C5.5: Perturbation Methods

Course synopsis

Recommended Prerequisites

Part A Differential Equations and Core Analysis (Complex Analysis). B5 courses are helpful but not officially required.

Overview

Perturbation methods underlie numerous applications of physical applied mathematics: including boundary layers in viscous flow, celestial mechanics, optics, shock waves, reaction-diffusion equations, and nonlinear oscillations. The aims of the course are to give a clear and systematic account of modern perturbation theory and to show how it can be applied to differential equations.

Synopsis (16 lectures)

Introduction to regular and singular perturbation theory: approximate roots of algebraic and transcendental equations. Asymptotic expansions and their properties. Asymptotic approximation of integrals, including Laplace's method, the method of stationary phase and the method of steepest descent. Matched asymptotic expansions and boundary layer theory. Multiple-scale perturbation theory. WKB theory and semiclassics.

Reading list

[1] E.J. Hinch, *Perturbation Methods* (Cambridge University Press, 1991), Chs. 1–3, 5–7.

[2] C.M. Bender & S.A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (Springer, 1999), Chs. 6, 7, 9–11.

[3] J. Kevorkian & J.D. Cole, *Perturbation Methods in Applied Mathematics* (Springer-Verlag, 1981), Chs. 1, 2.1–2.5, 3.1, 3.2, 3.6, 4.1, 5.2.

Authorship and acknowledgments

The author of these notes is **Jon Chapman**, with minor modifications by **Mason Porter**, **Philip Maini** and **Jim Oliver**. Please email comments and corrections to the course lecturer.

On the supplementary nature of these notes

These notes are supplementary to the lectures and should be viewed as being a part of the reading list. These notes are not meant to replace the lectures. Some of the material in these notes will be covered in a complementary way in lectures and in the model solutions to the problem sheet questions; some of the material covered in lectures is not covered in these notes and vice versa.

Contents

1 Introduction

Making precise approximations to solve equations is what distinguishes applied mathematicians from pure mathematicians, physicists and engineers. There are two methods for obtaining precise approximations: numerical methods and analytical (asymptotic) methods. These are not in competition but complement each other. Perturbation methods work when some parameter is large or small. Numerical methods work best when all parameters are order one. Agreement between the two methods is reassuring when doing research. Perturbation methods often give more physical insight. Finding perturbation approximations is more of an art than a science. It is difficult to give rules, only guidelines. Experience is valuable. Numerous additional worked examples may be found in *Perturbation Methods* by E.J. Hinch (Cambridge University Press, 1991, Chs. 1–3, 5–7) and *Advanced Mathematical Methods for Scientists and Engineers* by C.M. Bender and S.A. Orszag (Springer, 1999, Chs. 6, 7, 9–11).

2 Algebraic equations

Suppose we want to solve

$$x^2 + \epsilon x - 1 = 0$$

for x, where ϵ is a small parameter. The exact solutions are

$$x = -\frac{\epsilon}{2} \pm \sqrt{1 + \frac{\epsilon^2}{4}},$$

which we can expand using the binomial theorem:

$$x = \begin{cases} +1 - \frac{\epsilon}{2} + \frac{\epsilon^2}{8} - \frac{\epsilon^4}{128} + \cdots \\ -1 - \frac{\epsilon}{2} - \frac{\epsilon^2}{8} + \frac{\epsilon^4}{128} + \cdots \end{cases}$$

These expansions converge if $|\epsilon| < 2$. More important is that the truncated expansions give a good approximation to the roots when ϵ is small. For example, when $\epsilon = 0.1$:

$$\begin{array}{rcll} x & \sim & 1.0 & \text{1 term} \\ & & 0.95 & \text{2 terms} \\ & & 0.95125 & \text{3 terms} \\ & & 0.951249 & \text{4 terms} \\ \text{exact} & = & 0.95124922... & \end{array}$$

Here, we first found the exact solution, then approximated. Usually we need to make the approximation first, and then solve.

2.1 Iterative method

First, rearrange the equation so that it is in a form which can form the basis of an iterative process:

$$x = \pm\sqrt{1 - \epsilon x}.$$

Now, if we have an approximation to the positive root, x_n, say, a better approximation is given by

$$x_{n+1} = \sqrt{1 - \epsilon x_n}.$$

We need a starting point for the iteration: the solution when $\epsilon = 0$, $x_0 = 1$. After one iteration (on the positive root) we have

$$x_1 = \sqrt{1 - \epsilon}.$$

If we expand this as a binomial series we find

$$x_1 = 1 - \frac{\epsilon}{2} - \frac{\epsilon^2}{8} - \frac{\epsilon^3}{16} + \cdots .$$

We see that this is correct up to ϵ, but the ϵ^2 terms and higher are wrong. Hence we only need keep the first two terms

$$x_1 = 1 - \frac{\epsilon}{2} + \cdots .$$

Using this in the next iteration we have

$$x_2 = \sqrt{1 - \epsilon\left(1 - \frac{\epsilon}{2}\right)},$$

which can again be expanded to give

$$\begin{aligned} x_2 &= 1 - \frac{\epsilon}{2}\left(1 - \frac{\epsilon}{2}\right) - \frac{\epsilon^2}{8}\left(1 - \frac{\epsilon}{2}\right)^2 - \frac{\epsilon^3}{16}\left(1 - \frac{\epsilon}{2}\right)^3 + \cdots \\ &= 1 - \frac{\epsilon}{2} + \frac{\epsilon^2}{8} + \frac{\epsilon^3}{16} + \cdots . \end{aligned}$$

Now the ϵ^2 term is right, but the ϵ^3 term is still wrong. At each iteration more terms are correct, but more and more work is required. We can only check that a term is correct (without the exact solution) by proceeding to one more iteration and seeing if it changes.

The usual procedure is to place the dominant term of the equation on the x_{n+1} side (*i.e.*, the side that will give the new value), so that it can be calculated as a function of the terms on the x_n side (*i.e.*, the previously-obtained value). As we will see later, the identity of the dominant term can be adjusted by scaling.

2.2 Expansion method

First set $\epsilon = 0$ and find the unperturbed roots $x = \pm 1$ as in the iterative method. Now pose an expansion about one of these roots:

$$x = 1 + \epsilon x_1 + \epsilon^2 x_2 + \epsilon^3 x_3 + \cdots .$$

Substitute the expansion into the equation:

$$(1 + \epsilon x_1 + \epsilon^2 x_2 + \epsilon^3 x_3 + \cdots)^2 + \epsilon(1 + \epsilon x_1 + \epsilon^2 x_2 + \epsilon^3 x_3 + \cdots) - 1 = 0.$$

Expanding the first term

$$1 + 2x_1\epsilon + (x_1^2 + 2x_2)\epsilon^2 + (2x_1x_2 + 2x_3)\epsilon^3 + \cdots + \epsilon + \epsilon^2 x_1 + \epsilon^3 x_2 + \cdots - 1 = 0.$$

Now we equate coefficients of powers of ϵ.

At ϵ^0: $\quad 1 - 1 = 0.$

This level is automatically satisfied because we started the expansion with the correct value $x = 1$ at $\epsilon = 0$.

At ϵ^1: $\quad 2x_1 + 1 = 0, \quad$ *i.e.* $x_1 = -\frac{1}{2}$.

At ϵ^2: $\quad x_1^2 + 2x_2 + x_1 = 0, \quad$ *i.e.* $x_2 = \frac{1}{8}$,

where the previously determined value of x_1 is used.

At ϵ^3: $\quad 2x_1x_2 + 2x_3 + x_2 = 0, \quad$ *i.e.* $x_3 = 0$.

The expansion method is much easier that the iterative method when working to higher orders. However, it is necessary to assume the form of the expansion (in powers of ϵ).

2.3 Singular perturbations

Consider the problem:

$$\epsilon x^2 + x - 1 = 0.$$

When $\epsilon = 0$ there is just one root $x = 1$, but when $\epsilon \neq 0$ there are two roots. This is an example of a **singular perturbation** problem, in which the limit problem $\epsilon = 0$ differs in an important way from the limit $\epsilon \to 0$. The most interesting problems are often singular. Problems which are not singular are said to be **regular**.

To see what is happening let us look at the exact solutions

$$x = \frac{-1 \pm \sqrt{1+4\epsilon}}{2\epsilon},$$

and expand them for small ϵ (convergent if $|\epsilon| < 1/4$). The expansions of the two roots are

$$x = \begin{cases} 1 - \epsilon + 2\epsilon^2 - 5\epsilon^4 + \cdots \\ -\frac{1}{\epsilon} - 1 + \epsilon - 2\epsilon^2 + 5\epsilon^4 + \cdots \end{cases}$$

Thus the second root disappears to $x = -\infty$ as $\epsilon \to 0$.

We see that to capture the second root we need to start the expansion not with ϵ^0 but with ϵ^{-1}:

$$x = \frac{x_{-1}}{\epsilon} + x_0 + \epsilon x_1 + \cdots .$$

Substituting into the equation gives

$$\epsilon \left(\frac{x_{-1}}{\epsilon} + x_0 + \epsilon x_1 + \cdots \right)^2 + \left(\frac{x_{-1}}{\epsilon} + x_0 + \epsilon x_1 + \cdots \right) - 1 = 0.$$

Expanding the first term gives

$$\frac{1}{\epsilon} x_{-1}^2 + 2x_{-1}x_0 + \epsilon(2x_{-1}x_1 + x_0^2) + \cdots + \frac{1}{\epsilon} x_{-1} + x_0 + \epsilon x_1 + \cdots - 1 = 0.$$

Comparing coefficients of ϵ^n gives

At ϵ^{-1}: $\quad x_{-1}^2 + x_{-1} = 0, \quad$ *i.e.* $x_{-1} = -1$ or 0.

The root $x_{-1} = 0$ leads to the regular root, so we consider the singular root $x_{-1} = -1$.

At ϵ^0: $\quad 2x_{-1}x_0 + x_0 - 1 = 0, \quad$ *i.e.* $x_0 = -1$.

At ϵ^1: $\quad 2x_{-1}x_1 + x_0^2 + x_1 = 0, \quad$ *i.e.* $x_1 = 1$.

2.3.1 Rescaling the equation

Instead of starting the expansion with ϵ^{-1}, a very useful idea for singular problems is to rescale the variables before making the expansion. If we introduce the rescaling

$$x = \frac{X}{\epsilon}$$

into the originally singular equation we find that the equation for X,

$$X^2 + X - \epsilon = 0,$$

is regular. Thus the problem of finding the correct starting point for the expansion can be viewed as the problem of finding a suitable rescaling to regularise the singular problem.

2.4 Finding the right rescaling

Systematic approach: general rescaling

First pose a general rescaling with scaling factor $\delta(\epsilon)$:

$$x = \delta X,$$

in which X is strictly of order one as $\epsilon \to 0$. This gives

$$\epsilon\delta^2 X^2 + \delta X - 1 = 0.$$

Then consider the dominant balance in the equation as δ varies from very small to very large.

(i) $\delta \ll 1$. Then

$$\epsilon\delta^2 X^2 + \delta X - 1 = \text{small} + \text{small} - 1,$$

which cannot possibly balance the zero on the right-hand side of the equation. As δ is gradually increased the first term to break the domination of the -1 term is δX, which comes into play when $\delta = 1$.

(ii) $\delta = 1$. Now the left-hand side is

$$\epsilon\delta^2 X^2 + \delta X - 1 = \text{small} + X - 1.$$

This can balance the zero on the right-hand side, and produces the regular root $X = +1+$ small.

(iii) $1 \ll \delta \ll \epsilon^{-1}$. Now the term δX dominates the left-hand side, since upon dividing by δ,

$$\frac{\epsilon\delta^2 X^2 + \delta X - 1}{\delta} = \text{small} + X + \text{small}.$$

This can only balance the zero on the right-hand side if $X = 0$, but that violates the restriction that X is strictly of order one. As δ is further increased the dominance of δX is broken when the first term comes into play at $\delta = \epsilon^{-1}$.

(iv) $\delta = \epsilon^{-1}$. Now the left-hand side divided by $\epsilon\delta^2$ is

$$\frac{\epsilon\delta^2 X^2 + \delta X - 1}{\epsilon\delta^2} = X^2 + X + \text{small}.$$

This can balance the zero on the right-hand side and gives the singular root $X = -1+$ small. (Note that the solution $X = 0$ is not permitted since X has to be strictly of order one).

(v) $\delta \gg \epsilon^{-1}$. Finally, if δ is larger still then the left-hand side divided by $\epsilon\delta^2$ is dominated by the first term

$$\frac{\epsilon\delta^2 X^2 + \delta X - 1}{\epsilon\delta^2} = X^2 + \text{small} + \text{small},$$

which cannot balance the zero on the right-hand side with X strictly of order one.

Alternative approach: pairwise comparison

An alternative method is to compare terms pairwise, which is quicker when there are a small number of terms. To get a sensible answer from equating the left-hand side to zero we need at least two terms to be in balance (sometimes known as a **distinguished limit**). The possible combinations are the first and second terms, the first and third terms, or the second and third terms.

(i) First and second terms in balance. To have ϵx^2 and x in balance requires x to be of size ϵ^{-1}. Then these terms are both of size ϵ^{-1}, and dominate the remaining term -1, which is of size one. This leads to the singular root.

(ii) First and third terms in balance. To have ϵx^2 and -1 in balance requires x to be of size $\epsilon^{-1/2}$. Then these terms are both of size one, but the remaining term x is of size $\epsilon^{-1/2}$, so that this single term dominates and there is no sensible balance.

(iii) Second and third terms in balance. To have x and -1 in balance requires x to be of size one. Then these terms are both of size one, and dominate the remaining term which is size ϵ. This leads to the regular root.

2.5 Non-integral powers

Consider the quadratic equation

$$(1-\epsilon)x^2 - 2x + 1 = 0.$$

Setting $\epsilon = 0$ gives $x = 1$ as the double root (a sign of the danger to come). Proceeding as usual we pose the expansion

$$x = 1 + \epsilon x_1 + \epsilon^2 x_2 + \cdots .$$

Substituting into the equation

$$(1-\epsilon)\left(1 + \epsilon x_1 + \epsilon^2 x_2 + \cdots\right)^2 - 2\left(1 + \epsilon x_1 + \epsilon^2 x_2 + \cdots\right) + 1 = 0.$$

Expanding gives

$$1 + 2x_1\epsilon + (2x_2 + x_1^2)\epsilon^2 + \cdots - \epsilon - 2x_1\epsilon^2 + \cdots - 2 - 2x_1\epsilon - 2x_2\epsilon^2 + \cdots + 1 = 0.$$

Comparing coefficients of ϵ gives

At ϵ^0: $\quad 1 - 2 + 1 = 0,$

which is automatically satisfied because we started with the correct value $x = 1$ at $\epsilon = 0$.

At ϵ^1: $\quad 2x_1 - 1 - 2x_1 = 0,$

which cannot be satisfied by any value of x_1 (except $x_1 = \infty$ in some sense).

The cause of the difficulty is illustrated by looking at the exact solution

$$x = \frac{1}{1 \pm \epsilon^{1/2}}.$$

Expanding the largest root for small ϵ gives

$$x = 1 + \epsilon^{1/2} + \epsilon + \epsilon^{3/2} + \cdots .$$

We should have expanded in powers of $\epsilon^{1/2}$ instead of ϵ. This is what $x_1 = \infty$ is hinting at: the scaling on x_1 is too small. (In retrospect we could have guessed that an order $\epsilon^{1/2}$ change in x would be required to produce and order ϵ change in a function at its minimum.)

If we pose the expansion

$$x = 1 + \epsilon^{1/2}x_{1/2} + \epsilon x_1 + \cdots ,$$

and substitute into the equation

$$(1-\epsilon)\left(1 + \epsilon^{1/2}x_{1/2} + \epsilon x_1 + \cdots\right)^2 - 2\left(1 + \epsilon^{1/2}x_{1/2} + \epsilon x_1 + \cdots\right) + 1 = 0.$$

Expanding gives

$$1 + 2x_{1/2}\epsilon^{1/2} + (2x_1 + x_{1/2}^2)\epsilon + (2x_{3/2} + 2x_{1/2}x_1)\epsilon^{3/2} + \cdots - \epsilon - 2x_{1/2}\epsilon^{3/2} + \cdots$$
$$-2 - 2x_{1/2}\epsilon^{1/2} - 2x_1\epsilon - 2x_{3/2}\epsilon^{3/2} + \cdots + 1 = 0.$$

Comparing coefficients of ϵ we find that

At ϵ^0: $\quad 1 - 2 + 1 = 0,$

is automatically satisfied as usual and

At $\epsilon^{1/2}$: $\quad 2x_{1/2} - 2x_{1/2} = 0,$

is satisfied for all values of $x_{1/2}$. Slightly disturbing that $x_{1/2}$ is not determined but at least the expansion is consistent so far.

At ϵ^1: $\quad 2x_1 + x_{1/2}^2 - 1 - 2x_1 = 0,$

so that $x_{1/2} = \pm 1$ and x_1 is not determined at this level.

At $\epsilon^{3/2}$: $\quad 2x_{3/2} + 2x_{1/2}x_1 - 2x_{1/2} - 2x_{3/2} = 0,$

so that $x_1 = 1$ for both roots $x_{1/2}$, while $x_{3/2}$ is not determined.

2.6 Finding the right expansion sequence

How would we determine the expansion sequence if we did not have the exact solution to compare with? First pose a general expansion

$$x = 1 + \delta_1 x_1, \qquad \delta_1(\epsilon) \ll 1$$

and substitute this into the equation to get

$$(1-\epsilon)\left(1 + \delta_1 x_1\right)^2 - 2\left(1 + \delta_1 x_1\right) + 1 = 0.$$

Expanding

$$1 + 2\delta_1 x_1 + \delta_1^2 x_1^2 - \epsilon + 2\epsilon\delta_1 x_1 + \delta_1^2 \epsilon x_1^2 - 2 - 2\delta_1 x_1 + 1 = 0.$$

Simplifying leaves

$$\delta_1^2 x_1^2 - \epsilon + 2\epsilon\delta_1 x_1 + \delta_1^2 \epsilon x_1^2 = 0.$$

Now we play the dominant balance game again. Since $\epsilon\delta_1 \ll \epsilon$ the leading terms are $\delta_1^2 x_1^2$ and ϵ. Thus to get a sensible balance we need $\delta_1 = \epsilon^{1/2}$. With this value for δ_1 we equate coefficients of ϵ to get

$$x_1^2 - 1 = 0, \quad \textit{i.e.}\ x_1 = \pm 1.$$

To proceed to higher order we play the game again. Choosing $x_1 = 1$ for example, we now have

$$x = 1 + \epsilon^{1/2} + \delta_2 x_2, \qquad \delta_2(\epsilon) \ll \epsilon^{1/2}.$$

Substituting into the equation

$$(1-\epsilon)\left(1 + \epsilon^{1/2} + \delta_2 x_2\right)^2 - 2\left(1 + \epsilon^{1/2} + \delta_2 x_2\right) + 1 = 0.$$

Expanding

$$\begin{aligned}1 &+ 2\epsilon^{1/2} + \epsilon + 2\delta_2 x_2 + 2\epsilon^{1/2}\delta_2 x_2 + \delta_2^2 x_2^2 \\ &- \epsilon - 2\epsilon^{3/2} - \epsilon^2 - 2\epsilon\delta_2 x_2 - 2\epsilon^{3/2}\delta_2 x_2 - \epsilon\delta_2^2 x_2^2 - 2 - 2\epsilon^{1/2} - 2\delta_2 x_2 + 1 = 0.\end{aligned}$$

Simplifying leaves

$$2\epsilon^{1/2}\delta_2 x_2 + \delta_2^2 x_2^2 - 2\epsilon^{3/2} - \epsilon^2 - 2\epsilon\delta_2 x_2 - 2\epsilon^{3/2}\delta_2 x_2 - \epsilon\delta_2^2 x_2^2 = 0.$$

Since $\delta_2 \ll \epsilon^{1/2}$ the dominant term involving δ_2 is $2\epsilon^{1/2}\delta_2 x_2$. This must balance with $-2\epsilon^{3/2}$, giving $\delta_2 = \epsilon$ and $x_2 = 1$.

2.7 Iterative method

This is often very useful in cases where the expansion sequence is not known. Writing the original quadratic as

$$(x-1)^2 = \epsilon x^2,$$

we are led to the iterative process

$$x_{n+1} = 1 \pm \epsilon^{1/2} x_n.$$

Starting with $x_0 = 1$ the positive root gives

$$x_1 = 1 + \epsilon^{1/2}$$

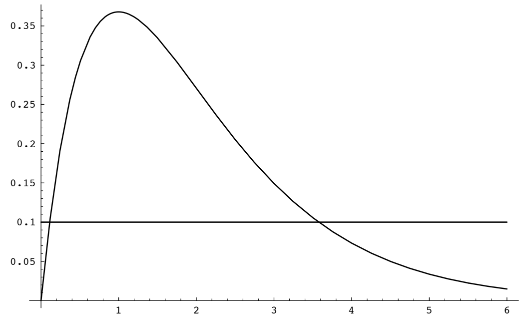
and

$$x_2 = 1 + \epsilon^{1/2} + \epsilon.$$

2.8 Logarithms

Consider the transcendental equation

$$xe^{-x} = \epsilon.$$



One root is near $x = 0$ and is easy to approximate. The other gets large as $\epsilon \to 0$ and is more difficult to find. Since the expansion sequence is not obvious we use the iterative procedure. Now, when $x = \log 1/\epsilon$, $xe^{-x} = \epsilon \log 1/\epsilon \gg \epsilon$. When $x = 2\log 1/\epsilon$, $xe^{-x} = 2\epsilon^2 \log 1/\epsilon \ll \epsilon$. Over this range the term x is slowly varying while e^{-x} is rapidly varying. This suggests rewriting the equation as

$$e^{-x} = \frac{\epsilon}{x}$$

giving the iterative scheme

$$x_{n+1} = \log(1/\epsilon) + \log x_n.$$

We have seen that the root lies roughly around $x = \log(1/\epsilon)$, so we start the iteration from $x_0 = \log(1/\epsilon)$. Then

$$x_1 = \log(1/\epsilon) + \log\log(1/\epsilon).$$

Then

$$\begin{aligned}
x_2 &= \log(1/\epsilon) + \log(\log(1/\epsilon) + \log\log(1/\epsilon)) \\
&= \log(1/\epsilon) + \log\log(1/\epsilon) + \log\left(1 + \frac{\log\log(1/\epsilon)}{\log(1/\epsilon)}\right) \\
&= \log(1/\epsilon) + \log\log(1/\epsilon) + \frac{\log\log(1/\epsilon)}{\log(1/\epsilon)} - \frac{1}{2}\left(\frac{\log\log(1/\epsilon)}{\log(1/\epsilon)}\right)^2 + \cdots .
\end{aligned}$$

Iterating again

$$
\begin{aligned}
x_3 &= \log(1/\epsilon) + \log\left(\log(1/\epsilon) + \log\log(1/\epsilon) + \frac{\log\log(1/\epsilon)}{\log(1/\epsilon)} - \frac{1}{2}\left(\frac{\log\log(1/\epsilon)}{\log(1/\epsilon)}\right)^2\right) \\
&= \log(1/\epsilon) + \log\log(1/\epsilon) + \log\left(1 + \frac{\log\log(1/\epsilon)}{\log(1/\epsilon)} + \frac{\log\log(1/\epsilon)}{(\log(1/\epsilon))^2} - \frac{1}{2}\frac{(\log\log(1/\epsilon))^2}{(\log(1/\epsilon))^3}\right) \\
&= \log(1/\epsilon) + \log\log(1/\epsilon) + \left(\frac{\log\log(1/\epsilon)}{\log(1/\epsilon)} + \frac{\log\log(1/\epsilon)}{(\log(1/\epsilon))^2} - \frac{1}{2}\frac{(\log\log(1/\epsilon))^2}{(\log(1/\epsilon))^3}\right) \\
&\quad - \frac{1}{2}\left(\frac{\log\log(1/\epsilon)}{\log(1/\epsilon)} + \frac{\log\log(1/\epsilon)}{(\log(1/\epsilon))^2} - \cdots\right)^2 + \frac{1}{3}\left(\frac{\log\log(1/\epsilon)}{\log(1/\epsilon)} + \cdots\right)^3 + \cdots \\
&= \log(1/\epsilon) + \log\log(1/\epsilon) + \frac{\log\log(1/\epsilon)}{\log(1/\epsilon)} + \\
&\quad \frac{-\frac{1}{2}(\log\log(1/\epsilon))^2 + \log\log(1/\epsilon)}{(\log(1/\epsilon))^2} + \frac{\frac{1}{3}(\log\log(1/\epsilon))^3 - \frac{3}{2}(\log\log(1/\epsilon))^2 + \cdots}{(\log(1/\epsilon))^3} + \cdots .
\end{aligned}
$$

Difficult sequence to guess. The appearance of $\log\epsilon$, and especially of $\log\log(1/\epsilon)$, means that very small values of ϵ are needed for the asymptotic expansion to be a good approximation. Normally we hope to do OK for $\epsilon = 0.5$, or at worst $\epsilon = 0.1$. However even when $\epsilon = 10^{-9}$, $\log\log(1/\epsilon)$ is only 3.

3 Asymptotic approximations

3.1 Definitions

Convergence A series $\sum_{n=0}^{\infty} f_n(z)$ is said to **converge** at a fixed value of z if given an arbitrary $\epsilon > 0$ it is possible to find a number $N_0(z, \epsilon)$ such that

$$\left|\sum_{n=M}^{N} f_n(z)\right| < \epsilon \qquad \text{for all } M,\, N > N_0.$$

A series $\sum_{n=0}^{\infty} f_n(z)$ is said to **converge** to a function $f(z)$ at a fixed value of z if given an arbitrary $\epsilon > 0$ it is possible to find a number $N_0(z, \epsilon)$ such that

$$\left|\sum_{n=0}^{N} f_n(z) - f(z)\right| < \epsilon \qquad \text{for all } N > N_0.$$

Thus a series converges if its terms decay sufficiently rapidly as $n \to \infty$.

The property of convergence is less useful in practice that we are often led to believe. Consider

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2}\, dt.$$

Since e^{-t^2} is analytic in the entire complex plane it can be expanded in a Taylor series

$$\sum_{n=0}^{\infty} \frac{(-t^2)^n}{n!}$$

which converges with an infinite radius of convergence (*i.e.* it converges for all t). This allows us to integrate term by term to get a series for erf(z) which also converges with an infinite radius of convergence:

$$
\begin{aligned}
\operatorname{erf}(z) &= \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n+1}}{(2n+1)n!} \\
&= \frac{2}{\sqrt{\pi}}\left(z - \frac{z^3}{3} + \frac{z^5}{10} - \frac{z^7}{42} + \frac{z^9}{216} - \frac{z^{11}}{1320} + \cdots\right).
\end{aligned}
$$

Taking eight terms in the series gives an accuracy of 10^{-5} up to $z = 1$. As z increases progressively more terms are needed to maintain this accuracy, *e.g.* 16 terms at $z = 2$, 31 terms at $z = 3$, 75 terms at $z = 5$. As well as requiring lots of terms, the intermediate terms get very large when z is large (there is lots of cancellation from positive and negative terms). Thus round-off errors come into play. A computer with a round-off error of 10^{-7} can give an answer accurate to only about 10^{-4} at $z = 3$ because the largest term is about 214. At $z = 5$ the largest term is 6.6×10^8, so that round-off error swamps the answer, and the computer gets it completely wrong.

The problem is that the truncated sums are very different from the converged limit—the approximation does not get better with each successive term (until we have a lot of terms).

An alternative approximation to erf (z) can be constructed by writing

$$\operatorname{erf}(z) = 1 - \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-t^2}\, dt$$

and integrating by parts to give

$$\int_z^\infty e^{-t^2}\, dt = \int_z^\infty \frac{2te^{-t^2}}{2t}\, dt = \frac{e^{-z^2}}{2z} - \int_z^\infty \frac{e^{-t^2}}{2t^2}\, dt.$$

Continuing integrating by parts gives

$$\operatorname{erf}(z) = 1 - \frac{e^{-z^2}}{z\sqrt{\pi}} \left(1 - \frac{1}{2z^2} + \frac{1.3}{(2z^2)^2} - \frac{1.3.5}{(2z^2)^3} + \cdots \right).$$

This series diverges for all z: it has radius of convergence zero. However, the truncated series is very useful. At $z = 2.5$ three terms give an accuracy of 10^{-5}. At $z = 3$ only two terms are necessary. The series has the important property that the leading term is almost correct, and the addition of each successive term gets us a bit closer to the answer, *i.e.* each of the corrections is of decreasing size (until they finally start to diverge). The series is an **asymptotic** series.

Asymptoticness A sequence $\{f_n(\epsilon)\}_{n\in\mathbb{N}_0}$ is said to be **asymptotic** if for all $n \geq 1$

$$\frac{f_n(\epsilon)}{f_{n-1}(\epsilon)} \to 0 \qquad \text{as } \epsilon \to 0.$$

A series $\sum_{n=0}^{\infty} f_n(\epsilon)$ is said to be an **asymptotic approximation** to (or asymptotic expansion of) a function $f(\epsilon)$ as $\epsilon \to 0$ if for all $N \geq 0$

$$\frac{f(\epsilon) - \sum_{n=0}^{N} f_n(\epsilon)}{f_N(\epsilon)} \to 0 \qquad \text{as } \epsilon \to 0,$$

i.e. the remainder is smaller than the last term included once ϵ is sufficiently small. We write

$$f \sim \sum_{n=0}^{\infty} f_n(\epsilon) \qquad \text{as } \epsilon \to 0.$$

Usually we don't worry about getting the whole series, just the first few terms.

Often the $f_n(\epsilon)$ are powers of ϵ multiplied by a coefficient, *i.e.*

$$f \sim \sum_{n=0}^{\infty} a_n \epsilon^n$$

which is called an **asymptotic power series**. Sometimes though, as we have already seen, fractional powers or logs may appear.

Order notation We write $f = O(g)$ as $\epsilon \to 0$ to mean that there exist constants $K > 0$ and $\epsilon_0 > 0$ such that

$$|f| < K|g| \quad \text{for all } \epsilon < \epsilon_0.$$

We write $f = o(g)$ as $\epsilon \to 0$ to mean

$$\frac{f}{g} \to 0 \quad \text{as } \epsilon \to 0.$$

Then $f_n(\epsilon)$ is an asymptotic sequence if $f_n = o(f_{n-1})$, and $f \sim \sum_{n=0}^{\infty} f_n$ if

$$f - \sum_{n=0}^{N} f_n = o(f_N) \quad \text{for all } N \geq 0.$$

Examples

- $\sin x = O(x)$ as $x \to 0$.
- $\sin x = O(1)$ as $x \to \infty$.
- $\sin x = O(1)$ as $x \to 0$. Note that quite often when dealing with simple powers often take order to be largest/smallest such power that works.
- $\log x = O(x)$ as $x \to \infty$.
- $\log x = o(x)$ as $x \to \infty$.
- $\log x = o(x^{-\delta})$ as $x \to 0$, for any $\delta > 0$.

3.2 Uniqueness and manipulation of asymptotic series

If a function posesses an asymptotic approximation in terms of an asymptotic sequence, then that approximation is **unique** for that particular sequence. Given the existence of an approximation $f \sim \sum_{n=0}^{\infty} a_n \delta_n(\epsilon)$ in terms of a given sequence $\{\delta_n(\epsilon)\}_{n \in \mathbb{N}_0}$, the coefficients can be evaluated inductively from

$$a_k = \lim_{\epsilon \to 0} \frac{f(\epsilon) - \sum_{n=0}^{k-1} a_n \delta_n(\epsilon)}{\delta_k(\epsilon)}.$$

Note that the uniqueness is for a given sequence. A single function may have many asypmtotic approximations, each in terms of a different sequence. For example

$$\begin{aligned}
\tan(\epsilon) &\sim \epsilon + \frac{\epsilon^3}{3} + \frac{2\epsilon^5}{15} + \cdots \\
&\sim \sin\epsilon + \frac{1}{2}(\sin\epsilon)^3 + \frac{3}{8}(\sin\epsilon)^5 + \cdots \\
&\sim \epsilon\cosh\left(\sqrt{\frac{2}{3}}\,\epsilon\right) + \frac{31}{270}\left(\epsilon\cosh\left(\sqrt{\frac{2}{3}}\,\epsilon\right)\right)^2 + \cdots .
\end{aligned}$$

Note also that the uniqueness is for a given function: two functions may share the same asymptotic approximation, because they differ by a quantity smaller than the last term included. For example

$$\begin{aligned}
\exp(\epsilon) &\sim \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} \quad \text{as } \epsilon \to 0, \\
\exp(\epsilon) + \exp(-1/\epsilon) &\sim \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} \quad \text{as } \epsilon \to 0^+
\end{aligned}$$

($\epsilon \to 0^+$ means as ϵ tends to zero through positive values). Two functions sharing the same asymptotic power series, as above, can only differ by a quantity which is not analytic, because two analytic functions with the same power series are identical.

Asymptotic approximations can be naively added, subtracted, multiplied or divided, resulting in the correct asymptotic expression for the sum, difference, product or quotient, perhaps based on an enlarged asymptotic sequence.

One asymptotic series can be substituted into another, although care is needed with exponentials. For example, if

$$f(z) = e^{z^2}, \qquad z(\epsilon) = \epsilon^{-1} + \epsilon,$$

then

$$f(z(\epsilon)) = e^{(\epsilon^{-1}+\epsilon)^2} \sim e^{-\epsilon^{-2}} e^2 \left(1 + \epsilon^2 + \frac{\epsilon^4}{2} + \cdots\right).$$

However, if only the leading term in z is used we get the wrong answer $\exp(-\epsilon^{-2})$, in error by a factor of e^2. To avoid this error exponents need to be calculated to $O(1)$, not just to leading order. Remember that cos and sin are exponentials as far as this is concerned.

Asymptotic expansions can be **integrated** term by term with respect to ϵ resulting in the correct asymptotic expansion of the integral. However, in general they may not be differentiated with safety. The trouble comes with terms like $\epsilon \cos(1/\epsilon)$ which has a derivative $O(1/\epsilon)$ rather than the expected $O(1)$. Such terms move higher up the expansion when integrated (safe), but lower down it when differentiated (unsafe). Thus when differentiating there is always the worry that neglected higher-order terms suddenly become important.

3.3 Numerical use of divergent series

Usually the first few terms in a series are enough to get the desired accuracy. However, if a more accurate representation is needed more terms can be taken. Clearly, if the series is divergent, as they often are, it makes no sense to keep including extra terms when they stop decreasing in magnitude and start to diverge. Truncating at the smallest term is known as **optimal** truncation.

3.4 Parametric expansions

So far we have been considering functions of a single variable as that variable tends to zero. Such problems often occur in ordinary and especially partial differential equations when considering far field behaviour for example, and there are known as coordinate expansions.

More common is for the solution of an equation to depend on more than one variable, $f(x, \epsilon)$ say. Often we have a differential equation in the independent variable x which contains a small parameter ϵ, hence the name parametric expansion. For functions of two variables the obvious generalisation is to allow the coefficients of the asymptotic expansion to be functions of the second variable:

$$f(x, \epsilon) \sim \sum_{n=0}^{\infty} a_n(x)\delta_n(\epsilon) \quad \text{as } \epsilon \to 0.$$

4 Asymptotic approximation of integrals

4.1 Integration by parts

We have already seen the use of integration by parts to obtain an asymptotic approximation of the error function. Here we show some more examples.

Example 1. Derivation of an asymptotic power series If $f(\epsilon)$ is differentiable near $\epsilon = 0$ then the local behaviour of $f(\epsilon)$ near 0 may be studied using integration by parts. We write

$$f(\epsilon) = f(0) + \int_0^\epsilon f'(x)\,dx.$$

Integrating by parts once gives

$$f(\epsilon) = f(0) + \big[(x-\epsilon)f'(x)\big]_0^\epsilon + \int_0^\epsilon (\epsilon - x) f''(x)\,dx.$$

Repeating $N-1$ times gives

$$f(\epsilon) = \sum_{n=0}^{N} \frac{\epsilon^n f^{(n)}(0)}{n!} + \frac{1}{N!}\int_0^\epsilon (\epsilon - x)^N f^{(N+1)}(x)\,dx.$$

If the remainder term exists for all N and sufficiently small $\epsilon > 0$ then

$$f(\epsilon) \sim \sum_{n=0}^{\infty} \frac{\epsilon^n f^{(n)}(0)}{n!} \qquad \text{as } \epsilon \to 0.$$

If the series converges then it is just the Taylor expansion of $f(\epsilon)$ about $\epsilon = 0$.

Example 2.

$$I(x) = \int_x^\infty e^{-t^4}\,dt.$$

As $x \to \infty$,

$$\begin{aligned} I(x) &= -\frac{1}{4}\int_x^\infty \frac{1}{t^3}\frac{d}{dt}(e^{-t^4})\,dt \\ &= \left[-\frac{e^{-t^4}}{4t^3}\right]_x^\infty - \frac{3}{4}\int_x^\infty \frac{1}{t^4}e^{-t^4}\,dt \\ &= \frac{e^{-x^4}}{4x^3} - \frac{3}{4}\int_x^\infty \frac{1}{t^4}e^{-t^4}\,dt. \end{aligned}$$

The first term is the leading-order asymptotic approximation because

$$\int_x^\infty \frac{1}{t^4}e^{-t^4}\,dt < \frac{1}{x^4}\int_x^\infty e^{-t^4}\,dt = \frac{1}{x^4}I(x) \ll I(x) \qquad \text{as } x \to \infty.$$

Further integration by parts gives more terms in the asymptotic series.

Example 3.

$$I(x) = \int_0^x t^{-1/2}e^{-t}\,dt.$$

Here we need to be more careful because the naive approach

$$I(x) = \left[-t^{-1/2}e^{-t}\right]_0^x - \frac{1}{2}\int_0^x t^{-3/2}e^{-t}\,dt$$

gives $\infty - \infty$. Instead we express $I(x)$ as the difference between two integrals

$$I(x) = \int_0^\infty t^{-1/2}e^{-t}\,dt - \int_x^\infty t^{-1/2}e^{-t}\,dt.$$

The first integral is finite, independent of x; it has the value $\Gamma(1/2) = \sqrt{\pi}$. The second may be integrated by parts successfully, because the contribution from the endpoint vanishes.

$$\begin{aligned}\int_0^x t^{-1/2}e^{-t}\,dt &= \sqrt{\pi} + \int_x^\infty t^{-1/2}\frac{d}{dt}(e^{-t})\,dt \\ &= \sqrt{\pi} - \frac{e^{-x}}{\sqrt{x}} + \frac{1}{2}\int_x^\infty t^{-3/2}e^{-t}\,dt.\end{aligned}$$

General rule: Integration by parts will not work if the contribution from one of the limits of integration is much larger than the size of the integral. Here $I(x)$ is finite for all $x > 0$, but at the endpoint $t = 0$ the integrand has a singularity, which gets worse on differentiating.

4.2 Failure of integration by parts

$$I(x) = \int_0^\infty e^{-xt^2}\,dt.$$

If we try integration by parts we find

$$\int_0^\infty e^{-xt^2}\,dt = \int_0^\infty \left(-\frac{1}{2xt}\right)\left(-2xte^{-xt^2}\right)\,dt = \left[\frac{e^{-xt^2}}{-2xt}\right]_0^\infty - \int_0^\infty \frac{1}{2xt^2}e^{-xt^2}\,dt.$$

The final integral does not exist, a sure sign that integration by parts has failed. In fact, $I(x)$ has the exact value $\sqrt{\pi}/(2\sqrt{x})$. Integration by parts could never pick up this fractional power, and is doomed to failure. Integration by parts will also not work when the dominant contribution to the integral comes from an interior point rather than an end point. While integration by parts is simple to use and gives an explicit error term that can often be rigorously bounded, it is of limited applicability and inflexible.

4.3 Laplace's method

Laplace's method is a general technique for obtaining the behaviour as $x \to +\infty$ of integrals of the form

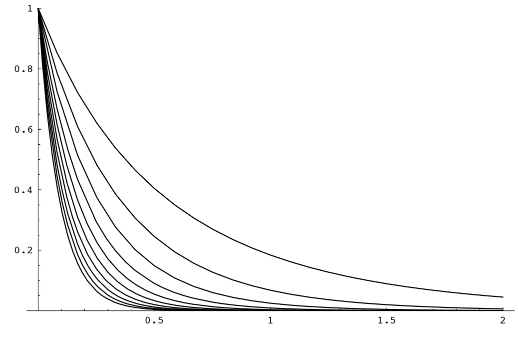
$$I(x) = \int_a^b f(t)e^{x\phi(t)}\,dt,$$

where $f(t)$ and $\phi(t)$ are real continuous functions.

Example Find the asymptotic behaviour of

$$I(x) = \int_0^{10} \frac{e^{-xt}}{(1+t)}\,dt$$

as $x \to +\infty$. The integrand is shown for $x = 1, \cdots, 10$.



As $x \to \infty$ the largest contribution to the integral comes from near $t = 0$ because this is where $-t$ is biggest. For values of t away from zero the integrand is exponentially small. So split the range of integration:

$$I(x) = \int_0^{\epsilon} \frac{e^{-xt}}{(1+t)}\,dt + \int_{\epsilon}^{10} \frac{e^{-xt}}{(1+t)}\,dt$$

where $x^{-1} \ll \epsilon \ll 1$. The second integral is $O(e^{-\epsilon x})$ which is exponentially small by comparison to the first, so we can neglect it. In the first integral t is small so we can Taylor expand $1/(1+t)$. The best way to be systematic is to change variable $xt = s$, giving

$$I(x) \sim \frac{1}{x} \int_0^{x\epsilon} \frac{e^{-s}}{(1+s/x)}\,ds.$$

Since $x\epsilon$ (the largest value of s) is $\ll x$ we Taylor expand $1/(1+s/x)$ to give

$$I(x) \sim \frac{1}{x} \int_0^{x\epsilon} e^{-s} \sum_{n=0}^{\infty} \frac{(-s)^n}{x^n}\,ds = \sum_{n=0}^{\infty} \frac{1}{x^{n+1}} \int_0^{x\epsilon} (-s)^n e^{-s}\,ds,$$

since the expansion is uniform on $0 < s < \epsilon x$. Finally, we can now replace the upper limit $x\epsilon$ by infinity in each sum, introducing only exponentially small errors again because integration by parts shows that

$$\int_{x\epsilon}^{\infty} s^n e^{-s}\,ds = O((x\epsilon)^n e^{-\epsilon x}).$$

Hence

$$I(x) \sim \sum_{n=0}^{\infty} \frac{1}{x^{n+1}} \int_0^{\infty} (-s)^n e^{-s}\,ds = \sum_{n=0}^{\infty} \frac{(-1)^n n!}{x^{n+1}} \qquad \text{as } x \to \infty.$$

4.4 Watson's lemma

The method of the example can be justified using Watson's lemma, which applies to integrals of the form

$$I(x) = \int_0^b f(t) e^{-xt}\,dt, \qquad b > 0.$$

Suppose $f(t)$ is continuous on the interval $0 \le t \le b$ and has the asymptotic series expansion

$$f(t) \sim t^{\alpha} \sum_{n=0}^{\infty} a_n t^{\beta n} \qquad \text{as } t \to 0+,$$

where $\alpha > -1$ and $\beta > 0$ so that the integral converges at $t = 0$. If $b = \infty$ it is also necessary that $f(t) \ll e^{ct}$ as $t \to +\infty$ for some positive constant c so that the integral converges at $t = \infty$. Then Watson's lemma states that

$$I(x) \sim \sum_{n=0}^{\infty} \frac{a_n \Gamma(\alpha + \beta n + 1)}{x^{\alpha+\beta n+1}} \qquad \text{as } x \to +\infty.$$

The derivation of Watson's Lemma is basically by the same method as in the example if the asymptotic series for f is uniformly convergent in a neighbourhood of the origin (as is often the case in practice). If this is not the case (as it is in general), then it is no longer possible to interchange the order of integration and summation: we work instead with a finite number of terms in the asymptotic expansion of f by writing, for each positive integer N,

$$f(t) = t^{\alpha} \sum_{n=0}^{N-1} a_n t^{\beta n} + \mathrm{O}(t^{\beta N}) \qquad \text{as } t \to 0+;$$

the result is then readily derived by showing that, for each positive integer N,

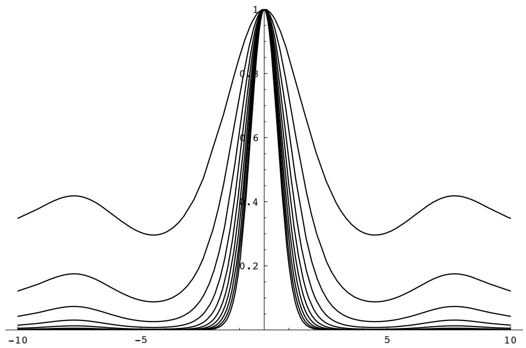
$$I(x) = \sum_{n=0}^{N-1} \frac{a_n \Gamma(\alpha + \beta n + 1)}{x^{\alpha+\beta n+1}} + \mathrm{O}\left(\frac{1}{x^{\alpha+\beta N+1}}\right) \qquad \text{as } x \to +\infty.$$

4.5 Asymptotic expansion of general Laplace integrals

Consider the integral

$$I(x) = \int_a^b f(t) e^{x\phi(t)}\, dt.$$

We have seen that the dominant contribution to the integral will come from the place where $\phi(t)$ is largest.



There are three cases to consider

1. The maximum is at $t = a$.
2. The maximum is at $t = b$.
3. The maximum is at some $t = c$, with $a < c < b$.

In each case the argument is as follows:

1. The dominant contribution to the integral comes from the near the maximum of ϕ. We can reduce the range of integration to this local contribution introducing only exponentially small errors.
2. Near this point we can expand ϕ and f in Taylor series.
3. After rescaling the integration variable, we can replace the integration limits by ∞ introducing only exponentially small errors.

Case 1: The maximum is at $t = a$. First we can split the integral into a local and nonlocal part:

$$I(x) = \int_a^{a+\epsilon} f(t) e^{x\phi(t)}\, dt + \int_{a+\epsilon}^b f(t) e^{x\phi(t)}\, dt,$$

where $x^{-1} \ll \epsilon \ll x^{-1/2}$ (we will see where these restrictions come from soon). The second integral is exponentially small compared to the first, since it is $O(e^{x\phi(a+\epsilon)})$ and $\phi(a+\epsilon) \sim \phi(a) + \epsilon\phi'(a)$. Thus the second integral is $O(e^{x\epsilon\phi'(a)})$ times the first (which we will see is $O(e^{x\phi(a)})$). This is why we need $x\epsilon \gg 1$ (remember that $\phi'(a) < 0$ since ϕ is maximum at $t = a$).

In the first it is OK to expand $\phi(t)$ and $f(t)$ as an asymptoptic series about $t = a$:

$$\phi(t) \sim \phi(a) + (t-a)\phi'(a) + \cdots, \qquad f(t) \sim f(a) + (t-a) f'(a) + \cdots.$$

Then

$$I(x) \sim \int_a^{a+\epsilon} \left(f(a) + (t-a) f'(a) + \cdots\right) e^{x\left(\phi(a) + (t-a)\phi'(a) + \frac{(t-a)^2}{2}\phi''(a) + \cdots\right)}\, dt$$

Now we rescale the integration variable to remove the x from the exponential, *i.e.* we set $x(t-a) = s$. Then

$$I(x) \sim \frac{e^{x\phi(a)}}{x} \int_0^{x\epsilon} \left(f(a) + \frac{s}{x} f'(a) + \cdots\right) e^{s\phi'(a) + \frac{s^2}{2x}\phi''(a) + \cdots}\, ds.$$

Note that $\phi'(a) < 0$, since ϕ is maximum at a. Now we can expand $e^{\frac{s^2}{2x}\phi''(a)+\cdots}$ as $x \to \infty$ as

$$1 + \frac{s^2}{2x}\phi''(a) + \cdots .$$

This is OK providing $(x\epsilon)^2/x \ll 1$ *i.e.* $\epsilon \ll x^{-1/2}$. This is where the other restriction on ϵ comes from. Keeping only the leading-order term we have

$$I(x) \sim \frac{f(a)e^{x\phi(a)}}{x} \int_0^{x\epsilon} e^{s\phi'(a)}\, ds.$$

Now we can replace the upper limit by infinity, introducing only exponentially small errors:

$$I(x) \sim \frac{f(a)e^{x\phi(a)}}{x} \int_0^{\infty} e^{s\phi'(a)}\, ds = -\frac{f(a)e^{x\phi(a)}}{x\phi'(a)}.$$

Case 2: The maximum is at $t = b$. A similar argument shows that

$$I(x) \sim \frac{f(b)e^{x\phi(b)}}{x\phi'(b)}.$$

Case 3: The maximum is at $t = c$, $a < c < b$. First we can split the integral into a local and nonlocal part:

$$I(x) = \int_a^{c-\epsilon} f(t)e^{x\phi(t)}\, dt + \int_{c-\epsilon}^{c+\epsilon} f(t)e^{x\phi(t)}\, dt + \int_{c+\epsilon}^{b} f(t)e^{x\phi(t)}\, dt,$$

where in this case we will see that we need $1/x^{1/2} \ll \epsilon \ll 1/x^{1/3}$ (we will see where these restrictions come from shortly). The first and last integrals are exponentially small compared to the second, since they are $O(e^{x\phi(c+\epsilon)})$. In this case $\phi(c+\epsilon) \sim \phi(c) + \frac{\epsilon^2}{2}\phi''(c)$ because ϕ has a maximum at the interior point $t = c$ so $\phi'(c) = 0$. This is why we need $x\epsilon^2 \gg 1$, *i.e.* $x^{-1/2} \ll \epsilon$.

In the second integral it is OK to expand $\phi(t)$ and $f(t)$ as an asymptoptic series about $t = c$:

$$\phi(t) \sim \phi(c) + \frac{(t-c)^2}{2}\phi''(c) + \frac{(t-c)^3}{6}\phi'''(c)\cdots, \qquad f(t) \sim f(c) + (t-c)f'(c) + \cdots .$$

Then

$$I(x) \sim \int_{c-\epsilon}^{c+\epsilon} \big(f(c) + (t-c)f'(c) + \cdots\big)\, e^{x(\phi(c)+\frac{(t-c)^2}{2}\phi''(c)+\frac{(t-c)^3}{6}\phi'''(c)+\cdots)}\, dt$$

Now we rescale the integration variable to remove the x from the exponential, *i.e.* we set $\sqrt{x}\,(t-c) = s$ (note the different scaling of the contributing region). Then

$$I(x) \sim \frac{e^{x\phi(c)}}{\sqrt{x}} \int_{-\sqrt{x}\,\epsilon}^{\sqrt{x}\,\epsilon} \Big(f(c) + \frac{s}{x}f'(c) + \cdots\Big)\, e^{\frac{s^2}{2}\phi''(c)+\frac{s^3}{6\sqrt{x}}\phi'''(c)+\cdots}\, ds.$$

Note that $\phi''(c) < 0$, since ϕ has a maximum at $t = c$. Now we can expand $e^{\frac{s^3}{6\sqrt{x}}\phi'''(c)+\cdots}$ as $x \to \infty$ as

$$1 + \frac{s^3}{6\sqrt{x}}\phi'''(c) + \cdots .$$

This is OK providing $(x^{1/2}\epsilon)^3/x^{1/2} \ll 1$, *i.e.* $\epsilon \ll x^{-1/3}$. This is where the other restriction on ϵ comes from. Keeping only the leading-order term we have

$$I(x) \sim \frac{f(c)e^{x\phi(c)}}{\sqrt{x}} \int_{-\sqrt{x}\,\epsilon}^{\sqrt{x}\,\epsilon} e^{\frac{s^2}{2}\phi''(c)}\, ds.$$

Now we can replace the upper and lower limits by $\pm\infty$, introducing only exponentially small errors:

$$I(x) \sim \frac{f(c)e^{x\phi(c)}}{\sqrt{x}} \int_{-\infty}^{\infty} e^{\frac{s^2}{2}\phi''(c)}\, ds = \frac{\sqrt{2\pi}\, f(c)e^{x\phi(c)}}{\sqrt{-x\phi''(c)}}.$$

4.6 Method of stationary phase

The method of stationary phase is used for problems in which the exponent ϕ is not real but purely imaginary, say $\phi(t) = i\psi(t)$, where $\psi(t)$ is real.

$$I(x) = \int_a^b f(t)e^{ix\psi(t)}\,dt.$$

Riemann-Lebesgue lemma If $\int_a^b |f(t)|\,dt < \infty$ and $\psi(t)$ is continuously differentiable for $a \le t \le b$ and not constant on any subinterval in $a \le t \le b$, then

$$\int_a^b f(t)e^{ix\psi(t)}\,dt \to 0 \qquad \text{as } x \to \infty.$$

Useful when using integration by parts.

Example

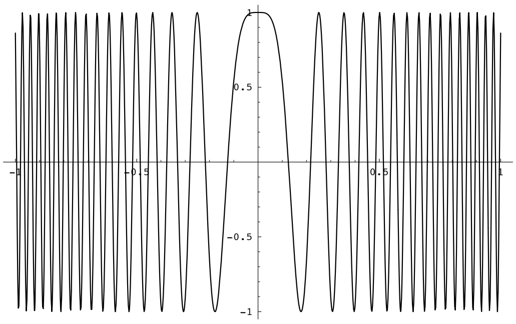
$$I(x) = \int_0^1 \frac{e^{ixt}}{1+t}\,dt.$$

Integrating by parts gives

$$I(x) = -\frac{ie^{ix}}{2x} + \frac{i}{x} - \frac{i}{x}\int_0^1 \frac{e^{ixt}}{(1+t)^2}\,dt.$$

The last integral is lower order by the Riemann-Lebesgue lemma.

Why is the Riemann-Lebesgue lemma true? Locally near any point $t = t_0$, $\psi(t) \sim \psi(t_0)+(t-t_0)\psi'(t_0)+\cdots$ and the period of oscillation is $\frac{2\pi}{x\psi'(t_0)}$. As $x \to \infty$ this is very small, $f(t)$ is almost constant, and the contribution from the "up" and "down" parts of the oscillation almost cancel out. (You can find a rigorous proof of the Riemann-Lebesgue lemma in analysis books.) However, this is not true if $\psi'(t_0) = 0$. In this case the integrand oscillates much more slowly near t_0, so that there is less cancellation. Here's a plot of $\mathrm{Re}(e^{100ix^2})$.



Suppose $\psi'(c) = 0$ with $a < c < b$, with $\psi'(t)$ being nonzero for $a \le t < c$ and $c < t \le b$. As for Laplace's method, we split the range of integration

$$I(x) = \int_a^{c-\epsilon} f(t)e^{ix\psi(t)}\,dt + \int_{c-\epsilon}^{c+\epsilon} f(t)e^{ix\psi(t)}\,dt + \int_{c+\epsilon}^{b} f(t)e^{ix\psi(t)}\,dt,$$

where $\epsilon \ll 1$. The first and third integrals are lower order. To show this we use integration by parts

$$\begin{aligned}\int_a^{c-\epsilon} f(t)e^{ix\psi(t)}\,dt &= \int_a^{c-\epsilon} \frac{f(t)}{ix\psi'(t)}\frac{d}{dt}\left(e^{ix\psi(t)}\right)\,dt \\ &= \left[\frac{f(t)}{ix\psi'(t)}e^{ix\psi(t)}\right]_a^{c-\epsilon} - \frac{1}{x}\int_a^{c-\epsilon} e^{ix\psi(t)}\frac{d}{dt}\left(\frac{f(t)}{i\psi'(t)}\right)\,dt.\end{aligned}$$

Providing the last integral exists it is lower order by the Riemann-Lebesgue lemma. The first integal is

$$O\left(\frac{1}{x\psi'(c-\epsilon)}\right) = O\left(\frac{1}{x\epsilon\psi''(c)}\right)$$

providing $\psi''(c) \neq 0$. For the second integral we expand ψ and f as an asymptoptic series about $t = c$

$$f(t) \sim f(c) + (t-c)f'(c) + \cdots, \qquad \psi(t) \sim \psi(c) + \frac{(t-c)^2}{2}\psi''(c) + \frac{(t-c)^3}{6}\psi'''(c) + \cdots.$$

Then

$$\int_{c-\epsilon}^{c+\epsilon} f(t)e^{ix\psi(t)}\,dt \sim \int_{c-\epsilon}^{c+\epsilon} \left(f(c) + (t-c)f'(c) + \cdots\right) e^{ix\left(\psi(c)+\frac{(t-c)^2}{2}\psi''(c)+\frac{(t-c)^3}{6}\psi'''(c)+\cdots\right)}\,dt.$$

As for Laplace's method, we change the integration variable so that the oscillation is on an order one scale by setting $x^{1/2}(t-c) = s$ to give

$$\int_{c-\epsilon}^{c+\epsilon} f(t)e^{ix\psi(t)}\,dt \sim \frac{e^{ix\psi(c)}}{x^{1/2}} \int_{-x^{1/2}\epsilon}^{x^{1/2}\epsilon} \left(f(c) + \frac{s}{x^{1/2}}f'(c) + \cdots\right) e^{i\frac{s^2}{2}\psi''(c)+i\frac{s^3}{6x^{1/2}}\psi'''(c)+\cdots}\,ds.$$

Now we can expand $e^{i\frac{s^3}{6x^{1/2}}\psi'''(c)+\cdots}$ as

$$1 + i\frac{s^3}{6x^{1/2}}\psi'''(c) + \cdots$$

so long as $\epsilon \ll x^{-1/3}$. The leading order term is

$$\int_{c-\epsilon}^{c+\epsilon} f(t)e^{ix\psi(t)}\,dt \sim \frac{f(c)e^{ix\psi(c)}}{x^{1/2}} \int_{-x^{1/2}\epsilon}^{x^{1/2}\epsilon} e^{i\frac{s^2}{2}\psi''(c)}\,ds.$$

Now we replace the limits of integration by $\pm\infty$, which introduces error terms of order $1/(x\epsilon)$ (check by integration by parts). Hence

$$\int_{c-\epsilon}^{c+\epsilon} f(t)e^{ix\psi(t)}\,dt \sim \frac{f(c)e^{ix\psi(c)}}{x^{1/2}} \int_{-\infty}^{\infty} e^{i\frac{s^2}{2}\psi''(c)}\,ds + O\left(\frac{1}{x\epsilon}\right) = \frac{\sqrt{2\pi}f(c)e^{ix\psi(c)}e^{\pm i\pi/4}}{x^{1/2}|\psi''(c)|^{1/2}} + O\left(\frac{1}{x\epsilon}\right)$$

where (contour integration reveals that) the factor $e^{+i\pi/4}$ is used if $\psi''(c) > 0$ and $e^{-i\pi/4}$ is used if $\psi''(c) < 0$. Thus we need $x^{-1/2} \gg (\epsilon x)^{-1}$, *i.e.* $\epsilon \gg x^{-1/2}$, as in Laplace's method. The error is the same order as the neglected first and third integrals. So finally

$$I(x) = \frac{\sqrt{2\pi}f(c)e^{ix\psi(c)}e^{\pm i\pi/4}}{x^{1/2}|\psi''(c)|^{1/2}} + O\left(\frac{1}{x\epsilon}\right)$$

as $x \to \infty$ with $x^{-1/2} \ll \epsilon \ll x^{-1/3}$.

Important notes

- The error terms are only algebraically small, not exponentially small as in Laplace's method.
- Higher-order corrections are very hard to get since they may come from the whole range of integration. This is in contrast to Laplace's method where the full asymptotic expansion depends only on the local region because the errors are exponentially small.

4.7 Method of steepest descents

Laplace's method and the method of stationary phase are really just special cases of the general method of steepest descents, which is for integrals

$$I(x) = \int_C f(t) e^{x\phi(t)}\, dt,$$

where $f(t)$ and $\phi(t)$ are **complex**, and C is some contour in the complex t-plane.

We might expect, based on Laplace's method, that the important contribution to the integral as $x \to +\infty$ comes from the place where $\mathrm{Re}(\phi)$ is maximum, at t_0 say, and that the integral is basically of size

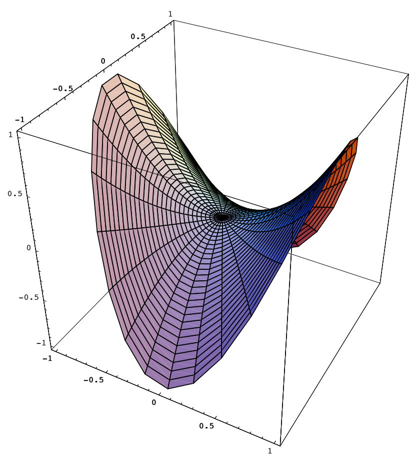
$$f(t_0) e^{x\phi(t_0)} \sqrt{\frac{2\pi}{-\lambda \phi''(t_0)}}$$

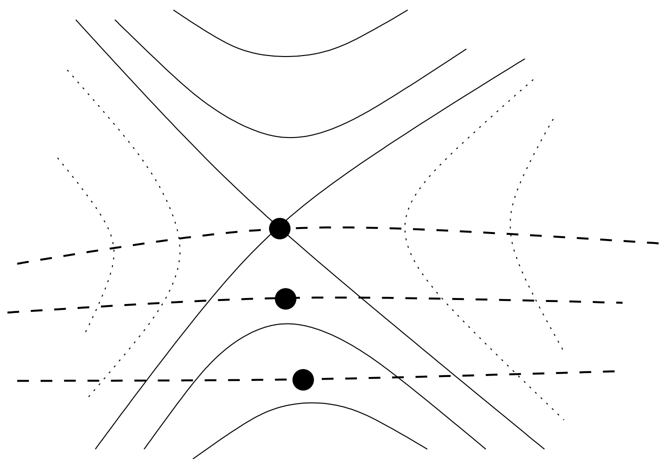
(where $'$ is the derivative along the path of integration). However, this estimate is **way too big**. The reason is that it ignores the rapid oscillation due to the imaginary part of ϕ, which causes cancellation exactly as in the method of stationary phase. We can see that the estimate above is wrong by deforming the contour a bit, which does not change the value of the integral, but which can change the maximum value of $\mathrm{Re}(\phi)$.

Now, since $\phi(t) = u(\xi,\eta) + \mathrm{i}v(\xi,\eta)$ is an analytic function of $t = \xi + \mathrm{i}\eta$, we have the Cauchy-Riemann equations

$$u_\xi = v_\eta, \quad u_\eta = -v_\xi.$$

Hence $\nabla^2 u = u_{\xi\xi} + u_{\eta\eta} = 0$. This means that u cannot have any maxima or minima in the (ξ,η)-plane, only saddle points (since a maximum or minimum would require $u_{\xi\xi}u_{\eta\eta} > 0$). Thus the landscape of u has hills ($u > 0$) and valleys ($u < 0$) at infinity, with saddle points which are the passes from one valley into another. By the Cauchy-Riemann equations the saddle points are where $d\phi/dt = 0$. If our contour is infinite it must tend to infinity in a valley (see *e.g.* surface plot of $u(\xi,\eta) = \eta^2 - \xi^2$ for $\phi(t) = -t^2$). By deforming the contour we can keep reducing the maximum value of u, until the contour goes through the saddle point which is the lowest that u gets (see *e.g.* contour plot of u in which solid lines are for positive values of u, dotted lines are for negative values of u, and the dashed lines show C being deformed through the saddle point).





But why do we know that this is the right value. Suppose we can deform the contour C into one in which v is constant. Then there is no oscillation in the integrand, and the Laplace-type argument will work. Now if v is constant on the path, then $\nabla v = (v_\xi, v_\eta)$ is perpendicular to the path. By the Cauchy-Riemann equations this means that $\nabla u = (u_\xi, u_\eta)$ is parallel to the path, so that the path follows the steepest directions on the surface of u. There is only one path on which v is constant which goes to a valley at $\pm\infty$ and this is the path through the saddle. A little thought shows that this has to be the case. Since

u is first increasing as we come up from one valley and then decreasing as we go off to another valley, we must go through a point where $du/ds = 0$, where s is distance along the path. Since v is constant, so that $dv/ds = 0$, everywhere on the path, we must go through a saddle point at which both $du/ds = 0$ and $dv/ds = 0$.

So the method of steepest descents is as follows:

(i) Deform the contour to be the steepest descent contour through the relevant saddle point(s).

(ii) Evaluate the local contribution from the saddle exactly as in Laplace's method.

(iii) Evaluate the local contribution from the end point(s) exactly as in Laplace's method.

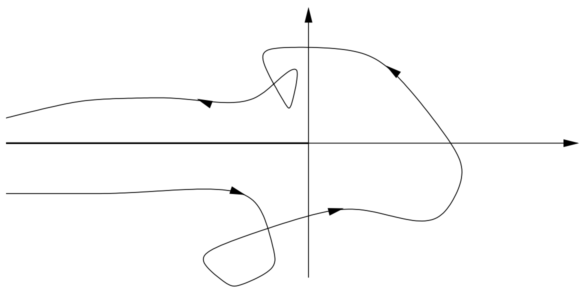
Remember that when deforming the contour we must include the contribution from any poles that we cross.

Of course, we could have chosen a path on which $u = \mathrm{Re}(\phi)$ was constant and applied the method of stationary phase. However, we have seen that Laplace's method is far superior in that it can generate all the terms in the asymptotic series: the neglected "tails" of the integral are exponentially small. In fact, the best way to generate higher order terms in a stationary phase integral is to deform to the steepest descent contour.

Example: Steepest descents on the gamma function Consider as $x \to \infty$ the gamma function $\Gamma(x)$, which may be defined by

$$\frac{1}{\Gamma(x)} = \frac{1}{2\pi i}\int_{C'} e^t t^{-x} dt\,,$$

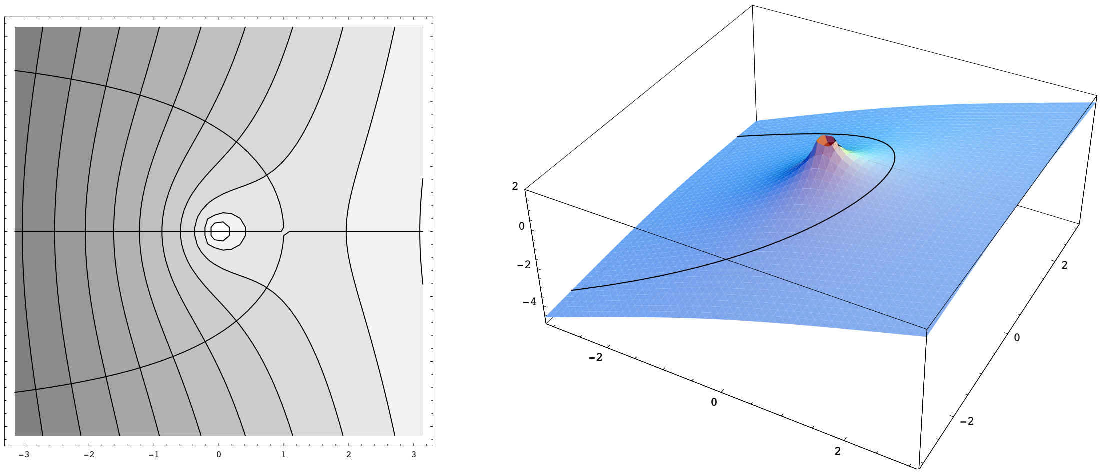
where C' is a contour which starts at $t = -\infty - ia$ ($a > 0$), encircles the branch cut that lies along the negative real axis, and then ends up at $t = -\infty + ib$ ($b > 0$).



This is a moveable saddle problem. Writing $e^t t^{-x} = e^{t - x\ln t}$ and differentiating the whole exponent with respect to t shows that there is a saddle point at $t = x$. Thus we begin by changing the moveable saddle to a fixed saddle by the change of variable $t = xs$ to give

$$\frac{1}{\Gamma(x)} = \frac{1}{2\pi i x^{x-1}}\int_C e^{x(s-\log s)}\,ds = \frac{1}{2\pi i x^{x-1}}\int_C e^{x\phi(s)}\,ds$$

where $\phi = s - \log s$ and C is the rescaled contour (which we could take to be the same as C' by the deformation theorem). The saddle point(s) are now at $\phi'(s) = 0$, *i.e.* $s = 1$.



We deform to the steepest descent contour $\mathrm{Im}(s) = \arg(s)$ ($|\arg(s)| < \pi$) through the saddle as illustrated. Having deformed to the steepest descent contour the procedure is exactly that for Laplace's method. The integral is split into a local contribution from near the saddle and the rest, which is exponentially smaller. For the local contribution ϕ is expanded in a Taylor series about the saddle point $s = 1$ giving

$$\phi(s) \sim 1 + \frac{(s-1)^2}{2} - \frac{(s-1)^3}{3} + \cdots$$

so that

$$\frac{1}{\Gamma(x)} \sim \frac{e^x}{2\pi i x^{x-1}} \int e^{\frac{x(s-1)^2}{2} - \frac{x(s-1)^3}{3} + \cdots}\, ds.$$

At this stage the integral is from $-\epsilon$ to ϵ along the steepest descent contour from the saddle $s = 1$. We then rescale the integration variable so that the quadratic term in the exponent is $O(1)$ by setting $\sqrt{x}\,(s-1) = u$, giving

$$\frac{1}{\Gamma(x)} \sim \frac{e^x}{2\pi i x^{x-1}\sqrt{x}} \int e^{\frac{u^2}{2} - \frac{u^3}{3\sqrt{x}} + \cdots}\, du,$$

where the integral is from $-x^{1/2}\epsilon$ to $x^{1/2}\epsilon$ along the steepest descent contour. We now expand $e^{-\frac{u^3}{3\sqrt{x}} + \cdots}$ keeping only the first term and replace the integration limits by $\pm\infty$ along the steepest descent contour (introducing only exponentially small errors), giving

$$\frac{1}{\Gamma(x)} \sim \frac{e^x}{2\pi i x^{x-1/2}} \int e^{\frac{u^2}{2}}\, du.$$

Now the steepest descent contour is locally parallel to the imaginary axis near to the saddle point $s = 1$, so we set $u = iv$. A comparison with the figure above tells us which way to integrate – in this case from $v = -\infty$ to $v = \infty$. Thus,

$$\frac{1}{\Gamma(x)} \sim \frac{e^x}{2\pi x^{x-1/2}} \int_{-\infty}^{\infty} e^{\frac{-v^2}{2}}\, dv = \frac{e^x}{\sqrt{2\pi}\, x^{x-1/2}},$$

i.e.

$$\Gamma(x) \sim \sqrt{2\pi}\, x^{x-1/2} e^{-x} \quad \text{as } x \to \infty.$$

Example: Steepest descents on the Airy function

1. Positive argument Consider as $x \to \infty$ the Airy function

$$\mathrm{Ai}(x) = \frac{1}{2\pi} \int_{C'} e^{i(t^3/3 + xt)}\, dt,$$

where C' is a contour that starts at infinity with $2\pi/3 < \arg(t) < \pi$ and ends at infinity with $0 < \arg(t) < \pi/3$. Note that the integrand decays at infinity where $\mathrm{Re}(it^3) < 0$, *i.e.* in the sectors defined by $0 < \arg(t) < \pi/3$, $2\pi/3 < \arg(t) < \pi$ and $4\pi/3 < \arg(t) < 5\pi/3$.

This is a moveable saddle problem. Differentiating the whole exponent shows that the saddle points are at $t = \pm ix^{1/2}$. Thus we rescale $t = x^{1/2}z$ to give

$$\mathrm{Ai}(x) = \frac{x^{1/2}}{2\pi}\int_C e^{ix^{3/2}(z^3/3+z)}\,dz = \frac{x^{1/2}}{2\pi}\int_C e^{x^{3/2}\phi(z)}\,dz,$$

where $\phi(z) = i(z^3/3 + z)$ and C is the rescaled contour, which we could take to be the same as C' by the deformation theorem and must start in the sector V_1 and end in the sector V_2 shown in figure 1(a).

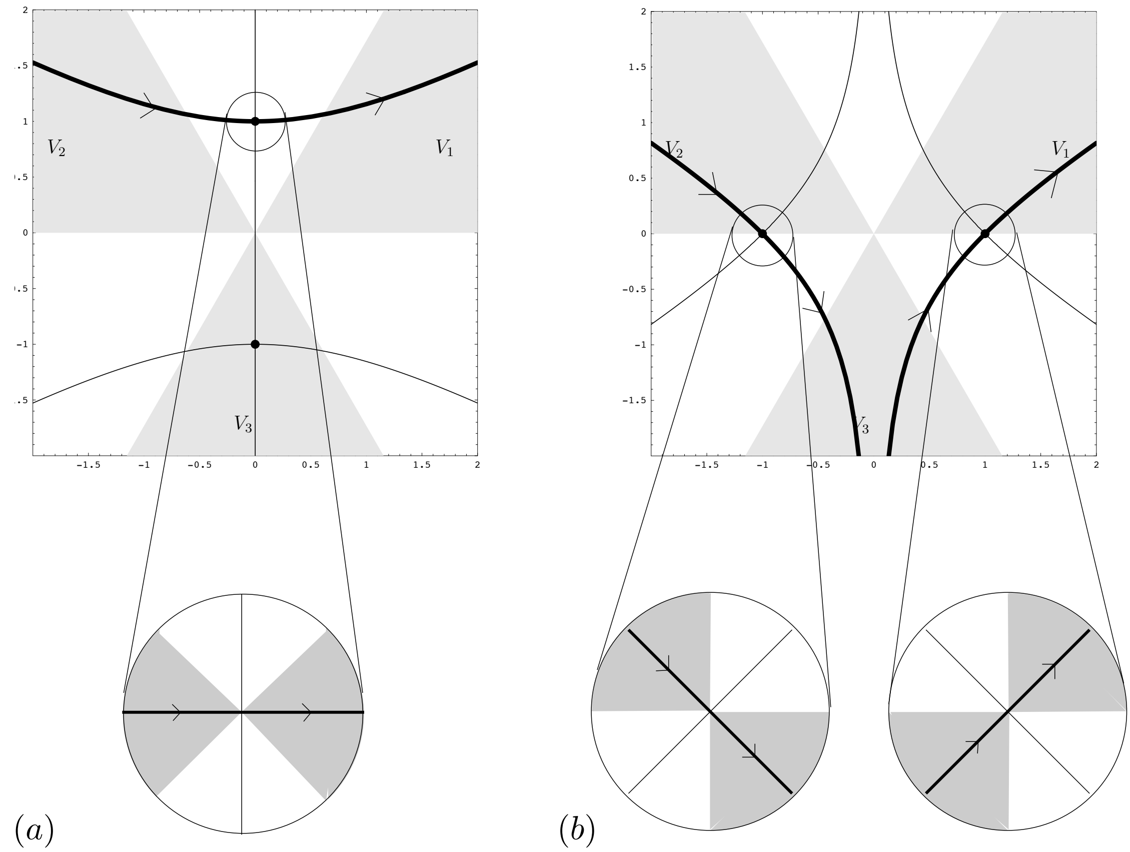


Figure 1: Steepest descent curves for (a) $x \to \infty$ and (b) $x \to -\infty$. Note that the shading shows the sectors which are valleys **at infinity** and is not supposed to be a contour plot of the magntitude of the integrand, but just an aid to determine the steepest descent (rather than ascent) contour.

The saddle points are the points where $\phi'(z) = 0$, *i.e.* $z = \pm i$. We deform the contour C to the steepest descent contour from V_2 to V_1, which goes through the saddle point $z = i$ but not the saddle point $z = -i$.

Having deformed to the steepest descent contour the procedure is exactly that for Laplace's method. The integral is split into a local contribution from near the saddle and the rest, which is exponentially smaller. For the local contribution ϕ is expanded in a Taylor series about the saddle point $z = i$ as

$$\phi(z) \sim -\frac{2}{3} - (z-i)^2 + \cdots,$$

so that

$$\mathrm{Ai}(x) \sim \frac{x^{1/2}e^{-2x^{3/2}/3}}{2\pi}\int e^{-x^{3/2}(z-i)^2+\cdots}\,dz.$$

At this stage the integral is from $-\epsilon$ to ϵ along the steepest descent contour from the saddle $z = i$. Now we change variable by setting $x^{3/4}(z - i) = u$ to give

$$\mathrm{Ai}(x) \sim \frac{e^{-2x^{3/2}/3}}{2\pi x^{1/4}} \int e^{-u^2+\cdots}\, du,$$

where the integral is from $-x^{3/4}\epsilon$ to $x^{3/4}\epsilon$ along the steepest descent contour. We now replace these limits by $\pm\infty$ along the steepest descent contour (introducing only exponentially small errors). Keeping only the leading order term we therefore have

$$\mathrm{Ai}(x) \sim \frac{e^{-2x^{3/2}/3}}{2\pi x^{1/4}} \int e^{-u^2}\, du,$$

where the integral goes to infinity along the steepest descent contour. The steepest descent contour is given by $-u^2$ real and negative, *i.e.* u real. A comparison with figure 1(a) tells us which way to integrate – in this case from $-\infty$ to ∞. Thus

$$\mathrm{Ai}(x) \sim \frac{e^{-2x^{3/2}/3}}{2\pi x^{1/4}} \int_{-\infty}^{\infty} e^{-u^2}\, du = \frac{e^{-2x^{3/2}/3}}{2\sqrt{\pi}\, x^{1/4}}.$$

2. Negative argument Consider as $x \to \infty$ the Airy function

$$\mathrm{Ai}(-x) = \frac{1}{2\pi} \int_{C'} e^{i(t^3/3 - xt)}\, dt,$$

with C' as before. As before, we rescale $t = x^{1/2} z$ to give

$$\mathrm{Ai}(-x) = \frac{x^{1/2}}{2\pi} \int_C e^{ix^{3/2}(z^3/3 - z)}\, dz = \frac{x^{1/2}}{2\pi} \int_C e^{x^{3/2}\phi(z)}\, dz,$$

where C is as before, but now $\phi(z) = i(z^3/3 - z)$. The saddle points are the points where $\phi'(z) = 0$, *i.e.* $z = \pm 1$. The steepest descent contour through $z = 1$ goes from V_3 to V_1. The steepest descent contour through $z = -1$ goes from V_3 to V_2 (see figure 1(b)). Thus we must deform the contour C to go from V_2 to V_3 through the saddle at $z = -1$, and then from V_3 to V_1 through the saddle at $z = 1$. Thus in this case both saddles will contribute to the integral.

Near $z = 1$ we expand ϕ as a Taylor series

$$\phi(z) \sim -\frac{2i}{3} + i(z-1)^2 + \cdots$$

to give

$$\frac{x^{1/2} e^{-2ix^{3/2}/3}}{2\pi} \int e^{ix^{3/2}(z-1)^2 + \cdots}\, dz.$$

We change variable by setting $x^{3/4}(z - 1) = u$ to give

$$\frac{e^{-i2x^{3/2}/3}}{2\pi x^{1/4}} \int e^{iu^2}\, du.$$

As usual we now replace the integration limits by $\pm\infty$ along the steepest descent contour. The steepest descent contour is given by iu^2 real and negative, *i.e.* $u = e^{i\pi/4} s$ with s real. A comparison with figure 1(b) tells us which way to integrate – in this case from $s = -\infty$ to $s = +\infty$. Thus the contribution from $z = 1$ is

$$\frac{e^{i\pi/4} e^{-2ix^{3/2}/3}}{2\pi x^{1/4}} \int_{-\infty}^{\infty} e^{-s^2}\, ds = \frac{e^{i\pi/4} e^{-2ix^{3/2}/3}}{2\sqrt{\pi}\, x^{1/4}}.$$

Near $z = -1$ we expand ϕ as a Taylor series

$$\phi(z) \sim \frac{2i}{3} - i(z+1)^2 + \cdots$$

to give

$$\frac{x^{1/2} e^{2ix^{3/2}/3}}{2\pi} \int e^{-ix^{3/2}(z+1)^2 + \cdots} \, dz.$$

We change variable by setting $x^{3/4}(z+1) = u$ to give

$$\frac{e^{2ix^{3/2}/3}}{2\pi x^{1/4}} \int e^{-iu^2} \, du.$$

As usual we now replace the integration limits by $\pm\infty$ along the steepest descent contour. The steepest descent contour is given by $-iu^2$ real and negative, *i.e.* $u = e^{3i\pi/4} s$ with s real. A comparison with figure 1(b) tells us which way to integrate – in this case from $s = \infty$ to $s = -\infty$. Thus the contribution from $z = -1$ is

$$\frac{e^{3i\pi/4} e^{2ix^{3/2}/3}}{2\pi x^{1/4}} \int_{\infty}^{-\infty} e^{-s^2} \, ds = \frac{e^{-i\pi/4} e^{2ix^{3/2}/3}}{2\sqrt{\pi}\, x^{1/4}}.$$

Adding together the two contributions we find

$$\text{Ai}(-x) \sim \frac{e^{i\pi/4} e^{-2ix^{3/2}/3}}{2\sqrt{\pi}\, x^{1/4}} + \frac{e^{-i\pi/4} e^{2ix^{3/2}/3}}{2\sqrt{\pi}\, x^{1/4}} = \frac{1}{\sqrt{\pi}\, x^{1/4}} \cos\left(\frac{\pi}{4} - \frac{2x^{3/2}}{3}\right).$$

4.8 Splitting the range of integration

We have seen in the previous examples how to split the range of integration into a local part, in which some functions may be approximated by Taylor series, and a global part, which in the previous cases was lower order. In general we may follow such a procedure, splitting the range of integration and using different approximations in each range.

Example 1

$$\int_0^1 \frac{dx}{(x+\epsilon)^{1/2}}.$$

On the one hand we would like to expand the integrand for small ϵ:

$$\frac{1}{(x+\epsilon)^{1/2}} \sim \frac{1}{x^{1/2}} - \frac{\epsilon}{2x^{3/2}} + \cdots.$$

However, such an expansion is only OK if $\epsilon \ll x$. Thus there are two regions to consider, $x = O(1)$ and $x = O(\epsilon)$.

- If $x = O(\epsilon)$ the integrand is $O(\epsilon^{-1/2})$ and contribution to the integral is therefore $O(\epsilon^{1/2})$.
- If $x = O(1)$ the integrand is $O(1)$ and contribution to the integral is therefore $O(1)$.

Thus we expect the global contribution to dominate.

We split the range of integration from 0 to δ and from δ to x, where $\epsilon \ll \delta \ll 1$. We write

$$\int_0^1 \frac{dx}{(x+\epsilon)^{1/2}} = \int_0^\delta \frac{dx}{(x+\epsilon)^{1/2}} + \int_\delta^1 \frac{dx}{(x+\epsilon)^{1/2}}.$$

In the first integral we rescale $x = \epsilon u$ to give

$$\int_0^1 \frac{dx}{(x+\epsilon)^{1/2}} = \int_0^{\delta/\epsilon} \frac{\epsilon^{1/2} du}{(u+1)^{1/2}} + \int_\delta^1 \frac{dx}{(x+\epsilon)^{1/2}}.$$

Now we are safe to Taylor series the second integrand. The first integral is

$$\int_0^{\delta/\epsilon} \frac{\epsilon^{1/2}du}{(u+1)^{1/2}} = -2\epsilon^{1/2} + 2(\epsilon+\delta)^{1/2}.$$

The second is

$$\begin{aligned}
\int_\delta^1 \frac{dx}{(x+\epsilon)^{1/2}} &\sim \int_\delta^1 \left(\frac{1}{x^{1/2}} - \frac{\epsilon}{2x^{3/2}} + \cdots\right) dx \\
&\sim 2 - 2\delta^{1/2} + \epsilon - \frac{\epsilon}{\delta^{1/2}} + \cdots .
\end{aligned}$$

Hence

$$\begin{aligned}
\int_0^1 \frac{dx}{(x+\epsilon)^{1/2}} &\sim -2\epsilon^{1/2} + 2(\epsilon+\delta)^{1/2} + 2 - 2\delta^{1/2} + \epsilon - \frac{\epsilon}{\delta^{1/2}} + \cdots \\
&\sim -2\epsilon^{1/2} + 2\delta^{1/2} + \frac{\epsilon}{\delta^{1/2}} + \cdots 2 - 2\delta^{1/2} + \epsilon - \frac{\epsilon}{\delta^{1/2}} + \cdots \\
&\sim 2 - 2\epsilon^{1/2} + \epsilon + \cdots ,
\end{aligned}$$

remembering that $\epsilon \ll \delta$. Notice that the final answer is independent of δ as it should be. We can check that our answer is right by comparing with the exact solution

$$2\left((1+\epsilon)^{1/2} - \epsilon^{1/2}\right) \sim 2 - 2\epsilon^{1/2} + \epsilon + \cdots .$$

Example 2

$$I = \int_0^{\pi/4} \frac{d\theta}{\epsilon^2 + \sin^2\theta}.$$

As before there are two regions, $\theta = O(1)$ and $\theta = O(\epsilon)$.

- If $\theta = O(\epsilon)$ the integrand is $O(\epsilon^{-2})$ and contribution to the integral is therefore $O(\epsilon^{-1})$.
- If $\theta = O(1)$ the integrand is $O(1)$ and contribution to the integral is therefore $O(1)$.

Thus we expect the local contribution to dominate.

As before we split the range of integration at δ, with $\epsilon \ll \delta \ll 1$.

$$I = \int_0^\delta \frac{d\theta}{\epsilon^2 + \sin^2\theta} + \int_\delta^{\pi/4} \frac{d\theta}{\epsilon^2 + \sin^2\theta}.$$

In the first integral we rescale $\theta = \epsilon u$ to give

$$I = \int_0^{\delta/\epsilon} \frac{\epsilon du}{\epsilon^2 + \sin^2(\epsilon u)} + \int_\delta^{\pi/4} \frac{d\theta}{\epsilon^2 + \sin^2\theta}.$$

Now in the first integral $\epsilon u \leq \delta \ll 1$ so we are safe to Taylor expand $\sin^2(\epsilon u)$, giving

$$\begin{aligned}
\int_0^{\delta/\epsilon} \frac{\epsilon du}{\epsilon^2 + \sin^2(\epsilon u)} &\sim \int_0^{\delta/\epsilon} \frac{\epsilon du}{\epsilon^2 + \epsilon^2 u^2 - \epsilon^4 u^4/3 + \cdots} \\
&\sim \int_0^{\delta/\epsilon} \left(\frac{1}{\epsilon(1+u^2)} + \frac{\epsilon u^4}{3(1+u^2)^2} + \cdots\right) du \\
&= \frac{1}{\epsilon}\tan^{-1}\frac{\delta}{\epsilon} + O(\epsilon) \\
&= \frac{\pi}{2\epsilon} - \frac{1}{\delta} + \cdots + O(\epsilon).
\end{aligned}$$

In the second integral we can expand the integrand in powers of ϵ to give

$$\int_{\delta}^{\pi/4} \frac{d\theta}{\epsilon^2 + \sin^2\theta} \sim \int_{\delta}^{\pi/4} \left(\frac{1}{\sin^2\theta} - \frac{\epsilon^2}{\sin^4\theta} + \cdots \right) d\theta = -1 + \cot(\delta) + O(\epsilon^2) \sim -1 + \frac{1}{\delta} + \cdots .$$

Hence

$$I \sim \frac{\pi}{2\epsilon} - 1 + \cdots .$$

5 Matched Asymptotic expansions

5.1 Singular Perturbations

If a differential equation $D_\epsilon y = 0$ has a small parameter ϵ in it (D_ϵ is the differential operator associated with this differential equation), it is natural to aim to use the solution of the limiting case $D_0 y = 0$ (corresponding to $\epsilon = 0$) as an approximaton for the solution of $D_\epsilon y = 0$. However, if ϵ multiplies the **highest** derivative of y, say $d^k y/dx^k$, a difficulty arises. The original $D_\epsilon y = 0$ is a k-th order equation with k boundary conditions. However, $D_0 y = 0$ only has order $\leq k-1$, so it cannot satisfy all of the boundary conditions (in general). This is called a **singular perturbation problem**: the operator D_ϵ is a singular perturbation of D_0.

Linear example Consider

$$\epsilon y'' + y' + y = 0,$$

$$y(0) = a, \quad y(1) = b,$$

where a and b are prescribed real constants. When $\epsilon = 0$, we have

$$y' + y = 0,$$

$$y(0) = a, \quad y(1) = b.$$

The solution is $y = Ae^{-x}$ which cannot satisfy both boundary conditions in general.

If y is the solution to $D_\epsilon y = 0$ then one possible behaviour in such cases is that

- over most of the range $\epsilon d^k y/dx^k$ is small, and y approximately obeys $D_0 y = 0$.
- in certain regions, often near the ends of the range, $\epsilon d^k y/dx^k$ is **not** small, and y adjusts itself to the boundary conditions (*i.e.* $d^k y/dx^k$ is large in some places).

In fluid dynamics these regions are known as boundary layers, in solid mechanics they are known as edge layers, in electrodymanics they are known as skin layers, etc.

A procedure for determining the solution of a singular perturbation problem with boundary layers is

(1) Determine the scaling of the boundary layers (*e.g.* $x \propto \epsilon$ or $\epsilon^{1/2}$ or $\ldots$.)

(2) Rescale the independent variable in the boundary layer (*e.g.* $x = \hat{x}\epsilon$ or $\hat{x}\epsilon^{1/2}$ or $\ldots$.)

(3) Find the asymptotic expansions of the solutions in the boundary layers and outside the boundary layers (the "inner" and "outer" solutions).

(4) Fix the arbitrary constants in these solutions by

  (a) inner solutions obey the boundary conditions.

  (b) "matching"—making the inner and outer solutions join up properly in the transition region between them.

This is the method of **matched asymptotic expansions**. (You will do something similar later when you examine turning points that arise in the WKB method.)

We will illustrate the procedure with our linear example. Note that this problem can be solved exactly. We will work as if we don't have any *a priori* knowledge about the solution - *i.e.* as though there may be boundary layers at either end, even though the boundary layer is actually only at $x = 0$.

Scaling Near $x = 0$ we let $x_L = x/\epsilon^\alpha$ (L = left-hand end; x_L = local variable for inspecting the boundary layer on the left.) Then we write $y(x) = y_L(x_L)$. Then

$$\begin{aligned} \frac{dy}{dx} &= \frac{dy_L}{dx_L}\frac{dx_L}{dx} = \epsilon^{-\alpha}\frac{dy_L}{dx_L}, \\ \frac{d^2y}{dx^2} &= \epsilon^{-2\alpha}\frac{d^2y_L}{dx_L^2}, \end{aligned}$$

so that

$$\epsilon^{1-2\alpha}\frac{d^2y_L}{dx_L^2} + \epsilon^{-\alpha}\frac{dy_L}{dx_L} + y_L = 0.$$

Now in the boundary layer d^2y_L/dx^2 is significant. We must increase α until this term balances the **largest** of the others in the equation. Hence we want

$$1 - 2\alpha = \min(-\alpha, 0), \qquad i.e.\ \alpha = 1.$$

So the boundary layer is of width ϵ. Note that if we choose $1 - 2\alpha = 0$, *i.e.* $\alpha = 1/2$, to balance the first and third terms, then the second term is $O(\epsilon^{-1/2})$ which is bigger than the other two. The boundary layer at the right is also of width ϵ.

So now we develop our asymptotic expansion as follows:

(1) Away from the ends of the interval ("the middle") we expand y as

$$y(x) = y_M(x) \sim y_{M0}(x) + \epsilon y_{M1}(x) + \cdots,$$

(2) near the left-hand end we rescale x by a factor of ϵ so we have $x_L = x/\epsilon$ and we expand

$$y(x) = y_L(x_L) \sim y_{L0}(x_L) + \epsilon y_{L1}(x_L) + \cdots,$$

(3) near the right-hand end we rescale $x-1$ by a factor of ϵ so we have $x_R = (x-1)/\epsilon \le 0$ and we expand

$$y(x) = y_R(x_R) \sim y_{R0}(x_R) + \epsilon y_{R1}(x_R) + \cdots.$$

Solution on left The equation in the inner variable reads

$$\frac{d^2y_L}{dx_L^2} + \frac{dy_L}{dx_L} + \epsilon y_L = 0.$$

Inserting the expansion and equating coefficients of powers of ϵ gives

$$\begin{aligned} O(1): &\qquad \frac{d^2y_{L0}}{dx_L^2} + \frac{dy_{L0}}{dx_L} = 0, \\ O(\epsilon): &\qquad \frac{d^2y_{L1}}{dx_L^2} + \frac{dy_{L1}}{dx_L} + y_{L0} = 0, \end{aligned}$$

etc. Hence

$$y_{L0} = A_{L0} + B_{L0}e^{-x_L}.$$

To satisfy $y(0) = a$ we have $A_{L0} + B_{L0} = a$.

Solution in middle The equation in the outer variable reads

$$\epsilon \frac{d^2 y_M}{dx^2} + \frac{dy_M}{dx} + y_M = 0.$$

Inserting the expansion and equating coefficients of powers of ϵ gives

$$\begin{aligned} O(1): &\quad \frac{dy_{M0}}{dx} + y_{M0} = 0, \\ O(\epsilon): &\quad \frac{d^2 y_{M0}}{dx^2} + \frac{dy_{M1}}{dx} + y_{M1} = 0, \end{aligned}$$

etc. Hence

$$y_{M0} = A_{M0} e^{-x}.$$

Solution on right The equation in the inner variable reads

$$\frac{d^2 y_R}{dx_R^2} + \frac{dy_R}{dx_R} + \epsilon y_R = 0.$$

Inserting the expansion and equating coefficients of powers of ϵ gives

$$\begin{aligned} O(1): &\quad \frac{d^2 y_{R0}}{dx_R^2} + \frac{dy_{R0}}{dx_R} = 0, \\ O(\epsilon): &\quad \frac{d^2 y_{R1}}{dx_R^2} + \frac{dy_{R1}}{dx_R} + y_{R0} = 0, \end{aligned}$$

etc. Hence

$$y_{R0} = A_{R0} + B_{R0} e^{-x_R}.$$

To satisfy $y(1) = b$ we have $A_{R0} + B_{R0} = b$.

So far we have 5 arbitrary constants and 2 equations. The other 3 equations will come by matching.

Matching Idea is that there is an "overlap" region where both expansions should hold and therefore be equal.

$$y_L(x_L) \sim y_M(x) \text{ as } x \to 0 \text{ and } x_L = x/\epsilon \to \infty.$$

One way is to introduce an "intermediate" scaling $\hat{x} = x/\epsilon^\alpha$, $0 < \alpha < 1$. Then with $\epsilon \to 0$ with $\hat{x}$ fixed we have $x = \epsilon^\alpha \hat{x} \to 0$ and $x_L = \epsilon^{\alpha-1}\hat{x} \to \infty$. Often it is easiest to choose some fixed α, say $\alpha = 1/2$. In this case, matching at the left-hand end we have

$$\begin{aligned} y_L &= A_{L0} + B_{L0} e^{-\epsilon^{\alpha-1}\hat{x}} + O(\epsilon), \\ &= A_{L0} + O(\epsilon), \end{aligned}$$

while

$$\begin{aligned} y_M &= A_{M0} e^{-\epsilon^\alpha \hat{x}} + O(\epsilon) \\ &= A_{M0} - \epsilon^\alpha \hat{x} A_{M0} + \cdots + O(\epsilon). \end{aligned}$$

For these to be the same as $\epsilon \to 0$ we need

$$A_{L0} = A_{M0}.$$

Thus the y values have to match: the outer limit of the inner problem needs to match with the inner limit of the outer problem.

Matching at the right-hand end we use the intermediate variable $\tilde{x} = (x-1)/\epsilon^\alpha \leq 0$ giving

$$y_R = A_{R0} + B_{R0} e^{-\epsilon^{\alpha-1}\tilde{x}} + O(\epsilon), \tag{1}$$

while

$$\begin{aligned} y_M &= A_{M0} e^{-1-\epsilon^\alpha \tilde{x}} + O(\epsilon) \\ &= A_{M0} e^{-1} - \epsilon^\alpha \tilde{x} A_{M0} e^{-1} + \cdots + O(\epsilon). \end{aligned}$$

Clearly to match we cannot have exponential growth in (1), so that

$$B_{R0} = 0, \qquad A_{R0} = A_{M0} e^{-1}.$$

Again the y values must match. Hence our 5 equations are

$$\begin{aligned} A_{L0} + B_{L0} &= a, \\ A_{R0} + B_{R0} &= b, \\ A_{L0} &= A_{M0}, \\ B_{R0} &= 0, \\ A_{R0} &= A_{M0} e^{-1}. \end{aligned}$$

Hence

$$A_{L0} = eb, \quad B_{L0} = a - eb, \quad A_{M0} = eb, \quad A_{R0} = b, \quad B_{R0} = 0,$$

and the solution in the three regions is given by

$$\begin{aligned} y_{L0} &= eb + (a-eb) e^{-x_L}, \\ y_{M0} &= ebe^{-x}, \\ y_{R0} &= b. \end{aligned}$$

There is no rapid variation in y in the right-hand boundary layer - we do not really need this boundary layer.

Composite Expansion To plot the solution for example we want a uniformly valid expansion. One way to construct a uniformly valid approximation is to add together the solution in the inner and outer regions, and then subtract the solution in the "overlap" region which has been counted twice. Write the inner solution in terms of outer variables and the outer in terms of inner variables and expand

$$\begin{aligned} y_{L0} &= eb + (a-eb) e^{-x/\epsilon} = eb + O(\epsilon), \\ y_{M0} &= ebe^{-\epsilon x_L} = eb + O(\epsilon). \end{aligned}$$

The common term which has been counted twice is eb. Hence the composite expansion is

$$y \sim eb + (a-eb) e^{-x_L} + ebe^{-x} - eb = (a-eb) e^{-x/\epsilon} + ebe^{-x}.$$

The error is $O(\epsilon)$ over the whole range of x.

Higher-order terms At order ϵ in each region

$$\begin{aligned} y_{L1} &= -ebx_L + (a-eb) x_L e^{-x_L} + A_{L1} + B_{L1} e^{-x_L}, \\ y_{M1} &= -ebxe^{-x} + A_{M1} e^{-x}, \\ y_{R1} &= -bx_R + A_{R1} + B_{R1} e^{-x_R}. \end{aligned}$$

Boundary conditions:

$$A_{L1} + B_{L1} = 0, \qquad A_{R1} + B_{R1} = 0.$$

Matching Follow same procedure. At left-hand end write inner and outer expansions in terms of $\hat{x}$:

$$\begin{aligned} y_L &= eb + (a-eb)e^{-\epsilon^{\alpha-1}\hat{x}} + \epsilon\left(-eb\epsilon^{\alpha-1}\hat{x} + (a-eb)\epsilon^{\alpha-1}\hat{x}e^{-\epsilon^{\alpha-1}\hat{x}} + A_{L1} + B_{L1}e^{-\epsilon^{\alpha-1}\hat{x}}\right)\epsilon + O(\epsilon^2), \\ &= eb - eb\epsilon^{\alpha}\hat{x} + A_{L1}\epsilon + O(\epsilon^2) \end{aligned}$$

while

$$\begin{aligned} y_M &= ebe^{-\epsilon^{\alpha}\hat{x}} + \epsilon\left(-eb\epsilon^{\alpha}\hat{x}e^{-\epsilon^{\alpha}\hat{x}} + A_{M1}e^{-\epsilon^{\alpha}\hat{x}}\right) + O(\epsilon^2) \\ &= eb - \epsilon^{\alpha}\hat{x}eb + \frac{\epsilon^{2\alpha}\hat{x}^2}{2}eb + \cdots \\ &\quad - eb\epsilon^{\alpha+1}\hat{x} + eb\epsilon^{2\alpha+1}\hat{x}^2 + A_{M1}\epsilon - A_{M1}\epsilon^{1+\alpha}\hat{x} + \cdots \\ &\quad + O(\epsilon^2). \end{aligned}$$

Matching we find that

$$A_{L1} = A_{M1}.$$

Note some terms jump order: $-\epsilon^{\alpha}\hat{x}eb$ comes from the inner expansion of the first-outer term, but from the outer expansion of the second-inner term. Note that in order for the neglected terms $\epsilon^{2\alpha}$ to be smaller than the last retained term ϵ we need $\alpha > 1/2$.

At the right-hand end using the intermediate variable $\tilde{x}$ we find

$$\begin{aligned} y_R &= b + \epsilon\left(-b\epsilon^{\alpha-1}\tilde{x} + A_{R1} + B_{R1}e^{-\epsilon^{\alpha-1}\tilde{x}}\right) + O(\epsilon^2), \\ &= b - b\epsilon^{\alpha}\tilde{x} + \epsilon A_{R1} + \epsilon B_{R1}e^{-\epsilon^{\alpha-1}\tilde{x}} + O(\epsilon^2), \end{aligned}$$

while

$$\begin{aligned} y_M &= ebe^{-1-\epsilon^{\alpha}\tilde{x}} + \epsilon\left(-eb(1+\epsilon^{\alpha}\tilde{x})e^{-1-\epsilon^{\alpha}\tilde{x}} + A_{M1}e^{-1-\epsilon^{\alpha}\tilde{x}}\right) + O(\epsilon^2) \\ &= b - \epsilon^{\alpha}\tilde{x}b + \frac{\epsilon^{2\alpha}\tilde{x}^2}{2}b + \cdots \\ &\quad - b(\epsilon + \epsilon^{\alpha+1}\tilde{x})\left(1 - \epsilon^{\alpha}\tilde{x} + \cdots\right) + \epsilon A_{M1}e^{-1}\left(1 - \epsilon^{\alpha}\tilde{x} + \cdots\right) \\ &\quad + O(\epsilon^2). \end{aligned}$$

Matching gives

$$B_{R1} = 0, \qquad A_{M1}e^{-1} - b = A_{R1}.$$

Hence we now have the 5 equations for 5 unknowns:

$$\begin{aligned} A_{L1} + B_{L1} &= 0, \\ A_{R1} + B_{R1} &= 0, \\ A_{L1} &= A_{M1}, \\ B_{R1} &= 0, \\ A_{M1}e^{-1} - b &= A_{R1}, \end{aligned}$$

with solution

$$A_{R1} = 0, \quad B_{R1} = 0, \quad A_{M1} = be, \quad A_{L1} = be, \quad B_{L1} = -be.$$

This gives

$$\begin{aligned} y_{L1} &= -ebx_L + (a-eb)x_Le^{-x_L} + eb - ebe^{-x_L}, \\ y_{M1} &= -ebxe^{-x} + ebe^{-x}, \\ y_{R1} &= -bx_R. \end{aligned}$$

Composite expansion Composite is $y_L + y_M -$ overlap. Write y_L in terms of the outer variable:

$$\begin{aligned} y_L &= eb + (a-eb)e^{-x/\epsilon} + \epsilon\left(-eb\frac{x}{\epsilon} + (a-eb)\frac{x}{\epsilon}e^{-x/\epsilon} + eb - ebe^{-x/\epsilon}\right) + \cdots \\ &= eb - ebx + \epsilon eb + O(\epsilon^2); \end{aligned}$$

and y_M in terms of the inner variables:

$$\begin{aligned} y_M &= ebe^{-\epsilon x_L} + \epsilon\left(-eb\epsilon x_L e^{-\epsilon x_L} + ebe^{-\epsilon x_L}\right) + \cdots \\ &= eb - eb\epsilon x_L + \epsilon eb + O(\epsilon^2). \end{aligned}$$

The common value in the overlap region is

$$eb - eb\epsilon x_L + \epsilon eb = eb - ebx + \epsilon eb.$$

Hence the composite expansion is

$$\begin{aligned} y &= eb + (a-eb)e^{-x/\epsilon} + \epsilon\left(-eb\frac{x}{\epsilon} + (a-eb)\frac{x}{\epsilon}e^{-x/\epsilon} + eb - ebe^{-x/\epsilon}\right) \\ &\quad + ebe^{-x} + \epsilon\left(-ebxe^{-x} + ebe^{-x}\right) - (eb - ebx + \epsilon eb) + O(\epsilon^2) \\ &= (a-eb)e^{-x/\epsilon} + (a-eb)xe^{-x/\epsilon} - \epsilon ebe^{-x/\epsilon} \\ &\quad + ebe^{-x} - \epsilon ebxe^{-x} + \epsilon ebe^{-x} + O(\epsilon^2). \end{aligned}$$

Van Dyke's matching rule Using the intermediate variable $\hat{x}$ is tiresome. Van Dyke's matching 'rule' usually works and is much more convenient. (However, at the end of the day it's a matter of keeping track of the size of each term, as you'll see in this discussion.) Van Dyke's rule is

$$(m \text{ term inner})(n \text{ term outer}) = (n \text{ term outer})(m \text{ term inner}).$$

I.e. in the outer variables expand to n terms; then switch to inner variables and reexpand to m terms. The result is the same as first expanding in the inner to m terms, then switching to outer variables and reexpanding to n terms.

Example revisited

$$(1\text{to}) = A_{M0}e^{-x}.$$

In inner variables this is

$$A_{M0}e^{-\epsilon x_L}.$$

Expanded this is

$$A_{M0} - A_{M0}\epsilon x_L + A_{M0}\frac{\epsilon^2 x_L^2}{2} + \cdots.$$

Hence

$$\begin{aligned} (1\text{ti})(1\text{to}) &= A_{M0}, \\ (2\text{ti})(1\text{to}) &= A_{M0} - A_{M0}\epsilon x_L, \end{aligned}$$

etc. Similarly

$$(1\text{ti}) = A_{L0} + B_{L0}e^{-x_L}.$$

In outer variables this is

$$A_{L0} + B_{L0}e^{-x/\epsilon}.$$

Expanded this is

$$A_{L0} + E.S.T.,$$

where $E.S.T.$ means "exponentially small term". Hence

$$\begin{aligned}(1\text{to})(1\text{ti}) &= A_{L0},\\ (2\text{to})(1\text{ti}) &= A_{L0},\end{aligned}$$

etc. So

$$(1\text{to})(1\text{ti}) = (1\text{ti})(1\text{to}) \Rightarrow A_{M0} = A_{L0}.$$

Warning: When using this matching rule you must treat log as $O(1)$ because of the size of logarithmic terms.

Choice of scaling revisited Near $x = 0$ we let $x_L = x/\epsilon^\alpha$, and $y(x) = y_L(x_L)$ so that

$$\epsilon^{1-2\alpha}\frac{d^2y_L}{dx_L^2} + \epsilon^{-\alpha}\frac{dy_L}{dx_L} + y_L = 0.$$

Now as we gradually increase α we find

	$\epsilon^{1-2\alpha}\frac{d^2y_L}{dx_L^2}$	$+$	$\epsilon^{-\alpha}\frac{dy_L}{dx_L}$	$+$	y_L	$=0$
$\alpha = 0$			....	balance	....	the outer
$0 < \alpha < 1$			dominant			the overlap
$\alpha = 1$	....	balance	....			the inner
$1 < \alpha$	dominant					the sub-inner

The inner and outer regions can be matched because they share a common term, which is dominant in the overlap region.

The potentially interesting scalings in an equation are those which balance two or more terms. Such scalings are sometimes called **distinguished limits**.

5.2 Where is the boundary layer?

To have the possibility of a non-trivial boundary layer we need some solution in the inner region which decays as we move towards the outer. In the problem we considered, the non-constant solution in the right-hand "boundary layer" grew exponentially as we moved to the outer, so there could never be a boundary layer at $x = 1$.

Boundary layers do not have to be at boundaries! There can be thin regions of high gradients in the interior of the domain (they are then sometimes called interior layers)

Example

$$\epsilon y'' + p(x)y' + q(x)y = 0 \quad \text{for } 0 < x < 1, \quad \text{with } y(0) = A\,, \quad y(1) = B,$$

where A, B are prescribed constants. If $p(x) > 0$ for all $x \in [0, 1]$, then we expect to find a boundary layer at $x = 0$. If $p(x) < 0$ for all $x \in [0, 1]$, then we expect to find a boundary layer at $x = 1$. If $p(x) = 0$ for some $x = x_0$, then there may be an interior layer at $x = x_0$.

Example

$$\epsilon^2 f'' + 2f(1 - f^2) = 0 \quad \text{for } -1 < x < 1, \quad \text{with } f(-1) = -1 \text{ and } \quad f(1) = 1.$$

The outer solution $f = 1$ is OK near $x = 1$, while the outer solution $f = -1$ is OK near $x = -1$. Somewhere there must be a transition between these two states. Rescale near $x = x_0$ by setting $x = x_0 + \epsilon X$ to give

$$\frac{d^2 f}{dX^2} + 2f(1 - f^2) = 0 \qquad \text{in } -\infty < X < \infty$$

$$\text{with} \quad f \to -1 \text{ as } X \to -\infty \qquad \text{and} \quad f \to 1 \text{ as } X \to +\infty,$$

This transition layer has solution

$$f = \tanh(X)$$

for any x_0. In this case the exact solution is

$$f \sim \tanh(x/\epsilon)$$

and the transition layer is near $x = 0$. This could be argued by symmetry. However, the position of the transition layer is exponentially sensitive to the boundary data. Finding it for other data is nontrivial. Not all transition layer problems are so hard.

5.3 Boundary layers in PDEs.

Example Consider the heat transfer from a cylinder in potential flow with small diffusion (high Peclet number). Thus we have to solve

$$\boldsymbol{u} \cdot \nabla T = \epsilon \nabla^2 T \quad \text{in } r \geq 1$$

where

$$\boldsymbol{u} = \nabla \phi, \qquad \phi = \left(r + \frac{1}{r}\right) \cos\theta,$$

with boundary conditions

$$T = 1 \text{ on } r = 1, \qquad T \to 0 \text{ as } r \to \infty.$$

Outer solution
Expand

$$T \sim T_0 + \epsilon T_1 + \cdots \quad \text{as } \epsilon \to 0,$$

Substitute in giving

$$\text{At } \epsilon^0: \qquad \boldsymbol{u} \cdot \nabla T_0 = 0.$$

Hence T_0 is constant on streamlines. Since all (almost all: not the cylinder itself or the wake) streamlines start at $x = -\infty$, where $T_0 = 0$, this means that $T_0 = 0$. Proceeding with the expansion gives $T_n = 0$ for all n.

There is a thermal boundary layer near the cylinder.

Inner solution
In cylindrical coordinates the equation is

$$\left(1 - \frac{1}{r^2}\right) \cos\theta \frac{\partial T}{\partial r} - \left(1 + \frac{1}{r^2}\right) \frac{\sin\theta}{r} \frac{\partial T}{\partial \theta} = \epsilon \left(\frac{\partial^2 T}{\partial r^2} + \frac{1}{r}\frac{\partial T}{\partial r} + \frac{1}{r^2}\frac{\partial^2 T}{\partial \theta^2} \right).$$

Need to scale r close to one so that diffusion becomes important. Set $r = 1 + \delta\rho$:

$$\left(1 - \frac{1}{(1+\delta\rho)^2}\right) \frac{\cos\theta}{\delta} \frac{\partial T}{\partial \rho} - \left(1 + \frac{1}{(1+\delta\rho)^2}\right) \frac{\sin\theta}{(1+\delta\rho)} \frac{\partial T}{\partial \theta} =$$
$$\epsilon \left(\frac{1}{\delta^2} \frac{\partial^2 T}{\partial \rho^2} + \frac{1}{\delta(1+\delta\rho)} \frac{\partial T}{\partial \rho} + \frac{1}{(1+\delta\rho)^2} \frac{\partial^2 T}{\partial \theta^2} \right).$$

i.e.

$$(2\delta\rho + \cdots) \frac{\cos\theta}{\delta} \frac{\partial T}{\partial \rho} - (2 + \cdots) \sin\theta \frac{\partial T}{\partial \theta} = \epsilon \left(\frac{1}{\delta^2} \frac{\partial^2 T}{\partial \rho^2} + \frac{1}{\delta} (1 + \cdots) \frac{\partial T}{\partial \rho} + (1 + \cdots) \frac{\partial^2 T}{\partial \theta^2} \right).$$

Hence $\delta = \epsilon^{1/2}$ and expanding $T \sim \widehat{T}_0(\rho, \theta) + \epsilon^{1/2} \widehat{T}_1(\rho, \theta) + \cdots$ as $\epsilon \to 0$ gives at leading order the boundary layer equation

$$2\rho \cos\theta \frac{\partial \widehat{T}_0}{\partial \rho} - 2 \sin\theta \frac{\partial \widehat{T}_0}{\partial \theta} = \frac{\partial^2 \widehat{T}_0}{\partial \rho^2},$$

with $\widehat{T}_0 = 1$ on $\rho = 0$ and $\widehat{T}_0 \to 0$ as $\rho \to \infty$. Lie group analysis shows that the solution is of similarity form:

$$\widehat{T}_0 = f(\eta), \quad \eta = \frac{\rho \sin\theta}{(1+\cos\theta)^{1/2}}.$$

This gives

$$\frac{2\rho\cos\theta\sin\theta}{(1+\cos\theta)^{1/2}} f' - 2\sin\theta\left(\frac{\rho\cos\theta}{(1+\cos\theta)^{1/2}} + \frac{\rho\sin^2\theta}{2(1+\cos\theta)^{3/2}}\right) f' = \frac{\sin^2\theta}{(1+\cos\theta)} f'',$$

which is

$$f'' + \eta f' = 0.$$

Hence

$$f = A\int_\eta^\infty e^{-u^2/2}\, du + B.$$

$f \to 0$ as $\eta \to \infty$ gives $B = 0$. $f = 1$ on $\eta = 0$ gives $A = \sqrt{2/\pi}$. Hence the boundary layer solution is

$$\widehat{T}_0 = \sqrt{\frac{2}{\pi}} \int_\eta^\infty e^{-u^2/2}\, du.$$

As $\rho \to \infty$ this decays exponentially; the solution in the outer region is exponentially small.

Note that the boundary layer solution works providing θ is not close to 0 or π. There is another inner region near each stagnation point. There is also a boundary layer in the wake, where $\theta = 0$ and $r > 1$. The streamline from here comes from the cylinder, not from infinity. Note that the heat loss $\partial T/\partial r$ is $O(1/\epsilon^{1/2})$. This is the reason for the wind chill factor.

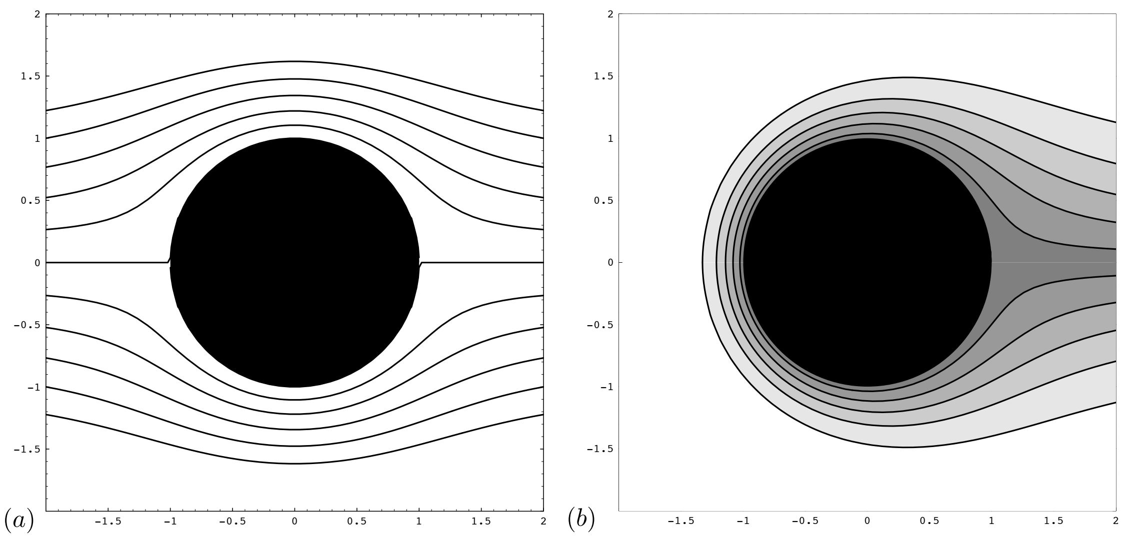


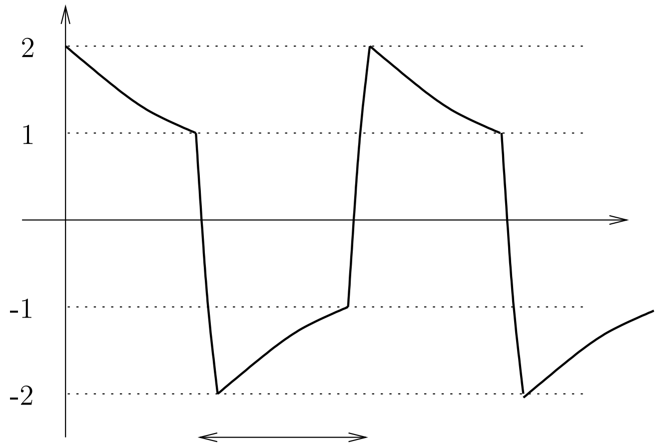
Figure 2: (a) Streamlines. (b) Isotherms.

5.4 Nonlinear oscillators

Example: Van del Pol oscillator Consider

$$\ddot{x} + \mu(x^2 - 1)\dot{x} + x = 0, \qquad \mu \gg 1.$$

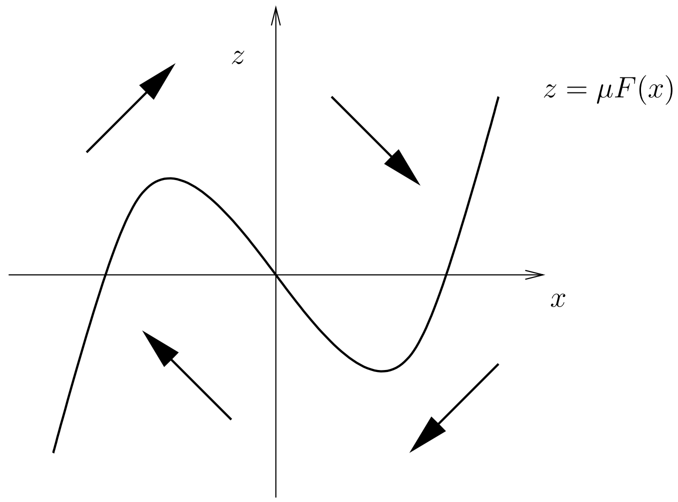
We shall show that the oscillation looks like



with long slow regions separated by rapid transitions. Such a solution is known as a "relaxation oscillation". We could proceed directly with m.a.e.s on the second order equation, as in Hinch 5.6, but to get a better understanding of what is going on we write the equation as a system of two first order equations

$$\begin{aligned}\dot{x} &= z-\mu\left(\frac{x^3}{3}-x\right)=z-\mu F(x),\\ \dot{z} &= -x.\end{aligned}$$

(An equation written in this way is said to be in Liénard form.)



Arrows indicate the general form of the motion for all μ. However, if $\mu \gg 1$ then $|\dot{x}| \gg |\dot{z}|$ except near the curve $z=\mu F(x)$. This indicates that z will be of size μ, so that it is sensible to rescale z with μ by setting $z=\mu y$, giving

$$\begin{aligned}\dot{x} &= \mu\left(y-\frac{x^3}{3}+x\right)=\mu(y-F(x)),\\ \mu\dot{y} &= -x.\end{aligned}$$

We see now that there are two timescales: x evolves on the fast timescale $t=O(\mu^{-1})$ (unless $y\approx F(x)$), while y evolves on the slow timscale $t=O(\mu)$.

Let us first consider the fast timescale by setting $t=\tau/\mu$. The equations become

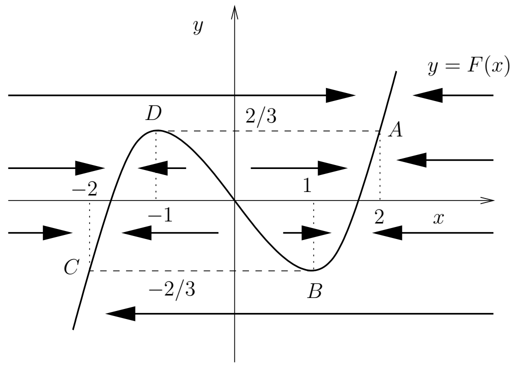
$$\begin{aligned}x_\tau &= y-\frac{x^3}{3}+x=y-F(x),\\ y_\tau &= -\frac{x}{\mu^2}.\end{aligned}$$

Expand x and y in inverse powers of μ as $x \sim x_0 + \mu^{-2}x_1 + \cdots$, $y \sim y_0 + \mu^{-2}y_1 + \cdots$ as $\mu \to \infty$. Inserting these expansions into the equations and equating coefficients of powers of μ we find at leading order

$$\begin{aligned} x_{0\tau} &= y_0 - \frac{x_0^3}{3} + x_0 = y_0 - F(x_0), \\ y_{0\tau} &= 0. \end{aligned}$$

Hence y_0 is constant on the fast timescale. Now, for a given initial y_0, x_0 tends to a root of $y_0 = F(x_0)$:

- if $y_0 > \frac{2}{3}$, the unique root between A and $+\infty$.
- if $y_0 < -\frac{2}{3}$, the unique root between C and $-\infty$.
- if $-\frac{2}{3} < y_0 < \frac{2}{3}$, $x_0 \to$ point on AB if it starts to the right of BD.
- if $-\frac{2}{3} < y_0 < \frac{2}{3}$, $x_0 \to$ point on CD if it starts to the left of BD.



Having reached the curve $y_0 = F(x_0)$ the solution comes to rest on the fast timescale, and thus begins to evolve on the slow timescale instead.

Let us scale onto the slow timescale by setting $t = \mu T$, giving

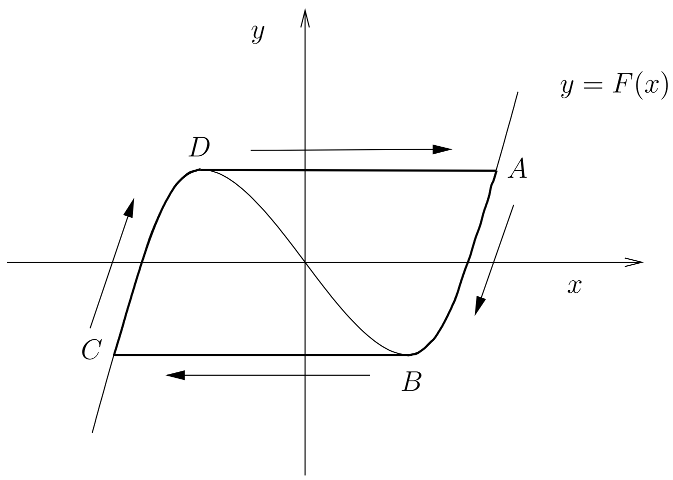
$$\begin{aligned} x_T &= \mu^2\left(y - \frac{x^3}{3} + x\right) = \mu^2(y - F(x)), \\ y_T &= -x. \end{aligned}$$

Again we expand x and y in inverse powers of μ as $x \sim x_0 + \mu^{-2}x_1 + \cdots$, $y \sim y_0 + \mu^{-2}y_1 + \cdots$ as $\mu \to \infty$. Inserting these expansions into the equations and equating coefficients of powers of μ we find at leading order

$$\begin{aligned} 0 &= y_0 - \frac{x_0^3}{3} + x_0 = y_0 - F(x_0), \\ y_{0T} &= -x_0. \end{aligned}$$

Hence the solution in the slow timescale stays on the curve $y_0 = F(x_0)$ but moves along it according to $y_{0T} = -x_0$.

Thus we have the following picture. A trajectory starting say from $(0, 1)$ quickly moves across to the branch $A\infty$. Then it remains close to the curve $y = F$ and since $\dot{y} = -x < 0$ it moves slowly down the curve. When it reaches B, if cannot keep going down and stay on the curve $y = F(x)$, so on the fast timescale (x, y) flies across horizontally to near C. Then $\dot{y} = -x > 0$ so (x, y) climbs slowly up on the curve $y = F(x)$ to D. Then (x, y) flies across horizontally to near A again and the motion becomes periodic.



During this oscillation, the main time is spent on AB and CD. The time taken to go from A to B is

$$T_{AB} = \int_A^B \frac{dy}{y_T} = -\int_A^B \frac{dy}{x} = \int_1^2 \frac{dy}{dx}\frac{dx}{x} = \int_1^2 \frac{(x^2-1)dx}{x} = \left[\frac{x^2}{2} - \log x\right]_1^2 = \left(\frac{3}{2} - \log 2\right).$$

Therefore period of oscillation $\approx \mu(3 - 2\log 2)$.

Solution by matched asymptotics

We start from $x = 2$ at $t = 0$.

Slow phase The slow time scale is $t = \mu T$ giving

$$\frac{1}{\mu^2}\frac{d^2x}{dT^2} + \frac{dx}{dT}(x^2-1) + x = 0.$$

This suggests an expansion

$$x \sim X_0 + \mu^{-2}X_1 + \cdots.$$

Substituting the expansion into the equation and equating coefficients of powers of μ:

At μ^0: $\quad \dfrac{dX_0}{dT}(X_0^2-1) + X_0 = 0 \quad$ with $X_0 = 2$ at $T = 0$,

with implicit solution

$$T = \log X_0 - \frac{X_0^2}{2} - \log 2 + 2.$$

This solution breaks down when $X_0 \to 1$ because $dX_0/dT = -X_0/(X_0^2-1) \to \infty$. The nature of the blow up is

$$T \sim -\log 2 + \frac{3}{2} - (X_0-1)^2 \quad \text{as } X_0 \to 1,$$

i.e.

$$X_0 \sim 1 + \left(\frac{3}{2} - \log 2 - T\right)^{1/2} \quad \text{as } T \to \frac{3}{2} - \log 2.$$

Proceeding to determine the next term in the expansion we find

At μ^{-2}: $\quad \dfrac{dX_1}{dT}(X_0^2-1) + 2\dfrac{dX_0}{dT}X_0X_1 + X_1 = -\dfrac{d^2X_0}{dT^2} \quad$ with $X_1 = 0$ at $T = 0$.

We could solve this for X_1, but the important thing is to determine the behaviour of X_1 as $T \to \frac{3}{2} - \log 2$, $X_0 \to 1$, which illustrates the breakdown of the asympototic series and indicate show to rescale in the transition region. With $X_0 \sim 1 + \left(\frac{3}{2} - \log 2 - T\right)^{1/2}$ we find

$$\frac{dX_1}{dT} 2\left(\frac{3}{2} - \log 2 - T\right)^{1/2} - \left(\frac{3}{2} - \log 2 - T\right)^{-1/2} X_1 + X_1 \sim \frac{1}{4}\left(\frac{3}{2} - \log 2 - T\right)^{-3/2},$$

giving

$$X_1 \sim \frac{1}{4}\left(\frac{3}{2} - \log 2 - T\right)^{-1} \qquad \text{as } T \to \frac{3}{2} - \log 2.$$

We see that X_1 blows up as $T \to \frac{3}{2} - \log 2$, so that $\mu^{-2}X_1$ ceases to be smaller than X_0 and the expansion ceases to be asymptotic. If we set $T = \frac{3}{2} - \log 2 + \delta s$ then $X_0 \sim 1 + \delta^{1/2}(-s)^{1/2}$, $\mu^{-2}X_1 \sim \frac{1}{4\mu^2}\delta^{-1}(-s)^{-1}$ as $\delta \to 0$. This means that X_1 becomes as large as $X_0 - 1$ when $\delta = \mu^{-4/3}$.

Transition phase We rescale $T = \frac{3}{2} - \log 2 + \mu^{-4/3}s$, $x = 1 + \mu^{-2/3}z$ (corresponding to $t = \mu(\frac{3}{2} - \log 2) + \mu^{-1/3}s$), giving

$$\frac{d^2z}{ds^2} + 2z\frac{dz}{ds} + 1 + \frac{1}{\mu^{2/3}}\left(z^2\frac{dz}{ds} + z\right) = 0.$$

Notice that each of the three terms in the original equation contributes to the leading order balance; this is characteristic of transition regions. The rescaled equation suggests the asymptotic expansion

$$z \sim z_0 + \mu^{-2/3}z_1 + \cdots.$$

Matching back to the slow phase we need $z \sim (-s)^{1/2} + \frac{1}{4}(-s)^{-1}$ as $s \to -\infty$.

At μ^0: $$\frac{d^2z_0}{ds^2} + 2z_0\frac{dz_0}{ds} + 1 = 0.$$

We can integrate once immediately to give

$$\frac{dz_0}{ds} + z_0^2 + s = a.$$

As $s \to -\infty$, $z_0 \sim (-s)^{1/2} + \frac{a}{2}(-s)^{-1/2} + \frac{1}{4}(-s)^{-1} + \cdots$. Hence matching gives $a = 0$. The Ricatti equation for z_0 can be linearised by setting $z_0 = \zeta'/\zeta$, giving the Airy equation

$$\zeta'' + s\zeta = 0.$$

So

$$\zeta = \alpha \mathrm{Ai}(-s) + \beta \mathrm{Bi}(-s).$$

As $s \to -\infty$,

$$\mathrm{Ai}(-s) \sim \frac{1}{2\sqrt{\pi}(-s)^{1/4}} \exp\left(-\frac{2}{3}(-s)^{3/2}\right), \qquad \mathrm{Bi}(-s) \sim \frac{1}{\sqrt{\pi}(-s)^{1/4}} \exp\left(\frac{2}{3}(-s)^{3/2}\right).$$

$$\frac{d}{ds}\mathrm{Ai}(-s) \sim (-s)^{1/2}\mathrm{Ai}(-s), \qquad \frac{d}{ds}\mathrm{Bi}(-s) \sim -(-s)^{1/2}\mathrm{Bi}(-s).$$

Hence $\beta = 0$ and

$$z_0 = \frac{\frac{d}{ds}\mathrm{Ai}(-s)}{\mathrm{Ai}(-s)} = -\frac{\mathrm{Ai}'(-s)}{\mathrm{Ai}(-s)}.$$

But $\mathrm{Ai}(-s) \to 0$, $z_0 \to -\infty$, as $s \to s_0 \approx 2.33811$. From the equation for z_0, if $|z_0| \to \infty$ at a finite value of s then

$$\frac{dz_0}{ds} + z_0^2 \sim 0,$$

so that

$$z_0 \sim -\frac{1}{s_0 - s}.$$

Including the correction term s gives

$$z_0 \sim -\frac{1}{s_0 - s} + \frac{s_0(s_0 - s)}{3} + \cdots .$$

Hence, rewritinh this in terms of x and t, as the breakdown is approached we have

$$x \sim 1 + \mu^{-2/3}\left[-\frac{1}{s_0 - \mu^{1/3}t} + \frac{s_0(s_0 - \mu^{1/3}t)}{3}\right].$$

The expansion ceases to be asymptotic when $\mu^{2/3}(s_0 - \mu^{1/3}t)$ is order one, *i.e.* $t = \mu^{-1/3}s_0 + O(\mu^{-1})$.

Fast phase The transition region suggests the scalings $t = \mu^{-1/3}s_0 + \mu^{-1}\tau$ for the fast phase. The governing equation becomes

$$\frac{d^2x}{d\tau^2} + (x^2 - 1)\frac{dx}{d\tau} + \mu^{-2}x = 0.$$

Matching backwards into the transition region gives

$$x \sim 1 + \frac{1}{\tau} - \frac{\tau s_0}{3\mu^{4/3}} \qquad \text{as } \tau \to -\infty.$$

This matching condition suggests an expansion of the form

$$x \sim x_0 + \mu^{-4/3}x_1 + \cdots .$$

At μ^0: $$\frac{d^2x_0}{d\tau^2} + (x_0^2 - 1)\frac{dx_0}{d\tau} = 0.$$

integrating once and choosing the constant of integration by matching gives

$$\frac{dx_0}{d\tau} + \frac{x_0^3}{3} - x_0 = -\frac{2}{3}.$$

Integrating again and matching backwards gives

$$\frac{1}{3}\log\left(\frac{2 + x_0}{1 - x_0}\right) + \frac{1}{1 - x_0} = -\tau.$$

The fast phase ends when $\tau \to \infty$ and $x_0 \sim -2 + 3e^{-3\tau-1}$. This is minus where we started, and the process repeats.

6 Multiple Scales

Of all asymptotic techniques, this is the one which is the most like a "black art". Problems characterised by having two processes, each with their own scales, acting simultaneously. Rapidly varying phase, slowly varying amplitude; modulated waves. Contrast with matched asymptotic expansions, where the two processes with different scales are acting in different regions.

Example: back to van der Pol oscillator

$$\ddot{x} + \epsilon\dot{x}(x^2 - 1) + x = 0.$$

Last time we looked at relaxation oscillations for large ϵ (called μ then). Here will will study with small $\epsilon > 0$ the initial value problem with initial conditions

$$x = 1, \quad \dot{x} = 0 \quad \text{at } t = 0.$$

Treating the problem as a regular perturbation expansion in ϵ gives

$$x(t, \epsilon) \sim \cos t + \epsilon \left[\frac{3}{8}(t\cos t - \sin t) - \frac{1}{32}(\sin 3t - 3\sin t)\right] + \cdots .$$

This expansion is valid for fixed t as $\epsilon \to 0$, but breaks down when $t \geq O(\epsilon^{-1})$, because of the resonant terms. When the second term in an expansion becomes as big as the first it is an indication that the expansion is breaking down.

The problem is that the damping term only changes the amplitude by an order one amount over a timescale of order ϵ^{-1}, by a slow accumulation of small effects. Thus the two processes on the two time scales are fast oscillation and slow damping.

We try to capture the behaviour on both these timescales by introducing **two** time variables:

$$\begin{aligned} \tau &= t \quad \text{— the fast time of the oscillation,} \\ T &= \epsilon t \quad \text{— the slow time of the amplitude drift.} \end{aligned}$$

We look for a solution of the form $x(t; \epsilon) = x(\tau, T; \epsilon)$ treating the variables τ and T as **independent**. We have

$$\frac{d}{dt} = \frac{d\tau}{dt}\frac{\partial}{\partial\tau} + \frac{dT}{dt}\frac{\partial}{\partial T} = \frac{\partial}{\partial\tau} + \epsilon\frac{\partial}{\partial T},$$

so that

$$\ddot{x} = x_{\tau\tau} + 2\epsilon x_{\tau T} + \epsilon^2 x_{TT}.$$

Then we expand

$$x(\tau, T; \epsilon) \sim x_0(\tau, T) + \epsilon x_1(\tau, T) + \cdots \quad \text{as } \epsilon \to 0.$$

At ϵ^0 we find

$$x_{0\tau\tau} + x_0 = 0 \quad \text{in } t \geq 0,$$

with

$$x_0 = 1, \quad x_{0\tau} = 0 \qquad \text{at } t = 0.$$

Hence

$$x_0 = R(T)\cos(\tau + \theta(T)).$$

Thus the amplitude and phase are constant as far as the fast timescale τ is concerned, but vary over the slow timescale T. Applying the initial conditions we require

$$R(0) = 1, \qquad \theta(0) = 0.$$

Apart from these conditions R and θ are arbitrary at present. Proceeding to order ϵ^1:

$$\begin{aligned} x_{1\tau\tau} + x_1 &= -x_{0\tau}(x_0^2 - 1) - 2x_{0\tau T} \qquad \text{in } t \geq 0 \\ &= 2R\theta_T\cos(\tau + \theta) + \left(2R_T + \frac{R^3}{4} - R\right)\sin(\tau + \theta) + \frac{R^3}{4}\sin 3(\tau + \theta). \end{aligned}$$

The initial conditions are

$$x_1 = 0, \qquad x_{1\tau} = -x_{oT} = -R_T \quad \text{at } t = 0.$$

Now, the $\sin 3(\tau+\theta)$ term is OK, but the $\sin(\tau+\theta)$ and $\cos(\tau+\theta)$ terms are resonant, and will give a response of the form $t\sin(\tau+\theta)$ and $t\cos(\tau+\theta)$. Thus the expansion will cease to be asymptotic again when $t = O(\epsilon^{-1})$. To keep the expansion asymptotic, we use the freedom we have in R and θ to eliminate these resonant terms (the so-called secularity or integrability or solvability condition of Poincaré), giving

$$\theta_T = 0, \qquad R_T = \frac{R(4-R^2)}{8}.$$

Using the initial conditions we therefore have

$$\theta = 0, \qquad R = \frac{2}{(1+3e^{-T})^{1/2}}.$$

Thus the amplitude of the oscillator drifts towards the value 2, which we found was a limit cycle. Thus, in particular, we have shown that the limit cycle is stable.

If we are interested in the correction x_1 we can now calculate it as

$$x_1 = -\frac{1}{32}R^3\sin 3\tau + S(T)\sin(\tau+\phi(T)),$$

with new amplitude and phase functions S and ϕ. These will be determined by a secularity condition on x_2, etc.

At higher orders we would find that a resonant forcing is impossible to avoid. In fact this is the case here in solving for x_1: we cannot avoid resonance in x_2. This can be avoided by introducing an additional slow timescale $T_2 = \epsilon^2 t$.

A simple example which illustrates the need for such a super slow time scale is the damped linear oscillator

$$\ddot{x} + 2\epsilon\dot{x} + x = 0$$

with solution

$$x = e^{-\epsilon t}\cos\left(\sqrt{1-\epsilon^2}\,t\right).$$

The amplitude drifts on the timescale ϵ^{-1}, while the phase drifts on the timescale ϵ^{-2}. In general, if we want the solution correct to $O(\epsilon^k)$ for times of $O(\epsilon^{k-n})$ then we need a hierarchy of n slow timescales.

Example: the van der Pol oscillator again

$$\ddot{x} + \epsilon\dot{x}(x^2-1) + x = 0.$$

In practice we often work directly with the variable t to save introducing the variable τ and make use of the complex representation of trigonometric functions to simplify the algebra. Thus, in seeking a multiple scales solution we begin by substituting

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \epsilon\frac{\partial}{\partial T}$$

to obtain

$$x_{tt} + 2\epsilon x_{tT} + \epsilon^2 x_{TT} + \epsilon\dot{x}(x^2-1) + x = 0.$$

Expanding

$$x \sim x_0(t,T) + \epsilon x_1(t,T) + \cdots \quad \text{as } \epsilon \to 0,$$

we obtain at leading order

$$x_{0tt} + x_0 = 0.$$

The general solution of this PDE has the complex representation

$$x_0 = \frac{1}{2}\left(A(T)e^{it} + \overline{A}(T)e^{-it}\right),$$

where A is an arbitrary complex function of T, $\overline{A}$ is the complex conjugate of A and the pre-factor of 1/2 has been introduced so that $|A(T)|$ is the slowly-varying amplitude and $\arg(A(T))$ is the slowly-varying phase, *e.g.* if $A(T) = R(T)e^{i\Theta(T)}$, where $R(T) \geq 0$, then $x_0 \equiv R(t)\cos(it + \Theta(T))$.

At $O(\epsilon^1)$, we obtain

$$\begin{aligned} x_{1tt} + x_1 &= -x_{0t}(x_0^2 - 1) - 2x_{0tT} \\ &= -\frac{1}{2}\left(iAe^{it} - i\overline{A}e^{-it}\right)\left(\frac{1}{4}(Ae^{it} + \overline{A}e^{-it})^2 - 1\right) - \left(iA_Te^{it} - i\overline{A}_Te^{-it}\right), \\ &= -i\left(\frac{dA}{dT} - \frac{A(4-|A|^2)}{8}\right)e^{it} + \text{ complex conjugate term } + \text{ non-secular terms.} \end{aligned}$$

Secular terms proportional to $e^{\pm it}$ are suppressed only if $A(T)$ satisfies the ODE

$$A_T = \frac{A(4-|A|^2)}{8}.$$

Substituting $A(T) = R(T)e^{i\Theta(T)}$, where $R(T) \geq 0$, we recover the ODEs

$$\Theta_T = 0, \quad R_T = \frac{R(4-R^2)}{8}.$$

7 The WKB method

(Liouville 1837, Green 1837, Horn 1899, Rayleigh 1912, Gans 1915, Jeffrey 1923, Wentzel 1926, Kramers 1926, Brillouin 1926, Langer 1931, Olver 1961, Meyer 1973. Notice how it's not named after the scientists who discovered it. The WKB method achieved prominence in the 20th century in its use in semiclassical analysis of quantum problems, among other areas.)

One example of a singular perturbation problem that does **not** have boundary layers is

$$\epsilon^2 y'' + y = 0.$$

It has **oscillatory** solutions and is typical of many problems arising from wave propagation, with ϵ = wavelength/size of region. So, for high-frequency propagation, ϵ is small and we need a way to deal asymptotically with such problems. The WKB method is such a method for linear wave propagation problems, and is illustrated by the equation

$$\epsilon^2 y'' + q(x)y = 0, \tag{2}$$

with $q(x) \neq 0$ in the region of interest.

Let us first see what happens if we try to solve the problem by multiple scales. Let $\epsilon X = x$ to give

$$\frac{d^2y}{dX^2} + q(\epsilon X)y = 0.$$

Thus we have an oscillator with a slowly varying frequency. We might be tempted to write $y = y(x, X)$, giving

$$\frac{\partial^2 y}{\partial X^2} + 2\epsilon\frac{d^2y}{\partial x \partial X} + \epsilon^2\frac{\partial^2 y}{\partial x^2} + q(x)y = 0. \tag{3}$$

Expanding $y \sim y_0 + \epsilon y_1 + \cdots$ gives at leading order

$$\frac{\partial^2 y_0}{\partial X^2} + q(x)y = 0.$$

Hence

$$y_0 = A(x)\cos(q(x)^{1/2}X + \theta(x)),$$

where $A(x)$ and $\theta(x)$ are arbitrary functions of x, to be determined by secularity conditions at next order. Equating coefficients of ϵ^1 in (3) gives

$$\frac{\partial^2 y_1}{\partial X^2} + 2\frac{d^2 y_0}{\partial x \partial X} + q(x) y_1 = 0,$$

i.e.

$$\begin{aligned}\frac{\partial^2 y_1}{\partial X^2} + q(x)y_1 &= 2\frac{\partial}{\partial x}\left(A(x)q(x)^{1/2}\sin(q(x)^{1/2}X + \theta(x))\right) \\ &= 2\frac{d}{dx}\left(Aq^{1/2}\right)\sin(q^{1/2}X+\theta) - 2Aq^{1/2}\left(X\frac{dq^{1/2}}{dx} + \frac{d\theta}{dx}\right)\cos(q^{1/2}X+\theta).\end{aligned}$$

The secularity condition says that there can be no multiple of $\cos(q^{1/2}X+\theta)$ or $\sin(q^{1/2}X+\theta)$ on the right-hand side. Hence

$$\frac{d}{dx}\left(Aq^{1/2}\right) = 0, \qquad X\frac{dq^{1/2}}{dx} + \frac{d\theta}{dx} = 0.$$

Here we see a problem though. The second secularity condition contains the fast scale X, and so cannot be satisfied since θ is a function of the slow scale x only. This will happen whenever the frequency of the fast oscillation depends on the slow scale.

Let us now return to (2). Instead of using multiple scales, we assume a WKB asymptotic expansion for y of the form

$$y = e^{i\phi(x)/\epsilon}A(x,\epsilon),$$

with

$$A(x,\epsilon) \sim \sum_{n=0}^{\infty} A_n(x)\epsilon^n.$$

This gives

$$y'' \sim e^{i\phi/\epsilon}\left(-\frac{(\phi')^2 A}{\epsilon^2} + \frac{2i\phi' A'}{\epsilon} + \frac{i\phi'' A}{\epsilon} + A''\right),$$

so that substituting the expansions into the equation gives at leading order ($O(\epsilon^0)$):

$$\phi'(x)^2 = q_0(x).$$

Hence

$$\phi = \pm\sqrt{q_0(x)}.$$

At order ϵ^1 we find

$$2\phi' A_0' + \phi'' A_0 = 0,$$

while at order ϵ^{n+1} for $n \geq 1$ we find

$$A_{n-1}'' + 2i\phi' A_n' + i\phi'' A_n = 0.$$

These are successive first-order **linear** equations for A_i. The first is

$$\frac{2A_0'}{A_0} + \frac{\phi''}{\phi'} = 0,$$

which we can integrate to

$$2\log A_0 + \log\phi' = \text{const.},$$

i.e.

$$A_0 = \frac{\alpha_0}{q_0(x)^{1/4}},$$

for some constant α_0. In a wave propagation problem this $A_0(x)$ gives the amplitude, and this equation corresponds to energy conservation.

The equation for A_n can be solved using an integrating factor giving

$$2i(\phi')^{1/2}\left((\phi')^{1/2}A_n\right)' = -A''_{n-1},$$

i.e.

$$A_n = \frac{i}{2(\phi')^{1/2}}\int \frac{A''_{n-1}}{(\phi')^{1/2}}\,dx;$$

the right-hand side is known.

Example The Legendre polynomial $P_n(x)$. If we let $y(\theta) = \sqrt{\sin\theta}P_n(\cos\theta)$ for $0 < \theta < \pi$ then the equation satisfied by y is

$$y'' + \left(n^2 + n + \frac{1}{4} + \frac{1}{4\sin^2\theta}\right)y = 0.$$

Let $\epsilon = 1/(n+1/2)$. Then

$$\epsilon^2 y'' + \left(1 + \frac{\epsilon^2}{4\sin^2\theta}\right)y = 0.$$

With the WKB ansatz $y = Ae^{i\phi/\epsilon}$ at leading order (ϵ^0)

$$(\phi')^2 = 1, \qquad \phi' = \pm 1, \qquad \phi = \pm\theta.$$

At order ϵ^1 we have

$$2\phi' A_0' + \phi'' A_0 = 0,$$

i.e.

$$A_0' = 0, \qquad A_0 = \alpha_0.$$

At order ϵ^2 we have

$$A_0'' + 2i\phi' A_1' + i\phi'' A_1 + \frac{1}{4\sin^2\theta}A_0 = 0,$$

i.e.

$$2iA_1' = \mp\frac{\alpha_0}{4\sin^2\theta}$$

so that

$$A_1 = \mp\frac{i\alpha_0\cot\theta}{8}.$$

Thus

$$\sqrt{\sin\theta}P_n(\cos\theta) \sim \hat{\alpha}_0\left(1 - \frac{i\cot\theta}{8(n+1/2)}\cdots\right)e^{i(n+1/2)\theta} + \hat{\beta}_0\left(1 + \frac{\cot\theta}{8(n+1/2)}\cdots\right)e^{-i(n+1/2)\theta}$$

as $n \to \infty$.

Example. An eigenvalue problem Find the large eigenvalues $\lambda \gg 1$ of the Sturm-Liouville problem

$$y'' + \lambda p(x)y = 0 \quad \text{for } 0 < x < 1, \quad \text{with } y(0) = 0,\ y(1) = 0,$$

where $p(x) > 0$ for $0 \le x \le 1$. Put $\lambda = \epsilon^{-2}$, where for $\lambda \gg 1$ we require $0 < \epsilon \ll 1$, so that

$$\epsilon^2 y'' + p(x)y = 0 \quad \text{for } 0 < x < 1, \quad \text{with } y(0) = 0,\ y(1) = 0.$$

Then with the WKB approximation $y \sim Ae^{i\phi/\epsilon}$ as $\epsilon \to 0^+$, we have

$$\phi' = \pm p^{1/2}, \quad A_0 \propto \frac{1}{(\phi')^{1/2}}.$$

If we fix $\phi(x) = +\int_0^x p(s)^{1/2}\,ds$ and $A_0(x) = p(x)^{-1/4}$, then the two independent solutions are given by

$$y_+ \sim A_0 e^{i\phi/\epsilon}, \quad y_- \sim A_0 e^{-i\phi/\epsilon}.$$

Hence, at leading order the general solution may be written in the form

$$y(x) \sim \alpha A_0(x) \cos\left(\frac{\phi(x)}{\epsilon}\right) + \beta A_0(x) \sin\left(\frac{\phi(x)}{\epsilon}\right)$$

as $\epsilon \to 0^+$, where α and β are arbitrary real constants. The boundary condition $y(0) = 0$ requires $\alpha = 0$, so that the boundary condition $y(1) = 0$ is satisfied at leading order only if

$$\beta A_0(1) \sin\left(\frac{\phi(1)}{\epsilon}\right) = \mathrm{o}(1) \quad \text{as } \epsilon \to 0^+.$$

Since $A_0(1) > 0$ and $\beta \neq 0$ for a nontrivial solution, we require

$$\sin\left(\frac{\phi(1)}{\epsilon}\right) = \mathrm{o}(1) \quad \text{as } \epsilon \to 0^+,$$

i.e.

$$\frac{\phi(1)}{\epsilon} \sim n\pi \quad \text{as } \epsilon \to 0^+,\ n \to \infty \text{ with } n \in \mathbb{N}.$$

The eigenvalues are therefore given approximately by

$$\epsilon_n \sim \frac{\phi(1)}{n\pi} = \frac{\int_0^1 \sqrt{p(x)}\,dx}{n\pi}$$

or

$$\lambda_n \sim \left(\frac{n\pi}{\int_0^1 \sqrt{p(x)}\,dx}\right)^2$$

as $n \to \infty$ with $n \in \mathbb{N}$.

Example. Turning points Find the large eigenvalues $\lambda \gg 1$ of the harmonic oscillator

$$-y'' + x^2 y = \lambda y \quad \text{for } -\infty < x < \infty, \quad \text{with } y \to 0 \text{ as } |x| \to \infty.$$

For $\lambda \gg 1$ again let $\epsilon = 1/\lambda$ and rescale $x = \epsilon^{-1/2}\bar{x}$ to give (dropping the bars)

$$\epsilon^2 y'' + (1 - x^2) y = 0 \quad \text{for } -\infty < x < \infty, \quad \text{with } y \to 0 \text{ as } |x| \to \infty.$$

Using the WKB *ansatz* we find

$$\phi' = \pm\sqrt{1 - x^2}, \qquad A_0 \propto \frac{1}{(1 - x^2)^{1/4}}.$$

Hence, the general solution has the expansion

$$y \sim \frac{\alpha_0}{(1 - x^2)^{1/4}} e^{\frac{i\phi(x)}{\epsilon}} + \frac{\beta_0}{(1 - x^2)^{1/4}} e^{-\frac{i\phi(x)}{\epsilon}} \quad \text{as } \epsilon \to 0^+, \tag{4}$$

where α_0 and β_0 are arbitrary complex constants and we have fixed $\phi(x) = \int_0^x (1 - s^2)^{1/2}\,ds$, but this approximation is only good for $|x| < 1$. When x is close to ± 1, $(1 - x^2)$ is small, and the WKB approximation breaks down. At these places $\phi' = 0$ (so they are known as turning points), and hence $A_0 = \infty$ (which indicates the breakdown). We must use a different expansion in the vicinity of each turning point (an "inner expansion") and match it with this "outer expansion".

Before we do the inner expansion, let us continue with the outer expansion for $|x| > 1$. Then we can still use WKB, and we find that, in $x > 1$ say

$$y \sim \frac{\alpha_1}{(x^2-1)^{1/4}} e^{-\frac{1}{\epsilon}\int_1^x (s^2-1)^{1/2}\,ds} + \frac{\beta_1}{(s^2-1)^{1/4}} e^{\frac{1}{\epsilon}\int_1^x (x^2-1)^{1/2}\,ds}, \tag{5}$$

where α_1 and β_1 are arbitrary real constants. Now we can apply the boundary condition at $x = +\infty$ to give $\beta_1 = 0$. The inner region near $x = 1$ will allow us to connect the coefficients α_0 and β_0 to α_1 and β_1. This will give us one condition on α_0 and β_0. The inner region near $x = -1$ will give us another.

Locally near $x = 1$ we rescale $x = 1 + \epsilon^{2/3}\hat{x}$, $y = \epsilon^{-1/6}\hat{y}(\hat{x})$ to give at leading order

$$\frac{d^2\hat{y}}{d\hat{x}^2} - 2\hat{x}\hat{y} = 0.$$

This is just the Airy equation. We want a solution which matches with (5) as $\hat{x} \to \infty$. This solution is

$$\hat{y} = C\mathrm{Ai}\left(2^{1/3}\hat{x}\right),$$

where Ai is the Airy function. It can be shown that as $\hat{x} \to \infty$

$$\epsilon^{-1/6}\hat{y}(\hat{x}) = \epsilon^{-1/6}C\mathrm{Ai}\left(2^{1/3}\hat{x}\right) \sim \frac{C}{2^{13/12}\sqrt{\pi}(\epsilon^{2/3}\hat{x})^{1/4}} e^{-2\sqrt{2}\hat{x}^{3/2}/3},$$

while the inner limit of (5) is

$$\frac{\alpha_1}{(2\epsilon^{2/3}\hat{x})^{1/4}} e^{-2\sqrt{2}\hat{x}^{3/2}/3}.$$

Hence, matching the two expansions gives

$$C = \frac{2^{13/12}\sqrt{\pi}\alpha_1}{2^{1/4}}.$$

Now as $\hat{x} \to -\infty$, it can be shown that

$$\epsilon^{-1/6}\hat{y}(\hat{x}) = \epsilon^{-1/6}C\mathrm{Ai}\left(-2^{1/3}\hat{x}\right) \sim -\frac{Ce^{-i\pi/4}}{2^{13/12}\sqrt{\pi}}\left(\frac{1}{(\epsilon^{2/3}\hat{x})^{1/4}} e^{-2\sqrt{2}i\hat{x}^{3/2}/3} - \frac{i}{(\epsilon^{2/3}\hat{x})^{1/4}} e^{2\sqrt{2}i\hat{x}^{3/2}/3}\right),$$

while the inner limit of (4) is

$$\frac{\alpha_0}{(2\epsilon^{2/3}\hat{x})^{1/4}} e^{-\frac{i}{\epsilon}\phi(1)} e^{-2\sqrt{2}i\hat{x}^{3/2}/3} + \frac{\beta_0}{(2\epsilon^{2/3}\hat{x})^{1/4}} e^{\frac{i}{\epsilon}\phi(1)} e^{2\sqrt{2}i\hat{x}^{3/2}/3}.$$

Matching the two expansions requires

$$\alpha_0 e^{-\frac{i}{\epsilon}\phi(1)} \sim -\frac{C2^{1/4}e^{-i\pi/4}}{2^{13/12}\sqrt{\pi}}, \qquad \beta_0 e^{\frac{i}{\epsilon}\phi(1)} \sim \frac{Ci2^{1/4}e^{-i\pi/4}}{2^{13/12}\sqrt{\pi}}$$

as $\epsilon \to 0^+$. For nontrivial solution (*i.e.* $C \neq 0$), we therefore require

$$\alpha_0 e^{-\frac{i}{\epsilon}\phi(1)} \sim i\beta_0 e^{\frac{i}{\epsilon}\phi(1)}$$

as $\epsilon \to 0^+$. Similarly, through a local analysis at $x = -1$ we find

$$\alpha_0 e^{-\frac{i}{\epsilon}\phi(-1)} \sim -i\beta_0 e^{\frac{i}{\epsilon}\phi(-1)}$$

as $\epsilon \to 0^+$. Hence, for a nonzero solution α_0, β_0 to exist we need

$$\frac{e^{-\frac{i}{\epsilon}\phi(1)}}{e^{-\frac{i}{\epsilon}\phi(-1)}} \sim \frac{ie^{\frac{i}{\epsilon}\phi(1)}}{-ie^{\frac{i}{\epsilon}\phi(-1)}},$$

giving

$$e^{-\frac{2i}{\epsilon}(\phi(1)-\phi(-1))+i\pi} \sim 1,$$

so that

$$\frac{2(\phi(1)-\phi(-1))}{\epsilon} \sim (2n+1)\pi \quad \text{as } \epsilon \to 0^+,\ n \to \infty \text{ with } n \in \mathbb{N}.$$

Hence, the eigenvalues are given approximately by

$$\epsilon_n \sim \frac{\phi(1)-\phi(-1)}{(n+1/2)\pi} = \frac{\int_{-1}^{1}\sqrt{1-x^2}\,dx}{(n+1/2)\pi} = \frac{1}{2n+1}$$

or

$$\lambda_n = \frac{1}{\epsilon_n} \sim 2n+1$$

as $n \to \infty$ with $n \in \mathbb{N}$. In fact these are **exact** in this case. The exact solutions are

$$y_n = e^{-x^2/2}H_n(x), \qquad \lambda_n = 2n+1.$$