Further Mathematical Methods

Dominic Vella and Peter Grindrod
based on notes by Jon Chapman and Andrew Krause

April 16, 2024

Contents

Preliminaries

This is a short, sharp, background narrative that sets the analysis of differential and integral equations within a function space setting. This is not examinable but provides a context and big picture to the methods introduced in the course.

You may have already met function spaces within finite element methods, quantum mechanics, and elsewhere.

Inner Product Spaces

A Hilbert space is a complete normed space, X, that is equipped with an **inner product**, that is a real valued, bi-linear, function

$$\langle u, v\rangle : X \times X \to \mathbb{R},$$

such that the norm on X is given by

$$||u|| = \langle u, u\rangle^{1/2}.$$

"Completeness" merely requires that all Cauchy sequences have a convergent subsequence. So the rationals are incomplete (we can easily have a sequence of rationals converging to an irrational number which by definition is not in the space;) whereas as $\mathbb{R}$ is complete (indeed we can define the reals as *the completion* of the rationals).

We all know the n-dimensional Euclidean space, $\mathbb{R}^n$: then the inner product is the familiar vector "dot" product, or scalar product, $\langle u, v\rangle = u^T.v$.

Non-zero elements u and v in X are **orthogonal** iff $\langle u, v\rangle = 0$.

Note that if we work in the complex extension, $\mathbb{C}^n$, we must define

$$\langle u, v\rangle = u^T.\bar{v},$$

and so on, where $\bar{v}$ denotes the complex conjugate of v.

Let us think primarily of real spaces (since we often want to consider real valued functions as solutions to applied problems) yet introduce ideas about the spectra (eigenvalues and so forth) within their complexification (we know real matrices have complex valued spectra and eigenvectors, and so on) .

The space of pth power ($p > 1$), complex valued ,integrable function on $\Omega = [a, b]$ are those functions f for which the following integral exists:

$$\int_\Omega |f(t)|^p dt.$$

This space is called $L_p[\Omega]$ and is equipped with the norm

$$||f||_p \equiv \left(\int_\Omega |f(t)|^p dt \right)^{1/p}.$$

For $p = 2$ $L_p[\Omega]$ is a Hilbert space with the inner product

$$\langle u, v \rangle = \int_\Omega u(t)\bar{v}(t)dt.$$

For $p \neq 2$ but positive it is a complete normed pace (these are usually called Banach spaces), but it has no inner product, since if we try to define

$$\langle u, v \rangle = \int_\Omega u(t)^{p/2}\bar{v}(t)^{p/2}dt$$

this isn't linear in u and $\bar{v}$.

$L_2[\Omega]$ generalises to cases where Ω is a non trivial subset on $\mathbb{R}^m$ (where $m = 1, 2, 3, \dots$), usually with a piecewise smooth boundary $\delta\Omega$.

We only really need to know about $\mathbb{C}^n$ and $L_2[\Omega]$ for now.

We meet will $L_2[\Omega]$ when we consider integral operators and differential operators, made up of a differential form (in one or more dimensions) and imposing appropriate boundary conditions. We also meet it in finite element methods, within numerical analysis, and in the definition of weak solutions to PDEs.

The general theory of Hilbert spaces is available in chapter 3 of E. Kreyzig, "Introduction for Functional Analysis", which is very far from introductory! It is freely available in pdf form https://physics.bme.hu/sites/physics.bme.hu/files/users/BMETE15AF53_kov/Kreyszig%20-%20Introductory%20Functional%20Analysis%20with%20Applications%20(1).pdf. Enjoy.

Lots of this is peppered throughout the book by P. Grindrod, "Patterns and Waves", including the Fredholm alternative and applications and more general spectral theory of differential operations in some specific applied math settings: freely available, see https://www.researchgate.net/publication/351122823_Patterns_And_Waves_The_Theory_and_Application_of_Reaction-Diffusion_Equations

Linear Operators and Equations

A linear operator L acting on a Hilbert space, X, has both a domain, $D(L)$ and range in X, so that

$$L : D(L) \subset X \to X.$$

We define the adjoint operator L^* on X to be the linear operator such that

$$\langle Lu, v \rangle = \langle u, L^*v \rangle$$

for all u and v in X for which these expressions are finite and well-defined. If you are given X and L then you should always work directly from this definition to find L^* (see the examples below).

We are guaranteed that an adjoint operator L^* exists, since for all v fixed $\langle Lu, v\rangle$ is a linear functional in u and there is a theorem (the the Riesz representation theorem, see Kryezig) that states it can be written as $\langle u, w\rangle$ for some $w \in X$, so we can set $L^*v = w$, defining L^* point-wise. Moreover $\langle Lu, v\rangle$ is clearly linear in v. It is obvious that $(L^*)^* = L$. In any example, such as the differential or integral operators discussed below, these issues are very straightforward. Their general utility is often very useful, though, for modellers (not to get sucked into "the weeds").

Some operators are self adjoint ($L = L^*$) in which case all of the spectrum lives on the real axis.

In the case of equations, we might be interested in solutions of an inhomogeneous equation:

$$Lu = g,$$

with g given; where we also have the related homogeneous equation

$$Lu = 0.$$

Similarly we can consider an adjoint equation

$$L^*u = h,$$

with h given; and we also have the related homogeneous adjoint equation

$$L^*u = 0.$$

These equations (and their possible solutions) are the catalyst for Fredholm Alternative theory.

In fact, "$Lu = g$ has a solution iff g is orthogonal to the null space of L^*".

Say this to yourself every day. If L and thus L^* are invertible (so there is no null space in either case - that is, no non-zero solutions to the homogeneous equations above) then a unique solution of $Lu = g$ exists. *Alternatively* if the null spaces are non-trivial then this provides a *solvability condition.*

In the case of $X = \mathbb{R}^n$ then L is merely multiplication by an $n \times n$ real matrix A, and and L^* corresponds to multiplication by the real matrix A^T.

Suppose $X = L_2([a, b])$ the space of square integrable functions over the interval $[a.b]$. Consider the Fredholm operator

$$Lu(x) = u(x) - \lambda \int_a^b K(x,t)u(t)dt,$$

where, say, u is real valued.

To find the adjoint operator in this example we consider any real valued function v and the inner product

$$\langle Lu, v\rangle = \int_a^b v(x)\left(u(x) - \lambda \int_a^b K(x,t)u(t)dt\right)dx = \int_a^b uv - \lambda \int_a^b \int_a^b v(x)K(x,t)u(t)dtdx$$

swapping x and t in the last integraland reversing the order of integration we see

$$\langle Lu, v\rangle = \int_a^b uv - \lambda. \int_a^b \int_a^b v(t)K(t,x)u(x)dtdx = \int_a^b y(x)\left(v(x) - \lambda \int_a^b K(t,x)v(t)dt\right) dx.$$

Hence we take

$$L^* v(x) = v(x) - \lambda \int_a^b K(t,x)v(t)dt,$$

so that $\langle Lu, v\rangle = \langle u, L^* v\rangle$, as required in the above definition of the adjoint.

A linear differential operator defined over some spatial domain, Ω say, is a linear differential form defined over Ω (with either ordinary or partial derivatives depending upon the dimension of Ω), together with some suitable homogeneous boundary conditions to be imposed on the (assumed) piecewise smooth boundary, $\delta\Omega$.

Suppose, for example, $X = L_2([a,b])$ the space of square integrable functions over the interval $[a,b]$, with its intergral inner product.

Consider the linear operator

$$Lu(x) = u''(x) + A(x)u'(x) + B(x)u(x) \quad a < x < b, \quad u(a) = 0,\ u(b) = 0,$$

for given real valued functions A and B.

We have (using integration by parts, twice for the first term and once for the middle term),

$$\langle Lu, v\rangle = \int_a^b v(x)\left(u''(x) + A(x)u'(x) + B(x)u(x)\right) dx$$

$$= \left[v(x)u'(x)\right]_a^b - \left[v'(x)u(x)\right]_a^b + [v(x)A(x)u(x)]_a^b + \int_a^b u(x)(v''(x) - (A(x)v(x))' + B(x)v(x))dx.$$

Applying the boundary conditions on u, then $0 = [v'(x)u(x)]_a^b = [v(x)A(x)u(x)]_a^b$, and we have

$$\langle Lu, v\rangle = \left[v(x)u'(x)\right]_a^b + \int_a^b u(x)\left(v''(x) - (A(x)v(x))' + B(x)v(x)\right) dx.$$

But this must be true for a wide choice of u in X, and so we must impose

$$L^* v = v''(x) - (A(x)v(x))' + B(x)v(x) \quad a < x < b, \quad v(a) = 0,\ v(b) = 0.$$

Then $\langle Lu, v\rangle = \langle u, L^* v\rangle$ as required.

We might wish to solve an inhomogeneous equation (these often arise in asymptotic approaches to bifurcation analysis):

$$Lu(x) = u''(x) + A(x)u'(x) + B(x)u(x) = g(x) \quad a < x < b, \quad u(a) = 0,\ u(b) = 0.$$

Or we might have a PDE where $x \in \Omega$ a suitable subset of $\mathbb{R}^m$ together with suitable boundary conditions. Then all of the above considerations will be in play.

If Ω is in $\mathbb{R}^m$ then the divergence theorem (sometimes called Green's theorem) does the job of the integration by parts and the operator L and its adjoint L^* contain partial derivatives. Thus the corresponding equations will be PDEs.

For example, consider the operator

$$Lu = \Delta u(\mathbf{x}) + a(\mathbf{x})u \quad \mathbf{x} \in \Omega \subset \mathbb{R}^m, \quad \mathbf{n}.\nabla u = 0 \;\; \mathbf{x} \in \delta\Omega.$$

Then, using the divergence theorem, and the no-flux (Neumann) boundary condition on u, we have

$$\int_\Omega \nabla.((\nabla u)v - (\nabla v)u)\, d\mathbf{x} = \int_{\delta\Omega} \mathbf{n}.((\nabla u)v - (\nabla v)u)\, dA = -\int_{\delta\Omega} \mathbf{n}.(\nabla v)u\, dA.$$

So

$$\int_\Omega (\nabla.\nabla u + \alpha u)v\, d\mathbf{x} - \int_\Omega (\nabla.\nabla v + \alpha v)u\, d\mathbf{x} = -\int_{\delta\Omega} \mathbf{n}.(\nabla v)u\, dA.$$

Thus, if

$$L^*v = \Delta v(\mathbf{x}) + a(\mathbf{x})v \quad \mathbf{x} \in \Omega \subset \mathbb{R}^m, \quad \mathbf{n}.\nabla v = 0 \; \mathbf{x} \in \delta\Omega,$$

that is $L^* = L$, we have

$$\langle Lu, v\rangle - \langle u, L^*v\rangle = 0.$$

Hence L is self-adjoint.

If you ever have a problem (an equation) to solve for a function $u(x)$ with inhomogeneous boundary conditions, then it is often rather useful to choose a function, $u_0(x)$ say, that satisfies the boundary given conditions, so that $u(x) = u_0(x) + \tilde{u}(x)$, and then $\tilde{u}(x)$ satisfies homogeneous boundary conditions. The substitution may well change the inhomogeneous part of the full equation. But it makes the whole process much easier, since the resulting differential form has homogeneous boundary conditions: hence it is a linear operator. And so all of the above applies.

For differential operators there are some important subtleties we have glossed over, since the domain of 2nd order operators L, like the one above, is dense in $L_2([a, b])$: it contains those $u \in L_2([a, b])$ for which $u'' \in L_2([a, b])$; and which satisfy the homogeneous boundary conditions. Nevertheless that domain, $D(L)$, is itself a vector space and inherits the integral inner product and thus all notions of orthogonality.

For Fredholm integral operators we just need a well-behaved kernel, $K \in L_2$ with respect to both arguments, in order for L to be well defined.

1 Introduction

In the Supplementary Applied Mathematics course you were introduced to several approaches for understanding and solving boundary value problems given by ordinary differential equations. The first half of this course provides further tools and perspectives on related problems in integral equations and the theory of linear operators, filling in some of the theoretical gaps in boundary value problem theory. The second half of this course will then explore the calculus of variations, and optimal control theory. Throughout, by way of example, we will also introduce you to some aspects of perturbation theory which arise naturally in the examples. We will not have time to cover any of these topics in detail, but you should be able to come away from this course having a basic understanding of the ideas, and the ability to continue learning about any of these methods on your own. Good resources for the material on integral equations and the Fredholm Alternative can be found in Chapters 1, 3, and 4 of *Principles Of Applied Mathematics: Transformation And Approximation* by James Keener. Further material on the calculus of variations and optimal control can be found in *Calculus of Variations and Optimal Control Theory: A Concise Introduction* by Daniel Liberzon. You should review some aspects of linear algebra, particularly the rank-nullity Theorem, the kernel and nullspace of a matrix, and how these ideas relate to eigenvalues and diagonalization.

1.1 Integral Equations

There are many different formulations of integral equations, but the following are four common nontrivial examples.

Volterra non-homogeneous

$$y(x) = f(x) + \int_a^x K(x,t)\, y(t)\, \mathrm{d}t, \qquad x \in [a,b].$$

Volterra homogeneous

$$y(x) = \int_a^x K(x,t)\, y(t)\, \mathrm{d}t, \qquad x \in [a,b].$$

Fredholm non-homogeneous

$$y(x) = f(x) + \lambda \int_a^b K(x,t)\, y(t)\, \mathrm{d}t, \qquad x \in [a,b].$$

Fredholm homogeneous

$$y(x) = \lambda \int_a^b K(x,t)\, y(t)\, \mathrm{d}t, \qquad x \in [a,b].$$

The function $K(x,t)$ is the *kernel* of the integral equation.

A value of λ for which the homogeneous Fredholm equation has a solution which is not identically zero is called an *eigenvalue*, and the corresponding non-zero solution $y(x)$ is an *eigenfunction*. As with boundary value problems, one can develop a spectral theory of the eigenvalues and eigenfunctions in order to express the solution to the inhomogeneous problem.

We remark that throughout we *always* consider $\lambda \neq 0$. Note that some literature will put the λ on the $y(x)$ on the left-hand side, or will relate the eigenvalues written each way via $\lambda = 1/\mu$; these conventions are unimportant as long as $\lambda \neq 0$ and one is careful where the λ appears. We will also only consider real solutions in this course, but the theory generalizes easily to the complex case. Finally, unless otherwise stated, all functions considered will be continuous. Rather than work out or present the general theory, we will focus on Fredholm equations of a particular type. First, we relate these operators to familiar boundary value problems.

1.1.1 Relationship with differential equations

Example 1. Consider the differential equation

$$y''(x) + \lambda y(x) = g(x),$$

where $\lambda > 0$ is constant and g is contonuous on $[a,b]$. Integrating from a to $x \in [a,b]$ gives

$$y'(x) - y'(a) + \lambda \int_a^x y(t)\,\mathrm{d}t = \int_a^x g(t)\,\mathrm{d}t.$$

Integrating again gives

$$y(x) - y(a) - y'(a)(x-a) + \lambda \int_a^x \int_a^u y(t)\,\mathrm{d}t\,\mathrm{d}u = \int_a^x \int_a^u g(t)\,\mathrm{d}t\,\mathrm{d}u.$$

Switching the order of integration gives

$$y(x) - y(a) - y'(a)(x-a) + \lambda \int_a^x (x-t)y(t)\,\mathrm{d}t = \int_a^x (x-t)g(t)\,\mathrm{d}t. \tag{1}$$

Initial conditions Suppose $y(a)$ and $y'(a)$ are given. Then we have a Volterra non-homogeneous integral equation with

$$K(x,t) = \lambda(t-x), \qquad f(x) = y(a) + y'(a)(x-a) + \int_a^x (x-t)g(t)\,\mathrm{d}t.$$

Boundary conditions Suppose $y(a)$ and $y(b)$ are given. Then, putting $x=b$ in (1)

$$y(b) - y(a) - y'(a)(b-a) + \lambda \int_a^b (b-t)y(t)\,\mathrm{d}t = \int_a^b (b-t)g(t)\,\mathrm{d}t,$$

so that

$$y'(a) = \frac{1}{b-a}\left(y(b) - y(a) + \lambda \int_a^b (b-t)y(t)\,\mathrm{d}t - \int_a^b (b-t)g(t)\,\mathrm{d}t\right).$$

On substituting into (1) and simplifying this gives the non-homogeneous Fredholm equation

$$y(x) = f(x) + \lambda \int_a^b K(x,t)\, y(t)\, \mathrm{d}t$$

where

$$f(x) = y(a) + \frac{(x-a)}{(b-a)}(y(b) - y(a)) + \frac{1}{b-a}\int_a^x (x-b)(t-a)g(t)\,\mathrm{d}t + \frac{1}{b-a}\int_x^b (x-a)(t-b)g(t)\,\mathrm{d}t$$

$$K(x,t) = \begin{cases} \dfrac{(t-a)(b-x)}{b-a} & a \le t \le x \le b, \\ \dfrac{(x-a)(b-t)}{b-a} & a \le x \le t \le b. \end{cases}$$

■

NB: This kernel should look familiar to you from the study of Green's functions for boundary value problems.

2 Fredholm Alternative

The Fredholm Alternative is often considered one of the most important Theorems in applied mathematics (competing with Taylor's Theorem, among others). It gives a notion of 'solvability criterion' for a wide range of linear operators, and has numerous applications in differential and integral equations. Here we will present it first in the finite-dimensional case of linear algebra, followed by the cases of integral and differential equations. For integral equations of a particular type, the proof of this Theorem demonstrates how to construct solutions in a manner analogous to the eigenfunction expansions for boundary value problems.

2.1 Matrices

Consider a linear equation of the form,

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \tag{2}$$

where $\boldsymbol{A}$ is an $m \times n$ real matrix, $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{b} \in \mathbb{R}^m$. Let $\boldsymbol{a_i}$ be the ith column of $\boldsymbol{A}$. Then we have a solvability condition as follows.

Proposition 1. The Fredholm Alternative for general matrices
Either

1. The system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ has a solution $\boldsymbol{x}$;

or

2. The system $\boldsymbol{A}^T\boldsymbol{v} = \boldsymbol{0}$ has a solution $\boldsymbol{v}$ with $\boldsymbol{v}^T\boldsymbol{b} \neq 0$;

Thus $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ has a solution $\boldsymbol{x}$ if and only if $\boldsymbol{v}^T\boldsymbol{b} = 0$ for every $\boldsymbol{v}$ in $\mathbb{R}^m$ such that $\boldsymbol{A}^T\boldsymbol{v} = \boldsymbol{0}$.

Proof.

$$\begin{aligned}
\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \text{ has a solution } \boldsymbol{x} \quad &\Leftrightarrow \quad \boldsymbol{b} \text{ is a linear combination of the columns of } \boldsymbol{A} \\
&\Leftrightarrow \quad \boldsymbol{b} \in \operatorname{span}(\{\boldsymbol{a_1}, \cdots, \boldsymbol{a_n}\}) \\
&\Leftrightarrow \quad \operatorname{span}(\boldsymbol{b}) \subseteq \operatorname{span}(\boldsymbol{a_1}, \cdots, \boldsymbol{a_n}) \\
&\Leftrightarrow \quad \operatorname{span}(\boldsymbol{b})^\perp \supseteq \operatorname{span}(\boldsymbol{a_1}, \cdots, \boldsymbol{a_n})^\perp \\
&\Leftrightarrow \quad \text{every vector } \boldsymbol{v} \text{ with each } \boldsymbol{a_i}^T\boldsymbol{v} = 0 \text{ also has } \boldsymbol{b}^T\boldsymbol{v} = 0.
\end{aligned}$$

Note that $B^\perp$ denotes all vectors perpendicular to every vector in the set B. □

To relate this to integral equations we need to consider square matrices. Then we can write

Proposition 2. The Fredholm Alternative for square matrices
Either

1. The system $\boldsymbol{Ax} = \boldsymbol{b}$ has a unique solution $\boldsymbol{x}$;

or

2. There exist nonzero solutions to the system $\boldsymbol{A}^T\boldsymbol{v} = \boldsymbol{0}$. In this case $\boldsymbol{Ax} = \boldsymbol{b}$ has a solution if and only if $\boldsymbol{v}^T\boldsymbol{b} = 0$ for every $\boldsymbol{v}$ such that $\boldsymbol{A}^T\boldsymbol{v} = \boldsymbol{0}$. Such a solution (if it exists) is not unique, since any null vector of A may be added to it.

2.2 Integral equations

We will consider the non-homogeneous Fredholm equation

$$y(x) = f(x) + \lambda \int_a^b K(x,t)\, y(t)\, \mathrm{d}t, \qquad x \in [a,b].$$

2.2.1 A simple case

Recap (from the "Preliminaries" section): consider the operator:

$$Lu = u(x) - \lambda \int_a^b K(x,t)u(t)dx,$$

defined within the Hilbert space $L_2[a,b]$, thne we have

$$\langle Lu, v\rangle = \int_a^b \left(v(x)u(x) - \lambda v(x) \int_a^b K(x,t)u(t)dt \right) dx$$

$$\langle Lu, v\rangle = \int_a^b v(x)u(x)dx - \lambda \int_a^b v(x) \int_a^b K(x,t)u(t)dtdx$$

$$\langle Lu, v\rangle = \int_a^b v(x)u(x)dx - \lambda \int_a^b \int_a^b v(t)K(t,x)u(x)dtdx$$

$$\langle Lu, v\rangle = \int_a^b \left(u(x)v(x) - \lambda u(x) \int_a^b v(t)K(t,x)dt \right) dx$$

$$= \int_a^b u(x) \left(v(x) - \lambda \int_a^b v(t)K(t,x)dt \right) dx$$

This is $\langle u, L^* v\rangle$, by definition. So in this case we have the adjoint operator:

$$L^* v = v(x) - \lambda \int_a^b K((t,x)v(t)\, dt$$

Now consider simplest case for K: that of a separable product of two functions g and h. We have

$$y(z) = f(x) + \lambda \int_a^b g(x)h(t)y(t)dt \quad (Ly = f). \tag{3}$$

Here f, the inhomogeneity, and g and h and λ (real) in the operator L are all given.

This can be written as

$$y(x) = f(x) + \lambda g(x)X$$

where X is a (real) linear functional of y, given by

$$X = \int_a^b h(t)y(t)dt,$$

(which is the inner product of y and h).

Similarly we may have an adjoint equation (using the adjoint operator, L^*),

$$\tilde{y}(z) = f(x) + \lambda \int_a^b g(x)h(t)\tilde{y}(t)dt \quad (L^*\tilde{y} = f). \tag{4}$$

This can be written as

$$\tilde{y}(x) = f(x) + \lambda h(x)Y$$

where Y is a (real) linear functional of $\tilde{y}$, given by

$$Y = \int_a^b g(t)\tilde{y}(t)dt.$$

Both the equation and the adjoint equation have homogeneous forms :

$$y(x) = \lambda g(x)X, \quad (Ly = 0), \tag{5}$$

and

$$\tilde{y}(x) = \lambda h(x)Y, \quad (L^*\tilde{y} = 0). \tag{6}$$

Notice that $L^*\tilde{y} = 0$, which is equation (6), has a solution, that is, there is a non-trivial null space for the adjoint operator, iff it is in the form $\tilde{y} = ch(x)$, for any constant c, and then (6) implies we must have

$$ch(x) = \lambda h(x)Y = \lambda h(x)c \int_a^b g(t)h(t)dt.$$

So there is a non trivial null space for the adjoint operator (spanned by $h(x)$) iff

$$1 = \lambda \int_a^b g(t)h(t)dt.$$

If (λ, g, h) do not satisfy this then there is no non-trivial null space.

Similarly, the equation $Ly = 0$, which is equation (5), has a solution, that is, there is a non-trivial null space for the operator L, in the form $y = cg(x)$ iff

$$1 = \lambda \int_a^b g(t)h(t)dt.$$

Now, returning to (3). Does it have a solution?

Multiplying by $h(x)$ and integrating we get

$$X = \int_a^b h(x)y(x) = \int_a^b f(x)h(x)dx + \int_a^b \lambda g(x)h(x)dx\ X$$

where X is as above.

So

$$X\left(1 - \lambda \int_a^b g(x)h(x)dx\right) = \int_a^b f(x)h(x)dx. \tag{7}$$

If $(1 - \lambda \int_a^b g(x)h(x)dx) \neq 0$, we can solve this last for X and hence

$$y(x) = f(x) + \lambda g(x)\frac{\int_a^b f(x)h(x)dx.}{(1 - \lambda \int_a^b g(x)h(x)dx)}.$$

Finally if $(1 - \lambda \int_a^b g(x)h(x)dx) = 0$ then from (7) the RHS musts be zero so that

$$0 = \int_a^b f(x)h(x)dx.$$

This says $f(x)$ must be orthogonal to $h(x)$, which spans the null space of L^*, the adjoint operator.

In that case, X is a constant and so we have

$$y = f(x) + \text{constant } g(x),$$

meaning we can add on any multiple of the function $g(x)$ which lies in the null space of the operator L. If desired you can check this is a solution (since f in y makes no contribution to X since it must already be orthogonal to h, in this case

Notice that the condition $(1 - \lambda \int_a^b g(x)h(x)dx) = 0$ is exactly the condition where both L^* and L each have non-trivial null spaces spanned by $h(x)$ and $g(x)$ respectively.

We can consider the adjoint equation (4) in te same way. The inhomogeneity will have to be orthogonal the any non-trivial null space of the adjoint of L^*: but $(L^*)^* = L$, so it will have to be orthogonal the any non-trivial null space of L.

Thus we arrive at the following conclusion.

Proposition 3. Fredholm Alternative (Degenerate Integral Kernel)
Either

1. There are unique solutions to both (3) and (4);

or

2. The are nonzero solutions to both (3) to (4). In this case there exists a solution to (3) iff f is orthogonal to the null space of the adjoint, L^*, which is spanned by the $h(x)$, the solution of (6): so iff
$$\int_a^b f(x)h(x)\,dx = 0.$$
If a solution of (3) does exist in this case then it is non-unique, since any nonzero solution of (5), which is thus in the null space of L, can be added.

If the solvability condition $\int_a^b f(x)h(x)\,\mathrm{d}x = 0$ is met then the general solution of (3) is

$$y(x) = f(x) + cg(x),$$

for all $c \in \mathbb{R}$.

2.3 Integral equations: general case

We consider the Fredholm equation

$$y(x) = f(x) + \lambda \int_a^b K(x,t)y(t)\,\mathrm{d}t, \qquad x \in [a,b] \tag{F}$$

along with the adjoint and homogeneous equations

$$y(x) = f(x) + \lambda \int_a^b K(t,x)y(t)\,\mathrm{d}t, \qquad x \in [a,b] \tag{F^T}$$

$$y(x) = \lambda \int_a^b K(x,t)y(t)\,\mathrm{d}t, \qquad x \in [a,b] \tag{H}$$

$$y(x) = \lambda \int_a^b K(t,x)y(t)\,\mathrm{d}t, \qquad x \in [a,b] \tag{H^T}$$

where $f : [a,b] \to \mathbb{R}$ and the kernel $K : [a,b]^2 \to \mathbb{R}$ are continuous and λ is constant.

Theorem 1. The Fredholm Alternative For each fixed λ exactly one of the following two statements is true. Either

1. The equation (F) has a unique continuous solution. In particular if $f \equiv 0$ on $[a,b]$ then $y \equiv 0$ on $[a,b]$. In this case (F^T) also has a unique continuous solution.

or

2. The equation (H) has a finite maximal linearly independent set of, say, r continuous solutions $y_1, \ldots, y_r$ ($r > 0$). In this case (H^T) also has a maximal linearly independent set of r continuous solutions $z_1, \ldots, z_r$ and (F) has a solution if and only if the solvability conditions

$$\int_a^b f(x) z_k(x)\,\mathrm{d}x = 0, \qquad k = 1, \ldots, r,$$

are all satisfied. When they are, the complete solution to (F) is given by

$$y(x) = g(x) + \sum_{i=1}^{r} c_i y_i(x), \qquad x \in [a, b],$$

where $c_1, \ldots, c_r$ are arbitrary constants and $g : [a, b] \to \mathbb{R}$ is any continuous solution to (F).

We sketch the proof of the theorem for the degenerate kernel

$$K(x, t) = \sum_{j=1}^{n} g_j(x) h_j(t), \qquad x, t \in [a, b].$$

Proof. We may assume that each of the sets $\{g_1, g_2, \cdots, g_n\}$ and $\{h_1, h_2, \cdots, h_n\}$ are linearly independent (otherwise express each element in terms of a linearly independent subset). Then we have

$$y(x) = f(x) + \lambda \sum_{j=1}^{n} X_j g_j(x), \qquad \text{where} \qquad X_j = \int_a^b h_j(t) y(t)\,\mathrm{d}t, \tag{F$_1$}$$

$$y(x) = f(x) + \lambda \sum_{j=1}^{n} Y_j h_j(x), \qquad \text{where} \qquad Y_j = \int_a^b g_j(t) y(t)\,\mathrm{d}t, \tag{F$_1^T$}$$

$$y(x) = \lambda \sum_{j=1}^{n} X_j g_j(x), \tag{H$_1$}$$

$$y(x) = \lambda \sum_{j=1}^{n} Y_j h_j(x). \tag{H$_1^T$}$$

Multiply (F_1) by $h_i(x)$ and integrate over x to give

$$\mu X_i - \sum_{j=1}^{n} a_{ij} X_j = b_i,$$

where

$$\mu = \frac{1}{\lambda}, \qquad a_{ij} = \int_a^b g_j(x) h_i(x)\,\mathrm{d}x, \qquad b_i = \mu \int_a^b f(x) h_i(x)\,\mathrm{d}x.$$

We may write this as

$$(\mu I - A)\,\mathbf{X} = \mathbf{b} \tag{F$_2$}$$

where $\mathbf{X} = (X_j)$ and $\mathbf{b} = (b_j)$ are column vectors, $A = (a_{ij})$ is a matrix, and I is the identity matrix. Similarly (F_1^T) becomes

$$(\mu I - A)^T \mathbf{Y} = \left(\mu I - A^T\right) \mathbf{Y} = \mathbf{c} \tag{F_2^T}$$

where A^T is the transpose of A and $\mathbf{c} = (c_i)$ with

$$c_i = \mu \int_a^b f(x) g_i(x)\,\mathrm{d}x.$$

Similarly (H_1) and (H_1^T) become

$$(\mu I - A)\,\mathbf{X} = \mathbf{0} \tag{H_2}$$

$$(\mu I - A)^T\,\mathbf{Y} = \mathbf{0} \tag{H_2^T}$$

Now we are back in the case of linear algebra. So, suppose that there are no nontrivial solutions to (H_2), i.e., that μ is not an eigenvalue of A. Then, since $\mu I - A$ is nonsingular, there are unique solutions to (F_2) and (F_2^T), thus (1) holds.

On the other hand, suppose μ is an eigenvalue of A with eigenspace of dimension r spanned by eigenvectors $\mathbf{X}^k$, $k = 1, \ldots, r$. Then the corresponding eigenspace of A^T is also of dimension r and spanned by $\mathbf{Y}^k$, $k = 1, \ldots, r$, say. Then

$$y_k(x) = \lambda \sum_{j=1}^{n} X_j^k g_j(x), \tag{8}$$

$$z_k(x) = \lambda \sum_{j=1}^{n} Y_j^k h_j(x), \tag{9}$$

form a maximal set of linearly independent solutions of (H) and (H^T) respectively. We know (F_2) has a solution if and only if

$$\mathbf{b}^T \mathbf{Y}^k = 0, \qquad k = 1, \ldots, r,$$

which is, noting from (9) that $\mathbf{Y}^k$ corresponds to the solution $z_k(x)$ of (H^T),

$$\sum_{j=1}^{n} \left(\mu \int_a^b f(x) h_j(x)\,\mathrm{d}x\right) \left(\int_a^b g_j(t) z_k(t)\,\mathrm{d}t\right) = 0.$$

Rearranging, this is

$$\int_a^b \left(\int_a^b \left(\sum_{j=1}^{n} g_j(t) h_j(x)\right) z_k(t)\,\mathrm{d}t\right) f(x)\,\mathrm{d}x = 0,$$

i.e.

$$\int_a^b \left(\int_a^b K(t,x) z_k(t)\,\mathrm{d}t\right) f(x)\,\mathrm{d}x = 0,$$

which gives

$$\int_a^b z_k(x) f(x)\, \mathrm{d}x = 0,$$

since z_k is a solution of (H^T).

□

This method of proof can be used to solve (F) for degenerate kernels.

Example 2. Solve the integral equation

$$y(x) = f(x) + \lambda \int_0^{2\pi} \sin(x+t) y(t)\, \mathrm{d}t,$$

in the two cases

(a) $f(x) = 1$;

(b) $f(x) = x$.

The equation may be written

$$\begin{aligned} y(x) &= f(x) + \lambda \int_0^{2\pi} (\sin x \cos t + \cos x \sin t)\, y(t)\, \mathrm{d}t \\ &= f(x) + \lambda X_1 \sin x + \lambda X_2 \cos x \end{aligned} \tag{10}$$

where

$$X_1 = \int_0^{2\pi} y(t) \cos t\, \mathrm{d}t, \qquad X_2 = \int_0^{2\pi} y(t) \sin t\, \mathrm{d}t.$$

Note that it is self-adjoint. Multiplying (10) by $\cos x$ (and $\sin x$) and integrating with respect to x gives

$$X_1 - \lambda\pi X_2 = \int_0^{2\pi} f(x) \cos x\, \mathrm{d}x, \tag{11}$$

$$X_2 - \lambda\pi X_1 = \int_0^{2\pi} f(x) \sin x\, \mathrm{d}x \tag{12}$$

since

$$\int_0^{2\pi} \cos^2 x\, \mathrm{d}x = \int_0^{2\pi} \sin^2 x\, \mathrm{d}x = \pi, \qquad \int_0^{2\pi} \cos x \sin x\, \mathrm{d}x = 0.$$

This system is invertible if the determinant of the coefficient matrix

$$\begin{vmatrix} 1 & -\lambda\pi \\ -\lambda\pi & 1 \end{vmatrix} = 1 - \lambda^2\pi^2 \neq 0.$$

In this case the (unique) solution is

$$X_1 = \frac{1}{1-\lambda^2\pi^2}\int_0^{2\pi} f(x)\left(\cos x + \lambda\pi \sin x\right)\,\mathrm{d}x,$$
$$X_2 = \frac{1}{1-\lambda^2\pi^2}\int_0^{2\pi} f(x)\left(\sin x + \lambda\pi \cos x\right)\,\mathrm{d}x.$$

Since

$$\int_0^{2\pi} x\sin x\,\mathrm{d}x = -2\pi, \qquad \int_0^{2\pi} \cos x\,\mathrm{d}x = \int_0^{2\pi} \sin x\,\mathrm{d}x = \int_0^{2\pi} x\cos x\,\mathrm{d}x = 0,$$

in case (a) we have $X_1 = X_2 = 0$ and therefore

$$y(x) = 1, \qquad x \in [0, 2\pi],$$

while in case (b) we have

$$y(x) = x - \frac{2\pi\lambda}{1-\lambda^2\pi^2}\left(\lambda\pi\sin x + \cos x\right),$$

provided $\lambda^2\pi^2 \neq 1$.

When $\lambda = 1/\pi$ the homogeneous version of (11)-(12) has solutions $X_1 = X_2$, while when $\lambda = -1/\pi$ it has solutions $X_1 = -X_2$. Thus the homogeneous version of (10) has solutions

$$y(x) = c(\sin x + \cos x) \qquad \text{when } \lambda = 1/\pi,$$
$$y(x) = d(\sin x - \cos x) \qquad \text{when } \lambda = -1/\pi,$$

where c and d are constants. Thus in order for solutions to exist we have the solvability conditions

$$\int_0^{2\pi} f(x)(\sin x + \cos x)\,\mathrm{d}x = 0, \qquad \text{when } \lambda = 1/\pi,$$

and

$$\int_0^{2\pi} f(x)(\sin x - \cos x)\,\mathrm{d}x = 0, \qquad \text{when } \lambda = -1/\pi.$$

In case (a) both conditions are met. Since $y = 1$ is a particular solution when $\lambda = \pm 1/\pi$, the general solution in this case is

$$y(x) = 1 + c(\sin x + \cos x) \qquad \text{when } \lambda = 1/\pi,$$
$$y(x) = 1 + d(\sin x - \cos x) \qquad \text{when } \lambda = -1/\pi,$$

where c and d are arbitrary constants.

In case (b) neither condition is met and there are no solutions when $\lambda = \pm 1/\pi$.

■

2.4 Linear ordinary differential equations

We are going to describe solvability conditions for linear ODE's analogous to those for linear algebraic equations. We will do this for the 2nd order real scalar case, and give the general version later.

Consider a differential operator

$$L[u] = \frac{d^2u}{dx^2} + \alpha(x)\frac{du}{dx} + \beta(x)u = u'' + \alpha u' + \beta u,$$

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

Primary problem

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

with 2 linear homogeneous boundary conditions on u and u' at $x = 0$, 1.

Adjoint problem

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

with 2 linear homogeneous boundary conditions on v and v' at $x = 0$, 1.

The solvability result is that

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

The adjoint differential operator is

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

This obeys the fundamental identity

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

a bilinear form in the *boundary values* of u and v. This bilinear form is non-singular if $B(u,v) = 0$ for all v implies $u = 0$. Equivalently

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

and $B(u,v)$ is non-singular if the central matrix is non-singular. Then if $u(1)$, $u'(1)$, $u(0)$, $u'(0)$ obey 2 linear homogeneous equations (the primary boundary conditions) then we shall need 2 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.

2.4.1 Examples of adjoints

Example 3. 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 4. 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 5. 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

$$\int_0^1 (vL[u] - uL^*[v])\ dx = B(u,v) = 0.$$

We then multiply the primary problem by v and integrate, and multiply the adjoint problem by u and integrate, then subtract the first from the second to get,

$$\int_0^1 (vL[u] - uL^*[v])\ dx = \int_0^1 vb\, dx = 0.$$

The harder part (if adjoint condition holds then a solution exists) requires 2 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 sometimes called the Fredholm Alternative.

2.4.2 Applications

Example 6. Primary:

$$u'' = b(x), \qquad u'(0) = u'(1) = 0.$$

Adjoint:

$$v'' = 0, \qquad v'(0) = v'(1) = 0.$$

There is a nontrivial solution of the adjoint, namely

$$v = 1.$$

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

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

■

Example 7. Find the asymptotic solution of the equation

$$\ddot{x} + (1+\epsilon)x = \cos t, \qquad x(0) = x(2\pi), \quad \dot{x}(0) = \dot{x}(2\pi), \tag{13}$$

as $\epsilon \to 0$. Suppose we try a perturbation expansion

$$x(t) \sim x_0(t) + \epsilon x_1(t) + \cdots . \tag{14}$$

Substituting into the equation gives

$$(\ddot{x}_0 + \epsilon\ddot{x}_1 + \cdots) + (1+\epsilon)(x_0 + \epsilon x_1 + \cdots) = \cos t.$$

Expanding the brackets gives

$$\ddot{x}_0 + x_0 + \epsilon(\ddot{x}_1 + x_0 + x_1) + \cdots = \cos t.$$

Equating coefficients of powers of ϵ gives

$$\ddot{x}_0 + x_0 = \cos t, \qquad \ddot{x}_1 + x_0 + x_1 = 0, \qquad \cdots$$

Thus the leading-order problem is

$$\ddot{x}_0 + x_0 = \cos t, \qquad x_0(0) = x_0(2\pi), \quad \dot{x}_0(0) = \dot{x}_0(2\pi). \tag{15}$$

Note that this is self-adjoint. Is there a solution? The homogeneous version

$$\ddot{x}_0 + x_0 = 0, \qquad x_0(0) = x_0(2\pi), \quad \dot{x}_0(0) = \dot{x}_0(2\pi),$$

has solutions

$$x_0 = \cos t, \qquad \text{and} \qquad x_0 = \sin t.$$

Since

$$\int_0^{2\pi} \cos^2 t \, \mathrm{d}t \neq 0$$

we conclude that (15) has no solution. This does not mean that (13) has no solution: it means that our expansion (14) was incorrect. In (13) we are forcing with a term that is almost resonant (it is resonant when $\epsilon = 0$). Thus we expect the response to be large. Let us try instead

$$x(t) \sim \frac{1}{\epsilon} x_0(t) + x_1(t) + \cdots . \tag{16}$$

Substituting into the equation gives

$$\ddot{x}_0 + x_0 + \epsilon(\ddot{x}_1 + x_0 + x_1) + \cdots = \epsilon \cos t.$$

Equating coefficients of powers of ϵ now gives

$$\ddot{x}_0 + x_0 = 0, \qquad \ddot{x}_1 + x_0 + x_1 = \cos t, \qquad \cdots$$

This time the leading-order problem is

$$\ddot{x}_0 + x_0 = 0, \qquad x_0(0) = x_0(2\pi), \quad \dot{x}_0(0) = \dot{x}_0(2\pi), \tag{17}$$

with solution

$$x_0 = A \cos t + B \sin t,$$

where A and B are arbitrary constants, undetermined at this stage. To determine A and B we need to consider the equation at next order. This is

$$\ddot{x}_1 + x_0 + x_1 = \cos t,$$

or, using our expression for x_0,

$$\ddot{x}_1 + x_1 = (1 - A)\cos t - B\sin t, \qquad x_1(0) = x_1(2\pi), \quad \dot{x}_1(0) = \dot{x}_1(2\pi).$$

Now we use the Fredholm alternative again. There is a solution for x_1 if and only if the right-hand side is orthogonal to the solutions $\cos t$ and $\sin t$ of the homogeneous problem. Multiplying by $\cos t$ and integrating gives

$$1 - A = 0 \qquad \Rightarrow \quad A = 1.$$

Multiplying by $\sin t$ and integrating gives

$$B = 0.$$

Thus the leading order solution is

$$x \sim \frac{1}{\epsilon}\cos t.$$

In fact, this leading-order solution *is* the exact solution of the original problem, and we can continue looking at higher-order terms to see they are all zero. While this is a linear problem (and hence we could have solved the problem directly), this example illustrates a powerful combined use of asymptotic methods and solvability conditions which is widely applicable for nonlinear systems, especially oscillators. The solvability theory always tells us something important in the case that $\epsilon = 0$ – namely, that there is no solution, as we saw above in the regular perturbation expansion.

■

Example 8. Consider the equation

$$\epsilon\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = u - u^3 + \epsilon, \qquad u'(-\infty) = 0, \quad u'(\infty) = 0,$$

(with $u(-\infty)$ close to -1 and $u(\infty)$ close to 1). Consider an expansion

$$u \sim u_0 + \epsilon u_1 + \cdots .$$

Then, at leading order (equating coefficents of ϵ^0)

$$-\frac{\partial^2 u_0}{\partial x^2} = u_0 - u_0^3, \qquad u_0(-\infty) = -1, \quad u_0(\infty) = 1.$$

The solution is

$$u_0 = \tanh\left(\frac{x - x_0(t)}{\sqrt{2}}\right),$$

where $x_0(t)$ is arbitrary. This is the solution to the steady problem with $\epsilon = 0$, but it can be translated arbitrarily. To determine x_0 we need to go to the next order. At first order (equating coefficients of ϵ^1)

$$\frac{\partial u_0}{\partial t} - \frac{\partial^2 u_1}{\partial x^2} = u_1 - 3u_0^2 u_1 + 1, \qquad u_1'(-\infty) = 0, \quad u_1'(\infty) = 0.$$

Rearranging

$$-\frac{\partial^2 u_1}{\partial x^2} - u_1 + 3u_0^2 u_1 = 1 - \frac{\partial u_0}{\partial t} = 1 + \frac{\mathrm{d}x_0}{\mathrm{d}t}\frac{\partial u_0}{\partial x}. \tag{18}$$

Now, since

$$-\frac{\partial^2 u_0}{\partial x^2} - u_0 + u_0^3 = 0,$$

differentiating gives

$$-\frac{\partial^3 u_0}{\partial x^3} - \frac{\partial u_0}{\partial x} + 3u_0^2\frac{\partial u_0}{\partial x} = 0.$$

Thus $u_1 = \partial u_0/\partial x$ satisfies the homogeneous version of (18). Therefore, by the Fredholm Alternative, the right-hand side must be orthogonal to $\partial u_0/\partial x$:

$$0 = \int_{-\infty}^{\infty}\left(1 + \frac{\mathrm{d}x_0}{\mathrm{d}t}\frac{\partial u_0}{\partial x}\right)\frac{\partial u_0}{\partial x} = [u_0]_{-\infty}^{\infty} + \frac{\mathrm{d}x_0}{\mathrm{d}t}\int_{-\infty}^{\infty}\left(\frac{\partial u_0}{\partial x}\right)^2 \mathrm{d}x = 2 + \frac{\mathrm{d}x_0}{\mathrm{d}t}\int_{-\infty}^{\infty}\left(\frac{\partial u_0}{\partial x}\right)^2 \mathrm{d}x.$$

Thus

$$\frac{\mathrm{d}x_0}{\mathrm{d}t} = -\frac{2}{\int_{-\infty}^{\infty}(\partial u_0/\partial x)^2\, \mathrm{d}x}.$$

■

Example 9. Consider the equation for $y(x)$:

$$y'' + Ty + y^3 = 0, \qquad y(0) = 0, \quad y(1) = 0. \tag{19}$$

Let us first consider the linearised equation:

$$y'' + Ty = 0, \qquad y(0) = 0, \quad y(1) = 0.$$

This is an eigenvalue problem: there are solutions only for particular values of T. After imposing the boundary condition at $x = 0$ we have

$$y = \sin\sqrt{T}x.$$

The condition at $x = 1$ then implies

$$\sin\sqrt{T} = 0 \qquad \Rightarrow \qquad T = n^2\pi^2.$$

Then the solution is

$$y = A\sin n\pi x, \qquad T = n^2\pi^2,$$

where A is arbitrary.

Let us see how the nonlinear term affects this calculation when we are close to the bifurcation point $T = n^2\pi^2$. Returning to (19) let us pose an expansion

$$y = \epsilon y_0 + \epsilon^3 y_1 + \cdots, \qquad T = T_0 + \epsilon^2 T_1 + \cdots.$$

Then, equating coefficients of ϵ^1:

$$y_0'' + T_0 y_0 = 0, \qquad y_0(0) = 0, \quad y_0(1) = 0,$$

so that

$$T_0 = n^2\pi^2, \qquad y_0 = A \sin n\pi x,$$

as above. The coefficient A is determined by proceeding to next order. Equating coefficients of ϵ^3:

$$y_1'' + T_0 y_1 + T_1 y_0 + y_0^3 = 0, \qquad y_1(0) = 0, \quad y_1(1) = 0.$$

Substituting in for y_0, T_0 gives

$$y_1'' + n^2\pi^2 y_1 = -AT_1 \sin n\pi x - A^3 \sin^3 n\pi x. \tag{20}$$

Now the homogeneous equation is satisfied by $\sin n\pi x$. Thus in order for there to be a solution for y_1, by the Fredholm Alternative the right-hand side must be orthoginonal to $\sin n\pi x$. Thus

$$0 = \int_0^1 AT_1 \sin^2 n\pi x + A^3 \sin^4 n\pi x \, \mathrm{d}x = \frac{AT_1}{2} + \frac{3A^3}{8}.$$

since

$$\int_0^1 \sin^2 n\pi x \, \mathrm{d}x = \frac{1}{2}, \qquad \int_0^1 \sin^2 n\pi x \, \mathrm{d}x = \frac{3}{8}.$$

As an alternative to evaluating the integrals we observe

$$\begin{aligned} \sin^3 n\pi x &= \left(\frac{1}{2\mathrm{i}}\left(\mathrm{e}^{\mathrm{i}n\pi x} - \mathrm{e}^{-\mathrm{i}n\pi x}\right)\right)^3 = -\frac{1}{8\mathrm{i}}\left(\mathrm{e}^{3\mathrm{i}n\pi x} - 3\mathrm{e}^{\mathrm{i}n\pi x} + 3\mathrm{e}^{\mathrm{i}n\pi x} - \mathrm{e}^{-3\mathrm{i}n\pi x}\right)^3 \\ &= -\frac{1}{4}\left(\sin 3n\pi x - 3\sin n\pi x\right). \end{aligned}$$

Thus the right-hand side of (20) is

$$-AT_1 \sin n\pi x + A^3 \frac{1}{4}\left(\sin 3n\pi x - 3 \sin n\pi x\right).$$

We know that $\sin 3n\pi x$ is orthogonal to $\sin n\pi x$. Thus we need the coefficient of $\sin n\pi x$ to vanish, i.e.

$$-AT_1 - \frac{3A^3}{4} = 0.$$

Thus the amplitude is

$$A = \sqrt{-\frac{4T_1}{3}}.$$

Note that this means that the branch of solutions exists for $T_1 < 0$, i.e. for T slightly less than the critical value $n^2\pi^2$.

■

Example 9 revisited
Here we look at a solution for the previous example where we do not presume the asymptotic seris scalings that are given above: ($T = T_0 + O(\epsilon^2)$ and $u = \epsilon u + O(\epsilon^3)$, etc).

Instead, let us suppose we do not anticipate those, and instead we expand both u and T as full regular series in integer powers of (small) ϵ. This should be the default form mst analysts. Only with experience and anticipation might we guess the right terms to put into teh expansion. The key to a regular expansion is that the unknown function or coefficient next term is uniformly bounded so that the whole $O(\epsilon^{\text{next}}$ term is genuinely of a higher order in ϵ than the whole preceding terms and thus negligible when contrasted with it as $\epsilon \to 0$.

Consider consider the following nonlinear equation for $u(x)$:

$$u'' + Tu + u^3 = 0 \quad 0 < x < \pi, \quad u(0) = u(1) = 0.$$

Introduce a small parameter ϵ so that $T = T_0 + \epsilon T_1 + \epsilon^2 T_2 + \ldots$, where the T_i's are real constants, independent of ϵ, to be determined (succesively) and consider small solutions as an asymptotic expansion

$$u = \epsilon\phi_1(x) + \epsilon^2\phi_2(x) + \epsilon^3\phi_3(x) + \ldots,$$

where the $\phi_i(x)$'s are real functions independent of ϵ to be determined (successively).

Substituting in and equating powers of ϵ we find to order ϵ,

$$\phi_1'' + T_0\phi_1 = 0 \quad \phi_1(0) = \phi_1(1) = 0.$$

So we must have $T_0 = n^2\pi^2$ and $\phi_1 = A \sin n\pi x$ for some real constant $A \neq 0$, and thus is non trivial for A non zero.

This equation can be written $L\phi_1 = 0$, where L is self-adjoint.

To order ϵ^2 we have

$$\phi_2'' + T_0\phi_2 = -T_1\phi_0 \quad \phi_2(0) = \phi_2(1) = 0.$$

The Fredholm Alternative says that this has a solution if and only if the RHS is orthogonal to $\sin n\pi x \propto \phi_1$, which spans the null space of L (which is self adjoint). Thus we must take $T_1 = 0$. Then wlog we can take $\phi_2 = 0$.

To order ϵ^3 we have

$$\phi_3'' + n^2\phi_3 = -T_2 A \sin n\pi x - A^3 \sin^3 n\pi x \quad \phi_3(0) = \phi_3(1) = 0.$$

The Fredholm Alternative says that this has a solution if and only if the RHS is orthogonal to $\sin n\pi x \propto \phi_1$ (which spans the null space of the adjoint operator): hence we have

$$-T_2 \int_0^\pi \sin^2 n\pi x dx = A^2 \int_0^\pi \sin^4 n\pi x dx.$$

So $A = \pm\sqrt{-T_2 \int_0^\pi \sin^2 n\pi x dx / \int_0^\pi \sin^4 n\pi x dx} = \pm\sqrt{-4T_2/3}$, as required. A is real iff $T_2 \leq 0$.

WLOG we can take $T_2 = -1 < 0$ and absorb $|T_2|$ into the definition of ϵ^2.

Hence we have

$$T = n^2\pi^2 - \epsilon^2 + O(\epsilon^3)$$

and

$$u(x) = \epsilon A \sin n\pi x + O(\epsilon^3)$$

with $A = 2/\sqrt{3}$ given as above: there is a pitchfork bifurcation at $T = n^2\pi^2$.

Of course what is T_2 here is written as T_1 in the above earlier version of this solution ; and u_3 here is y_1 in that version .

2.4.3 Generalisation

Suppose $\boldsymbol{u}$ is a vector of complex-valued functions, obeying a higher-order primary problem

Primary

$$L[\boldsymbol{u}] = \boldsymbol{b}(x) \qquad \text{on } 0 \le x \le 1,$$

with primary boundary conditions on $\boldsymbol{u}$ at $x = 0,\, 1$, where the primary differential form on the domin is

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

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

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

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

Then

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

Adjoint

$$L^*[\boldsymbol{v}] = 0,$$

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

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

The fundamental identity is

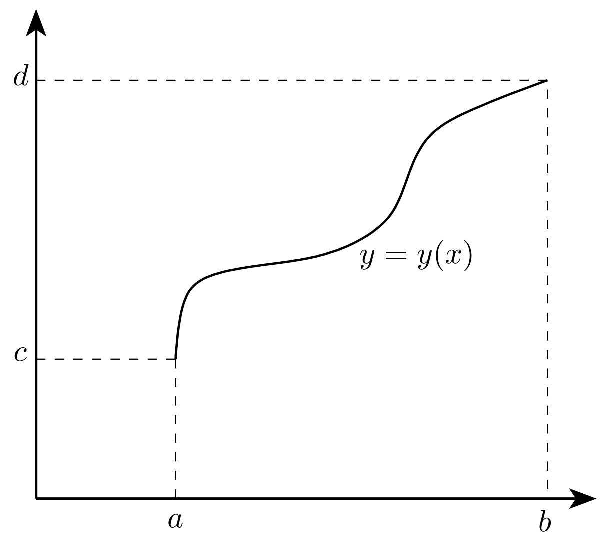
$$\begin{aligned}
\langle \boldsymbol{v}, L[\boldsymbol{u}]\rangle - \langle L^*[\boldsymbol{v}], \boldsymbol{u}\rangle &= \int_0^1 \sum_r \Big(\boldsymbol{v}^* \boldsymbol{A_r} \boldsymbol{u}^{(r)} - (-1)^r (\boldsymbol{v}^* \boldsymbol{A_r^*})^{(r)} \boldsymbol{u}\Big)\, dx \\
&= \left[\sum_r \boldsymbol{v}^*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

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 Calculus of variations

3.1 The main idea

Start with a **simple example**: consider a plane curve joining two points (a, c) and (b, d) and given by the smooth graph $y = y(x)$.



NB this disallows some slopes.

Define the **functional**

$$J[y] = \int_a^b (y'(x))^2 \, \mathrm{d}x.$$

NB $J : V \to \mathbb{R}$, where V is a suitable **function space**, e.g. the set $C^2[a, b]$ of twice continuously differentiable functions $y(x)$ defined on $[a, b]$, satisfying $y(a) = c$ and $y(b) = d$ [we won't dwell much on the strict conditions on $y(x)$].

Now we ask: which function $y(x) \in V$ **minimises** the functional $J[y]$?

To answer this, let $y(x)$ be the desired *extremal* function which minimises $J[y]$. Then any admissable perturbation about $y(x)$ should *increase* J. So consider $J[y + \epsilon\eta]$, where $\eta \in C^2[a, b]$ with $\eta(a) = \eta(b) = 0$. Now

$$\begin{aligned} J[y + \epsilon\eta] &= \int_a^b (y'(x) + \epsilon\eta'(x))^2 \, \mathrm{d}x \\ &= J[y] + 2\epsilon \int_a^b y'(x)\eta'(x) \, \mathrm{d}x + \epsilon^2 \int_a^b (\eta'(x))^2 \, \mathrm{d}x. \end{aligned}$$

We want this to have a **minimum** when $\epsilon = 0$, and a necessary condition is

$$\int_a^b y'(x)\eta'(x) \, \mathrm{d}x = 0.$$

[Then the coefficient of $\epsilon^2 \geq 0$ so it *is* a minimum not a maximum.]

Now integrate by parts to give

$$\underbrace{\big[y'(x)\eta(x)\big]_a^b}_{=0 \text{ since } \eta(a)=\eta(b)=0} - \int_a^b \eta(x)y''(x)\,\mathrm{d}x = 0.$$

We deduce that

$$\int_a^b \eta(x)y''(x)\,\mathrm{d}x = 0$$

for all $\eta \in C^2[a,b]$ with $\eta(a) = \eta(b) = 0$.

Fundamental Lemma of Calculus of Variations (FLCV)
If

$$\int_a^b \eta(x)\phi(x)\,\mathrm{d}x = 0 \qquad \forall \eta \in C^2[a,b] \text{ with } \eta(a) = \eta(b) = 0,$$

and ϕ is **continuous**, then

$$\phi(x) \equiv 0 \qquad \text{on } [a,b].$$

Hence we find that the function $y(x)$ that minimises $J[y]$ satisfies

$$y''(x) \equiv 0,$$

i.e.

$$y = Ax + B = c + \frac{(d-c)}{(b-a)}(x-a),$$

which is a **straight line** from (a,c) to (b,d).

Possible motivations
(i) 1-d flow of electricity through a semiconductor. $\phi(x)$ = electric potential (voltage).

The energy dissipated (as heat) in the medium is given by

$$J[\phi] = \int_0^l \sigma(x)(\phi'(x))^2\,\mathrm{d}x,$$

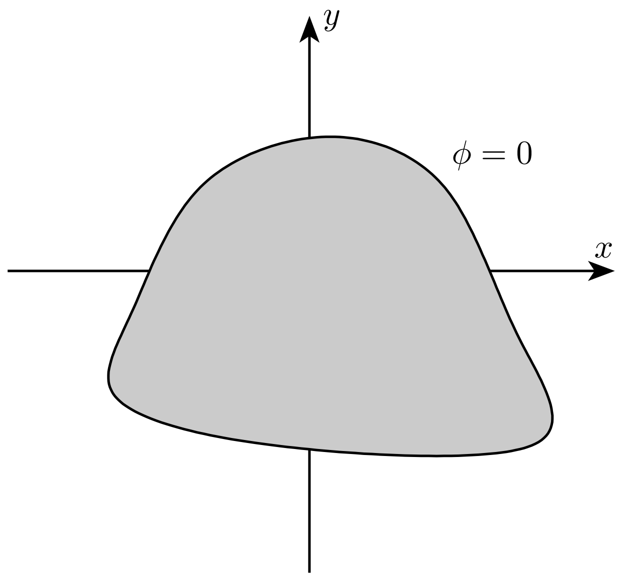
where $\sigma(x)$ is the ocnductivity of the medium. So dissipation is minimised when ϕ satisfies

$$\frac{\mathrm{d}}{\mathrm{d}x}(\sigma(x)\phi'(x)) = 0.$$

(ii) Drive from A to B in a given time T. Let your position at time t be $x(t)$. Then $x(0) = a$, $x(T) = b$. Suppose there is a frictional resistance $k\dot{x}(t)$. Then the work done against friction during the journey is

$$J[x] = \int_0^T k\dot{x}(t)^2\,\mathrm{d}t.$$

This suggests that driving at **constant speed** ($\ddot{x} = 0$) is the most efficient.



3.2 Generalization: A class of problems

This simple example falls into a class of problems: to minimise or maximise a functional

$$J[y] = \int_a^b F(x, y(x), y'(x))\,\mathrm{d}x$$

(where $F(x, y, y')$ is given) over all $y \in C^2[a, b]$ satisfying $y(a) = c$, $y(b) = d$.

Let $y(x)$ be an extremal function and perturb:

$$J[y + \epsilon\eta] = \int_a^b F(x, y + \epsilon\eta, y' + \epsilon\eta')\,\mathrm{d}x,$$

where $\eta \in C^2[a, b]$ with $\eta(a) = \eta(b) = 0$. Expand using Taylor's theorem:

$$J[y + \epsilon\eta] \quad = \quad J[y] + \epsilon \int_a^b \left(\eta \frac{\partial F}{\partial y}(x, y, y') + \eta' \frac{\partial F}{\partial y'}(x, y, y') \right) \mathrm{d}x + O(\epsilon^2).$$

NB here we treat x, y and y' as independent variables.

At an extremal we must have

$$\int_a^b \left(\eta \frac{\partial F}{\partial y} + \eta' \frac{\partial F}{\partial y'} \right) \mathrm{d}x = 0.$$

Integrate by parts:

$$\int_a^b \eta \left(\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \left(\frac{\partial F}{\partial y'} \right) \right) \mathrm{d}x + \underbrace{\left[\eta \frac{\partial F}{\partial y'} \right]_a^b}_{=0 \text{ since } \eta(a)=\eta(b)=0} = 0.$$

Since this is true for all $\eta \in C^2[a,b]$ with $\eta(a) = \eta(b) = 0$ by the FLCV we have Euler's equation (basic equation of Calculus of Variations):

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right) - \frac{\partial F}{\partial y} = 0}$$

NB $\mathrm{d}/\mathrm{d}x$ not $\partial/\partial x$.

Examples

(i) In our previous example $F(x, y, y') = (y')^2$. This gives

$$\frac{\mathrm{d}}{\mathrm{d}x}(2y') = 0, \qquad \text{i.e.} \qquad y'' = 0.$$

(ii) Curve of minimum length joining (a, c) to (b, d). Length

$$J[y] = \int_a^b \sqrt{1 + (y')^2}\,\mathrm{d}x,$$

subject to $y(a) = c$, $y(b) = d$. Then

$$F = \sqrt{1 + (y')^2}, \qquad \frac{\partial F}{\partial y} = 0, \qquad \frac{\partial F}{\partial y'} = \frac{y'}{\sqrt{1 + (y')^2}}.$$

So Euler's equation is

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right) = \frac{y''}{(1 + (y')^2)^{3/2}} = 0.$$

Thus $y'' = 0$ so $y = Ax + B$. Linear (again). Thus

$$y(x) = c + \frac{(d - c)}{(b - a)}(x - a),$$

a straight line, as expected.

3.3 Extensions

3.3.1 Natural boundary conditions

This time let

$$J[y] = \int_a^b F(x, y, y')\,\mathrm{d}x$$

where $y(a) = c$ but $y(b)$ is NOT prescribed. Again let $y(x)$ be an extremal of $J[y]$ and consider $y + \epsilon\eta$, where $\eta(a) = 0$ but $\eta(b)$ is *arbitrary*. Then

$$\begin{aligned} J[y + \epsilon\eta] &= \int_a^b F(x, y, y')\,\mathrm{d}x \\ &\sim J[y] + \epsilon \int_a^b \left(\eta \frac{\partial F}{\partial y} + \eta' \frac{\partial F}{\partial y'}\right)\mathrm{d}x + O(\epsilon^2). \end{aligned}$$

At an extremal, we must have

$$\int_a^b \left(\eta \frac{\partial F}{\partial y} + \eta' \frac{\partial F}{\partial y'} \right) \mathrm{d}x = 0$$

$$\Rightarrow \qquad \int_a^b \eta \left(\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \left(\frac{\partial F}{\partial y'} \right) \right) \mathrm{d}x + \left[\eta \frac{\partial F}{\partial y'} \right]_b^a = 0.$$

This is true for all $\eta \in C^2[a,b]$ satisfying $\eta(a) = 0$. *In particular* it is true for all $\eta \in C^2[a,b]$ satisfying $\eta(a) = \eta(b) = 0$, so

$$\int_a^b \eta \left(\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \left(\frac{\partial F}{\partial y'} \right) \right) \mathrm{d}x + \left[\eta \frac{\partial F}{\partial y'} \right]_b^a = 0, \qquad \forall \eta \in C^2[a,b] \text{ such that } \eta(a) = 0.$$

Then FLCV $\Rightarrow$ Euler's equation again. Now we are left with

$$\left[\eta \frac{\partial F}{\partial y'} \right]_b^a = 0 = \eta(b) \left. \frac{\partial F}{\partial y'} \right|_{x=b}.$$

Since $\eta(b)$ is arbitrary we must have

$$\boxed{\frac{\partial F}{\partial y'} = 0 \qquad \text{at } x = b.}$$

This is the **natural boundary condition** applied at any boundary where no boundary conditions are prescribed in advance.

Trivial Example

Minimise the length

$$J[y] = \int_a^b \sqrt{1 + (y')^2} \, \mathrm{d}x$$

subject to $y(a) = c$ but $y(b)$ kept free.

Euler equation is

$$y'' = 0 \qquad \Rightarrow \qquad y = Ax + B.$$

Boundary conditions.

Imposed boundary condition

$$y(a) = c.$$

Natural boundary condition

$$\frac{\partial F}{\partial y'} = \frac{y'}{\sqrt{1 + (y')^2}} = 0 \qquad \text{at } x = b.$$

Thus $y'(b) = 0$. Thus $A = 0$ and $y' \equiv 0$, i.e.

$$y = c$$

as expected.

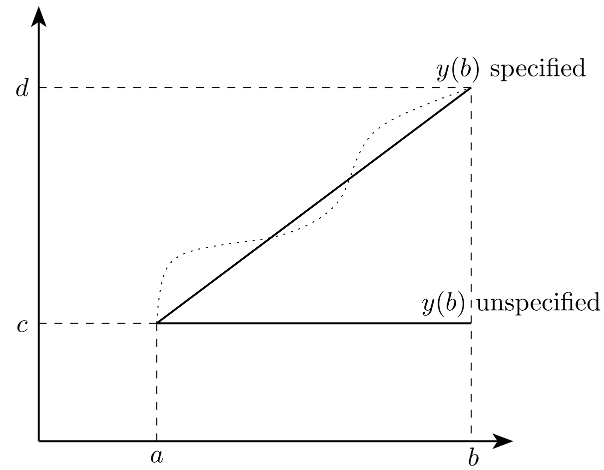


Figure 1:

3.3.2 Generalisation to higher derivatives

Suppose we want to minimise

$$J[y] = \int_a^b F(x, y, y', y'')\,\mathrm{d}x$$

subject to $y(a) = c$, $y(b) = d$, $y'(a) = m$, $y'(b) = n$. Perturbing y to $y + \epsilon\eta$ and linearising in η gives

$$J[y+\epsilon\eta] \quad \sim \quad J[y] + \epsilon \int_a^b \left(\eta \frac{\partial F}{\partial y} + \eta' \frac{\partial F}{\partial y'} + \eta'' \frac{\partial F}{\partial y''} \right) \mathrm{d}x + O(\epsilon^2).$$

At an extremal we have

$$\int_a^b \left(\eta \frac{\partial F}{\partial y} + \eta' \frac{\partial F}{\partial y'} + \eta'' \frac{\partial F}{\partial y''} \right) \mathrm{d}x = 0$$

$$\Rightarrow \qquad \int_a^b \left(\eta \frac{\partial F}{\partial y} - \eta \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right) + \eta \frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{\partial F}{\partial y''}\right) \right) \mathrm{d}x + \left[\eta \frac{\partial F}{\partial y'} + \eta' \frac{\partial F}{\partial y''} - \eta \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y''}\right) \right]_a^b = 0$$

Thus the Euler equation is

$$\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right) + \frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{\partial F}{\partial y''}\right) = 0.$$

This generalises in the obvious way.

3.3.3 More dependent variables

Suppose we want to minimise

$$J[y, z] = \int_a^b F(x, y, y', z, z')\,\mathrm{d}x$$

subject to $y(a) = c$, $y(b) = d$, $z(a) = m$, $z(b) = n$. We perturb and consider $J[y + \epsilon\eta, z + \delta\xi]$. Since we can vary η and ξ independently we get an Euler equation for each variable:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right) &= \frac{\partial F}{\partial y}, \\ \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial z'}\right) &= \frac{\partial F}{\partial z}. \end{aligned}$$

These will be coupled in general.

3.4 Constraints

3.4.1 Pointwise constraints

Once we have more dependent variables we can consider pointwise constraints of the form

$$G(y, z) = 0. \tag{21}$$

The condition for stationarity is

$$\int_a^b \left(\left[\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right) \right] \eta + \left[\frac{\partial F}{\partial z} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial z'}\right) \right] \xi \right) \mathrm{d}x = 0. \tag{22}$$

However, now η and ξ cannot be assigned arbitrarily because of the constraint (21). Taylor expanding (21) gives

$$\frac{\partial G}{\partial y}\eta + \frac{\partial G}{\partial z}\xi = 0.$$

Multiply by a Lagrange multiplier λ (which in this case is a function of x) and integrate to give

$$\int_a^b \left(\lambda\frac{\partial G}{\partial y}\eta + \lambda\frac{\partial G}{\partial z}\xi \right) \mathrm{d}x = 0.$$

Subtract this from (22) to give

$$\int_a^b \left(\left[\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right) - \lambda\frac{\partial G}{\partial y} \right] \eta + \left[\frac{\partial F}{\partial z} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial z'}\right) - \lambda\frac{\partial G}{\partial z} \right] \xi \right) \mathrm{d}x = 0.$$

Now suppose we choose λ so that the coefficient of η vanishes. Then since ξ can be chosen arbitrarily its coefficient must also vanish. Thus

$$\begin{aligned} \frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right) - \lambda\frac{\partial G}{\partial y} &= 0, \\ \frac{\partial F}{\partial z} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial z'}\right) - \lambda\frac{\partial G}{\partial z} &= 0. \end{aligned}$$

These two equations and (21) form three equations for y, z and λ. Note that this is the same as minimising $F - \lambda G$, as G does not depend on y' or z'.

3.4.2 Integral Constraints

Suppose we have to minimise or maximise a functional

$$J[y] = \int_a^b F(x, y, y')\,\mathrm{d}x$$

subject to $y(a) = c$ and $y(b) = d$ [can easily generalise to natural boundary conditions] and y has to satisfy the constraint

$$K[y] = \int_a^b G(x, y, y') = C \qquad \text{(constant)}.$$

Example
The minimal length curve enclosing a given area

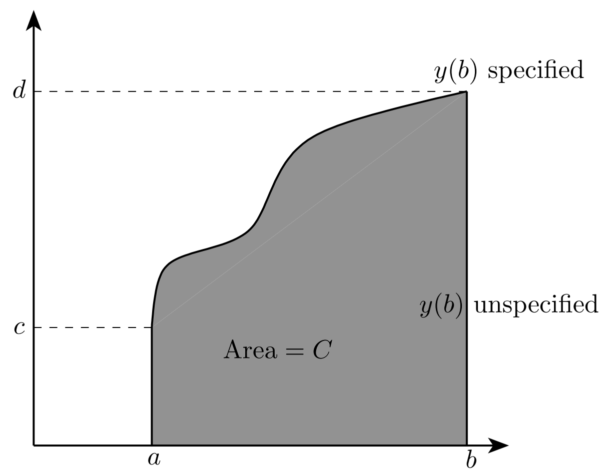


Figure 2:

$$\min J[y] = \int_a^b \sqrt{1 + (y')^2}\,\mathrm{d}x$$

subject to $y(a) = c$, $y(b) = d$ and

$$K[y] = \int_a^b y(x)\,\mathrm{d}x = C.$$

Now if we perturb about the extremal $y(x)$ then

$$\begin{aligned} K[y + \epsilon\eta] &\sim K[y] + \epsilon \int_a^b \left(\eta \frac{\partial G}{\partial y} + \eta' \frac{\partial G}{\partial y'} \right) \mathrm{d}x + O(\epsilon^2) \\ &\sim C + \epsilon \int_a^b \left(\eta \frac{\partial G}{\partial y} + \eta' \frac{\partial G}{\partial y'} \right) \mathrm{d}x + O(\epsilon^2) \\ &= C, \end{aligned}$$

so η is **not** arbitrary. It has to satisfy

$$\int_a^b \left(\eta\frac{\partial G}{\partial y} + \eta'\frac{\partial G}{\partial y'}\right)\,\mathrm{d}x = 0.$$

A trick to get around this problem is to add **two** perturbation functions, ξ and η satisfying $\xi(a) = \eta(a) = \xi(b) = \eta(b) = 0$. Then

$$\begin{aligned} K[y+\epsilon\eta+\delta\xi] &\sim K[y] + \epsilon\int_a^b \left(\eta\frac{\partial G}{\partial y} + \eta'\frac{\partial G}{\partial y'}\right)\,\mathrm{d}x + \delta\int_a^b \left(\xi\frac{\partial G}{\partial y} + \xi'\frac{\partial G}{\partial y'}\right)\,\mathrm{d}x + O(\epsilon^2) \\ &= C. \end{aligned} \tag{23}$$

The idea now is to *fix* the function $\xi(x)$ and, for any subsequently chosen $\eta(x)$, then to determine δ as a function of ϵ in such a way that (23) is satisfied. Thus η will be arbitrary, but we have to choose $\delta(\epsilon)$ in the right way. In order to be able to choose such a δ we need

$$\left.\frac{\partial K}{\partial \delta}\right|_{\delta=0,\epsilon=0} = \int_a^b \left(\xi\frac{\partial G}{\partial y} + \xi'\frac{\partial G}{\partial y'}\right)\,\mathrm{d}x = \int_a^b \xi\left(\frac{\partial G}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial G}{\partial y'}\right)\right)\,\mathrm{d}x \neq 0.$$

Provided

$$\frac{\partial G}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial G}{\partial y'}\right) \not\equiv 0$$

(cases where this is zero are degenerate and uninteresting), we can certainly choose ξ so that this is true. Let us choose such a ξ. Then, for *any* subsequent choice of η, we can determine δ as a function of ϵ so that (23) is satisfied. Now

$$\begin{aligned} J[y+\epsilon\eta+\delta(\epsilon)\xi] &\sim J[y] + \epsilon\int_a^b \left(\eta\frac{\partial F}{\partial y} + \eta'\frac{\partial F}{\partial y'}\right)\,\mathrm{d}x + \delta(\epsilon)\int_a^b \left(\xi\frac{\partial F}{\partial y} + \xi'\frac{\partial F}{\partial y'}\right)\,\mathrm{d}x \\ &\quad + O(\epsilon^2) \\ &\sim J[y] + \epsilon\int_a^b \left(\eta\frac{\partial F}{\partial y} + \eta'\frac{\partial F}{\partial y'}\right)\,\mathrm{d}x + \epsilon\frac{\mathrm{d}\delta}{\mathrm{d}\epsilon}(0)\int_a^b \left(\xi\frac{\partial F}{\partial y} + \xi'\frac{\partial F}{\partial y'}\right)\,\mathrm{d}x \\ &\quad + O(\epsilon^2). \end{aligned}$$

Since y is an extremal we must have

$$\int_a^b \eta\left(\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right)\right)\,\mathrm{d}x + \frac{\mathrm{d}\delta}{\mathrm{d}\epsilon}(0)\int_a^b \xi\left(\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right)\right)\,\mathrm{d}x = 0. \tag{24}$$

Similarly, integrating by parts in (23) gives

$$\int_a^b \eta\left(\frac{\partial G}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial G}{\partial y'}\right)\right)\,\mathrm{d}x + \frac{\mathrm{d}\delta}{\mathrm{d}\epsilon}(0)\int_a^b \xi\left(\frac{\partial G}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial G}{\partial y'}\right)\right)\,\mathrm{d}x = 0. \tag{25}$$

Solving (25) for $\mathrm{d}\delta/\mathrm{d}\epsilon$ and substituting into (24) gives

$$\int_a^b \eta\left(\frac{\partial}{\partial y}(F-\lambda G) - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'} - \lambda\frac{\partial G}{\partial y'}\right)\right)\,\mathrm{d}x = 0, \tag{26}$$

where λ is a constant, defined as the ratio of two definite integrals involving the arbitrary **fixed** function ξ:

$$\lambda = \frac{\displaystyle\int_a^b \xi \left(\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right)\right) \mathrm{d}x}{\displaystyle\int_a^b \xi \left(\frac{\partial G}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial G}{\partial y'}\right)\right) \mathrm{d}x}.$$

Since (26) is true for *any* $\eta \in C^2[a,b]$ satisfying $\eta(a) = \eta(b) = 0$ the FLVC implies that $F - \lambda G$ satisfies Euler's equation:

$$\boxed{\frac{\partial}{\partial y}(F - \lambda G) - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial}{\partial y'}(F - \lambda G)\right) = 0.}$$

λ is called a **Lagrange multiplier** and is fixed by satisfying the constraint

$$\boxed{\int_a^b G(x, y, y')\,\mathrm{d}x = C.}$$

This can also be thought of (and is taught in many books as) introducing a new functional (e.g. for $C = 0$)

$$\hat{J} = \int_a^b F(x, y, y') - \lambda G(x, y, y')\,\mathrm{d}x$$

and minimising over y. Then λ is determined from the constraint

$$\int_a^b G(x, y, y')\,\mathrm{d}x = C.$$

Simple Example

Minimise

$$\int_0^1 (y'(x))^2\,dx$$

over all $C^2[0,1]$ functions satisfying $y(0) = y(1) = 0$ and

$$\int_0^1 y(x)\,\mathrm{d}x = 1.$$

So $F = (y')^2$, $G = y$, giving

$$-\lambda - \frac{\mathrm{d}}{\mathrm{d}x}(2y') = 0$$

so that

$$y(x) = -\frac{\lambda x^2}{4} + Ax + B = -\frac{\lambda x(x-1)}{4},$$

after imposing the boundary conditions. Then fix λ by imposing the constraint

$$\int_0^1 y(x)\,\mathrm{d}x = \frac{\lambda}{24} = 1.$$

Thus $\lambda = 24$ and

$$y(x) = 6x(1-x).$$

3.5 More independent variables

Consider

$$J[\phi] = \iint_D F(x, \phi, \phi_x, \phi_y)\, \mathrm{d}x\, \mathrm{d}y$$

where $\phi = \phi(x, y)$, D is a region of the (x, y) plane, and ϕ satisfies $\phi = 0$ on ∂D.

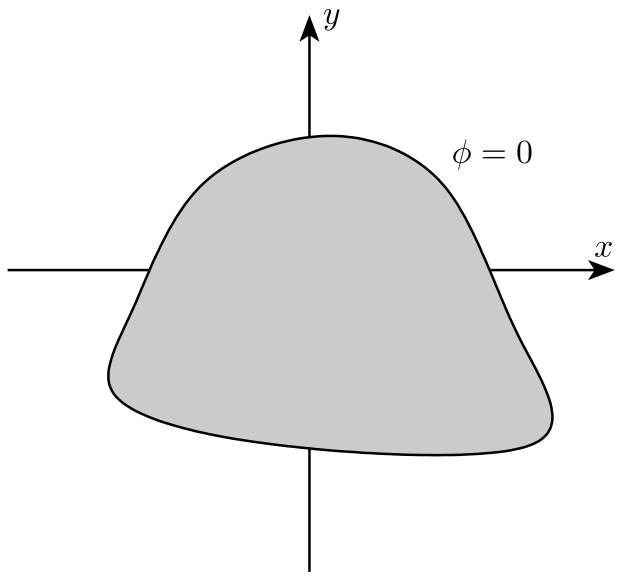


Figure 3:

$$J[\phi + \epsilon\eta] = J[\phi] + \epsilon \iint_D \left(\eta \frac{\partial F}{\partial \phi} + \eta_x \frac{\partial F}{\partial \phi_x} + \eta_y \frac{\partial F}{\partial \phi_y} \right) \mathrm{d}x\, \mathrm{d}y$$

Now instead of integration by parts we need to use Green's Theorem. From the identity

$$\nabla \cdot (\eta \mathbf{f}) = \nabla \eta \cdot \mathbf{f} + \eta\, \nabla \cdot \mathbf{f},$$

we find

$$\iint_D (\nabla \eta \cdot \mathbf{f} + \eta\, \nabla \cdot \mathbf{f})\, \mathrm{d}x\, \mathrm{d}y = \int_{\partial D} \eta\, \mathbf{f} \cdot \mathbf{n}\, \mathrm{d}s.$$

Thus, with

$$\mathbf{f} = \left(\frac{\partial F}{\partial \phi_x}, \frac{\partial F}{\partial \phi_y} \right),$$

we find

$$\begin{aligned} \iint_D \left(\eta_x \frac{\partial F}{\partial \phi_x} + \eta_y \frac{\partial F}{\partial \phi_y} \right) \mathrm{d}x\, \mathrm{d}y \quad &= \quad -\iint_D \left(\eta \frac{\partial}{\partial x} \left(\frac{\partial F}{\partial \phi_x} \right) + \eta \frac{\partial}{\partial y} \left(\frac{\partial F}{\partial \phi_y} \right) \right) \mathrm{d}x\, \mathrm{d}y \\ &\quad + \int_{\partial D} \eta \left(\frac{\partial F}{\partial \phi_x} n_x + \frac{\partial F}{\partial \phi_y} n_y \right) \mathrm{d}s. \end{aligned}$$

Thus

$$\iint_D \left(\eta \frac{\partial F}{\partial \phi} - \eta \frac{\partial}{\partial x} \left(\frac{\partial F}{\partial \phi_x} \right) - \eta \frac{\partial}{\partial y} \left(\frac{\partial F}{\partial \phi_y} \right) \right) \mathrm{d}x \, \mathrm{d}y = 0$$

for all η so the Euler equation (or Euler-Lagrange equation) is

$$\frac{\partial}{\partial x} \left(\frac{\partial F}{\partial \phi_x} \right) + \frac{\partial}{\partial y} \left(\frac{\partial F}{\partial \phi_y} \right) = \frac{\partial F}{\partial \phi}.$$

This is now a p.d.e.

3.6 Vector fields: two dependent variables and two independent variables

Notation. Here we will write all vectors as column vectors: so use $(u, v)^T$ to mean $\begin{pmatrix} u \\ v \end{pmatrix}$. This is a good practice to keep especially when you start to consider vector calculus and the nabla operations.

Let Ω be a bounded convex domain in two dimensions, with a piecewise smooth boundary denoted by $\delta\Omega$. Consider the functional:

$$J(u, v) = \int_\Omega F(x, y, u(x, y), u_x(x, y), u_y(x, y), v(x, y), v_x(x, y), v_y(x, y)) \, dxdy,$$

defined for the C^2 vector field $(u(x, y), v(x, y))^T$ on Ω, satisfying boundary conditions where $(u, v)^T$ is given and continuous on the bpoundary, $\delta\Omega$.
Find a pair of partial differential equations that the extremal vector field $(u, v)^T$ must satisfy.

Solution
Suppose that $(u, v)^T$ is an extremal.
Consider $J(u + \eta, v + \psi)$ where both η and ψ vanish on the boundary, $\delta\Omega$, then

$$J(u + \eta, v + \psi) = \int_\Omega F \, dxdy + \int_\Omega F_u \eta + F_{u_x} \eta_x + F_{u_y} \eta_y + F_v \psi + F_{v_x} \psi_x + F_{v_y} \psi_y \, dxdy,$$

Here F and its partial derivates are all evaluated at

$$(x, y, u(x, y), u_x(x, y), u_y(x, y), v(x, y), v_x(x, y), v_y(x, u)).$$

We will write $(F_{u_x}, F_{u_y})^T$ to denote the column vector field over Ω.
The second integral must vanish at an extremal. It is equal to

$$\int_\Omega F_u \eta + F_v \psi + \begin{pmatrix} F_{u_x} \\ F_{u_y} \end{pmatrix} . \nabla \eta + \begin{pmatrix} F_{v_x} \\ F_{v_y} \end{pmatrix} . \nabla \psi \, dxdy,$$

$$= \int_\Omega \left(F_u - \nabla . \begin{pmatrix} F_{u_x} \\ F_{u_y} \end{pmatrix} \right) \eta + \left(F_v - \nabla . \begin{pmatrix} F_{v_x} \\ F_{v_y} \end{pmatrix} \right) \psi \, dxdy,$$

This is true by the Divergence theorem (Green's theorem), since we have the identities

$$\nabla . \left(\eta \begin{pmatrix} F_{u_x} \\ F_{u_y} \end{pmatrix} \right) = \eta \nabla . \begin{pmatrix} F_{u_x} \\ F_{u_y} \end{pmatrix} + \nabla \eta . \begin{pmatrix} F_{u_x} \\ F_{u_y} \end{pmatrix}.$$

So

$$\int_{\delta\Omega} \left(\eta \begin{pmatrix} F_{u_x} \\ F_{u_y} \end{pmatrix} \right) . \mathbf{n} \, dS + \int_{\Omega} \left(F_u - \nabla . \begin{pmatrix} F_{u_x} \\ F_{u_y} \end{pmatrix} \right) \eta \, dxdy$$
$$= \int_{\Omega} \eta F_u + \begin{pmatrix} F_{u_x} \\ F_{u_y} \end{pmatrix} . \nabla \eta \, dxdy$$

and the integrand within the boundary integral (from the Divergence theorem) vanishes due to the boundary condition on η. Similar for v and ψ.

Hence we have a pair of scalar partial differential equations

$$F_u - \nabla . \begin{pmatrix} F_{u_x} \\ F_{u_y} \end{pmatrix} = 0 \quad F_v - \nabla . \begin{pmatrix} F_{v_x} \\ F_{v_y} \end{pmatrix} = 0 \qquad (x, y)^T \in \Omega.$$

So if, say, $F(x, u, u, v, u_x, u_u, v_x, v_y) = \nabla u . \nabla u + \nabla v . \nabla v$ then we have

$$0 = 2\nabla . \nabla u = 2\Delta u \text{ and } 0 = 2\nabla . \nabla v = 2\Delta v.$$

3.7 The Hamiltonian

Suppose a function $y(x)$ satisfies Euler's equation

$$\frac{\mathrm{d}}{\mathrm{d}x} \left(\frac{\partial F}{\partial y'} \right) = \frac{\partial F}{\partial y}$$

for some function $F(x, y, y')$. Note that

$$\begin{aligned} \frac{\mathrm{d}F}{\mathrm{d}x} &= \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} \frac{\mathrm{d}y}{\mathrm{d}x} + \frac{\partial F}{\partial y'} \frac{\mathrm{d}y'}{\mathrm{d}x} \\ &= \frac{\partial F}{\partial x} + \frac{\mathrm{d}}{\mathrm{d}x} \left(\frac{\partial F}{\partial y'} \right) \frac{\mathrm{d}y}{\mathrm{d}x} + \frac{\partial F}{\partial y'} \frac{\mathrm{d}^2 y}{\mathrm{d}x^2} \\ &= \frac{\partial F}{\partial x} + \frac{\mathrm{d}}{\mathrm{d}x} \left(\frac{\mathrm{d}y}{\mathrm{d}x} \frac{\partial F}{\partial y'} \right) = \frac{\partial F}{\partial x} + \frac{\mathrm{d}}{\mathrm{d}x} \left(y' \frac{\partial F}{\partial y'} \right). \end{aligned}$$

Therefore, if we define the **Hamiltonian**

$$H = y' \frac{\partial F}{\partial y'} - F,$$

then

$$\frac{\mathrm{d}H}{\mathrm{d}x} = -\frac{\partial F}{\partial x}.$$

If F does not depend explicitly on x (the problem is *autonomous*) then

$$\frac{\partial F}{\partial x} = 0$$

and hence $H =$ constant. In this case H is a *conserved quantity* (often identifyable as energy).

Example
Suppose

$$F = \sqrt{1 + (y')^2} + y^2.$$

The Euler equation is

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{y'}{\sqrt{1 + (y')^2}}\right) = 2y.$$

This is not very nice...but

$$H = \frac{(y')^2}{\sqrt{1 + (y')^2}} - \sqrt{1 + (y')^2} - y^2 = -\frac{1}{\sqrt{1 + (y')^2}} - y^2 = \text{constant},$$

gives a *first integral* of the o.d.e.

We can transform the Euler equation into *canonical form* by changing independent variables. Think of F and H as functions of (x, p, q) instead of (x, y, y'), where

$$q = y, \qquad p = \frac{\partial F}{\partial y'};$$

p is known as the generalised momentum. Then, then definition of H is

$$H = py' - F$$

(where y' is a function of x, p, q) and Euler's equation is

$$\frac{\mathrm{d}p}{\mathrm{d}x} = \frac{\partial F}{\partial y}.$$

So

$$\frac{\partial H}{\partial y'} = p + y'\frac{\partial p}{\partial y'} - \frac{\partial F}{\partial y'} = y'\frac{\partial p}{\partial y'}$$

by the Chain rule, since $p = \partial F/\partial y'$. But

$$\frac{\partial H}{\partial y'} = \frac{\partial H}{\partial q}\frac{\partial q}{\partial y'} + \frac{\partial H}{\partial p}\frac{\partial p}{\partial y'} = \frac{\partial H}{\partial p}\frac{\partial p}{\partial y'}.$$

Thus

$$y' = \frac{\mathrm{d}q}{\mathrm{d}x} = \frac{\partial H}{\partial p}.$$

Also

$$\frac{\mathrm{d}p}{\mathrm{d}x} = \frac{\partial F}{\partial y} = \frac{\partial}{\partial y}\big(py' - H\big) = y'\frac{\partial p}{\partial y} - \frac{\partial H}{\partial y}.$$

But

$$\frac{\partial H}{\partial y} = \frac{\partial H}{\partial p}\frac{\partial p}{\partial y} + \frac{\partial H}{\partial q}\frac{\partial q}{\partial y} = y'\frac{\partial p}{\partial y} + \frac{\partial H}{\partial q}.$$

Thus

$$\frac{\mathrm{d}p}{\mathrm{d}x} = y'\frac{\partial p}{\partial y} - y'\frac{\partial p}{\partial y} - \frac{\partial H}{\partial q} = -\frac{\partial H}{\partial q}.$$

Thus

$$\boxed{\frac{\mathrm{d}p}{\mathrm{d}x} = -\frac{\partial H}{\partial q} \qquad\qquad \frac{\mathrm{d}q}{\mathrm{d}x} = \frac{\partial H}{\partial p}.}$$

These are **Hamilton's equations**. Note that

$$\frac{\mathrm{d}H}{\mathrm{d}x} = \frac{\partial H}{\partial x} + \frac{\partial H}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}x} + \frac{\partial H}{\partial q}\frac{\mathrm{d}q}{\mathrm{d}x} = \frac{\partial H}{\partial x} + \frac{\mathrm{d}q}{\mathrm{d}x}\frac{\mathrm{d}p}{\mathrm{d}x} - \frac{\mathrm{d}p}{\mathrm{d}x}\frac{\mathrm{d}q}{\mathrm{d}x} = \frac{\partial H}{\partial x}.$$

Thus if

$$\frac{\partial H}{\partial x} = 0$$

then H is conserved as expected.

Free boundaries

Minimise

$$J[y,b] = \int_a^b F(x,y,y')\,\mathrm{d}x$$

subject to $y(a) = c$, $y(b) = d$ where b is **unspecified**.

$$\begin{aligned} J[y+\epsilon\eta; b+\epsilon\beta] &= \int_a^{b+\epsilon\beta} F(x, y+\epsilon\eta, y'+\epsilon\eta')\,\mathrm{d}x \\ &= J[y,b] + \epsilon\left\{\int_a^b \left(\eta\frac{\partial F}{\partial y} + \eta'\frac{\partial F}{\partial y'}\right)\mathrm{d}x + \beta F(b, y(b), y'(b))\right\} + O(\epsilon^2). \end{aligned}$$

Taylor expanding the boundary condition

$$\begin{aligned} d &= y(b+\epsilon\beta) + \eta(b+\epsilon\beta) \\ &= y(b) + \epsilon\left[\beta y'(b) + \eta(b)\right] + O(\epsilon^2) \\ &= d + \epsilon\left[\beta y'(b) + \eta(b)\right] + O(\epsilon^2). \end{aligned}$$

Thus

$$\eta(a) = 0, \qquad \eta(b) = -\beta y'(b).$$

At an extremal

$$\int_a^b \left(\eta\frac{\partial F}{\partial y} + \eta'\frac{\partial F}{\partial y'}\right)\mathrm{d}x + \beta F(b, y(b), y'(b)) = 0.$$

Integrate by parts to give

$$\beta F(b, y(b), y'(b)) + \int_a^b \eta\left[\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right)\right]\mathrm{d}x + \left[\eta\frac{\partial F}{\partial y'}\right]_a^b = 0.$$

Hence

$$\beta \left[F - y' \frac{\partial F}{\partial y'} \right]_{x=b} + \int_a^b \eta \left[\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x} \left(\frac{\partial F}{\partial y'} \right) \right] \mathrm{d}x = 0.$$

Thus FLCV gives us Euler's equation and the extra free boundary condition

$$F = y' \frac{\partial F}{\partial y'} \qquad \text{at } x = b$$

(i.e. $H = 0$).

Example
minimise

$$J[y, b] = \int_0^b \left(\frac{1}{2} \left(y'\right)^2 + \frac{1}{2} y^2 + 1 \right) \mathrm{d}x$$

subject to $y(0) = 0$, $y(b) = 1$. Euler's equation is

$$y'' = y.$$

Solving and applying the boundary conditions gives

$$y = \frac{\sinh x}{\sinh b}.$$

The extra free boundary condition is

$$\frac{1}{2} \left(y'\right)^2 - \frac{1}{2} y^2 - 1 = 0 \text{ at } x = b.$$

This gives

$$\frac{\cosh^2 b}{\sinh^2 b} = 3 \quad \Rightarrow \quad b = \tanh^{-1} \left(\frac{1}{\sqrt{3}} \right).$$

CHECK

$$J \left[\frac{\sinh x}{\sinh b}, b \right] = \frac{1}{2} \coth b + b.$$

This is minimised when

$$-\frac{1}{2} \mathrm{cosech}^2 b + 1 = 0 \quad \Rightarrow \quad b = \tanh^{-1} \left(\frac{1}{\sqrt{3}} \right).$$

OR note that

$$H = \frac{1}{2} \left(y'\right)^2 - \frac{1}{2} y^2 - 1 = \text{constant (autonomous)} = 0$$

by the free boundary condition. Hence

$$y' = \sqrt{y^2 + 2}.$$

Thus

$$x = \int \frac{\mathrm{d}y}{\sqrt{y^2+2}} = \sinh^{-1}\left(\frac{y}{\sqrt{2}}\right).$$

Thus

$$y = \sqrt{2}\sinh x.$$

Then the boundary condition $y(b) = 1$ gives

$$b = \sinh^{-1}\left(\frac{1}{\sqrt{2}}\right) = \tanh^{-1}\left(\frac{1}{\sqrt{3}}\right).$$

4 Optimal control

Example
Suppose $x(t)$ satisfies the differential equation

$$\dot{x} = u + x,$$

where $u(t)$ is our control variable. Suppose we want to vary u so as to control x. For example, starting from $x(0) = a$ we may wish to arrive at $x(T) = 0$.

Is this possible? Yes! Just choose *any* function $x(t)$ satisfying the initial and final condition and then read off the required control as

$$u = \dot{x} - x.$$

However, in practice there may be bounds on the achievable u, e.g. $-1 \leq u \leq 1$. This will leads to bounds on the initial condition for which the desired final condition is achievable. In the example if $u \leq 1$ then the maximum achievable value of $x(T) - x(0)$ occurs when

$$\dot{x} - x = 1 \quad \Rightarrow \quad x = -1 + A\mathrm{e}^t = -1 + (1+a)\mathrm{e}^t.$$

Then $x(T) = 0$ gives $a = -1 + \mathrm{e}^{-T}$. The problem is controllable only if a is greater than this value.

We may wish to find the control which *minimises* a cost function. For example, the work done agaist friction may be

$$\int_0^T u\dot{x}\,\mathrm{d}t.$$

Thus we may want to define the cost function as

$$C = \int_0^T u(u+x)\,\mathrm{d}t$$

and ask for the control which achieves the goal and minimises $C[x, u]$.

4.1 General setting

So, in general we may find the following optimal control problem:

$$\text{minimise } C[x,u] = \int_0^T h(t,x,u)\,\mathrm{d}t,$$

over all controls $u(t)$ satisfying the control problem

$$\dot{x} = f(t,x,u), \qquad x(0) = a, \quad x(T) = b.$$

We wish to "control" the solution x so as to arrive at the far boundary condition at the lowest possible cost.
This now resembles a variational problem, with the control functional acting as a **constraint**.

Notation: in this section we will use subscripts to denote partial derivatives and a dot to denote a full time derivative (d/dt), where no confusion arises.

Let us approach it by perturbing about the extremal functions $x(t)$ and $u(t)$:

$$C[x + \epsilon\xi, u + \epsilon\eta] = C[x, u] + \epsilon \int_0^T (\xi h_x + \eta h_u) \, \mathrm{d}t + O(\epsilon^2),$$

while

$$\dot{x} + \epsilon\dot{\xi} = f(t, x, u) + \epsilon (\xi f_x + \eta f_u) + O(\epsilon^2),$$

and $\xi(0) = 0 = \xi(T)$.

Since $\dot{x} = f(t, x, u)$ for an extremal function we need

$$\int_0^T (\xi h_x + \eta h_u) \, \mathrm{d}t = 0,$$

for all ξ and η satisfying

$$\dot{\xi} = \xi f_x + \eta f_u,$$

with $\xi(0) = \xi(T) = 0$. We require $\partial f/\partial u \neq 0$, otherwise the control u has no influence on the problem. Then we can solve for

$$\eta = \left(\dot{\xi} - \xi f_x\right) \Big/ f_u,$$

and plug it into the integral

$$\int_0^T \left(\xi h_x + \left(\dot{\xi} - \xi f_x\right)(h_u/f_u)\right) \mathrm{d}t = 0.$$

As usual integrate by parts to give

$$\int_0^T \xi \left(h_x - f_x(h_u/f_u) - \frac{\mathrm{d}}{\mathrm{d}t}(h_u/f_u) \right) \mathrm{d}t + [\xi(h_u/f_u)]_0^T = 0$$

Since $\xi(0) = \xi(T) = 0$ the boundary term is zero. Hence we find that x and u have to satisfy the o.d.e.

$$\frac{\mathrm{d}}{\mathrm{d}t}(h_u/f_u) = h_x - f_x\left((h_u/f_u)\right).$$

This o.d.e. is coupled with the control problem

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f,$$

with $x(0) = a$, $x(T) = b$. In principle two coupled first order o.d.e.s with two boundary conditions gives a unique solution.

Return to the example

$$f = u + x, \qquad h = u(u + x),$$

so that we get

$$\frac{\mathrm{d}}{\mathrm{d}t}(2u + x) = u - (2u + x),$$

i.e.

$$2\dot{u} + \dot{x} = -(u + x),$$

along with the control problem

$$\dot{x} = u + x.$$

Adding gives

$$\dot{u} + \dot{x} = 0,$$

so that $u + x = A$ constant. Then $\dot{x} = A$, $x(0) = a$, $x(T) = 0$ gives

$$x = a + At, \qquad A = -\frac{a}{T},$$

so that

$$x = a\left(1 - \frac{t}{T}\right),$$

(constant velocity is the most efficient), and the optimal control is

$$u = a\left(\frac{t}{T} - 1 - \frac{1}{T}\right).$$

Note the existence of a first integral which facilitated the solution of this example. As in the calculus of variations this will be generally true for *autonomous* problems.

Returning to the **general case.**
We define the **Hamiltonian**:

$$H(t, x, u) = f\ (h_u/f_u) - h,$$

where we have the dynamics

$$\frac{\mathrm{d}}{\mathrm{d}t}(h_u/f_u) = h_x - f_x\,(h_u/f_u).$$

together with

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f.$$

We will show that when the system is autonomous (there being no explicit dependence upon t, and it is thus translatable with respect to time) there is a first integral for the *motion*, a conserved quantity.

To ease the notation we will use subscripts to denote partial differentiation of f and h and $\dot{f}$, etc to denote df/dt and other full time derivatives.

We have

$$H(t,x,u) = f\left(\frac{h_u}{f_u}\right) - h.$$

Then direct differentiation yields

$$\dot{H} = \dot{f}\left(\frac{h_u}{f_u}\right) + f\left(\dot{\overline{\frac{h_u}{f_u}}}\right) - h_t - h_x f - h_u \dot{u}.$$

So substituting for $\left(\dot{\overline{\frac{h_u}{f_u}}}\right) = h_x - f_x\left(\frac{h_u}{f_u}\right)$,

$$\dot{H} = (f_t + f_x f + f_u \dot{u})\left(\frac{h_u}{f_u}\right) + f\left(h_x - f_x\left(\frac{h_u}{f_u}\right)\right) - h_t - h_x f - h_u \dot{u}.$$

Thus

$$\dot{H} = f_t\left(\frac{h_u}{f_u}\right) - h_t$$

Equivalently

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \left(\frac{\partial h}{\partial u}\Big/\frac{\partial f}{\partial u}\right)\frac{\partial f}{\partial t} - \frac{\partial h}{\partial t}.$$

So if the problem is autonomous, then

$$\frac{\partial f}{\partial t} = \frac{\partial h}{\partial t} = 0$$

and H is conserved.

In the above example $f = u + x$, $h = u(u + x)$ and

$$H = (u + x)(2u + x) - u(u + x) = (u + x)^2$$

which is conserved as we found before.

Example 2
Solve

$$\dot{x} = x + u, \qquad x(0) = 0, \quad x(1) = 1,$$

where u is chosen to minimise

$$\int_0^1 u^2\,\mathrm{d}t.$$

Now $f = x + u$, $h = u^2$, so the Hamiltonian is

$$H = (x + u) \times 2u - u^2 = u^2 + 2xu.$$

Completing the square

$$(u+x)^2 = H^2 + x^2 \quad \Rightarrow \quad u + x = \pm\sqrt{H + x^2}.$$

Choose the plus sign ($\dot{x}$ should be positive from the initial and final conditions) to give

$$\dot{x} = \sqrt{H + x^2}.$$

Therefore

$$t = \int \frac{\mathrm{d}x}{\sqrt{H + x^2}} = \sinh^{-1}\left(\frac{x}{\sqrt{H}}\right).$$

Therefore

$$x = \sqrt{H}\sinh t.$$

The final condition $x(1) = 1$ determines H, to gives

$$x = \frac{\sinh t}{\sinh 1}, \qquad H = \mathrm{cosech}^2 1, \qquad u = \frac{\mathrm{e}^{-t}}{\sinh 1}.$$

4.2 The Pontryagin Maximum Principle (Non-examinable)

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 minumum 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**

$$H(x, t, p) = \max\{H_0(x, t, u, p) : u \in U_f(x, t)\} = H_0(x, t, u_0(x, t, p), p). \tag{27}$$

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 start 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

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

Then

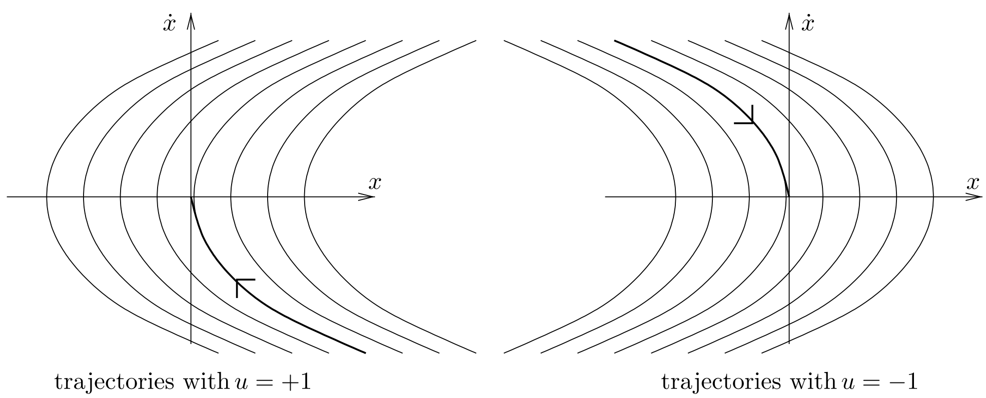
$$\begin{aligned} \dot{p}_1 &= -\frac{\partial H}{\partial x_1} = 0, \\ \dot{p}_2 &= -\frac{\partial H}{\partial x_2} = -p_1. \end{aligned}$$

We already see that

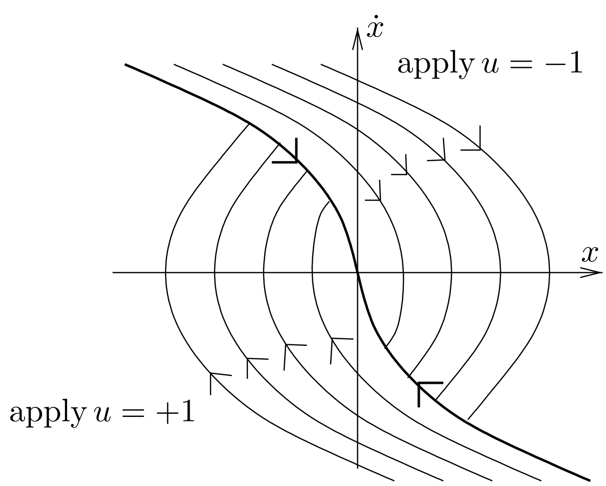
(a) the optimal trajectory will use $u = \pm 1$ only: "bang-bang" control.

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

What does this mean in the phase plane? 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 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 and 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. Depends on z being in the interior of the set over which f is defined and on f, g being differentiable in z. Does not depend on any differentiability in x or y. $\square$

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, f 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 $\delta \to 0$:

$$h(x,t,u) + F_x(x,t)f(x,t,u) + F_t \le 0, \tag{28}$$

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 (28) for almost all t, and again for simplicity assume it holds everywhere. So (28) 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

$$H(x,t,F_x(x,t)) = -F_t(x,t). \tag{29}$$

Now assume that H is C^1 and F is C^2. To derive the $\dot{p}$ equation, first note that by (27)

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

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

$$f(x,t,u_0(x,t,p)) = H_p(x,t,p). \tag{30}$$

Then the derivative of p along the optimal trajectory is

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

But by (29)

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

by (30). 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 (29). So (iii) is satisfied.

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