1 Averaging-Overview

A large class of perturbed problems can be posed in the standard form

$$\frac{dx}{dt} = \varepsilon F(x,\theta;\varepsilon) \qquad x \in \mathbb{R}^n \tag{1}$$

$$\frac{d\theta}{dt} = \omega(\theta) + \varepsilon G(x,\theta;\varepsilon) \qquad \theta \in \mathbb{R}^m \tag{2}$$

where both F and G are 2π-periodic in θ_i, $\theta = (\theta_1, ...\theta_m)$.

If $F(x,t;\varepsilon)$ is 2π-periodic in t an alternate "standard form" is

$$\frac{dx}{dt} = \varepsilon F(x,t;\varepsilon) \qquad x \in \mathbb{R}^n \qquad x(0) = x_0 \tag{3}$$

The connection can be seen if one rewrites (3) as the system

$$\frac{dx}{dt} = \varepsilon F(x,\theta;\varepsilon) \tag{4}$$

$$\frac{d\theta}{dt} = 1 \tag{5}$$

For (3) define the "average"

$$\hat{F}(x,\varepsilon) \equiv \int_0^{2\pi} F(x,t,\varepsilon)dt \tag{6}$$

Let y satisfy the initial value problem

$$\frac{dy}{dt} = \varepsilon \hat{F}(y,\varepsilon) \qquad y(0) = x_0 \tag{7}$$

Then "averaging" theorems prove, under certain conditions,

$$\| x(t) - y(t) \| = O(\varepsilon) \qquad t = O\left(\varepsilon^{-1}\right) \tag{8}$$

The method of averaging thus involves

a) A transformation to standard form

b) Solving the averaged equations

Example: Nearly harmonic

$$\ddot{x} + x = \varepsilon f(x, \dot{x}) \qquad \dot{(\ )} = \frac{d}{dt}(\ ) \tag{9}$$

Unperturbed solution is

$$x_0(t) = A_0 \cos(t + \phi_0) \tag{10}$$

where A_0 and ϕ_0 are constants. Use a variation of parameters based on the amplitude-phase form of the solution $x_0(t)$:

$$x(t) = A(t)\cos(t + \phi(t)) \equiv A\cos\psi \tag{11}$$

Equation (9) imposes one condition relating the two functions A, ϕ. We impose a second condition on A and ϕ which guarantees

$$\dot{x}(t) = -A(t)\sin(t + \phi(t)) = -A\sin\psi \quad , \tag{12}$$

that is, a condition which makes differentiation of x appear as if A and ϕ were constant. Explicitly, we are therefore imposing the condition

$$\frac{d}{dt} A\cos\psi = -A\sin\psi \tag{13}$$

Using (11)-(12) in (9), it is readily seen that

$$-\dot{A}\sin\psi - A\dot{\phi}\cos\psi = \varepsilon f(A\cos\psi, -A\sin\psi) \tag{14}$$

Expanding out the second condition (13) one gets

$$\dot{A}\cos\psi - \dot{\phi}A\sin\psi = 0 \tag{15}$$

Solving (14)-(15) for $\dot{A}$ and $\dot{\phi}$ yields

$$\dot{A} = \varepsilon F_1(A, \phi, t) \equiv -\varepsilon \ \sin\psi \ f(A\cos\psi, -A\sin\psi) \tag{16}$$

$$\dot{\phi} = \varepsilon F_2(A, \phi, t) \equiv -\varepsilon \ \frac{1}{A} \ \cos\psi \ f(A\cos\psi, -A\sin\psi) \tag{17}$$

Clearly, F_k is 2π-periodic in t, thus (16)-(17) is in standard form.

Example: weakly nonlinear problems

Let

$$\dot{x} = A(t)x + \varepsilon g(x,t) \qquad x(0) = x_0 \in \mathbb{R}^n \tag{18}$$

The problem is said to be weakly nonlinear because the unperturbed problem

$$\dot{z} = A(t)z \tag{19}$$

is linear. The unperturbed problem has n independent solutions $z_i(t)$ from which a fundamental solution matrix

$$\Psi(t) = [z_1, ...z_n] \tag{20}$$

can be formed. Furthermore, Ψ is invertible since the columns are independent. If we define y via

$$x = \Psi(t)y \tag{21}$$

and substitute (21) into (25) we get

$$\dot{x} = \dot{\Psi}y + \Psi\dot{y} = A\Psi y + \varepsilon g(\Psi y, t) \tag{22}$$

But, since $\dot{\Psi} = A\Psi$ and Ψ is invertible,

$$\dot{y} = \varepsilon \ \Psi^{-1}(t) \ g(\Psi y, t) \tag{23}$$

If all z_i and g are T-periodic in t, the latter system is in standard form.

Example: strongly nonlinear problems

$$\dot{x} = f(x,t) + \varepsilon g(x,t,\varepsilon) \qquad x(0) = x_0 \in \mathbb{R}^n \tag{24}$$

Is "strongly nonlinear" because the unperturbed problem is nonlinear. Assume the unperturbed problem

$$\dot{y} = f(y,t) \qquad y(0) = Z \tag{25}$$

has a (explicitly) known solution

$$y = Y(t, Z) \tag{26}$$

Seek a solution of the perturbed problem of the form

$$x = Y(t, z(t)) \tag{27}$$

where $z(t) = (z_1(t), ...z_n(t))$ is some vector valued function. Substitution of this expression into (24) yields

$$\frac{\partial y}{\partial t} + \frac{\partial y}{\partial z}\frac{dz}{dt} = f(y,t) + \varepsilon g(y,t,\varepsilon) \tag{28}$$

where

$$\frac{\partial y}{\partial z} = \left[\frac{\partial Y_i}{\partial Z_j}\right]_{ij} \tag{29}$$

Since

$$\frac{\partial y}{\partial t} = f(y,t) \tag{30}$$

this simplifies to

$$\frac{\partial y}{\partial z}\frac{dz}{dt} = \varepsilon g(y,t,\varepsilon) \tag{31}$$

If $\frac{\partial y}{\partial z}$ is nonsingular, then

$$\frac{dz}{dt} = \varepsilon \left(\frac{\partial y}{\partial z}\right)^{-1} g(y,t,\varepsilon) \tag{32}$$

may be in standard form, depending on f, g.

Example: strongly nonlinear Kuzmak-Luke

Kuzmak and Luke studied the perturbed Hamiltonian problem

$$\ddot{x} + g(x) = \varepsilon f(x, \dot{x}) \qquad , \qquad \dot{(\ )} = \frac{d}{dt}(\ ) \tag{33}$$

This problem can be posed as the system

$$\begin{aligned} \dot{x}_1 &= x_2 & (34) \\ \dot{x}_2 &= -g(x_1) + \varepsilon f(x_1, x_2) & (35) \end{aligned}$$

The leading problem is integrable:

$$\begin{aligned} \dot{y}_1 &= y_2 & (36) \\ \dot{y}_2 &= -g(y_1) & (37) \end{aligned}$$

Its Hamiltonian $H(y_1, \dot{y}_1)$ is constant and equal to the energy E

$$H(y_1, \dot{y}_1) = E = \frac{1}{2}\dot{y}_1^2 + V(y_1) \qquad , \qquad V(y) = \int^y g(s)\ ds \tag{38}$$

Thus, the solution is given implicitly by

$$t - \phi = \int_0^{y_1} \frac{ds}{\sqrt{2(E - V(s))}} \tag{39}$$

Depending on g and initial conditions, the solution $y_1(t)$ will be periodic with period T. We restrict our attention to these cases.

Specifying initial conditions for (y_1, y_2) is equivalent to specifying the values of the constants E and ϕ. Thus, if the integral in (39) can be evaluated and inverted there are functions Y_1, Y_2 such that

$$\begin{aligned} y_1(t) &= Y_1(t, E, \phi) & (40) \\ y_2(t) &= Y_2(t, E, \phi) & (41) \end{aligned}$$

The $y_k(t)$ are T-periodic in t and there is a function $\bar{T}$ such that $T = \bar{T}(E)$:

$$\bar{T} = \int_{C(E)} \frac{ds}{\sqrt{2(E - V(s))}} \tag{42}$$

where $C(E)$ is the closed curve in the (y_1, y_2)-plane corresponding to the periodic orbit with energy E.

Since this problem is strongly nonlinear, the methods of the previous section should apply. Here $Z = (E, \phi)$. Thus, we now impose a time dependence on Z to transform the model to standard form. In particular,

$$\begin{aligned} x_1(t) &= Y_1(t, E(t), \phi(t)) && (43) \\ x_2(t) &= Y_2(t, E(t), \phi(t)) && (44) \end{aligned}$$

Thus,

$$\begin{aligned} \dot{x}_1(t) &= \frac{\partial Y_1}{\partial t} + \frac{\partial Y_1}{\partial E}\dot{E} + \frac{\partial Y_1}{\partial \phi}\dot{\phi} = Y_2 && (45) \\ \dot{x}_2(t) &= \frac{\partial Y_2}{\partial t} + \frac{\partial Y_2}{\partial E}\dot{E} + \frac{\partial Y_2}{\partial \phi}\dot{\phi} = -g(Y1) + \varepsilon f(Y_1, Y_2) && (46) \end{aligned}$$

Using the fact that (Y_1, Y_2) solves the leading problem this reduces to

$$\begin{aligned} \frac{\partial Y_1}{\partial E}\dot{E} + \frac{\partial Y_1}{\partial \phi}\dot{\phi} &= 0 && (47) \\ \frac{\partial Y_2}{\partial E}\dot{E} + \frac{\partial Y_2}{\partial \phi}\dot{\phi} &= \varepsilon f(Y_1, Y_2) && (48) \end{aligned}$$

Solving these for $\dot{E}$ and $\dot{\phi}$ it is clear there are functions F_1 and F_2 such that

$$\begin{aligned} \dot{E} &= \varepsilon F_1(E, \phi, t) && (49) \\ \dot{\phi} &= \varepsilon F_2(E, \phi, t) && (50) \end{aligned}$$

However, recall that Y_k are periodic in t but have a period $T = \bar{T}(E)$. That is, (49)-(50) are not in standard form since the functions F_k do not have a fixed period (2π). This example illustrates that not all variation of parameter based transformations of oscillatory problems will yield a standard form. Here, the method of averaging cannot be applied to (49)-(50). All is not lost, however. There are different ways to transform perturbed hamiltonian problems to standard form using "canonical" transformations. In particular, the choice of "action" J rather than E essentially solves the dilema. Action is defined as

$$J = J(E) = \int_{C(E)} \sqrt{2(E - V(s))} \quad , \qquad (51)$$

is a function of E and geometrically is the area enclosed by the periodic orbit. It is a long story, but by introducing an "angle" variable $Q = E'(J)t$, it can be shown that equations for (J, Q) analagous to (49)-(50) are in fact in standard form.

Van der Pol Equation

$$\ddot{x} + \varepsilon(x^2 - 1)\dot{x} + x = 0 \tag{52}$$

Transformation to standard form via

$$x(t) = A(t)\cos(\psi) \quad , \quad \psi = t + \phi(t) \tag{53}$$

$$\dot{x}(t) = -A(t)\sin(\psi) \tag{54}$$

Yields the standard form:

$$\dot{A} = \varepsilon \sin(\psi)(1 - A^2\cos^2(\psi))(A\sin(\psi)) \tag{55}$$

$$\dot{\phi} = \varepsilon \frac{1}{A}\cos(\psi)(1 - A^2\cos^2(\psi))(A\sin(\psi)) \tag{56}$$

Define the average notation

$$< * >= \frac{1}{2\pi}\int_0^{2\pi} * dt \tag{57}$$

and noting

$$\begin{array}{ll} < \sin^2(\psi) >= \frac{1}{2} & < \sin(\psi)\cos(\psi) >= 0 \\ < \sin^2(\psi)\cos^2(\psi) >= \frac{1}{8} & < \sin(\psi)\cos^3(\psi) >= 0 \end{array} \tag{58}$$

we obtain the averaged equations

$$\dot{\hat{A}} = \frac{\varepsilon}{8}\hat{A}(4 - \hat{A}^2) \tag{59}$$

$$\dot{\hat{\phi}} = 0 \tag{60}$$

The solution of these eqautions will of course depend on initial conditions for $A(0), \phi(0)$. But notice if $\hat{A}(0) = 2$, $\hat{A}(t) = 2$ for all t. Since $A = \hat{A}$ for $t = O(\varepsilon^{-1})$, we obtain an asymptotic approximation to the limit cycle of the original problem, namely,

$$x(t) \sim 2\cos(t) \qquad t = O(\varepsilon^{-1}) \tag{61}$$

Moreover, from the sign of the derivative $\dot{\hat{A}}$ for $\hat{A} > 2$ and $\hat{A} < 2$ we observe the limit cycle is stable. Though technically we have only found an asymptotic approximation the limit cycle, the use of averaging in this manner can be formalized to proving the existence of limit cycles for sufficiently small ε.

Mathieu Equation

$$\ddot{x} + (1 + 2\varepsilon cos(2t))x = 0 \qquad x(0) = a \qquad \dot{x}(0) = 0 \tag{62}$$

This equation models oscillators that are subject to frequency modulation. Here, the frequency $\omega^2 = (1 + 2\varepsilon\cos(2t))$ is not a function of a slow time. Conversion to standard form does not use the the amplitude-phase forms but rather

$$x(t) = y_1(t)\cos(t) + y_2(t)\sin(t) \qquad , \tag{63}$$

$$\dot{x}(t) = -y_1(t)\sin(t) + y_2(t)\cos(t) \tag{64}$$

To transform the system to standard form, substitute (63) into (62) to obtain one equation for $\dot{y}_1$ and $\dot{y}_2$. A second equation is obtained by requiring the time derivative of the right side of (63) to equal the right side of (64). Solving for $\dot{y}_1$ and $\dot{y}_2$ we obtain the standard form:

$$\dot{y}_1 = 2\varepsilon\sin(t)\cos(2t)(y_1\cos(t) + y_2\sin(t)) \qquad y_1(0) = a \tag{65}$$

$$\dot{y}_2 = 2\varepsilon\cos(t)\cos(2t)(y_1\cos(t) + y_2\sin(t)) \qquad y_2(0) = 0 \tag{66}$$

The averaged equations are then

$$\dot{\hat{y}}_1 = -\frac{1}{2}\hat{y}_2 \qquad \hat{y}_1(0) = a \tag{67}$$

$$\dot{\hat{y}}_2 = -\frac{1}{2}\hat{y}_1 \qquad \hat{y}_2(0) = 0 \tag{68}$$

Solving these one finds

$$\hat{y}_1(t) = a\cosh(\varepsilon t) \qquad , \qquad \hat{y}_2(t) = -a\sinh(\varepsilon t) \tag{69}$$

so that

$$x(t) = \hat{y}_1(t)\cos(t) + \hat{y}_2(t)\sin(t) + O(\varepsilon) \tag{70}$$

for $t = O(\varepsilon^{-1})$.

An Averaging Theorem

There are many many averaging theorems. Here we state and prove one taken from Verhulst (1985). First, let

$$\dot{x} = \varepsilon f(x,t) + \varepsilon^2 g(x,t,\varepsilon) \quad , \quad x(0) = x_0 \in \mathbb{R}^n \tag{71}$$

$$\dot{y} = \varepsilon \hat{f}(y) \qquad y(0) = x_0 \in \mathbb{R}^n \tag{72}$$

where f is T-periodic in t (and T is ε-independent) and

$$\hat{f}(y) = \frac{1}{T}\int_0^T f(y,t)dt \tag{73}$$

Theorem *Let $x, y, x_0 \in D \subset \mathbb{R}^n$, $t \geq 0$ and assume*

(A1) $f, g, \frac{\partial f}{\partial x}$ are defined, continuous and bounded by an ε-independent constant M on $D \times \mathbb{R}^+$,

(A2) f is Lipschitz continuous for $x \in D$,

(A3) $y(t)$ is contained in a proper subset of D,

Then, $x(t) - y(t) = O(\varepsilon)$ for $t = O(\varepsilon^{-1})$.

Proof: Define

$$u(x,t) = \int_0^t (f(x,s) - \hat{f}(x))ds \tag{74}$$

Since u is T-periodic in t

$$\| u(x,t) \| \leq \max_{t\in[0,T]} \| \int_0^T f(x,s)ds - \int_0^T \hat{f}(x)ds \| \tag{75}$$

$$\leq \max_{t\in[0,T]} \| \int_0^T f(x,s)ds \| + \max_{t\in[0,T]} \| \int_0^T \hat{f}(x)ds \| \tag{76}$$

$$\leq 2TM \tag{77}$$

making u uniformly bounded on $D \times \mathbb{R}^+$.

Next, define the "near-identity" transformation

$$x(t) = z(t) + \varepsilon u(z(t),t) \tag{78}$$

Since $\| u(x,t) \| \leq 2MT$ uniformly, $x(t) - z(t) = O(\varepsilon)$. Now note,

$$\dot{x} = \dot{z} + \varepsilon u_t(z,t) + \varepsilon u_x(z,t)\dot{z} = \varepsilon f(z+\varepsilon u, t) + \varepsilon^2 g(z + \varepsilon u, t, \varepsilon) \tag{79}$$

But, since $u_t(z,t) = f(z,t) - \hat{f}(z)$, we find

$$(1 + \varepsilon u_x)\dot{z} = \varepsilon \hat{f} + \varepsilon R \tag{80}$$

where the remainder term R is

$$R = f(z + \varepsilon u, t) - f(z, t) + \varepsilon g(z + \varepsilon u, t, \varepsilon) \tag{81}$$

Since f_x is uniformly bounded, so is u_x and

$$(I + \varepsilon u_x)^{-1} = I - \varepsilon u_x + \varepsilon^2 S \tag{82}$$

where S is uniformly bounded on $\Omega = D \times \mathbb{R}^+$. Thus,

$$\dot{z} = \varepsilon(I - \varepsilon u_x + \varepsilon^2 S)(\hat{f} + R) = \varepsilon \hat{f}(z) + \varepsilon P \tag{83}$$

where

$$P = R + \varepsilon(\hat{f} + R)(\varepsilon S - u_x) \tag{84}$$

We wish to show $P = O(\varepsilon)$ uniformly on Ω. Since f is Lipschitz continuous and u is bounded, there is a constant L such that

$$\| f(z + \varepsilon u, t) - f(z, t) \| \leq \varepsilon L \| u \| \leq 2\varepsilon LMT \tag{85}$$

Thus,

$$\| R \| \leq \varepsilon(2LMT + M) \tag{86}$$

from which we deduce $R = O(\varepsilon)$ uniformly on Ω. Since S and u_x are uniformly bounded by some constant $K \in \mathbb{R}$, there is some constant C for sufficiently small ε such that

$$\| P \| \leq \varepsilon(2LMT + M) + \varepsilon(\varepsilon + 1)K(MT + 2LMT) \leq \varepsilon C \tag{87}$$

showing $P = O(\varepsilon)$ uniformly on Ω. Letting $\tilde{t} = \varepsilon t$, we then see that the equation for $\dot{z}$ can be written

$$\frac{dz}{d\tilde{t}} = \hat{f}(z) + P \tag{88}$$

Using another theorem in Verhulst (1985), any system of the form (88) with $z(0) = x_0$ with $P = O(\varepsilon)$ uniformly on Ω must then have $\| z - y \| = O(\varepsilon)$ for $\tilde{t} = O(1)$. ∎