Adaptive Perturbation Theory I: Quantum Mechanics[*]

Marvin Weinstein

Stanford Linear Accelerator Center, Stanford University, Stanford, California 94309

(Dated: September 29, 2005)

Adaptive perturbation is a new method for perturbatively computing the eigenvalues and eigenstates of quantum mechanical Hamiltonians that heretofore were not believed to be treatable by such methods. The novel feature of adaptive perturbation theory is that it decomposes a given Hamiltonian, H, into an unperturbed part and a perturbation in a way which extracts the leading non-perturbative behavior of the problem exactly. This paper introduces the method in the context of the pure anharmonic oscillator and then goes on to apply it to the case of tunneling between both symmetric and asymmetric minima. It concludes with an introduction to the extension of these methods to the discussion of a quantum field theory. A more complete discussion of this issue will be given in the second paper in this series, and it will show how to use the method of adaptive perturbation theory to non-perturbatively extract the structure of mass, wavefunction and coupling constant renormalization.

PACS numbers: 03.65-w,03.70+k,11.10

I. INTRODUCTION

Perturbation theory, one of the most useful tools in the physicist's toolbox, is the method most commonly used to analyze problems in quantum mechanics and field theory. It is so useful that it is applied even when the perturbation clearly isn't small enough for the expansion to converge. One then typically tries to extract meaningful results by applying some summation technique to the resulting series, often with mixed success. What is really needed, is a method that extracts the leading non-perturbative behavior of the quantity to be computed, and then corrects this leading approximation in a systematic manner. This paper introduces a simple technique that works in just this way. I call this method *adaptive perturbation theory*.

I will limit the bulk of this paper to discussing how adaptive perturbation theory works on problems in quantum mechanics. In the second paper in this series I will show how to apply an appropriately generalized version of adaptive perturbation theory to a quantum field theory. In particular, I will then show how to extract, in a non-perturbative context, the general structure of mass, wavefunction and coupling constant renormalization.

II. EXAMPLE 1: THE PURE ANHARMONIC OSCILLATOR

To introduce the ideas behind *adaptive perturbation theory*, I begin by studying the pure anharmonic oscillator, which seems naively hard to treat perturbatively because it is apparently devoid of a small expansion parameter. The pure anharmonic oscillator is defined by a Hamiltonian of the form

$$H = \frac{1}{2}\,p^2 + \frac{1}{6}\,\lambda\,x^4, \tag{1}$$

where the operators x and p satisfy the usual canonical commutation relation $[x,p] = i$.

Since H contains only the terms p^2 and $\lambda\,x^4$, there does not seem to be a way to perturbatively compute the eigenstates and eigenvalues of H based upon a division of H into an unperturbed Hamiltonian H_0, whose eigenstates are known, and a perturbation V. Nevertheless, I will now show that this can be done. The trick is to define a decomposition of the Hamiltonian in Eq. 1 into an unperturbed (i.e., solvable) part and a perturbation, in a way that doesn't depend upon λ. The choice of decomposition is what is *adapted* to the problem at hand. Thus, the decomposition, and therefore the perturbation theory for each eigenstate, will be different.

To begin the process of decomposing H, I introduce a one-parameter family of annihilation and creation operators, $A_\gamma^\dagger$ and A_γ, as follows:

$$x = \frac{1}{\sqrt{2\gamma}} \left(A_\gamma^\dagger + A_\gamma \right),$$
$$p = i \sqrt{\frac{\gamma}{2}} \left(A_\gamma^\dagger - A_\gamma \right). \tag{2}$$

Note that given the commutation relations of x and p, the commutation relation $[A_\gamma, A_\gamma^\dagger] = 1$ holds independent of γ. In terms of these operators H can be rewritten as

$$
\begin{aligned}
H &= \frac{\gamma}{4} \left(-A_\gamma^{\dagger\,2} - A_\gamma^{\,2} + 2 A_\gamma^\dagger A_\gamma + 1 \right) \\
&\quad + \frac{\lambda}{4\gamma^2} \left(A_\gamma^{\dagger\,2} A_\gamma^{\,2} + 2 A_\gamma^\dagger A_\gamma + 1 + \frac{1}{6}(A_\gamma^{\dagger\,4} + A_\gamma^{\,4}) + \frac{2}{3}(A_\gamma^{\dagger\,3} A_\gamma + A_\gamma^\dagger A_\gamma^{\,3}) + (A_\gamma^{\dagger\,2} + A_\gamma^{\,2}) \right) \\
&= \left(\frac{\gamma}{4} + \frac{\lambda}{4\gamma^2} \right) (2 N_\gamma + 1) + \frac{\lambda}{4\gamma^2} N_\gamma (N_\gamma - 1) + \left(\frac{\gamma}{4} - \frac{\lambda}{2\gamma^2} \right) (A_\gamma^{\dagger\,2} + A_\gamma^{\,2}) \\
&\quad + \frac{\lambda}{4\gamma^2} \left(\frac{1}{6}(A_\gamma^{\dagger\,4} + A_\gamma^{\,4}) + \frac{2}{3}(A_\gamma^{\dagger\,3} A_\gamma + A_\gamma^\dagger A_\gamma^{\,3}) \right) \tag{3}
\end{aligned}
$$

where I have defined the γ-dependent number operator to be

$$N_\gamma = A_\gamma^\dagger A_\gamma. \tag{4}$$

At this point I define a γ-dependent decomposition of H into an unperturbed part and a perturbation as follows:

$$
\begin{aligned}
H &= H_0(\gamma) + V(\gamma), \\
H_0(\gamma) &= \left(\frac{\gamma}{4} + \frac{\lambda}{4\gamma^2} \right) (2 N_\gamma + 1) + \frac{\lambda}{4\gamma^2} N_\gamma (N_\gamma - 1), \\
V(\gamma) &= \left(\frac{\gamma}{4} - \frac{\lambda}{2\gamma^2} \right) (A_\gamma^{\dagger\,2} + A_\gamma^{\,2}) \\
&\quad + \frac{\lambda}{4\gamma^2} \left(\frac{1}{6}(A_\gamma^{\dagger\,4} + A_\gamma^{\,4}) + \frac{2}{3}(A_\gamma^{\dagger\,3} A_\gamma + A_\gamma^\dagger A_\gamma^{\,3}) \right). \tag{5}
\end{aligned}
$$

In other words, the unperturbed Hamiltonian is defined to be that part of H which is diagonal in the number operator N_γ and the perturbation is defined to be everything else.

If we think of the Hamiltonian as a large matrix in the basis of Fock-space states generated by applying $A_\gamma^\dagger$ to the state $|0_\gamma\rangle$, then $H_0(\gamma)$ is the diagonal part of the matrix and the perturbation is the matrix obtained by setting all terms on the diagonal to zero.

Although it isn't conventional to include the term $\frac{\lambda}{4\gamma^2} N_\gamma (N_\gamma - 1)$ in the definition of the unperturbed Hamiltonian, $H_0(\gamma)$, it makes sense to do so for two reasons. First, this part of the Hamiltonian is diagonal in N_γ and thus it is just as easy to deal with as the part proportional to the first power of N_γ. Second, including it in the unperturbed Hamiltonian puts a term proportional to $N_\gamma^{\,2}$ in the energy denominators which appear in the perturbation expansion based upon this decomposition. This suppresses the $n!$ growth of the usual perturbation expansion. The only apparent drawback to putting a term proportional to $N_\gamma^{\,2}$ into $H_0(\gamma)$ is that it makes the energies of the unperturbed levels grow like n^2, whereas the correct answer is $n^{4/3}$. While this is not a disaster, it seems to say that getting the correct answer from a straightforward perturbative approach will be difficult. This observation leads us to add the *adaptive* element to *adaptive perturbation theory*.

The essential point is that I have not as yet committed to a particular value of the parameter γ in Eqs. 5-5. The trick consists of adapting the choice of γ to the computation at hand. Thus, to compute the energy of the Nth eigenstate I will choose a value, $\gamma(N)$, which varies as a function of N. The appropriate value of $\gamma(N)$ is fixed by a simple variational calculation.

The variational calculation, which serves as the starting point for any adaptive perturbation theory computation, begins by defining a γ-dependent family of Fock-states, $|N_\gamma\rangle$. First we define the γ-dependent vacuum state, $|0_\gamma\rangle$, by the condition

$$A_\gamma |0_\gamma\rangle = 0. \tag{6}$$

Next, we define the n-particle state built on this vacuum state such that

$$N_\gamma|n_\gamma\rangle = n|n_\gamma\rangle, \tag{7}$$

by

$$|n_\gamma\rangle = \frac{1}{\sqrt{n!}} A_\gamma^{\dagger\,n}|0_\gamma\rangle. \tag{8}$$

In terms of functions of x, the state $|0_\gamma\rangle$ is just a gaussian of whose width is determined by γ. Similarly, the state $|n_\gamma\rangle$ is the appropriate n^{th} order Hermite polynomial obtained by applying the γ-dependent differential operator

$$A_\gamma^\dagger = \sqrt{\frac{\gamma}{2}}\,x - \frac{1}{\sqrt{2\gamma}}\frac{d}{dx} \tag{9}$$

to this gaussian. (It is worth noting that the specific functional form of the state $|n_\gamma\rangle$ depends upon γ, but the number of times this function takes the value zero does not.)

The value of γ used to define the adaptive perturbation theory for the nth level of the anharmonic oscillator is determined by requiring that it minimize the expectation value

$$E_n(\gamma) = \langle n_\gamma|H|n_\gamma\rangle. \tag{10}$$

Eqs. 5-5 show that this expectation value is equal to

$$E_n(\gamma) = \langle N_\gamma|H_0(\gamma)|n\rangle = \left(\frac{\gamma}{4} + \frac{\lambda}{4\gamma^2}\right)(2\,n+1) + \frac{\lambda}{4\gamma^2}n\,(n-1). \tag{11}$$

Minimizing $E_n(\gamma)$ with respect to γ leads to the equation

$$\gamma = \lambda^{1/3}\left(\frac{2\,(n^2+n+1)}{2n+1}\right)^{1/3}. \tag{12}$$

At this point all we have to do is substitute this value into Eq. 10, to see why we chose γ to minimize $E_n(\gamma)$. If we do this, we obtain

$$E_n(\gamma)_{\min} = \frac{3}{8}\lambda^{1/3}\,(2n+1)^{2/3}\left(2n^2+2(n+1)\right)^{1/3} \tag{13}$$

which, for large n, behaves as $\lambda^{1/3}\,n^{4/3}$, which is the correct answer.

The fact that all energies scale as $\lambda^{1/3}$ is an easily obtained exact result and so, the non-trivial part of variational computation is the derivation of the dependence of the energy on n. To see why all energies are proportional to $\lambda^{1/3}$ it suffices to make the following canonical transformation

$$x \to \frac{x}{\lambda^{1/6}} \quad ; \quad p \to \lambda^{1/6}p. \tag{14}$$

In terms of these operators, the Hamiltonian of the pure anharmonic oscillator becomes

$$H = \lambda^{1/3}\left(\frac{1}{2}p^2 + \frac{1}{6}x^4\right), \tag{15}$$

thus proving the claim. A comparison of the variational computation and the result of a second-order perturbation theory for $\lambda = 1$ and widely differing values of N is given in Table I. As advertised, we see that the adaptive perturbation theory for each level converges rapidly, independent of λ and N.

A. Why does the expansion converge?

At this point I should address the question of why adaptive perturbation theory converges so rapidly.

First, observe that once I have done the variational calculation the perturbation theory computation is done for states defined in terms of a single fixed value of γ. Thus, all states which appear in the perturbative expansion are orthogonal to one another and the perturbation series is well-defined. Second, for the values of γ associated

with a given state $|n_0\rangle$, the unperturbed energies for states having another occupation number, n, all have a term $n^2/\gamma^2 \approx n^2/n_0^{2/3}$. It is only for values of n that are near n_0 that the energy denominators behave as $n^{4/3}$. For values of $n >> n_0$ the factor γ does not suppress the n^2 behavior. This is what suppresses the $n!$-growth of the perturbation expansion, since now the factor of $\sqrt{(n+1)(n+2)(n+3)(n+4)}$, which comes from applying $A_\gamma^{\dagger\,4}$ to a state $|N_\gamma\rangle$, is canceled by the $1/n^2$ appearing in the energy denominator, instead of the usual $1/n$. I'll make no attempt to prove this assertion, instead I refer the reader to a pair of very nice papers by I. G. Halliday and P. Suranyi[1, 2] where, by a very different technique which shares this same feature, they give a proof that the fact that the unperturbed energies are quadratic in N leads to a convergent perturbation expansion.

Finally, I should comment that it isn't a problem that the variationally determined states $|N_\gamma\rangle$, for different occupation numbers n, are not orthogonal to one another. First, there is a simple formula that expresses the state $|0_\gamma\rangle$ for one choice of γ as a sum over the states of definite occupation number for a different value of γ (see Appendix A). Thus, if one wishes to do a non-perturbative diagonalization of the full Hamiltonian truncated to many different variationally determined states, one can either orthogonalize these states by hand, or use the formula for the overlap of these states and solve the so-called *relative eigenvalue problem*

$$H|\psi\rangle = \epsilon\, M|\psi\rangle. \tag{16}$$

Note, that if $|n_{\gamma(n)}\rangle$ stands for one of the variationally constructed states, then the Hermitian matrix M is given by

$$M_{n,m} = \langle M_{\gamma(n)}|M_{\gamma(m)}\rangle. \tag{17}$$

The second point is that since the perturbation expansion for any state $|n_{\gamma(n)}\rangle$ converges, eventually the perturbatively corrected states $|n_{\gamma(n)}\rangle$ become orthogonal to one another. This follows from the fact that eigenstates of a Hermitian Hamiltonian having distinct eigenvalues are orthogonal to one another.

B. What does this have to do with quasi-particles?

An interesting corollary to the adaptive perturbation theory technique is that it provides an explicit realization of the quasi-particle picture underlying much of many-body theory. What we have shown is that no matter how large the underlying coupling, the physics of the states near a given n_0-particle state can be accurately described in terms of perturbatively coupled eigenstates of an appropriately chosen harmonic oscillator. Of course, as we have seen, the appropriately chosen harmonic oscillator picture changes as n_0 changes. Furthermore, the perturbatively coupled states $|n_{\gamma(n_0)}\rangle$, for $n \approx n_0$, correspond to states containing an infinite number of particles, if we choose as a basis those states that correspond to a significantly different value of n_0.

III. EXAMPLE 2: ADDING A MASS TERM

An obvious modification of the pure anharmonic oscillator is to add a positive mass term; i.e., let H be

$$H = \frac{1}{2}p^2 + \frac{m^2}{2}x^2 + \frac{1}{6}\lambda\, x^4. \tag{18}$$

This problem is dealt with in exactly the same way. First, introduce γ-dependent annihilation and creation operators, to arrive at a slightly modified formula for the expectation value in state $|N_\gamma\rangle$; *i.e.*,

$$E_n(\gamma) = \left(\frac{\gamma}{4} + \frac{m^2}{4\gamma} + \frac{\lambda}{4\gamma^2}\right)(2\,n+1) + \frac{\lambda}{4\gamma^2}n\,(n-1). \tag{19}$$

TABLE I: A comparison of the zeroth order and second order perturbation results, for the energy of the Nth level of the pure anharmonic oscillator, to the exact answer for $\lambda = 1$ and widely varying values of N.

λ	N	Variational	2nd Order Perturbation	Exact	Variational % Err	Perturbative % Error
1.0	0	0.375	0.3712	0.3676	0.02	0.0098
1.0	1	1.334	1.3195	1.3173	0.01	0.0017
1.0	10	17.257	17.508	17.4228	-0.009	0.0049
1.0	40	104.317	105.888	105.360	-0.009	0.0050

Once again, minimizing $E_n(\gamma)$ with respect to γ leads to the equation,

$$\gamma^3 - m^2\gamma - 2\lambda\left(\frac{n^2 + n + 1}{2n + 1}\right) = 0. \tag{20}$$

Clearly, for small λ and small n, the solution to this equation is $\gamma \approx m$. In other words, γ is what it would be for the usual perturbative expansion. Note, however, that no matter how small λ is, the term $\lambda n(n-1)$ eventually dominates the energy denominator and shuts off the $N!$-growth of the perturbation expansion. Also, for large enough N, it is clear that Eq. 19 is eventually dominated by the term proportional to λ, and for those states the energy, once again, goes like $\lambda^{1/3} n^{4/3}$.

As in the previous case, numerical studies show that independent of λ, m and n by second order in the perturbation theory determined by the variational calculation, the eigenvalues of H are computed to an accuracy of better than one percent.

IV. EXAMPLE 3: THE DOUBLE WELL

It should be clear from the previous discussion that adaptive perturbation theory can be applied to any Hamiltonian of the form

$$H = \frac{p^2}{2} + \frac{m^2}{2}x^2 + \lambda\, x^p. \tag{21}$$

The problem becomes more interesting when the mass term is chosen to be negative and the potential energy develops two distinct minima. (This is interesting in the case of a field theory because it is associated with the sponataneous breaking of the discrete symmetry $\phi \to -\phi$.) I will now show how to apply adaptive perturbation theory to this problem, since obtaining accurate answers requires generalizing the procedure in a way that will be very important in the next paper in this series.

Up to an irrelevant constant, the most general negative mass version of the anharmonic oscillator can be written as:

$$H = \frac{1}{2}p^2 + \frac{1}{6}\lambda\left(x^2 - f^2\right)^2. \tag{22}$$

Clearly, for a non-vanishing value of f^2, this potential has two minima located at $x = \pm f$. Thus, we would expect that the best gaussian approximation to the ground state of this system can't be a gaussian centered at the origin; classical intuition would imply that it is a gaussian centered about another point, $x = c$. In other words, if $|0_\gamma\rangle$ is a state centered at $x = 0$, it is better to adopt a trial state of the form

$$|c, \gamma\rangle = e^{-icp}|0_\gamma\rangle. \tag{23}$$

Since

$$e^{icp}\, x\, e^{-icp} = x + c, \tag{24}$$

it follows that computing the expectation value of the Hamiltonian, Eq. 22, in the state specified in Eq. 23, is the same as computing the expectation value of the Hamiltonian obtained by replacing the operator x by $x + c$, in the state $|0_\gamma\rangle$. In the language of the previous sections, this is equivalent to introducing the annihilation and creation operators $A_\gamma^\dagger$ and A_γ as follows:

$$\begin{aligned} x &= \frac{1}{\sqrt{2\gamma}}\left(A_\gamma^\dagger + A_\gamma\right) + c, \\ p &= i\sqrt{\frac{\gamma}{2}}\left(A_\gamma^\dagger - A_\gamma\right). \end{aligned} \tag{25}$$

With this definition, the normal ordered form of the Hamiltonian, Eq. 22, becomes

$$\begin{aligned} H = H(\gamma, c) &= \left(\frac{\gamma}{4} + \frac{\lambda f^4}{6} - \frac{\lambda f^2 c^2}{3} + \frac{\lambda c^4}{6} + \frac{\lambda c^2}{2\gamma} + \frac{\lambda}{8\gamma^2} - \frac{\lambda f^2}{6\gamma}\right) + \frac{\lambda c}{3}\sqrt{\frac{2}{\gamma}}\left(c^2 - f^2 + \frac{3}{2\gamma}\right)\left(A_\gamma^\dagger + A_\gamma\right) \\ &+ \left(\frac{\lambda c^2}{\gamma} + \frac{\lambda}{2\gamma^2} + \frac{\gamma}{2} - \frac{\lambda f^2}{3\gamma}\right)N_\gamma + \frac{\lambda}{4\gamma^2}N_\gamma(N_\gamma - 1) + \left(-\frac{\lambda f^2}{6\gamma} + \frac{\lambda}{4\gamma^2} + \frac{\lambda c^2}{2\gamma} - \frac{\gamma}{4}\right)\left(A_\gamma^{\dagger\,2} + A_\gamma^{\,2}\right) \\ &+ \frac{\lambda c}{3\sqrt{2}\gamma^{3/2}}\left(A_\gamma^{\dagger\,3} + A_\gamma^{\,3}\right) + \frac{\lambda}{6\gamma^2}\left(A_\gamma^{\dagger\,3}A_\gamma + A_\gamma^\dagger A_\gamma^{\,3}\right) + \frac{\lambda}{24\gamma^2}\left(A_\gamma^{\dagger\,4} + A_\gamma^{\,4}\right). \end{aligned} \tag{26}$$

The expectation value of this Hamiltonian in the state $|0_\gamma\rangle$ is

$$\mathcal{E}(c,\gamma) = \left(\frac{\gamma}{4} + \frac{\lambda f^4}{6} - \frac{\lambda f^2 c^2}{3} + \frac{\lambda c^4}{6} + \frac{\lambda c^2}{2\gamma} + \frac{\lambda}{8\gamma^2} - \frac{\lambda f^2}{6\gamma}\right), \tag{27}$$

which should be minimized with respect to both γ and c in order to define the starting point of the adaptive perturbation theory computation.

Fig. 1 shows $\mathcal{E}(\gamma,c)$ for a particular choice of λ and f^2 and it exhibits what is a generic feature of any such plot, namely that there are three minima. To get a better feeling for what the three dimensional plot is showing us, it is convenient to hold c fixed and solve for the value of γ that minimizes $\mathcal{E}(\gamma,c)$; call this $\gamma(c)$, and then plot $\mathcal{E}(\gamma(c),c)$ for various values of λ and f^2. Three such plots are shown in Fig. 2. The first plot is for a value of f which is large enough so that the lowest energy is obtained for a gaussian shifted either to the right or left by the amount $c = \pm c_{\min}$. While there is a local minimum at $c = 0$ it has a higher energy than the shifted states. Things change as one lowers the value of f. Thus, the second plot shows that for lower f the shifted wavefunctions and the one centered at zero are essentially degenerate in energy. For a slightly smaller value of f things reverse and the unshifted wavefunction has a lower energy than the shifted ones.
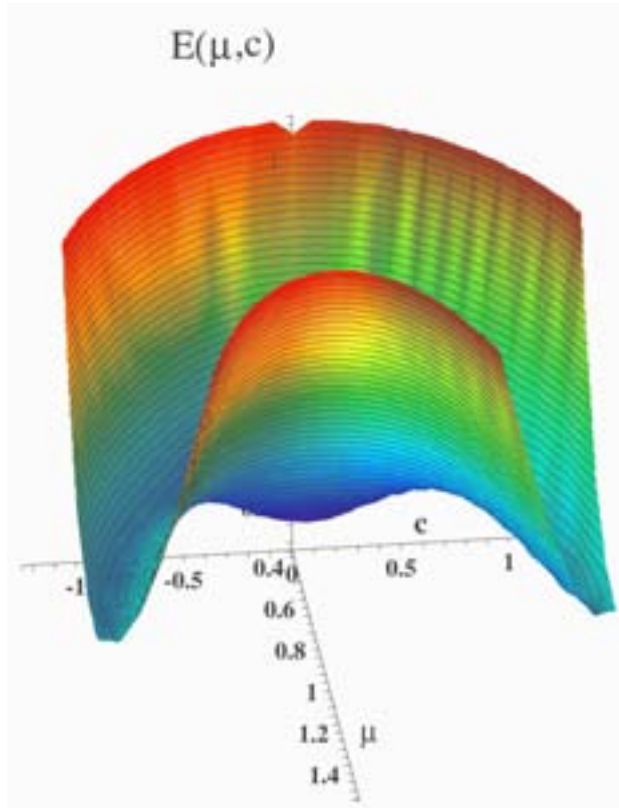

FIG. 1: Effective potential as a function of γ and c showing three local minima.

While, there would seem to be nothing wrong with the situation shown in Fig. 2, it is problematic when one applies the same sort of analysis to negative mass ϕ^4 field theory in $1+1$-dimensions. In this case, if one were to add a term like $J\phi$ to the Hamiltonian, this result would imply the existence of a first order phase transition when J reached some finite value. At this point the expectation value of ϕ in the ground-state would jump discontinuously from zero to a non-zero value. It has been rigorously shown that such a first order phase transition at a non-vanishing value of J cannot occur[4].

For this reason it is important to understand the origin of the secondary minimum at $c = 0$ and how to avoid it ever becoming lower in energy than the flanking minima at $c = \pm c_{\min}$. Fortunately, the origin of the problem is easy to understand and fixing it is not difficult. The key is that the Hamiltonian in Eq. 26 contains a term linear in c and linear in $A_\gamma^\dagger + A_\gamma$. Taking the effect of this term into account removes, or at least minimizes the importance of, the secondary minimum at $c = 0$. In order to allow this term to play a role in the variational calculation we have to adopt
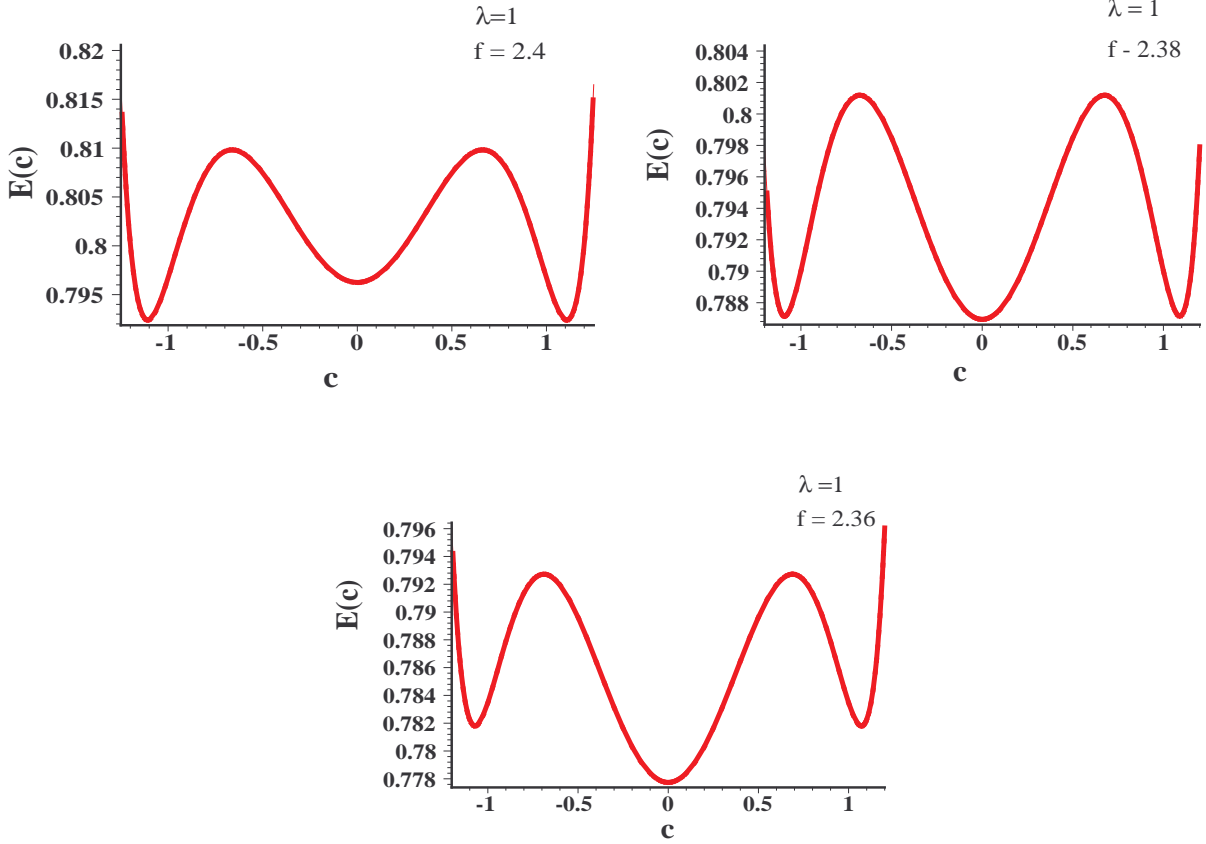
FIG. 2: Effective potential as a function of c, showing two degenerate local minima at $c = \pm c_{\min}$ and one minimum at $c = 0$. The three plots are for $\lambda = 1$ and different values of f and show how the minimum at $c = 0$ changes from lying higher than the minima at $c = \pm c_{\min}$ to being the global minimum. If this were field theory this would be characteristic of a first order phase transition.

a slightly more general form of the variational wavefunction; namely, we try a state of the form

$$\psi(\gamma, c, \alpha) = e^{icp} \left(\cos(\alpha) \, |0_\gamma\rangle + \sin(\alpha) \, |1_\gamma\rangle \right). \tag{28}$$

The easiest way to minimize the expectation value of the Hamiltonian, Eq. 26, in such a state is to fix the values of γ and c and solve for the value of α which minimizes this expression. Then, substitute this value, $\alpha(\gamma, c)$, and minimize the resulting equation respect to γ and c. Minimizing the expectation value of the Hamiltonian when γ and c are held fixed is equivalent to diagonalizing the 2×2-matrix

$$H_2(0, \gamma) = \begin{pmatrix} \langle 0_\gamma | H(\gamma, c) | 0_\gamma \rangle & \langle 0_\gamma | H(\gamma, c) | 1_\gamma \rangle \\ \langle 1_\gamma | H(\gamma, c) | 0_\gamma \rangle & \langle 1_\gamma | H(\gamma, c) | 1_\gamma \rangle \end{pmatrix}. \tag{29}$$

It is simple to show that the lowest eigenvalue of this matrix is $\mathcal{E}(\gamma, c)$; i.e.,

$$\mathcal{E}(\gamma, c) = \frac{\langle 0_\gamma | H(\gamma, c) | 0_\gamma \rangle + \langle 1_\gamma | H(\gamma, c) | 1_\gamma \rangle}{2} - \sqrt{\left(\frac{\langle 1_\gamma | H(\gamma, c) | 1_\gamma \rangle - \langle 0_\gamma | H(\gamma, c) | 0_\gamma \rangle}{2} \right)^2 + \langle 0_\gamma | H(\gamma, c) | 1_\gamma \rangle^2}. \tag{30}$$

The eigenstate corresponding to this eigenvalue has the form

$$|\psi_\gamma(c)\rangle = \frac{1}{\sqrt{1 + x(c)^2}} |0_\gamma\rangle - \frac{x(c)}{\sqrt{1 + x(c)^2}} |1_\gamma\rangle, \tag{31}$$

where

$$x(c) = \frac{\sqrt{(\langle 0_\gamma|H(\gamma,c)|0_\gamma\rangle - \langle 1_\gamma|H(\gamma,c)|1_\gamma\rangle)^2 + 4\left(\langle 0_\gamma|H(\gamma,c)|1_\gamma\rangle\right)^2} + \left(\langle 0_\gamma|H(\gamma,c)|0_\gamma\rangle - \langle 1_\gamma|H(\gamma,c)|1_\gamma\rangle\right)}{2\langle 0_\gamma|H(\gamma,c)|1_\gamma\rangle}. \tag{32}$$

Note that $x(c)$ is an odd function of c. The next step is to minimize $\mathcal{E}(\gamma, c)$ with respect to γ and c. Fig. 3 shows what a typical plot looks like for finite values of f^2. For some values of λ and f one can discern a secondary minimum near $c = 0$, but it always lies higher than the minima corresponding to non-zero values of c.

It may seem surprising that although $c_{\min} \to 0$ as $f \to 0$, it doesn't vanish at $f = 0$. Neither does the expectation value of the operator x, although it is much, much smaller than $c_{\min}$. This is because the energy is improved by moving the gaussian slightly off $x = 0$ and widening the symmetric wavefunction, admixing a small amount of the $N = 2$ state. Note, that if we rewrite the Hamiltonian, introducing a factor V as follows,

$$H = \frac{1}{2V}p^2 + \frac{V}{6}\lambda\left(x^2 - f^2\right)^2, \tag{33}$$

and let $V \to \infty$, then $c_{\min} \to 0$ as $f \to 0$. This happens because V acts as a classical *mass* for the oscillator and, one expects that for infinite mass, the wavefunction of the oscillator should become a narrower and narrower and lie at the classical minimum of the potential.

Returning to the case $V = 1$ and non-vanishing f, the fact that the two variational wavefunctions $|\gamma, +c_{\min}\rangle$ and $|\gamma, -c_{\min}\rangle$ have the same energy means that no matter how small the overlap $\langle \gamma, -c_{\min}|H|\gamma, +c_{\min}\rangle$ they will mix maximally. Thus it is possible to do better by considering the symmetric and anti-symmetric wavefunctions

$$\begin{aligned}
|\psi_{\text{even}}\rangle &= e^{icp}\left(\cos(\alpha)|0_\gamma\rangle + \sin(\alpha)|1_\gamma\rangle\right) + e^{-icp}\left(\cos(\alpha)|0_\gamma\rangle - \sin(\alpha)|1_\gamma\rangle\right) \\
|\psi_{\text{odd}}\rangle &= e^{icp}\left(\cos(\alpha)|0_\gamma\rangle + \sin(\alpha)|1_\gamma\rangle\right) - e^{-icp}\left(\cos(\alpha)|0_\gamma\rangle - \sin(\alpha)|1_\gamma\rangle\right),
\end{aligned} \tag{34}$$

and computing the expectation value of the Hamiltonian in each of these states. The formulas needed to carry out this computation appear in Appendix B.

The most important results are:

1. For large f^2 the energy splitting between the even and odd states is, as expected, exponentially small;

2. This splitting grows as f^2 tends towards zero and, in the limit of $f^2 = 0$, becomes of order unity.

These results say that while for large f it is appropriate to think of the splitting as due to tunneling between the degenerate minima, this is not the case when $f \to 0$. This is true even though $c_{\min}$ stays finite. The reason $c_{\min}$ stays finite in the limit $f \to 0$ is that a non-vanishing value of $c_{\min}$ produces a state that is an admixture of $n = 2$ and higher n states with the $n = 0$ state. In fact, due to this admixing this calculation produces energies for both the ground state and first excited state of the pure anharmonic oscillator that are equal in accuracy to our earlier second order adaptive perturbation theory calculation.
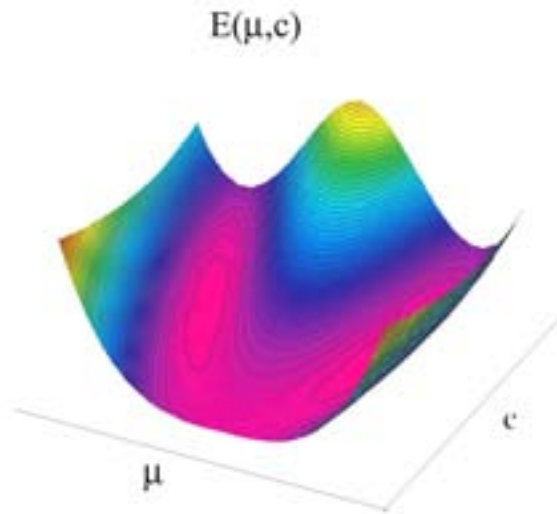

FIG. 3: Effective potential as a function o f c, one global minimum.

A. Larger n

For larger values of n, things change, even for $f \gg 0$. Generically, what happens is that as n grows $c_{\min}$ tends to zero, but doesn't quite get there. Thus, as for the case $n = 0$, there are still two degenerate minima corresponding to equal and opposite values of $c_{\min}$ and so one can lower the energy by forming the states

$$|\psi_{\text{even}}\rangle = e^{icp} \left(\cos(\alpha) |n_\gamma\rangle + \sin(\alpha) |n + 1_\gamma\rangle \right) + e^{-icp} \left(\cos(\alpha) |n_\gamma\rangle - \sin(\alpha) |n + 1_\gamma\rangle \right) \tag{35}$$

and

$$|\psi_{\text{odd}}\rangle = e^{icp} \left(\cos(\alpha) |n_\gamma\rangle + \sin(\alpha) |n + 1_\gamma\rangle \right) - e^{-icp} \left(\cos(\alpha) |N_\gamma\rangle - \sin(\alpha) |n + 1_\gamma\rangle \right). \tag{36}$$

The result of such a computation is to show that as N_γ grows the splitting between these states grows. This follows directly from a computation of the overlap

$$\langle \psi_{\text{odd}} | H | \psi_{\text{even}} \rangle, \tag{37}$$

which, using the formulas in Appendix B, can be shown to grow as a function of n. Eventually, no matter what value we assign to f (so long as it is finite) there is a value of n for which the splitting between the states $|\psi_{\text{even}}\rangle$ and $|\psi_{\text{odd}}\rangle$ becomes of order unity. This is the point at which it makes no sense to talk about tunneling between states defined on one or the other side of the potential barrier.

V. TUNNELING BETWEEN ASYMMETRIC MINIMA

In the preceding section I showed how to use adaptive perturbation theory to compute the tunneling between degenerate minima in a symmetric potential. Clearly, while the adaptive perturbation theory discussion is a rather simple way to get accurate results, it is by no means the only method that can be used. However, if one wishes to consider the tunneling in a very asymmetric potential (see Fig. 4) then the familiar techniques are much more difficult to apply. In particular, if one envisions a problem in which one starts in the lowest state concentrated in the left hand well and wants to know how it evolves in time, then I know of no simple techniques which allow one to compute the behavior accurately. In this section I show how to use the techniques of adaptive perturbation theory, plus one additional trick, to obtain a simple picture of how things change in time.
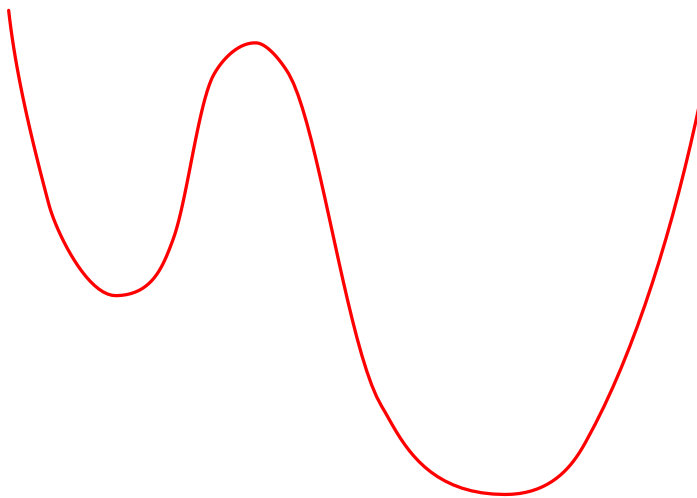

FIG. 4: Asymmetric potential well.

From the preceding discussion, it should be clear that given a specific form for the potential shown in Fig. 4, we can use the adaptive trick of rewriting things in terms of shifted Fock states to approximately compute the lowest

lying state in the left well and a finite but sufficiently large number of n-particle states in the right well. These computations, augmented by perturbation theory, take care of computing the effects which have nothing to do with the tunneling. The tunneling contributions come from computing the transition matrix elements between the lowest lying state shifted to the left of the origin and all the states corresponding to a shift to the right of the origin. This matrix element is computed as in the case of the symmetric potential discussed in the preceding section. Assuming we have carried out this step to the desired degree of accuracy, and that the minima are widely separated, so that the transition matrix element is small, the next step is to compute the time evolution of the state on the left when the tunneling is taken into account.

A. The Simplified Problem

It is important to observe that, at this point, we have effectively mapped the original problem into a simpler one; namely, that of a Hamiltonian H which can be decomposed into $H = H_0 + V$, where H_0 has a discrete set of eigenstates $|\psi_j\rangle$ with energies E_j, $1 < j < N$, (representing the levels in the right hand well) and a single state $|\psi_0\rangle$, with energy E_0 (representing the lowest lying state of the left hand well). Moreover, given the potential in Fig. 4, it will be assumed that E_0 lies somewhere in the range of the E_j, *i.e.*, $E_1 < E_0 < E_N$, and that it is not degenerate with any of the E_j's (I will discuss relaxing the last assumption at a later point). This problem is essentially equivalent to the Wigner-Weisskopf?? problem of a discrete state decaying to a continuum, except that now it is a discrete state decaying to a bunch of discrete states.

Since I have assumed that adaptive perturbation theory was used to determine the E_j's and E_0, the perturbation V will be obtained by computing the transition matrix element between the state $|\psi_0\rangle$ and the states $|\psi_j\rangle$. In other words, the only non-vanishing matrix elements of V will be $V_{0j} = \langle\psi_0|V|\psi_j\rangle$ for $1 \le j \le N$. We now wish to discuss how to compute the probability of finding the system in the state $|\psi_0\rangle$ or $|\psi_j\rangle$ at any time t, assuming it is, at time $t = 0$, in the state $|\psi_0\rangle$. This is, of course, equivalent to computing $\langle\psi_0|e^{-itH}|\psi_0\rangle$ or $\langle\psi_j|e^{-itH}|\psi_0\rangle$.

B. The Resolvent Operator

For reasons which will become clear in a moment, the best way to calculate $\langle\psi_0|e^{itH}|\psi_0\rangle$ for $t > 0$, is in terms of the resolvent operator $1/(H - z)$. This is because the matrix elements of the resolvent operator are directly related to the Laplace transform of the function in question. Explicitly, consider the Laplace transform

$$\psi_{00}(z) = \int_0^\infty dt\, e^{-zt}\, \langle\psi_j|e^{-itH}|\psi_0\rangle, \tag{38}$$

where $z = c + i\omega$ for fixed $c > 0$ and $-\infty \le \omega \le \infty$. Because, in our simplified problem, the eigenvalues of H are real and finite in number, this function is obviously well defined for z in the right-half complex z-plane, and for $t \ge 0$ the inverse transform is given by the contour integral

$$\langle\psi_j|e^{-itH}|\psi_0\rangle = \frac{1}{2\pi i} \int_{c-i\infty}^{c+\infty} \psi_{00}(z)e^{zt}dz. \tag{39}$$

The function is defined in the full complex plane by analytic continuation.

Carrying out the integration in Eq. 38 yields

$$\psi_{00}(z) = -i\langle\psi|\frac{1}{H - iz}|\psi\rangle. \tag{40}$$

Because the resolvent operator only has poles at the discrete eigenvalues of the operator H, it is convenient to substitute $z \to iz$ and redefine the contour in Eq. 39 to obtain

$$\langle\psi|e^{itH}|\psi\rangle = -\int_{\infty+ic}^{-\infty+ic} dz\, \langle\psi|\frac{1}{H - z}|\psi\rangle e^{-itz} \tag{41}$$

where we have assumed $t \ge 0$. Obviously, the same thing can be done for two different states; *i.e.*,

$$\langle\chi|e^{itH}|\psi\rangle = -\int_{\infty+ic}^{-\infty+ic} dz\, \langle\chi|\frac{1}{H - z}|\psi\rangle e^{-itz} \tag{42}$$

The reason for focusing on the resolvent operator is that it will be easy to compute the relevant matrix elements of the resolvent operator exactly. This is because the resolvent operator satisfies the equation

$$\frac{1}{H-z} = \frac{1}{H_0-z} - \frac{1}{H_0-z} V \frac{1}{H-z}, \tag{43}$$

and therefore the matrix element of interest satisfies the corresponding equation

$$\langle\psi|\frac{1}{H-z}|\psi\rangle = \langle\psi|\frac{1}{H_0-z}|\psi\rangle - \langle\psi|\frac{1}{H_0-z} V \frac{1}{H-z}|\psi\rangle, \tag{44}$$

which can be solved by iteration to obtain the formal solution

$$\langle\psi|\frac{1}{H-z}|\psi\rangle = \langle\psi|\frac{1}{H_0-z} \sum_{n=0}^{\infty} (-V \frac{1}{H_0-z})^n|\psi\rangle. \tag{45}$$

For our problem, this sum can be evaluated in closed form because the state $|\psi_0\rangle$ is an eigenstate of H_0 with eigenvalue E_0 and the only non-vanishing matrix elements of V are of the form $V_{0j} = \langle\psi_0|V|\psi_n\rangle$. From this it follows that the sum in Eq. 45 reduces to a geometric series, which is easily summed to give

$$\langle\psi_0|\frac{1}{H-z}|\psi_0\rangle = \frac{1}{E_0-z} \sum_{p=0}^{\infty} \left(\sum_j \frac{\langle\psi_0|V|\psi_j\rangle\langle\psi_j|V|\psi_0\rangle}{(E_0-z)(E_j-z)} \right)^p = \frac{1}{E_0-z-\sum_j \frac{|\langle\psi_0|V|\psi_j\rangle||^2}{E_j-z}}. \tag{46}$$

To cast this into a more useful form multiply both the numerator and denominator of the expression by $\prod_j(E_j - z)$ and rewrite it a rational polynomial. Since the poles of the resolvent operator correspond to the zeros of the denominator, we see that we can rewrite the resulting expression as

$$\langle\psi_0|\frac{1}{H-z}|\psi_0\rangle = \frac{1}{E_0-z-\sum_j \frac{|V_{0j}|^2}{E_j-z}} = \frac{\prod_{j=1}^{N}(E_j-z)}{\prod_{j=0}^{N}(\bar{E}_j-z)}, \tag{47}$$

where we denote the eigenvalues of H as $\bar{E}_j$. Clearly, the $\bar{E}_j$'s can always be written in terms of the corresponding eigenvalues of H_0 as

$$\bar{E}_j = E_j + \delta_j, \tag{48}$$

and thus it is the quantities δ_j which remain to be determined.

Eq. 47 is exact, independent of the size of V. However, since the matrix elements of V are assumed to be quite small, it is useful to rewrite it in a way that allows us to compute the δ_j's perturbatively. To do this, begin by rewriting the second term in Eq. 47 as

$$\begin{aligned}
\prod_{j=0}^{N}(\bar{E}_j-z) = {}& \prod_{j=0}^{N}(E_j-z) - |V_{01}|^2 \prod_{j=2}^{N}(E_j-z) \\
& - |V_{02}|^2(E_1-z)\prod_{j=2}^{N}(E_j-z) - \ldots - |V_{0p}|^2 \prod_{j=1}^{p-1}(E_j-z) \prod_{l=p+1}^{N}(E_l-z) \\
& - \ldots - |V_{0N}|^2 \prod_{j=1}^{N-1}(E_j-z).
\end{aligned} \tag{49}$$

Since $|V_{0j}|^2 \ll 1$ this implies $\delta_j \approx |V_{0j}|^2$. Thus, for example, setting $j = 0$ and choosing $z = E_0 + \delta_0$, the equation becomes

$$\begin{aligned}
0 = {}& -\delta_0 \prod_{j=1}^{N}(E_j-E_0) - |V_{01}|^2 \prod_{j=2}^{N} -|V_{02}|^2(E_1-E_0)\prod_{j=3}^{N} \\
& - \ldots - |V_{0p}|^2 \prod_{j=1}^{p-1}(E_j-E_0) \prod_{l=p+1}^{N}(E_l-E_0) - \ldots - |V_{0N}|^2 \prod_{j=1}^{N-1}(E_j-E_0),
\end{aligned} \tag{50}$$

or, in other words,

$$\delta_0 = -\sum_{j=1}^{N} \frac{|V_{0j}|^2}{E_j - E_0}. \tag{51}$$

The other energy shifts are obtained in a similar manner. For example, if we set $z = \bar{E}_1 = E_1 + \delta_1$ then, to lowest order in δ_1 or $|V_{0j}|^2$, the equation becomes

$$0 = (E_0 - E_1)(-\delta_1) \prod_{j=2}^{N}(E_j - E_1). - |V_{01}|^2 \prod_{j=2}^{N}(E_j - E_1) \tag{52}$$

This is true because the other terms take the form

$$-|V_{02}|^2(-\delta_1) \prod_{j=3}^{N}(E_j - E_1) - \ldots - |V_{0m}|^2(-\delta_1) \prod_{j=2}^{m-1}(E_j - E_1) \prod_{l=m+1}^{N}(E_j - E_1), \tag{53}$$

which are all of second order in the perturbation. Thus, finally, we have

$$\delta_1 = \frac{|V_{01}|^2}{E_0 - E_1}. \tag{54}$$

Similarly, we for all other values of j we obtain

$$\delta_j = -\frac{|V_{0j}|^2}{E_0 - E_j}. \tag{55}$$

(Note, I have implicitly assumed that the spacing between the levels E_j is larger than the typical value of $|V_{0j}|^2$ and that E_0 is not degenerate with any of the E_j's. The second restriction is easy to remove, since if $E_0 \approx E_j$ for some j, then all we have to treat the computation of δ_0 and δ_j more carefully, solving a quadratic equation for the two roots. All the other δ_j's remain unchanged. If however the spectrum of the E_j's becomes dense, then one has to do a more careful job for the levels which lie within a distance of order $|V_{0j}|^2$ of E_0.)

These results are half of what we need to reconstruct the full time dependence of $\langle \psi_0 | e^{itH} | \psi_0 \rangle$. From Eq. 39 we see that since the matrix element of the resolvent operator is a sum of simple poles, we can get the full time dependence of the matrix element by closing the contour, once we know the residues of the poles. Fortunately, given our previous discussion, it is easy to compute the residues of the poles. First, we rewrite Eq. 47 as

$$\langle \psi_0 | \frac{1}{H - z} | \psi_0 \rangle = \frac{1}{E_0 - z - \sum_j \frac{|V_{0j}|^2}{E_j - z}} = \frac{\prod_{j=1}^{N}(E_j - z)}{\prod_{j=0}^{N}(\bar{E}_j - z)} = \frac{\prod_{j=1}^{N}(\bar{E}_j - z - \delta_j)}{\prod_{s=0}^{N}(\bar{E}_s - z)}, \tag{56}$$

and expand the result to first order in δ_j. This yields the equation

$$\langle \psi_0 | \frac{1}{H - z} | \psi_0 \rangle = \frac{1}{\bar{E}_0 - z} \left(1 - \sum_{j=1}^{N} \frac{\delta_j}{\bar{E}_j - z} \right). \tag{57}$$

Next, we note that Eq. 57 explicitly shows that there is a pole at every value $z = \bar{E}_j$, so it becomes a simple matter to compute the residues to the same order in the perturbation. Letting $z \to E_0$ we obtain the result

$$r_0 = -\left(1 - \sum_{j=1}^{N} \frac{\delta_j}{\bar{E}_j - \bar{E}_0} \right), \tag{58}$$

and, similarly, letting $z \to \bar{E}_j$ we obtain

$$r_j = -\frac{\delta_j}{\bar{E}_j - \bar{E}_0}. \tag{59}$$

Combining these results, we see that the amplitude for finding the particle in the state $|\psi_0\rangle$, at any time t, is given by

$$\langle \psi_0 | e^{i H t} | \psi_0 \rangle = \sum_{j=0}^{N} -r_j e^{-i \bar{E}_j t}. \tag{60}$$

It follows from evaluating Eq. 60 at $t = 0$ that

$$r_0 + \sum_{j=1}^{N} r_j = -1, \tag{61}$$

which is easily shown to be true.

To get the amplitude for making a transition to state $|\psi_m\rangle$, we have to compute

$$\langle \psi_m | \frac{1}{H - z} | \psi_0 \rangle. \tag{62}$$

However, since V only causes a transition from $|\psi_m\rangle$ to $|\psi_0\rangle$, it follows from the integral equation

$$\frac{1}{H - z} = \frac{1}{H_0 - z} - \frac{1}{H_0 - z} V \frac{1}{H - z} \tag{63}$$

that

$$\langle \psi_m | \frac{1}{H - z} | \psi_0 \rangle = \frac{V_{m0}}{E_m - z} \left(\frac{\prod_{j=1}^{N}(E_j - z)}{\prod_{l=0}^{N}(\bar{E}_l - z)} \right). \tag{64}$$

Canceling the factors of $(E_j - z)$ we obtain

$$\langle \psi_m | \frac{1}{H - z} | \psi_0 \rangle = V_{m0} \frac{\prod_{j=1}^{m-1}(\bar{E}_j - z - \delta_j) \prod_{l=m+1}^{N}(\bar{E}_l - z - \delta_l)}{\prod_{k=0}^{N}(\bar{E}_k - z)}. \tag{65}$$

Expanding this expression to first order in the δ's we obtain

$$\begin{aligned}
\tilde{r}_0 &= -\frac{V_{m0}}{\bar{E}_m - \bar{E}_0} + \sum_{j=1}^{m-1} \frac{\delta_j}{(\bar{E}_m - \bar{E}_0)(\bar{E}_j - \bar{E}_0)} \sum_{l=m+1}^{N} \frac{\delta_l}{(\bar{E}_m - \bar{E}_0)(\bar{E}_l - \bar{E}_0)} \\
\tilde{r}_m &= -\frac{V_{m0}}{\bar{E}_0 - \bar{E}_m} + \sum_{j=1}^{m-1} \frac{\delta_j}{(\bar{E}_0 - \bar{E}_m)(\bar{E}_j - \bar{E}_m)} \sum_{l=m+1}^{N} \frac{\delta_l}{(\bar{E}_0 - \bar{E}_l)(\bar{E}_l - \bar{E}_m)} \\
\tilde{r}_j &= \frac{\delta_j}{(\bar{E}_0 - \bar{E}_j)(\bar{E}_m - \bar{E}_j)}, \quad \text{where } j \neq 0 \text{and } j \neq m
\end{aligned} \tag{66}$$

It is important to check that this solution gives the correct boundary condition that at $t = 0$; i.e., $\langle \psi_m | e^{i H t} | \psi_0 \rangle = 0$. Clearly, this is the same as requiring that the sum of the residues $\tilde{r}_j$ vanish. Applying the preceding result we see that

$$\begin{aligned}
\tilde{r}_0 + \tilde{r}_m &= \sum_{j=1}^{m-1} \frac{\delta_j}{\bar{E}_m - \bar{E}_0} \left(\frac{1}{\bar{E}_j - \bar{E}_0} - \frac{1}{\bar{E}_j - \bar{E}_m} \right) + \sum_{j=m+1}^{N} \frac{\delta_j}{\bar{E}_m - \bar{E}_0} \left(\frac{1}{\bar{E}_j - \bar{E}_0} - \frac{1}{\bar{E}_j - \bar{E}_m} \right) \\
&= \sum_{j=1}^{m-1} \frac{-\delta_j}{(\bar{E}_j - \bar{E}_0)(\bar{E}_j - \bar{E}_m)} + \sum_{l=m+1}^{N} \frac{-\delta_l}{(\bar{E}_l - \bar{E}_0)(\bar{E}_l - \bar{E}_m)},
\end{aligned} \tag{67}$$

and therefore,

$$\tilde{r}_0 + \tilde{r}_m + \sum_{j=1}^{m-1} \tilde{r}_j + \sum_{l=m+1}^{N} \tilde{r}_l = 0, \tag{68}$$

which is the desired result.

Given these formulas, we can now write the time-dependent state $|\psi_0(t)\rangle$ as

$$|\psi_0(t)\rangle = e^{itH}|\psi_0\rangle = \sum_{m=0}^{N} \langle\psi_m|e^{itH}|\psi_0\rangle|\psi_m\rangle. \tag{69}$$

It is easy to see that in the case where the eigenvalues of H_0 are widely spaced on the scale of $|V_{0j}|^2$, the largest residues, and therefore the largest contribution to $|\psi_0(t)\rangle$, comes from $|\psi_0\rangle$ and the state $|\psi_m\rangle$ for which $E_m - E_0$ takes the smallest value. If, however, one takes the limit in which the levels in the right well become dense, as would be the case if we made the right hand minimum very flat and very wide, then things change. The formula Eq. 46 remains exact, but now the sum over levels is best approximated by an integral over a continuous variable E; *i.e.*,

$$\langle\psi_0|\frac{1}{H-z}|\psi_0\rangle = \frac{1}{E_0-z}\sum_{p=0}^{\infty}\left(\int_{E_{min}}^{E_{max}} \frac{\langle\psi_0|V|\psi(E)\rangle\langle\psi(E)|V|\psi_0\rangle}{(E_0-z)(E_j-z)}\right)^p = \frac{1}{E_0 - z - \int_{E_{min}}^{E_{max}} \frac{|\langle\psi_0|V|\psi(E)\rangle||^2}{E-z}}. \tag{70}$$

Now, since E_0 lies in the region of integration, we see immediately from the fact that the integral develops an imaginary part for $E_{min} \leq z \leq E_{max}$ that the resolvent operator no longer has a pole at real $z \approx E_0$. Instead, the zero of the denominator, and thus the pole of the resolvent operator, moves off to a complex value of z on the second sheet of the function. Now, in order to compute the time dependence of the state one has to deform the contour for the inverse transform and pick up this pole, giving the usual exponential decay of the state $|\psi_0\rangle$, plus a background integral that does not decay exponentially in t for large t.

VI.   A LOOK AHEAD TO FIELD THEORY

Before concluding it is worth taking a quick look ahead to the case of an interacting scalar field theory, to see how the techniques of adaptive perturbation theory can be used to non-perturbatively derive the structure of renormalization. In this section I will focus on a simple calculation that gives a non-perturbative approximation to mass renormalization and leave the discussion of how one can capture the structure of wavefunction and coupling constant renormalization to the next paper.

For the sake of this introduction to the application of the techniques of adaptive perturbation theory to a field theory, I will focus on the case of a self-interacting scalar field, with a Hamiltonian of the form

$$H = \int d^n x \left[\frac{1}{2}\Pi_\phi(x)^2 + \frac{1}{2}\nabla\phi(x)\cdot\nabla\phi(x) + \frac{\lambda}{6}\phi(x)^4\right], \tag{71}$$

where the spatial dimension n can be arbitrary. In what follows I assume that everything is defined in a finite volume, V, so that momenta are discrete. This isn't necessary, but it will make all of the manipulations which follow well-defined. Given this assumption, the Fourier transforms of the field and its conjugate momentum are defined to be

$$\phi(\vec{k}) = \frac{1}{\sqrt{V}}\int d^n x e^{-i\vec{k}\cdot\vec{x}}\phi(x) \quad ; \quad \Pi_\phi(\vec{k}) = \frac{1}{\sqrt{V}}\int d^n x e^{-i\vec{k}\cdot\vec{x}}\Pi_\phi(x). \tag{72}$$

In terms of these operators the Hamiltonian can be rewritten as

$$H = \sum_{\vec{k}}\left[\frac{\Pi_\phi(-\vec{k})\Pi_\phi(\vec{k})}{2} + \frac{(\vec{k}\cdot\vec{k}+m^2)}{2}\phi(\vec{k})\,\phi(-\vec{k})\right] + \frac{\lambda}{6}\frac{1}{V}\sum_{\vec{k}_i}\phi(\vec{k}_1)\,\phi(\vec{k}_2)\phi(\vec{k}_3)\,\phi(\vec{k}_4)\delta(\sum_{i=1}^{4}\vec{k}_i). \tag{73}$$

At this point I set up the adaptive perturbation theory calculation by defining $\gamma(k)$-dependent annihilation and creations operators as follows,

$$\phi(\vec{k}) = i\sqrt{\frac{\gamma(\vec{k})}{2}}(\mathcal{A}(-\vec{k})^\dagger - \mathcal{A}(\vec{k})) \quad ; \quad \Pi(\vec{k}) = \frac{(\mathcal{A}(-\vec{k})^\dagger + \mathcal{A}(\vec{k}))}{\sqrt{2\,\gamma(\vec{k})}}, \tag{74}$$

and define the associated *vacuum* state for this choice of $\gamma(\vec{k})$'s,

$$|vac\rangle = \prod_k |0_{\gamma(k)}\rangle, \tag{75}$$

by the condition that it be annihilated by all the $\mathcal{A}(\vec{k})$'s.

Adaptive perturbation theory around this vacuum state is now determined by minimizing the vacuum expectation value of the Hamiltonian with respect to all of the $\gamma(\vec{k})$'s. It follows directly from these definitions that the function to be minimized is

$$\langle vac| H |vac\rangle = \sum_k \left[\frac{\gamma(\vec{k})}{4} + \frac{(\vec{k} \cdot \vec{k} + m^2)}{4\,\gamma(\vec{k})} \right] + \frac{\lambda}{4V} \left[\sum_{\vec{k}} \frac{1}{\gamma(\vec{k})} \right]^2. \tag{76}$$

Obviously, if the range of the momenta appearing in these sums are unrestricted, these expressions diverge, as happens in all field theories. In the context of perturbation theory one usually deals with this problem by regulating the integrals and then adding counterterms to the Lagrangian to cancel these divergences. Since I wish to discuss this theory non-perturbatively, I will adopt a different strategy; namely, I will render the theory well defined by restricting the operators $\phi(\vec{k})$ and $\Pi_\phi(\vec{k})$ to be finite in number. This can be done in a variety of ways. One is to put the theory on a lattice by defining the momentum components to be $k_\mu(j) = 2\pi j/L_\mu a$, where L_μ is the size of the lattice in the μ direction, a is the lattice spacing and j is an integer running from $-L_\mu/a \le j \le L_\mu/a$. Another way to accomplish the same thing, is to keep only those operators $\phi(\vec{k})$ and $\Pi_\phi(\vec{k})$ for which $\vec{k} \cdot \vec{k} \le \Lambda$. Since nothing I will say from this point on requires that I specify a particular scheme, I avoid making any detailed assumptions.

Minimizing Eq. 76 with respect to each $\gamma(\vec{k})$ yields

$$\frac{1}{4} - \frac{(\vec{k} \cdot \vec{k} + m^2)}{4\,\gamma(k)^2} - \frac{\lambda}{2\gamma(\vec{k})^2} \left[\frac{1}{V} \sum_{\vec{k}'} \frac{1}{\gamma(\vec{k}')} \right] = 0, \tag{77}$$

which can be rearranged to give

$$\gamma(k)^2 = \vec{k} \cdot \vec{k} + m^2 + 2\,\lambda \left[\frac{1}{V} \sum_{\vec{k}'} \frac{1}{\gamma(\vec{k}')} \right]. \tag{78}$$

Setting $\vec{k} = 0$ we obtain

$$\gamma(0)^2 = m^2 + 2\,\lambda \left[\frac{1}{V} \sum_{\vec{k}'} \frac{1}{\gamma(\vec{k}')} \right], \tag{79}$$

which can be substituted into Eq. 78 to give

$$\gamma(\vec{k})^2 = \vec{k} \cdot \vec{k} + \gamma(0)^2. \tag{80}$$

Substituting this expression into the equation for $\gamma(0)$, it becomes the non-perturbative equation

$$\gamma(0)^2 = m^2 + 2\,\lambda \frac{1}{V} \sum_{\vec{k}'} \frac{1}{\sqrt{\vec{k}' \cdot \vec{k}' + \gamma(0)^2}}, \tag{81}$$

which is reminiscent of the Nambu Jona-Lasinio equation. This equation obviously has a solution, since the formula for the vacuum energy clearly diverges both at $\gamma(0) = 0$ and $\gamma(0) = \infty$ and is finite in between. At this point we can take the infinite volume limit and replace the sum over $\vec{k}$'s by an integral up to some cut-off Λ. In this case the equation to be solved becomes

$$\gamma(0)^2 - m^2 = 2\,\lambda \frac{2^p}{(2\pi)^p} \int_0^\Lambda k^{p-1} dk \frac{1}{\sqrt{\vec{k} \cdot \vec{k} + \gamma(0)^2}}. \tag{82}$$

The simplest way to understand the general properties of the solution to this equation is to graph the left and right hand sides as a function of $\gamma(0)$ and see where they cross. Clearly, the graph of the left hand side of Eq. 82 is a parabola that intersects the vertical axis at $-m^2$. The graph of the right hand side of the equation starts at some large (or in 1+1-dimension infinite) value and then drops monotonically to zero as $\gamma(0) \to \infty$. Generically, since for finite $\gamma(0) \ll \Lambda$, the integral is proportional to Λ^{p-1} for $p > 1$, and $\ln(\Lambda)$ for $p = 1$, the point where the parabola

crosses the graph of the integral will be for large $\gamma(0)$. The only way to have the graphs cross for a value of $\gamma(0)$ that is small with respect to Λ is to make m^2 large and negative. In fact, one should really think of Eq. 82 not as an equation for $\gamma(0)$, but rather as an equation for m^2; *i.e.*, rewrite it as

$$m^2 = \gamma(0)^2 - 2\lambda \frac{2^p}{(2\pi)^p} \int_0^\Lambda k^{p-1} dk \frac{1}{\sqrt{\vec{k} \cdot \vec{k} + \gamma(0)^2}} \tag{83}$$

and then arbitrarily pick a value of $\gamma(0)$ and use this equation to determine for what value of m^2 the chosen value of $\gamma(0)$ will minimize the ground state energy density. This is, of course, nothing but a non-perturbative way of determining the leading mass renormalization counter term.

While this discussion is amusing, it does nothing to show what one has to do to similarly understand wavefunction and coupling constant renormalization in a non-perturbative context. That discussion will be the subject of the second paper in this series.

VII. REPRISE

It was my purpose in this paper to introduce the notion of *adaptive perturbation theory* as a way of perturbatively computing quantities that one heretofore didn't know how to compute in any simple way. I focused on problems in ordinary quantum mechanics, such as the pure anharmonic oscillator and the problem of computing the eigenstates of generic double-well problems, as well as computing tunneling between states in distinct wells, in order to exhibit the versatility of the technique. Finally, I discussed the simplest way to begin to apply these techniques to a field theory in order to obtain a better behaved perturbation expansion independent of the size of the interaction term. In addition, I tried to indicate how the same approach would allow one to non-perturbatively approach the computation of renormalized quantities.

In the next paper I will show how to extend these techniques to non-perturbatively capture the full structure of renormalization. I will also show how something related to the running coupling constant appears and explain why this approach essentially forces one to discuss the computation of scattering amplitudes in a way that is intimately related to the approach used in the parton model.

The key point I would like the reader to take away from this discussion is that, in an interacting theory, the perturbation theory one uses to compute an interesting quantity must be adapted to the problem at hand. Although a variational computation based on energy considerations is the tool used to define the appropriate parameters to be used when dividing the Hamiltonian up into an *unperturbed* part and and an *interaction term*, it is the perturbation theory based upon this division which is crucial to computing quantities other than the energy of a state. This observation takes us beyond more usual variational methods which cannot be usefully applied to compute matrix elements of general operators, or scattering amplitudes in a field theory.

VIII. ACKNOWLEDGEMENTS

APPENDIX A. SOME OPERATOR IDENTITES

This appendix is devoted to reminding the reader how to derive useful normal ordering identities that are useful for carrying out the computations I have already described, as well as the computations in Appendix B.

Theorem:

$$e^{\alpha(A_\gamma^\dagger + A_\gamma)} = e^{\alpha^2/2} e^{\alpha A_\gamma^\dagger} e^{\alpha A_\gamma} \tag{84}$$

Proof: Assume

$$e^{\alpha(A_\gamma^\dagger + A_\gamma)} = e^{f(\alpha)} e^{g(\alpha)A_\gamma^\dagger} e^{h(\alpha)A_\gamma} \tag{85}$$

and differentiate with respect to α to obtain

$$(A_\gamma^\dagger + A_\gamma)\, e^{\alpha(A_\gamma^\dagger + A_\gamma)} = \left[\frac{d}{d\alpha}f(\alpha) + \frac{d}{d\alpha}g(\alpha)A_\gamma^\dagger + \frac{d}{d\alpha}h(\alpha)\, e^{g(\alpha)A_\gamma^\dagger}\, e^{h(\alpha)A_\gamma}\, e^{-g(\alpha)A_\gamma^\dagger} \right] e^{f(\alpha)}\, e^{g(\alpha)A_\gamma^\dagger}\, e^{h(\alpha)A_\gamma}. \tag{86}$$

Successive differentiation of the term $e^{g(\alpha)A_\gamma^\dagger}\, e^{h(\alpha)A_\gamma}\, e^{-g(\alpha)A_\gamma^\dagger}$ with respect to α gives the identity

$$e^{g(\alpha)A_\gamma^\dagger}\, e^{h(\alpha)A_\gamma}\, e^{-g(\alpha)A_\gamma^\dagger} = \sum_n \frac{(g(\alpha))^n}{n!} [A_\gamma^\dagger, [A_\gamma^\dagger, [\ldots, [A_\gamma^\dagger, A_\gamma]]] \ldots]_n. \tag{87}$$

Observing that $[A_\gamma^\dagger, A_\gamma] = -1$ allows us to rewrite this as

$$e^{g(\alpha)A_\gamma^\dagger}\, e^{h(\alpha)A_\gamma}\, e^{-g(\alpha)A_\gamma^\dagger} = A_\gamma - g(\alpha) \tag{88}$$

Substitute this into Eq. 86 and cancel the factor of $e^{\alpha(A_\gamma^\dagger + A_\gamma)}$ against the term $e^{f(\alpha)}\, e^{g(\alpha)A_\gamma^\dagger}\, e^{h(\alpha)A_\gamma}$ on the right hand side of the equation to obtain

$$(A_\gamma^\dagger + A_\gamma) = \frac{d}{d\alpha}f(\alpha) + \frac{d}{d\alpha}g(\alpha)A_\gamma^\dagger + \frac{d}{d\alpha}h(\alpha)\, (A_\gamma - g(\alpha))\,. \tag{89}$$

Equating coefficients of $A_\gamma^\dagger$, A_γ and the unit operator yields the following differential equations:

$$\frac{d}{d\alpha}g(\alpha) = 1 \quad ; \quad \frac{d}{d\alpha}h(\alpha) = 1 \quad ; \quad \frac{d}{d\alpha}f(\alpha) = g(\alpha)\frac{d}{d\alpha}h(\alpha) \tag{90}$$

The initial conditions $f(0) = 0$, $g(0) = 0$ and $h(0) = 0$ follow from the fact that $e^{\alpha(A_\gamma^\dagger + A_\gamma)}$ is the identity operator for $\alpha = 0$. Given these boundary conditions, it follows directly that

$$g(\alpha) = \alpha \quad ; \quad h(\alpha) = \alpha \quad ; \quad f(\alpha) = \frac{\alpha^2}{2}, \tag{91}$$

which is what we wished to prove.

A similar calculation proves the following result:

Theorem:

$$e^{\alpha(A_\gamma^\dagger - A_\gamma)} = e^{-\alpha^2/2}\, e^{\alpha A_\gamma^\dagger}\, e^{-\alpha A_\gamma} \tag{92}$$

This identity is used to compute, among other things, the overlap $\langle 0_\gamma | e^{-icp} | 0_\gamma \rangle$ as follows:

$$\langle 0_\gamma | e^{-icp} | 0_\gamma \rangle = \langle 0_\gamma | e^{c\sqrt{\frac{\gamma}{2}}(A_\gamma^\dagger - A_\gamma)} | 0_\gamma \rangle = e^{-c^2\gamma/4} \langle 0_\gamma | e^{c\sqrt{\frac{\gamma}{2}}A_\gamma^\dagger}\, e^{c\sqrt{\frac{\gamma}{2}}A_\gamma} | 0_\gamma \rangle = e^{-c^2\gamma/4}, \tag{93}$$

since, by assumption, $A_\gamma | 0_\gamma \rangle = 0$

I mentioned that there was a simple expression for the state $|0_\gamma\rangle$ in terms of the states built upon $|0_{\gamma'}\rangle$. The explicit form of this relation is the following:

Theorem:

$$|0_\gamma\rangle = \sqrt{1 - \tanh(\beta)^2}\, e^{-\frac{1}{2}\tanh(\beta)\, A_\gamma^{\dagger\,2}} |0_{\gamma'}\rangle, \tag{94}$$

where $\tanh(\beta) = (\gamma - \gamma')/(\gamma + \gamma')$.

Proof: Given that we introduced that γ-dependent annihilation and creation operators as

$$x = \frac{A_\gamma^\dagger + A_\gamma}{\sqrt{2\gamma}} \quad ; \quad p = i\sqrt{\frac{2}{\gamma}}(A_\gamma^\dagger - A_\gamma) \tag{95}$$

it follows that for two different values of γ

$$A_\gamma = \frac{1}{2} \frac{(\gamma + \gamma')}{\sqrt{\gamma \gamma'}} A_{\gamma'} + \frac{1}{2} \frac{(\gamma - \gamma')}{\sqrt{\gamma \gamma'}} A_{\gamma'}^\dagger. \tag{96}$$

By definition $A_\gamma |0_\gamma\rangle = 0$. Thus we have

$$\frac{1}{2} \frac{(\gamma + \gamma')}{\sqrt{\gamma \gamma'}} A_{\gamma'} |0_\gamma\rangle + \frac{1}{2} \frac{(\gamma - \gamma')}{\sqrt{\gamma \gamma'}} A_{\gamma'}^\dagger |0_\gamma\rangle = 0. \tag{97}$$

Now, consider a state of the form

$$e^{\lambda A_{\gamma'}^\dagger{}^2} |0_{\gamma'}\rangle. \tag{98}$$

Requiring that A_γ annihilates this state is equivalent to the statement

$$\frac{1}{2} \frac{(\gamma + \gamma')}{\sqrt{\gamma \gamma'}} A_{\gamma'} e^{\lambda (A_{\gamma'}^\dagger)^2} |0_{\gamma'}\rangle + \frac{1}{2} \frac{(\gamma - \gamma')}{\sqrt{\gamma \gamma'}} A_{\gamma'}^\dagger e^{\lambda (A_{\gamma'}^\dagger)^2} |0_{\gamma'}\rangle = 0. \tag{99}$$

Moving the factor $e^{\lambda A_{\gamma'}^\dagger{}^2}$ to the left of each term we obtain

$$e^{\lambda A_{\gamma'}^\dagger{}^2} \left[\frac{1}{2} \frac{(\gamma + \gamma')}{\sqrt{\gamma \gamma'}} e^{-\lambda (A_{\gamma'}^\dagger)^2} A_{\gamma'} e^{\lambda (A_{\gamma'}^\dagger)^2} + \frac{1}{2} \frac{(\gamma - \gamma')}{\sqrt{\gamma \gamma'}} A_{\gamma'}^\dagger \right] |0_\gamma'\rangle = 0. \tag{100}$$

Once again, expanding the term $e^{-\lambda A_{\gamma'}^\dagger{}^2} A_{\gamma'} e^{\lambda A_{\gamma'}^\dagger{}^2}$ in multiple commutators we obtain the results that A_γ will annihilate the state if and only if

$$\lambda \frac{(\gamma + \gamma')}{\sqrt{\gamma \gamma'}} + \frac{1}{2} \frac{(\gamma - \gamma')}{\sqrt{\gamma \gamma'}} = 0, \tag{101}$$

or, in other words, if

$$\lambda = \frac{1}{2} \left(\frac{\gamma - \gamma'}{\gamma + \gamma'} \right) = \tanh(\beta). \tag{102}$$

This is what had to be proved.

Finally, the normalization factor come from the identity

$$e^{\lambda A_{\gamma'}^\dagger{}^2} |0_{\gamma'}\rangle = \sum_{n=0}^\infty \frac{\lambda^n}{n!} \sqrt{2n!} |2n_{\gamma'}\rangle \tag{103}$$

The norm squared of this state is given by the sum

$$\sum_{n=0}^\infty \frac{\lambda^{2n}}{n!^2} 2n! = \frac{1}{\sqrt{1 - 4\lambda^2}}, \tag{104}$$

so the normalized state is

$$|0_\gamma\rangle = (1 - \tanh(\beta)^2)^{1/4} e^{\frac{1}{2} \tanh(\beta) A_{\gamma'}^\dagger{}^2} |0_{\gamma'}\rangle. \tag{105}$$

It is worth noting that another way to derive this result is to observe that

$$e^{\frac{\beta}{2}(A^{\dagger 2} - A^2)} \tag{106}$$

is a unitary operator that satisfies the relation

$$e^{\frac{\beta}{2}(A^{\dagger 2} - A^2)} = e^{-\frac{1}{2} \tanh(\beta) A^{\dagger 2}} e^{-\ln(\cosh(\beta))(A^\dagger A + \frac{1}{2})} e^{\frac{1}{2} \tanh(\beta) A^{\dagger 2}}. \tag{107}$$

This result is proven in much the same way as the other normal ordering relations. Putting back the appropriate γ' dependence in this equation and applying the operator to the state $|0_{\gamma'}\rangle$ gives the desired result.

APPENDIX B. DETAILS OF DOUBLE WELL CALCULATION

As we already noted, the two degenerate variational states for non-vanishing f^2 are defined as:

$$|\chi(c,\gamma)\rangle = e^{-icp}|\psi_\gamma(c)\rangle \tag{108}$$

and

$$|\chi(-c,\gamma)\rangle = e^{icp}|\psi_\gamma(-c)\rangle \tag{109}$$

(assuming $c > 0$ throughout).

Once one has found the best choice of variational parameters, still lower energy states are obtained by forming the even and odd parity combinations:

$$|\psi_{\text{even}}(\gamma)\rangle = \frac{1}{N_{\text{even}}}\left(|\chi(c,\gamma)\rangle + |\chi(-c,\gamma)\rangle\right) \tag{110}$$

and

$$|\psi_{\text{odd}}(\gamma)\rangle = \frac{1}{N_{\text{odd}}}\left(|\chi(c,\gamma)\rangle - |\chi(-c,\gamma)\rangle\right), \tag{111}$$

where N_{even} and N_{odd} are normalization factors.

To evaluate the normalization factors N_{even} and N_{odd} we need to compute the scalar product $\langle\chi(c,\gamma)|\chi(-c,\gamma)\rangle$ as well as overlap elements $\langle\chi(c,\gamma)|H|\chi(-c,\gamma)\rangle$. The diagonal scalar products are obviously unity and the diagonal expectation values of the Hamiltonian are taken to be the values that minimize $\mathcal{E}(c,\gamma)$.

Expressing p as a sum of annihilation and creation operators and using the identity for computing the normal ordered exponential, Eq. 92, allows one to rewrite the overlap as

$$\langle\chi(-c,\gamma)|\chi(c,\gamma)\rangle = \langle\psi_{-c}(\gamma)|e^{-2icp}|\psi_c(\gamma)\rangle = e^{-c^2\gamma}\langle\psi_{-c}(\gamma)|e^{c\sqrt{2\gamma}A_\gamma^\dagger}e^{-c\sqrt{2\gamma}A_\gamma}|\psi_c(\gamma)\rangle = e^{-c^2\gamma}\langle\phi_{-c}(\gamma)|\phi_c(\gamma)\rangle, \tag{112}$$

where the states $\langle\phi_c(\gamma)|$ and $|\phi_{-c}(\gamma)\rangle$ are defined in terms of the states

$$|\psi_c(\gamma)\rangle = \frac{1}{\sqrt{1+x(c)^2}}|\psi_0(\gamma)\rangle - \frac{x(c)}{\sqrt{1+x(c)^2}}|\psi_1(\gamma)\rangle \tag{113}$$

and

$$|\psi_{-c}(\gamma)\rangle = \frac{1}{\sqrt{1+x(c)^2}}|\psi_0(\gamma)\rangle + \frac{x(c)}{\sqrt{1+x(c)^2}}|\psi_1(\gamma)\rangle, \tag{114}$$

(where $x(c)$ is given by the formula for positive c), to be

$$|\phi_c(\gamma)\rangle = e^{-c\sqrt{2\gamma}A_\gamma}|\psi_c(\gamma)\rangle \tag{115}$$

and

$$|\phi_{-c}(\gamma)\rangle = e^{c\sqrt{2\gamma}A_\gamma}|\psi_{-c}(\gamma)\rangle. \tag{116}$$

Since the state $|\psi_c(\gamma)\rangle$ contains only $n = 0$ and $n = 1$ states, only the first two terms in the expansion of the exponential act. Thus, it follows that

$$|\phi_c(\gamma)\rangle = e^{-c\sqrt{2\gamma}A_\gamma}|\psi_c(\gamma)\rangle = \left(\frac{1+c\sqrt{2\gamma}\,x(c)}{\sqrt{1+x(c)^2}}\right)|\psi_0(\gamma)\rangle - \frac{x(c)}{\sqrt{1+x(c)^2}}|\psi_1(\gamma)\rangle \tag{117}$$

and

$$|\phi_{-c}(\gamma)\rangle = e^{c\sqrt{2\gamma}\,A_\gamma}|\psi_{-c}(\gamma)\rangle = \left(\frac{1 + c\sqrt{2\gamma}\,x(c)}{\sqrt{1 + x(c)^2}}\right)|\psi_0(\gamma)\rangle + \frac{x(c)}{\sqrt{1 + x(c)^2}}|\psi_1(\gamma)\rangle, \tag{118}$$

so the overlap is

$$\langle\chi(-c,\gamma)|\chi(c,\gamma)\rangle = e^{-c^2\gamma}\langle\phi_{-c}(\gamma)|\phi_c(\gamma)\rangle = e^{-c^2\gamma}\frac{\left(1 + c\sqrt{2\gamma}\,x(c)\right)^2 - x(c)^2}{1 + x(c)^2}. \tag{119}$$

Thus, the normalization for the state $|\psi_{\text{even}}(\gamma)\rangle$ is

$$N_{\text{even}} = \sqrt{(1 + \langle\chi(-c,\gamma)|\chi(c,\gamma)\rangle)}. \tag{120}$$

Similarly

$$N_{\text{odd}} = \sqrt{2(1 - \langle\chi(-c,\gamma)|\chi(c,\gamma)\rangle)}. \tag{121}$$

Finally, to compute the formula for the expectation value of H in the state $|\psi_{\text{even}}(c,\gamma)\rangle$ or $|\psi_{\text{odd}}(c,\gamma)\rangle$, we need the following identities:

$$\begin{aligned}
e^{-icp}\,H(p,x)\,e^{icp} &= H(p,x-c) \\
e^{icp}\,H(p,x)\,e^{-icp} &= H(p,x+c) \\
e^{-icp}\,H(p,x)\,e^{-icp} &= H(p,x-c)\,e^{-2icp} = e^{-c^2\gamma}\,H(p,x-c)\,e^{c\sqrt{2\gamma}A_\gamma^\dagger}\,e^{-c\sqrt{2\gamma}A_\gamma} \\
&= e^{-c^2\gamma}\,e^{c\sqrt{2\gamma}A_\gamma^\dagger}\,H(p-ic\gamma,x)\,e^{-c\sqrt{2\gamma}A_\gamma} \\
e^{icp}\,H(p,x)\,e^{icp} &= H(p,x+c)\,e^{2icp} = e^{-c^2\gamma}\,H(p,x+c)\,e^{c\sqrt{2\gamma}A_\gamma^\dagger}\,e^{c\sqrt{2\gamma}A_\gamma} \\
&= e^{-c^2\gamma}\,e^{-c\sqrt{2\gamma}A_\gamma^\dagger}\,H(p+ic\gamma,x)\,e^{c\sqrt{2\gamma}A_\gamma}.
\end{aligned} \tag{122}$$

We then need to evaluate these expressions between the states $|\psi_c(\gamma)\rangle$ and $|\psi_{-c}(\gamma)\rangle$. Clearly, the first two terms are equal to $\mathcal{E}(\lambda, f^2, c, \gamma)_{\text{min}}$, which is the minimum value of $\mathcal{E}(\lambda, f^2, c, \gamma)$ obtained by minimizing with respect to c and γ, since the minima are degenerate.

Given these identities, it is straightforward to evaluate

$$\begin{aligned}
\langle\psi_{\text{even}}|H|\psi_{\text{even}}\rangle = \frac{1}{N_{\text{even}}^2} \quad &[\langle\chi(c,\gamma)|H(p,x)|\chi(c,\gamma)\rangle + \langle\chi(-c,\gamma)|H(p,x)|\chi(-c,\gamma)\rangle \\
&+ \langle\chi(-c,\gamma)|H(p,x)|\chi(c,\gamma)\rangle + \langle\chi(c,\gamma)|H(p,x)|\chi(-c,\gamma)\rangle].
\end{aligned} \tag{123}$$

Clearly, the first two terms evaluate to the function $\mathcal{E}(\lambda, f^2, c, \gamma)$ and the remaining terms are

$$\langle\chi(-c,\gamma)|H(p,x)|\chi(c,\gamma)\rangle = e^{-c^2\gamma}\langle(1 + c\sqrt{2\gamma}A_\gamma)\psi_{-c}(\gamma)|H(p-ic\gamma,x)|(1 - c\sqrt{2\gamma}A_\gamma)\psi_c(\gamma)\rangle \tag{124}$$

and

$$\langle\chi(c,\gamma)|H(p,x)|\chi(-c,\gamma)\rangle = e^{-c^2\gamma}\langle(1 - c\sqrt{2\gamma}A_\gamma)\psi_c(\gamma)||H(p+ic\gamma,x)|(1 + c\sqrt{2\gamma}A_\gamma)\psi_{-c}(\gamma)\rangle. \tag{125}$$

Explicit substitution of the shifted definition of p in H yields

$$H(p - ic\gamma, x) = H(p,x) - ic\gamma p - \frac{c^2}{2}\gamma^2 \tag{126}$$

and

$$H(p + ic\gamma, x) = H(p,x) + ic\gamma p - \frac{c^2}{2}\gamma^2. \tag{127}$$

To relate the two cross-terms it is helpful to observe that

$$\langle\chi(-c,\gamma)|O|\chi(c,\gamma)\rangle = \langle\chi(c,\gamma)|O^\dagger|\chi(-c,\gamma)\rangle^*. \tag{128}$$

Since $i\,p$ is anti-hermitian, we get a minus sign for that term and plus signs for all the others. Note that the expectation value of $H(p,x)$ that appears in the cross terms is obtained from Eq. 29 setting $c = 0$ and evaluating the result between the states $\langle \phi_{-c}(\gamma)|$ and $|\phi_c(\gamma)\rangle$.

The result for the expectation value of the Hamiltonian in the variational states $|\psi_{\text{even}}\rangle$ and $|\psi_{\text{odd}}\rangle$ is:

$$
\begin{aligned}
\langle \psi_{\text{even}}|H|\psi_{\text{even}}\rangle \;=\;& \frac{2}{N_{\text{even}}^2}\left[\mathcal{E}(\lambda, f^2, c, \gamma) + \frac{e^{-c^2\gamma}}{1 + x(c)^2}\left((1 + \sqrt{2\gamma}\,c\,x(c))^2 \left(\frac{\gamma}{4} + \frac{\lambda f^4}{6} + \frac{\lambda}{8\gamma^2} - \frac{\lambda f^2}{6\gamma} - \frac{c^2\gamma^2}{2} \right) \right. \right. \\
&+ \left. \left. 2\,x(c)\,(1 + \sqrt{2\gamma}\,c\,x(c))\,c\,\gamma\,\sqrt{\frac{\gamma}{2}} - x(c)^2\left(\frac{5\lambda}{8\gamma^2} + \frac{3\gamma}{4} - \frac{\lambda f^2}{2\gamma} + \frac{\lambda f^4}{6} - \frac{c^2\gamma^2}{2} \right) \right) \right]
\end{aligned}
\tag{129}
$$

and

$$
\begin{aligned}
\langle \psi_{\text{odd}}|H|\psi_{\text{odd}}\rangle \;=\;& \frac{2}{N_{\text{even}}^2}\left[\mathcal{E}(\lambda, f^2, c, \gamma) - \frac{e^{-c^2\gamma}}{1 + x(c)^2}\left((1 + \sqrt{2\gamma}\,c\,x(c))^2 \left(\frac{\gamma}{4} + \frac{\lambda f^4}{6} + \frac{\lambda}{8\gamma^2} - \frac{\lambda f^2}{6\gamma} - \frac{c^2\gamma^2}{2} \right) \right. \right. \\
&+ \left. \left. 2\,x(c)\,(1 + \sqrt{2\gamma}\,c\,x(c))\,c\,\gamma\,\sqrt{\frac{\gamma}{2}} - x(c)^2\left(\frac{5\lambda}{8\gamma^2} + \frac{3\gamma}{4} - \frac{\lambda f^2}{2\gamma} + \frac{\lambda f^4}{6} - \frac{c^2\gamma^2}{2} \right) \right) \right]
\end{aligned}
\tag{130}
$$

Clearly, so long as the factor $e^{-c^2\gamma}$ is small, the splitting between the even and odd states is quite small. Thus the tunneling between the left and right states is slow. When we go to smaller values of f or larger values of n, the tunneling rate increases, and at $f = 0$, minimizing $\langle \psi_{\text{even}}|H|\psi_{\text{even}}\rangle$ and $\langle \psi_{\text{odd}}|H|\psi_{\text{odd}}\rangle$ with respect to γ and c leads to energies that agree in accuracy with the second order adaptive perturbation theory calculation for the pure anharmonic oscillator, which was done without using the shifted wavefunction. Thus, as expected, including the shift includes the effects of the second order computation.

---

[1] I. G. Halliday and P. Suranyi, Phys. Lett. B **85**, 421 (1979).
[2] I. G. Halliday and P. Suranyi, Phys. Rev. D **21**, 1529 (1980).
[3] V. F. Weisskopf and E. P. Wigner, Z. Phys. **63**, 54 (1930); **65**, 18 (1930).
[4] B. Simon and R. Griffiths,Comm. Math, Phys.,bf 33, 145, (1973)