined and has a closed graph in $\hat{\mathcal{D}} \times \mathcal{H}$ since the graph is closed in $\mathcal{H} \times \mathcal{H}$ with a weaker topology. Hence by the closed graph theorem,

$$\|(H_0 + \kappa V)\psi\| \leqslant a_1|||\psi||| \quad \text{and} \quad \|(H_0 - \kappa V)\psi\| \leqslant a_2|||\psi|||$$

Hence,

$$\begin{aligned}\|V\psi\| &\leqslant \frac{1}{2\kappa}[\|(H_0 + \kappa V)\psi\| + \|(H_0 - \kappa V)\psi\|] \\ &\leqslant \frac{1}{2\kappa}(a_1 + a_2)|||\psi|||\end{aligned}$$

hence, the condition (b) holds!
$\leftarrow$: Now let us assume that relative bound conditions (a) and (b) hold. Then for $\psi \in \mathcal{D}(H_0)$,

$$\begin{aligned}\|H_0\psi\| &\leqslant \|(H_0 + \kappa V)\psi\| + |\kappa|\|V\psi\| \\ &\leqslant \|(H_0 + \kappa V)\psi\| + |\kappa|a\|H_0\psi\| + |\kappa|b\|\psi\|\end{aligned}$$

So, if κ small enough such that $|\kappa| < 1/a$, then

$$\|H_0\psi\| \leqslant \frac{1}{1-|\kappa|a}\|(H_0+\kappa V)\psi\| + \frac{b|\kappa|}{1-|\kappa|a}\|\psi\|$$

So, consider a sequence $\psi_n \to \psi$ in $\mathcal{H}$ with $\psi_n \in \mathcal{D}(H_0)$ and $(H_0+\kappa V)\psi_n$ Cauchy, then $H_0\psi$ is Cauchy by the above inequality. Hence $\psi \in \mathcal{D}(H_0)$. Hence, $H_0 + \kappa V$ is closed on $\mathcal{D}(H_0)$, independent of κ. Clearly $(H_0+\kappa V)\psi$ is analytic in κ for $\psi \in \mathcal{D}(H_0)$. □

Let us conclude this entire section by giving an example unrelated to our main application for brevity. This is the *Helium atom.*

Example 2. *It can be shown that if $V << H_0$ and $W << H_0$, then $W << H_0 + V$. So, letting*

$$H_0 = -\Delta_1 - \Delta_2 - \frac{2}{r_1} - \frac{2}{r_2}, \qquad V = \frac{1}{|r_1 - r_2|} \tag{43}$$

on $L^2(\mathbb{R}^6)$, we see that H_0+ is an analytic family of type (A). In the approximation of infinite nuclear mass, $H_0 + V$ is the helium atom Hamiltonian.

The single particle operator

$$h = -\Delta_1 - \frac{2}{r_1}$$

is an operator with an exactly solvable eigenvalue problem. Its lowest eigenvalue is $E = -1$. Notice then that $H_0 = h \otimes 1 + 1 \otimes h$ on $L^2(\mathbb{R}^3) \otimes L^2(\mathbb{R}^3) = L^2(\mathbb{R}^6)$, so by other work its ground state energy is -2. Hence, for $|\kappa|$ small, the ground state energy $E(\kappa)$ is analytic with the Taylor coefficients at $\kappa = 0$ given by the Rayleigh-Schrödinger coefficients/formula given by (18) and the following computations.

In practical applications, one is interested in the ground state energy $E(1)$ of the helium atom. Two situations arise:

1) *Is the radius of convergence for the Rayleigh-Schrödinger series for $E(\kappa)$ around $\kappa = 0$ is bigger than 1? This is projected to be true but a difficult question to answer.*

2) *When $\beta = 1$, even if the series is convergent, a large number of terms are required to approximate $E(1)$ well. This is a problem because the higher-order terms are very difficult to compute. The first order approximation value $V_{00} = \langle \Omega_0, V\Omega_0 \rangle$ for $E(1) - E(0)$ is off from experimental values by 15%. With other methods better accuracy than 1% can be obtained (See [1]). Compared to these methods however, in the regime of κ small, perturbation theory is more accurate.*

3 The Feshbach-Schur Method

We now begin to address the main scenario of this talk: the time-independent Born-Oppenheimer problem. The main idea behind this approach is to use the Feshbach-Schur map to construct an isospectral Hamiltonian: ie. another operator that by construction must share an eigenvalue at 0. Solving the eigenvalue problem of this isospectral Hamiltonian will yield a solution for our original one.

The advantage of approaches such as this one is that they transform the problem of determining (some of) the eigenvalues of a self-adjoint operator H on a Hilbert space H into the solution of a non-linear equation in a *smaller* Hilbert space K. Often the space K is finite-dimensional and the solution is found by algebraic means. When the spectrum of the Hamiltonian is at least partly continuous some refinements are needed to make this argument in a rigorous proof. See [6] for a rigorous introduction to the Feshbach method, and [9] and [10] for rigorous results on the time-independent theory by constructing a locally isospectral effective Hamiltonian.

3.1 The Feshbach-Schur Map

In this section we define the Feshbach-Schur map for a self-adjoint Hamiltonian H acting on a Hilbert space $\mathcal{H}$. Let P and $\overline{P}$ be orthogonal projections (ie. P and $\overline{P}$ are self-adjoint operators with $P^2 = P$, $\overline{P}^2 = \overline{P}$ and $P+\overline{P} = 1$.) We require that we choose these projections so that

I) We have Ran $P \subset \mathcal{D}(H)$ and $H_{\overline{P}} := \overline{P}H\overline{P}|_{\text{Ran }\overline{P}}$ is invertible.

II) The following conditions hold for the operator $R_{\overline{P}} = \overline{P}H_{\overline{P}}^{-1}\overline{P}$,

$$\|R_{\overline{P}}\| < \infty, \quad \|PHR_{\overline{P}}\| < \infty, \quad \|R_{\overline{P}}HP\| < \infty \tag{44}$$

In doing so we can define the *Feshbach-Schur operator/map*:

$$F_P(H) := P(H - HR_{\overline{P}}H)P|_{\text{Ran }P} \tag{45}$$

The main result is that this operator is well-defined and isospectral at 0 with H. Moreover, under these conditions the self-adjointness of $F_P(H)$ follows from that of H (though we shall not prove this.)

Theorem 9 (Isospectrality Theorem)**.** *Assume that properties 1.I) and 1.II) hold. Then the operators H and $F_P(H)$ are isospectral at* 0*, in the sense that:*

a) $0 \in \sigma(H)$ *if and only if* $0 \in \sigma(F_P(H))$

b) $H\psi = 0$ *if and only if* $F_P(H)\phi = 0$

where ψ *and* ϕ *are related by* $\phi = P\psi$ *and* $\psi = Q\phi$*, where* Q *is a bounded operator known as the resolvent and is defined by:*

$$Q = Q(H) = P - R_{\overline{P}}HP$$

We shall not record the proof of this here. See [4].

3.2 Main Result

In order to apply our method to the time-independent Born-Oppenheimer problem, we make some approximations. From now on we shall refer to the spaces $L^2(\mathbb{R}^{3(N+M)})$ as L^2, $L^2(\mathbb{R}^{3N})$ as L^2_x, and so on. Our main result is the following:

Theorem 10. *E_0 is the ground state energy of H_{mol} if and only if it solves the equation $\lambda = \nu_0(\lambda)$, where $\nu_0(\lambda)$ is the ground state energy of the effective nuclear Hamiltonian $H_e(\lambda)$ as defined below:*

$$H_e(\lambda) := H_e + v + T(\lambda) \tag{46}$$

on $L^2(\mathbb{R}^{3M})$, where $v = O(\kappa)$ and $T(\lambda) = O(\kappa^2)$ act on $L^2(\mathbb{R}^{3N})$: $T(\lambda)$ as in in eq (59) and v is given by (assuming ψ_y is real, which happens if the ground state is non-degenerate)

$$v(y) := \sum_1^M \frac{1}{2m_j} \int |\nabla_{y_j}\psi_y|^2 dx \tag{47}$$

and by (62) in the case when it is complex. To second order E_0 is the ground state energy of $H_{eff} := H_e + v$. In the complex ψ_y case this operator can be written as:

$$H_{eff} = -\sum_1^M \frac{1}{2m_j}(\nabla_{y_j} - iA_j)^2 + E(y) + \tilde{v} \tag{48}$$

where $A_j = i\langle\psi_y, \nabla_{y_j}\psi_y\rangle$ and

$$\tilde{v} = -\sum_1^M \frac{1}{2m_j} c\|[P, \nabla_{y_j}]\psi_y\|^2 \tag{49}$$

We define the ground state of the operator $H_e(y)$ to be $\psi_y(x)$ with ground state energy $E(y)$ (We assume for simplicity that the ground state energy $E(y)$ is non-degenerate.) Minimizing this energy with respect to y would yield the equilibrium positions of the nuclei, ie. the shape of the molecule. It has been shown (see [1] or [4]) that the function $E(y)$ increases as $|y| \to \infty$. Furthermore it is bounded above by 0. It hence has some minimum $\lambda_0 := \min_y E(y)$ (which may not be unique.)

We define P to be the orthogonal projection onto the ground state of H_e, ie.

$$(P\Psi)(x,y) = \psi_y(x) \int \overline{\psi_y(x)}\Psi(x,y)dx \tag{50}$$

where the ground state $\psi_y(x)$ is normalized as $\int |\psi_y(x)|^2 dx = 1$. This is in fact an orthogonal projection on L^2 is easy to show. We further define $\overline{P}$ to be the orthogonal complement to P, ie. $P + \overline{P} = 1$.

For technical reasons we will modify the original Hamiltonian by replacing the point nuclear charges with smeared ones (this is often done in the physics and computational literature. Formally these smeared charges are referred to as form-factors.) This replaces the Coulomb interaction potential $V(x,y)$ in (7) by a potential which is differentiable as many times as the form-factors are. Note that the Coulomb potential created by a point charge e at a point y can be written as:

$$V(x) = \frac{e}{|x-y|} = \int \frac{e}{|x-z|}\delta(y-z)dz$$

To remove the Coulomb singularity, the singular charge density $e\delta(y-z)$ is "smeared" by non-singular H^k one, $\rho(y-z)$ (say a Gaussian) to obtain

$$V'(x) = \int \frac{\rho(y-z)}{|x-z|}dz$$

for which we shall abuse notation and refer to the form-factor as $V(x,y)$ from now on. All the necessary conditions of self-adjointness, etc hold in this case.

In order to prove the next lemma we will need regularity results on the ground state $\psi_y(x)$ in both variables. This is one of the main reasons for replacing our potential by form-factors. In the variable x, the ground state is H^2 (this follows by a variational principle). To address the regularity in parameter y we shall appeal to the perturbation theory of Section 2 and the following notion of *k-differentiable in the resolvent sense*.

Consider a family of operators $\{H(y)\}$ indexed by a parameter y. We say the family $\{H(y)\}$ *is differentiable in the resolvent sense* if

(i) For every y, $H(y)$ is a closed operator with non-empty resolvent.

(ii) For all y_0, in each connected component R of the resolvent $\rho(H(y_0))$, there exists a $\lambda_0 \in R$ such that $\lambda_0 \in R$ for all y near y_0, and the resolvent $(H(y) - \lambda_0)^{-1}$ is a k-differentiable operator-valued function of y, near y_0.

This is an analogue of analytic in the sense of Kato! In fact we can make our form-factor smooth. We record our differentiability result here:

Theorem 11. *Let $\{H(y)\}$ be a family of operators k-times differentiable in the resolvent sense, indexed by $y \in \mathbb{R}^m$. Let E_0 be a non-degenerate discrete eigenvalue of $H(y_0)$. Then for y near y_0 there is exactly one point $E(y)$ of $\sigma(E(y))$ near E_0 and this point is isolated and non-degenerate. Furthermore $E(y)$ has an eigenvector ψ_y, and for y near y_0 both are k-times differentiable functions of y. If $H(y)$ is self-adjoint then ψ_y can be normalized. In particular, we will have*

$$\sup_y \|\partial_y^\alpha \psi_y\|_{H^k(x)} \leqslant C \tag{51}$$

for index $|\alpha| \leqslant k$.

Proof. The proof of this theorem follows from the Kato-Rellich theorem Theorem 6. The differentiability condition (51) follows from the equations

$$E(y) = \frac{\langle \psi_0, H(y)P(y)\psi_0 \rangle}{\langle \psi_0, P(y)\psi_0 \rangle}$$

$$\psi_y = \langle \psi_{y_0}, P(y)\psi_{y_0} \rangle^{-1/2} P(y)\psi_0$$

since for some y near y_0, $P(y)\psi_0 \neq 0$ since $P(y)\psi_0 \to P(y_0)\psi_0 = \psi_0$. This follows from an application of the 2^{nd} resolvent identity. To illustrate how the differentiability condition arises let us differentiate once. Let $n = \langle \psi_{y_0}, P(y)\psi_{y_0} \rangle$. Then,

$$\partial_y \psi_y = n^{-1/2} \partial_y P(y)\psi_0 - \frac{1}{2} n^{-3/2} (\partial_y n) P(y)\psi_0$$

where $\partial_y n = \langle \psi_{y_0}, \partial_y P(y)\psi_{y_0} \rangle$. It then follows that this is L^2_x-bounded using the fact that $\partial P(y)$ is. The bounds for higher derivatives follow by similar calculations. □

We note of course that $H_e(y)$ (with the form-factors replacing the Coulomb potential $V(x, y)$) is k-differentiable in the resolvent sense for $V'(x, y)$ k-differentiable. This follows from the 2^{nd} resolvent identity (6):

$$\frac{d}{dy}(H_e(y) - z)^{-1} = -(H_e(y) - z)^{-1} \frac{d}{dy} H_e(y) (H_e(y) - z)^{-1}$$

So we apply this theorem to obtain local differentiability around the minimizer of the ground state energy y_0.

We now seek to apply the Feshbach-Schur method using H_{mol} and the orthogonal projection (50). First we must verify that the necessary conditions hold.

Lemma 4. *The following conditions required for the Feshbach-Schur map (see **REFERENCE**) hold for the operator $H = H_{mol} - \lambda$ and the projection P defined in (50):*

I) We have Ran $P \subset \mathcal{D}(H)$ and $H_{\overline{P}} := \overline{P}H\overline{P}|_{Ran\ \overline{P}}$ is invertible.

II) The following conditions hold for the operator $R_{\overline{P}} = \overline{P}H_{\overline{P}}^{-1}\overline{P}$,

$$\|R_{\overline{P}}\| < \infty, \quad \|PHR_{\overline{P}}\| < \infty, \quad \|R_{\overline{P}}HP\| < \infty \tag{52}$$

Proof of Lemma 4. Let us first show that the operator $H_{\overline{P}} := \overline{P}(H-\lambda)\overline{P}|_{\mathrm{Ran}\ \overline{P}}$ is invertible. Note that P is an orthogonal projection and hence Ran P is a closed subspace and hence also a Hilbert space. Let $E'(y)$ be the first excited state energy of $H_e(y)$ above $E(y)$ (so that $E'(y) > E(y)\ \forall y$.) In particular this implies that $\min_y E'(y) > \min_y E(y) = \lambda_0$.
Now recall the decomposition of $H_{mol} = H_e + T_N$. Recalling the fiber integral operators and the fact that $[H_e(y), P(y)] = 0$ (as $P(y)$ is an eigenprojection of $H_e(y)$), we conclude that H_e commutes with the projection P and hence also with $\overline{P}$.

Being a sum of negative Laplacian operators, T_N is a positive operator itself, i.e. $T_N \geqslant 0$. Recall that when we bound operators $A \geqslant B$ we are bounding them in expectation, ie. $\langle (A-B)\phi, \phi\rangle \geqslant 0$. Hence,

$$\begin{aligned}
\overline{P}H_{mol}\overline{P} &\geqslant H_e\overline{P} \\
&\geqslant \int^{\oplus} H_e(y)\overline{P}(y)dy \\
&> \int^{\oplus} E'(y)\overline{P}(y)dy > \lambda_0\overline{P}
\end{aligned} \tag{53}$$

Thus we conclude that the operator $H_{\overline{P}}^0 := \overline{P}(H_{mol} - \lambda_0)\overline{P}$ is invertible as 0 is not part of the spectrum. Choosing λ close enough to λ_0, we can bound:

$$\begin{aligned}
\|H_{\overline{P}}^0 - H_{\overline{P}}\| &= \|\overline{P}(H_{mol} - \lambda_0)\overline{P} - \overline{P}(H_{mol} - \lambda)\overline{P}\| \\
&= \|\overline{P}(\lambda_0 - \lambda)\overline{P}\| \\
&= |\lambda_0 - \lambda| < \delta
\end{aligned}$$

And hence choosing λ such that $\delta < \dfrac{1}{\|H_{\overline{P}}^{0\,-1}\|}$, by a standard Neumann series argument we have that $H_{\overline{P}}$ is invertible.

Now let us show the bounds (52) in II). Recall $R_{\overline{P}} = \overline{P}H_{\overline{P}}^{-1}\overline{P}$. The first bound is trivial, since by (53) the operator $H_{\overline{P}}$ is invertible on Ran $\overline{P}$ and hence $R_{\overline{P}} = \overline{P}H_{\overline{P}}^{-1}\overline{P}$ is bounded.

We now show that $PH_{mol}R_{\overline{P}}$ is a bounded operator from $L^2 \to L^2$. In particular this implies that $R_{\overline{P}}H_{mol}P$ is bounded too since $R_{\overline{P}}H_{mol}P = (PH_{mol}R_{\overline{P}})^*$.

First observe that the H_e contribution vanishes, as $PH_eR_{\overline{P}} = E(y)PR_{\overline{P}}$ and $\overline{P}P = 0 = P\overline{P}$. Hence we are only looking at

$$PH_{mol}R_{\overline{P}}\Psi = PT_NR_{\overline{P}}\Psi \tag{54}$$

To simplify notation let us set $2m_j = 1$ for all j, so that $T_N = -\Delta_y$. For the moment assume that the term $R_{\overline{P}} : L^2 \to H^2$ is bounded (we shall show this last.) Writing $\Delta_y = \nabla_y^2$ we introduce the commutator, compute it, and eliminate $P\overline{P} = 0$ terms:

$$PT_NR_{\overline{P}}\Psi = 2(\nabla_y P)\nabla_y R_{\overline{P}} + (\Delta_y P)R_{\overline{P}}$$

We can compute:

$$(\nabla_y P)\Psi = \nabla_y\psi_y\langle\psi_y, \Psi\rangle_x + \psi_y\langle\nabla_y\psi_y, \Psi\rangle_x$$

Similarly,

$$\Delta_y P\Psi = \Delta_y\psi_y\langle\psi_y, \Psi\rangle_x + 2\nabla_y\psi_y\langle\nabla_y\psi_y, \Psi\rangle_x + \psi_y\langle\Delta_j\psi_y, \Psi\rangle_x$$

We can bound each term here using our regularity results for ψ_y, for example

$$\begin{aligned}
\|\nabla_y\psi_y\langle\psi_y, \Psi\rangle_x\|_{L^2} &= \int\int|\nabla_y\psi_y(x)|^2dx|\int\overline{\psi_y(x)}\Psi(x,y)dx|^2dy \\
&\leqslant \|\nabla_y\psi_y\|_{L^2(x)}^2\int(\int|\psi_y(x)|^2dx)^2(\int|\Psi(x,y)|^2dx)^2dy \\
&\leqslant \|\nabla_y\psi_y\|_{L^2(x)}^2\|\psi_y\|_{L^2(x)}^2\|\Psi\|_{L^2}^2
\end{aligned}$$

Hence the operators $(\nabla_y P)$ and $(\Delta_y P)$ are bounded.

Finally, we need to justify that $\nabla_y R_{\overline{P}} = \nabla_y\overline{P}(\overline{P}(H_{mol} - \lambda)\overline{P})^{-1}\overline{P}$ is bounded. We can write

$$\|\nabla_y R_{\overline{P}}\Psi\|^2 = \langle R_{\overline{P}}\Psi, -\Delta_y R_{\overline{P}}\Psi\rangle \leqslant \|R_{\overline{P}}\Psi\|\|\overline{P}\Delta_y\overline{P}R_{\overline{P}}\|$$

The fact that both terms are bounded will follow from the fact that $R_{\overline{P}}$ is bounded from $L^2 \to H^2$. This in turns follows in two steps. Notice that $R_{\overline{P}}$ acts on the closed subspace

Ran $\overline{P}$, so it must be shown that $R_{\overline{P}} : L^2 \to \overline{P}H^2$ and that $\overline{P}H^2 \subset H^2$. The latter follows from similar computations to the bounds for $\Delta_y P$ done above (ie. using equation (51)) and further using the regularity of $\psi_y(x)$ in x, ie. $\psi_y(x) \in H^2(x)$. We also note that the same computations will show Ran $P \subset \mathcal{D}(H_{mol})$ hence satisfying condition I).

Recalling (53) we can similarly obtain:

$$\langle \overline{P}(H_{mol} - \lambda)\overline{P}\phi, \phi\rangle > \langle \overline{P}(\lambda_0 - \lambda + T_N)\overline{P}\phi, \phi\rangle > 0 \tag{55}$$

Choosing λ such that $\lambda_0 - \lambda > 0$, notice that the operator $\overline{P}(\lambda_0 - \lambda + T_N)\overline{P}$ does not contain 0 in its spectrum and hence it is invertible (we can bound it below by $\lambda_0 - \lambda$). As $\overline{P}H^2 \subset H^2$ it is bounded operator from $H^2 \to L^2$. We can then bound the norms, as

$$\begin{aligned}\|\overline{P}(H_{mol} - \lambda)\overline{P}\|^2 &= \sup_{\|\phi\|=1} \langle \overline{P}(H_{mol} - \lambda)\overline{P}\phi, \phi\rangle \\ &> \sup_{\|\phi\|=1} \langle \overline{P}(\lambda_0 - \lambda + T_N)\overline{P}\phi, \phi\rangle = \|\overline{P}(\lambda_0 - \lambda + T_N)\overline{P}\|^2\end{aligned}$$

Hence $\|(\overline{P}(H_{mol} - \lambda)\overline{P})^{-1}\| < \|(\overline{P}(\lambda_0 - \lambda + T_N)\overline{P})^{-1}\|$ and thus $R_{\overline{P}}$ is a bounded operator. Hence all the terms in II) are bounded and the proof concludes. □

As the pre-requisite conditions are satisfied, we can apply the Isospectrality Theorem (9) to conclude that the operator $F_P(H_{mol})$ is isospectral to H_{mol} at 0. Before using this conclusion we will re-write the expression for $F_P(H_{mol} - \lambda)$. First,

$$\begin{aligned}R_{\overline{P}} &= \overline{P}[(\overline{P}H_{mol}\overline{P} - \lambda\overline{P})|_{\text{Ran }\overline{P}}]^{-1}\overline{P} \\ &= \overline{P}(\overline{P}H_{mol}\overline{P} - \lambda)^{-1}\overline{P}\end{aligned} \tag{56}$$

keeping in mind the domain of $\overline{P}H_{mol}\overline{P} - \lambda$ as being Ran $\overline{P}$. Thus,

$$F_P(H_{mol} - \lambda) = P(H_{mol} - \lambda - H_{mol}R_{\overline{P}}H_{mol})P|_{\text{Ran } P}$$

Observing that $PH_e\overline{P} = 0 = \overline{P}H_eP$ simplify the term:

$$H_{mol}R_{\overline{P}}H_{mol} = \kappa^2 W R_{\overline{P}} W$$

where recall $\kappa W = T_N$. Hence keeping in mind that F_P acts on Ran P and omitting this from the equation, we arrive at

$$F_P(H_{mol} - \lambda) := H_P - \kappa^2 U(\lambda) - \lambda \tag{57}$$

where $H_P := PH_{mol}P$ and $U(\lambda) = PWR_{\overline{P}}WP$, and $R_{\overline{P}}$ is as defined in (56).

Now let us appeal to the Isospectrality theorem. Noting that $\lambda \in \sigma(H)$ is equivalent to $0 \in \sigma(H - \lambda)$, we obtain:

$$\lambda \in \sigma_d(H_{mol}) \iff 0 \in \sigma_d(F_P(H_{mol} - \lambda)) \tag{58}$$

We explicitly evaluate the right-hand side of (57) in the following proposition.

Proposition 2. *The operator (57) can be formally computed as:*

$$F_P(H_{mol} - \lambda)\Psi = \psi_y(x)(E(y) + T_N + v + T(\lambda))f(y) - \lambda\Psi \tag{59}$$

acting on $\Psi \in H^2(\mathbb{R}^{3(N+M)})$*, and:*

$$\begin{aligned} v &:= -\sum_1^M \frac{1}{2m_j}(a_{jy_j} +_{y_j} a_j - \|_{y_j}\psi_y\|_x^2), \\ f(y) &:= \int \overline{\psi_y(x)}\Psi(x,y)dx, \end{aligned} \tag{60}$$

where $a_j := \psi_{y,y_j}\,\psi_{yx}$.

Let us remark on the contribution of the terms. $T(\lambda)$ is a second order differential operator that is not computed explicitly but arises from the term $U(\lambda)$ in (57). We note that formally the term $E(y)$ is of 0^{th} order in κ, the term $H_{eff} = T_N + v$ (where v is defined below) is of first order, and the term $T(\lambda)$ is formally of second order.

Proof of Proposition (2). First note that by defining $f(y)$ as in (60) we can then write $P\Psi = \psi_y f$ for $\Psi \in H^2(\mathbb{R}^{3(N+M)})$. Let us begin by computing the first term in (57), the term $PH_{mol}P$. First, compute

$$\begin{aligned} (H_{mol}P\Psi(x,y) &= (H_e + T_{\text{nucl}})\psi_y(x)f(y) \\ &= E(y)\psi_y(x)f(y) + T_N\psi_y(x)f(y) \\ &\quad - \psi_y(x)T_{\text{nucl}}f(y) + \psi_y(x)T_{\text{nucl}}f(y) \\ &= \psi_y(x)(E(y) + T_N)f(y) + [T_N, \psi_y(x)]f(y) \\ &= [\psi_y(x)H_e + W_x]f(y,t) \end{aligned} \tag{61}$$

where we define $W_x = [T_N, \psi_y(x)]$. Hence, the term $PH_{mol}P$ becomes:

$$PH_{mol}P\Psi = \psi_y(T_N + E(y) + v)f$$

where we have further defined $v := \langle \psi_y, W_x \rangle_x$. Now let us further evaluate this v term. We shall do so with the help of the following lemma.

For ψ_y complex, the operator v defined is of the form

$$v = -\sum_1^M \frac{1}{2m_j}(a_{jy_j} +_{y_j} a_j - \|_{y_j}\psi_y\|_x^2) \tag{62}$$

where $a_j := \psi_{y,y_j}\,\psi_{yx}$ and $\|\cdot\|_z$ stands for the L^2-norm in the z variable. If ψ_y is purely real then v is the operator of multiplication by the function

$$v(y) := \sum_1^M \frac{1}{2m_j}\int |_{y_j}\psi_y|^2 dx. \tag{63}$$

Proof of Lemma. Let us first compute the commutator in the operator family W_x. Computing against an arbitrary function f,

$$\begin{aligned} W_x(f) &= [T_{\text{nucl}}, \psi_y(x)](f) \\ &= -\sum_1^M \frac{1}{2m_j}(\Delta_{y_j}(\psi_y f) - \psi_y \Delta_{y_j}(f)) \\ &= -\sum_1^M \frac{1}{2m_j}(\psi_y \Delta_{y_j}(f) + 2_{y_j}\psi_{yy_j}(f) + \Delta_{y_j}\psi_y(f) - \Delta_{y_j}\psi_y(f)) \\ &= -\sum_1^M \frac{1}{2m_j}(\Delta_{y_j}\psi_y + 2(_{y_j}\psi_y)_{y_j})(f) \end{aligned}$$

Plugging this into the definition of v gives

$$\begin{aligned} v &= -\int \overline{\psi_y(x)}[\sum_1^M \frac{1}{2m_j}(\Delta_{y_j}\psi_y + 2(_{y_j}\psi_y)_{y_j})]dx \\ &= -\sum_1^M \frac{1}{2m_j}(\psi_{y,y_j}\,\psi_{yy_j} + \psi_{y,y_j}\,(_{y_j}\psi_y)). \end{aligned} \tag{64}$$

Since $\psi_{y,y_j}\,(_{y_j}\psi_y) =_{y_j} \psi_{y,y_j}\,\psi_y -_{y_j} \psi_{y,y_j}\,\psi_y$, (which for ψ real reduces to the vector calculus identity $\psi\Delta\psi = (\overline{\psi}\psi) - |\psi|^2$), and defining $a_j := \psi_{y,y_j}\,\psi_y$, we obtain (62).

We see that for ψ_y real,

$$a_j = \psi_{y,y_j}\,\psi_{yx} = \int dx\psi_{yy_j}\psi_y = \int dx(\psi_{yy_j}\psi_y) = \frac{1}{2}{}_{y_j}\int dx\psi_y^2$$

and therefore $a_j = 0$, which implies (63). □

Hence we obtain the term $E(y) + T_N + v$ on the right-hand side of (57), since the non-degeneracy assumption of the ground state ψ_y implies that ψ_y is purely real. This concludes the computation for the $PH_{mol}P$ term.

Now let us turn our attention to the $U(\lambda)$ term and briefly comment on it. Let us examine it in more detail. Recalling that $\kappa W = T_N$:

$$\kappa^2 U(\lambda) = PT_N\overline{P}(\overline{P}H_{mol}\overline{P} - \lambda)^{-1}\overline{P}T_N P$$

We can hence compute

$$\overline{P}T_N P\Psi = \overline{P}(H_{mol} - \lambda)P\Psi$$

and using equation (61) we can write

$$\overline{P}T_N P\Psi = Kf(y)$$

where $K := \overline{P}W_x = W_x - v$. Hence we do not compute it explicitly but remark that then $\kappa^2 U(\lambda)$ is formally of order $O(\kappa^2)$ and is of the form:

$$U(\lambda)\Psi = \psi_y T(\lambda) f(y)$$

□

The above is a formal computation since indeed the domain of $F_P(H_{mol} - \lambda)$ is defined to be Ran P. The latter however can easily be identified with $L^2(\mathbb{R}^{3M})$. Let us slightly recast the isospectrality problem by defining

$$H_e(\lambda) = F_P(H_{mol} - \lambda) + \lambda P \tag{65}$$

so that it becomes:

$$\lambda \in \sigma_d(H_{mol}) \iff \lambda \in \sigma_d(H_e(\lambda)) \tag{66}$$

Let δ be the gap between the ground state energy $E(y)$ and the rest of the spectrum of $H_e(y)$. Then for $\lambda \leqslant E(y) + \delta/2$ the operator $H_e(\lambda)$ acts on $L^2(\mathbb{R}^{3M})$ in the following form:

$$H_e(\lambda) := T_N + E_\kappa(y, \lambda) \tag{67}$$

where $E_\kappa(y, \lambda)$ can be computed to higher orders of κ. It is simple to see how the above arises from (59) as $f(y) \in L^2(\mathbb{R}^{3M})$. We have

$$E_\kappa(y, \lambda) = E(y) + v + T(\lambda)$$

Hence up to second order the eigenvalue problem for $H_e(\lambda)$ recovers the effective Schrödinger equations obtained in the heuristic derivation for the BO approximation. We can now prove our main Theorem 10.

Proof of Theorem 10. Everything preceeding equation (48) follows directly from Proposition 2 and equations (66) and (67). To obtain equation (48) recall the definition of v in equation (62). We can then write:

$$v = -\sum_{1}^{M} \frac{1}{2m_j}(a\nabla_{y_j} + \nabla_{y_j} a_j + a_j^2 + \tilde{v}) \tag{68}$$

where

$$\tilde{v} = -\sum_{1}^{M} \frac{1}{2m_j}[\int |\nabla_{y_j}\psi_y|^2 dx + a_j^2] \tag{69}$$

where recall $a_j := \langle \psi_y, \nabla_{y_j}\psi_y \rangle$. Hence introducing (68) into $H_{eff} = H_e + v$ we clearly obtain (48) with $\tilde{v}$ given by (69). Now let us re-arrange the latter into (49). First we insert the partition of unity $P + \overline{P} = 1$ into the term:

$$\int |\nabla_{y_j}\psi_y|^2 dx = \langle \nabla_{y_j}\psi_y, (P + \overline{P})\nabla_{y_j}\psi_y \rangle$$

By the definition of P:

$$\langle \nabla_{y_j}\psi_y, P\nabla_{y_j}\psi_y \rangle = \langle \nabla_{y_j}\psi_y, \psi_y \rangle \langle \psi_y, \nabla_{y_j}\psi_y \rangle$$

Note that $\langle \nabla_{y_j}\psi_y, \psi_y \rangle = \nabla_{y_j}\langle \psi_y, \psi_y \rangle - \langle \psi_y, \nabla_{y_j}\psi_y \rangle$ and by the normalization condition, $\nabla_{y_j}\langle \psi_y, \psi_y \rangle = 0$. Hence

$$\langle \nabla_{y_j}\psi_y, P\nabla_{y_j}\psi_y \rangle = -a_j^2 \tag{70}$$

Using that $\overline{P}\psi_y = 0$ we conclude $\overline{P}\nabla_{y_j}\psi_y = [P, \nabla_{y_j}]\psi_y$ we can write

$$\langle \nabla_{y_j}\psi_y, \overline{P}\nabla_{y_j}\psi_y \rangle = \|[P, \nabla_{y_j}]\psi_y\|^2 \tag{71}$$

Combining (70) and (71) we obtain

$$\int |\nabla_{y_j}\psi_y|^2 dx = -a_j^2 + \|[P, \nabla_{y_j}]\psi_y\|^2$$

which yields (49) after being substituted in (69). □

4 Brief Summary of Talk

1. Motivation: Understanding the spectrum of Schrödinger operators is key to understanding the dynamics of the corresponding quantum system. *Perturbation theory* is a framework is studying the spectrum of an unknown operator of the form $H(\kappa) = H_0 + \kappa V$, as a function of a the *coupling constant* κ.

2. The discrete spectrum (eigenvalues) under *regular perturbations* (for families of operators $H(\kappa)$ analytic in the Kato sense) is charaterized by the Kato-Rellich theorem. The energy levels $E(\kappa)$ (eigenvalues) will be analytic functions in κ. The most common energy level of study is the ground state (the lowest eigenvalue).

3. It is sufficient in the case of the Schrödinger operator $H(\kappa)$ for V to be H_0-bounded for the family $H(\kappa)$ to be analytic in the Kato sense. However, more general problems may feature completely unbounded perturbations V.

4. The Feshbach-Schur method is a powerful technique that allows us to estimate the eigenvalues of $H(\kappa)$ using the isospectrality property of the Feshbach map, which projects the spectral problem of a linear operator on a larger (infinite-dimensional) space to a spectral problem for a (generally) nonlinear operator on a smaller (finite-dimensional) space. The latter may be easier to solve!

5. The Born-Oppenheimer framework suggests a way for the Feshbach-Schur method to be applied to the time-independent problem of molecular dynamics. We illustrate an application of this after making some simplifying assumptions.

References

[1] Reed. M., Simon. B .: *Analysis of Operators.* Methods of Modern Mathematical Physics. Vol 4 Academic Press (1978)

[2] Reed. M., Simon. B .: *Fourier Analysis. Self- Adjointness.* Methods of Modern Mathematical Physics. Vol 2 Academic Press (1978)

[3] Hislop, P.D., and Sigal, I.M.: *Introduction to spectral theory.* Applied Mathematical Sciences 133, NY: Springer, 1996

[4] Gustafson, S.J., and Sigal, I.M.: *Mathematical Concepts of Quantum Mechanics* 3rd ed., Springer Universitext, 2020

[5] Lubich, C.: *From Quantum to Classical Molecular Dynamics: Reduced Models and Numerical Analysis*. Zurich Lectures in Advanced Mathematics. European Mathematical Society, 2008

[6] Bach, V., Fröhlich, J., and Sigal, I.M., *Quantum electrodynamics of confined nonrelativistic particles.* Adv. Math. 137 (1998) 205-298, and 137 (1998) 299-395.

[7] Tuefel, S., *Adiabatic Perturbation Theory in Quantum Dynamics.* Springer-Verlag Berlin Heidelberg 2003

[8] M. Born and R. Oppenheimer. Zur Quantentheorie der Molekeln, Ann. Phys. (Leipzig) 84, 457–484 (1927).

[9] Combes, J.-M., Duclos P., and Seiler, R.: *The Born-Oppenheimer approximation, in: Rigorous Atomic and Molecular Physics* (eds. G. Velo, A. Wightman), New York, Plenum, 185–212 (1981).

[10] Klein M., Martinez A., Seiler R., and Wang X.P., *On the Born-Oppenheimer expansion for polyatomic molecules,* Commun. Math. Phys. 143, 607–639 (1992)

[11] Szabo A. and Ostlund N.S., *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Revised ed., Dover Publ., New York, 1996.

[12] Le Bris C. (ed.), *Handbook of Numerical Analysis,* Special Volume Computational Chemistry, Elsevier, Amsterdam, 2003.

[13] Kato T., *Fundamental Properties of Hamiltonian Operators of Schrodinger Type,* Transactions of the American Mathematical Society, 1951, Pg. 195-211 http://www.jstor.org/stable/1990366