The quantum 2D-harmonic oscillator in 1:1 resonance with time-dependent perturbation[1]

Averaging applied to slowly varying quantum systems

Richard Huveneers

Department of Mathematics
University of Utrecht, P.O.Box 80.010
3508 TA Utrecht, The Netherlands

PACS numbers : 02.30.Hq, 02.30.Mv, 02.70.Dh, 02.70.Rw
Keywords : oscillator, resonance, perturbation, time-dependent, averaging, adiabatic

[1]This research was supported by NWO-grant nr. 611-306-537

Abstract

In this paper we will study the 2D-harmonic oscillator in 1:1 resonance. We add a general third- and fourth order perturbation which is *slowly* time-dependent, and are interested in the resulting interaction between the unperturbed orbits (states).
By treating this system quantummechanically we get a *linear* system (Schrödinger's equation) on the infinite dimensional space $L^2(\mathbb{R}^2)$. We show that this system can be reduced to a large finite dimensional system by bounding the perturbation suitably at infinity.
By applying perturbation theory and averaging we are able to split this large system into smaller subsystems. We keep track of all error terms and show explicitly under which conditions they can be neglected.
By applying adiabatic perturbation theory we finally transform the positive time-axis into a finite time-interval.
We conclude by indicating how the theory can be extended to other resonances and to the 3D harmonic oscillator.

1 Introduction

1.1 Classical formulation of the problem

Hamilton systems with two degrees of freedom near stable equilibrium points have been studied extensively during the last decades, with the result that most of these systems are well understood both theoretically and numerically. The main tools which are applied in these studies are normal form transformation, center-manifold theory, KAM-theory, averaging, bifurcation analysis and numerics (Guckenheimer and Holmes [1], Arnold [2], Lazutkin [3], Sanders and Verhulst [4] and Hairer et al [5]).
In almost all studies the Hamiltonian is taken either time-independent or time-periodic, mostly to reduce the dimension of phase space by one. It is therefore hardly known what happens when the Hamiltonian is time-dependent in a non-periodic way.

The motivation for this study is that many physical objects exhibit some form of symmetry. Most galaxies for instance, have axes or planes of symmetry. These symmetric equilibrium configurations are generally the outcome of the evolution from an asymmetric state. We would like to trace the effect of the asymmetries.

A problem is that studies of the evolution of actual physical systems are difficult and so relatively rare. We propose therefore to ignore, at least for the time being, the actual physical mechanisms and to consider systems described by a Hamiltonian of the form

$$\mathcal{H}(\vec{p},\vec{q},\epsilon t) = \mathcal{H}_s(\vec{p},\vec{q}) + f(\delta t)\mathcal{H}_a(\vec{p},\vec{q}) \tag{1.1}$$

where $\mathcal{H}_s$ is the part of the Hamiltonian which is symmetric in some sense; $\mathcal{H}_a$ is the asymmetric part which is slowly vanishing as we put

$$f(0)=1,\ f'(t)<0,\ \lim_{t\to\infty} f(t)=0,\ 0<\delta\ll 1 \tag{1.2}$$

To study the dynamics induced by the Hamiltonian (1.1) is still a formidable problem. A first step in this field has been made by Van den Broek [6], who considered a 2D-harmonic oscillator with a third order perturbation which slowly becomes symmetric with respect to one of the coordinates:

$$\begin{aligned}\mathcal{H}(p_x,p_z,x,z) &= \tfrac{1}{2m}(p_x^2+p_z^2) + \tfrac{m}{2}(\omega_1^2x^2+\omega_2^2z^2)\\ &\quad - (\tfrac{a_1}{3}x^3+a_2xz^2) - f(\delta t)(a_3x^2z+\tfrac{a_4}{3}z^3)\end{aligned} \tag{1.3}$$

The space-coordinates are called x and z for historical reasons.
This system was studied in an ε-neighbourhood of the origin of phase space.

The first remark we should make is that the nonresonant case (ω_1 and ω_2 rationally independent) is not very interesting, since normal form theory (which can easily be extended to time-dependent perturbations) gives an accurate approximation of the exact solution (the accuracy depends on how far the system is from low order resonances and on the timescale).
There is only one first order resonance in two degrees of freedom systems, the 1:2 resonance (ω_1:ω_2 = 1:2), to which Van den Broek devoted most of his study.
At second order, there are two resonances, the 1:1 and the 1:3 resonance. In this article we will focus our attention on the 1:1 resonance, so we put $\omega_1=\omega_2=\omega$.
Note that, already in the time-independent case, the calculations for a second order resonance are longer and more difficult than for a first order resonance.

Since the 1:1 resonance is a second order resonance, its effect will appear in the fourth order terms of the Hamiltonian in the normal form calculation. As the perturbation to the 2D-harmonic oscillator

should be regarded as the start of a Taylor-series, we have to add also general fourth order terms, since it makes no sense computing the effect of the third order terms to fourth order while neglecting other fourth order contributions:

$$\begin{aligned} \mathcal{H}(p_x,p_z,x,z) &= \tfrac{1}{2m}(p_x^2+p_z^2) + \tfrac{k}{2}(x^2+z^2) \\ &+ a_{30}x^3+a_{12}xz^2+a_{40}x^4+a_{22}x^2z^2+a_{04}z^4 \\ &+ f(\delta t)\{a_{21}x^2z+a_{03}z^3+a_{31}x^3z+a_{13}xz^3\} \end{aligned} \tag{1.4}$$

Where $k=m\omega^2$ and we have renamed the coefficients in a more transparent way.
We will study system (1.4) from the initial time t=0 up to the final time $t=\delta^{-1}$.
To do this, we will need the following condition on f in our calculations:

$$f(1) \ll 1 \tag{1.5}$$

This condition merely prescribes the scaling of time: At $t=\delta^{-1}$ we want the system to be in it's final state. This puts only a *small* condition on f(t). If this condition is not satisfied, f(t) decays too slowly, so that the system may never reach a final state; It may have interesting dynamics all the way to infinity. In this case one first has to study system (1.4) on a shorter timescale, and then study a simplified system (i.e. with the local dynamics removed) on the long timescale (averaging is a nice example of this).

In writing down system (1.4), we have neglected the fifth- and higher order terms of the Taylor-series of the perturbation. We introduce the parameter ε, being the radius of the 4-dimensional ball around the origin of phasespace within which this is a good approximation of the Hamiltonian. Looking at Hamilton's equations, one can see immediately that fifth order terms can only be neglected if:

$$\frac{1}{k}\max\{|a_{05}|,|a_{14}|,|a_{23}|,|a_{32}|,|a_{41}|,|a_{50}|\}\ \epsilon^3\frac{\omega}{\delta} \ll 1 \tag{1.6}$$

Note that this is a *lower* bound on δ.
If δ does not satisfy this condition, the dynamics take place on such a long timescale that fifth order terms have an order one contribution to it.

An application of system (1.4) can be found in astrophysics and (by the similar structure of Newton's and Coulomb's law) in electrodynamics. By applying certain simple transformations (see for example Verhulst [7]), it is shown that system (1.4) describes the motion of a particle in an axisymmetric potential with a symmetry plane perpendicular to the symmetry axis. With the perturbation we can describe either the influence of a passing neighbouring starsystem, or the influence of the starsystem which generates the potential, becoming symmetric with respect to the z-plane slowly in time (Babin [8] showed for a class of partial differential equations that in this case all solutions will gradually obtain the same symmetry).
Furthermore it should be clear that this theory can be applied to any system with a truncated Taylor-series given by (1.4), i.e. a 2D-harmonic oscillator in 1:1 resonance with a perturbation which becomes symmetric with respect to one variable slowly in time.

1.2 Quantum formulation of the problem

We now turn to the quantummechanical treatment of this system. We will do this by substituting $\frac{\hbar}{i}\vec{\nabla}$ for $\vec{p}$ in system (1.4) and considering Schrödinger's equation.
The correspondence between classical- and quantum systems has not been established satisfactorily (see for instance Gutzwiller [9]). However, here we assume that this transition makes sense.
Apart from this, Schrödinger's equation with a time-dependent potential constitutes an interesting class of quantum systems in itself.

Possible applications include the passing of a big and a small atom, an asymmetric charge distribution inside the core of an atom, the motion of 1 electron in a multi-electron atom, and a slowly varying external forcefield applied to an atom.
At this point we must also refer to another application of the theory of averaging to quantum mechanics given by Sanders [13].

To make all observables dimensionless and to have terms of different order show up explicitly, we perform the following transformation, after which we will silently omit the bars again:

$$\begin{aligned} x &= \frac{\bar{x}}{\alpha} \\ z &= \frac{\bar{z}}{\alpha} \\ t &= \bar{t}\,\frac{2\pi}{\omega} \\ \delta &= \bar{\delta}\,\frac{\omega}{2\pi} \\ a_{ij} &= \bar{a}_{ij}\,\frac{k}{\epsilon^{i+j-2}} \end{aligned} \tag{1.7}$$

with $\alpha = \left(\frac{mk}{\hbar^2}\right)^{\frac{1}{4}}$.

So we will consider Schrödinger's equation

$$i\frac{1}{2\pi}\frac{\partial}{\partial t}\Psi(x,z,t) = \mathcal{H}\,\Psi(x,z,t) \tag{1.8}$$

with

$$\begin{aligned} \mathcal{H} &= \mathcal{H}^{(0)} + \frac{1}{\alpha\epsilon}\mathcal{H}^{(1)} + \frac{1}{\alpha^2\epsilon^2}\mathcal{H}^{(2)} \\ \mathcal{H}^{(0)} &= -\frac{1}{2}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}\right) + \frac{1}{2}(x^2+z^2) \\ \mathcal{H}^{(1)} &= a_{30}x^3 + a_{12}xz^2 + f(\delta t)\left\{a_{21}x^2z + a_{03}z^3\right\} \\ \mathcal{H}^{(2)} &= a_{40}x^4 + a_{22}x^2z^2 + a_{04}z^4 + f(\delta t)\left\{a_{31}x^3z + a_{13}xz^3\right\} \end{aligned} \tag{1.9}$$

where we have grouped terms of order m+2 in $\mathcal{H}^{(m)}$.
We immediately notice that we now have to consider a linear partial differential equation instead of a nonlinear ordinary differential equation. Another direct consequence of quantum mechanics is that we have a natural *lower*-bound on ε, since it makes no sense considering system (1.9) in a neighbourhood of the origin lying within the groundstate of $\mathcal{H}^{(0)}$. We therefore have

$$\epsilon \gg \frac{1}{\alpha} \tag{1.10}$$

This shows explicitly that the factor $(\alpha\varepsilon)^{-1}$ which we use to distinguish terms of different order, is indeed much smaller than 1.
In the realistic case that ε lies within the range of validity of Hamilton's equations we have $\alpha\varepsilon \approx 10^{15}$.

1.3 Discussion of the problem

It's clear that the quantummechanical treatment has some calculational advantages above the classical treatment, like the linearity of Schrödinger's equation and (as we will show) the separate treatment of the two small parameters $(\alpha\varepsilon)^{-1}$ and δ. Of course there are also a few (serious) complications, like that we now have two solve a partial differential equation on the infinite dimensional space $L^2(\mathbb{R}^2)$. Moreover, the unbounded behaviour of the perturbation at infinity introduces some serious methematical problems (as we will show in the next section).
It turns out that a proper bounding of the perturbation at infinity solves all three problems at once, in the sense that we end up with a linear system of coupled ordinary differential equations with a large, but finite, dimension.
In order to do any computations, we have to reduce the dimension further. By applying perturbation theory (Merzbacher [10]) and averaging we are able to split the large system into smaller subsystems. We keep track of all error terms and show explicitly under which conditions they can be neglected.
The resulting subsystems are small enough to be studied systematically. Up to this stage we did not make any use of the slow time-dependence of the perturbation. By applying adiabatic perturbation theory we are able to transform the positive time-axis into a finite time-interval. From the final system we can draw some important conclusions.

2 Bounding the perturbation

In this section we will solve the just mentioned problems regarding the behaviour of the third- and fourth order terms at infinity, and the infinite dimension of system (1.9).
The first problem arises because a quantummechanical particle senses the potential at classically forbidden regions. In the present system the third- and fourth order terms generically cause directions in the x-z plane, in which the potential tends to minus infinity, with the inevitable result that the particle will make the one-way trip to infinity *even* if it has to pass classically forbidden regions (this phenomenon is known as quantum tunnelling).
We stress that this is only a mathematical problem, since this tunnelling effect takes place on a much larger time-scale than the one on which the asymmetric terms in the perturbation vanish. If we would not put a bound on the perturbation, our expressions would get a lot longer, and we would not be able to let time tend to infinity.
So, regarding the third- and fourth order terms as a perturbation to the 2D-harmonic oscillator, we will bound the behaviour of the perturbation towards infinity. As far as we can see, there are three ways of doing so, the third and least obvious of which turns out to be most useful. All three ways are based on the idea that since the third- and fourth order terms should be regarded as terms of a Taylor-series, their value is only meaningful in a neighbourhood of the origin. Therefore, we are allowed to change their value away from the origin by which we can bound them towards infinity. Although the three ways of bounding the behaviour of the perturbation towards infinity are physically equivalent, they differ largely in the forthcoming calculations.
The most obvious way is to truncate the perturbation at a distance $\alpha\varepsilon$ from the origin. This means that we would set the perturbation equal to zero for $x^2+z^2>\alpha^2\varepsilon^2$.
Using this we would have to deal with error-functions everywhere in our calculations, which would be extremely messy.
A better way is to bound the perturbation by multiplying it with a factor $\exp(-(x^2+z^2)/(\alpha^2\varepsilon^2))$, which would clearly (just expand the exponential) not effect the third- and fourth order terms.
Using this, we would have to deal with finite series in a very small parameter, which would allow us to proceed our calculations further than when using the first approach. After some steps however the expressions again become very large and far from transparent.
The third and calculationally the best way, is to bound the perturbation by its action on the eigenfunctions of the unperturbed system:

$$\psi^{(0)}_{nm}(x,z) = \frac{1}{\sqrt{\pi 2^n (n-m)!m!}}\, e^{-\frac{1}{2}(x^2+z^2)}\, H_m(x)\, H_{n-m}(z) \qquad (2.1)$$

with n=0,1,2,... and m=0,1,...,n and where H_m are the Hermite polynomials.

The corresponding eigenvalues are given by $E_n^{(0)} = n+\frac{1}{2}$. So the index m denotes the degeneracy; the energy level $E_n^{(0)}$ has multiplicity n+1.

The perturbation is now bounded by setting its action on all high unperturbed eigenfunctions equal to zero. In other words, we take a high cutoff energy level N_{cut} (typically $(\alpha\varepsilon)^2$) and make the perturbation kill all higher energy levels, i.e. replace $\mathcal{H}^{(1)}$ by $\bar{\mathcal{H}}^{(1)}$ whose action on a wavefunction Ψ is defined by:

$$\bar{\mathcal{H}}^{(1)}\Psi = P_{N_{cut}}\mathcal{H}^{(1)}P_{N_{cut}}\Psi \tag{2.2}$$

where the linear projection operator $P_{N_{cut}}$ is defined by:

$$P_{N_{cut}}\psi_{nm}^{(0)} = \begin{cases} \psi_{nm}^{(0)} & \text{for } n \le N_{cut} \\ 0 & \text{for } n > N_{cut} \end{cases} \tag{2.3}$$

This truncation scheme is justified by the fact that higer order eigenfunctions are concentrated further away from the origin (more quantitatively, $\left\langle \psi_{nm}^{(0)} \middle| x^2+z^2 \middle| \psi_{nm}^{(0)} \right\rangle \sim n$). The projection operator $P_{N_{cut}}$ kills any interaction with these high order eigenfunctions.
Note that these states will still be propagated, since the projection operator is not applied to the unperturbed Hamiltonian. Variants of this projection method have been successfully applied to many infinite dimensional problems.

Similarly, we replace $\mathcal{H}^{(2)}$ by $\bar{\mathcal{H}}^{(2)}$.
Note that $\bar{\mathcal{H}}^{(1)}$ and $\bar{\mathcal{H}}^{(2)}$ can not be represented by a function in coordinatespace.
We will soon show that this way of bounding the perturbation has great calculational advantages. Therefore, we will use it in the rest of this article. From now on we will omit the bars on the perturbation, which is implicitly understood to be bounded in this way. This will not cause any confusion since the interesting dynamics takes place in a subspace on which the action of the bounded and the unbounded perturbation are equal, as we will show.

The relation between the cutoff size N_{cut} and the initial value problem (1.4) is determined by the *range* of the unperturbed eigenfunctions (2.1). Using (3.2) one easily verifies that the expectation value of x^2+z^2 in the eigenstate $\psi_{nm}^{(0)}(x,z)$ is proportional with n. So in order to bound the perturbation at a distance $\alpha\varepsilon$ from the origin (i.e. O(1) in the original variables) N_{cut} will typically be of order $(\alpha\varepsilon)^2$.

Let us demonstrate the idea with an example.
For simplicity consider the 1D-harmonic oscillator

$$\left(-\frac{1}{2}\frac{\partial^2}{\partial x^2} + \tfrac{1}{2}x^2 - E\right)\psi(x) = 0 \tag{2.4}$$

with perturbation x^3. We applied the unbounded- and bounded perturbation (with N_{cut}=15) to the wavepacket $\psi(x)=\frac{1}{1+x^2}$. This gives the following picture:

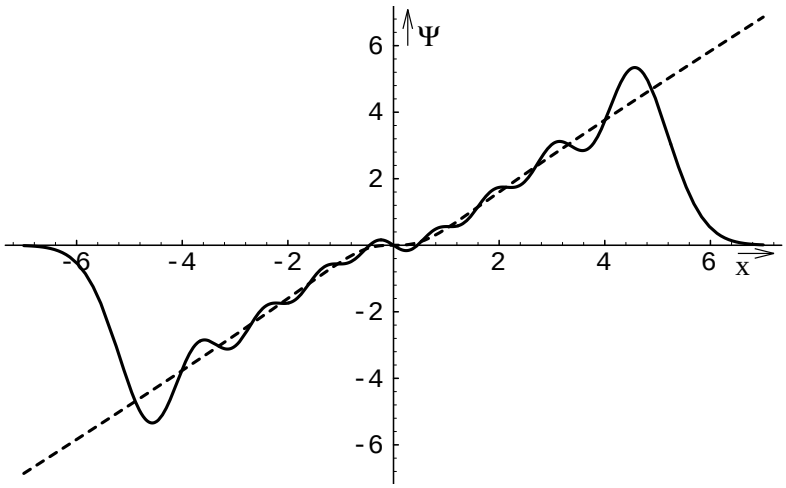

Figure 2.1: The action of the bounded and the unbounded perturbation

The dashed (solid) line represents the unboundedly (respectively boundedly) perturbed wavepacket.
We see that the bounded perturbation indeed approximates the unbounded perturbation in the neighbourhood of the origin. In practice the approximation will be much better since $\alpha\varepsilon$ and thus N_{cut} is very large.

Another big advantage of bounding the perturbation in this way, is that it solves our second problem (the infinite dimension of system (1.9)) almost trivially, because the *unperturbed* eigenfunctions (2.1) with $n>N_{cut}$ are now also eigenfunctions of the *perturbed bounded* Hamiltonian with the same eigenvalues. Moreover, the space spanned by the unperturbed eigenfunctions with $n\leq N_{cut}$ is invariant under the perturbed bounded Hamiltonian.
Combining these two statements we obtain that expanding the exact solution to Schrödinger's equation in the unperturbed eigenfunctions reduces the dimension of system (1.9) from infinite to $\frac{1}{2}(N_{cut}+1)(N_{cut}+2)$. We will show this explicitly in the next section.
So we have solved system (1.9) explicitly for "almost all" dimensions.
We call these dimensions the trivial dimensions of the system.

Note however that the remaining number of dimensions is still very large, though *finite*.
We are left with a $\frac{1}{2}(N_{cut}+1)(N_{cut}+2)$-dimensional *linear* system of *ordinary* differential equations which is *independent* of the infinite number of trivial dimensions.
Note that although the remaining number of dimensions is now *finite*, it is still too large to do any calculations. We will show that the reduced system can be split up in small, independent subsystems.

3 Splitting in subsystems: the concept

The physical idea behind the splitting in subsystems is much simpler than the mathematical proof. The idea is this:
For each t we can calculate the "instantaneous" energy spectrum of the Hamiltonian. Since the applied perturbation is small, the energy spectrum will show great similarity with the energy spectrum of the unperturbed 2D-harmonic oscillator, i.e. the instantaneous energy-levels will lie clustered around the unperturbed energy levels $E_n^{(0)}$, with the number of levels clustered around an unperturbed level being equal to its degeneracy (n+1).
Since our Hamiltonian depends slowly on time, the energy levels will change also *slowly* in time, with the result that no transitions are allowed between two different clusters simply because the energy gap is too big. However, transitions within a cluster are allowed, because these energy levels lie much closer to each other.

The result is that the dynamics within each cluster evolves independently of the other clusters, thus splitting up our system drastically.
The reader might sense some trouble in the case that the perturbation does not remove the degeneracy completely, i.e. if two instantaneous energy levels are (almost) equal, since in that case the transition rate between those two levels can become very large, resulting in very rapid dynamics. In the sequel we shall carry out the underlying mathematical analysis. The degeneration problem will be dealt with in section 7.

In our calculations we will need the matrix elements of the perturbation with respect to the unperturbed basis. These matrix elements can be computed efficiently using the following algorithm:

Define $ME(p,m_1,m_2) = \left\langle \psi^{(0)}_{m_2}(x) \middle| x^p \middle| \psi^{(0)}_{m_1}(x) \right\rangle$, for all nonnegative integers m_1, m_2 and p, where the usual Dirac-bracket notation is employed, and the $\psi^{(0)}_m$ are the eigenfunctions of the 1D-harmonic oscillator:

$$\psi^{(0)}_m(x) = \sqrt{\frac{1}{\sqrt{\pi}2^m m!}} e^{-\frac{1}{2}x^2} H_m(x) \tag{3.1}$$

Using the well-known recursion relations for the Hermite-polynomials, one verifies the recursion relation for ME:

$$ME(p,m_1,m_2) = \sqrt{\frac{m_1}{2}} ME(p-1,m_1-1,m_2) + \sqrt{\frac{m_1+1}{2}} ME(p-1,m_1+1,m_2) \tag{3.2}$$

where the first term vanishes for m_1=0.
Adding the trivial initial value $ME(0,m_1,m_2) = \delta_{m_1 m_2}$, we have determined all $ME(p,m_1,m_2)$.

Since $\psi^{(0)}_{nm}(x,z) = \psi^{(0)}_m(x)\psi^{(0)}_{n-m}(z)$, the matrix elements of $\mathcal{H}^{(0)}$, $\mathcal{H}^{(1)}$ and $\mathcal{H}^{(2)}$ are readily calculated.
Note that the simplicity of the expressions for the matrix elements is again due to the particular way of bounding the perturbation.

4 Calculation of the matrix elements

We introduce the following straightforward notation:

$$\begin{aligned}
\mathcal{H}^{(0)}_{n_1 m_1, n_2 m_2} &= \left\langle \psi^{(0)}_{n_2 m_2} \middle| \mathcal{H}^{(0)} \middle| \psi^{(0)}_{n_1 m_1} \right\rangle \\
\mathcal{H}^{(1)}_{n_1 m_1, n_2 m_2} &= \left\langle \psi^{(0)}_{n_2 m_2} \middle| \mathcal{H}^{(1)} \middle| \psi^{(0)}_{n_1 m_1} \right\rangle \\
\mathcal{H}^{(2)}_{n_1 m_1, n_2 m_2} &= \left\langle \psi^{(0)}_{n_2 m_2} \middle| \mathcal{H}^{(2)} \middle| \psi^{(0)}_{n_1 m_1} \right\rangle
\end{aligned} \tag{4.1}$$

Explicit calculation, using the recursion formula, gives the following expressions (to avoid errors, this calculation and most of the calculations that will follow were carried out by Mathematica on a Sun workstation (typical calculation time: 15 minutes)):

$$\mathcal{H}^{(0)}_{n_1 m_1, n_2 m_2} = (n_1+1)\delta_{n_1 n_2}\delta_{m_1 m_2} \tag{4.2}$$

$$
\begin{aligned}
\mathcal{H}^{(1)}_{n_1 m_1, n_2 m_2} * 2^{\frac{3}{2}} &= a_{30}\sqrt{(-2+m_2)(-1+m_2)m_2}\ \delta_{m_1,-3+m_2}\ \delta_{n_1,-3+n_2} \\
&+ a_{21}\sqrt{(-1+m_2)m_2(-m_2+n_2)}\ f(\delta t)\ \delta_{m_1,-2+m_2}\ \delta_{n_1,-3+n_2} \\
&+ a_{21}\sqrt{(-1+m_2)m_2(1-m_2+n_2)}\ f(\delta t)\ \delta_{m_1,-2+m_2}\ \delta_{n_1,-1+n_2} \\
&+ a_{12}\sqrt{m_2(-1-m_2+n_2)(-m_2+n_2)}\ \delta_{m_1,-1+m_2}\ \delta_{n_1,-3+n_2} \\
&+ \sqrt{m_2}\ (a_{12}-2a_{12}m_2+3a_{30}m_2+2a_{12}n_2)\ \delta_{m_1,-1+m_2}\ \delta_{n_1,-1+n_2} \\
&+ a_{12}\sqrt{m_2(1-m_2+n_2)(2-m_2+n_2)}\ \delta_{m_1,-1+m_2}\ \delta_{n_1,1+n_2} \\
&+ a_{03}\sqrt{(-2-m_2+n_2)(-1-m_2+n_2)(-m_2+n_2)}\ f(\delta t)\ \delta_{m_1,m_2}\ \delta_{n_1,-3+n_2} \\
&+ \sqrt{-m_2+n_2}\ (a_{21}-3a_{03}m_2+2a_{21}m_2+3a_{03}n_2)\ f(\delta t)\ \delta_{m_1,m_2}\ \delta_{n_1,-1+n_2} \\
&+ \sqrt{1-m_2+n_2}\ (a_{21}+2a_{21}m_2+3a_{03}(1-m_2+n_2))\ f(\delta t)\ \delta_{m_1,m_2}\ \delta_{n_1,1+n_2} \\
&+ a_{03}\sqrt{(1-m_2+n_2)(2-m_2+n_2)(3-m_2+n_2)}\ f(\delta t)\ \delta_{m_1,m_2}\ \delta_{n_1,3+n_2} \\
&+ a_{12}\sqrt{(1+m_2)(-1-m_2+n_2)(-m_2+n_2)}\ \delta_{m_1,1+m_2}\ \delta_{n_1,-1+n_2} \\
&+ \sqrt{1+m_2}\ (a_{12}+3a_{30}-2a_{12}m_2+3a_{30}m_2+2a_{12}n_2)\ \delta_{m_1,1+m_2}\ \delta_{n_1,1+n_2} \\
&+ a_{12}\sqrt{(1+m_2)(1-m_2+n_2)(2-m_2+n_2)}\ \delta_{m_1,1+m_2}\ \delta_{n_1,3+n_2} \\
&+ a_{21}\sqrt{(1+m_2)(2+m_2)(-m_2+n_2)}\ f(\delta t)\ \delta_{m_1,2+m_2}\ \delta_{n_1,1+n_2} \\
&+ a_{21}\sqrt{(1+m_2)(2+m_2)(1-m_2+n_2)}\ f(\delta t)\ \delta_{m_1,2+m_2}\ \delta_{n_1,3+n_2} \\
&+ a_{30}\sqrt{(1+m_2)(2+m_2)(3+m_2)}\ \delta_{m_1,3+m_2}\ \delta_{n_1,3+n_2}
\end{aligned}
\tag{4.3}
$$

if n_1 and n_2 are both smaller than or equal to N_{cut}, otherwise the matrix element is zero.

$$
\begin{aligned}
\mathcal{H}^{(2)}_{n_1 m_1, n_2 m_2} \;*\; 4 \;=\;& a_{40}\sqrt{(-3+m_2)(-2+m_2)(-1+m_2)m_2}\;\delta_{m_1,-4+m_2}\;\delta_{n_1,-4+n_2} \\
&+ a_{31}\sqrt{(-2+m_2)(-1+m_2)m_2(-m_2+n_2)}\;f(\delta t)\;\delta_{m_1,-3+m_2}\;\delta_{n_1,-4+n_2} \\
&+ a_{31}\sqrt{(-2+m_2)(-1+m_2)m_2(1-m_2+n_2)}\;f(\delta t)\;\delta_{m_1,-3+m_2}\;\delta_{n_1,-2+n_2} \\
&+ a_{22}\sqrt{(-1+m_2)m_2(-1-m_2+n_2)(-m_2+n_2)}\;\delta_{m_1,-2+m_2}\;\delta_{n_1,-4+n_2} \\
&+ \sqrt{(-1+m_2)m_2}\;(a_{22}-2a_{40}-2a_{22}m_2+4a_{40}m_2+2a_{22}n_2)\;\delta_{m_1,-2+m_2}\;\delta_{n_1,-2+n_2} \\
&+ a_{22}\sqrt{(-1+m_2)m_2(1-m_2+n_2)(2-m_2+n_2)}\;\delta_{m_1,-2+m_2}\;\delta_{n_1,n_2} \\
&+ a_{13}\sqrt{m_2(-2-m_2+n_2)(-1-m_2+n_2)(-m_2+n_2)}\;f(\delta t)\;\delta_{m_1,-1+m_2}\;\delta_{n_1,-4+n_2} \\
&+ 3\sqrt{m_2(-m_2+n_2)}\;(-a_{13}m_2+a_{31}m_2+a_{13}n_2)\;f(\delta t)\;\delta_{m_1,-1+m_2}\;\delta_{n_1,-2+n_2} \\
&+ 3\sqrt{m_2(1-m_2+n_2)}\;(a_{31}m_2+a_{13}(1-m_2+n_2))\;f(\delta t)\;\delta_{m_1,-1+m_2}\;\delta_{n_1,n_2} \\
&+ a_{13}\sqrt{m_2(1-m_2+n_2)(2-m_2+n_2)(3-m_2+n_2)}\;f(\delta t)\;\delta_{m_1,-1+m_2}\;\delta_{n_1,2+n_2} \\
&+ a_{04}\sqrt{(-3-m_2+n_2)(-2-m_2+n_2)(-1-m_2+n_2)(-m_2+n_2)}\;\delta_{m_1,m_2}\;\delta_{n_1,-4+n_2} \\
&+ \sqrt{(-1-m_2+n_2)(-m_2+n_2)}\;(-2a_{04}+a_{22}-4a_{04}m_2+2a_{22}m_2+4a_{04}n_2)\;\delta_{m_1,m_2}\;\delta_{n_1,-2+n_2} \\
&+ (3a_{04}+a_{22}+3a_{40}-6a_{04}m_2+6a_{40}m_2+6a_{04}m_2^2-4a_{22}m_2^2+6a_{40}m_2^2+6a_{04}n_2 \\
&\quad +2a_{22}n_2-12a_{04}m_2n_2+4a_{22}m_2n_2+6a_{04}n_2^2)\;\delta_{m_1,m_2}\;\delta_{n_1,n_2} \\
&+ \sqrt{(1-m_2+n_2)(2-m_2+n_2)}\;(6a_{04}+a_{22}-4a_{04}m_2+2a_{22}m_2+4a_{04}n_2)\;\delta_{m_1,m_2}\;\delta_{n_1,2+n_2} \\
&+ a_{04}\sqrt{(1-m_2+n_2)(2-m_2+n_2)(3-m_2+n_2)(4-m_2+n_2)}\;\delta_{m_1,m_2}\;\delta_{n_1,4+n_2} \\
&+ a_{13}\sqrt{(1+m_2)(-2-m_2+n_2)(-1-m_2+n_2)(-m_2+n_2)}\;f(\delta t)\;\delta_{m_1,1+m_2}\;\delta_{n_1,-2+n_2} \\
&+ 3\sqrt{(1+m_2)(-m_2+n_2)}\;(a_{31}-a_{13}m_2+a_{31}m_2+a_{13}n_2)\;f(\delta t)\;\delta_{m_1,1+m_2}\;\delta_{n_1,n_2} \\
&+ 3\sqrt{(1+m_2)(1-m_2+n_2)}\;(a_{13}+a_{31}-a_{13}m_2+a_{31}m_2+a_{13}n_2)\;f(\delta t)\;\delta_{m_1,1+m_2}\;\delta_{n_1,2+n_2} \\
&+ a_{13}\sqrt{(1+m_2)(1-m_2+n_2)(2-m_2+n_2)(3-m_2+n_2)}\;f(\delta t)\;\delta_{m_1,1+m_2}\;\delta_{n_1,4+n_2} \\
&+ a_{22}\sqrt{(1+m_2)(2+m_2)(-1-m_2+n_2)(-m_2+n_2)}\;\delta_{m_1,2+m_2}\;\delta_{n_1,n_2} \\
&+ \sqrt{(1+m_2)(2+m_2)}\;(a_{22}+6a_{40}-2a_{22}m_2+4a_{40}m_2+2a_{22}n_2)\;\delta_{m_1,2+m_2}\;\delta_{n_1,2+n_2} \\
&+ a_{22}\sqrt{(1+m_2)(2+m_2)(1-m_2+n_2)(2-m_2+n_2)}\;\delta_{m_1,2+m_2}\;\delta_{n_1,4+n_2} \\
&+ a_{31}\sqrt{(1+m_2)(2+m_2)(3+m_2)(-m_2+n_2)}\;f(\delta t)\;\delta_{m_1,3+m_2}\;\delta_{n_1,2+n_2} \\
&+ a_{31}\sqrt{(1+m_2)(2+m_2)(3+m_2)(1-m_2+n_2)}\;f(\delta t)\;\delta_{m_1,3+m_2}\;\delta_{n_1,4+n_2} \\
&+ a_{40}\sqrt{(1+m_2)(2+m_2)(3+m_2)(4+m_2)}\;\delta_{m_1,4+m_2}\;\delta_{n_1,4+n_2}
\end{aligned}
\tag{4.4}
$$

if n_1 and n_2 are both smaller than or equal to N_{cut}, otherwise the matrix element is zero.

Four very important conclusions (for making the error estimates which will show up soon) can be drawn directly from these formulas:

$\mathcal{H}^{(1)}_{n_1 m_1, n_2 m_2}$ is zero if $n_1 - n_2 \notin \{-3,-1,1,3\}$ or $n_1 > N_{cut}$ or $n_2 > N_{cut}$.

$\mathcal{H}^{(2)}_{n_1 m_1, n_2 m_2}$ is zero if $n_1 - n_2 \notin \{-4,-2,0,2,4\}$ or $n_1 > N_{cut}$ or $n_2 > N_{cut}$.

For each pair (n_1,m_1) there are exactly 16 pairs (n_2,m_2) such that $\mathcal{H}^{(1)}_{n_1 m_1, n_2 m_2}$ is not zero.

For each pair (n_1,m_1) there are exactly 25 pairs (n_2,m_2) such that $\mathcal{H}^{(2)}_{n_1 m_1, n_2 m_2}$ is not zero.

Expanding the solution $\Psi(x,z,t)$ to Schrödinger's equation in the $\psi^{(0)}_{nm}$,

$$\Psi(x,z,t) = \sum_{n=0}^{\infty}\sum_{m=0}^{n} C_{nm}(t)\psi^{(0)}_{nm}(x,z) \tag{4.6}$$

and substituting this expansion into Schrödinger's equation, and taking the inner product with $\psi^{(0)}_{n_1m_1}$, we get

$$i\frac{1}{2\pi}\dot{C}_{n_1m_1} = (n_1+1)C_{n_1m_1} + \frac{1}{\alpha\epsilon}\sum_{n_2=0}^{\infty}\sum_{m_2=0}^{n_2} C_{n_2m_2}\left(\mathcal{H}^{(1)}_{n_1m_1,n_2m_2} + \frac{1}{\alpha\epsilon}\mathcal{H}^{(2)}_{n_1m_1,n_2m_2}\right) \tag{4.7}$$

So if $n_1>N_{cut}$ we get $\dot{C}_{n_1m_1} = -2\pi i(n_1+1)C_{n_1m_1} \Rightarrow C_{n_1m_1}(t) = C_{n_1m_1}(0)e^{-2\pi i(n_1+1)t}$. Moreover, the coefficients $C_{n_1m_1}$ with $n_1>N_{cut}$ do not appear in the coupled differential equations for the coefficients $C_{n_1m_1}$ with $n_1\leq N_{cut}$.
This demonstrates explicitly the finit dimensionality of our system. Thanks to our way of bounding the perturbation we are able to solve explicitly all (infinitely many) other dimensions.
So we are left with a ½$(N_{cut}+1)(N_{cut}+2)$-dimensional coupled system of linear ordinary differential equations.

5 Proof of validity of splitting in subsystems

We will now show that this system can be split up in smaller subsystems. The set of all subsystems is *not* entirely equivalent to the full system, but we will show that the introduced error remains small on the long timescale δ^{-1}.
To do this, we write the matrix representation of the Hamiltonian with respect to the unperturbed eigenfunctions schematically as

$$\mathcal{H} = \begin{pmatrix}
\mathcal{H}^0 & * & * & * & * & & & & & & & & \\
* & \mathcal{H}^1 & * & * & * & * & & & & & & & \\
* & * & \mathcal{H}^2 & * & * & * & * & & & & & & \\
* & * & * & \mathcal{H}^3 & * & * & * & * & & & \varnothing & & \\
* & * & * & * & \mathcal{H}^4 & * & * & * & * & & & & \\
 & * & * & * & * & \mathcal{H}^5 & * & * & * & * & & & \\
 & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & & \\
 & & & * & * & * & * & \mathcal{H}^{N_{cut}-5} & * & * & * & * & \\
 & & & & * & * & * & * & \mathcal{H}^{N_{cut}-4} & * & * & * & * \\
 & & \varnothing & & & * & * & * & * & \mathcal{H}^{N_{cut}-3} & * & * & * \\
 & & & & & & * & * & * & * & \mathcal{H}^{N_{cut}-2} & * & * \\
 & & & & & & & * & * & * & * & \mathcal{H}^{N_{cut}-1} & * \\
 & & & & & & & & * & * & * & * & \mathcal{H}^{N_{cut}}
\end{pmatrix} \tag{5.1}$$

In this representation the $\mathcal{H}^n$ are composed of the contribution of $\mathcal{H}^{(0)}$ restricted to the subspace spanned by the $\psi^{(0)}_{nm}$ (m=0,1,..,n), and the *much* smaller contribution of $\mathcal{H}^{(2)}$ (because of the factor $(\alpha\varepsilon)^{-2}$) and are thus (n+1)x(n+1) (almost) diagonal matrices with n+1 on the diagonal.
The stars (*) represent the other contributions of $\mathcal{H}^{(1)}$ and $\mathcal{H}^{(2)}$.

Using the same schematic representation, we claim that the solution of Schrödinger's equation will hardly change if we replace $\mathcal{H}$ by

$$\bar{\mathcal{H}} = \frac{1}{\alpha^2\epsilon^2}\begin{pmatrix} M^0 & & & & & \\ & M^1 & & & \varnothing & \\ & & M^2 & & & \\ & & & \ddots & & \\ & & & & M^{N_{cut}-2} & \\ & \varnothing & & & & M^{N_{cut}-1} \\ & & & & & & M^{N_{cut}} \end{pmatrix} \tag{5.2}$$

where M^n is again an (n+1)x(n+1) matrix whose elements are given by

$$M^n_{m_1 m_2} = \mathcal{H}^{(2)}_{nm_1,nm_2} + \sum_{\substack{n_3=0 \\ n_3\neq n}}^{\infty} \sum_{m_3=0}^{n_3} \mathcal{H}^{(1)}_{nm_1,n_3m_3} \mathcal{H}^{(1)}_{n_3m_3,nm_2} \frac{1}{n-n_3} \tag{5.3}$$

This replacement is what we call *splitting in subsystems*, since the Hamiltonian has become block diagonal.

Remark 1: Using (4.3) and (4.4) we can calculate the above sum explicitly, yielding

$$\begin{aligned}
&M^n_{m_1m_2} * 24 \\
&= 6\sqrt{(-1+m_2)m_2(1-m_2+n)(2-m_2+n)}\ (-2a_{12}^2+a_{22}+a_{12}a_{30}+(a_{03}a_{21}-2a_{21}^2)f(\delta t)^2)\ \delta_{m_1,-2+m_2} \\
&+ 2\sqrt{m_2(1-m_2+n)}\ (-30a_{03}a_{12}+9a_{13}-10a_{12}a_{21}+30a_{03}a_{12}m_2-9a_{13}m_2-30a_{21}a_{30}m_2 \\
&\quad +9a_{31}m_2-30a_{03}a_{12}n+9a_{13}n-10a_{12}a_{21}n)\ f(\delta t)\ \delta_{m_1,-1+m_2} \\
&+ (18a_{04}-5a_{12}^2+6a_{22}-18a_{12}a_{30}-33a_{30}^2+18a_{40}-36a_{04}m_2+10a_{12}^2m_2-90a_{30}^2m_2+36a_{40}m_2 \\
&\quad +36a_{04}m_2^2+6a_{12}^2m_2^2-24a_{22}m_2^2+72a_{12}a_{30}m_2^2-90a_{30}^2m_2^2+36a_{40}m_2^2+36a_{04}n-18a_{12}^2n \\
&\quad +12a_{22}n-36a_{12}a_{30}n-72a_{04}m_2n+4a_{12}^2m_2n+24a_{22}m_2n-72a_{12}a_{30}m_2n+36a_{04}n^2-10a_{12}^2n^2 \\
&\quad +(-33a_{03}^2-18a_{03}a_{21}-5a_{21}^2+90a_{03}^2m_2-10a_{21}^2m_2-90a_{03}^2m_2^2+72a_{03}a_{21}m_2^2+6a_{21}^2m_2^2-90a_{03}^2n \\
&\qquad -36a_{03}a_{21}n-8a_{21}^2n+180a_{03}^2m_2n-72a_{03}a_{21}m_2n-16a_{21}^2m_2n-90a_{03}^2n^2)\ f(\delta t)^2)\ \delta_{m_1,m_2} \\
&+ 2\sqrt{(1+m_2)(-m_2+n)}\ (-10a_{12}a_{21}-30a_{21}a_{30}+9a_{31}+30a_{03}a_{12}m_2-9a_{13}m_2-30a_{21}a_{30}m_2 \\
&\quad +9a_{31}m_2-30a_{03}a_{12}n+9a_{13}n-10a_{12}a_{21}n)\ f(\delta t)\ \delta_{m_1,1+m_2} \\
&+ 6\sqrt{(1+m_2)(2+m_2)(-1-m_2+n)(-m_2+n)}\ (-2a_{12}^2+a_{22}+a_{12}a_{30}+(a_{03}a_{21}-2a_{21}^2)f(\delta t)^2)\ \delta_{m_1,2+m_2}
\end{aligned} \tag{5.4}$$

Remark 2: Readers familiar with second order averaging (see Sanders and Verhulst [4]) will recognize the structure of this expression. Indeed, a large part of the proof is devoted to second order averaging, *while keeping track of the full error terms*.
Unfortunately, first order averaging gives only trivial results. This is a direct consequence of the unperturbed harmonic oscillator being in 1:1 resonance, in agreement with the classical result (Van den Broek [6]), so we have to do all these computations.

Remark 3: In our discussion of bounding the perturbation, we used bars to distinguish bounded from unbounded operators. Then we dropped the bars again, implicitly assuming all operators to be bounded. So there should be no confusion by our usage of the bar in a totally different context.

We will prove this claim by approximating both the solution of Schrödinger's equation in matrix form:

$$i\frac{1}{2\pi}\frac{\partial}{\partial t}C = \mathcal{H}C,\ C(0) = C_0 \tag{5.5}$$

and the solution of Schrödinger's equation with $\mathcal{H}$ replaced by $\bar{\mathcal{H}}$:

$$i\frac{1}{2\pi}\frac{\partial}{\partial t}\bar{C} = \bar{\mathcal{H}}\bar{C},\ \bar{C}(0) = C(0) = C_0 \tag{5.6}$$

where $C(t)$ and $\bar{C}(t)$ are ½(N_{cut}+1)(N_{cut}+2)-dimensional vectors.

We will show that the difference $C-\bar{C}$ remains small up to t=δ^{-1}, by calculating an upper bound on the introduced error.

We stress that we really have to calculate such an upper bound, since we have two small parameters in our system, $(\alpha\varepsilon)^{-1}$ and δ. An error term of the form $O((\alpha\varepsilon)^{-1})$ is not very meaningful since we now by (1.6) that the error will become infinitely large if we let δ tend to zero. An extra complication arises because some of the constants that occur in the calculations are rather large. In order to be able to distinguish the "small" parameter ranges of $(\alpha\varepsilon)^{-1}$ and δ, we need an explicit upper bound on the introduced error.

The approximation of C and $\bar{C}$ up to t=δ^{-1} is made in timesteps $\Delta t=1$, i.e. we first approximate $C(1)$ and $\bar{C}(1)$ with a certain error, then $C(2)$ and $\bar{C}(2)$ using $C(1)$, $\bar{C}(1)$ and the previous error, and so on up to t=δ^{-1}.

At this point we want to remark that without loss of generality we can take δ^{-1} to be an integer, since if this is not the case we extend the final time (and thus decrease δ) until this condition is satisfied.

On the "small" time-intervals of size 1, C and $\bar{C}$ are approximated by formal expansion in the small parameter $(\alpha\varepsilon)^{-1}$ up to second order. From (4.2), (4.3) and (4.4) it's easily seen that the matrix elements of $\mathcal{H}^{(0)}$, $\mathcal{H}^{(1)}$ and $\mathcal{H}^{(2)}$ are O(1), $O((\alpha\varepsilon)^{-1})$ respectively $O((\alpha\varepsilon)^{-2})$.

Also, any time-dependence of $\mathcal{H}^{(1)}$ and $\mathcal{H}^{(2)}$ on these small time-intervals is removed, i.e. wherever t occurs in $\mathcal{H}^{(1)}$ and $\mathcal{H}^{(2)}$ we substitute the initial time of the time-interval. This gives rise to an $O(\delta)$ contribution to the error.

So we first calculate an $O((\alpha\varepsilon)^{-1})$ approximation of $C(t)$ for t$\in$[0,1], by neglecting all $O((\alpha\varepsilon)^{-1})$ (and $O(\delta)$) contributions. We also calculate an upper bound (the error) on the effect of the neglected terms.

We substitute this expression into Schrödinger's equation and repeat the calculation (now neglecting all $O((\alpha\varepsilon)^{-2})$ contributions) to obtain an $O((\alpha\varepsilon)^{-2})$ approximation of $C(t)$.

Again we substitute this expression in Schrödinger's equation and (neglecting all $O((\alpha\varepsilon)^{-3})$ contributions) we integrate Schrödinger's equation from t=0 to t=1, to obtain an $O((\alpha\varepsilon)^{-3})$ approximation of $C(1)$.

This process is equivalent to second order averaging, except that we now keep track of the error terms.

Basically the same procedure is applied to approximate $\bar{C}(1)$. There's one difference, because $\bar{\mathcal{H}}$ is $O((\alpha\varepsilon)^{-2})$, so in the first step we neglect $O((\alpha\varepsilon)^{-2})$ terms and in the second step already $O((\alpha\varepsilon)^{-4})$ terms.

The error terms in these calculations are very lengthy expressions. Moreover, they are implicit, meaning that they depend on the unknown solution of Schrödinger's equation. Because of that we calculate an upper bound on the error terms using simple arithmetic, i.e. taking absolute value, applying the triangular inequality repeatedly and substituting

$$M_c := \max\{|a_{03}|,|a_{12}|,|a_{21}|,|a_{30}|,|a_{04}|,|a_{13}|,|a_{22}|,|a_{31}|,|a_{40}|\} \tag{5.7}$$

for each $|a_{ij}|$ and substituting

$$M_{Df} := \max_{s\in[0,1]} |f'(s)| \tag{5.8}$$

for each $|f'(s)|$.
Wherever the unknown solution of Schrödinger's equation occurs, we use the Cauchy-Schwarz inequality and the fact that solutions of Schrödinger's equation have norm 1 for all t.
The resulting equations from these calculations are:

$$\begin{aligned} C_{nm_1}(1) &= C_{nm_1}(0) + \frac{2\pi}{i}\frac{1}{\alpha^2\epsilon^2}\sum_{m_2=0}^{n} M^n_{m_1m_2}(0)C_{nm_2}(0) \\ &\pm \left\{\frac{\delta}{\alpha\epsilon}12(7+n)^{3/2}M_cM_{Df} + \frac{1}{\alpha^3\epsilon^3}4100(7+n)^{7/2}(1+10M_c(7+n))M_c^2\right\} \end{aligned} \tag{5.9}$$

with n=0,1,2,...,N_{cut} and m=0,1,...,n, and where the ± sign indicates that the expression immediately following it is an upper bound on the error of the expression(s) preceeding it.

$$\begin{aligned} \bar{C}_{nm_1}(1) &= \bar{C}_{nm_1}(0) + \frac{2\pi}{i}\frac{1}{\alpha^2\epsilon^2}\sum_{m_2=0}^{n} M^n_{m_1m_2}(0)\bar{C}_{nm_2}(0) \\ &\pm \left\{\frac{\delta}{\alpha^2\epsilon^2}15(1+n)^2M_c(1+4M_c)M_{Df} + \frac{1}{\alpha^4\epsilon^4}(1+n)^4M_c^2(1+4M_c)^2940\right\} \end{aligned} \tag{5.10}$$

To obtain the error terms we had to make two assumptions with respect to n, knowing

$$\begin{aligned} \frac{1}{\alpha^2\epsilon^2}(n+7) &< \frac{2}{35} \\ \frac{1}{\alpha\epsilon}M_c(n+7)^{3/2} &< \frac{1}{200} \end{aligned} \tag{5.11}$$

The constants at the right side of these inequalities are (within certain bounds) arbitrary. Making them smaller will restrict the number of subsystems one is able to study but will give a (slightly) smaller upperbound on the error and vica versa. For the particular values we have chosen, a few terms in the calculation of the error of different *order* have equal *magnitude* for the maximum values of n.

The first assumption merely bounds the neighbourhood of the origin we are allowed to consider. This is probably more clear in a classical view. Since x^2+z^2 is the classical counterpart of $(n+1)\alpha^{-2}$, this assumption says $\sqrt{x^2+z^2} < \frac{1}{4}\epsilon$.
If this assumption is not satisfied, one is considering an energy level to far away from the origin.
Similarly, the second assumption says that the perturbation must be small. More explicitly, the contribution of $\mathcal{H}^{(1)}$ to the solution of Schrödinger's equation must be small on a time-interval of size 1.
Another way of saying this is that the matrix elements of the (n,m) rows of $\mathcal{H}^{(1)}$ must be small with respect to the unperturbed energy level spacing, for all m=0,1,...,n.
Note that this does *not* restrict N_{cut}. It only restricts the value of n, the subsystem under consideration (to be more precise, the quantumnumber n indicates the subsystem clustered around the n-th unperturbed (degenerate) energy level).
In practice, the second assumption will imply the first one. Only for extreme values of M_c this might not be the case, which is why we state both of them.
Moreover the second assumption will mostly be satisfied (compare with (1.6)), since otherwise fifth order contributions to the Hamiltonian (1.9) (which we neglected) are likely to have a non-neglectable influence, so this treatment would be invalid anyway.

The error term of $C_{nm}(1)$ consists of two terms. The first term represents the neglected time-dependence of the perturbation, which is small because δ«1, and because of (5.11). The second term represents the effect of the effect of the effect of the third order perturbation, which is of third order and generally neglected in second order avaraging. If this term is not small enough, second order averaging is not accurate enough, and one has to fall back on a more accurate approximation. This will occur only for *high* energy levels.

The error term of $\bar{C}_{nm}(1)$ consists again of two terms, of which the first one again represents the neglected time-dependence of the perturbation. The second term now represents the effect of the effect of the fourth order perturbation, which is of fourth order and again generally neglected in second order averaging.

We now define e_n to be an upper bound on $\left|C_{nm}-\bar{C}_{nm}\right|$ for all m=0,1,…,n.
Taking the difference of (5.9) and (5.10) and again using (5.11) we get:

$$\begin{aligned} e_n(t_0+1) \le\ & e_n(t_0)\left|1+2\pi i\frac{1}{\alpha^2\epsilon^2}M_M(n)\sqrt{5}\right| + \left\{\frac{\delta}{\alpha\epsilon}(16+15M_c)(7+n)^{3/2}M_cM_{Df}\right. \\ & \left. + \frac{1}{\alpha^3\epsilon^3}(41000(7+n)M_c+3760(M_c+2)^2)(7+n)^{7/2}M_c^2\right\} \end{aligned} \tag{5.12}$$

with $M_M(n)$ an upper bound on the length of the vectors formed by the rows of M^n. Using (5.4) we can calculate $M_M(n)$

$$M_M(n) = (n+1)^2M_c(\tfrac{1}{4}+M_c)\sqrt{340} \tag{5.13}$$

Combining the two previous expressions and applying the result recursively, starting with $e_n(t_0=0)=0$, we finally get

$$\begin{aligned} e_n(\tfrac{1}{\delta}) \le\ & \left(1+\frac{33600}{\alpha^4\epsilon^4}(1+n)^4M_c^2(\tfrac{1}{4}+M_c)^2\right)^{\frac{1}{\delta}}\left\{\frac{1}{\alpha\epsilon}(16+15M_c)(7+n)^{3/2}M_cM_{Df}\right. \\ & \left. + \frac{1}{\delta\alpha^3\epsilon^3}(41000(7+n)M_c+3760(M_c+2)^2)(7+n)^{7/2}M_c^2\right\} \\ \le\ & \exp\left(\frac{33600}{\delta\alpha^4\epsilon^4}(1+n)^4M_c^2(\tfrac{1}{4}+M_c)^2\right)\left\{\frac{1}{\alpha\epsilon}(16+15M_c)(7+n)^{3/2}M_cM_{Df}\right. \\ & \left. + \frac{1}{\delta\alpha^3\epsilon^3}(41000(7+n)M_c+3760(M_c+2)^2)(7+n)^{7/2}M_c^2\right\} \end{aligned} \tag{5.14}$$

Remark 1: We recall that this is an upper bound on the error introduced by approximating $\mathcal{H}$ with $\bar{\mathcal{H}}$. The actual error will be smaller.

Remark 2: The combination of (5.14) and (5.11) is the quantum analogy of (1.6) in the sense that it restricts the "small" parameter range of δ from below, i.e. δ is only allowed to take those values for which $e_n(\delta^{-1})$ is sufficiently small.
Substituting some realistic values: $M_c=10$, $M_{Df}=100$, n=1000, $\alpha\varepsilon=10^{15}$ and $\delta=10^{-20}$ we get: $e_{1000}(10^{10}) = 1.4\ 10^{-4}$.

Remark 3: Because of assumptions (5.11) we are not allowed to let n tend to infinity, i.e. take the classical limit. This is also clear from (5.14). The error would become infinitely large.

6 Discussion

The error estimate (5.14) contains a few rather large numerical constants, which are the result of rough estimates on higher order contributions to the error. As such their contribution is very small unless n approaches its upper bound prescribed by (5.11), at which point terms of different order have comparable magnitude.
For instance, of the two terms inside the curly braces in (5.14), the first term will usually be dominant (due to the power of $\alpha\varepsilon$). In the calculations great care has been taken not to overestimate this term. Moreover, from the smallness of the second term, it follows immediately that the argument of the exponent is even much smaller, so the value of the exponent is very close to 1.

There are also a few non-numerical constants, like M_c and M_{Df} in (5.14), which can be large (the realistic values given are not upper bounds) but fixed. Although they do dictate the error estimate in an obvious way, they are constant from an asymptotical point of view (i.e. they do not depend on ε and δ). Their effect is limited to determining the upper bounds of n and $\frac{1}{\delta}$.

In short, for non-extreme values of n (determined by (5.11)) and $\frac{1}{\delta}$, formula (5.14) reduces to

$$e_n \lesssim \frac{1}{\alpha\epsilon}(16+15M_c)(7+n)^{3/2}M_c M_{Df} \tag{6.1}$$

7 Transformation to the instantaneous eigenbasis

Now that we have established the validity of the splitting in subsystems, we can study each subsystem separately. This is done most efficiently by making use of the slow time-dependence of the Hamiltonian. Indeed, if the Hamiltonian were not time-dependent at all, we could write down the explicit solution at once, since system (4.7) is linear.

Each subsystem is of the following form, where the index n references the specific subsystem under consideration (n=0,1,...,N_{cut}):

$$i\frac{1}{2\pi}\dot{C}_{nm_1} = \frac{1}{\alpha^2\epsilon^2}\sum_{m_2=0}^{n} M^n_{m_1m_2}(f(\delta t))C_{nm_2} \tag{7.1}$$

with m_1=0,1,...,n and where we have written explicitly the slow time-dependence of the matrix elements (they only depend on t through f).

Since M^n is a Hermitian matrix, it can be diagonalized for each value of t. This gives us n+1 *real orthonormal* eigenvectors $\vec{v}_{nm}(\delta t)$ (m=0,1,...,n) with corresponding *real* eigenvalues $E^{(2)}_{nm}(\delta t)$.

The eigenvalues are called $E^{(2)}_{nm}$, since using second order perturbation theory, one can show that these correspond with the second order energy correction (the first order energy correction is zero, due to the 1:1 resonance of the unperturbed system).

Under the assumption that f is analytic, Kato [11] (Chap. II, §6) proved that both $\vec{v}_{nm}(\delta t)$ and $E^{(2)}_{nm}(\delta t)$ can be chosen analytically in time for all t.
We are thus allowed to perform the following unitary coordinate transformation (the transformation to the instantaneous eigenbasis):

$$C_{nm_1} = \sum_{m_2=0}^{n} \left(\vec{v}_{nm_2}\right)_{m_1} \bar{C}_{nm_2} e^{-2\pi i \frac{1}{\alpha^2\epsilon^2}\int_0^t E^{(2)}_{nm_2}(f(\delta s))ds}$$
$$\bar{C}_{nm_1} = e^{2\pi i \frac{1}{\alpha^2\epsilon^2}\int_0^t E^{(2)}_{nm_1}(f(\delta s))ds} \sum_{m_2=0}^{n} \left(\vec{v}_{nm_1}\right)_{m_2} C_{nm_2} \tag{7.2}$$

Substituting this into (7.1) we get

$$\frac{d}{dt}\bar{C}_{nm_1} = \sum_{m_2=0}^{n} \left(\frac{d}{dt}\vec{v}_{nm_1}\right)\cdot\vec{v}_{nm_2}\bar{C}_{nm_2} e^{2\pi i \frac{1}{\alpha^2\epsilon^2}\int_0^t \left(E^{(2)}_{nm_1}(f(\delta s))-E^{(2)}_{nm_2}(f(\delta(s))\right)ds} \tag{7.3}$$

By differentiating $M^n\vec{v}_{nm_1} = E^{(2)}_{nm_1}\vec{v}_{nm_1}$ with respect to t and taking the inner product with an arbitrary eigenvector $\vec{v}_{nm_2}$ (with $E^{(2)}_{nm_1} \neq E^{(2)}_{nm_2}$) we get

$$\left(\frac{d}{dt}\vec{v}_{nm_1}\right)\cdot\vec{v}_{nm_2} = \frac{\left(\frac{d}{dt}M^n(f(\delta t))\right)\vec{v}_{nm_1}\cdot\vec{v}_{nm_2}}{E^{(2)}_{nm_1}-E^{(2)}_{nm_2}} \tag{7.4}$$

Combining the two previous equations we finally get

$$\frac{d}{dt}\bar{C}_{nm_1} = \sum_{\substack{m_2=0 \\ m_2\neq m_1}}^{n} \frac{\left(\frac{d}{dt}M^n(f(\delta t))\right)\vec{v}_{nm_1}\cdot\vec{v}_{nm_2}}{E^{(2)}_{nm_1}-E^{(2)}_{nm_2}}\bar{C}_{nm_2} e^{2\pi i \frac{1}{\alpha^2\epsilon^2}\int_0^t \left(E^{(2)}_{nm_1}(f(\delta s))-E^{(2)}_{nm_2}(f(\delta s))\right)ds} \tag{7.5}$$

Remark 1: In the summation over m_2, the case $m_2=m_1$ gives a zero denominator. This is not a problem since it follows from (7.3) and the orthonormality of the set $\vec{v}_{nm}$ that this case can be omitted from the summation.

Remark 2: The denominator becomes also zero if for some value of t, two eigenvalues become equal. This is also no problem, because the theorem of Kato [11] we mentioned before implies that the entire fraction behaves regularly as t passes through this value.

What we are really interested in, is in what way the function f and the parameter δ influence $C_{nm}(\delta^{-1})$. This is analyzed best by transforming to the new time τ defined by

$$\tau(t) = f(\delta t) \tag{7.6}$$

Note that this implies that τ runs from 1 to 0.

Using that $d\tau=\delta f'(\delta t)dt$ and that both M^n and $E^{(2)}_{nm}$ depend only on t through f, (7.5) is transformed into

$$\frac{d}{d\tau}\bar{C}_{nm_1} = \sum_{\substack{m_2=0 \\ m_2\neq m_1}}^{n} \frac{\left(\frac{d}{d\tau}M^{(n)}(\tau)\right)\vec{v}_{nm_1}\cdot\vec{v}_{nm_2}}{E^{(2)}_{nm_1}-E^{(2)}_{nm_2}}\bar{C}_{nm_2} e^{2\pi i \frac{1}{\delta\alpha^2\epsilon^2}\int_1^\tau \frac{E^{(2)}_{nm_1}(s)-E^{(2)}_{nm_2}(s)}{f'(f^{-1}(s))}ds} \tag{7.7}$$

where all quantities are now understood to depend on τ instead of f(δt).

We can now immediately draw three important conclusions:

1. The only influence of the function f on the solution $\bar{C}_{nm}(t)$ is that it determines the importance $(f'(f^{-1}(\tau))^{-1})$ of the degree of asymmetry (τ).

2. It follows directly from system (7.1) that if

$$\frac{2\pi}{\alpha^2\epsilon^2} \sup_{\substack{t\geq 0 \\ m_1=0,1,..,n}} \left| \sum_{m_2=0}^{n} M^n_{m_1 m_2} C_{n m_2} \right| \ll \delta \tag{7.8}$$

we get

$$\left| C_{nm}(\tfrac{1}{\delta}) - C_{nm}(0) \right| \ll 1 \tag{7.9}$$

Using (5.13) we find that the above condition is satisfied if

$$\delta \gg \frac{116}{\alpha^2\epsilon^2} M_c(\tfrac{1}{4}+M_c)(n+1)^2 \tag{7.10}$$

So this limiting case is fairly uninteresting; C_{nm} hardly changes at all.
This is in agreement with the results of Van den Broek [6], who showed that there is no interesting dynamics for the special case $\delta=\varepsilon\gg\varepsilon^2$.

3. The other limiting case, i.e.

$$\delta \ll \frac{116}{\alpha^2\epsilon^2} M_c(\tfrac{1}{4}+M_c)(n+1)^2 \tag{7.11}$$

turns out to be also uninteresting. In this case

$$\left| \bar{C}_{nm}(\tfrac{1}{\delta}) - \bar{C}_{nm}(0) \right| \ll 1 \tag{7.12}$$

We remark that the $C_{nm}(\delta^{-1})$ can be computed easily from this, using the explicit relationship (7.2) between $\bar{C}_{nm}(\frac{1}{\delta})$ and $C_{nm}(\frac{1}{\delta})$.
To obtain (7.12) we introduce the dummy variable $T=\delta t$ into system (7.5) and add the equation $\dot{T}=\delta$.
Replacing all occurences of δt by T and noting that

$$\begin{aligned} \int_0^t \left(E^{(2)}_{nm_1}(f(\delta s)) - E^{(2)}_{nm_2}(f(\delta s)) \right) ds &= t \, \frac{1}{\delta t} \int_0^{\delta t} \left(E^{(2)}_{nm_1}(f(s')) - E^{(2)}_{nm_2}(f(s')) \right) ds' \\ &= t \, F_{m_1 m_2}(T) \end{aligned} \tag{7.13}$$

We obtain a system suitable for general non-periodic first order averaging (Sanders and Verhulst [4]). It's easy to show that the averaged equations are $\frac{d}{dt}\bar{C}_{nm}=0$, so we have proven our statement.
Care has to be taken in the averaging process for those times for which $F_{m_1 m_2}(T)=0$ for some

m_1 and m_2. This problem can be circumvented by excluding time-intervals of size $\frac{\eta}{\delta}$ around these critical times and letting η tend to zero at the end of the averaging process.

So the interesting dynamics takes place for values of δ in between these two limiting cases. In this way δ can be regarded as being a control parameter guiding the system from one uninteresting case to another, through a highly interesting area (if this is not the case, we are studying quite a dull system :-)).
Except for low values of n, this range has to be studied numerically.

For n=0 we get $\frac{d}{dt}\bar{C}_{00}=0$, and since the relation between $\bar{C}_{00}$ and C_{00} is explicitly known, this case is nearly trivial: The groundstate undergoes a time-dependent phaseshift:

$$C_{00}(t) = C_{00}(0)\, e^{-2\pi i \frac{1}{\alpha^2\epsilon^2}\int_0^t E_{00}^{(2)}(f(\delta s))ds} \tag{7.14}$$

with

$$E_{00}^{(2)}(\tau) = \frac{1}{24}(18a_{04}-5a_{12}^2+6a_{22}-18a_{12}a_{30}-33a_{30}^2+18a_{40} + (-33a_{03}^2-18a_{03}a_{21}-5a_{21}^2)\ \tau^2) \tag{7.15}$$

8 Conclusion

We believe to have paved the way for a systematic study (i.e. by studying each subsystem separately) of the 2D quantum harmonic oscillator in 1:1 resonance with a slowly varying perturbation (we never made use of the limiting symmetry properties, so we don't have to restrict ourselves to that class of perturbations).

The study of the first non-trivial sybsystem (n=1), a complicated *linear* time-dependent system of two coupled differential equations, reveals some interesting results. We succeeded in deriving an explicit mapping between initial and final states for the case of exponential decay, which is a powerful tool in the analysis of the subsystem. This explicit mapping shows interesting dynamics precisely for the indicated values of δ. We hope to report on this in a subsequent paper.

8.1 Extensions to related problems

The starting point of this research was to analyse a model problem in such a way that the results could be applied to more genral problems.

The theory is easily extendible to the 1:2 resonance, since this resonance is known to be active already at first order, so the proof of the splitting in subsystems greatly simplifies, since we only have to apply first order averaging techniques.
The only thing that is (slightly) more complicated for the 1:2 resonance is the degeneracy of the unperturbed energy levels; There is a difference in the degeneracy between even and odd values of n.

The theory can also be extended to the 1:3 resonance, since this resonance is known to be active also at second order, so the proof of the splitting in subsystems will hardly change, again apart from a slight complication with respect to the degeneracy.

To prove the splitting for higher order resonances, a lot more work has to be done, at least to obtain a similar error estimate. The extra work arises because one has to apply a third- or higher order averaging process (Verhulst [12]). It's not hard to show that eventually such a process will prove the

splitting, so one can simply study the subsystems without worrying about the introduced error, which is guaranteed to be bounded and small for low values of n.
The only way to get an expression for it, is to go through the laborious high order averaging process.

Under the same restriction (i.e. depending on the order of the resonance) this theory can also be extended to study the low order resonances of the 3D quantum harmonic oscillator with a slowly time-dependent perturbation.
Although the calculation of the error estimate will be longer (due to the extra degree of freedom), there will be no technical complications.
This should also be clear from the fact that the same physical reasoning applies to the 3D case that applied to the 2D case.

Finally, it should be clear that the methods developed in this article can be applied to any quantum perturbation problem for which the unperturbed problem has (some) discrete energylevels and the perturbation is slowly time-dependent and bounded towards infinity, although the calculations will be more elaborate than the ones presented here.

References

[1] J.M. Guckenheimer and P.J. Holmes, Nonlinear oscillations, dynamical systems, and bifurcations of vector fields, Springer Verlag, New York (1990).

[2] V.I. Arnold, Geometrical methods in the theory of ordinary differential equations, Springer Verlag, New York (1983).

[3] V.F. Lazutkin, KAM theory and semiclassical approximations to eigenfunctions, Springer Verlag, Berlin (1993).

[4] J.A. Sanders and F. Verhulst, Averaging methods in nonlinear dynamical systems, Springer Verlag, New York (1985).

[5] E. Hairer et al, Solving ordinary differential equations, Springer Verlag, Berlin (1991).

[6] B. van den Broek, Studies in nonlinear resonance; Applications of averaging, thesis, University of Utrecht (1988).

[7] F. Verhulst, Discrete symmetric dynamical systems at the main resonances with applications to axi-symmetric galaxies, Phil. Trans. R. Soc. London **290** (1979) 435.

[8] A.V. Babin, Symmetry of instabilities and attractors for scalar equations in symmetric domains, preprint, Moscow institute of engeneers of transportation (1994).

[9] M.C. Gutzwiller, Chaos in classical and quantum mechanics, Springer Verlag, New York (1991).

[10] E. Merzbacher, Quantum mechanics, Wiley, New York (1970).

[11] T. Kato, Perturbation theory for linear operators, Springer Verlag, Berlin (1980).

[12] F. Verhulst, A note on higher order averaging, Int. J. Non-Linear Mech. **23** (1988) 341.

[13] J. Sanders, Second quantization and averaging: Fermi resonance, J. Chem. Phys. **74** (10) (1981) 5733

List of captions for illustrations

Set of copies of illustrations

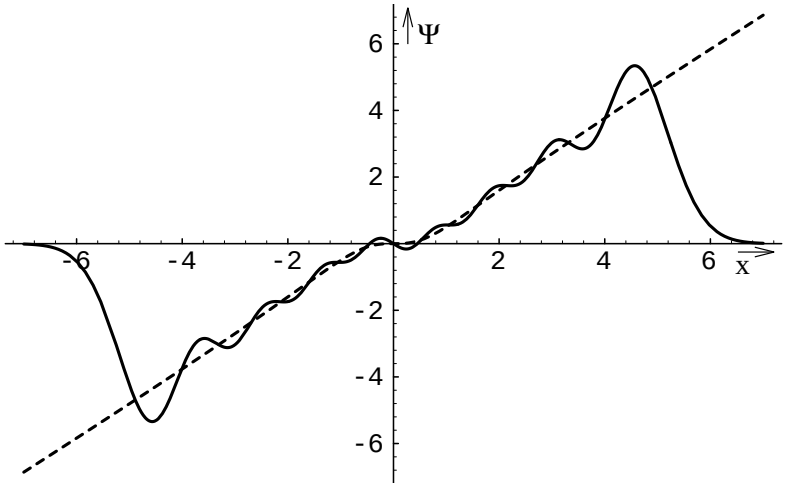

Figure 2.1: The action of the bounded and the unbounded perturbation