WEAKLY NONLINEAR OSCILLATIONS WITH ANALYTIC FORCING

SETH MUSSER

Abstract. Weakly nonlinear oscillations exhibit a wide range of phenomena not seen in simple linear oscillations. This paper considers weakly nonlinear oscillations with an analytic forcing term, and attempts to understand various quantitative solution methods for this problem. We give a quantitative demonstration of the failure of regular perturbation theory, and use this failure to motivate investigation into two-timing and averaging theory. We see that both two-timing and averaging theory give initially identical approximations of a solution, in fact giving excellent approximations even when the nonlinear effects cease to be weak. Finally, we see that the results given by two-timing and averaging theory are physically valid only when the exact solution is bounded.

Contents

1. Introduction

In many physical scenarios an equation will arise which is of the form

$$\ddot{x} + \epsilon h(x, \dot{x}) + x = 0 \text{ where } 0 < \epsilon \ll 1.$$

The initial conditions can be given by $x(0) = y_0$ and $\dot{x}(0) = \dot{y}_0$. In addition we will make the requirement that $h(x, y)$ be an analytic function of x and y. For most of the paper the $h(x, y)$s considered will be convergent for all $x, y \in \mathbb{R}$. Note that the requirement that h be analytic is not too restrictive a condition, as in most well studied models of oscillators h is assumed to be analytic. Examples of such cases might be a spring with small damping, $h(x, y) = 2y$, or perhaps a pendulum with moderate amplitude oscillations where $h(x, y) = -x^3$ and $\epsilon = \frac{1}{6}$. In any case, for a nonlinear h this equation quickly becomes intractable if one attempts to use techniques which work for linear ordinary differential equations. In the absence of exact solutions it becomes a natural thing to try an approximate method, such as regular perturbation theory. However, whatever its successes in other areas of nonlinear differential equations, regular perturbation theory fails miserably when applied

Date: August 29, 2014.

to this problem. This can largely be attributed to its method of solution, which generates unbounded terms in the expansion of $x(t,\epsilon)$ because of resonance in the solutions. In fact, if $h(x,y)=\sigma(x)$ or $h(x,y)=\gamma(y)$, where σ and γ are not constant and also analytic, then regular perturbation theory will have produced an unbounded term by at least the third term of its expansion. Obviously, this is a problem as there exist many hs for which the solution actually deceases (ex: $h(x,y)=2y$), something totally misrepresented by the unboundedness of regular perturbation theory.

Seeing the failure of regular perturbation theory, it becomes natural to look for other approximations which preserve the uniform convergence of $x(t,\epsilon)$. Two (and more)-timing is such an approximation. It succeeds by realizing that there are several natural time scales in such a problem: a fast time, t, which controls the oscillation of solutions; a slow time, ϵt, which controls their major decay and/or large frequency shifts; and a super slow time, $\epsilon^2 t$, which controls the decay and/or frequency shifts of problems with even hs, etc. It then makes the somewhat ad hoc assumption that these different time scales can be treated as independent of one another. This transforms the expansion of $x(t,\epsilon)$ from a series of total differential equations to a series of partial differential equations in each time scale. Though this separation of time into several variables is fundamentally non-rigorous, it does give enough freedom to eliminate any unbounded terms which might arise, a la regular perturbation theory. In doing so it produces very useful differential equations on the amplitude and frequency of solutions. An example using the Van der Pol oscillator will show that these approximations are fantastically accurate, even when $\epsilon=O(1)$.

Those desiring either a more physically motivated or more rigorous approximation can turn to averaging theory. It uses the concepts of energy within an oscillator and the oscillator's mean displacement per cycle as a way to derive differential equations governing the amplitude and frequency of the oscillator. At least to $O(\epsilon)$ these differential equations agree with the ones derived through two-timing, validating the more ad hoc methodology used by two-timing. In fact, the physical perspective given by averaging theory allows us to understand intuitively the difference between even and odd σ and γ, as far as this pertains to the development of unbounded terms in their respective expansions of $x(t,\epsilon)$. However, despite averaging theory being more physically motivated and more rigorous, it becomes clear that it is non-trivial to extend its results to $O(\epsilon^2)$ or further. Two-timing's expansion, however, provides a very (computationally) clear roadmap to expanding any solution to $O(\epsilon^n)\forall n\in\mathbb{N}$, merely by considering n independent timescales $t,\epsilon t,\cdots,\epsilon^n t$.

This being said, even to $O(\epsilon)$ there are certain hs which will produce problems for two-timing and averaging theory. We show the method of scaling that Kevorkian and Cole propose to highlight where these problems might arise. However, by applying a second two timing argument to the normalized solution, we can see that a broader spectrum of hs will create an approximation which is sensitive to initial conditions. This leads us to conclude that, physically, the application of two-timing is extremely useful, provided it is constrained to trajectories that do not head to infinity as t goes to infinity.

2. Regular Perturbation Theory

We know that if x is a solution of the above equation it will be a function of ϵ. This function can be formally expanded as $x(t,\epsilon)=\sum_{n=0}^{\infty}\epsilon^n x_n(t)$. We can also expand h about $(x_0,\dot{x}_0)$. Then we have that $h(x,\dot{x})=h(x_0,\dot{x}_0)+\epsilon\left[\frac{\partial h}{\partial x}(x_0,\dot{x}_0)x_1+\frac{\partial h}{\partial \dot{x}}(x_0,\dot{x}_0)\dot{x}_1\right]+O(\epsilon^2)$.

Substituting these expressions into the first equation, and collecting powers of ϵ gives:

$$\ddot{x}_0 + x_0 = 0 \tag{1}$$

$$\ddot{x}_1 + x_1 = -h(x_0, \dot{x}_0) \tag{2}$$

$$\ddot{x}_2 + x_2 = -\left[\frac{\partial h}{\partial x}(x_0, \dot{x}_0)x_1 + \frac{\partial h}{\partial \dot{x}}(x_0, \dot{x}_0)\dot{x}_1\right], \tag{3}$$

where we can separate powers because we require the equations hold for all small ϵ. Note that requiring the solution be valid for all small ϵ also implies $x_0(0) = y_0$, $\dot{x}_0(0) = \dot{y}_0$ and $x_i(0) = 0, \dot{x}_i(0) = 0$ for $i \neq 0$.

Before discussion of regular perturbation theory we should be specific about what it is we want from an approximate solution. We look for one which will converge uniformly, as this will allow us to compute only a finite number of terms yet still have an idea of the long-term behaviour of the system (as the uniform convergence justifies the neglect of higher order terms for all times t).

Definition 2.1. Let a secular term $x_i(t)$ in the expansion of $x(t, \epsilon)$ be defined as one such that $\limsup_{t\to\infty} |x_i(t)| = \infty$.

Lemma 2.2. The solution to (1) is $x_0(t) = A\cos(t + \phi)$ where $A = \sqrt{y_0^2 + \dot{y}_0^2}$, $\tan\phi = -\frac{\dot{y}_0}{y_0}$.

Proof. We see that $\ddot{x}_0 = -A\cos(t + \phi) = -x_0$, no matter the values of A and ϕ. Plugging in the initial conditions $x(0) = y_0$ and $\dot{x}(0) = \dot{y}_0$ gives the values of A and ϕ. □

Example 2.3. Consider the weakly damped harmonic oscillator, whose amplitude x is given by an equation of the form $\ddot{x} + 2\epsilon\dot{x} + x = 0$. Multiplying by $\dot{x}$ and integrating with respect to time shows that $V(x, \dot{x}, t) = \frac{1}{2}(\dot{x})^2 + \frac{1}{2}x^2 + 2\epsilon\int_0^t (\dot{x})^2 \mathrm{d}t'$ is constant on every trajectory. Thus if the oscillator's initial conditions are $y_0, \dot{y}_0$, then we must have that $V(x, \dot{x}, t) = \frac{1}{2}y_0^2 + \frac{1}{2}\dot{y}_0^2$ for all t. But since each term of $V(x, \dot{x}, t)$ is greater than or equal to zero, this means that $|x| \leq \sqrt{\dot{y}_0^2 + y_0^2}$ for all t.

Now consider attempting to solve the problem with an ordinary perturbative approach. From Lemma 2.2 we have that $x_0(t) = A\cos(t + \phi)$ where A and ϕ are given in Lemma 2.2. But then (2) implies that $\ddot{x}_1 + x_1 = A\sin(t + \phi)$. One can check that a solution to this is $x_1(t) = \frac{A}{2}(\sin(t + \phi) - t\cos(t + \phi))$, where we have used that $x_1(0) = 0, \dot{x}_1(0) = 0$. Thus, we have that $x(t, \epsilon) = A\cos(t + \phi) + \epsilon\frac{A}{2}(\sin(t + \phi) - t\cos(t + \phi)) + O(\epsilon^2)$. Notice that this solution is only valid for $t \ll O(\epsilon^{-1})$, as otherwise the secular term will dominate, in contradiction to the boundedness relation derived above. Indeed, not only does the growth of this term contradict the boundedness relation above, but our hopes for the series of $x(t, \epsilon)$ to be uniformly convergent.

In fact, this equation can be solved explicitly to give

$$x(t, \epsilon) = A\cos\phi\; e^{-\epsilon t}\left(\cos\left(\sqrt{1 - \epsilon^2}t\right) + \frac{\epsilon - \tan\phi}{\sqrt{1 - \epsilon^2}}\sin\left(\sqrt{1 - \epsilon^2}t\right)\right)$$

where the first two terms in its Taylor expansion are $A\cos(t+\phi)+\epsilon\frac{A}{2}(\sin(t+\phi)-t\cos(t+\phi))$.

As noted, the appearance of a secular term means that the expansion is no longer uniformly convergent for all times t. In this case, regular perturbation theory has failed to give a solution which satisfies the condition that we stated; in fact failing by the second term of the series. As in this case, regular perturbation theory often gives an expansion which converges to a solution for all t only if an infinite number of terms are computed. Therefore, it is a natural question to ask for what $h(x, y)$ secular terms occur in the series for $x(t, \epsilon)$, and if they do, where in the series they appear.

First, a few lemmas are required.

Lemma 2.4. The equation $\ddot{x} + x = A\cos(\omega t + \phi)$ only has secular solutions if $\omega = 1$, with the same result holding if cos is replaced with sin.

Proof. One can check that the general solution is given by:

$$x(t) = c_0 \cos t + c_1 \sin t + \begin{cases} \frac{A}{1-\omega^2}\cos(\omega t + \phi) & \text{if } \omega \neq 1 \\ \frac{A}{2} t \sin(t + \phi) & \text{if } \omega = 1 \end{cases},$$

where $c_0, c_1 \in \mathbb{R}$ are chosen to satisfy initial conditions.

If instead we have $\ddot{x} + x = A\sin(\omega t + \phi)$, then the general solution is given by

$$x(t) = c_0 \cos t + c_1 \sin t + \begin{cases} \frac{A}{1-\omega^2}\sin(\omega t + \phi) & \text{if } \omega \neq 1 \\ -\frac{A}{2} t \cos(t + \phi) & \text{if } \omega = 1 \end{cases},$$

where, again, $c_0, c_1 \in \mathbb{R}$ are chosen to satisfy intial conditions. $\square$

This is an extremely useful thing to know, as it can be combined with the Fourier series decomposition of $h(x_0, \dot{x}_0)$ to get an understanding of exactly when secular terms occur. Note that we can Fourier transform $h(x_0, \dot{x}_0)$ because, per our assumption, it is analytic and it is a periodic function of $t + \phi$ (since x_0 and $\dot{x}_0$ are periodic functions of $t + \phi$). Let:

$$h(x_0, \dot{x}_0) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(n\theta) + b_n \sin(n\theta) \text{ where } \theta = t + \phi$$

then by Lemma 2.4, and (2), we see that x_1 will have a secular term if, and only if, a_1 or b_1 is not zero. We now naturally turn to computing the Fourier series of some functions of x_0 and $\dot{x}_0$.

Lemma 2.5. Let $f\colon \mathbb{R} \to \mathbb{R}$ be a function such that $f(\theta + 2\pi) = f(\theta)$. If $f(\theta + \pi) = -f(\theta)$ *or* if $f(-\theta) = -f(\theta)$ then $\int_0^{2\pi} f(\theta)\mathrm{d}\theta = 0$.

Proof. Simple manipulations of integrals reveal this. $\square$

Definition 2.6. Let:

$$f(m, p, n) := \frac{1}{\pi}\int_0^{2\pi} \cos^m \theta \sin^p \theta \cos(n\theta)\mathrm{d}\theta \text{ for } n, m, p \in \mathbb{N} \cup \{0\}$$

$$g(m, p, n) := \frac{1}{\pi}\int_0^{2\pi} \cos^m \theta \sin^p \theta \sin(n\theta)\mathrm{d}\theta \text{ for } n, m, p \in \mathbb{N} \cup \{0\}.$$

Lemma 2.7. We have that $f(m, p, n) = 0$ unless $m, p,$ and n are all even, *or* p is even and m and n are odd. Similarily, we have that $g(m, p, n) = 0$ unless m is even and p and n are odd, *or* n is even and p and m are odd. It is also the case that $f(2m, 2p, 0) > 0$ for all $m, p \in \mathbb{N} \cup \{0\}$.

Proof. Let $f(\theta) := \cos^m\theta\sin^p\theta\cos(n\theta)$ and $g(\theta) := \cos^m\theta\sin^p\theta\sin(n\theta)$. We have that $f(-\theta) = (-1)^p f(\theta)$ and $g(-\theta) = (-1)^{p+1}g(\theta)$. So if p is odd then $f(m,p,n) = 0$ by Proposition 2.5, and similarily, if p is even then $g(m,p,n) = 0$. Additionally, $f(\theta+\pi) = (-1)^{m+p+n}f(\theta)$ and $g(\theta+\pi) = (-1)^{m+p+n}g(\theta)$. Then since p must be even for $f(m,p,n)$ to be nonzero, by Proposition 2.5 we have that n and m must both be even or both be odd for $f(m,p,n)$ to be nonzero (since if $n+p+m$ is odd then $f(m,p,n) = 0$). Similarily, we see that n and m must have opposite parity for $g(m,p,n)$ to be nonzero.

Now let $p(\theta) := \cos^m\theta\sin^p\theta$. Then $f(2m,2p,0) = \frac{1}{\pi}\int_0^{2\pi}[p(\theta)]^2\mathrm{d}\theta$. We see that p is continuous, not identically zero, and therefore $f(2m,2p,0)$ is strictly positive. $\square$

Remark 2.8. Note that the above conditions given in Lemma 2.7 are necessary, but not sufficient, for f and g to be nonzero; for example, it can be shown that $f(2,2,2) = 0$ even though $m, p,$ and n are all even.

One final lemma and then it is time to see in what cases regular perturbation theory fails our criteria for a good approximation.

Definition 2.9. Let $\begin{pmatrix} n \\ k \end{pmatrix} := \frac{n!}{k!(n-k!)}$ when $n \geq k$ and $n,k \in \mathbb{N}\cup\{0\}$ and $\begin{pmatrix} n \\ k \end{pmatrix} := 0$ if $n < k$ or if $k \notin \mathbb{N}\cup\{0\}$ or $n \notin \mathbb{N}\cup\{0\}$.

Lemma 2.10. We have that $f(m,0,n) = 2^{1-m}\begin{pmatrix} m \\ \frac{m+n}{2} \end{pmatrix}$ and $g(m,0,n) = 0$ for all $m,n \in \mathbb{N}\cup\{0\}$. We also have that $f(0,2p,2n) = (-1)^n 2^{1-2p}\begin{pmatrix} 2p \\ p+n \end{pmatrix}$ and $g(0,2p+1,2n+1) = (-1)^n 2^{-2p}\begin{pmatrix} 2p+1 \\ p-n \end{pmatrix}$ for all $p,n \in \mathbb{N}\cup\{0\}$ and that $f(0,p,n) = 0$ if p or n is odd and $g(0,p,n) = 0$ if p or n is even. Furthermore, $g(1,2k+1,2n) = (-1)^{n+1}2^{1-2k}\frac{2n}{2k+2}\begin{pmatrix} 2k+2 \\ k+n+1 \end{pmatrix}$

Proof. Since in the expression $g(m,0,n)$, p is in the form $p = 2k$, where $k = 0$, then p is even and by Lemma 2.7 this means that $g(m,0,n) = 0$ regardless of the values of m or n. It is also known that $\cos\theta = 2^{-1}(e^{i\theta}+e^{-i\theta})$ where $i^2 = -1$. Then we have that:

$$\begin{aligned}
f(m,0,n) =& \frac{1}{\pi}\int_0^{2\pi}\cos^m\theta\cos(n\theta)\mathrm{d}\theta \\
=& \frac{1}{2^{m+1}\pi}\int_0^{2\pi}(e^{i\theta}+e^{-i\theta})^m(e^{in\theta}+e^{-in\theta})\mathrm{d}\theta \\
=& \frac{1}{2^{m+1}\pi}\int_0^{2\pi}\left[\left(e^{in\theta}+e^{-in\theta}\right)\sum_{k=0}^{m}\begin{pmatrix} m \\ k \end{pmatrix}e^{i(m-2k)\theta}\right]\mathrm{d}\theta \text{ by the Binomial Exp. Thm.} \\
=& \frac{1}{2^{m+1}\pi}\sum_{k=0}^{m}\begin{pmatrix} m \\ k \end{pmatrix}\left[\int_0^{2\pi}e^{i(m+n-2k)\theta}\mathrm{d}\theta + \int_0^{2\pi}e^{i(m-n-2k)\theta}\mathrm{d}\theta\right]
\end{aligned}$$

where summation and integration could be interchanged because the sum was finite, and integration is a linear operator. Note that $\int_0^{2\pi}e^{ip\theta}\mathrm{d}\theta = -\frac{i}{p}\left[e^{ip\theta}\right]_0^{2\pi} = 0$ if $p \neq 0$, but if $p = 0$ then the integral is equal to 2π. Thus, the only nonzero contributions to the sum will be when $m+n-2k=0$ and $m-n-2k=0$. This also implies that if $(m\pm n)/2$ is not an integer, then there will be no k such that this occurs, and the entire sum will be zero. Thus:

$$f(m,0,n) = \frac{1}{2^{m+1}\pi}\left[2\pi\binom{m}{\frac{m+n}{2}} + 2\pi\binom{m}{\frac{m-n}{2}}\right].$$

It can be checked that $\binom{m}{\frac{m+n}{2}} = \binom{m}{\frac{m-n}{2}}$. Then having taken care of the case where $\frac{m\pm n}{2} \notin \mathbb{N}$ by setting the binomial coefficient to zero if this is true, we have that $f(m,0,n) = 2^{1-m}\binom{m}{\frac{m+n}{2}}$.

Since $m=0$ is even, then by Lemma 2.7, if p and n are not both even $f(0,p,n)=0$. Also by Lemma 2.7, if p and n are not both odd then $g(0,p,n)=0$. Finally, by realizing $\sin\theta = -2^{-1}i(e^{i\theta}-e^{-i\theta})$ and using identical arguments we can arrive at the expressions given above for $f(0,2p,2n)$, $g(0,2p+1,2n+1)$, and $g(1,2k+1,2n)$. □

Since we have assumed $h(x,y)$ to be analytic, it can be written in the form $h(x,y) = \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} d_{ij}x^iy^j$ where $d_{ij}\in\mathbb{R}$ for all $i,j\in\mathbb{N}\cup\{0\}$. Referring back to Lemma 2.2, we see that no matter the $h(x,y)$ we have $x_0(t) = A\cos(t+\phi)$. To simplify calculation $t+\phi := \theta$, where this is simple substitution since $\mathrm{d}t = \mathrm{d}\theta$. Therefore, in the equation for $\ddot{x}_1 + x_1 = -h(x_0,\dot{x}_0)$, if h is analytic we have:

$$\frac{\mathrm{d}^2x_1}{\mathrm{d}\theta^2} + x_1 = \sum_{j=0}^{\infty}\sum_{i=0}^{\infty}(-1)^{j+1}A^{i+j}d_{ij}\cos^i\theta\sin^j\theta \tag{4}$$

$$= \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n\cos(n\theta) + b_n\sin(n\theta) \text{ where } a_i, b_i \text{ Fourier coeff. of } -h(x_0,\dot{x}_0). \tag{5}$$

As mentioned before, x_1 will have a secular term only when $a_1, b_1 = 0$. Therefore, to see what sort of analytic h will not give secular terms, we may calculate a_1 and b_1 and set them equal to zero. Hopefully, once this is done we will have a meaningful condition on d_{ij}. In fact we do get such a result, provided that we require $x_1(t)$ not have a secular term for *all* initial conditions. However, as discussed x_1's secular terms are directly dependent on h. So if x_1 was without secular terms only for certain initial conditions, this means that h would depend in some way on initial conditions. To have our forcing function dependent on intial conditions is anathema to the physical picture we hope to build, so we can see it would be a kind of cheating to not require $x_1(t)$ to be non-secular for all initial conditions.

Theorem 2.11. If h is analytic of the form $h(x,y) = \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} d_{ij}x^iy^j$ and $x_1(t)$ does not have secular terms for any given initial conditions, then:

$$\sum_{i=0}^{l} d_{(2i+1)(2l-2i)}f(2i+2, 2l-2i, 0) = 0 \quad \forall l\in\mathbb{N}\cup\{0\} \tag{6}$$

$$\sum_{j=0}^{l} d_{(2l-2j)(2j+1)}f(2l-2j, 2j+2, 0) = 0 \quad \forall l\in\mathbb{N}\cup\{0\}. \tag{7}$$

Proof. The conditions (6) and (7) on d_{ij} arise simply from computing a_1 and b_1 from (4) and setting them equal to zero (and by Lemma 2.4 getting rid of any secular terms in x_1). The calculation that follows is somewhat tedious, but the result is hopefully rewarding enough.

$$\begin{aligned}
0 := a_1 =& \frac{1}{\pi}\int_0^{2\pi}\left[\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}(-1)^{j+1}A^{i+j}d_{ij}\cos^i\theta\sin^j\theta\right]\cos\theta \mathrm{d}\theta \\
&= \sum_{j=0}^{\infty}\sum_{i=0}^{\infty}(-1)^{j+1}A^{i+j}d_{ij}f(i+1,j,0)^1 \\
&= \sum_{j=0}^{\infty}\sum_{i=0}^{\infty}(-1)^{2j+1}A^{2(i+j)+1}d_{(2i+1)(2j)}f(2i+2,2j,0) \text{ by Lemma 2.4} \\
&= \sum_{l=0}^{\infty}\sum_{i=0}^{l}A^{2l+1}d_{(2i+1)(2l-2i)}f(2i+2,2l-2i,0) \text{ by setting l=j+i}
\end{aligned}$$

One can view this as a power series for A, where $0 := \sum_{l=0}^{\infty} v_l A^{2l+1}$, and the v_ls are given by (6). We require that the series $\sum_{l=0}^{\infty} v_l A^{2l+1}$ is zero for all values of $A \in \mathbb{R}$ (or, equivalently $a_1 = 0$ for all initial conditions). However, this implies that all of its derivatives are zero, and thus that $v_l = 0$ for all $l \in \mathbb{N} \cup \{0\}$. Then we have condition (6) on the d_{ij}. A similar computation requiring $b_1 := 0$ gives condition (7). □

Corollary 2.12. If all d_{ij} terms have the same sign, then $d_{ij} = 0$ if i and j are not of the same parity. More importantly, if $h(x, y)$ is analytic and has no cross terms so that $h(x, y) = \sigma(x) + \gamma(y)$, then σ and γ must both be even functions.

Proof. Suppose that all $d_{ij} \geq 0$. By Lemma 2.7 $f(2m, 2p, 0) > 0$ for all $m, p \in \mathbb{N} \cup \{0\}$. So then $0 \leq d_{(2i_0+1)(2l-2i_0)}f(2i_0+2, 2l-2i_0, 0) \leq \sum_{i=0}^{l} d_{(2i+1)(2l-2j)}f(2i+2, 2l-2i, 0) = 0$ where i_0 is any natural number less than or equal to l. Dividing by the positive $f(2i_0+2, 2l-2i_0, 0)$ we see that $d_{(2i_0+1)(2l-2i_0)} = 0$ for every $l \in \mathbb{N} \cup \{0\}$, and every $i_0 \leq l$. Picking $i_0 = n$ and $l = n + k \geq i_0$ reveals that $d_{(2n+1)(2k)} = 0$ for every $n, k \in \mathbb{N} \cup \{0\}$. Using condition (7) we can see that $d_{(2k)(2n+1)} = 0$ is also true for every $n, k \in \mathbb{N} \cup \{0\}$. The same argument holds, with flipped inequalities, if $d_{ij} \leq 0$. Thus, if all d_{ij} have the same sign, then $d_{ij} = 0$ if i and j are not the same parity and $x_1(t)$ is required to not contain any secular terms.

Finally, let $d_{ij} = 0$ if *both* j and i are greater than or equal to one. In other words let $h(x, y)$ be analytic, so that $h(x, y) = \sum_{j=0}^{\infty} d_{0j}y^j + \sum_{i=1}^{\infty} d_{i0}x^i := \gamma(y) + \sigma(x)$. But then condition (6) becomes $d_{(2l+1)(0)}f(2l+2, 0, 0) = 0$ for all $l \in \mathbb{N}\cup\{0\}$ and condition (7) becomes $d_{(0)(2l+1)}f(0, 2l + 2, 0) = 0$ for all $l \in \mathbb{N} \cup \{0\}$. Since $f(2m, 2p, 0) > 0$ this means that all the odd terms of σ and γ must be zero. Thus, if $h(x, y) = \sigma(x) + \gamma(y)$ then $\sigma(x) = \sigma(-x)$ and $\gamma(-y) = \gamma(y)$. □

Even though conditions (6) and (7) appear at first glance to be fairly algebraically intractable, they are very useful for most physical phenomena. Three of the most important equations of the form we have been considering are: the Van der Pol oscillator, the unforced Duffing equation, and the damped harmonic oscillator. Respectively they have the form: $h_V(x, y) = y(x^2 - 1)$, $h_D(x, y) = x^3, h_H(x, y) = 2y$. Both h_V and h_H have nonzero d_{01}, and thus fail condition (7). Additionally, h_D has no cross terms, but is an odd function of x. Thus, attempting to use regular perturbation theory for any of the three will give a secular

[1] because the series converges uniformly in A (true since h is analytic)

term in $x_1(t)$. Then for any finite expansion of $x(t,\epsilon)$ the approximation given by regular perturbation theory will only be valid for $t \ll O(\epsilon^{-1})$. So then regular perturbation theory fails miserably to approximate arguably the most important weakly nonlinear equations. However, what if it is possible to construct some $h(x,y)$ such that it never has a secular term in any $x_i(t)$? This would certainly be redeeming for regular perturbation theory. A simple example will suggest that this is not the case.

Example 2.13. Consider $h(x,y) = \sigma(x) = x^2$. Then we have that $\frac{\mathrm{d}^2 x_1}{\mathrm{d}\theta^2} + x_1 = -A^2\cos^2\theta = -\frac{A^2}{2}(1+\cos(2\theta))$. By Lemma 2.4 this gives:

$$x_1(t) = c_0\cos(t+\phi) + c_1\sin(t+\phi) - \frac{A^2}{2} + \frac{A^2}{6}\cos(2t+2\phi),$$

where c_0 and c_1 are chosen so that $x_1(0) = 0$ and $\dot{x}_1(0) = 0$. But then from (3) this means that:

$$\begin{aligned}
\frac{\mathrm{d}^2 x_2}{\mathrm{d}\theta^2} + x_2 =& -2x_0x_1 \\
=& A^3\cos\theta - \frac{A^3}{3}\cos\theta\cos(2\theta) - 2Ac_0\cos^2\theta - 2Ac_1\sin\theta\cos\theta \\
=& -Ac_0 + \frac{5A^3}{6}\cos\theta - Ac_0\cos(2\theta) - Ac_1\sin(2\theta) - \frac{A^3}{6}\cos(2\theta).
\end{aligned}$$

So then we see that $\frac{5A^3}{6}\cos\theta$ generates a secular term, ensuring that $x_2(t)$ has a secular term. So in this case even though $h(x,y)$ manages to avoid a secular term in $x_1(t)$, it will still arise by $x_2(t)$.

Theorem 2.14. Suppose that $h(x,y) = \sigma(x)$ or $h(x,y) = \gamma(y)$ is a function such that neither $x_1(t)$ nor $x_2(t)$ contains secular terms. If $h(x,y)$ is analytic, then it is constant.

Proof. For ease of notation, let $L_p(x) = \frac{\mathrm{d}^2 x}{\mathrm{d}p^2} + x$. Let $h(x,y) = \sigma(x)$ where σ is analytic. Then $\sigma(x) = \sum_{i=0}^{\infty} b_i x^i$. We know from Corollary 2.12 that if $x_1(t)$ does not contain any secular terms we can write $\sigma(x) = \sum_{i=0}^{\infty} d_i x^{2i}$, where $b_{2i} = d_i$. Therefore, (2) becomes $L_\theta(x_1) = -\sum_{m=0}^{\infty} A^{2m} d_m \cos^{2m}\theta$. Lemma 2.10 allows us to Fourier transform this to:

$$\begin{aligned}
L_\theta(x_1) =& -\sum_{n=0}^{\infty}\sum_{m=0}^{\infty} 2^{1-\delta_{n0}-2m} A^{2m} d_m \binom{2m}{m+n} \cos(2n\theta)^2 \\
x_1(\theta) =& \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \frac{2^{1-\delta_{n0}-2m}}{4n^2-1}\binom{2m}{m+n} A^{2m} d_m \cos(2n\theta) + \cos\theta \sum_{m=0}^{\infty} 2^{-2m}\binom{2m}{m} A^{2m} d_m \\
L_\theta(x_2) =& -\sigma'(x_0)x_1 \text{ by (3)} \\
=& \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\sum_{i=0}^{\infty} \frac{(4i+4)}{2^{2m+\delta_{n0}}(1-4n^2)}\binom{2m}{m+n} A^{2(i+m)+1} d_{i+1} d_m \cos^{2i+1}\theta\cos(2n\theta) \\
& -\sigma'(x_0(\theta))\cos\theta \sum_{m=0}^{\infty} 2^{-2m}\binom{2m}{m} A^{2m} d_m
\end{aligned}$$

Note that since $x_0(\theta) = A\cos\theta$, and since $\sigma'(x)$ is an odd function of x,

$$\begin{aligned}
\sigma'(x_0(\theta+\pi)) &= -\,\sigma'(x_0(\theta)) \\
\implies \cos^2(\theta+\pi)\sigma'(x_0(\theta+\pi)) &= -\cos^2\theta\sigma'(x_0(\theta)) \\
\sin(\theta+\pi)\cos(\theta+\pi)\sigma'(x_0(\theta+\pi)) &= -\sin\theta\cos\theta\sigma'(x_0(\theta)).
\end{aligned}$$

Then we have that the last term contributes nothing to the a_1 or b_1 of the term on the RHS of the last equation. By Lemma 2.7 we see that $b_1 = 0$. We require that $a_1 = 0$; by Lemma 2.10 this gives:

$$a_1 = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\sum_{i=0}^{\infty} \frac{2i+2}{2^{2(i+m)+\delta_{n0}}(1-4n^2)} \binom{2m}{m+n}\binom{2i+2}{i+n+1} A^{2(m+i)+1} d_{i+1} d_m := 0 \tag{8}$$

$$\implies \sum_{n=0}^{\infty}\sum_{i=0}^{l} \frac{2i+2}{2^{\delta_{n0}}(1-4n^2)} \binom{2l-2i}{l-i+n}\binom{2i+2}{i+n+1} d_{i+1} d_{l-i} = 0 \quad \forall l \in \mathbb{N}\cup\{0\}. \tag{9}$$

The implication above was arrived at by setting $m + i = l$ and realizing the coefficients of all powers of A must be zero for this to be true for all A. Also note that if $n > l$, then no matter the value of i, $(2l-2i, l-i+n) = 0$, thus the sum becomes finite over both i and n.

Let $\Delta_{(i+1)(l-i)} = d_{i+1} d_{l-i}$ and let:

$$\mu_{(i+1)(l-i)} := \sum_{n=0}^{l} \frac{2i+2}{2^{\delta_{n0}}(1-4n^2)} \binom{2l-2i}{l-i+n}\binom{2i+2}{i+n+1}$$

We want to show that for every l and every i, $\mu_{(i+1)(l-i)} > 0$. We notice that the only negative terms appears in the sum when $n > 0$. Thus, we need to show:

$$\left|\sum_{n=1}^{l} \frac{1}{1-4n^2} \binom{2l-2i}{l-i+n}\binom{2i+2}{i+n+1}\right| < \frac{1}{2}\binom{2l-2i}{l-i}\binom{2i+2}{i+1}.$$

Since $\binom{2k}{k} > \binom{2k}{k+l}$ for every k and l, we have:

$$\begin{aligned}
\left|\sum_{n=1}^{l} \frac{1}{1-4n^2} \binom{2l-2i}{l-i+n}\binom{2i+2}{i+n+1}\right| &< \left|\sum_{n=1}^{l} \frac{1}{1-4n^2} \binom{2l-2i}{l-i}\binom{2i+2}{i+1}\right| \\
&< \binom{2l-2i}{l-i}\binom{2i+2}{i+1} \sum_{n=1}^{\infty} \frac{1}{4n^2-1} \\
&= \frac{1}{2}\binom{2l-2i}{l-i}\binom{2i+2}{i+1}.
\end{aligned}$$

Note that we used the fact that $\frac{2}{4n^2-1} = \frac{1}{2n-1} - \frac{1}{2n+1}$, so that $\sum_{n=1}^{l} \frac{1}{4n^2-1} = \frac{l}{2l+1}$. This holds for every l and i, so that

$$(9) \implies \sum_{i=0}^{l} \mu_{(i+1)(l-i)} \Delta_{(i+1)(l-i)} = 0 \quad \forall l \in \mathbb{N}\cup\{0\},$$

[2] The Kronecker delta, or $\delta_{ij} := \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$ is used to incorporate $\frac{a_0}{2}$ into the sum.

with $\mu_{(i+1)(l-i)} > 0$ no matter the i or l. Suppose that $d_0 \neq 0$. Then using the above sum when $l = 0$ and dividing by $\mu_{10}d_0$ reveals $d_1 = 0$. Now let $P(n)$ be the proposition that $d_k = 0$ for all $k \leq n$ such that $k \neq 0$. Assuming $P(n)$ and considering $l = n$ gives that $\mu_{(n+1)(0)}\Delta_{(n+1)(0)} = 0$ is the only relavent term in the above sum. Dividing through by the nonzero μ and d_0 means that $d_{n+1} = 0$, or $P(n+1)$ is true. Thus by induction $d_k = 0$ for all $k \in \mathbb{N}$.

Alternatively, suppose that $d_0 = 0$. Then one has that $\mu_{11}d_1^2 + \mu_{20}d_0d_2 = \mu_{11}d_1^2 = 0$, or $d_1 = 0$ (with $l = 1$). Let $l = 2n+1$ and let $P(n)$ once again be the proposition that $d_k = 0$ for all $k \leq n$ such that $k \neq 0$. Suppose that $2n - i + 1 \geq n + 1$ and that $i + 1 \geq n + 1$. Then $i \leq n$ and $i \geq n$ ensuring that $i = n$. Note that if either $i+1$ or $2n-i+2$ is less than $n+1$ then $\Delta_{(i+1)(2n-i+1)} = 0$ by assumption of $P(n)$. Thus if $P(n)$ is true, taking $l = 2n+1$ gives $\mu_{(n+1)(n+1)}d_{n+1}^2 = 0$, and dividing gives $P(n+1)$. Then by induction $d_k = 0$ for all $k \in \mathbb{N} \cup \{0\}$.

A similar argument can be used if $h(x, y) = \gamma(y)$ with γ analytic. In this case $\mu_{(i+1)(l-i)} = -\sum_{n=1}^{l-i} \frac{4n^2}{2^{\delta_{n0}}(1-4n^2)} \binom{2l-2i}{l-i+n}\binom{2i+2}{i+n+1}$. It is clear that $\mu < 0$ in this case. Since the sum is nonexistant for $\Delta_{(l+1)(0)}$, the second induction argument will work no matter the d_0 value. Then we have shown that if $h(x, y)$ is an analytic even function of just x or y, and is not trivially constant, then it avoids secular terms in its expansion out to x_1, but will certainly acquire them if expanded to x_2. $\square$

3. Two (and more) - Timing

Seeing the spectacular failure of regular perturbation theory to address weakly nonlinear oscillations, we must search for an approximation which satisfies the conditions we gave at the beginning of the paper. Referring back to Example 2.1, we see that the exact solution featured a "slow" time ϵt and a "fast" time t. Heuristically we may guess that this is true for most weakly nonlinear oscillations; there is a fast time during which the system oscillates and a slow time during which the system decays. We expect this because, since $|\epsilon| \ll 1$, the nonlinear behaviour is mostly suppresed, so that given a small time t_0 the system behaves like a harmonic oscillator, but given a long time $t_0\epsilon^{-1}$ it reveals its nonlinear colours. Seeking a way to quantify these heuristic assumptions we make the apparently very ad hoc suggestion that t, ϵt, $\cdots$, and $\epsilon^n t$ are independent variables. Initially, this method seems suspicious, so we shall test it.

Consider only two time scales $T_0 := t$ and $T_1 := \epsilon t$. Since we choose to treat these as independent variables we have that:

$$
\begin{aligned}
\frac{\mathrm{d}x}{\mathrm{d}t} &= \frac{\partial x}{\partial T_0}\frac{\partial T_0}{\partial t} + \frac{\partial x}{\partial T_1}\frac{\partial T_1}{\partial t} \\
&= \frac{\partial x}{\partial T_0} + \epsilon\frac{\partial x}{\partial T_1} \\
\implies \quad \frac{\mathrm{d}^2x}{\mathrm{d}t^2} &= \frac{\partial^2 x}{\partial T_0^2} + 2\epsilon\frac{\partial^2 x}{\partial T_1 \partial T_0} + \epsilon^2\frac{\partial^2 x}{\partial T_1^2}.
\end{aligned}
$$

Substituting these into the original expression, expanding h about $\left(x_0, \frac{\partial x_0}{\partial T_0}\right)$, and collecting powers of ϵ gives:

$$\begin{aligned}
&\frac{\partial^2 x_0}{\partial T_0^2} + x_0 = 0 &(10)\\
&\frac{\partial^2 x_1}{\partial T_0^2} + x_1 = -2\frac{\partial^2 x_0}{\partial T_1 \partial T_0} - h\left(x_0, \frac{\partial x_0}{\partial T_0}\right), &(11)
\end{aligned}$$

where we have ignored the equations governing $x_i(t)$ for $i \geq 2$. An example will illuminate the effectiveness of the method, even for large ϵ.

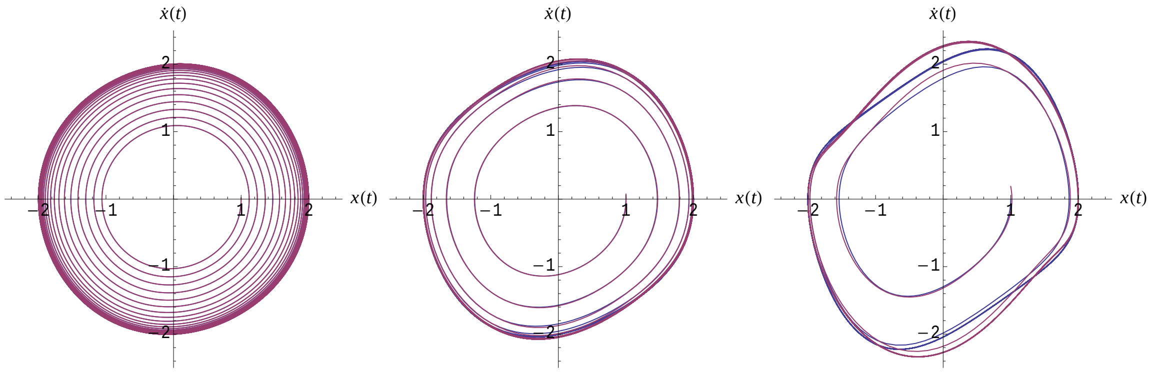
Example 3.1. Suppose that we have the Van der Pol oscillator, $h(x, y) = y(x^2 - 1)$, with initial conditions $x(0) = 1$ and $\frac{\partial x}{\partial T_0}(0) = 0$. By (10) we have that $x_0(T_0) = A(T_1)\cos(T_0 + \phi(T_1))$, where $\frac{\partial A}{\partial T_0} = \frac{\partial \phi}{\partial T_0} = 0$. Substitution into (11) gives:

(12)
$$\frac{\partial^2 x_1}{\partial T_0^2} + x_1 = \left(2\frac{\partial A}{\partial T_1} - A + \frac{1}{4}A^3\right)\sin(T_0 + \phi) + 2A\frac{\partial \phi}{\partial T_1}\cos(T_0 + \phi) + \frac{1}{4}A^3\sin(3T_0 + 3\phi).$$

Now we see why two-timing is so useful. It allows us to choose $A(T_1)$ so that the secular term in x_1 can be avoided. Making that choice in this case gives the equations $A' = \frac{1}{2}A - \frac{1}{8}A^3$ and $A\phi' = 0$, where the prime indicates partial differentiation with respect to T_1. However, since A and ϕ are constant under T_0, and since we are only considering two time scales, the partial differential equations above are actually total differential equations with respect to T_1. Integrating the equation for A using partial fractions gives that $A(T_1) = \frac{2}{\sqrt{1+ce^{-T_1}}}$ where $c = \frac{4}{A(0)^2} - 1$. Since $A(0) \neq 0$ this is never zero, so $A\phi' = 0$ implies $\phi' = 0$, or ϕ is constant under both T_0 and T_1. From the initial conditions it is clear that $\phi = 0$ and $A(0) = 1$. Then $x_0(t) = \frac{2}{\sqrt{1+3e^{-\epsilon t}}}\cos t$, since $T_1 = \epsilon t$. Since A and ϕ are now uniquely determined, it is easy to calculate x_1 from (12). Using Lemma 2.4 we have $x_1(t) = \frac{3\sin t - \sin(3t)}{4(1+3^{-\epsilon t})^{3/2}}$, where we have ensured that $x_1(0) = 0$ and $\frac{\partial x_1}{\partial T_0}(0) = 0$. So then we have that:

$$x(t, \epsilon) = \frac{2\cos t}{\sqrt{1 + 3e^{-\epsilon t}}} + \epsilon\frac{3\sin t - \sin(3t)}{4(1 + 3e^{-\epsilon t})^{3/2}} + O(\epsilon^2)$$

Below is a series of plots in the phase plane. The red trajectory is the approximation, whereas the blue trajectory was numerically integrated by Mathematica.



(A) Small $\epsilon = .05$ (B) Medium $\epsilon = .2$ (C) Large $\epsilon = .5$

FIGURE 1. Phase plane comparisons for Van der Pol oscillator

What is spectacular about the figures shown above is the stunning range of ϵ over which they are accurate. Even when ϵ is on the order of 1, far greater than the restriction we made in the beginning of the paper that $|\epsilon| \ll 1$, we see that the approximation is still very accurate.

Heartened by how surprisingly accurate two-timing is, even over a large range of ϵ, we would like to move to more general results. Let $b_1(A) = \frac{1}{\pi}\int_0^{2\pi} \sin\theta \; h\left(x_0, \frac{\partial x_0}{\partial T_0}\right) \mathrm{d}\theta$ and $a_1(A) = \frac{1}{\pi}\int_0^{2\pi} \cos\theta \; h\left(x_0, \frac{\partial x_0}{\partial T_0}\right) \mathrm{d}\theta$, where $\theta = T_0 + \phi(T_1)$ is constant over T_0. It is clear from Example 3.1 above that if we require

$$\frac{\partial A}{\partial T_1} = \frac{b_1(A)}{2} \text{ and } A\frac{\partial \phi}{\partial T_1} = \frac{a_1(A)}{2} \tag{13}$$

that this will eliminate any secular terms from x_1.

4. Physical Perspective and Averaging Methods

Another approach to approximating solutions for problems of this form are those found from averaging theory. Averaging theory has the added pro of containing more physical insight, at least in the $O(\epsilon)$ terms. It is also more easily made rigorous. However, it is harder than two timing to extend to other equation types, such as partial differential equations. Averaging theory's usual approach is to convert the equation of weakly nonlinear oscillation into two equations in polar form in the phase plane. From these the $O(\epsilon)$ decay or growth of trajectories and the $O(\epsilon)$ period differences are discerned. We will take an equivalent approach, but one that will hopefully give more physical insight.

Definition 4.1. Define $\mathcal{E}(t) = \frac{1}{2}(\dot{x}(t))^2 + \frac{1}{2}(x(t))^2$ to be the energy possessed by the oscillator at time t, where $x(t)$ is its displacement from the origin.

Proposition 4.2. Suppose that the weakly nonlinear equations give a periodic solution with $T = 2\pi + O(\epsilon)$. Then $\frac{\mathcal{E}(T) - \mathcal{E}(0)}{T} = \epsilon A b_1(A)/2 + O(\epsilon^2)$. If we regard the average energy change per cycle as equal to its instantaneous rate of change we recover $\frac{\mathrm{d}A}{\mathrm{d}t} = \epsilon b_1(A)/2 + O(\epsilon^2)$.

Proof. For now let the period of the solution be T. We see from multiplying the oscillation equation by $\dot{x}$ and integrating over time that:

$$\begin{aligned}
0 =& \ddot{x}\dot{x} + x\dot{x} + \epsilon h(x,\dot{x})\dot{x} \\
0 =& \int_0^T \ddot{x}\dot{x} + x\dot{x} + \epsilon h(x,\dot{x})\dot{x} \ \mathrm{d}t \\
0 =& \frac{1}{2}\left([\dot{x}(T)]^2 - [\dot{x}(0)]^2 + [x(T)]^2 - [x(0)]^2\right) + \epsilon \int_0^T h(x,\dot{x})\dot{x} \ \mathrm{d}t' \\
0 =& \mathcal{E}(T) - \mathcal{E}(0) + \epsilon \int_0^T h(x,\dot{x})\dot{x}\mathrm{d}t' \\
\frac{\mathcal{E}(T) - \mathcal{E}(0)}{T} =& -\frac{\epsilon}{T}\int_0^T h(x_0,\dot{x}_0)\dot{x}_0 \ \mathrm{d}t + O(\epsilon^2) \text{ where } T = 2\pi + O(\epsilon) \\
=& \epsilon A \frac{b_1(A)}{2} + O(\epsilon^2) \text{ where } A \text{ treated as constant} \\
\frac{\mathrm{d}\mathcal{E}}{\mathrm{d}t} =& \frac{\mathcal{E}(T) - \mathcal{E}(0)}{T} \ \textit{by assumption} \\
A^2 =& (\dot{x})^2 + x^2 = 2\mathcal{E} \\
\implies A\frac{\mathrm{d}A}{\mathrm{d}t} =& \epsilon A \frac{b_1(A)}{2} + O(\epsilon^2)
\end{aligned}$$

Note that if is $h(x,y)$ analytic and converges for all x, y then, by Theorem 2.11, $b_1(A)$ is an odd power series of A that converges for all A. But then $\frac{\mathrm{d}\mathcal{E}}{\mathrm{d}t} = \epsilon f(\mathcal{E})$ where f is a power series that converges for all $\mathcal{E}$ and is such that $f(0) = 0$. Since f converges for all $\mathcal{E}$ it is continuously differentiable for all $\mathcal{E}$. Then, by the Existance and Uniqueness theorem and $f(0) = 0$, if $\mathcal{E}(t_0) = 0$ then $\mathcal{E}(t) = 0$ for all times t. This means that if $\mathcal{E}(t_0) \neq 0$ for some t_0, then $\mathcal{E}(t) > 0$ for all times t. Since $A^2 = 2\mathcal{E}$ and $b_1(0) = 0$ we can divide through the last equation by A, no matter the initial conditions.

This gives us the identical equation for A that we derived from two-timing. Nevertheless, though this equation's origin was slightly mysterious in the context of two-timing, here we directly see that to $O(\epsilon)$ it is b_1 that controls the change in energy of the system. □

Proposition 4.3. Suppose that $x(t,\epsilon) = A\cos([1+\epsilon\omega]t) + c = A\cos(t+\phi) + c$, where $c = \frac{\epsilon}{2\pi}\int_0^{2\pi} h(A\cos\theta, -A\sin\theta)\mathrm{d}\theta = \epsilon\frac{a_0}{2}$ is the 'center' of the oscillations and the constant term of x_1. Then we can see that $A\frac{\mathrm{d}\phi}{\mathrm{d}t} = \epsilon a_1(A)/2 + O(\epsilon^2)$.

Proof. Note that since we have assumed $x(t,\epsilon)$ has the form

$$A\cos([1+\epsilon\omega]t) + c \textit{ and } A\cos(t+\phi) + c$$

then $\frac{\mathrm{d}\phi}{\mathrm{d}t}$ must be equal to $\epsilon\omega$. We also see that:

$$\begin{aligned}
\frac{1}{2\pi}\int_0^{2\pi} C\cos([1+\epsilon\omega]t)\mathrm{d}t =& \frac{C}{2\pi(1+\epsilon\omega)}\sin(2\pi\epsilon\omega) \\
=& \frac{C}{2\pi}\left(1 - \epsilon\omega + O(\epsilon^2)\right)\left(2\pi\epsilon\omega + O(\epsilon^3)\right) \\
=& C\epsilon\omega + O(\epsilon^2).
\end{aligned}$$

Then, by the assumptions on $x(t,\epsilon)$, we have that $\frac{1}{2\pi}\int_0^{2\pi} x(t,\epsilon)\mathrm{d}t = A\epsilon\omega + c + O(\epsilon^2) = A\frac{\mathrm{d}\phi}{\mathrm{d}t} + c + O(\epsilon^2)$. Then we have that:

$$\begin{aligned}
\frac{1}{2\pi}\int_0^{2\pi} x(t,\epsilon)\mathrm{d}t =& \frac{1}{2\pi}\int_0^{2\pi} x_0 \ \mathrm{d}\theta + \epsilon\frac{1}{2\pi}\int_0^{2\pi} x_1 \ \mathrm{d}\theta + O(\epsilon^2) \\
=& \epsilon\frac{a_0}{2} + \epsilon\frac{1}{2\pi}\int_0^{2\pi} a_1\frac{\theta}{2}\sin\theta - b_1\frac{\theta}{2}\cos\theta \ \mathrm{d}\theta + O(\epsilon^2) \\
=& \epsilon\left(\frac{a_1}{2} + \frac{a_0}{2}\right) + O(\epsilon^2) \\
\implies A\frac{\mathrm{d}\phi}{\mathrm{d}t} =& A\epsilon\omega = \epsilon\frac{a_1(A)}{2} + O(\epsilon^2).
\end{aligned}$$

Note that the rest of the terms in x_1 canceled because they were of the form $\cos(n\theta)$ or $\sin(n\theta)$ for $n \in \mathbb{N}\setminus\{1\}$. It is now more physically obvious that the function of a_0 is to shift the center, while the function of a_1 is to shift the frequency. $\square$

Looking to the figure on the right we see confirmation of this. The red function represents the difference between a numerically integrated trajectory for $h_{\text{red}}(x,y) = x^3$ with initial conditions $\dot{x}_0 = 0$ and $x_0 = 1$ and $\cos x$. The blue function represents the difference between a numerically integrated trajectory for $h_{\text{blue}}(x,y) = x^2$ with identical inital conditions and $\cos x$ (where both are adjusted by their 'center'). Notice that an even forcing function dumps more energy into either positive or negative x values, only affecting the 'center' of the function rather than the frequency, at least as seen in the diagram, to $O(\epsilon^2)$.

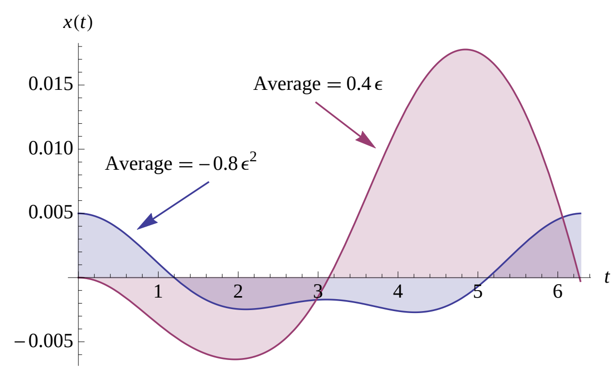


FIGURE 2. Changes from a_1 with $h_{\text{blue}}(x,y) = x^2$, $h_{\text{red}}(x,y) = x^3$, and $\epsilon = .01$

In a similar manner, the y^2 damping does not shift the energy by a significant amount. It merely dumps more energy into movement in one direction, rather than the other. By Proposition 4.3 this acts only to change the center of oscillations. Thus we see intuitively that even functions' approximations do not fail until $t \approx O\left(\frac{1}{\epsilon^2}\right)$ because they don't shift total frequency or energy (at least to $O(\epsilon)$), instead they shift frequency from one side to the other, or energy from one direction to another. This has the effect of only changing the 'center' of oscillations, not something that creates secular terms like those arising from odd forcing functions.

5. Problems with Two-Timing

Now that we see the equivalence of two-timing and averaging theory to $O(\epsilon)$ in $\frac{\mathrm{d}A}{\mathrm{d}t}$ and $\frac{\mathrm{d}\phi}{\mathrm{d}t}$, we can begin to address difficulties with the methods. Example 3.1 highlights well the first of these. It was by no means easy to find $A(T_1)$ in the above example, and in fact (13) can give very complicated non-linear equations for A and ϕ, not greatly simplifying the problem at all. However, we have tailored our approximation to be qualitatively clear after only a finite number of terms, as opposed to the infinite number required by regular perturbation theory, so it is not surprising that this increase in clarity comes with a corresponding increase in difficulty. Thus, we have our first problem, which can be annoying, but is hardly fatal to the method; difficulty in evaluating $A(T_1)$ and $\phi(T_1)$.

The second, more fundamental, problem is best illustrated with another example.

Example 5.1. Let $h(x, y) = -\frac{8}{3}y^3$, with initial conditions $x(0) = 1$ and $\dot{x}(0) = 0$. Then $a_1(A) = 0$ and $b_1(A) = 2A^3$. Solving (13) gives that $x(t) = \frac{1}{\sqrt{1-2\epsilon t}} \cos t + O(\epsilon)$. Thus, two-timing gives a solution which goes to infinity as $t \to \frac{1}{2\epsilon}$ (for small ϵ).

The blue trajectory shown on the right is the numerically integrated trajectory for this problem, whereas the red functions represent the envelope predicted by two timing $\left(\pm\frac{1}{\sqrt{1-2\epsilon t}}\right)$. Due to sensitivity in the numerical integration it is difficult to tell whether or not the exact solution goes to infinity in finite time, as the approximation predicts. However, following the lead of J.D. Cole and J. Kevorkian, it is possible to introduce a somewhat ad hoc rescaling of variables to suggest that two-timing fails in this case.

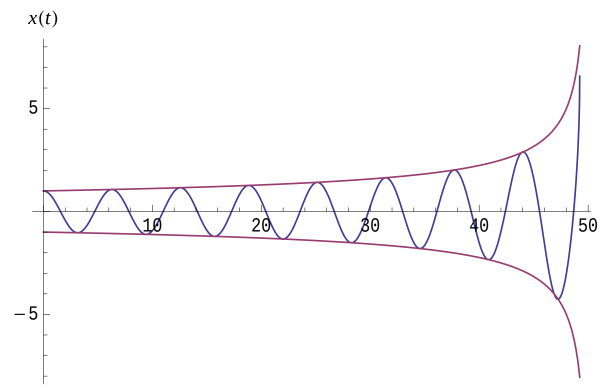


Figure 3. Failure of Two-Timing $h(x, y) = -\frac{8}{3}y^3$ and $\epsilon = .01$

Let $\tilde{x} = \epsilon^{\alpha} x$, where α is contant in time. Then our weakly nonlinear oscillator equation is transformed to:

$$\epsilon^{-\alpha}\ddot{\tilde{x}} - \frac{8}{3}\epsilon^{1-3\alpha}(\dot{\tilde{x}})^3 + \epsilon^{-\alpha}\tilde{x} = 0.$$

Let $\alpha = \frac{1}{2}$, and this transforms to $\ddot{\tilde{x}} - \frac{8}{3}(\dot{\tilde{x}})^3 + \tilde{x} = 0$, which is the fully nonlinear problem. In other words, if $x = O(\epsilon^{-1/2})$, then we are no longer dealing with a perturbation problem, but one that is fully nonlinear. Consider the above example, and let $x = \sqrt{\frac{a}{\epsilon}}$, where a is some constant. The approximation reveals that this x value will be reached when $t = \frac{1}{2\epsilon} - \frac{1}{2a}$. Then, unsurprisingly considering the rescaling, when $x = O(\epsilon^{-1/2})$ we have that t is very close to the asymptote of the approximation. So as the approximation is blowing up in finite time, we are no longer able to discern whether or not it is a valid approximation because the problem has transformed to a fully nonlinear one. By looking at the graph we can also see that the rescaling is correct in predicting that the approximation will be good at least until $x = O(\epsilon^{-1/2})$.

Physically, however, there is a deeper problem with this example. As can be seen by looking at the envelope function predicted, if the initial conditions given are even slightly off then the approximation will blow up before or after the exact solution begins to really increase. This means that for a period of time the difference between the approximation and the exact solution will be infinite. Thus the correctness of this approximation appears to be very sensitive to intial conditions. In fact this provides a clearer test of when two-timing should be taken with a grain of salt, than does scaling.

Example 5.2 (Extended 5.1)**.** Let $\lambda = \frac{x}{A(T_1)}$, and suppose that λ is in a nearly periodic orbit centered at 0, with radius B. Then to $O(\epsilon)$, $\frac{\mathrm{d}B}{\mathrm{d}t} = \frac{\epsilon}{2A}\left(b_1(AB) - b_1(A)B\right)$. Let $f(A) = \frac{\mathrm{d}b_1}{\mathrm{d}A} - \frac{b_1(A)}{A}$. If $f(A) > 0$ then the solution exhibits the sensitivity to intial conditions discussed above, whereas if $f(A) < 0$ then the approximation converges to the exact solution even when given slightly different initial values than the exact solution (at least to $O(\epsilon)$).

Proof. We have that $\dot{x} = \dot{A}\lambda + A\dot{\lambda}$ and $\ddot{x} = \ddot{A}\lambda + 2\dot{A}\dot{\lambda} + A\ddot{\lambda}$. We also know from averaging theory that $\dot{A} = \epsilon b_1(A)/2 + O(\epsilon^2)$. This means that $\ddot{A} = \epsilon/2 \frac{\partial b_1}{\partial A}\frac{\mathrm{d}A}{\mathrm{d}t} + O(\epsilon^2) = \epsilon^2/4\frac{\partial b_1}{\partial A} b_1(A) = O(\epsilon^2)$. Inserting these into the original equation gives:

$$A\ddot{\lambda} + \epsilon b_1(A)\dot{\lambda} + \epsilon h(A\lambda, A\dot{\lambda} + \epsilon b_1(A)\lambda/2) + A\lambda + O(\epsilon^2) = 0.$$

However, since h is independent of ϵ, the $\dot{A}\lambda$ term within h will lead to, at most, $O(\epsilon^2)$ contributions. But then we have

$$A\ddot{\lambda} + \epsilon\left[b_1(A)\dot{\lambda} + h(A\lambda, A\dot{\lambda})\right] + A\lambda + O(\epsilon^2) = 0$$

From Proposition 4.2 we established that if $h(x, y)$ is analytic, then A would not be zero unless it was initially zero, so we can divide the equation through by this term. Then if we assume that λ is nearly periodic, with radius B, two-timing and averaging theory give us that:

$$\begin{aligned}\frac{\mathrm{d}B}{\mathrm{d}t} =& \frac{\epsilon}{2A}\frac{1}{\pi}\int_0^{2\pi} h(AB\cos\theta, -AB\sin\theta)\sin\theta - b_1(A)B\sin^2\theta \,\mathrm{d}\theta + O(\epsilon^2)\\ =& \frac{\epsilon}{2A}\left(b_1(AB) - b_1(A)B\right) + O(\epsilon^2) \text{ where } A, B \text{ treated as constant.}\end{aligned}$$

Note that the series expansions of both methods ensured that ignoring the $O(\epsilon^2)$ terms in the differential equation was justified. Inserting $B = 1$, we see that at least to $O(\epsilon)$ this gives an equilibrium point at 1. Seeing that we deliberately ensure $A(0) = x(0)$, and thus that $\lambda(0) = 1$, we see that the solution begins on this equilibrium point. However, we are unsure whether or not any $O(\epsilon^2)$ changes will perturb the solution, so we seek to understand the stability of the equilibrium point when $B = 1$. Thus we differentiate with respect to B to see that if:

$$\begin{aligned}f(A) :=& \epsilon\left[\frac{\mathrm{d}b_1}{\mathrm{d}(AB)}(AB) - \frac{b_1(A)}{A}\right]_{B=1}\\ =& \epsilon\left[\frac{\mathrm{d}b_1}{\mathrm{d}A} - \frac{b_1(A)}{A}\right] > 0 \ B = 1 \text{ unstable}\\ f(A) <& 0 \ B = 1 \text{ stable.}\end{aligned}$$

Note that we are perfectly conscious that A is a function of time, the reason we nonetheless arrive at this conclusion is due to its extremely slow change. For those still squeamish, hopefully Example 5.3 will demonstrate the usefulness of $f(A)$, as well as our lack of pretense about its accuracy. ∎

Consider $b_1(A) = A^3$, as in the earlier example. Then we see that $f(A) = 2\epsilon A^2 > 0$ (since A is never zero unless it is initially zero) and therefore that $B = 1$ is unstable. Therefore, if the exact solution represents some physical phenomenon, and we measure the initial amplitude (for use in our approximation) slightly imperfectly, we have that $B \neq 1$

initially. However, in this case, this small error will compound rapidly, since $B = 1$ is an unstable equilibrium point for $\frac{\mathrm{d}B}{\mathrm{d}t}$. We give one final example to illustrate the usefulness of considering λ and $f(A)$.

Example 5.3. Let $h(x, y) = 2y^3\left(x^2 - \frac{2}{3}\right)$ be the equation for a 'modified' Van der Pol oscillator, one with a cubic damping term instead of a linear damping term. Furthermore, let $x(0) = 1$ and $\dot{x}(0) = 0$. It can be checked that $b_1(A) = A^3\left(1 - \frac{1}{4}A^2\right)$, so that it is clear that just as in the ordinary Van der Pol oscillator, $A = 2$ is a stable equilibrium solution. This then gives that $f(A) = \epsilon A^2(2 - A^2)$. But what are we to make of this? Since A is itself a function of time, this seems to imply that the stability of $B = 1$ changes as time goes on. We want to check whether or not this is true.

On the right is shown a test of this.

We have used Mathematica to numerically compute $A(t)$, with initial value $A(0) = \frac{9}{10}$. Since $\frac{\mathrm{d}A}{\mathrm{d}t} = \epsilon b_1(A)/2 + O(\epsilon^2)$, we expect A to grow from its initial value at $A(0) = .9$ and approach the steady state solution $A = 2$ (as mentioned above, $A = 2$ is a stable equilibrium solution). But above, we saw that $f(A) = \epsilon A^2(2 - A^2)$. Since we expect $f(A)$ to govern the stability of $B = 1$, we see that initially $f(A(0)) = .96\epsilon > 0$ or that at $t = 0$, $B = 1$ is unstable. However, as $t \to \infty$ we noticed that $A \to 2$, so $\lim_{t\to\infty} f(A(t)) = -8\epsilon < 0$, or for large times $B = 1$ is stable. We expect this change in stability to occur when $A \approx \sqrt{2}$, since $f(\sqrt{2}) = 0$.

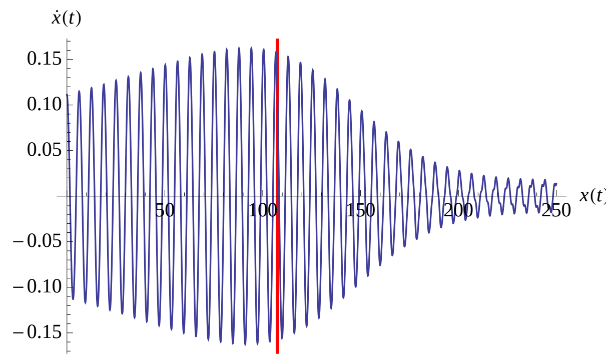


Figure 4. Change of B
$h(x, y) = 2y^3\left(x^2 - \frac{2}{3}\right)$ and $\epsilon = .01$

To test this let t^* be the time such that, for our numerically computed $A(t)$, we have $A(t^*) = 2$. We have also used Mathematica to numerically compute $x(t)$ from the oscillator equation itself, where we have specified that $x(0) = 1$ and $\dot{x}(0) = 0$. Since we also computed $A(t)$, we can then plot $\lambda(t) = x(t)/A(t)$. We notice that given the initial conditions on x and A, we have $\lambda(0) = \frac{10}{9}$ and $\dot{\lambda}(0) = 0$. Given Example 5.2 we then expect that $\lambda(t) \approx B(t)\cos t$, where to $O(\epsilon)$, $B(t)$ is given by $\dot{B} = \epsilon\left(b_1(AB) - b_1(A)B\right)/2A$ and $B(0) = \frac{10}{9}$. But since we are considering the stability of $B = 1$, it might be useful to consider $\lambda(t) - \cos t \approx (B(t) - 1)\cos t$. This is what is plotted in blue in Figure 4. The red line is placed at $t = t^*$, or at the time when we expect $B = 1$ to transform from an unstable equilibrium to a stable equilibrium.

From the figure we can see that, as expected, we initially have that $B(t)$ is increasing, as it is repelled by the unstable equilibrium at $B = 1$. Also as expected, we have that as $t \to \infty$, $B(t) \to 1$ since $B = 1$ has become a stable attractor for large times. In fact, the switch between $B = 1$ being stable and being unstable occured only slightly before $t = t^*$, when we expected it to occur. However, this slight discrepancy hints that $f(A)$ is not very useful quantitatively to predict the exact A value for such a shift to occur. Indeed, there is no reason it should be, in deriving it we more or less assumed that A was constant. The reason it worked as well as it did was likely because B was so close to 1. However, with all the uncertainty surrounding $f(A)$ we can still use it to derive at least one useful relation, by only looking at it for enormous solutions and already stable solutions.

Corollary 5.4. Suppose that $b_1(A)$ produces a trajectory which goes to infinity, and $\lim_{A\to\infty} A^{-p}b_1(A) = k$ for some $k > 0 \in \mathbb{R}$, where $p \in \mathbb{R}$ is such that $p > 1$. Then $B = 1$ is unstable. Suppose instead that $b_1(A)$ produces a trajectory which approaches a stable equilibrium point. Then $B = 1$ is stable, at least close to this point.

Proof. The first assumption gives us that as $t \to \infty$, $A \to \infty$. But then we see that since $\lim_{A\to\infty} A^{-p}b_1(A) = k$ for some nonzero k, that as $t \to \infty$, $b_1(A)$ goes to kA^p. But then we see that $\lim_{A\to\infty} f(A) = k(p-1)\epsilon \lim_{A\to\infty} A^{p-1}$. Since A began and stays positive, and $p > 1$ we see that $\lim_{A\to\infty} f(A) > 0$. Thus, we see that as $t \to \infty$, $B = 1$ becomes more and more unstable (since $f(A)$ converges to some function of A).

Let A^* be a stable equilibrium point of $\frac{\mathrm{d}A}{\mathrm{d}t}$. This means that $b_1(A^*) = 0$ and $\epsilon\frac{\mathrm{d}b_1}{\mathrm{d}A}(A^*) < 0$. However, we then immediately see that for such a point $f(A^*) = \epsilon\frac{\mathrm{d}b_1}{\mathrm{d}A}(A^*) < 0$. Thus, as seen in Example 5.3, as A approaches a stable equilibrium point, $B = 1$ itself becomes a stable equilibrium point. ☐

This Corollary allows us to see that for most trajectories, if h is such that the trajectory is pushed to infinity, then it will ensure that a slight mistake in initial conditions will doom the approximation (the notable exception is forcing proportional to the first power of oscillator velocity). It also allows us to see that for most cases, two-timing gives excellent approximations. Provided that A approaches an equilibrium point, Bs convergence to 1 is almost guaranteed.

Further research might be conducted into getting a more exact sense of when $B = 1$ is a stable point and investigating more fully cases like Example 5.3. Some other interesting cases might also be considered. Suppose that $b_1(A) = -\sin A$, and suppose that $A(0) = \pi-\delta$ while $x(0) = \pi + \delta$, where $0 < \delta \ll 1$. Then $B(0) = 1 + 2\delta/\pi + O(\delta^2)$. However, because A and x started on opposing sides of an unstable equilibrium point, we see that $A \to 0$, whereas $x \to 2\pi$. This means that $B \to \infty$, despite the difference of only 2δ in initial conditions. Research might be conducted into more fully understanding when unstable equilibria affect this sensitivity to initial conditions. Finally, this paper focused greatly on the difference between even and odd forcing functions. Much effort was devoted to understanding the amplitude and frequency shifts given by odd forcing functions, whereas even forcing functions were more or less neglected. A more computationally inclined researcher might expand our two timing considerations to the super-slow time $T_2 := \epsilon^2 t$ to find differential equations governing the $O(\epsilon^2)$ changes in amplitude and frequency of an evenly forced oscillator.

Resources. All graphics in this paper were created using Wolfram Mathematica.

Acknowledgments. It is a pleasure to thank my mentor, Ben Seeger, for his help with this paper. His suggestions for avenues of exploration, his helpful criticisms of the source material and derivations done on my own, and his leniency with the level of rigor he expected allowed me to explore this problem's many interesting solution methodologies to the depth I desired. I would also like to thank Peter May for setting up the REU program, allowing me to learn a great deal of math in a short amount of time and get paid to do it! Finally, I'd

like to thank my parents for supporting me during the program, in particular, and during my academic life, in general.

References

[1] D.W. Jordan and P. Smith. Nonlinear Ordinary Differential Equations. Oxford University Press, 1999.

[2] J. Kevorkian and J.D. Cole. Multiple Scale and Singular Perturbation Methods. Springer-Verlag, 1996.

[3] J. Kevorkian. The Uniformly Valid Asymptotic Approximations to the Solutions of Certain Non-linear Ordinary Differential Equations. http://thesis.library.caltech.edu/5116/1/Kevorkian

[4] James D. Meiss. Differential Dynamical Systems. Society for Industrial and Applied Mathematics, 2007.

[5] Steven H. Strogatz. Nonlinear Dynamics and Chaos. Perseus Books Publishing, 1994.