Lectures 1 and 2

1 Background

1.1 Different kinds of models

Mathematical modelling of a system/phenomenon produces a mathematical model which may take the form of, for example, one of

(a) difference equation

$$\mathbf{u}_i = f(\mathbf{u}_{i-1}),$$

where $\mathbf{u}_i$ = vector of quantities describing the state at time i,

(b) ordinary differential equation (ode) - the unknowns are functions of a single variable, $u_i = u_i(x)$ say, $i = 1, \ldots, n$, and

$$\frac{d\mathbf{u}}{dt} = f(\mathbf{u}), \qquad \text{i.e. } \frac{du_i}{dx} = f_i(x, u_1(x), \cdots, u_n(x)) \quad i = 1, \ldots, n$$

(c) integral equation, e.g. the Fredholm integral equation

$$u(x) - \lambda \int_a^b K(x, y)u(y)\, dy = f(x),$$

(d) partial differential equation (pde) - the unknowns are functions of more than one variable, say $u_i = u_i(x, y)$ or $u_i = u_i(x, y, z, t)$, for example $u = u(x, t)$ and

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0,$$

(e) combinations of any of these, e.g. for $t > 0$ and the delay term $u(t-1)$,

$$\frac{du}{dt}(t) = -\frac{u(t-1)}{t}.$$

The model may be linear (e.g. (c), (e)) or nonlinear (e.g. (a), (b), (d)).

Linear: all equations and boundary conditions involve the u_i only through terms proportional to the u_i. E.g.

$$\begin{aligned} u''(t) + 3e^t u'(t) + u(t-1) &= t^2, \text{ for } t \geq 1, \\ u(t) = 1 & \qquad \text{for } 0 \leq t \leq 1. \end{aligned}$$

Nonlinear: the problem is not linear, and so for example

$$u_t + uu_x = 0, \qquad \text{or} \qquad u_{tt} = u_{xx} + \sin u \text{ for } t > 0, \text{ all } x.$$

Often the model will contain many constants (ρ, g, etc.) and equations, and we try to simplify it in various ways. This is because we want to discover what is essential or important in the model. Usually this leads to both better understanding of the model and more efficient computation of solutions. Our first simplification follows from the principle that the model should be independent of the way that we measure things.

1.2 Dimensions, units, scaling

We define the dimensions of mass, length, and time to be the symbols M, L, and T and then any quantity that is defined as a product or ratio of other quantities has the product or ratio of their dimensions. So for example in mechanics we have, where $[X]$ means dimension of X,

$$
\begin{array}{rrcl}
 & [\text{length}] & = & L, \\
 & [\text{area}] & = & L^2, \\
 & [\text{volume}] & = & L^3, \\
 & [\text{density}] & = & M/L^3, \\
 & [\text{velocity}] & = & L/T, \\
 & [\text{acceleration}] & = & L/T^2, \\
 & [\text{momentum}] & = & ML/T, \\
 & [\text{force}] & = & ML/T^2, \\
\text{force/area} & [\text{pressure}] & = & M/LT^2.
\end{array}
$$

(In the SI system, the natural unit of pressure is kg m^{-1} s^{-2}: in fact this has its own name Pascal (Pa). In financial/biological modelling we need to include units of money, cash flow, cell density, cell movements, etc. In electromagnetic/thermal problems we also include charge Q and temperature K in degree centigrade or in Kelvin as fundamental dimensions along side M, L, T.)

Change of units If say a mass has numerical value x_1 when the unit of mass is m_1, and x_2 when the unit of mass is m_2, then

$$m_1 x_1 = m_2 x_2, \text{ so that } \frac{x_1}{x_2} = \frac{m_2}{m_1}.$$

If [quantity] $= M^\alpha L^\beta T^\gamma$ then its numerical values q transform by

$$\frac{q_1}{q_2} = \left(\frac{m_2}{m_1}\right)^\alpha \left(\frac{l_2}{l_1}\right)^\beta \left(\frac{t_2}{t_1}\right)^\gamma .$$

Consistency A valid equation between physical quantities must have all terms of the **same** dimension (otherwise the equation would not hold in a different system of units). For example Newton's laws and Bernouilli's equation may be written

$$F = ma \quad \text{and} \quad p + \frac{\rho u^2}{2} = p_0,$$

so that

$$[F] = \frac{ML}{T^2} = [m][a] \quad \text{and} \quad \left[\frac{\rho u^2}{2}\right] = \frac{M}{L^3}\left(\frac{L}{T}\right)^2 = \frac{M}{LT^2} = [p]\,.$$

Scaling/nondimensionalization Once we have a mathematical model of the physical situation, then if say x measures distance along a bar of length l, we can express the equation in terms of $\hat{x} = x/l$, which is dimensionless, i.e. $[\hat{x}] = 1$. Similarly for any physical quantity q we can let $\hat{q} = q/q_0$ where q_0 is some suitably chosen constant having the same dimensions as q. We do this in such a way as to **reduce** the parameters. E.g. vibrations in a beam with fixed ends may be modelled by

$$\begin{aligned} \rho \frac{\partial^2 y}{\partial t^2} &= T_0 \frac{\partial^2 y}{\partial x^2} - B \frac{\partial^4 y}{\partial x^4}, & (1) \\ y(0,t) = \frac{\partial y}{\partial x}(0,t) &= y(l,t) = \frac{\partial y}{\partial x}(l,t) = 0. & (2) \end{aligned}$$

The original system has four physical constants or parameters: ρ, T_0, B, l. These need to be measured by experiments for a practical situation. How many cases should we solve to understand the solution ?

$$\begin{aligned} \rho: &\text{ mass/unit length} & [\rho] &= ML^{-1}, \\ T_0: &\text{ tension} \times \text{area} & [T_0] &= MLT^{-2}, \\ B: &\text{ bending stiffness} & [B] &= ML^3T^{-2}, \\ l: &\text{ length} & [l] &= L. \end{aligned}$$

Let

$$\hat{x} = \frac{x}{l}, \qquad \hat{t} = \frac{t}{t_0}, \qquad \hat{y} = \frac{y}{y_0},$$

where t_0 is a representative timescale (which we will shortly determine) and y_0 is a representative scale for the unknown solution $y(x,t)$, the deflection of the beam. Then

$$\begin{aligned} \frac{\partial^2 \hat{y}}{\partial \hat{t}^2} &= \frac{t_0^2}{y_0} \frac{\partial^2 y}{\partial t^2} \\ &= \frac{t_0^2}{\rho y_0} \left(T_0 \frac{\partial^2 y}{\partial x^2} - B \frac{\partial^4 y}{\partial x^4} \right) \\ &= \frac{t_0^2}{\rho} \left(\frac{T_0}{l^2} \frac{\partial^2 \hat{y}}{\partial \hat{x}^2} - \frac{B}{l^4} \frac{\partial^4 \hat{y}}{\partial \hat{x}^4} \right) \\ &= \frac{t_0^2 T_0}{\rho l^2} \frac{\partial^2 \hat{y}}{\partial \hat{x}^2} - \frac{t_0^2 B}{\rho l^4} \frac{\partial^4 \hat{y}}{\partial \hat{x}^4}. \end{aligned}$$

Now we choose t_0 to eliminate one of these parameters (i.e. set it to equal 1). We set

$$t_0 = l\sqrt{\frac{\rho}{T_0}}, \qquad \text{so that} \qquad \hat{t} = \frac{t}{l}\sqrt{\frac{T_0}{\rho}}$$

(check that $[t_0] = T$ and therefore $\hat{t}$ is dimensionless). Then

$$\begin{aligned} \frac{\partial^2 \hat{y}}{\partial \hat{t}^2} &= \frac{\partial^2 \hat{y}}{\partial \hat{x}^2} - \beta \frac{\partial^4 \hat{y}}{\partial \hat{x}^4}, & (3) \\ \hat{y}(0,\hat{t}) = \frac{\partial \hat{y}}{\partial \hat{x}}(0,\hat{t}) &= \hat{y}(1,\hat{t}) = \frac{\partial \hat{y}}{\partial \hat{x}}(1,\hat{t}) = 0, & (4) \end{aligned}$$

where

$$\beta = \frac{B}{T_0 l^2}.$$

There is really only one parameter in the problem. Since $\hat{y}$, $\hat{t}$, $\hat{x}$ are dimensionless, β is dimensionless (check): a value of $\beta = 10^{-2}$ or 100 is intrinsic to the problem, not an accident due to the system of units used. Advantages of this include:

1. Simplifies equation: may reveal similarities with other equations modelling other physical situations (e.g. finance and heat conduction).
2. If β is large or small it can be useful to consider the limiting behaviour $\beta \to \infty$ or $\beta \to 0$. E.g. if $\beta = 10^{-3}$ we can usually neglect the $\beta \partial^4 \hat{y}/\partial \hat{x}^4$ term, while if $\beta = 10^3$ we cannot. Note we **cannot** tell this just from the value of B, because that value depends on the system of units used, or the experimental set-up.
3. Numerical solution generally goes better if quantities have been rescaled to be of order one (usually known as preconditioning).
4. Numerical solutions only have to be computed and plotted for a range of β values, not for combinations of independent choices of ρ, T_0, B and l.
5. Hence we get better understanding, and usually can suggest better experiments to verify the model and discover the underlying behaviour.

So we shall consider equations more like (3) than (1). They look artificial, but arise as rescaled versions of real problems. We call this nondimensionalising.

With all kinds of mathematical models we need to ask

- Does a solution exist ?
- Is it unique ?
- Does it depend continuously on the data specified in the problem ?

If any answer is "no", then we reconsider that model. In this first course we shall mainly consider ode's. These may arise

(a) directly as a model of a physical system, or

(b) indirectly as a subproblem when analysing some larger model.

When we have a well-posed and nondimensionalised model, we then want to consider its robustness and stability.

2 Linear ode's

We first revise some basic facts about linear problems, since often these provide the basis for approximation of more complicated problems.

2.1 The Fredholm Alternative

Preliminary observation about linear equations

$$A\mathbf{u} = \mathbf{b}. \tag{5}$$

A is an $m \times n$ real matrix, $\mathbf{u}$ is a vector in $\mathbb{R}^n$, $\mathbf{b}$ in $\mathbb{R}^m$. There is a solvability condition as follows: (5) has a solution $\mathbf{u} \Leftrightarrow \mathbf{v} \cdot \mathbf{b} := \mathbf{v}^T\mathbf{b} = 0$ for every $\mathbf{v}$ in $\mathbb{R}^m$ such that $A^T\mathbf{v} = 0$.

Proof Using $\mathcal{L}(\mathbf{a}_1, \cdots, \mathbf{a}_n) =$ the linear span of $\mathbf{a}_1, \cdots, \mathbf{a}_n$, we have for $A = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$,

$$\begin{aligned}
A\mathbf{u} = \mathbf{b} \text{ has a solution } \mathbf{u} \quad &\Leftrightarrow \quad \mathbf{b} \text{ is a linear combination of the columns of } A \\
&\Leftrightarrow \quad \mathbf{b} \in \mathcal{L}(\mathbf{a}_1, \cdots, \mathbf{a}_n) \\
&\Leftrightarrow \quad \mathcal{L}\mathbf{b} \subseteq \mathcal{L}(\mathbf{a}_1, \cdots, \mathbf{a}_n) \\
&\Leftrightarrow \quad \mathcal{L}\mathbf{b}^\perp \supseteq \mathcal{L}(\mathbf{a}_1, \cdots, \mathbf{a}_n)^\perp \\
&\Leftrightarrow \quad \text{every vector } \mathbf{v} \text{ with each } \mathbf{a}_i^T\mathbf{v} = 0 \text{ also has } \mathbf{v}^T\mathbf{b} = 0.
\end{aligned}$$

□

We are going to describe solvability conditions for linear ode's analogous to those for linear algebraic equations. We will do this first for the second order real scalar case, and give the general version later.

Consider a differential operator

$$L[u] := \frac{d^2u(x)}{dx^2} + \alpha(x)\frac{du(x)}{dx} + \beta(x)u(x) = u'' + \alpha u' + \beta u, \text{ for } 0 < x < 1,$$

where $\alpha(x)$, $\beta(x)$ are continuous real-valued functions on $[0, 1]$. We are going to consider:

Primary problem

$$L[u] = b(x) \qquad \text{for } 0 < x < 1,$$

with two linear homogeneous boundary conditions on u and/or u' at $x = 0$ and/or 1.

Adjoint problem

$$L^*[v] = 0 \qquad \text{for } 0 < x < 1,$$

with two linear homogeneous boundary conditions on v and v' at $x = 0$, 1. The adjoint differential operator is

$$L^*[v] = v'' - (\alpha v)' + \beta v.$$

The solvability result (usually known as the Fredholm alternative) is that

$$\text{Primary has a solution } u \Leftrightarrow \int_0^1 v(x)b(x)\,dx = 0 \quad \text{for every solution } v \text{ of the Adjoint.}$$

We first note the fundamental identity

$$\begin{aligned}
\int_0^1 \left(vL[u] - uL^*[v]\right)\,dx &= \int_0^1 v(u'' + \alpha u' + \beta u) - u(v'' - (\alpha v)' + \beta v)\,dx \\
&= \left[vu' - uv' + \alpha uv\right]_0^1 \\
&= B(u, v),
\end{aligned}$$

where $B(u, v)$ is a bilinear form in the *boundary values* of u and v.

This bilinear form must be non-singular in the same sense that $B(u, v) = 0$ for all v implies $u = 0$. Equivalently, we can write

$$B(u, v) = \begin{pmatrix} v(1) & v'(1) & v(0) & v'(0) \end{pmatrix} \begin{pmatrix} \alpha(1) & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & -\alpha(0) & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} u(1) \\ u'(1) \\ u(0) \\ u'(0) \end{pmatrix}$$

and $B(u, v)$ is non-singular if the central matrix is non-singular.

If $u(1)$, $u'(1)$, $u(0)$, $u'(0)$ obey two linear homogeneous equations (the primary boundary conditions) then we shall need two linear homogeneous equations on $v(1)$, $v'(1)$, $v(0)$, $v'(0)$ to force $B(u, v) = 0$ (there are 2 degrees of freedom left). These conditions on v are the adjoint boundary conditions.

2.2 Examples

Example 1 Suppose the primary boundary conditions are

$$u(0) = 0, \qquad u'(0) = 0 \qquad\qquad \text{Primary boundary conditions}$$

(an initial value problem, IVP). Then

$$B(u, v) = v(1)u'(1) - v'(1)u(1) + \alpha(1)v(1)u(1).$$

To force this to vanish (for arbitrary $u(1)$, $u'(1)$) we must have

$$v(1) = 0, \qquad v'(1) = 0 \qquad\qquad \text{Adjoint boundary conditions.}$$

Example 2 Suppose the primary boundary conditions are

$$u(0) = 0, \qquad u(1) = 0 \qquad\qquad \text{Primary boundary conditions}$$

(a boundary value problem, BVP). Then

$$B(u, v) = v(1)u'(1) - v(0)u'(0).$$

To force this to vanish (for arbitrary $u'(0)$, $u'(1)$) we must have

$$v(0) = 0, \qquad v(1) = 0 \qquad\qquad \text{Adjoint boundary conditions.}$$

Example 3 Suppose the primary boundary conditions are

$$u(0) = u(1), \qquad u'(1) = 0 \qquad\qquad \text{Primary boundary conditions}$$

(a generalised boundary value problem). Then

$$B(u, v) = -v'(1)u(0) + \alpha(1)v(1)u(0) - v(0)u'(0) + v'(0)u(0) - \alpha(0)v(0)u(0).$$

To force this to vanish (for arbitrary $u(0)$, $u'(0)$) we must have

$$v(0) = 0, \qquad v'(1) - \alpha(1)v(1) = v'(0) \qquad\qquad \text{Adjoint boundary conditions.}$$

Easy part of proof If a solution u of the primary problem exists, and v is any solution of the adjoint problem, then

$$\int_0^1 (vL[u] - uL^*[v]) \, dx = B(u, v),$$

i.e. for u (v) satisfying our primary (adjoint) problem we have

$$\int_0^1 vb \, dx = 0.$$

The harder part (if the adjoint condition holds then a solution exists) requires two steps

1. Convert the ode problem to an integral equation by using a Green's function.
2. Use the "Fredholm Alternative" theory of integral equations to write down solvability conditions for the integral equation.

This is why the solvability condition for ode's is usually called the Fredholm alternative.

Example of solvability condition
Primary problem:

$$u''(x) = b(x) \text{ for } 0 < x < 1, \qquad u'(0) = u'(1) = 0.$$

Adjoint problem:

$$v''(x) = 0 \text{ for } 0 < x < 1, \qquad v'(0) = v'(1) = 0.$$

There is a nontrivial solution of the adjoint, namely

$$v = \text{ constant, say } v = 1.$$

Hence there is a solution of the primary if and only if

$$\int_0^1 b(x)\, dx = 0.$$

Generalisation Suppose u is a vector of complex-valued functions, obeying a higher-order primary problem.

Primary

$$L[u] = b(x) \qquad \text{for } 0 < x < 1,$$

with primary boundary conditions on u at $x = 0$, 1, where

$$L[u] = \sum_{r=0}^{k} A_r(x) \frac{d^r u}{dx^r} = \sum_{r=0}^{k} A_r(x) u^{(r)},$$

where the $A_r(x)$ are matrices, continuous in x, and b is a vector of continuous functions. To state the adjoint problem we introduce some notation.

1. $A^* =$ conjugate of transpose of A [like A' in matlab]
2. If v is a vector of continuous functions (same order as b) then define an inner product

$$\langle v, b \rangle = \int_0^1 v(x)^* b(x)\, dx = \sum_i \int_0^1 \overline{v_i(x)} b_i(x)\, dx.$$

Then

$$\text{Primary has a solution } u \Leftrightarrow \langle v, b \rangle = 0 \quad \text{for every solution } v \text{ of the Adjoint problem.}$$

Adjoint

$$L^*[v] = 0,$$

with adjoint boundary conditions on v at $x = 0$, 1. The adjoint differential operator is

$$L^*[v] = \sum_{r=0}^{k} (-1)^r \left(A_r^* v\right)^{(r)}.$$

The fundamental identity is

$$\begin{aligned}
\langle v, L[u]\rangle - \langle L^*[v], u\rangle &= \int_0^1 \sum_r \left(v^* A_r u^{(r)} - (-1)^r (v^* A_r)^{(r)} u\right)\, dx \\
&= \left[\sum_r v^* A_r u^{(r-1)} - (v^* A_r)' u^{(r-2)} + \cdots + (-1)^{r-1} (v^* A_r)^{(r-1)} u\right]_0^1 \\
&= B(u, v).
\end{aligned}$$

This B is used to construct the adjoint boundary conditions exactly as in the basic case considered earlier (B is a Hermitian form now). The easy part of the proof is just as before.

3 Green's functions and Distributions

First we refer back to our basic ode

$$L[u] := u'' + \alpha(x)u' + \beta(x)u.$$

If the problem

$$\begin{cases} L[u] = 0, & \text{for } 0 < x < 1, \\ \text{Primary boundary conditions on } u, \end{cases}$$

has **only** the solution $u = 0$, **then** there is a "Green's function" $G(x, y)$ such that the problem

$$\begin{cases} L[u] = f(x), & \text{for } 0 < x < 1, \\ \text{Primary boundary conditions on } u, \end{cases}$$

has the **unique solution**

$$u(x) = \int_0^1 G(x, y) f(y)\, dy.$$

Lectures 3 and 4

Construction of a Green's function to solve

$$\begin{aligned} u''(x) + \alpha(x)u'(x) + \beta(x)u(x) &= f(x) \qquad \text{for } 0 < x < 1, \\ p_{11}u(1) + p_{12}u'(1) + p_{13}u(0) + p_{14}u'(0) &= 0, \\ p_{21}u(1) + p_{22}u'(1) + p_{23}u(0) + p_{24}u'(0) &= 0. \end{aligned}$$

Here $\alpha(x), \beta(x)$ are continuous functions on $[0,1]$ and p_{ij} are constants. Since we have a second order linear ode there is a two dimensional vector space of solutions of the homogeneous equation

$$u'' + \alpha u' + \beta u = 0 \qquad \text{for } 0 < x < 1.$$

Our convention throughout here is that in general $u(x)$ is twice continuously differentiable in $(0,1)$ and has continuous derivatives on $[0,1]$.

If we impose a single homogeneous boundary condition then there will still be a one-dimensional space of such solutions, so let u_1 be any function obeying the homogeneous ode and the first boundary condition, and let u_2 be any function obeying the homogeneous ode and the second boundary condition. By hypothesis there is no non-zero solution of the homogeneous ode and **both** boundary conditions, so u_1 does not satisfy the second boundary condition, and u_2 does not satisfy the first boundary condition. Hence u_1 and u_2 are linearly independent. Let $w = u_1u_2' - u_2u_1' \neq 0$ on $[0,1]$.

(If $w(x_0) = 0$ then

$$\begin{pmatrix} u_1(x_0) & u_2(x_0) \\ u_1'(x_0) & u_2'(x_0) \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

has a non-zero solution (A_1, A_2). Then $u = A_1u_1 + A_2u_2$ has $u(x_0) = 0$ and $u'(x_0) = 0$ and obeys $u'' + \alpha u' + \beta u = 0$. Therefore by the uniqueness theorem for solutions of our ode we have $u \equiv 0$. But u_1 and u_2 are linearly independent - contradiction. Therefore $w(x_0) \neq 0$.)

Now let

$$u_0(x) = u_2(x)\int_0^x \frac{u_1(y)f(y)\,dy}{w(y)} - u_1(x)\int_0^x \frac{u_2(y)f(y)\,dy}{w(y)}.$$

This is a particular integral. Calculation shows that

$$\begin{aligned} u_0' &= u_2'(x)\int_0^x \frac{u_1(y)f(y)\,dy}{w(y)} + \frac{u_2(x)u_1(x)f(x)}{w(x)} - u_1'(x)\int_0^x \frac{u_2(y)f(y)\,dy}{w(y)} - \frac{u_1(x)u_2(x)f(x)}{w(x)} \\ &= u_2'(x)\int_0^x \frac{u_1(y)f(y)\,dy}{w(y)} - u_1'(x)\int_0^x \frac{u_2(y)f(y)\,dy}{w(y)}, \\ u_0'' &= u_2''(x)\int_0^x \frac{u_1(y)f(y)\,dy}{w(y)} + \frac{u_2'(x)u_1(x)f(x)}{w(x)} - u_1''(x)\int_0^x \frac{u_2(y)f(y)\,dy}{w(y)} - \frac{u_1'(x)u_2(x)f(x)}{w(x)} \\ &= u_2''(x)\int_0^x \frac{u_1(y)f(y)\,dy}{w(y)} - u_1''(x)\int_0^x \frac{u_2(y)f(y)\,dy}{w(y)} + f(x). \end{aligned}$$

Hence

$$\begin{aligned} u_0'' + \alpha u_0' + \beta u_0 &= (u_2'' + \alpha u_2' + \beta u_2)\int_0^x \frac{u_1(y)f(y)\,dy}{w(y)} - (u_1'' + \alpha u_1' + \beta u_1)\int_0^x \frac{u_2(y)f(y)\,dy}{w(y)} + f(x) \\ &= f(x). \end{aligned}$$

So the general solution of our ode is $u = u_0 + A_1 u_1 + A_2 u_2$ with A_1, A_2 arbitrary. To make u obey the first boundary condition choose A_2 correctly. To make u obey the second boundary condition choose A_1 correctly. This provides a unique u. We will show it has the required form. First define

$$u_0(x) = \int_0^1 G_0(x,y) f(y)\, dy,$$

with

$$G_0(x,y) = \begin{cases} \dfrac{u_2(x)u_1(y) - u_1(x)u_2(y)}{w(y)} & \text{if } y \le x, \\ 0 & \text{if } y \ge x. \end{cases}$$

Here we choose $G_0 = 0$ for $x \le y$ so that

$$u_0(x) \quad = \quad \int_0^x G_0(x,y) f(y) dy$$

only depends on the values of $f(y)$ for $y \le x$, i.e. a causality dependence if x is time. (Later we make a more symmetric argument.)
Then, since u_1 satisfies the first boundary condition, u will also satisfy this boundary condition if we choose A_2 so that

$$p_{11}u_0(1) + p_{12}u_0'(1) + p_{13}u_0(0) + p_{14}u_0'(0) + A_2(p_{11}u_2(1) + p_{12}u_2'(1) + p_{13}u_2(0) + p_{14}u_2'(0)) = 0,$$

i.e.

$$A_2 = -\frac{p_{11}u_0(1) + p_{12}u_0'(1) + p_{13}u_0(0) + p_{14}u_0'(0)}{p_{11}u_2(1) + p_{12}u_2'(1) + p_{13}u_2(0) + p_{14}u_2'(0)}.$$

Note that the denominator is not zero since otherwise u_2 would satisfy both boundary conditions. But

$$\begin{aligned} u_0(1) &= \int_0^1 G_0(1,y) f(y)\, dy, \\ u_0'(1) &= \int_0^1 \frac{\partial G_0}{\partial x}(1,y) f(y)\, dy, \\ u_0(0) &= \int_0^1 G_0(0,y) f(y)\, dy, \\ u_0'(0) &= \int_0^1 \frac{\partial G_0}{\partial x}(0,y) f(y)\, dy. \end{aligned}$$

Therefore

$$A_2 = \int_0^1 \alpha_2(y) f(y)\, dy,$$

with

$$\alpha_2 = -\frac{p_{11}G_0(1,y) + p_{12}\frac{\partial G_0}{\partial x}(1,y) + p_{13}G_0(0,y) + p_{14}\frac{\partial G_0}{\partial x}(0,y)}{p_{11}u_2(1) + p_{12}u_2'(1) + p_{13}u_2(0) + p_{14}u_2'(0)}.$$

Similarly we find

$$A_1 = \int_0^1 \alpha_1(y) f(y)\, dy,$$

with

$$\alpha_1 = -\frac{p_{21}G_0(1,y) + p_{22}\frac{\partial G_0}{\partial x}(1,y) + p_{23}G_0(0,y) + p_{24}\frac{\partial G_0}{\partial x}(0,y)}{p_{21}u_1(1) + p_{22}u_1'(1) + p_{23}u_1(0) + p_{24}u_1'(0)}.$$

Thus

$$u = \int_0^1 G(x,y) f(y)\, dy$$

as required, where

$$G(x,y) = G_0(x,y) + \alpha_1(y)u_1(x) + \alpha_2(y)u_2(x).$$

In fact, for a given y we can check that

- $G(x,y)$ obeys the homogeneous boundary conditions at $x = 0$ and $x = 1$.
- $G(x,y)$ obeys the homogeneous differential equation except at $x = y$.
- $G(x,y)$ is continuous at $x = y$.
- $\left[\dfrac{\partial G(x,y)}{\partial x}\right]_{x=y-}^{x=y+} = 1.$

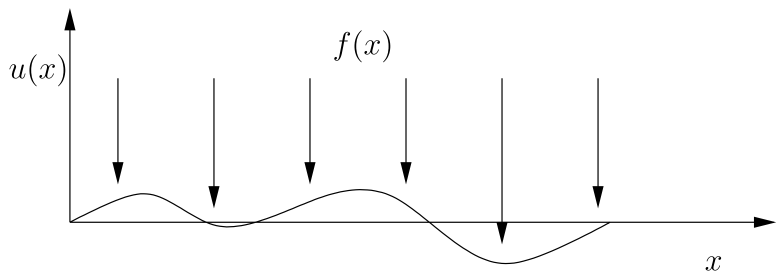
These four conditions can be used to construct the Greens function.

The integral operator defined by the Green's function is the inverse of the differential operator together with the boundary conditions.

Example

$$\begin{aligned} u'' &= f \qquad \text{on } [0,1], \\ u(0) &= 0, \\ u(1) &= 0. \end{aligned}$$

(Represents equation of a stretched string with a force $f(x)$ per unit length.)



Following the procedure above $u_1(x) = x$, $u_2(x) = 1 - x$, $w = -1$. Then

$$G_0(x,y) = \begin{cases} -(1-x)y + x(1-y) & \text{if } y \le x, \\ 0 & \text{if } y \ge x. \end{cases}$$

We need

$$u_0(0) + A_2 = 0, \qquad u_0(1) + A_1 = 0,$$

so that

$$\begin{aligned} A_2 &= -\int_0^1 G_0(0,y)f(y)\,dy, \\ A_1 &= -\int_0^1 G_0(1,y)f(y)\,dy. \end{aligned}$$

Hence

$$\begin{aligned} G(x,y) &= G_0(x,y) - G_0(0,y)u_2(x) - G_0(1,y)u_1(x) \\ &= \begin{cases} -y(1-x) & \text{if } y \le x, \\ -x(1-y) & \text{if } y \ge x. \end{cases} \\ &= -\min(x,y)(1-\max(x,y)). \end{aligned}$$

Or (easier) from the four conditions on G we know that

$$\begin{aligned} G(x,y) = a(y)u_1(x) = a(y)x \qquad &\text{if } x \le y, \\ G(x,y) = b(y)u_2(x) = b(y)(1-x) \qquad &\text{if } x \ge y, \end{aligned}$$

to satisfy equation and boundary conditions. The final two conditions determine a and b:

$$\begin{aligned} G(y,y) &= a(y)y = b(y)(1-y), \\ \left[\frac{\partial G(x,y)}{\partial x}\right]_{x=y-}^{x=y+} &= -b(y) - a(y) = 1. \end{aligned}$$

Hence $a(y) = -(1-y)$, $b(y) = -y$. Now, numerically discretize at $x_i = i/n$, $i = 0, \dots, n$. Write $u_i = u(x_i)$, $f_i = f(x_i)$. Then we write

$$u''(x_i) \to \frac{u_{i+1} - 2u_i + u_{i-1}}{(1/n)^2}.$$

The boundary conditions give $u_0 = u_n = 0$. The ode BVP becomes the tridiagonal set of linear equations

$$\begin{pmatrix} -2n^2 & n^2 & 0 & \cdots & 0 \\ n^2 & \cdot & \cdot & 0 & \vdots \\ 0 & \cdot & \cdot & \cdot & 0 \\ \vdots & 0 & \cdot & \cdot & n^2 \\ 0 & \cdots & 0 & n^2 & -2n^2 \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_{n-1} \end{pmatrix} = \begin{pmatrix} f_1 \\ \vdots \\ f_{n-1} \end{pmatrix}$$

i.e.

$$L_n \mathbf{u} = \mathbf{f}.$$

The integral representation

$$u(x) = \int_0^1 G(x,y)f(y)\,dy,$$

with the trapezium rule, becomes

$$u_i = \sum_j \frac{G(x_i,x_j)f_j}{n} = \sum_j G(x_i,x_j)f_j\Delta x,$$

i.e.

$$\mathbf{u} = G_n\mathbf{f},$$

where

$$(G_n)_{ij} = -\frac{\min(i,j)}{n}\left(1-\frac{\max(i,j)}{n}\right)\frac{1}{n}, \qquad (i,j = 1,\ldots,n-1).$$

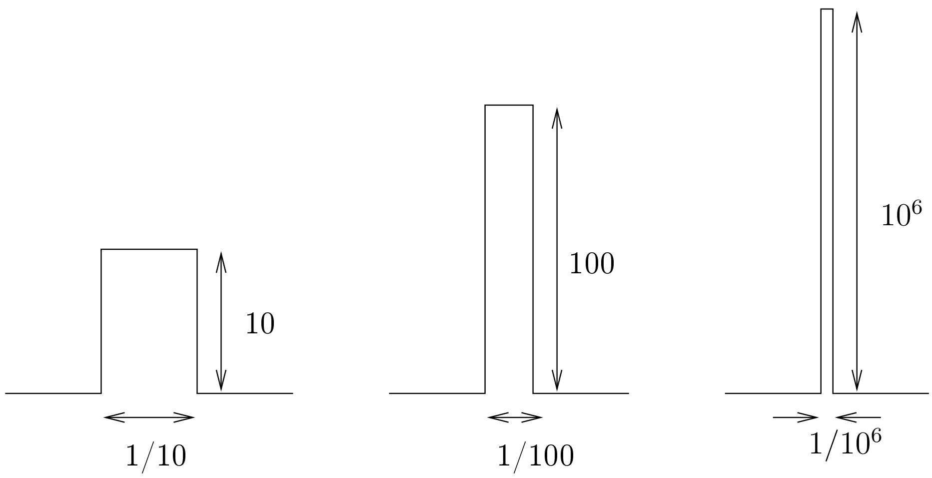
L_n and G_n are **exact** inverse matrices. In the discretized case $(G_n)_{ij}$ = response in u at x_i to unit input in f at x_j. Thus $G(x_i, x_j)$ is the response in u at x_i to a unit input in $f_j\Delta x$ at x_j.

This looks OK, in that we regard the force per unit length $f(x)$ as replaced by a sum of point inputs $f_j\Delta x$ at x_j, then that point force input $\to G(x,x_j)f_j\Delta x$ at x. Then sum over $j \leftrightarrow$ integrate over y.

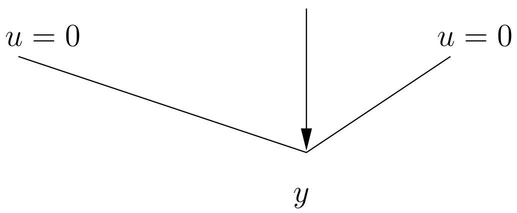
But, as $\Delta x \to 0$, the point force is concentrated over zero length, as if

$$\frac{\partial^2 G}{\partial x^2} = \delta(x - x_j)$$

where the RHS is a function that is zero except at $x = x_j$ and has a unit integral—which is impossible. $\delta(x-y)$ is "the force per unit length when a unit point force is applied at y". It is the limit of



In the example above $u'' = f$, $G(x,y)$ is really the response to a unit point force at $x = y$:



Note that $\int \delta(x-y)\,dx = 1$, so

$$\frac{\partial^2 G}{\partial x^2} = \delta(x-y) \qquad \to \qquad \left[\frac{\partial G}{\partial x}\right]_{x=y-}^{x=y+} = 1.$$

Formal intro Dirac observed that we could get a lot of correct results by dealing with ode's like that above with "functions" like $\delta(x-y)$ on the RHS, integrating them by parts etc as if they were genuine functions. How can this be made into a respectable mathematical theory ? Clearly we want to generalise our concept of "function" to include such things as $\delta(x-y)$ and then define allowed ways of working with generalised functions, differentiating them, integrating them, etc.

There are various ways of doing this (just as there are various ways of constructing the reals from the rationals).

(a) Generalised functions as limits of sequences of genuine functions. See J. Lighthill, F.A. and Gen. Funcs.

(b) Consider a large space of "distributions" and show that genuine functions are distributions.

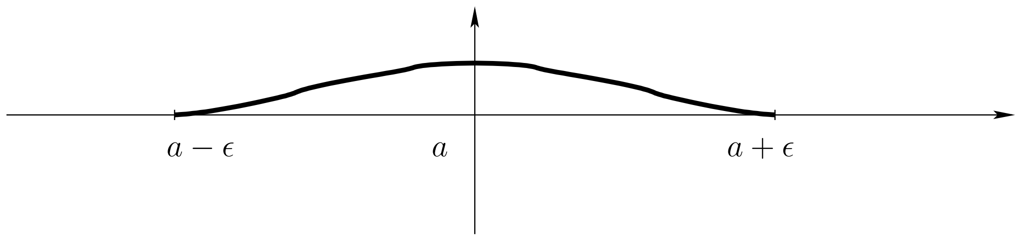
Let's use (b).

4 Distributions and Generalised Functions: Formal definitions

A **test function** in $\mathbb{R}$ is a smooth function $\phi : \mathbb{R} \to \mathbb{C}$ that is differentiable any number of times, and vanishes outside some bounded interval. Thus $\phi \in C^\infty(\mathbb{R})$ and has **compact support** (support$(\phi) = \overline{\{x : \phi(x) \neq 0\}}$ is compact if and only if for some X supp$(\phi) \subseteq [-X, X]$, i.e. if and only if $\phi(x) = 0$ for $|x| \geq X$). Usually candidates for ϕ (mollifiers) include

$$\begin{cases} \exp\left(\frac{-c}{\epsilon^2-(x-a)^2}\right) & \text{for } a-\epsilon < x < a+\epsilon, \\ 0 & \text{otherwise,} \end{cases}$$

for any constant $c > 0$.



A **distribution** is a linear map $u : C_0^\infty(\mathbb{R}) \to \mathbb{C}$ with the property that for all X there exists some C and N such that

$$|\langle u, \phi \rangle| \leq C \sum_{n \leq N} \max \left| \frac{\partial^n \phi}{\partial x^n} \right|.$$

for all ϕ with supp$(\phi) \subseteq [-X, X]$. (This condition can be checked easily in examples and is needed for convergence arguments, e.g. if there is a sequence of distributions u_j with $\langle u_j, \phi \rangle \to \langle u, \phi \rangle$ for all $\phi \in C_0^\infty$ then u is a distribution, and we say u_j converges (weakly) to u in the distributional sense.) Let $\mathcal{D}'$ be the space of such distributions u.

Continuous functions as distributions Suppose $f(x)$ is a continuous function : $\mathbb{R} \to \mathbb{C}$. Then the map

$$\phi \to \int_{-\infty}^{\infty} f(x)\phi(x)\,dx \qquad \text{is}$$

(i) well-defined [since $\phi(x) = 0$ for all $|x| \geq X$, so no convergence problems arise for the integral],

(ii) linear in ϕ,

(iii) obeys our inequality above because

$$\left|\int_{-\infty}^{\infty} f(x)\phi(x)\,dx\right| \leq \left(\int_{-X}^{X} |f(x)|\,dx\right) \max(|\phi|).$$

So $f(x)$ defines a distribution in this way, and this distribution is denoted by the same symbol f as the function.

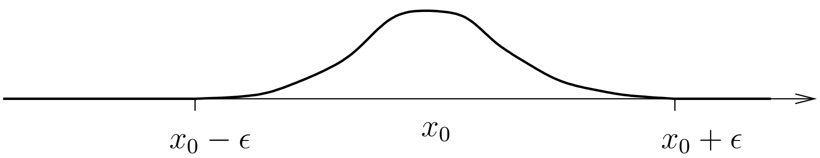
Distinct functions define distinct distributions: We need to show that if $f \neq g$, f and g both continuous, then there is **some** $\phi \in C_0^\infty(\mathbb{R})$ such that

$$\int_{-\infty}^{\infty} f(x)\phi(x)\,dx \neq \int_{-\infty}^{\infty} g(x)\phi(x)\,dx,$$

i.e.

$$\int_{-\infty}^{\infty} (f(x) - g(x))\phi(x)\,dx \neq 0. \tag{6}$$

Choose x_0 such that $f(x_0) \neq g(x_0)$. Then choose sufficiently small $\epsilon > 0$ such that $f - g$ is of fixed sign on $(x_0 - \epsilon, x_0 + \epsilon)$, and let $\phi \in C_0^\infty(\mathbb{R})$ be > 0 on $(x_0 - \epsilon, x_0 + \epsilon)$ and zero outside. Then (6) will hold.



So we now have a space $\mathcal{D}'$ of distributions, containing continuous functions as a special case. We can also make $f \in \mathcal{D}'$ if $f \in L^1_{loc}$, i.e. $f \in L^1[-X, X] \quad \forall X$, i.e. $\int_{-X}^{X} |f| dx < \infty \quad \forall X$. E.g. $f = \log|x|$ defines a distribution.

Examples of distributions that are not continuous functions

(1) The linear map $\delta : \mathbb{C}_0^\infty \to \mathbb{C}$ defined by $\langle \delta, \phi \rangle = \phi(0)$ is a standard distribution. It is linear and

$$|\langle u, \phi \rangle| \leq \max |\phi|.$$

It is known as the δ-distribution or δ-function. But there is no continuous or locally integrable function such that

$$\int_{-\infty}^{\infty} f(x)\phi(x)\,dx = \phi(0).$$

If there were, the left-hand side could be altered by changing ϕ in regions not including a, so that then $f = 0$ except at a (since the right-hand side would not change). But then using the ordinary definition of the integration

$$\int_{-\infty}^{\infty} f(x)\phi(x)\,dx = 0.$$

(2) Let $a \in \mathbb{R}$, $n \geq 0$,

$$\langle u, \phi \rangle = \left. \frac{d^n \phi}{d\phi^n} \right|_{x=a} = \phi^{(n)}(a).$$

This is linear, and

$$|\langle u, \phi \rangle| \leq \max |\phi^{(n)}|,$$

so it is a distribution. In fact this distribution is $(-1)^n \delta^{(n)}(x-a)$ in some sense (which can be seen if we integrate by parts).

(3) Let f be continuous or locally integrable and define

$$\langle u, \phi \rangle = \int_{-\infty}^{\infty} f(x) \frac{d^n \phi}{dx^n} \, dx.$$

Then

$$|\langle u, \phi \rangle| \leq \left(\int_{-X}^{X} f(x) \, dx \right) \max \left| \frac{d^n \phi}{dx^n} \right|,$$

so this is again a distribution.

Operations on distributions:

Linearity If u_1 and u_2 are distributions, α_1, $\alpha_2 \in \mathbb{C}$, then $\alpha_1 u_1 + \alpha_2 u_2$ is the distribution defined by

$$\langle \alpha_1 u_1 + \alpha_2 u_2, \phi \rangle = \alpha_1 \langle u_1, \phi \rangle + \alpha_2 \langle u_2, \phi \rangle.$$

$\mathcal{D}'$ becomes a $\mathbb{C}$-vector space in this way.

Differentiation Note that if f and f' are continuous then, for a choice of X that depends on the choice of ϕ,

$$\langle f', \phi \rangle = \int_{-\infty}^{\infty} f'(x)\phi(x) \, dx = \int_{-X}^{X} f'(x)\phi(x) \, dx = [f(x)\phi(x)]_{-X}^{X} - \int_{-X}^{X} f(x)\phi'(x) \, dx = -\langle f, \phi' \rangle.$$

So if u is any distribution we **define** its derivative u' by

$$\langle u', \phi \rangle = -\langle u, \phi' \rangle.$$

Multiplication by C^∞ functions If $g \in C^\infty(\mathbb{R})$ and $u \in \mathcal{D}'$ then we define $gu \in \mathcal{D}'$ by

$$\langle gu, \phi \rangle = \langle u, g\phi \rangle.$$

(Note that when $g \in C^\infty$, $\phi \in C_0^\infty$, then $g\phi \in C_0^\infty$.)

Notation We often write $\langle u, \phi \rangle$ as $\int_{-\infty}^{\infty} u(x)\phi(x) \, dx$, even when the distribution u does **not** arise from a continuous function. Consequently

(a) need to remember that "$u_1 = u_2$ as distributions" does not mean "$u_1(x) = u_2(x)$ for all x"; rather it means that

$$\langle u_1, \phi \rangle = \langle u_2, \phi \rangle \qquad \text{for all } \phi \in C_0^\infty(\mathbb{R}),$$

i.e.

$$\int_{-\infty}^{\infty} u_1(x)\phi(x) \, dx = \int_{-\infty}^{\infty} u_2(x)\phi(x) \, dx. \tag{7}$$

In checking (7) integration by parts can validly be used because of how u' is defined.

Translation along the axis If u is a distribution, then $u(x-a)$ is defined by

$$\int_{-\infty}^{\infty} u(x-a)\phi(x)\,dx = \int_{-\infty}^{\infty} u(x)\phi(x+a)\,dx.$$

E.g. the distribution $\delta(x-a)$ is such that

$$\int_{-\infty}^{\infty} \delta(x-a)\phi(x)\,dx = \int_{-\infty}^{\infty} \delta(x)\phi(x+a)\,dx = \phi(a),$$

by the definition of $\delta(x)$.

Green's functions as seen by distribution theory

For

$$L[u] = u'' + \alpha u' + \beta u = b,$$

plus linear homogeneous boundary conditions on u at 0 and 1, the Greens function $G(x,y)$ is characterised by

$$L[G] = \delta(x-y) \qquad \text{(Here we keep } y \text{ fixed and regard } G \text{ as a function of } x\text{)} \tag{8}$$

with G obeying the same boundary conditions as u. To show that this works, let G have these properties and let

$$u(x) = \int_0^1 G(x,y)b(y)\,dy.$$

Then u linearly depends on $G(x,y)$, so u always obeys the same linear homogeneous boundary conditions as G does. Moreover

$$\begin{aligned} L[u] &= \int_0^1 L[G]b(y)\,dy \qquad \text{since the derivatives in } L \text{ act on } x, \\ &= \int_0^1 \delta(x-y)b(y)\,dy = b(x). \end{aligned}$$

Check (properly) that $L[G] = \delta(x-y)$ along the following lines:

$$\begin{aligned} \int_{-\infty}^{\infty} L[G]\phi\,dx &= \int_{-\infty}^{\infty} (G''\phi + \alpha G'\phi + \beta G\phi)\,dx \\ &= \int_{-\infty}^{\infty} (G\phi'' - G(\alpha\phi)' + \beta G\phi)\,dx \qquad \text{(definition of distributional derivative)} \\ &= \lim_{\epsilon\to 0}\left(\left(\int_{-\infty}^{y-\epsilon} + \int_{y+\epsilon}^{\infty}\right)(G\phi'' - G(\alpha\phi)' + \beta G\phi)\,dx\right) \qquad \text{(since } G \text{ is continuous)} \\ &= \lim_{\epsilon\to 0}\left(\left(\int_{-\infty}^{y-\epsilon} + \int_{y+\epsilon}^{\infty}\right)(G''\phi + \alpha G'\phi + \beta G\phi)\,dx + [G'\phi - G\phi' + G\alpha\phi]_{x=y-\epsilon}^{x=y+\epsilon}\right) \\ &\qquad\qquad \text{(integration by parts on each subrange,} \\ &\qquad\qquad \text{and then use the classical form of } L[G]=0 \text{ there)} \\ &= \phi(y). \end{aligned}$$

We have shown that if

$$L[G] = \delta(x-y) \qquad \text{with the boundary conditions,}$$

then

$$L[u] = b(x) \qquad \text{with the boundary conditions}$$

is solved by

$$u = \int_0^1 G(x, y) b(y)\, dy.$$

Think of this as follows:

$$b(x) = \int_0^1 b(y) \delta(x - y)\, dy.$$

This represents the function $b(x)$ as a linear combination (integral) of the functions $\delta(x - y)$, taken with weighting $b(y)$. $G(x, y)$ is the response of the system represented by L and boundary conditions to a $\delta(x-y)$ on the RHS. Therefore, since the system responds **linearly** to its inputs, the response to $b(x)$ is a linear combination of the $G(x, y)$ taken with weighting $b(y)$. G is also called the **Fundamental solution** of the system. Solving for a second order ode $L[u] = u'' + \alpha u' + \beta u$,

$$\begin{array}{rcll} L[G] & = & \delta(x - y) & \text{in } [0, 1], \\ G & & \text{obeys the boundary conditions at } 0, 1, \end{array}$$

construct G obeying

$$\begin{array}{rcll} L[G] & = & 0 \quad \text{in } [0, y) & \text{and } G \text{ obeys the boundary condition at } 0, \\ L[G] & = & 0 \quad \text{in } (y, 1] & \text{and } G \text{ obeys the boundary condition at } 1, \\ G & & \text{continuous at } x = y, & \\ \left[\dfrac{\partial G(x, y)}{\partial x}\right]_{x=y-}^{x=y+} & = & 1. & \end{array}$$

Then this discontinuity in G' makes G'' have a δ-function singularity at $x = y$.

Generalisation

We have described $\mathcal{D}' = \mathcal{D}'(\mathbb{R})$, distributions on $\mathbb{R}$. If Ω is an open set in $\mathbb{R}^n$, (e.g. $\mathbb{R}^n$ itself) we define

$$C_0^\infty(\Omega) = \{\phi : \Omega \to \mathbb{C}, C^\infty \text{ with support}(\phi) \text{ compact } \subseteq \Omega\}.$$

We say u is a distribution on Ω, $u \in \mathcal{D}'(\Omega)$, if u is linear from $C_0^\infty(\Omega) \to \mathbb{C}$, taking ϕ to $\langle u, \phi \rangle$ and such that for all compact $K \subseteq \Omega$, there exist C and N such that

$$|\langle u, \phi \rangle| \leq C \sum_{i_1 + \cdots + i_n \leq N} \max \left| \left(\frac{\partial}{\partial x_1}\right)^{i_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{i_n} \phi \right|$$

for all ϕ such that $\mathrm{supp}(\phi) \subseteq K$. The δ-distribution on $\mathbb{R}^n$ is defined as on $\mathbb{R}$ by

$$\langle u, \phi \rangle = \phi(\mathbf{0}).$$

If u_1 and u_2 are in $\mathcal{D}'(\mathbb{R})$ then we can define a distribution $v \in \mathcal{D}'(\mathbb{R}^2)$ by $v(x, y) = u_1(x) u_2(y)$ i.e.

$$\langle v, \phi \rangle = \int \int v(x, y) \phi(x, y)\, dx\, dy = \int u_2(y) \left(\int u_1(x) \phi(x, y)\, dx \right) dy.$$

We can think of $\delta(x,y)$ in two dimensions as $\delta(x)\delta(y)$ since

$$\begin{aligned} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta(x)\delta(y)\phi(x,y)\,dx\,dy &= \int_{-\infty}^{\infty} \delta(y)\phi(0,y)\,dy \qquad \text{integrate over } x \text{ first} \\ &= \phi(0,0) \qquad \text{then integrate over } y. \end{aligned}$$

If u, $v \in \mathcal{D}'(\Omega)$ and $\Omega_1 \subset \Omega$, we say that $u = v$ on Ω_1 if

$$\langle u, \phi\rangle = \langle v, \phi\rangle \text{ for all } \phi \in C_0^\infty(\Omega_1).$$

$u = 0$ on a neighbourhood of x means what it says. If every point of Ω_1 has a neighbourhood on which $u = 0$ then $u = 0$ on Ω_1.
Proof Let $\phi \in C_0^\infty(\Omega_1)$. Then $K = \text{supp}(\phi)$ is compact $\subset \Omega_1$, and for all $x \in K$, $u = 0$ on some U_x neighbourhood of x. But $K \subset \cup U_i$, and $u = 0$ on $\cup U_i$.

Convergence of distributions

$$u_j \to u \text{ in } \mathcal{D}'(\Omega) \text{ iff } \langle u_j, \phi\rangle \to \langle u, \phi\rangle \text{ for all } \phi.$$

Weak convergence.
Note: earlier we introduced the adjoint operators L, L^*, where

$$\begin{aligned} Lu &:= a_2(x)\tfrac{d^2u}{dx^2} + a_1(x)\tfrac{du}{dx} + a_0(x)u, && \text{for } 0 < x < 1, \\ L^*v &:= \tfrac{d^2}{dx^2}(a_2(x)v) - \tfrac{d}{dx}(a_1(x)v) + a_0(x)v, && \text{for } 0 < x < 1, \end{aligned}$$

the matrix operator $\mathcal{B}(u) = 0$ gives the boundary conditions on u, $\mathcal{B}^*(v) = 0$ gives the boundary conditions on v (remember $\mathcal{B}^*$ is usually not the conjugate transpose of $\mathcal{B}$, but is obtained from the integral identity using integration by parts).
N.B. $B(u,v) =$ bilinear form involving boundary conditions, must equal 0 for all u, v satisfying $\mathcal{B}(u) = 0 = \mathcal{B}^*(v)$. Finally

$$\langle u, v\rangle = \int_0^1 u(x)v(x)dx.$$

Then

$$L^*G^*(x,y) = \delta(x-y)$$

defines the adjoint Green's function, and

$$u(x) = \int_0^1 G^*(x,y)f(y)dy$$

solves

$$Lu = f \quad \text{ on } 0 < x < 1, \quad \mathcal{B}(u) = 0.$$

Note

$$\begin{aligned} G^*(x,y) &= \langle \delta(\hat{x} - x), G^*(\hat{x}, y)\rangle_{\hat{x}} \\ &= \langle L_x G(\hat{x}, x), G^*(\hat{x}, y)\rangle_{\hat{x}} \\ &= \langle G(\hat{x}, x), L_x^* G^*(\hat{x}, y)\rangle_{\hat{x}} \\ &= \langle G(\hat{x}, x), \delta(\hat{x} - y)\rangle_{\hat{x}} \\ &= G(y,x) \end{aligned}$$

(using $\langle Lu, v\rangle = \langle u, L^*v\rangle$, etc.) Hence $G^*(x,y) = G(y,x)$, which is why in practice we mostly use $G(y,x)$. In particular, for self-adjoint problems, this implies that $G^*(x,y) = G(x,y)$, and so $G(x,y) = G(y,x)$. Thus G is symmetric in its arguments for self-adjoint problems, which often simplifies the solution procedure.

Lectures 5 and 6

5 Nonlinear systems of ode's

We shall consider systems in the form

$$\frac{du_i}{dt} = f_i(u_1, \cdots, u_n, t), \qquad (i = 1, \cdots, n),$$

which we write as

$$\frac{d\mathbf{u}}{dt} = \mathbf{f}(\mathbf{u}, t) \tag{9}$$

with vector $\mathbf{u}$ and $\mathbf{f}$.

Picard's Theorem System (9) with initial condition $\mathbf{u}(t_0) = \mathbf{u}_0 \in \mathbb{R}^n$ has a unique solution if $\mathbf{f}$ is continuous and obeys a Lipschitz condition

$$|\mathbf{f}(\mathbf{u}, t) - \mathbf{f}(\mathbf{u}', t)| \leq K|\mathbf{u} - \mathbf{u}'|.$$

Remarks

(1) A higher-order system, e.g.

$$\frac{d^2\mathbf{y}}{dt^2} = \mathbf{g}\left(\mathbf{y}, \frac{d\mathbf{y}}{dt}, t\right)$$

can be written in the form (9) by letting

$$\mathbf{v} = \frac{d\mathbf{y}}{dt}, \qquad \mathbf{u} = \begin{pmatrix} \mathbf{y} \\ \mathbf{v} \end{pmatrix}, \qquad \frac{d\mathbf{u}}{dt} = \begin{pmatrix} \mathbf{v} \\ \mathbf{g}(\mathbf{y}, \mathbf{v}, t) \end{pmatrix} = \mathbf{f}(\mathbf{u}, t),$$

so if $\mathbf{y} \in \mathbb{R}^k$ then $\mathbf{u} \in R^{2k}$ and the initial conditions needed are $\mathbf{y}(t_0) = \mathbf{y}_0$ and $\mathbf{v}(t_0) = \mathbf{v}_0$, position and velocity.

(2) If $\mathbf{f}$ is C^1, i.e. all $\partial f_i/\partial u_j$ exist and are continuous, then for $|\mathbf{u}|$, $|\mathbf{u}'| \leq R$ the Lipschitz condition holds

$$|\mathbf{f}(\mathbf{u}, t) - \mathbf{f}(\mathbf{u}', t)| \leq \left(\max_{|\mathbf{u}| \leq R} \left|\left|\frac{\partial f_i}{\partial u_j}\right|\right|\right) |\mathbf{u} - \mathbf{u}'|.$$

Hence solutions exist and are unique while they remain in $|\mathbf{u}| \leq R$.

One check to make on a model is that to be realistic it generally should ensure that $\mathbf{u}$ remains bounded as $t \to +\infty$, and this will then ensure uniqueness of the solution for all $t \geq t_0$. Otherwise solution will only exist until $\mathbf{u}(t) \to \infty$ or some singularity of $\mathbf{f}$.

(3) If $\mathbf{f}$ is C^m in $(\mathbf{u}, t)$ then the solution $\mathbf{u}$ is C^m in $\mathbf{u}_0$, t. If $\mathbf{f}$ is C^m in some parameters, so is the solution.

(4) The system is called "autonomous" if $\mathbf{f} = \mathbf{f}(\mathbf{u})$. The curve traced out in $\mathbb{R}^n$ by $\mathbf{u}(t)$ for $t \geq t_0$ is called "the trajectory starting from $\mathbf{u}_0$".

5.1 Discontinuous systems

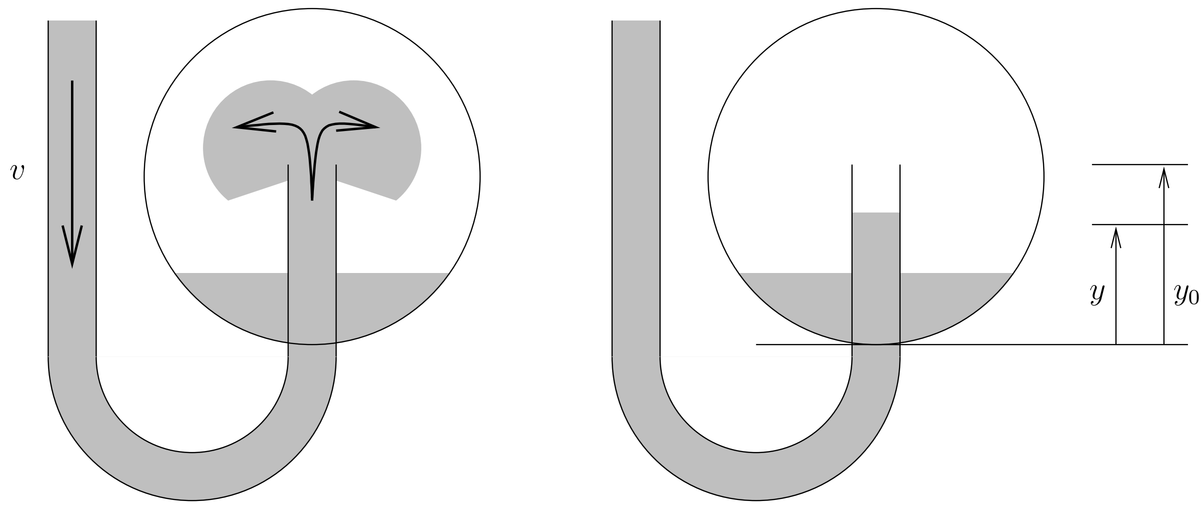
Realistic models

(a) Sometimes have discontinuous or non-Lipschitz $\mathbf{f}$.

(b) Should still have a unique solution in $t > t_0$.

(c) May not have a unique solution in $t < t_0$.

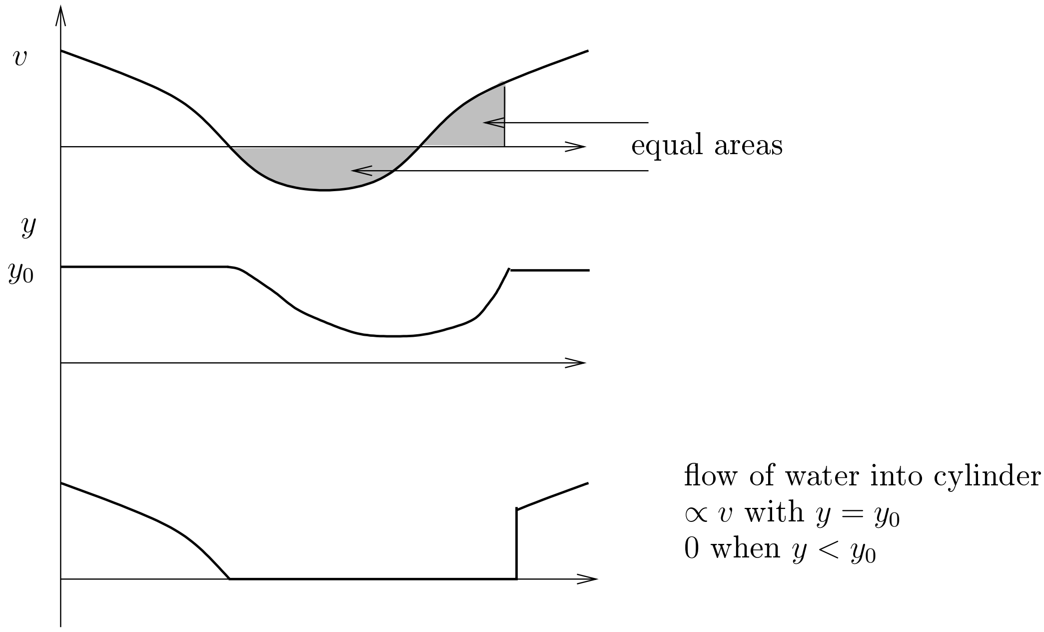
Example A pipe supplies water to a cylinder, but the flow occasionally reverses.



$$\frac{dy}{dt} = \begin{cases} v & \text{except} \\ 0 & \text{if } y = y_0 \text{ and } v \geq 0 \end{cases}$$

Discontinuous in y, v. Couple with dv/dt determined by pressure difference, friction, inertia of water in pipe.

The typical motion involves

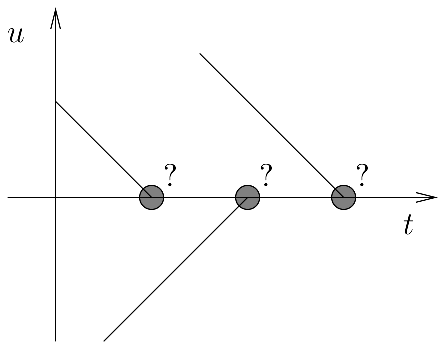


This system does have a unique solution for t increasing, but does not for t decreasing. Cases like this are dealt with computationally as

$$\dot{\mathbf{u}} = \begin{cases} \mathbf{f}_1(\mathbf{u}, t) & \text{for } \mathbf{g}(\mathbf{u}, t) < 0, \\ \mathbf{f}_2(\mathbf{u}, t) & \text{for } \mathbf{g}(\mathbf{u}, t) \geq 0 \end{cases}$$

with $\mathbf{f}_1$ and $\mathbf{f}_2$ continuous. Integrate $\mathbf{f}_1$ up to the point where $\mathbf{g}$ reaches 0, then switch to $\mathbf{f}_2$ and repeat. Matlab has facilities to do this. But note that it is easy to write down discontinuous models that have no solution, e.g. $u \in \mathbb{R}$,

$$\dot{u} = \begin{cases} -1 & \text{for } u > 0, \\ 1 & \text{for } u \leq 0 \end{cases}$$



5.2 General principles for autonomous equations $\dot{\mathbf{u}} = \mathbf{f}(\mathbf{u})$

We consider cases where $\mathbf{f}$ is continuous and Lipschitz and assume that any discontinuities are patched up correctly. We write $\mathbf{u}(t) = \Phi(t, \mathbf{u}_0)$ for the solution starting from $\mathbf{u}_0$ at $t = 0$.

Equilibrium points are $\alpha \in \mathbb{R}^n$ such that $\mathbf{f}(\alpha) = 0$. Then $\Phi(t, \alpha) = \alpha$ for all t. If $\mathbf{f}$ is differentiable at α, the **stability** of such a point is determined by the eigenvalues of $A = D\mathbf{f}(\alpha)$, i.e.

$$A = (a_{ij}) \qquad a_{ij} = \left.\frac{\partial f_i}{\partial u_j}\right|_\alpha .$$

If every eigenvalue λ of A has $\mathrm{Re}(\lambda) < 0$ then α is stable in the sense that every trajectory starting near enough to α remains near α and converges to α as $t \to \infty$.

Lyapunov criterion If $V(\mathbf{u})$ is a real-valued C^1 function on $\mathbb{R}^n$ then we define

$$\dot{V}(\mathbf{u}) = \nabla V(\mathbf{u}) \cdot \mathbf{f}(\mathbf{u}) = \sum_i \frac{\partial V}{\partial u_i} f_i = \text{rate of change of } V \text{ along a trajectory.}$$

If $D = \{\mathbf{u} : V(\mathbf{u}) \leq k\}$ is a bounded set containing α and $\dot{V} < 0$ on D except at α, then every trajectory starting in D remains in D and converges to α as $t \to \infty$. We say "D is part of the basin of attraction of α". This provides a way of finding explicitly a region that converges to α.

Proof Suppose some trajectory escapes from D. At its last point in D it has $V \leq k$ and $\dot{V} \geq 0$. But this can only happen at α. But trajectory from α stays there. $\square$

Lyanunov Stability The critical point α is said to be Lyapunov stable if given any neighbourhood U of α there exists a neighbourhood V of α such that $\mathbf{u}(0) \in V \Rightarrow \mathbf{u}(t) \in U$ for all $t \geq 0$.

Closed trajectories If $\mathbf{u}_0$ is not an equilibrium point, but $\Phi(t, \mathbf{u}_0) = \mathbf{u}_0$ for some $t > 0$ then the trajectory starting from $\mathbf{u}_0$ returns to $\mathbf{u}_0$ at some $t > 0$ and we call it a **closed trajectory**. The least such t is called the **period**.

Note: A linear system e.g. $\ddot{u} + a\dot{u} + bu = 0$, has closed trajectories only in special circumstances ($a = 0, b > 0$) and when they exist the amplitude is arbitrary (i.e. they have lots of closed trajectories). Nonlinear systems have closed trajectories "more often" and the amplitude is fixed by the nonlinearity.

5.3 Phase-plane methods (for $n = 2$)

When $n = 2$ we represent the start $\mathbf{u}$ by points in $\mathbb{R}^2$ and trajectories by curves, with arrows in the direction of increasing t. Note that the curves can only intersect if $\dot{u}_1 = 0 = \dot{u}_2$ by writing (if say $\dot{u}_2 \neq 0$) $\frac{du_1}{du_2} = \frac{f_1(u_1,u_2)}{f_2(u_1,u_2)}$ and then using $\frac{f_1}{f_2}$ Lipschitz to imply that u_1 is (uniquely) a well-defined function of u_2. (If curves intersect, then consider the uniqueness theorem for the initial data being the point of intersection.) Hence the equilibrium points α where $f(\alpha) = 0$ "geometrically" organise the solution curves in the (u_1, u_2) plane. The only other "geometrically" interesting feature is the existence of periodic solutions.

Poincaré-Bendixson Theorem Suppose D is a bounded closed region in $\mathbb{R}^2$ containing no equilibrium points and such that every trajectory starting in D remains in D. Then D contains a closed trajectory.

Bendixson-Dulac Suppose D is a simply connected region of $\mathbb{R}^2$ and $\nabla \cdot (\rho \mathbf{f}) > 0$ in D for some scalar function ρ. Then D contains no closed trajectory.

These results depend strongly on the topology of $\mathbb{R}^2$: none holds in $\mathbb{R}^n$ for $n \geq 3$.

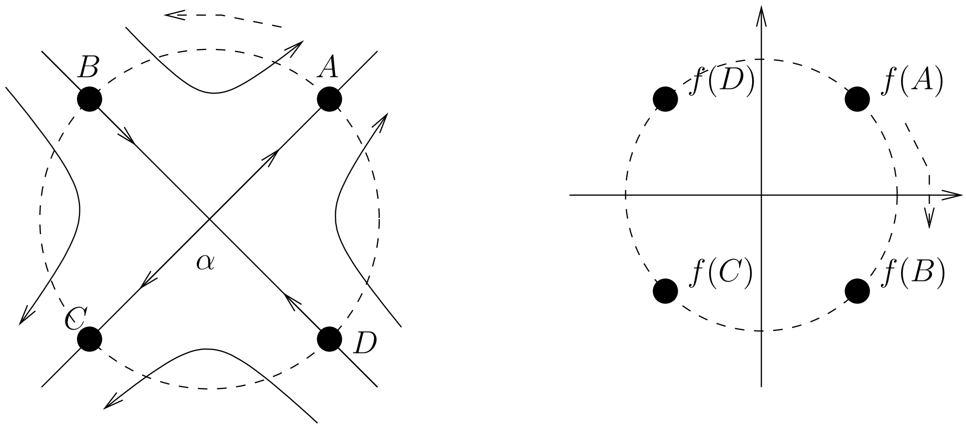
Poincaré Index for an isolated equilibrium point α is

$$I(\alpha) = \text{number of times } \mathbf{f}(\mathbf{u}) \text{ encircles 0 anticlockwise when } \mathbf{u} \text{ encircles } \alpha \text{ anticlockwise.}$$

If a closed trajectory encloses equilibrium points $\alpha_1, \ldots, \alpha_k$, then $\sum_1^k I(\alpha_i) = 1$.
Example. A saddle point is when

$$A = \left.\left(\frac{\partial f_i}{\partial u_j}\right)\right|_{\mathbf{u}=\alpha} \qquad \text{has eigenvalues } \lambda_1 < 0 < \lambda_2.$$

Trajectories near α look like



$\mathbf{f}(\mathbf{u})$ encircles 0 clockwise. So a closed trajectory cannot contain **only** a saddle. We now consider some useful special cases for applications.

5.4 Nonlinear Oscillators

Consider the 2 dimensional system

$$\ddot{u} + f(u)\dot{u} + g(u) = 0.$$

Let

$$F(u) = \int_0^u f(\xi)\, d\xi, \qquad \text{so } \dot{F}(u) = f(u)\dot{u}.$$

Let $v = \dot{u} + F(u)$. Then

$$\begin{aligned} \dot{u} &= v - F(u), \\ \dot{v} &= \ddot{u} + f(u)\dot{u} = -g(u). \end{aligned}$$

This is known as the Liénard form. If

(i) $g(u)\text{sign}(u) > 0$ for large $|u|$ (restoring force at large $|u|$),

(ii) $F(u)\text{sign}(u) \to \infty$ as $|u| \to \infty$ (positive damping at large u)

then all trajectories are bounded and all eventually enter a fixed bounded region.
If also

(iii) $g(u)\text{sign}(u) > 0$ for **all** $u \neq 0$ and $f(u)\text{sign}(u) < 0$ for **small** $u \neq 0$ (negative damping at small u)

then a closed trajectory exists. If also

(iv) f even, g odd, and $F(u) < 0$ on $(0, u_0)$, $F(u) > 0$ on (u_0, ∞)

then the closed trajectory is unique.

5.5 Simple examples

(i) Suppose we have a linear problem (force $\approx$ frictional resistance + spring restoring)

$$\ddot{u} + a\dot{u} + bu = 0.$$

(Note that we can always change the timescale ($t = T\hat{t}$) so that

$$\frac{1}{T^2}\ddot{u} + \frac{1}{T}a\dot{u} + bu = 0 \quad \text{for } u(\hat{t}),$$

and so we can choose $T = \frac{1}{\sqrt{|b|}}$ to scale b so that $b^2 = 1$.)
Then

$$\begin{aligned} \dot{u} &= v, \\ \dot{v} &= \ddot{u} = -av - bu \end{aligned}$$

are the phase plane equations for orbits $(u(t), v(t) = \dot{u}(t)) \in \mathbb{R}^2$. See Figures 1 and 2 for an illustration of the phase plane. Orbits spiral in if $a, b > 0$ and spiral out if $a < 0, b > 0$.

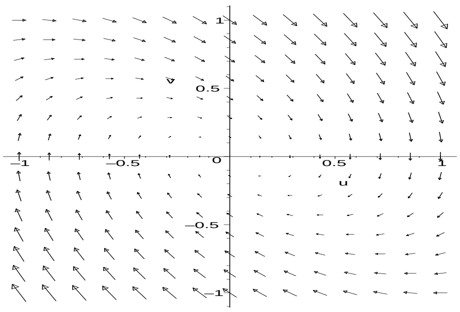


Figure 1: Phase plane (a=1, b=1).

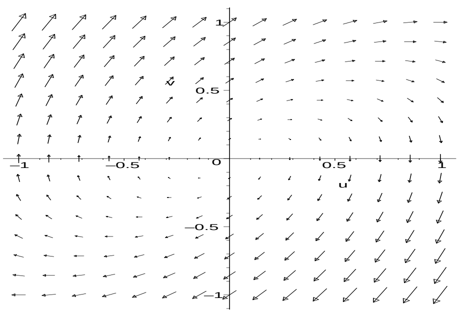


Figure 2: Phase plane (a=-1, b=1).

Note that

$$\frac{dv}{du} = \frac{\dot{v}}{\dot{u}} = -\frac{av + bu}{v} = -\frac{aw + b}{w}$$

can be integrated by writing $w := \frac{v}{u}$, as it is homogeneous.

(ii) Suppose we have

$$\ddot{\theta} = -\sin\theta.$$

Then

$$\dot{u} = \dot{\theta} = v, \qquad \dot{v} = \ddot{\theta} = -\sin u,$$

and (for $v \neq 0$)

$$\frac{dv}{du} = \frac{\dot{v}}{\dot{u}} = -\frac{\sin u}{v}$$

which implies $vdv = -\sin u du$, i.e.

$$\frac{1}{2}v^2 = \cos u + \text{ constant.}$$

This gives the orbits in the (u, v) plane.
Conservation of energy is

$$\frac{1}{2}v^2 - \cos u = E.$$

$(0, 0)$ is a centre, and $((2n+1)\pi, 0)$ with $n \in \mathbb{Z}$ are saddle points. See Figure 3 for an illustration.

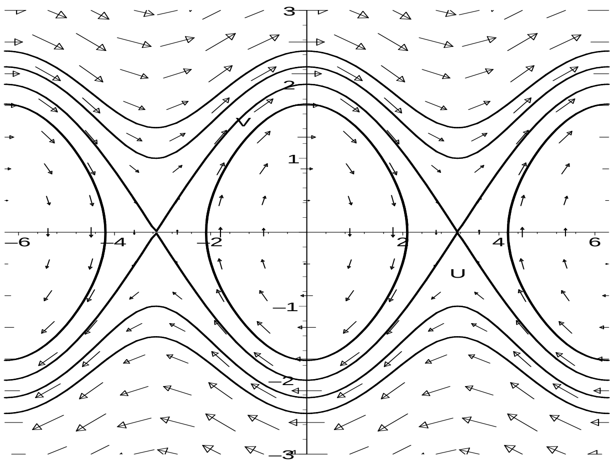


Figure 3: Phase plane and orbits.

Lectures 7 and 8

The van der Pol equation

Consider

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

In terms of the Liénard form $g(u) = u$, $f(u) = \mu(u^2 - 1)$, $F(u) = \mu(u^3/3 - u)$. All the conditions (i)-(iv) of the earlier paragraph hold and there is a unique closed trajectory. We shall estimate its position and period, for μ large and for μ small.

Small μ Suppose there is a periodic solution with period $2\pi/\omega$. We aim to approximate u and ω for small μ. Rescale time so that the period is 2π, i.e. set $T = \omega t$, giving

$$\omega^2 \ddot{u} + \mu\omega(u^2 - 1)\dot{u} + u = 0. \tag{10}$$

Fix the origin of time by choosing $\dot{u}(0) = 0$. We expand

$$\begin{aligned} \omega &= \omega_0 + \mu\omega_1 + \mu^2\omega_2 + \cdots, \\ u &= u_0 + \mu u_1 + \mu^2 u_2 + \cdots, \end{aligned}$$

i.e.

$$\begin{aligned} \omega &= \omega_0 + O(\mu), \\ \omega &= \omega_0 + \mu\omega_1 + O(\mu^2), \\ \omega &= \omega_0 + \mu\omega_1 + \mu^2\omega_2 + O(\mu^3), \end{aligned}$$

etc., as $\mu \to 0$, and similarly for u. These series are useful (even if they diverge!). In (10) we have

$$\begin{aligned} &(\omega_0 + \mu\omega_1 + \mu^2\omega_2 + \cdots)^2(\ddot{u}_0 + \mu\ddot{u}_1 + \mu^2\ddot{u}_2 + \cdots) \\ &\quad + \mu\left((u_0 + \mu u_1 + \mu^2 u_2 + \cdots)^2 - 1\right)(\dot{u}_0 + \mu\dot{u}_1 + \mu^2\dot{u}_2 + \cdots)(\omega_0 + \mu\omega_1 + \mu^2\omega_2 + \cdots) \\ &\quad + (u_0 + \mu u_1 + \mu^2 u_2 + \cdots) = 0. \end{aligned}$$

Equate coefficients of powers of μ. $O(1)$:

$$\omega_0^2 \ddot{u}_0 + u_0 = 0.$$

Hence $u_0 = A_0 \cos T/\omega_0$ (say). But the period in new time is 2π. Hence $\omega_0 = 1$. Now $O(\mu)$:

$$2\omega_1 \ddot{u}_0 + \ddot{u}_1 + (u_0^2 - 1)\dot{u}_0 + u_1 = 0,$$

i.e.

$$\begin{aligned} \ddot{u}_1 + u_1 &= 2\omega_1 A_0 \cos T + (A_0^2 \cos^2 T - 1)A_0 \sin T \\ &= 2\omega_1 A_0 \cos T - A_0 \sin T + A_0^3 \left(\frac{\sin T + \sin 3T}{4}\right), \end{aligned}$$

since $\cos^2 T \sin T = 1/2 \sin 2T \cos T = 1/4(\sin T + \sin 3T)$. Then u_1 will exhibit **resonance** unless there are no terms in $\cos T$, $\sin T$ on the RHS. If the solution for u_1 has terms like $T \sin T$ it will not be periodic. So, if we want $A_0 \neq 0$ we must have $\omega_1 = 0$ and $A_0 = 2$. Thus we have the leading-order amplitude of the limit cycle.

Then

$$\ddot{u}_1 + u_1 = 2\sin 3T,$$

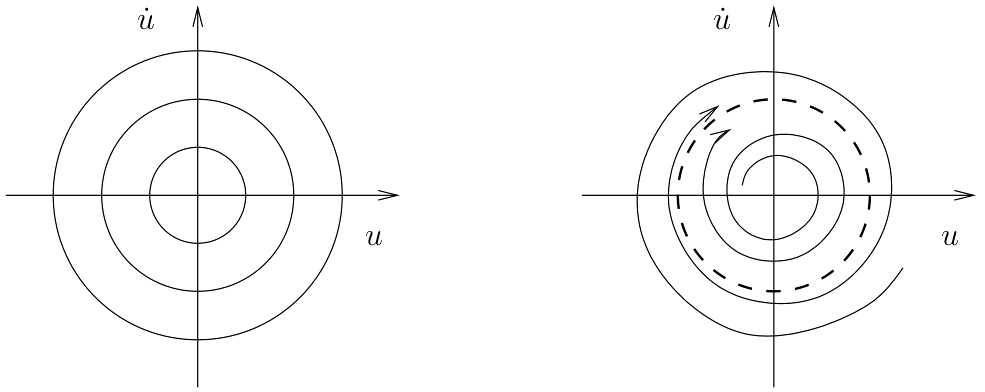
so

$$u_1 = -\frac{\sin 3T}{4} + A_1 \cos T + B_1 \sin T.$$

Then $\dot{u}_1(0) = 0$ gives $B = 3/4$. Hence

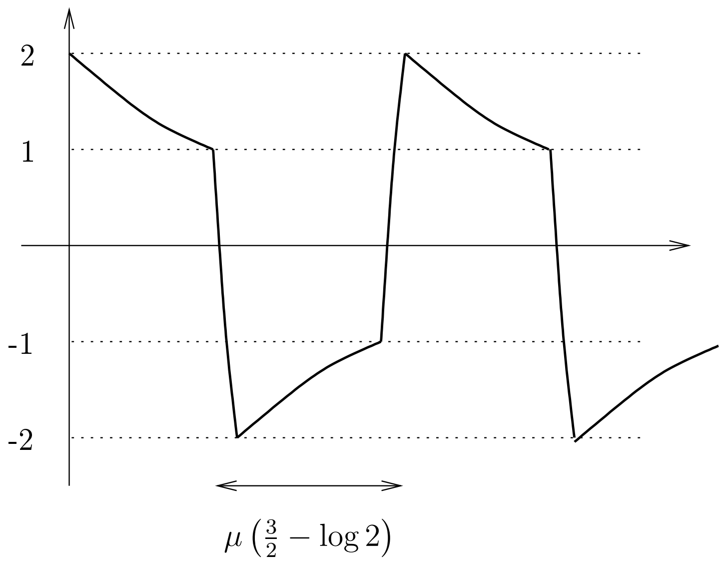
$$u_1 = A_1 \cos T + \frac{3\sin T - \sin 3T}{4} = A_1 \cos T + \sin^3 T.$$

Continue to μ^2? Resonance condition on u_2 determines ω_2 and A_1. Etc.



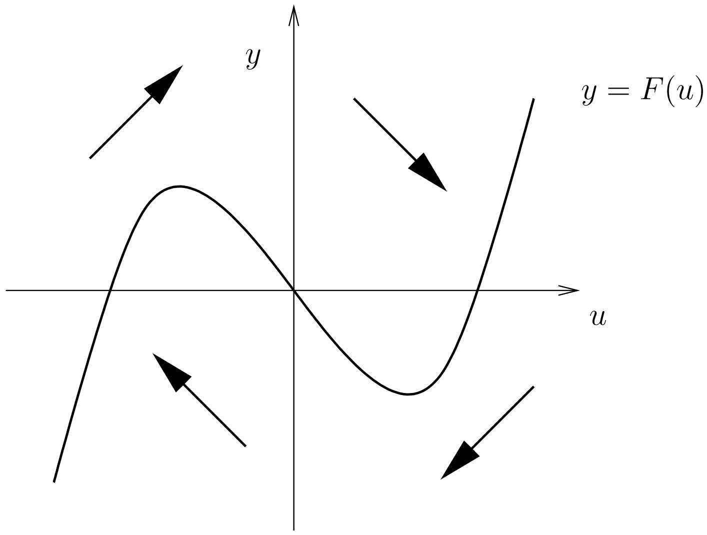
$\mu = 0$ $\qquad\qquad$ $\mu > 0$: only $A = 2$ survives

Large μ We shall show that the oscillation looks like



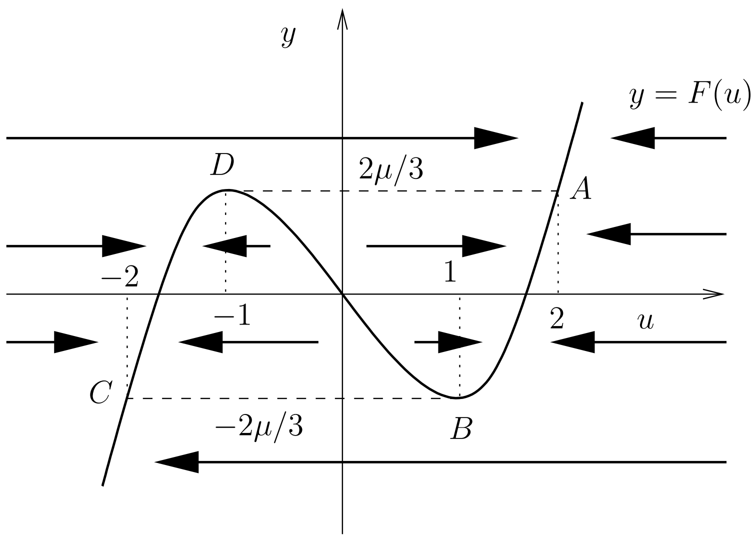
with long slow regions separated by rapid transitions. "Relaxation Oscillations". Write the equation in Liénard form as

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



Arrows indicate the general form of the motion for all μ. However, if $\mu \gg 1$ then $|\dot{u}| \gg |\dot{y}|$ except near the curve $y = F(u)$. So for a given initial y, u tends very rapidly to a root of $F(u) = y$:

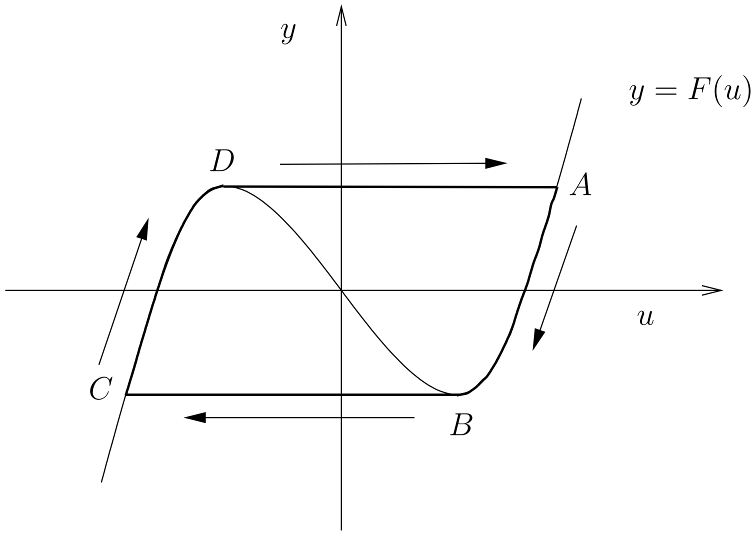
- if $y > \dfrac{2\mu}{3}$, the unique root between A and $+\infty$.
- if $y < -\dfrac{2\mu}{3}$, the unique root between C and $-\infty$.
- if $-\dfrac{2\mu}{3} < y < \dfrac{2\mu}{3}$, $u \to$ point on AD if it starts to the right of BD.
- if $-\dfrac{2\mu}{3} < y < \dfrac{2\mu}{3}$, $u \to$ point on BC if it starts to the left of BD.



A trajectory starting say from $(0, \mu)$ quickly moves across to the branch $A\infty$. Then it remains close to the curve $y = F$ and since $\dot{y} = -u$ it moves slowly down the curve. When it reaches B, (u, y) flies across horizontally to near C. Then $\dot{y} = -u > 0$ so (u, y) climbs up close to $y = F(u)$ to D. Then (u, 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

$$t_{AB} = \int_A^B \frac{dy}{\dot{y}} = -\int_A^B \frac{dy}{u} = \int_1^2 \frac{dy}{du}\frac{du}{u} = \int_1^2 \frac{\mu(u^2-1)du}{u} = \mu\left[\frac{u^2}{2} - \log u\right]_1^2 = \mu\left(\frac{3}{2} - \log 2\right).$$

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



6 Singular perturbations and Matched Asymptotic Expansions

6.1 Asymptotic expansions

Order notation We say that $f(a)$ is of order $g(a)$ as $a \to a_0$ if there exist two constants K, δ such that $|f(a)| \leq Kg(a)$ for $0 < |a - a_0| \leq \delta$. We write

$$f(a) = O(g(a)) \text{ as } a \to a_0.$$

We say that $f(a)$ is of smaller order than $g(a)$ as $a \to a_0$ if $f(a)/g(a) \to 0$ as $a \to a_0$, we write

$$f(a) = o(g(a)) \text{ as } a \to a_0.$$

(1) K is called the "implied constant in the $O(g(a))$ term".

(2) a may be real/complex, a finite value or infinity. The limit may be through integers $n \to \infty$.

(3) f and g need not be defined at $a = a_0$.

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, we often take order to be the 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 $\delta > 0$.

Suppose $f(\epsilon)$ is defined for $\epsilon > 0$ and suppose we have a sequence of approximations to it:

$$\begin{aligned} f(\epsilon) - f_0 &= O(\epsilon) \text{ as } \epsilon \to 0, \\ f(\epsilon) - f_0 - \epsilon f_1 &= O(\epsilon^2) \text{ as } \epsilon \to 0, \\ f(\epsilon) - f_0 - \epsilon f_1 - \cdots - \epsilon^N f_N &= O(\epsilon^{N+1}) \text{ as } \epsilon \to 0. \end{aligned}$$

Then we write

$$f(\epsilon) \sim \sum_0^\infty f_n \epsilon^n \text{ as } \epsilon \to 0.$$

We say "f has the asymptotic expansion $\sum_0^\infty f_n\epsilon^n$ as $\epsilon \to 0$". We have a sequence of useful approximations to $f(\epsilon)$ **even if** $\sum_0^\infty f_n\epsilon^n$ is actually divergent for every $\epsilon > 0$.

Important:

$$f(\epsilon) \sim \sum_0^\infty f_n \epsilon^n \text{ as } \epsilon \to 0$$

is not saying anything about the infinite sum; in particular it is not saying that the infinite sum converges. Rather it is an assertion about each finite sum. If we write the partial sums as $S_N = f_0 + \ldots + f_N\epsilon^N$ then

$$\text{ordinary convergence } \to \quad \Leftrightarrow \quad |S_N - f(\epsilon)| \to 0 \text{ as } N \to \infty \quad \text{for fixed } \epsilon.$$

$$\text{asymptotic } \sim \quad \Leftrightarrow \quad |S_N - f(\epsilon)| = O(\epsilon^{N+1}) \text{ as } \epsilon \to 0 \quad \text{for any fixed } N.$$

Example

$$f(\epsilon) = \int_0^\infty \frac{e^{-t}}{1+\epsilon t}\, dt \sim \sum_{n=0}^\infty (-1)^n n! \epsilon^n.$$

(This is shown by repeated integration by parts.)

If f depends on some other variable x (scalar or vector), then the f_n will depend on x and we say

$$f(x,\epsilon) \sim \sum_0^\infty f_n(x)\epsilon^n \text{ as } \epsilon \to 0,$$

to mean that for each N

$$f(x,\epsilon) - \sum_0^N f_n(x)\epsilon^n = O(\epsilon^{N+1}) \qquad \text{as } \epsilon \to 0.$$

If the implied constant in $O(\epsilon^{N+1})$ is uniform in some range of x, we say it is a uniform asymptotic expansion in that range.

Variations on this

(1) The limit may be $a \to a_0$ rather than $\epsilon \to 0$.

(2) The sequence 1, ϵ, $\epsilon^2, \ldots$ may be replaced by any $\phi_n(\epsilon)$ as long as $\phi_{n+1}(\epsilon) = o(\phi_n(\epsilon))$ as $\epsilon \to 0$. Then

$$f(x,\epsilon) \sim \sum_0^\infty f_n(x)\phi_n(\epsilon) \qquad \text{as } \epsilon \to 0$$

means that for all N

$$f(x,\epsilon) - \sum_0^N f_n(x)\phi_n(\epsilon) = O(\phi_{N+1}(\epsilon)) \text{ as } \epsilon \to 0.$$

For example $\log\epsilon$ often crops up.

Worrying example Let $I(x)$ be defined for $\lambda > 0$ by

$$I(\lambda) = \int_0^\pi \frac{\cos\lambda t}{1+t^2}\, dt.$$

We can show by integration by parts that

$$I(\lambda) = \frac{\sin\pi\lambda}{\lambda(1+\pi^2)} - \frac{2\pi\cos\pi\lambda}{\lambda^2(1+\pi^2)^2} - \frac{(6\pi^2-2)\sin\pi\lambda}{\lambda^3(1+\pi^2)^3} + \frac{24(\pi^2-1)\cos\pi\lambda}{\lambda^4(1+\pi^2)^4} + \cdots.$$

When $\lambda = 10$ the asymptotic expansion of $I(\lambda)$ gives

$$I(10) \approx 0 - 0.000531805 + 0 + 0.00000479083 + \cdots$$

suggesting $I(10)$ is well approximated by -0.0005270. But $I(10) \approx -0.0004558$.

If we had done instead

$$J(\lambda) = \int_\pi^\infty \frac{\cos\lambda t}{1+t^2}\, dt$$

we find the same expansion

$$J(\lambda) = -\frac{\sin\pi\lambda}{\lambda(1+\pi^2)} + \frac{2\pi\cos\pi\lambda}{\lambda^2(1+\pi^2)^2} + \frac{(6\pi^2-2)\sin\pi\lambda}{\lambda^3(1+\pi^2)^3} - \frac{24(\pi^2-1)\cos\pi\lambda}{\lambda^4(1+\pi^2)^4} + \cdots;$$

the expansion suggests $J(10) \approx 0.0005270$. This time numerically we find $J(10) \approx 0.0005271$.

The reason is

$$\int_0^\infty \frac{\cos\lambda t}{1+t^2}\, dt = \frac{1}{2}\int_{-\infty}^\infty \frac{\cos\lambda t}{1+t^2}\, dt = \frac{1}{2}\text{Real}\left(\int_{-\infty}^\infty \frac{e^{i\lambda t}}{1+t^2}\, dt\right) = \frac{1}{2}2\pi i\frac{e^{-\lambda}}{2i} = \frac{\pi e^{-\lambda}}{2}$$

using the calculus of residues. The difference between I and J is exponentially small. In calculating $I(10)$ we find

$$-0.0005270 + \frac{\pi e^{-10}}{2} \approx -0.0004558.$$

But this term will never appear in the asymptotic expansion.

6.2 Singular Perturbations

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

Linear example

$$D_\epsilon : \quad \begin{cases} \epsilon u'' + u' + u = 0, \\ u(0) = a, \quad u(1) = b. \end{cases}$$

When $\epsilon = 0$, we have

$$D_0 : \quad \begin{cases} u' + u = 0, \\ u(0) = a, \quad u(1) = b. \end{cases}$$

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

If u is the solution to D_ϵ then one possible behaviour in such cases is that

- over most of the range $\epsilon d^k u/dx^k$ is small, and u approximately obeys D_0,
- in certain regions, often near the ends of the range, $\epsilon d^k u/dx^k$ is **not** small, and u adjusts itself to the boundary conditions (i.e. $d^k u/dx^k$ 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 electrodynamics 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 the rules

 (a) inner solutions obey the boundary conditions;

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

This is the method of **matched asymptotic expansions**.

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 $u(x) = u_L(x_L)$. Then

$$\begin{aligned} \frac{du}{dx} &= \frac{du_L}{dx_L}\frac{dx_L}{dx} = \epsilon^{-\alpha}\frac{du_L}{dx_L}, \\ \frac{d^2u}{dx^2} &= \epsilon^{-2\alpha}\frac{d^2u_L}{dx_L^2}, \end{aligned}$$

so that

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

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

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

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 u as

$$u(x) = u_M(x) \sim u_{M0}(x) + \epsilon u_{M1}(x) + \cdots,$$

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

$$u(x) = u_L(x_L) \sim u_{L0}(x) + \epsilon u_{L1}(x) + \cdots,$$

(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

$$u(x) = u_R(x_R) \sim u_{R0}(x) + \epsilon u_{R1}(x) + \cdots.$$

Solution on left The equation in the inner variable reads

$$\frac{d^2u_L}{dx_L^2} + \frac{du_L}{dx_L} + \epsilon u_L = 0.$$

Inserting the expansion and equating coefficients of powers of ϵ gives

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

etc. Hence

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

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

Solution in middle The equation in the outer variable reads

$$\epsilon\frac{d^2u_M}{dx^2}+\frac{du_M}{dx}+u_M=0.$$

Inserting the expansion and equating coefficients of powers of ϵ gives

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

etc. Hence

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

Solution on right The equation in the inner variable reads

$$\frac{d^2u_R}{dx_R^2}+\frac{du_R}{dx_R}+\epsilon u_R=0.$$

Inserting the expansion and equating coefficients of powers of ϵ gives

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

etc. Hence

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

To satisfy $u(1)=b$ 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.

Lectures 9 and 10

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

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

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

$$\begin{aligned} u_L &= A_{L0} + B_{L0}e^{-\epsilon^{\alpha-1}\hat{x}} + O(\epsilon), \\ &\sim A_{L0} + E.S.T. + O(\epsilon), \end{aligned}$$

where $E.S.T.$ stands for exponentially small terms, while

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

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

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

Thus the u 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

$$u_R = A_{R0} + B_{R0}e^{-\epsilon^{\alpha-1}\tilde{x}} + O(\epsilon), \tag{11}$$

while

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

Clearly to match we cannot have exponential growth in (11), so that

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

Again the u 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

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

There is no rapid variation in u 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 take off 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

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

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

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

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

Higher-order terms At order ϵ in each region

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

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 $\hat{x}$:

$$\begin{aligned} u_L &= eb + (a - eb)e^{-\epsilon^{\alpha-1}\hat{x}} + \epsilon\left(-eb\epsilon^{\alpha-1}\hat{x} + (a - eb)\epsilon^{\alpha-1}\hat{x}e^{-\epsilon^{\alpha-1}\hat{x}} + A_{L1} + B_{L1}e^{-\epsilon^{\alpha-1}\hat{x}}\right)\epsilon + O(\epsilon^2), \\ &\sim eb + E.S.T. \\ &\quad - eb\epsilon^{\alpha}\hat{x} + A_{L1}\epsilon + E.S.T. \\ &\quad + O(\epsilon^2) \end{aligned}$$

while

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

Matching we find that

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

(note some terms jump order: $-\epsilon^{\alpha}\hat{x}eb$ comes from inner expansion of first outer term, but from outer expansion of second inner term.) Note that in order for the neglected terms ϵ^{2al} 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

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

while

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

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

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

Composite expansion Composite is $u_L + u_M -$ overlap. Write u_L in terms of outer variable:

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

and u_M in terms of inner variables

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

The common value in the overlap region is

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

Hence the composite expansion is

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

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. 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 revisiting

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

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

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

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 $O(1)$.

Other equations If we have say

$$\epsilon u'' + u' + \sin u = 0,$$

essentially the same procedure works.

$$\begin{aligned} u &= u_0 + \epsilon u_1 + \cdots, \\ \sin u &= \sin u_0 + \left(\epsilon u_1 + \epsilon^2 u_2 + \cdots\right)\cos u_0 - \left(\epsilon u_1 + \epsilon^2 u_2 + \cdots\right)^2 \frac{\sin u_0}{2} + \cdots \\ &= \sin u_0 + \epsilon(u_1 \cos u_0) + \epsilon^2\left(u_2 \cos u_0 - \frac{u_1^2 \sin u_0}{2}\right) + \cdots. \end{aligned}$$

So, for instance, the differential equations in the middle region are

$$\begin{aligned} u_0' + \sin u_0 &= 0, \\ u_0'' + u_1' + u_1 \cos u_0 &= 0, \\ u_1'' + u_2' + u_2 \cos u_0 - \frac{u_1^2 \sin u_0}{2} &= 0, \end{aligned}$$

etc. Note that the first of these is nonlinear, but all the rest are linear (but with non-constant coefficients).

7 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. Similar to WKB—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

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

Last time we looked for **periodic** solution for small ϵ. Here we will study the **initial value problem**

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

Treating the problem as a regular perturbation expansion in ϵ gives

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

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 $\frac{3}{8}\epsilon t \cos t$, etc. When the second term in an expansion becomes as big as the first (here when $\epsilon t = \frac{8}{3} = O(1)$) 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 long 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:

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

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

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

so that

$$\ddot{u} = u_{\tau\tau} + 2\epsilon u_{\tau T} + \epsilon^2 u_{TT}.$$

Then we expand

$$u(\tau, T;\epsilon) = u_0(\tau, T) + \epsilon u_1(\tau, T) + \cdots .$$

At ϵ^0 we find

$$u_{0\tau\tau} + u_0 = 0 \quad \text{in } t \geq 0,$$

with

$$u_0 = 1, \quad u_{0\tau} = 0 \qquad \text{at } t = 0.$$

Hence

$$u_0 = R(T)\cos(\tau + \theta(T)).$$

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:

$$\begin{aligned} u_{1\tau\tau} + u_1 &= -u_{0\tau}(u_0^2 - 1) - 2u_{0\tau T} \qquad \text{in } t \geq 0 \\ &= 2R\theta_T \cos(\tau + \theta) + \left(2R_T + \frac{R^3}{4} - R\right)\sin(\tau + \theta) + \frac{R^3}{4}\sin 3(\tau + \theta). \end{aligned}$$

The initial conditions are

$$u_1 = 0, \qquad u_{1\tau} = -u_{0T} = -R_T \quad \text{at } t = 0.$$

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 u_1 we can now calculate it as

$$u_1 = -\frac{1}{32}R^3 \sin 3\tau + S(T)\sin(\tau + \phi(T)),$$

with new amplitude and phase functions S and ϕ. These will be determined by a secularity condition on u_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 u_1: we cannot avoid resonance in u_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

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

with solution

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

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. Some may be better handled by strained coordinates.

8 The WKBJ method (Additional notes)

(Liouville 1837, Green 1837, Horn 1899, Rayleigh 1912, Gans 1915, Jeffrey 1923, Wentzel 1926, Kramers 1926, Brillouin 1926, Langer 1931, Olver 1961, Meyer 1973.)

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

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

It has **oscillatory** solutions and is typical of many problems arising from wave propagation, with $\epsilon =$ 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

$$\epsilon^2 u'' + q(x)u = 0, \tag{12}$$

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

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

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

$$\frac{\partial^2 u}{\partial X^2} + 2\epsilon \frac{d^2 u}{\partial x \partial X} + \epsilon^2 \frac{\partial^2 u}{\partial X^2} + q(x)u = 0. \tag{13}$$

Expanding $u \sim u_0 + \epsilon u_1 + \cdots$ gives at leading order

$$\frac{\partial^2 u_0}{\partial X^2} + q(x)u = 0.$$

Hence

$$u_0 = A(x)\cos(q(x)^{1/2}X + \theta(x)),$$

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 (13) gives

$$\frac{\partial^2 u_1}{\partial X^2} + 2\frac{d^2 u_0}{\partial x \partial X} + +q(x)u_1 = 0,$$

i.e.

$$\begin{aligned} \frac{\partial^2 u_1}{\partial X^2} + q(x)u_1 &= 2\frac{d}{dx}\left(A(x)q(x)^{1/2}\sin(q(x)^{1/2}X + \theta(x))\right) \\ &= 2\frac{d}{dx}\left(Aq^{1/2}\right)\sin(q^{1/2}X + \theta) - 2Aq^{1/2}\left(X\frac{dq^{1/2}}{dx} + \frac{d\theta}{dx}\right)\cos(q^{1/2}X + \theta). \end{aligned}$$

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 can not 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 (12). Instead of using multiple scales, we assume an asymptotic expansion for u of the form

$$u \sim e^{i\phi(x)/\epsilon} A(x,\epsilon), \qquad \text{with } A(x,\epsilon) \sim \sum_{n=0}^{\infty} A_n(x)\epsilon^n.$$

This gives

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

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.

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 $u(\theta) = \sqrt{\sin\theta}P_n(\cos\theta)$ for $0 < \theta < \pi$ then the equation satisfied by u is

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

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

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

With the WKBJ ansatz $u = Ae^{i\phi/\epsilon}$ 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 $n \to \infty$.

Another example Consider Airy's equation

$$u'' - xu = 0, \qquad x < 0.$$

For large x we rescale $x = \epsilon^{-2/3}\bar{x}$ to give (on dropping the bars)

$$\epsilon^2 u'' - xu = 0, \qquad x < 0.$$

With the WKBJ ansatz $u = Ae^{i\phi/\epsilon}$ at leading order (ϵ^0) we have

$$(\phi')^2 = -x,$$

so

$$\phi' = \pm(-x)^{1/2}, \qquad \phi = \pm\frac{2}{3}(-x)^{3/2}.$$

At first order (ϵ^1) we have

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

so that
$$A_0 = \frac{\alpha_0}{(\phi')^{1/2}} = \frac{\alpha_0}{(-x)^{1/4}}.$$
Thus
$$u \sim \frac{\alpha_0}{(-x)^{1/4}} \exp\left(\frac{2ik^2}{3}(-x)^{3/2}\right) + \frac{\beta_0}{(-x)^{1/4}} \exp\left(-\frac{2ik^2}{3}(-x)^{3/2}\right) \qquad \text{for } x < 0.$$
Can also do this in $x > 0$. Either change ansatz to $Ae^{k\phi}$ or just allow ϕ to be complex to find
$$u \sim \frac{\lambda_0}{x^{1/4}} \exp\left(\frac{2k^2}{3}x^{3/2}\right) + \frac{\mu_0}{x^{1/4}} \exp\left(-\frac{2k^2}{3}x^{3/2}\right) \qquad \text{for } x > 0.$$
There is a connection formula relating the coefficients α_0 and β_0 to λ_0 and μ_0.

Eigenvalue problems
$$u'' + \lambda p(x)u = 0, \qquad u(0) = 0, \quad u(1) = 0.$$
Find the eigenvalues λ. For large λ put $\lambda = \epsilon^{-2}$, so that
$$\epsilon^2 u'' + p(x)u = 0, \qquad u(0) = 0, \quad u(1) = 0.$$
Then with the WKB approximation $u \sim Ae^{i\phi/\epsilon}$ we have
$$\phi' = \pm p^{1/2}, \qquad A_0 = \frac{\alpha_0}{(\phi')^{1/2}}.$$
If we fix $\phi = +\int p^{1/2}\,dx$, $A_0 = p^{-1/4}$, then the two independent solutions are given by
$$u_+ \sim A_0 e^{i\phi/\epsilon}, \qquad u_- \sim A_0 e^{-i\phi/\epsilon}.$$
For the general solution
$$u = \alpha u_+ + \beta u_-$$
to satisfy boundary conditions we need
$$\begin{aligned} \alpha A_0(0)e^{i\phi(0)/\epsilon} + \beta A_0(0)e^{-i\phi(0)/\epsilon} &\sim 0, \\ \alpha A_0(1)e^{i\phi(1)/\epsilon} + \beta A_0(1)e^{-i\phi(1)/\epsilon} &\sim 0. \end{aligned}$$
There is a nonzero solution for α and β if and only if the determinant is zero, i.e.
$$\begin{vmatrix} e^{i\phi(0)/\epsilon}A_0(0) & e^{-i\phi(0)/\epsilon}A_0(0) \\ e^{i\phi(1)/\epsilon}A_0(1) & e^{-i\phi(1)/\epsilon}A_0(1) \end{vmatrix} \sim A_0(0)A_0(1)\left(e^{i(\phi(0)-\phi(1))/\epsilon} - e^{i(-\phi(0)+\phi(1))/\epsilon}\right) \sim 0,$$
i.e.
$$e^{2i(\phi(0)-\phi(1))/\epsilon} \sim 1, \qquad \text{or } (\phi(0) - \phi(1)) = n\pi\epsilon.$$
Hence the eigenvalues are given approximately by
$$\epsilon_n = \frac{\int_0^1 \sqrt{p(x)}\,dx}{n\pi}.$$
or
$$\lambda_n = \left(\frac{n\pi}{\int_0^1 \sqrt{p(x)}\,dx}\right)^2.$$
Interpretation in terms of integer number of wavelengths.

Complex ϕ makes this interesting...

Turning points Example

$$-u'' + x^2 u = \lambda u, \qquad u \to 0 \text{ as } |x| \to \infty.$$

Harmonic oscillator. $\lambda =$ energy level. For large λ again let $\epsilon = 1/\lambda$ and rescale $x = \epsilon^{-1/2}\bar{x}$ to give

$$\epsilon^2 u'' + (1 - x^2)u = 0, \qquad u \to 0 \text{ as } |x| \to \infty.$$

Then we the standard WKBJ ansatz we find

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

Hence the solution is

$$u \sim \frac{\alpha_0}{(1 - x^2)^{1/4}} e^{\frac{i}{\epsilon}\int^x (1-x^2)^{1/2}\,dx} + \frac{\beta_0}{(1 - x^2)^{1/4}} e^{-\frac{i}{\epsilon}\int^x (1-x^2)^{1/2}\,dx}. \tag{14}$$

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

$$u \sim \frac{\alpha_1}{(x^2 - 1)^{1/4}} e^{-\frac{1}{\epsilon}\int_1^x (x^2-1)^{1/2}\,dx} + \frac{\beta_1}{(x^2 - 1)^{1/4}} e^{\frac{1}{\epsilon}\int_1^x (x^2-1)^{1/2}\,dx}. \tag{15}$$

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 $x = 1 + \epsilon^{1/2}\hat{x}$ to give

$$u'' - 2\hat{x}u = 0.$$

This is just the airy equation. We want a solution which matches with (15) as $\hat{x} \to \infty$. This solution is

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

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

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

while the inner limit of (15) is

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

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

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

while the inner limit of (14) is

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

Hence, matching the two expansions gives

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

and so

$$\alpha_0 e^{-\frac{i}{\epsilon}\phi(1)} - i\beta_0 e^{\frac{i}{\epsilon}\phi(1)} = 0.$$

Similarly, through a local analysis at $x = -1$ we find

$$\alpha_0 e^{-\frac{i}{\epsilon}\phi(-1)} + i\beta_0 e^{\frac{i}{\epsilon}\phi(-1)} = 0.$$

Hence, for a nonzero solution α_0, β_0 to exist we need

$$\begin{vmatrix} e^{-\frac{i}{\epsilon}\phi(1)} & -ie^{\frac{i}{\epsilon}\phi(1)} \\ e^{-\frac{i}{\epsilon}\phi(-1)} & ie^{\frac{i}{\epsilon}\phi(-1)} \end{vmatrix} = e^{-\frac{i}{\epsilon}(\phi(1)-\phi(-1))+i\pi/2} - e^{\frac{i}{\epsilon}(\phi(1)-\phi(-1))-i\pi/2} = 0.$$

Hence

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

so that

$$2(\phi(1)-\phi(-1)) = (2n+1)\pi\epsilon_n, \qquad n \in \mathbb{Z},$$

so that the eigenvalues λ_n are given by

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

In fact these are **exact** in this case. The exact solutions are

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

Lectures 11 and 12

Hamiltonian Systems Hamiltonian system is

$$\dot{x} = \frac{\partial H}{\partial y}, \qquad \dot{y} = -\frac{\partial H}{\partial x},$$

where $H(x,y)$ is the Hamiltonian and x and y are vectors. Then, along a phase path

$$\frac{dH}{dt} = \frac{\partial H}{\partial x}\dot{x} + \frac{\partial H}{\partial y}\dot{y} = 0,$$

so the level curves of H are the phase paths.

Variational methods

Many problems in applied mathematics arise as finding the least energy or least cost solutions of a variational problem.

Example Suppose A is a smooth compact area in $\mathbb{R}^2$ and $\phi(x,y)$ is to be chosen to minimize

$$J(\phi) = \int_A \left(\frac{\partial \phi}{\partial x} - y\right)^2 + \left(\frac{\partial \phi}{\partial y} + x\right)^2 dA.$$

[Torsion problem: ϕ = warping function for a beam of cross-section A, variational J_{min} = first torsion constant of the beam.]

The following are equivalent

(i) ϕ minimises J.

(ii) for all continuous functions ψ,

$$\int_A \left(\frac{\partial \phi}{\partial x} - y\right)\frac{\partial \psi}{\partial x} + \left(\frac{\partial \phi}{\partial y} + x\right)\frac{\partial \psi}{\partial y}\, dA = 0.$$

(iii)

$$\nabla^2\phi = 0 \quad \text{in } A, \qquad \text{and } \frac{\partial \phi}{\partial n} = y\frac{\partial x}{\partial n} - x\frac{\partial y}{\partial n} \quad \text{on } \partial A.$$

Suppose ϕ minimises J. Then let ψ be arbitrary and $t \in \mathbb{R}$.

$$J(\phi) \le J(\phi + t\psi) = J(\phi) + ct + O(t^2).$$

If $c \neq 0$ then considering t small and of opposite sign to c, we would have a contradiction. Hence $c = 0$. But c = coefficient of t in $J(\phi + t\psi) = 2\times$ integral stated in (ii). For reverse implication note that quadratic term is $t^2 \int |\nabla\psi|^2\, dA$.

For (iii) note that (ii) is

$$\int_A \left[\frac{\partial}{\partial x}\left(\psi\left(\frac{\partial \phi}{\partial x} - y\right)\right) + \frac{\partial}{\partial y}\left(\psi\left(\frac{\partial \phi}{\partial y} + x\right)\right) - \psi\nabla^2\phi\right] dA$$
$$= \int_{\partial A} \psi\left(\frac{\partial \phi}{\partial x} - y\right) dy - \psi\left(\frac{\partial \phi}{\partial y} + x\right) dx - \int_A \psi\nabla^2\phi\, dA.$$

Hence, at the optimum

$$\nabla^2\phi = 0 \text{ in } A \qquad \text{and } \left(\frac{\partial\phi}{\partial x} - y\right)\frac{dy}{ds} = \psi\left(\frac{\partial\phi}{\partial y} + x\right)\frac{dx}{ds}$$

which can be put in the form stated.

Optimal control of deterministic systems: The Pontryagin Maximum Principle

Form of problem : The state vector x of a system obeys

$$\dot{x} = f(x,t,u), \qquad x, f \in \mathbb{R}^n,$$

where u is a **control** which we are free to choose subject to $u(t) \in U_f(x(t),t)$, the set of **feasible** controls which may depend on x and t. We have to choose u in such a way as to maximise (or minimise) some "gain" function

$$\int_0^T h(x,t,u)\,dt, \qquad h(x,t,u) \in \mathbb{R}.$$

Boundary conditions: typically $x(0)$ given; T and $x(T)$ may both be given, or T fixed but $x(T)$ free, or $x(T)$ fixed but T free, etc. E.g. if $x(T)$ is given, T is free and $h = 1$ we have a minimum time control, $\int_0^T h\,dt = T =$ time to get between specified end states.

Procedure (Pontryagin Maximum Principle) Introduce a vector $p \in \mathbb{R}^n$ and define

$$H_0(x,t,u,p) = h(x,t,u) + pf(x,t,u) \equiv h(x,t,u) + \sum_i p_i f_i(x,t,u).$$

(The "pre-Hamiltonian".) Let $u_0(x,t,p)$ be the value of u in $U_f(x,t)$ that maximises $H_0(x,t,u,p)$, and let the **Hamiltonian**

$$H(x,t,p) = \max\{H_0(x,t,u,p) : u \in U_f(x,t)\} = H_0(x,t,u_0(x,t,p),p). \tag{16}$$

Then the optimal trajectory is found by solving

$$\begin{aligned} \dot{x} &= f(x,y,u_0(x,t,p)) = \frac{\partial H}{\partial p} \qquad \text{(if max is attained)}, \\ \dot{p} &= -\frac{\partial H}{\partial x} \end{aligned}$$

(system of $2n$ ode's), subject to

(i) the given value of $x(0)$,

(ii) given value of $x(T)$, or $p(T) = 0$ if $x(T)$ is free,

(iii) given value of T, or $H = 0$ at T if T is free.

Notes

(i) Can replace max with min throughout.

(ii) p is called the dual variable vector, adjoint, co-state.

Example Suppose $x \in \mathbb{R}$, $\ddot{x} = u$, and u is restricted by $-1 \leq u \leq 1$, and from some initial state you have to reach $x(T) = 0$ in minimum time T. Take the state vector to be

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x \\ \dot{x} \end{pmatrix},$$

so the differential equations are

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= u, \end{aligned}$$

i.e.

$$f = \begin{pmatrix} x_2 \\ u \end{pmatrix}.$$

We want to minimise

$$T = \int_0^T 1\, dt, \qquad \text{so take } h = 1,$$

and we have

$$(x_1, x_2) = \begin{cases} (x_0, \dot{x}_0) & \text{as } t = 0, \\ (0, 0) & \text{as } t = T. \end{cases}$$

Then

$$H_0 = h + pf = 1 + p_1 x_2 + p_2 u,$$

where p_1,p_2 are conjugate to x_1 and x_2 respectively. Hence

$$H = \min H_0 = 1 + p_1 x_2 - |p_2|, \qquad u_0 = -\mathrm{sign}(p_2).$$

Then

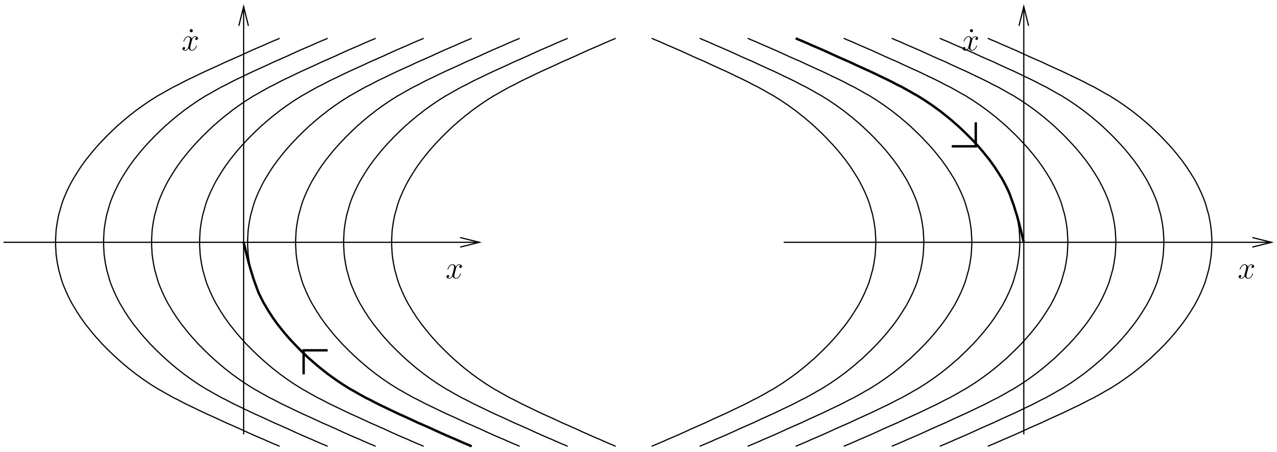
$$\begin{aligned} \dot{p}_1 &= -\frac{\partial H}{\partial x_1} = 0, \\ \dot{p}_2 &= -\frac{\partial H}{\partial x_2} = -p_1. \end{aligned}$$

We already see that

(a) the optimal trajectory will use $u = \pm 1$ only: “bang-bang” control.

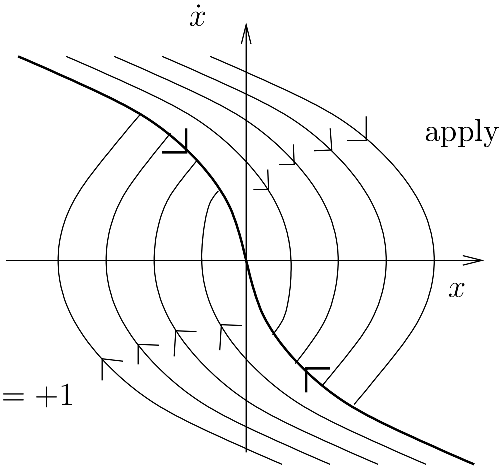
(b) u changes between ± 1 **at most once** on the optimal trajectory (since p_2 monotonic and $u_0 = -\mathrm{sign}(p_2)$).

What does this mean in the phase plane?



trajectories with $u = +1$ trajectories with $u = -1$

After the first switch, we must be on one of the dark paths. These are called the switching locus. Hence we must follow one family until we hit the switching locus and the other until $x = \dot{x} = 0$.



This is the "time optimal" control. (Can also find p_1 and p_2 etc by using the boundary conditions, but using (a) and (b) and the phase plane is easier.)

Note u does not vary continuously on the optimal trajectory. There is a discontinuity in $\ddot{x}$ where u changes sign. In some applications u may not be a real variable or vector at all, e.g. sound insulation: a board is to be built of layers of different materials subject to constraints on weight, thickness, cost, so as to minimise sound coming through. This is optimal control

$$\begin{aligned} t &\to x, \\ u &\to \text{what material used at } x, \\ \text{state} &\to \text{displacement/stress at } x \end{aligned}$$

(all assumed $\propto e^{i\omega t}$). Consequently we should prove the Pontryagin Maximum Principle (P.M.P.) by a method not assuming continuity in u.

"Proof" (Why the method usually works) First we prove the following
Lemma Suppose

$$g(y,z) = \max\{f(x,y,z) : x \in X(y)\} = f(x_0(y,z),y,z), \qquad x_0(y,z) \in X(y).$$

Then

$$\frac{\partial g}{\partial z}(y,z) = \frac{\partial f}{\partial z}(x_0(y,z),y,z).$$

Proof If f is differentiable

$$\frac{\partial g}{\partial z}(y,z) = \frac{\partial f}{\partial z}(x_0(y,z),y,z) + \frac{\partial f}{\partial x}(x_0(y,z),y,z)\frac{\partial x_0(y,z)}{\partial z}.$$

But x_0 defined to be maximum implies that

$$\frac{\partial f}{\partial x}(x_0(y,z),y,z) = 0.$$

But this inequality holds even if f is not differentiable in x. We have

$$f(x_0(y,z),y,z') \le g(y,z')$$

with equality at $z = z'$. Hence the z' derivatives are equal at $z = z'$, i.e. the required result. This depends on z being in the interior of the set over which f is defined and on f, g being differentiable in z. It does not depend on any differentiability in x or y. □

Now, to prove the Pontryagin maximum principle we have to show that there is a p defined on the optimal trajectory such that

(i) the optimal control u is the value maximising H;

(ii) $\dot{p} = -\dfrac{\partial H}{\partial x}$ on the optimal trajectory;

(ii) the boundary conditions hold.

Define

$$F(\xi,\tau) = \sup \int_\tau^T h(x,t,u)\,dt \qquad \text{starting from } x(\tau) = \xi,$$

(subject to $\dot{x} = f$, $u \in U_f(x,t)$ etc.). Then $F(x(0),0)$ is the required maximum. Assume f, h are continuous in (x,t) and F is C^1. We are going to show that $p = F_x$ (i.e. $p_i = \partial F/\partial x_i$) is the required function.

From the point (x,t) one possible control is to hold u constant (some value in $U_f(x,t)$) for small time δ, and then apply the optimal control from where you reach $(x_1, t+\delta)$. Here $x_1 = x + f(x,t,u)\delta + o(\delta)$, so $h(x,t,u)\delta + o(\delta) + F(x_1, t+\delta) \le F(x,t)$. Subtract $F(x,t)$, divide by δ and let $\delta \to 0$:

$$h(x,t,u) + F_x(x,t)f(x,t,u) + F_t \le 0, \tag{17}$$

for all $u \in U_f(x,t)$. (i.e.

$$h(x,t,u) + \sum_i \frac{\partial F}{\partial x_i} f_i + F_t \le 0.$$

If we integrate this inequality along any feasible trajectory (optimal or not) we have

$$\int_0^T h(x,t,u)\,dt + F(x(T),T) - F(x(0),0) \le 0,$$

i.e.

$$\int_0^T h(x,t,u)\,dt \le F(x(0),0).$$

(Remember $h \ge 0$.) This equation also clearly follows from the definition of F, since $F(x(0),0)$ is the supremum of the left-hand side over all possible controls. However, this definition of F means that there are controls that get arbitrarily close in this inequality. For simplicity, assume equality is **attained** for some optimal control. Then, for the optimal trajectory, equality holds in (17) for almost all t, and again for simplicity assume it holds everywhere. So (17) says

$$H_0(x,t,u,F_x(x,t)) + F_t(x,t) \le 0,$$

for all u, with equality for the optimal trajectory. Hence the optimal control does maximise H_0 for $p = F_x$ [(i) is satisfied], and we also see that the maximised value is

$$H(x,t,F_x(x,t)) = -F_t(x,t). \tag{18}$$

Now assume that H is C^1 and F is C^2. To derive the $\dot{p}$ equation, first note that by (16)

$$H_0(x,t,u_0(x,t,p),p') \le H(x,t,p'),$$

with equality at $p' = p$. Hence (by the previous Lemma) the p'-derivatives must agree at p, so

$$f(x,t,u_0(x,t,p)) = H_p(x,t,p). \tag{19}$$

Then the derivative of p along the optimal trajectory is

$$\dot{p} = \frac{d}{dt}\left(F_x(x,t)\right) = F_{xt}(x,t) + F_{xx} f(x,t,u_0(x,t,F_x)).$$

But by (18)

$$\begin{aligned} F_{xt}(x,t) &= -\frac{\partial}{\partial x}\left(H(x,t,F_x)\right) \\ &= -H_x(x,t,F_x) - H_p(x,t,F_x)F_{xx} \\ &= -H_x(x,t,F_x) - f(x,t,F_x)F_{xx} \end{aligned}$$

by (19). So

$$\dot{p} = -H_x(x,t,F_x)$$

as required [(ii) is satisfied].

For the boundary conditions, note that if $x(T)$ is free then $F(x,T) \equiv 0$ for all x, so $p = \partial F/\partial x = 0$ at T. If T is free but $x(T) = x_T$ is fixed, then $F(x_T,T) \equiv 0$, so

$$0 = \left.\frac{\partial F}{\partial t}\right|_T = -H(T)$$

by (18). So (iii) is satisfied.

Note If max is replaced by min, all inequalities are reversed and the "proof" is still OK.

Relationship of the P.M.P. to other areas of mathematics

Calculus of variations

Suppose we want to maximise

$$\int_0^T h(t, x, \dot{x})\, dt,$$

subject to

$$\int_0^T f(t, x, \dot{x})\, dt = c,$$

with T, $x(0)$, and $x(T)$ fixed. To formulate this problem as an optimal control problem, take $\dot{x} = u$ as the control and let $x_1 = x$, $x_2 = \int_0^t f(\tau, x, \dot{x})\, d\tau$. Then

$$\begin{aligned}
\dot{x}_1 &= \dot{x} = u, \\
\dot{x}_2 &= f(t, x_1, u),
\end{aligned}$$

and we need to get from

$$\begin{pmatrix} x(0) \\ 0 \end{pmatrix} \text{ to } \begin{pmatrix} x(T) \\ c \end{pmatrix}.$$

We have

$$H_0 = h(t, x_1, u) + p_1 u + p_2 f(t, x_1, u).$$

It is maximised by u such that

$$\frac{\partial h}{\partial u} + p_1 + p_2 \frac{\partial f}{\partial u} = 0.$$

The p_i obey

$$\begin{aligned}
\frac{dp_1}{dt} &= -\frac{\partial H}{\partial x_1} = -\frac{\partial h}{\partial x_1} - p_2 \frac{\partial f}{\partial x_1}, && (20) \\
\frac{dp_2}{dt} &= -\frac{\partial H}{\partial x_2} = 0. && (21)
\end{aligned}$$

So from (21) p_2 is constant, and if we call it $-\lambda$, then (20) is (on substituting for $p_1 = -\left(\frac{\partial h}{\partial u} - \lambda \frac{\partial f}{\partial u}\right)$ with $u = \dot{x}_1 = \dot{x}$)

$$\frac{d}{dt}\left(\frac{\partial h}{\partial \dot{x}} - \lambda \frac{\partial f}{\partial \dot{x}}\right) = \frac{\partial h}{\partial x} - \lambda \frac{\partial f}{\partial x},$$

the Euler-Lagrange equation with Lagrange multiplier λ.

Adjoints of linear ODEs

Suppose we want to know whether

$$y'' + \alpha y' + \beta y = b(x),$$

$$y(0) = a, \qquad y'(1) = c,$$

has a solution. One way to find out is to minimise

$$\int_0^1 \frac{1}{2}\left(y'' + \alpha y' + \beta y - b\right)^2 \, dx$$

subject to $y(0) = a$, $y'(1) = c$, and see if the minimum is 0. Regard y'' as the control u:

$$\frac{d}{dx}\begin{pmatrix} y \\ y' \end{pmatrix} = \begin{pmatrix} y' \\ u \end{pmatrix}.$$

We find

$$H_0 = \frac{1}{2}\left(u + \alpha y' + \beta y - b\right)^2 + py' + qu,$$

where p and q are conjugate variables to y and y' respectively. The boundary conditions are

$$y'(0) \text{ free } \Rightarrow q(0) = 0,$$

$$y(1) \text{ free } \Rightarrow p(1) = 0.$$

Now H_0 is minimised at

$$u + \alpha y' + \beta y - b + q = 0. \tag{22}$$

Hence

$$H = \frac{q^2}{2} + py' + q(q - \alpha y' - \beta y + b) = py' - \frac{q^2}{2} - q(\alpha y' + \beta y - b).$$

The conjugate variables p and q satisfy

$$\begin{aligned} p' &= -\frac{\partial H}{\partial y} = q\beta, \\ q' &= -\frac{\partial H}{\partial y'} = -p + \alpha q. \end{aligned}$$

So

$$q'' - (\alpha q)' = -p' = -\beta q,$$

i.e. q obeys the adjoint differential equation and boundary conditions.

Also, for any q_1 satisfying adjoint problem,

$$\int_0^1 (\alpha y' + \beta y + u) q_1 \, dx = \left[y' q_1 + y p_1\right] = 0,$$

by primary and adjoint boundary conditions. So (22)

$$b = (\alpha y' + \beta y + u) + q,$$

is actually the orthogonal decomposition of $b(x)$ as $q(x)$ solving the adjoint $+(y'' + \alpha y' + \beta y)$ orthogonal to the space of adjoint solutions.

Now,

primary problem has a solution $\Leftrightarrow q = 0$ (22) $\Leftrightarrow b$ is orthogonal to the space of adjoint solutions.

Singular control problems

The derivation of the PMP assumes

(1) the derivatives of H, f exist: Sometimes they don't, but often that is just at isolated points, or on a switching locus, and no great harm occurs.

(2) the optimum represented by f is attained. This can fail, even if u takes values in a closed set, etc.

E.g. suppose we have $\ddot{x} = u$, $-1 \le u \le 1$, and we want to control the state to $x(T) = 0$, $\dot{x}(T) = 0$ minimising

$$\int_0^T |u|\, dt.$$

Set up as before

$$H_0 = |u| + p_1 x_2 + p_2 u,$$

and now minimum H_0 occurs when

$$\begin{array}{rl} \text{if } p_2 > 1, & u_0 = -1, \\ \text{if } p_2 = 1, & u_0 \in [-1, 0] \\ \text{if } |p_2| < 1, & u_0 = 0, \\ \text{if } p_2 = -1, & u_0 \in [0, 1], \\ \text{if } p_2 < -1, & u_0 = 1. \end{array}$$

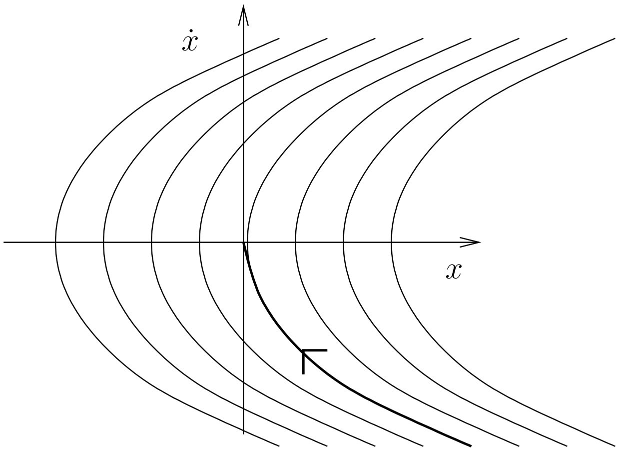
Hence

$$H = \min H_0 = \begin{cases} p_1 x_2 & \text{if } |p_2| \le 1, \\ p_1 x_2 + 1 - p_2 u & \text{if } p_2 > 1, \\ p_1 x_2 + 1 + p_2 u & \text{if } p_2 < -1. \end{cases}$$

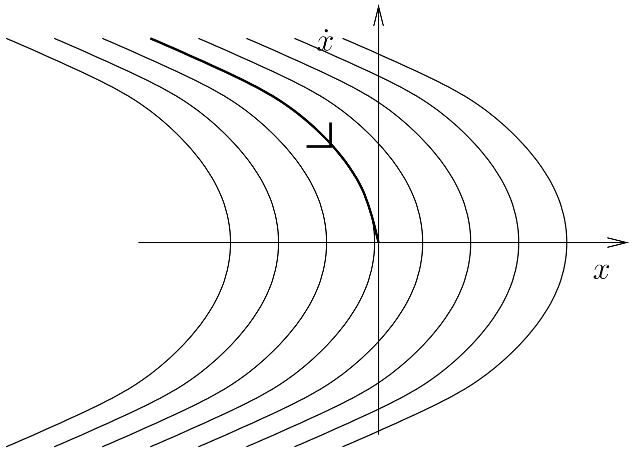
Then

$$\begin{aligned} \dot{p}_1 &= -\frac{\partial H}{\partial x_1} = 0, \\ \dot{p}_2 &= -\frac{\partial H}{\partial x_2} = -p_1. \end{aligned}$$

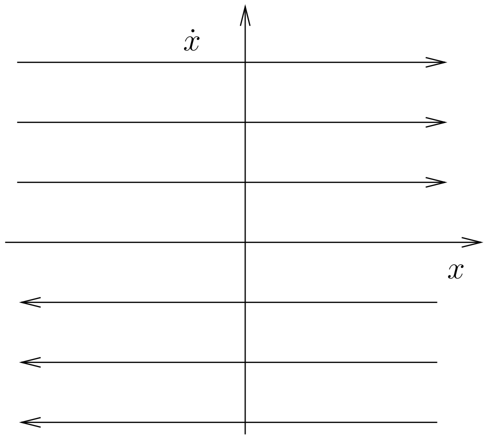
p_2 is still monotonic. The phase plane sketching is OK as far as



trajectories with $u = +1$



trajectories with $u = -1$



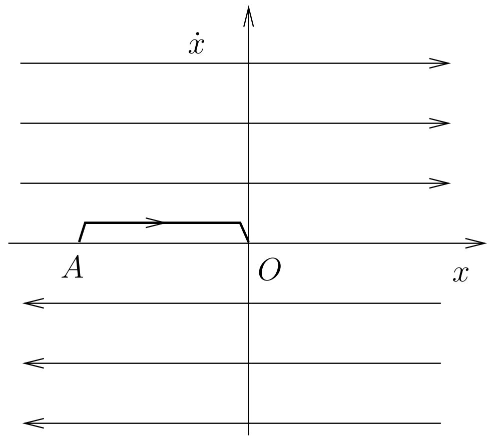
trajectories with $u = 0$

but a difficulty arises: From A we can reach O by: apply $u = +1$ for a small time ϵ; after that the velocity $\dot{x} = \epsilon$ and we switch off ($u = 0$) and wait. When x reaches $-\epsilon^2/2$ we apply $u = -1$ for time ϵ and reach $(0, 0)$. Then

$$\int_0^T |u| = 2\epsilon.$$

So it is arbitrarily small, but **cannot** be done with

$$\int_0^T |u| = 0.$$



Other kinds of singularity can also occur.

Lectures 13 and 14

Eigenfunction Expansions

Suppose $L[u]$ is a linear differential operator of order k with real coefficients and with coefficient of $d^k u/dx^k$ one sign, and with its boundary conditions forms a self-adjoint differential operator, i.e. if u and v are each real and obey the boundary conditions then

$$\langle v, L[u]\rangle = \langle L[v], u\rangle.$$

Note, this requires not just $L = L^*$ but the adjoint must have the same boundary conditions as the primary. E.g.

$$L[u] = (pu')' + qu \qquad \text{with } u(0) = u(1) = 0,$$

or

$$L[u] = u'''' \qquad \text{with } u(0) = u'(0) = u''(1) = u'''(1) = 0,$$

(clamped-free bending beam).

$$\int_0^1 vu''''\,dx = \left[vu''' - v'u'' + v''u' - v'''u\right]_0^1 + \int_0^1 v''''u\,dx$$

First two terms vanish at 1 because of u and at 0 because at v. Second two terms vanish at 1 because of v and at 0 because at u.

Then we say λ is an eigenvalue of L if

$$L[u] = \lambda u$$

has non-zero solutions u (obeying the boundary conditions). Such functions u are called eigenfunctions for λ.

Basic properties

(1) All the eigenvalues are real, and they form a sequence λ_n such that $|\lambda_n| \to \infty$ as $n \to \infty$.

(2) If $L[u_1] = \lambda_1 u_1$ and $L[u_2] = \lambda_2 u_2$ and $\lambda_1 \neq \lambda_2$ then

$$\langle u_1, u_2\rangle = 0 \qquad i.e. \int \overline{u_1} u_2\,dx = 0.$$

(3) For each eigenvalue λ, choose an orthonormal basis of the eigenspace, so we then have a sequence λ_1, $\lambda_2, \ldots$ and corresponding orthogonal u_1, $u_2, \ldots$ where if the λ-eigenspace has dimension k then λ occurs k times in the sequence (λ_i). If u is C^k and obeys the boundary conditions then

$$u = \sum c_n u_n \qquad \text{where } c_n = \langle u_n, u\rangle.$$

Proofs Some are elementary, some depend on Hilbert-Schmidt theory for self-adjoint kernels.

(1)
$$\bar{\lambda}\langle u, u\rangle = \langle \lambda u, u\rangle = \langle L[u], u\rangle = \langle u, L[u]\rangle = \langle u, \lambda u\rangle = \lambda\langle u, u\rangle.$$

Since $\langle u, u\rangle \neq 0$ we must have $\lambda = \bar{\lambda}$, i.e. λ is real.

(2)
$$\lambda_1\langle u_1, u_2\rangle = \langle \lambda_1 u_1, u_2\rangle = \langle L[u_1], u_2\rangle = \langle u_1, L[u_2]\rangle = \langle u_1, \lambda_2 u_2\rangle = \lambda_2\langle u_1, u_2\rangle.$$

Since $\lambda_1 \neq \lambda_2$ we must have $\langle u_1, u_2\rangle = 0$.

Since the eigenvalues λ are real, we can choose eigenfunctions u such that $L[u] = \lambda u$ to be real-valued. We usually do this, but we will write the formulae to be valid in general.

The **existence** of eigenfunctions and most other properties are shown by transforming the ode to an integral equation and using the Hilbert-Schmidt theory of integral equations with self-adjoint kernels.

Outline First, choose some c, real **not** an eigenvalue. Then the system $L[u] - cu = 0$, u obeying the boundary conditions has only the solution $u = 0$. So a Green's function G_c exists such that

$$\begin{aligned} L[u] - cu &= f, \\ u &\quad \text{obeys the boundary conditions} \end{aligned}$$

has the solution
$$u(x) = \int_0^1 G_c(x, y) f(y)\, dy$$
which we write as $u = G_c[f]$. Then show G_c is **symmetric** real kernel. Let f_1, f_2 be any continuous functions and $u_i = G_c[f_i]$, so u_i obey the boundary conditions and $L[u_i] - cu_i = f_i$. Then

$$\langle L[u_1], u_2\rangle = \langle u_1, L[u_2]\rangle,$$

i.e.
$$\langle cu_1 + f_1, u_2\rangle = \langle u_1, cu_2 + f_2\rangle.$$

After cancelling $c\langle u_1, u_2\rangle$ we are left with

$$\langle f_1, G_c[f_2]\rangle = \langle G_c[f_1], f_2\rangle,$$

i.e.
$$\int\!\!\int f_1(x) f_2(y) G_c(x, y)\, dy\, dx = \int\!\!\int f_2(x) f_1(y) G_c(x, y)\, dy\, dx = \int\!\!\int f_1(x) f_2(y) G_c(y, x)\, dx\, dy.$$

This holds for all continuous functions f_1, f_2 so $G_c(x, y) = G_c(y, x)$. Hence Hilbert-Schmidt theory applies to G_c and it has discrete characteristic values μ_n, $|\mu_n| \to \infty$, and the characteristic equation

$$u = \mu G_c[u],$$

has a finite-dimensional characteristic space. But $u = \mu G_c[u] = G_c[\mu u]$ if and only if u obeys the boundary conditions and $L[u] - cu = \mu u$, i.e. $L[u] = (\mu + c)u = \lambda u$. So the characteristic values μ_n of G_c are in 1 to 1 correspondence with eigenvalues $\lambda_n = \mu_n + c$ of the ode system, and the eigenspaces are finite-dimensional.

Construction of a complete system of eigenvalues and eigenfunctions Arrange the eigenvalues in increasing order of $|\lambda|$. If the λ-eigenspace has dimension d, repeat λ d times in the sequence. This produces a sequence $\lambda_1, \lambda_2, \ldots$. Choose an **orthonormal** basis of d eigenfunctions for the λ-eigenspace, and put them underneath the d entries λ:

$$\begin{array}{cccc} \lambda_1 & \lambda_2 & \lambda_3 & \cdots \\ u_1 & u_2 & u_3 & \cdots \end{array}$$

Then we have

$$\begin{aligned} L[u_i] &= \lambda_i u_i, \\ \langle u_i, u_j \rangle &= 0 \qquad \text{for } i \neq j, \\ \langle u_i, u_i \rangle &= ||u_i||^2 = \int_0^1 |u_i(x)|^2\,dx = 1 \qquad \text{(normalisation)}. \end{aligned}$$

Expansions Suppose u is C^k and obeys the boundary conditions. Then

$$u(x) = \sum_1^\infty a_n u_n(x),$$

where

$$a_n = \langle u_n, u \rangle,$$

and the series is uniformly convergent.

Proof is again by Hilbert-Schmidt theory. Let $f = L[u] - cu$, so f is continuous and $u = G_c[f]$. But Hilbert-Schmidt theorem then tells us there is a uniformly convergent expansion

$$G_c[f] = \sum_1^\infty \frac{\langle u_n, f \rangle u_n}{\mu_n},$$

where

$$\begin{aligned} \langle u_n, f \rangle &= \langle u_n, L[u] - cu \rangle \\ &= \langle L[u_n], u \rangle - c\langle u_n, u \rangle \\ &= (\lambda_n - c)\langle u_n, u \rangle \\ &= \mu_n \langle u_n, u \rangle. \end{aligned}$$

Note we have shown that

$$L[u] - cu = f, \qquad \text{with } u \text{ obeying the boundary conditions}$$

has solution

$$u(x) = \sum_1^\infty \frac{\langle u_n, f \rangle u_n(x)}{\lambda_n - c},$$

for c not an eigenvalue.

These expansions exist under weaker conditions on u. Generally if u is continuous then pointwise convergence a.e. If $u \in L^2$ then mean square convergence of eigenfunction expansion.

Archetypal example

$$L[u] = -u'', \qquad u(0) = u(1) = 0.$$

Then $\lambda_n = n^2\pi^2$, $n \geq 1$; $u_n = \sqrt{2}\sin n\pi x$. Fourier sine series.

Use of such expansions

(1) They do for the general self-adjoint problem what Fourier series do for the stretched string.

(2) Enable exact series solution of some problems

(3) Provide a way of **thinking** about the problem (modes etc.)

(4) Provide a **basis** in which related problems simplify e.g. when small damping is added, small changes to boundary conditions, etc.

Lectures 15 and 16

9 The Fourier Transform

Let $f \in L^1(\mathbb{R})$, i.e. $f : \mathbb{R} \to \mathbb{C}$ is measurable and

$$||f||_1 = \int_{-\infty}^{\infty} |f(x)|\, dx \qquad \text{is finite.}$$

Then its Fourier transform $\hat{f} : \mathbb{R} \to \mathbb{C}$ is defined by

$$\hat{f}(k) = \int_{-\infty}^{\infty} f(x) e^{ikx}\, dx.$$

Then

$$|\hat{f}(k)| \leq \int_{-\infty}^{\infty} |f(x)|\, dx,$$

so $||\hat{f}||_\infty \leq ||f||_1$. The Fourier transform is a **linear**, continuous map from L^1 to L^∞. Other basic properties:

(i) $\hat{f}(k)$ is continuous in k and $|\hat{f}(k)| \to 0$ as $k \to \infty$.

(ii) If $f_1(x) = f(x-a)$ then $\hat{f}_1(k) = e^{ika}\hat{f}(k)$.

(iii) If $f_2(x) = f(ax)$, $a \neq 0$, then $\hat{f}_2(k) = \dfrac{1}{|a|}\hat{f}\left(\dfrac{k}{a}\right)$.

(iv) If $f_3(x) = xf(x) \in L^1$ then $\hat{f}_3$ is differentiable and $\hat{f}_3(k) = -i\hat{f}'(k)$.

(v) If f is absolutely continuous and $f_4 = f' \in L^1$ then $\hat{f}_4(k) = -ik\hat{f}(k)$.

Proofs (i) Riemann-Lebesgue Lemma. (ii) and (iii) trivial.

For (iv) need to show differentiation under the integral sign is OK.

$$\frac{\hat{f}(k+h) - \hat{f}(k)}{h} = \int_{-\infty}^{\infty} f(x) e^{ikx} \left(\frac{e^{ihx}-1}{h}\right) dx.$$

As $h \to 0$, $\dfrac{e^{ihx}-1}{h} \to ix$, and $\left|\dfrac{e^{ihx}-1}{h}\right| \leq |x|$, so we have dominated convergence, to $i\hat{f}_3(k)$.

For (v) need to justify integration by parts.

$$\int_a^b f'(x) e^{ikx}\, dx = \Big[f(x)e^{ikx}\Big]_a^b - ik \int_a^b f(x) e^{ikx}\, dx.$$

Let $b = b_n \to +\infty$, $a = a_n \to -\infty$, through a sequence of values at which $f(b_n) \to 0$ [If no such sequence exists then $|f(x)| \geq \epsilon > 0$ for all large x and $f \notin L^1$.] Then $\hat{f}_4(k) = -ik\hat{f}(k)$.

Inversion If $\hat{f} \in L^1$ and f is continuous at x then

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f} e^{-ikx}\, dk.$$

Proof If f, $g \in L^1$ then $f(y)g(k)e^{iky} \in L^1(\mathbb{R})$ and taking its integral in either order we see

$$\int_{-\infty}^{\infty} f(y)\hat{g}(y)\,dy = \int_{-\infty}^{\infty} \hat{f}(k)g(k)\,dk. \tag{23}$$

Let

$$g(k) = g_\epsilon(k) = \frac{e^{-ikx-\epsilon k^2}}{2\pi} \in L^1 \quad \text{for } \epsilon > 0.$$

Let $\epsilon \to 0$. Then

$$\hat{g}_\epsilon(y) = \int_{-\infty}^{\infty} \frac{e^{-ikx-\epsilon k^2+iky}}{2\pi}\,dk = \frac{1}{2\pi}\sqrt{\frac{\pi}{\epsilon}}e^{-(x-y)^2/4\epsilon},$$

(pdf of Gaussian mean x variance 2ϵ).

$$\left[\text{ use } \int_{-\infty}^{\infty} e^{-\alpha k^2+\beta k}\,dk = \sqrt{\frac{\pi}{\alpha}}e^{\beta^2/4\alpha}, \qquad \text{for } \mathrm{Re}(\alpha) \geq 0, \quad \alpha \neq 0.\right]$$

As $\epsilon \to 0$ this tends to $\delta(x-y)$ as a distribution.

On LHS of (23)

$$\begin{aligned}
&\left|\int_{-\infty}^{\infty} f(y)\hat{g}_\epsilon(y)\,dy - f(x)\right| \\
\leq\ & \int_{|y-x|>\delta} |f(y)||\hat{g}_\epsilon(y)|\,dy + \left|\int_{|y-x|<\delta}(f(y)-f(x))\hat{g}_\epsilon(y)\,dy\right| + |f(x)|\left|1 - \int_{|y-x|<\delta}\hat{g}_\epsilon(y)\,dy\right| \\
\leq\ & ||f||_1\frac{e^{-\delta^2/4\epsilon}}{2\sqrt{\pi\epsilon}} + \sup\{|f(y)-f(x)| : |y-x| \leq \delta\} + |f(x)|\frac{2C\sqrt{\epsilon}e^{-\delta^2/4\epsilon}}{\sqrt{\pi}\delta} \\
<\ & \eta \qquad \text{for any prescribed } \eta
\end{aligned}$$

[Choose δ small enough for $\sup|f(y)-f(x)| : |y-x| \leq \delta$ to be small, then choose ϵ small enough.]

On RHS of (23)

$$\hat{f}(k)g_\epsilon(k) \to \frac{\hat{f}(k)e^{-ikx}}{2\pi},$$

and convergence is dominated by $|\hat{f}(k)| \in L^1$.

Importance of this: a large class of f are representable as linear combinations (integrals) of functions e^{-ikx}, which it is easy to differentiate.

Convolution If f, $g \in L^1$ then $|f(x-y)||g(y)| \in L^1(\mathbb{R}^2)$, so

$$h(x) = \int_{-\infty}^{\infty} f(x-y)g(y)\,dy$$

exists for almost all x; it is called the **convolution** of f and g, written $h = f * g$, and $h \in L^1$ with

$$\int_{-\infty}^{\infty} |h(x)|\,dx \leq ||f||_1||g||_1.$$

Also $f(x-y)g(y)e^{ikx} \in L^1(\mathbb{R}^2)$ and so

$$\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(x-y)g(y)e^{ikx}\,dx\right)dy = \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(x-y)g(y)e^{ikx}\,dy\right)dx,$$

i.e.

$$\int_{-\infty}^{\infty} \hat{f}(k)e^{iky}g(y)\,dy = \int_{-\infty}^{\infty} h(x)e^{ikx}\,dx,$$

i.e.

$$\hat{f}(k)\hat{g}(k) = \hat{h}(k) \qquad \text{for all } k.$$

Notes Conventions differ. If

$$f \notin L^1 \text{ but } \int_{-R}^{R} f(x)e^{ikx}\,dx \to \hat{f}(k) \text{ as } R \to \infty$$

we use this as the Fourier transform of f. Similarly we may use Cauchy principal value for finite singularities of f. Care needed in interpretation of results though.

Application 1 Solve

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{\partial^2 u}{\partial x^2} \text{ on } \mathbb{R}, \\ u(x,0) &= u_0(x) \qquad \text{given, } C^2 \text{ with } \int_{-\infty}^{\infty} |u_0|\,dx \text{ finite,} \end{aligned}$$

and the solution required to be Fourier transformable (no heat is supplied at infinity.) [heat conduction in an infinite bar; $u =$ temperature.] Then

$$\frac{\partial \hat{u}}{\partial t} = -k^2\hat{u},$$

so that

$$\hat{u}(k,t) = e^{-k^2 t}\hat{u}(k,0) = e^{-k^2 t}\hat{u_0}(k).$$

Therefore $u(x,t)$ is the convolution of u_0 with $E_t(x)$ where $E_t(x)$ is such that $\hat{E_t} = e^{-k^2 t}$. Inversion

$$\begin{aligned} E_t(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-k^2 t}e^{-ikx}\,dk \\ &= \frac{1}{2\pi}\sqrt{\frac{\pi}{t}}e^{-x^2/4t}. \end{aligned}$$

So

$$u(x,t) = \int_{-\infty}^{\infty} u_0(y)\frac{e^{-(x-y)^2/4t}}{2\sqrt{\pi t}}\,dy.$$

In fact E_t solves the problem

$$\begin{aligned} \frac{\partial E_t}{\partial t} &= \frac{\partial^2 E_t}{\partial x^2} \text{ on } \mathbb{R}, \\ E_t(x,0) &= \delta(x), \end{aligned}$$

and is the fundamental solution. Compare to Green's functions. More on this next term.

Application 2 Solve

$$\begin{aligned}\frac{\partial^2 u}{\partial t^2} &= -\frac{\partial^4 u}{\partial x^4} \text{ on } \mathbb{R},\\ \left.\begin{aligned} u(x,0) &= u_0(x)\\ \tfrac{\partial u}{\partial t}(x,0) &= v_0(x)\end{aligned}\right\} &\text{ both supposed } C^4 \text{ amd } L^1.\end{aligned}$$

[vibrations of an infinite beam.] Then

$$\frac{\partial^2 \hat{u}}{\partial t^2} = -k^4 \hat{u}.$$

So

$$\hat{u}(k,t) = \hat{u}_0(k)\cos(k^2 t) + \frac{\hat{v}_0(k)\sin(k^2 t)}{k^2}.$$

These can again be expressed as convolutions:

$$u = u_0 * f + v_0 * g,$$

with

$$\hat{f}(k,t) = \cos(k^2 t), \qquad \hat{g}(k,t) = \frac{\sin(k^2 t)}{k^2},$$

where $f = f(x,t)$, $g = g(x,t)$ and the convolutions are in x. Inverting

$$\begin{aligned} f(x,t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \cos(k^2 t) e^{-ikx}\, dk\\ &= \frac{1}{2\pi}\frac{1}{2}\left(\sqrt{\frac{\pi}{(-it)}}e^{(-ix)^2/4(-it)} + \sqrt{\frac{\pi}{(it)}}e^{(-ix)^2/4it}\right)\\ &= \frac{1}{4\pi}\sqrt{\frac{\pi}{t}}\left(e^{i\pi/4 - ix^2/4t} + \text{c.c.}\right)\\ &= \frac{1}{2\sqrt{\pi t}}\cos\left(\frac{\pi}{4} - \frac{x^2}{4t}\right).\end{aligned}$$

[NB this is not in L^1 but the note applies.] Now $\frac{\partial \hat{g}}{\partial t} = \hat{f}$. Hence

$$g(x,t) = \int_0^t \frac{1}{2\sqrt{\pi\tau}}\cos\left(\frac{\pi}{4} - \frac{x^2}{4\tau}\right)\, d\tau.$$

Write $\xi = x^2/4t$, $y = t/\tau$. Then

$$g(x,t) = \frac{\sqrt{t}}{2\sqrt{\pi}} B\left(\frac{x^2}{4t}\right),$$

where

$$B(\xi) = \int_1^\infty \frac{\cos(\pi/4 - \xi y)}{y^{3/2}}\, dy.$$

Fourier transform on L^2

$$L^2 = \{f : \mathbb{R} \to \mathbb{C}, \text{ measurable, with } \int_{-\infty}^{\infty} |f(x)|^2 \, dx \text{ finite}\}$$

(includes continuous functions with $\int_{-\infty}^{\infty} |f(x)|^2 \, dx$ finite.) If f and $\hat{f}$ are in $L^1 \cap L^2$ and the inversion theorem holds for f then

$$\begin{aligned}
\int_{-\infty}^{\infty} |f(x)|^2 \, dx &= \int_{-\infty}^{\infty} \overline{f(x)} f(x) \, dx \\
&= \int_{-\infty}^{\infty} \overline{f(x)} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(k) e^{-ikx} \, dk \right) dx \qquad \text{(inversion)} \\
&= \int_{-\infty}^{\infty} \frac{1}{2\pi} \hat{f}(k) \left(\int_{-\infty}^{\infty} \overline{f(x)} e^{-ikx} \, dx \right) dk \qquad \text{(Fubini)} \\
&= \int_{-\infty}^{\infty} \frac{1}{2\pi} \hat{f}(k) \overline{\hat{f}(x)} \, dk \qquad \text{(Fourier transform)}
\end{aligned}$$

Hence

$$\int_{-\infty}^{\infty} |f(x)|^2 \, dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{f}(k)|^2 \, dk \qquad \text{(Parseval's formula)}$$

Thinking of $f(x)$ as a linear combination of $\hat{f}(k)e^{-ikx}$ (integral) then Parseval's formula expresses "total energy of the function" = sum of the energy in each $\hat{f}(k)e^{-ikx}$. Like energy of a vibrating system = sum of energy in each mode.

In fact the Fourier transform can be defined on the whole of L^2 (not just the intersection with L^1), obeys Parsevals formula, and we still write

$$\hat{f}(k) = \int_{-\infty}^{\infty} f(x) e^{ikx} \, dx, \qquad f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(k) e^{-ikx} \, dk$$

even though these integrals are not absolutely convergent.

Relation with the Discrete Fourier Transform

Let $N \geq 1$ and let a_n be a sequence of N complex numbers, indexed by the integers n in $-N/2 < n \leq N/2$. The discrete Fourier transform is the sequence

$$b_r = \sum_n a_n e^{2\pi i r n/N} \qquad \text{for } -\frac{N}{2} < r \leq \frac{N}{2},$$

and the inversion formula is

$$a_n = \frac{1}{N} \sum_r b_r e^{-2\pi i r n/N} \qquad \text{for } -\frac{N}{2} < n \leq \frac{N}{2}.$$

[Sums over the stated range.] Let $\Delta x \Delta k = 2\pi/N$ and $x_n = n\Delta x$, $k_r = r\Delta k$. Then

$$x_n k_r = \frac{2\pi n r}{N}.$$

If $a_n = f(x_n)$ then

$$b_r = \sum_n f(x_n) e^{ix_n k_r} \approx \int_{-\frac{N\Delta x}{2}}^{\frac{N\Delta x}{2}} \frac{f(x) e^{ixk_r}}{\Delta x}\, dx \approx \frac{\hat{f}(k_r)}{\Delta x},$$

provided Δx is small enough to see the oscillations in f, and $N\Delta x$ is large enough to see the behaviour of f at infinity. Then discrete Fourier inversion is

$$\begin{aligned} f(x_n) = a_n &= \frac{1}{N} \sum_r b_r e^{-2\pi irn/N} \\ &\approx \frac{1}{N} \sum_r \frac{\hat{f}(k_r)}{\Delta x} e^{-ik_r x_n} \\ &\approx \frac{1}{N\Delta x \Delta k} \int_{-\frac{N\Delta k}{2}}^{\frac{N\Delta k}{2}} \hat{f}(k) e^{-ikx_n}\, dk \\ &\approx \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(k) e^{-ikx_n}\, dk. \end{aligned}$$

So same formula, but difficult to prove anything this way. Typically choose Δk and $\Delta x \propto 1/\sqrt{N}$ and N large. Can choose $\Delta k = \Delta x = \sqrt{(2\pi/N)}$ for symmetry, but this is not necessary. Better to use any knowledge of lengthscales etc to choose Δk, Δx.

Generalisation to higher dimensions

If $f \in L^1(\mathbb{R}^n)$, i.e.

$$\int \cdots \int |f(x_1, \cdots, x_n)|\, dx_1 \cdots dx_n \qquad \text{converges},$$

then we can take a Fourier transform in each x_i and get

$$\hat{f}(k_1, \cdots, k_n) = \int \cdots \int f(x_1, \cdots, x_n) e^{ik_1 x_1 + \cdots + ik_n x_n}\, dx_1 \cdots dx_n$$

which we write

$$\hat{f}(\mathbf{k}) = \int_{\mathbb{R}^n} f(\mathbf{x})\, e^{i\mathbf{k}\cdot\mathbf{x}}\, d\mathbf{x}.$$

The inversion formula is

$$f(\mathbf{x}) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \hat{f}(\mathbf{k})\, e^{-i\mathbf{k}\cdot\mathbf{x}}\, d\mathbf{k}.$$

If $f = f(\mathbf{x}, t)$ then we may use different conventions for time, typically

$$\begin{aligned} \hat{f}(\mathbf{k}, \omega) &= \int \int f(\mathbf{x}, t) e^{i\mathbf{k}\cdot\mathbf{x}} e^{-i\omega t}\, d\mathbf{x}\, dt, \\ f(\mathbf{x}, t) &= \frac{1}{(2\pi)^{n+1}} \int \int \hat{f}(\mathbf{k}, \omega) e^{i(\omega t - \mathbf{k}\cdot\mathbf{x})}\, d\mathbf{k}\, d\omega, \end{aligned}$$

represents f as a linear combination of sinusoidal waves. Terminology: If x is a space variable, transform variable k is a wavenumber (radians per metre); if t is a time variable, transform variable ω is a frequency (radians per second).