Mathematical Methods

J.Pearson

May 23, 2008

Abstract

These are a set of notes I have made, based on lectures given by N.Walet at the University of Manchester Jan-June '08. Please e-mail me with any comments/corrections: jap@watering.co.uk.

Contents

1 Linear Algebra **1**

1.1 Linear Vector Spaces . . . 1

1.2 Linear Independance & Basis Vectors . . . 3

1.3 Scalar Product . . . 4

1.4 Orthonormalisation . . . 4

1.4.1 Gramm-Schmidt Procedure . . . 5

1.4.2 Norm-Matrix Method . . . 5

1.4.3 Summary . . . 5

1.5 Function Spaces . . . 6

1.5.1 Example: Fourier Basis . . . 6

1.5.2 The δ Function . . . 7

1.5.3 Normalisable Function Spaces . . . 9

2 Operators, Eigenvectors & Eigenvalues **10**

2.1 Linear Operators . . . 10

2.1.1 Domain, Codomain & Range . . . 10

2.1.2 Matrix Representation of Linear Operators . . . 11

2.1.3 Adjoint and Hermitian Operators . . . 12

2.2 Eigenvalue Problems . . . 13

2.2.1 Proof: Real Eigenvalues of Hermitian Operators . . . 13

2.2.2 Proof: Orthogonal Eigenfunctions . . . 14

2.2.3 Example: Diagonalise A Matrix . . . 14

2.2.4 Generalised Eigenvalue Problem . . . 15

2.3 Sturm-Lioville Equations . . . 16

2.3.1 Bringing Equations to SL Form . . . 16

2.3.2 A Useful Result . . . 17

2.3.3 Second Solutions . . . 18

2.3.4 Classifying Singular Points . . . 18

2.4 Series Solutions . . . 19

2.4.1 Example: Quantum Harmonic Oscillator . . . 19
2.5 Legendre Polynomials . . . 22
2.5.1 Generating Function . . . 22
2.5.2 Application: Expand $\frac{1}{|r_1-r_2|}$ . . . 24
2.5.3 Normalisation . . . 25
2.6 Gram-Schmidt for Functions . . . 25

3 Green Functions **27**
3.0.1 Example: Electrostatics . . . 27
3.1 The Eigenstate Method . . . 28
3.1.1 Example: Vibrating String . . . 30
3.2 The Continuity Method: Strictly 1D . . . 31
3.2.1 Example: Vibrating String . . . 31
3.3 Quantum Mechanical Scattering . . . 33
3.3.1 The Born Approximation . . . 34
3.4 Time Dependent Wave Equation . . . 35
3.4.1 Solution for Green functions by Fourier Transforms . . . 36
3.5 Reducing the Dimension . . . 38

4 Variational Calculus **39**
4.1 Functionals . . . 39
4.2 Stationary Points . . . 39
4.3 Special Cases . . . 40
4.3.1 No y-dependance . . . 40
4.3.2 Example: Distance Between Two Points . . . 41
4.3.3 No x-dependance . . . 42
4.3.4 Example: Snell's Law . . . 42
4.3.5 Example: Brachistochrone . . . 44
4.4 Variable Endpoints . . . 45
4.4.1 Example: Brachistochrone . . . 46
4.5 More Than One Function . . . 47

4.5.1 Hamilton's Principle of Least Action 47
4.6 Field Theory 48
4.7 Constrained Variational Problems 52
4.7.1 Lagrange's Method of Undetermined Multipliers 52
4.7.2 Generalisation to Functionals 54
4.7.3 Example: The Catenary 55
4.7.4 Eigenvalue Problems 56
4.7.5 Rayleigh-Ritz Method 57

1 Linear Algebra

1.1 Linear Vector Spaces

A linear vector space V is a set of objects (i.e. vectors) with two operations: addition and multiplication. Addition is defined between two vectors, and multiplication by a scalar, which is $\in S$. Typically, the set of which S is in, is $\mathbb{R}, \mathbb{C}$. Now, we say that the operations must satisfy the following conditions:

- V is closed under addition and multiplication:

  $\forall a, b \in V$ then $a + b \in V$

  $\forall a \in V, \lambda \in S$ then $\lambda a \in V$

- Addition is commutative and associative:

  $\forall a, b \in V$, then $a + b = b + a$, that is, commutative.

  $\forall a, b, c \in V$, then $(a + b) + c = a + (b + c)$, that is, associative.

- There exists an identity element e for both addition and multiplication, so that $\forall a \in V$, $a \cdot e = e \cdot a = a$. This element is different for both operations:

  For addition, the identity is the null vector 0; so that $0 \in V$ so that $\forall a \in V$, $a + 0 = 0 + a = a$.

  For multiplication, use the unit element. $1 \in S$ so that $\forall a \in V$, $1a = a1 = a$.

- Multiplication is distributive and associative for addition of both vectors and scalars:

  $\forall a, b, \in V$ and $\lambda \in S$, then $\lambda(a + b) = \lambda a + \lambda b$.

  $\forall a \in V$ and $\lambda, \mu \in S$, then $(\lambda + \mu)a = \lambda a + \mu a$.

  $\forall a \in V$ and $\lambda, \mu \in S$, then $\lambda(\mu a) = (\lambda\mu)a$. That is, associative.

- For addition, the inverse element is $-a$, so that $\forall a \in V$, then $a + (-a) = (-a) + a = 0$, which is the identity element for the addition set.

Example Such a set of vectors could be $V = \mathbb{R}^3$:

$$\boldsymbol{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

And by inspection of all the above axioms, we see that V is a vector space over all $S \in \mathbb{R}$.

Example The space of 2D complex spinors:

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Is a vector space over all $S \in \mathbb{C}$. However, we see that the space of spinors of length 1 is not a vector space, as it is not closed under addition. To see this, consider two vectors $v_1, v_2 \in$the spinor space of length 1. However, upon addition of the two, and squaring, we find that the sum is not necessarily of length one. Hence, the set is not closed. Therefore, not a vector space.

Example Consider the space of all functions $f(x)$ that are square-integrable. That is:

$$\int_{-\infty}^{+\infty} |f(x)|^2 \, dx < \infty$$

The set of such functions is denoted $\mathcal{L}^2(\mathbb{R})$.
Where we let $f(x) : \mathbb{R} \mapsto \mathbb{C}$; that is, the function has the effect of going from the set of real numbers to the set of complex. This is quite a 'broad restriction'. To see that this space is closed, consider the product of two such functions $f, g \in \mathcal{L}^2(\mathbb{R})$. They will require:

$$\int |f(x)|^2 + |g(x)|^2 + 2\Re(fg^*) \, dx < \infty$$

It is not immediately obvious that this is the case. The first two terms are finite, as that was stated in the definition of f, g. Upon inspection, we see that the third term must also decay at $\pi\infty$. Hence, confirmed that $\mathcal{L}^2(\mathbb{R})$ is a vector space.

Let us ask: how can we show that the null vector 0 is unique? How can we know that for any $a \in V$, there is only one inverse $(-a)$? We consider the first question, but the second follows in the exact same argument.
Suppose that there are 2 null vectors 0 and $\oslash$. Then, each must satisfy:

$$a + 0 = a \qquad a + \oslash = a$$

Hence:

$$a + 0 = a + \oslash$$

Now, to each side add on the inverse of a, i.e. $(-a)$:

$$a + 0 + (-a) = a + \oslash + (-a)$$

Now, by the definition of the inverse of a, we see that $a + (-a) = 0$. Hence:

$$0 + 0 = 0 + \oslash$$

Therefore:

$$0 = \oslash$$

Hence, only one null vector. It is an identical argument to show that the inverse is unique.

In the entire section above, I have left off the 'boldface' type for vectors. It is implied when I state that a component is an element of a particular set.

1.2 Linear Independance & Basis Vectors

A set of vectors $\boldsymbol{a}, \boldsymbol{b}, \ldots, \boldsymbol{u}$ is called independent if:

$$\lambda \boldsymbol{a} + \mu \boldsymbol{b} + \ldots + \sigma \boldsymbol{u} = \boldsymbol{0}$$

Has only one solution:

$$\lambda = \mu = \ldots = 0$$

That is, you cannot express any one vector as the sum of any other.

The dimension n of a space V is the largest possible number of independent vectors. Any set of n independent vectors in an n-dimensional space thus forms a complete set of basis vectors. Hence, any vector can be written as the sum of basis vectors:

$$\boldsymbol{x} = x_1 \boldsymbol{e}_1 + x_2 \boldsymbol{e}_2 + \ldots + x_n \boldsymbol{e}_n$$

The x_i are the components of $\boldsymbol{x}$ in the basis $\{\boldsymbol{e}_i\}_{i=1}^n$; this notation reads: 'the set of all basis vectors'.

Example Show that the vectors $(1,1,0), (1,0,1), (0,1,1)$ are linearly independant, and find the components of (x,y,z) in this basis.

To show that they are linearly independant, we must show that the only solution to:

$$\alpha(1,1,0) + \beta(1,0,1) + \gamma(0,1,1) = (0,0,0)$$

Is $\alpha = \beta = \gamma = 0$. To do this, we can write the simultaneous equations, and put it into matrix form:

$$\begin{aligned} \alpha + \beta &= 0 \\ \alpha + \gamma &= 0 \\ \beta + \gamma &= 0 \end{aligned}$$

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

And solve. To show that a solution exists, we must have that $\det M \neq 0$, which is the case (and can be checked, and seen to be -2). This condition arises from the computation of the inverse: the solution involves the inverse, which involves finding the reciprocal of the matrix (which cannot be done if the matrix has zero determinant). Via gaussian elimination, we find that the only solution is indeed $\alpha = \beta = \gamma = 0$.
For the second part, we wish to compute x_1, x_2, x_3, where:

$$(x,y,z) = x_1(1,1,0) + x_2(1,0,1) + x_3(0,1,1)$$

So, we have the system of equations, which we solve via Gaussian elimination:

$$\begin{aligned} x_1 + x_2 &= x \\ x_1 + x_3 &= y \\ x_2 + x_3 &= z \end{aligned}$$

Which can easily be found to have solutions:

$$x_1 = \tfrac{1}{2}(x+y-z) \quad x_2 = \tfrac{1}{2}(x+z-y) \quad x_3 = \tfrac{1}{2}(y+z-x)$$

Hence found.

To show that the so called 'decomposition' $\boldsymbol{x} = x_i \boldsymbol{e}_i$ (summation implied) is unique, we suppose that it is not. That is, we suppose we may also write $\boldsymbol{x} = y_i \boldsymbol{e}_i$. Hence:

$$\boldsymbol{x} = \boldsymbol{x} \quad \Rightarrow \quad x_i \boldsymbol{e}_i = y_i \boldsymbol{e}_i \quad \Rightarrow \quad (x_i - y_i)\boldsymbol{e}_i = 0$$

Which we see has solutions $x_i = y_i$. Hence, the decomposition is unique.

Example The Fourier series of a function defined on $[-\pi, \pi]$ is given by:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + b_n \sin nx$$

Hence, for Fourier series, the basis functions are:

$$1, \{\sin nx, \cos nx\} \quad n = 1, \ldots, \infty$$

It is very hard however, to show that this basis is complete. This is a general complication in infinite-dimensional spaces.

1.3 Scalar Product

This is sometimes referred to as the 'inner product'. The notation we shall use for $\boldsymbol{a} \cdot \boldsymbol{b}$ is:

$$(\boldsymbol{a}, \boldsymbol{b})$$

We have the following relations:

$$\begin{aligned} (\boldsymbol{a}, \boldsymbol{b}) &= (\boldsymbol{b}, \boldsymbol{a})^* \\ (\boldsymbol{a}, \lambda\boldsymbol{b} + \mu\boldsymbol{c}) &= \lambda(\boldsymbol{a}, \boldsymbol{b}) + \mu(\boldsymbol{a}, \boldsymbol{c}) \\ (\boldsymbol{a}, \boldsymbol{a}) &\geq 0 \end{aligned}$$

Where we equality (in the final relation) if and only if $\boldsymbol{a} = 0$. We define the length, or the *norm* of a vector as:

$$||\boldsymbol{a}|| = \sqrt{(\boldsymbol{a}, \boldsymbol{a})}$$

If $(\boldsymbol{a}, \boldsymbol{b}) = 0$, then the two vectors are orthogonal.

1.4 Orthonormalisation

We can find an orthonormal basis from some non-orthonormal basis. There are 2 ways to do this:

1.4.1 Gramm-Schmidt Procedure

This procedure is an algorithm, to find the orthonormal basis vector $\boldsymbol{e}'_m$, from a non-orthonormal vector $\boldsymbol{e}_m$, provided that all $\boldsymbol{e}'_{m-1}$ basis vectors are. This last condition is satisfied, as the procedure is recursive. The algorithm is below:

$$\boldsymbol{e}''_m = \boldsymbol{e}_m - \sum_{i=1}^{m-1} (\boldsymbol{e}'_i, \boldsymbol{e}_m)\boldsymbol{e}'_i \tag{1.1}$$

$$\boldsymbol{e}'_m = \frac{\boldsymbol{e}''_m}{||\boldsymbol{e}''_m||} \tag{1.2}$$

So, we see that the first step will pick out all basis vectors which are not orthonormal to $\boldsymbol{e}_m$, then multiply the 'overlap' by the vector being examined (the $\boldsymbol{e}_i$). This is then subtracted from the non-orthonormal vector, to give something with is now orthogonal. The final step normalises the orthogonal vector.

1.4.2 Norm-Matrix Method

A second approach is to use an overlap matrix N, whose elements are given by:

$$N_{ij} = (\boldsymbol{e}_i, \boldsymbol{e}_j)$$

Now, if we can find matrix powers (which we will come to in a moment), then you can find the i^{th} orthogonal basis vector $\boldsymbol{e}'^{(i)}$ from the non-orthogonal basis vector $\boldsymbol{e}^{(i)}$ via:

$$\boldsymbol{e}'^{(i)}_k = N^{-\frac{1}{2}}_{kl} \boldsymbol{e}^{(i)}_l$$

Now, let:

$$N = O\text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)O^T$$

Where O is the matrix of eigenvectors, where its columns are composed of eigenvectors $O_{ij} = e^i_j$. Hence, we have that $O^{-1} = O^T$. Hence, $OO^T = I$. The middle matrix has diagonal entries, which are the eigenvalues. Thus, we define matrix powers:

$$N^{-\frac{1}{2}} = O\text{diag}(\lambda_1^{-\frac{1}{2}}, \lambda_2^{-\frac{1}{2}}, \ldots, \lambda_n^{-\frac{1}{2}})O^T$$

Also, by writing out the seperate definition for each term, we can show:

$$N^{-\frac{1}{2}} N N^{-\frac{1}{2}} = I$$

1.4.3 Summary

So, we have that the 'dot product' between two functions is denoted:

$$(\psi, \phi) = \int \psi^* \phi \, d\tau$$

Orthogonality is the statement:

$$(\boldsymbol{e}_i, \boldsymbol{e}_j) = \delta_{ij}$$

i.e. that any two basis vectors are orthogonal, unless they are the same. In another notation, this is actually:

$$(\boldsymbol{e}_i)_a (\boldsymbol{e}_j)_a = \delta_{ij}$$

Where the index a runs over the components of the basis vector $\boldsymbol{e}_i$.

1.5 Function Spaces

Here, we have a continuous label, as opposed to the previous discrete label.

1.5.1 Example: Fourier Basis

Consider the following basis vector, for a function $f : \mathbb{R} \mapsto \mathbb{C}$:

$$\phi_k(x) = \frac{1}{\sqrt{2\pi}} e^{ikx}$$

Our analogy to the discrete case is that any function can be made out of a sum of basis states, each multiplied by some amplitude. Hence, for the continuous basis, we integrate:

$$f(x) = \int_{-\infty}^{\infty} \phi_k(x) \tilde{f}(k) dk \tag{1.3}$$

The expansion coefficient (the 'amplitude') is just the Fourier transform:

$$\tilde{f}(k) = \int_{-\infty}^{\infty} \phi_k^*(x) f(x) dx \tag{1.4}$$

So, putting (1.4) into (1.3), noting that all 'x's in (1.4) are not the same as those in (1.3), hence, we put primes on them to distinguish:

$$f(x) = \int_{-\infty}^{\infty} \phi_k(x) \left(\int_{-\infty}^{\infty} \phi_k^*(x') f(x') dx' \right) dk$$

Rearranging slightly:

$$f(x) = \int_{-\infty}^{\infty} f(x') \left(\int_{-\infty}^{\infty} \phi_k^*(x') \phi_k(x) dk \right) dx'$$

We define the bit in brackets to be the Dirac-δ function:

$$\delta(x - x') \equiv \int_{-\infty}^{\infty} \phi_k^*(x') \phi_k(x) dk \tag{1.5}$$

This is the expression of completeness.
If we now substitute our basis function $\phi_k(x)$ in:

$$\delta(x - x') = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x-x')} dk$$

Thus, we see that we have the interpretation that the delta-function is the *fourier transform of 1.*

Let us do the other interesting thing: put (1.3) into (1.4), changing all k's in (1.3) to primed:

$$\tilde{f}(k) = \int_{-\infty}^{\infty} \phi_k^*(x) \left(\int_{-\infty}^{\infty} \phi_{k'}(x) \tilde{f}(k') dk' \right) dx$$

Again, let us rearrange the integrals:

$$\tilde{f}(k) = \int_{-\infty}^{\infty} \tilde{f}(k') \left(\int_{-\infty}^{\infty} \phi_{k'}(x) \phi_k^*(x) dx \right) dk'$$

Again, the bracketed part is the delta-function:

$$\delta(k - k') = \int_{-\infty}^{\infty} \phi_{k'}(x) \phi_k^*(x) dx$$

This is the equivalent of the *orthonormality relation.* So, we see that unless the indices on the basis are the same, we get zero out. We have not yet gone through the properties of the delta-function, but a simple statement to make is that $\delta(0) = 1, \delta(x \neq 0) = 0$.

1.5.2 The δ Function

The dirac delta is defined by:

$$f(x) = \int_{-\infty}^{\infty} f(x') \delta(x - x') dx' \tag{1.6}$$

Basically, the integral sweeps over all possible values of x', but the only one that contributes is the one at $x' = x$, which is the result of the integral.

We can think of the delta-function as an infinite spike at the origin, whose area is unity. That is:

$$\int_{-\infty}^{\infty} \delta(x - x') dx' = 1$$

So, how about constant within the delta-function? Some useful relations:

Consider:

$$\delta(a(x - x')) = \frac{1}{|a|} \delta(x - x')$$

Let us then evaluate:

$$\int_{-\infty}^{\infty} f(x') \delta(a(x - x')) dx'$$

Let us change variables:

$$y = ax \quad y' = ax'$$

So:

$$\int_{-\infty \times a}^{\infty \times a} f(\tfrac{y'}{a}) \delta(y - y') dy'/a$$

Notice, if $a < 0$, then the integral is negative; so, let us take the modulus of a. Hence:

$$\frac{1}{|a|}\int_{-\infty}^{\infty} f(\tfrac{y'}{a})\delta(y-y')dy' = \frac{1}{|a|}f(\tfrac{y}{a}) = \frac{1}{|a|}f(x)$$

Thus:

$$\int_{-\infty}^{\infty} f(x')\delta(a(x-x'))dx' = \frac{1}{|a|}f(x) \tag{1.7}$$

So, let us consider:

$$\delta(g(x))$$

So, we need to find the zeros of $g(x)$. Let them be at x_i; as we know that the delta-function is non-zero when its argument is zero. So, we have that $g(x_i) = 0$. Let us now do a Taylor expansion about the zero:

$$g(x_i + \epsilon) = g(x_i) + (x - x_i)g'(x_i) + \ldots$$

Note, the first term is zero, by definition. Hence:

$$g(x) \approx (x - x_i)g'(x_i)$$

Near a zero. So, we have that:

$$\delta(g(x)) = \delta(g'(x_i)(x - x_i))$$

Near a zero. This is just like we had before, except we have more than one place the argument of the delta-function is zero. So, we must add the contributions up from all the zeros. Hence:

$$\delta(g(x)) = \sum_i \frac{\delta(x - x_i)}{|g'(x_i)|} \tag{1.8}$$

Where x_i is a zero of the function $g(x)$.

Example Calculate:

$$\int_{-\infty}^{\infty} f(x)\delta(x^2 - c^2t^2)\,dx'$$

We have that $g(x) = x^2 - c^2t^2$, and hence its zeros are at $x = \pm ct$. So, $g'(x) = 2x$. Thus, we have that the integral above is just:

$$\begin{aligned}\int_{-\infty}^{\infty} f(x)\delta(x^2 - c^2t^2)\,dx' &= \frac{1}{|2ct|}\left\{\int_{-\infty}^{\infty} f(x)\delta(x - ct)\,dx' + \int_{-\infty}^{\infty} f(x)\delta(x + ct)\,dx'\right\} \\ &= \frac{1}{2c|t|}\{f(ct) + f(-ct)\}\end{aligned}$$

Where the factor of $\frac{1}{2c|t|}$ has come from evaluating $|g'(x)|$ at each zero of $g(x)$.

1.5.3 Normalisable Function Spaces

A function space which is normalisable is called a *Hilbert space* $\mathcal{H}$.

The Fourier basis is not normalisable; so is not a Hilbert space.

We must have a discrete basis to have a Hilbert space:

$$(\phi_n, \phi_m) = \delta_{nm} = \int_{-\infty}^{\infty} \phi_n^* \phi_m \, dx$$

Completeness would imply:

$$\psi(x) = \sum_n c_n \phi_n(x) \quad \forall \psi(x) \in \mathcal{H}$$

So:

$$(\phi_m, \psi) = \sum c_n (\phi_m, \phi_n) = \sum c_n \delta_{nm} = c_m$$

That is:

$$c_m = (\phi_m, \psi) \tag{1.9}$$

Putting this expression for the coefficients into that for completeness:

$$\psi(x) = \sum_n \int_{-\infty}^{\infty} \phi_n^*(x') \psi(x') \phi_n(x) dx'$$

Which has the dirac-delta expression within:

$$\psi(x) = \int_{-\infty}^{\infty} \delta(x - x') \psi(x') dx'$$

Where:

$$\delta(x - x') = \sum \phi_n^*(x') \phi_n(x)$$

Which is the statement of completeness.

An example in QM is that given two solutions to the TDSE, a third is found by a linear superposition: $\psi(x,t) = \lambda\psi_1 + \mu\psi_2$. We also know that the eigenfunctions of any Hermitian operator (such as the Hamiltonian) form a complete set. So:

$$\psi(x,t) = \sum c_n(t) \phi_n(x)$$

For example, for the SHO:

$$\phi(n(x) = e^{-\frac{x^2}{2b^2}} H_n(\tfrac{x}{b}) \quad E_n = (n + \tfrac{1}{2})\hbar\omega$$

Then:

$$\psi(x,t) = \sum c_n e^{-i(n+\frac{1}{2})\hbar\omega} \phi_n(x)$$

2 Operators, Eigenvectors & Eigenvalues

2.1 Linear Operators

An operator L is linear if, upon acting on $\boldsymbol{a}, \boldsymbol{b} \in V$, we get:

$$L(\lambda \boldsymbol{a} + \mu \boldsymbol{b}) = \lambda(L\boldsymbol{a}) + \mu(L\boldsymbol{b})$$

For $\lambda, \mu \in S$. Examples of such operators:

- Matrix multiplication by a fixed matrix;
- Differentiation:
$$Lf(x) = \frac{d}{dx} f(x)$$
- Integration:
$$\begin{aligned} (L_1 f)(x) &= \int_0^x f(x')dx' \\ (L_2 f)(x) &= \int_0^1 G(x, x')f(x')dx' \end{aligned}$$

2.1.1 Domain, Codomain & Range

If L maps a function f onto g, such that $Lf = g$, then the vector space of all functions f is called the *domain* of L; and the vector space of all functions g is the *codomain* of L. The functions mapped by Lf are called the *range*.

So, the range is a subset of the codomain. If L is invertible, then typically, the range is the same as the codomain.

Examples Consider the operator $L : \mathbb{C}^n \mapsto \mathbb{C}$, where $L(\boldsymbol{a}) \equiv (\boldsymbol{b}, \boldsymbol{a})$. Then, the range is the same as the codomain; which is $\mathbb{C}$. That is, any complex number can come out of the operation.

Consider the operator:

$$L = \begin{pmatrix} 3 & 2 & 1 \\ 6 & 4 & 2 \end{pmatrix}$$

Then, its domain is $\mathbb{R}^3$, codomain $\mathbb{R}^2$. That is $L : \mathbb{R}^3 \mapsto \mathbb{R}^2$. It maps from the space of 3-vectors to the space of 2-vectors. Infact, its range is not the entire space $\mathbb{R}^2$, due to the form of L. The range is $\lambda \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, where $\lambda \in \mathbb{R}$. To see this, just consider the action of the operator on a vector.

Another example is the 3D grad ∇. It maps a space of scalar fields to a space of vector fields.

2.1.2 Matrix Representation of Linear Operators

Consider some operator L which transforms between the domain and codomain, each having a separate set of basis vectors. That is:

$$L : \{\boldsymbol{e}_i\} \mapsto \{\boldsymbol{u}_i\}$$

So, expanding a general vector $\boldsymbol{x}$ in the domain, and $\boldsymbol{y}$ in the codomain, in terms of some coefficient and its relevant basis:

$$\boldsymbol{x} = \sum_i x_i \boldsymbol{e}_i \qquad \boldsymbol{y} = \sum_j y_j \boldsymbol{u}_j$$

So, our transformation:

$$\boldsymbol{y} = L\boldsymbol{x}$$

That is:

$$\begin{aligned} \sum_j y_j \boldsymbol{u}_j &= L \sum_i x_i \boldsymbol{e}_i \\ &= \sum_i x_i L\boldsymbol{e}_i \end{aligned}$$

By linearity. Now, dotting both sides with $\boldsymbol{u}_j$:

$$\sum_j y_j \boldsymbol{u}_j \cdot \boldsymbol{u}_j = \sum_i x_i L\boldsymbol{e}_i \cdot \boldsymbol{u}_j$$

Which is just:

$$y_j = \sum_i x_i L_{ji}$$

Where:

$$L_{ji} \equiv (\boldsymbol{u}_j, L\boldsymbol{e}_i) \tag{2.1}$$

If the bases are orthonormal.

Example Find the matrix representation of $\frac{d}{dx}$ on the space of real functions, on the interval $[-\pi, \pi]$.

Now, since the domain and codomain coincide, we can use the same bases are the same. Let us use the Fourier basis: $1, \{\cos nx, \sin nx\}_{n=1}^{\infty}$. So, let us write the action of the operator on our basis:

$$\frac{d}{dx}\begin{pmatrix} 1 \\ \sin x \\ \cos x \\ \sin 2x \\ \cos 2x \\ \vdots \end{pmatrix} = \begin{pmatrix} 0 \\ \cos x \\ -\sin x \\ 2\cos 2x \\ -2\sin 2x \\ \vdots \end{pmatrix} = M^T \begin{pmatrix} 1 \\ \sin x \\ \cos x \\ \sin 2x \\ \cos 2x \\ \vdots \end{pmatrix}$$

So, the matrix representation of the differential operator is, by inspection of the above:

$$M^T = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \dots & \\ 0 & 0 & 1 & 0 & \dots & & \\ 0 & -1 & 0 & 0 & \dots & & \\ 0 & 0 & 0 & 0 & 2 & 0 & \dots \\ 0 & 0 & 0 & -2 & 0\dots & & \end{pmatrix}$$

To just fill in a few of the infinite elements.

As another example, we consider the TISE. In 1D, we have the eiegenequation:

$$\hat{H}\psi(x) = E\psi(x)$$

Suppose we use a basis $\{\phi_i\}_{i=1}^{\infty}$, where $\phi_i \in \mathcal{H}$. Hence:

$$\psi = \sum_j c_j \phi_j$$

We get the matrix from:

$$\sum_j H_{ij} c_j = E c_i$$

Where:

$$H_{ij} = (\phi_i, H\phi_j) = \int \phi_i^* \hat{H} \phi_j \, dx$$

Notice though, that $H_{ij} = H_{ij}^\dagger$.

2.1.3 Adjoint and Hermitian Operators

The Hermitian, or adjoint, coinjugate of matrix is the complex conjugate of its transpose. That is:

$$(M_{ij})^\dagger = (M_{ji})^*$$

For example:

$$\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}^\dagger = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} = - \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

Which we term 'anti-Hermitian'. Also, consider:

$$\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}^\dagger = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$$

Which we term an 'Hermitian operator'.

Notice, the inner-product:

$$(u, v) = u^\dagger v$$

So that, we write it as a matrix product:

$$(u, v) = (u_1^*, u_2^*, \dots, u_n^*) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = u^\dagger v$$

The Hermitian conjugate of a matrix can be defined as:

$$(\boldsymbol{a}, M\boldsymbol{b}) = (M^\dagger \boldsymbol{a}, \boldsymbol{b}) \tag{2.2}$$

So that, if the codomain is C, and domain D:

$$M : D \mapsto C \qquad M^\dagger : C \mapsto D$$

Now, we can see this:

$$\begin{aligned}(\boldsymbol{a}, M\boldsymbol{b}) &= a_i^* M_{ij} b_j \\ &= a_i^* (M_{ji}^\dagger)^* b_j \\ &= (a_i M_{ji}^\dagger)^* b_j \\ &= (M_{ji}^\dagger a_i)^* b_j \\ &= (M^\dagger \boldsymbol{a}, \boldsymbol{b})\end{aligned}$$

An hermitian operator (which is the same as self-adjoint) is such that:

$$M^\dagger = M \tag{2.3}$$

So that the domain is the same as the codomain.

2.2 Eigenvalue Problems

An eigenvalue problem is such that:

$$L\boldsymbol{a} = \lambda \boldsymbol{a}$$

Theorems For a Hermitian operator, all eigenvalues are real, and eigenfunctions corresponding to different eigenvalues are orthogonal.

2.2.1 Proof: Real Eigenvalues of Hermitian Operators

So, we have that:

$$(\boldsymbol{a}, L\boldsymbol{a}) = (\boldsymbol{a}, \lambda \boldsymbol{a}) = \lambda (\boldsymbol{a}, \boldsymbol{a})$$

And, also:

$$(L\boldsymbol{a}, \boldsymbol{a}) = (\lambda \boldsymbol{a}, \boldsymbol{a}) = \lambda^* (\boldsymbol{a}, \boldsymbol{a})$$

Now, by defintion, $(\boldsymbol{a}, L\boldsymbol{a}) = (L\boldsymbol{a}, \boldsymbol{a})$ for hermitian operators. Hence, let us subtract the above two equations:

$$0 = (\lambda - \lambda^*)(\boldsymbol{a}, \boldsymbol{a})$$

Hence, as $(\boldsymbol{a}, \boldsymbol{a}) > 0$, we have:

$$\lambda = \lambda^* \tag{2.4}$$

2.2.2 Proof: Orthogonal Eigenfunctions

Here, let us start with two eigenequations:

$$L\boldsymbol{a} = \lambda\boldsymbol{a} \qquad L\boldsymbol{b} = \mu\boldsymbol{b}$$

Hence:

$$(\boldsymbol{b}, L\boldsymbol{a}) = \lambda(\boldsymbol{b}, \boldsymbol{a}) \qquad (L\boldsymbol{b}, \boldsymbol{a}) = \mu(\boldsymbol{b}, \boldsymbol{a})$$

However, due to Hermiticity, the above two equations are equal. Hence:

$$0 = (\lambda - \mu)(\boldsymbol{b}, \boldsymbol{a})$$

Hence, if $\lambda \neq \mu$, we have:

$$(\boldsymbol{b}, \boldsymbol{a}) = 0 \tag{2.5}$$

2.2.3 Example: Diagonalise A Matrix

Find the eigenvalues/vectors of:

$$M = \begin{pmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{pmatrix}$$

So, we find its eigenvalues by the standard determinant method:

$$\begin{vmatrix} 4-\lambda & 1 & 1 \\ 1 & 4-\lambda & 1 \\ 1 & 1 & 4-\lambda \end{vmatrix} = 0$$

From which we get to the charateristic equation:

$$(\lambda - 3)^2(\lambda - 6) = 0$$

From which we see that 2 of the 3 eigenvalues are degenerate.
We shall first find the non-denegerate eigenvalues' eigenvector, by solving the following:

$$\begin{pmatrix} 4-6 & 1 & 1 \\ 1 & 4-6 & 1 \\ 1 & 1 & 4-6 \end{pmatrix} \begin{pmatrix} e_3^1 \\ e_3^2 \\ e_3^3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Which has solution:

$$\boldsymbol{e}_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

For $\lambda = 3$, we have to solve:

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} e^1 \\ e^2 \\ e^3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Now, we can do this by choosing two vectors which are orthogonal to $\boldsymbol{e}_3$. Such as:

$$\boldsymbol{e}_1 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \qquad \boldsymbol{e}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

Now, to diagonalise the matrix , we form $M = O^T diag(\lambda) O$; where O is the matrix of eigenvectors:

$$O = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \end{pmatrix} \qquad \Rightarrow \qquad O^T = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{pmatrix}$$

Infact, we have written the columns as the components of $\boldsymbol{e}_3, \boldsymbol{e}_1, \boldsymbol{e}_2$, respectively. We have that $diag(\lambda)$ has components which are just the eigenvalues:

$$diag(\lambda) = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Hence:

$$M = O^T diag(\lambda) O = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{pmatrix} \begin{pmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \end{pmatrix}$$

2.2.4 Generalised Eigenvalue Problem

$$Ly(x) = \lambda \rho(x) y(x) \tag{2.6}$$

Where we call $\rho(x)$ the *weight function*.
If $u(x)$ and $v(x)$ are two eigenfunctions for different eigenvalues, and if L is hermitian, then:

$$(u, v)_\rho \equiv \int \rho(x) u^*(x) v(x)\, dx = 0$$

Note, hermiticity is defined with respect to (u, v). That is:

$$(v, Lu) = (v, \lambda \rho u) = \lambda (v, u)_\rho$$

and:

$$(Lv, u) = (\mu \rho v, u) = \mu (v, u)_\rho$$

So that, as $(Lv, u) = (v, Lu)$, we have:

$$(\lambda - \mu)(v, u)_\rho = 0$$

2.3 Sturm-Lioville Equations

If we have a generalised eigenvalue equation:

$$Ly(x) = \lambda\rho(x)y(x)$$

Where L is of the Sturm-Lioville (SL) form. That is:

$$L = -\frac{d}{dx}\left(p(x)\frac{d}{dx}\right) + q(x) \tag{2.7}$$

Where $p(x) > 0$. So, we have that the SL equation is therefore:

$$-\frac{d}{dx}\left(p(x)\frac{d}{dx}y(x)\right) + q(x)y(x) - \lambda\rho(x)y(x) = 0 \tag{2.8}$$

That is, expanding out the derivatives:

$$-p(x)\frac{d^2y}{dx^2} - \frac{dp}{dx}\frac{dy}{dx} + q(x)y(x) - \lambda\rho(x)y(x) = 0 \tag{2.9}$$

2.3.1 Bringing Equations to SL Form

Let us start with a very general second-order differential equation:

$$y''(x) + \alpha(x)y'(x) + \beta(x)y(x) + \lambda\tau(x)y(x) = 0$$

From now on, we will supress the x-argument. Now, multiply the above by $-p(x) \equiv -p$. That is:

$$-py'' - p\alpha y' - p\beta y - p\lambda\tau y = 0$$

Let us now compare the coefficients of this equation, with that of the SL equation (2.9). We hence see:

$$-p\alpha = -\frac{dp}{dx} \equiv -p'$$

Thus:

$$\alpha = \frac{p'}{p}$$

That is:

$$\alpha = \frac{d\ln p}{dx}$$

Hence:

$$\ln p = \int^x \alpha dx'$$

Thus we have that:

$$p(x) = e^{\int^x \alpha dx'} \tag{2.10}$$

And that $p(x) > 0$. Also, by comparison, we have that $p\tau = \rho$. Hence, $\tau > 0$.

Hence, we have found that if $\tau(x) > 0$, and if we multiply the equation by $p(x)$, which we find above; we can get any such differential equation into SL form.

2.3.2 A Useful Result

Now, let us think of two functions $u(x), v(x)$, where $x \in [a, b]$. Now, let us write:

$$vLu - uLv$$

Let us write each down, long hand:

$$\begin{aligned} vLu &= -v\frac{d}{dx}p\frac{d}{dx}u + vqu = -v[pu']' + vqu \\ uLv &= -u\frac{d}{dx}p\frac{d}{dx}v + uqv = u[pv']' + uqv \end{aligned}$$

Hence:

$$vLu - uLv = -v[pu']' + u[pv']'$$

That is, by the product rule:

$$vLu - uLv = -vpu'' - vp'u' + upv'' + up'v'$$

Now, if let us both add and subtract $v'pu'$; that is:

$$vLu - uLv = -vpu'' - vp'u' + upv'' + up'v' - v'pu' + v'pu'$$

So, by inspection, we see that this is just:

$$vLu - uLv = [-vpu' + upv']'$$

That is:

$$vLu - uLv = [p(uv' - vu')]' \tag{2.11}$$

Note, sometimes, the expression $uv' - vu'$ is sometimes called the 'Wronskian', and is actually just a determinant:

$$uv' - vu' = \begin{vmatrix} u & v \\ u' & v' \end{vmatrix}$$

However, getting back to the problem. We hence have:

$$\int_a^b (vLu - uLv)dx = [p(uv' - vu')]_a^b \tag{2.12}$$

Now, if the operator L is Hermitian, we hence have:

$$\int_a^b (vLu - uLv)dx = 0$$

That is then just:

$$p(b)[uv' - vu'](b) - p(a)[uv' - vu'](a) = 0$$

If $p(a) = p(b) = 0$, then we call a, b *singular points.*

2.3.3 Second Solutions

Now, our eigenvalue problem $Lu = \lambda\rho u$ is second order (since L is a second order differential operator). Hence, there is two solutions to such an equation, for the same value of λ. Suppose that it is $Lv = \lambda\rho v$. So:

$$Lu = \lambda\rho u \qquad Lv = \lambda\rho v$$

Now, multiplying the first by v, and the second by u gives:

$$vLu = v\lambda\rho u \qquad uLv = u\lambda\rho v$$

Subtracting the two:

$$vLu - uLv = v\lambda\rho u - u\lambda\rho v = 0$$

That is:

$$vLu - uLv = 0$$

Now, let us use our derived relation:

$$[p(uv' - vu')]' = vLu - uLv$$

Which is just 0, for us. Hence, that is:

$$p(uv' - vu') = c$$

Where c is some constant. Hence:

$$uv' - vu' = \frac{c}{p(x)}$$

Now, as u is known, this is just a first order ODE, for v. Let us substitute $v = uw$, for some function w. Hence, the above is just:

$$uuw' + uu'w - uwu' = \frac{c}{p}$$

That is just:

$$uuw' = u^2w' = \frac{c}{p} \quad \Rightarrow \quad w' = \frac{c}{pu^2}$$

Hence:

$$w(x) = c\int^x \frac{1}{p(x')u(x')^2}dx'$$

Now, these solutions do not exist if $p(x) = 0$. That is, the solutions diverge. However, we have shown that $p(x) = 0$ on the boundaries. So that, we had that the SL operator was defined on $[a, b]$, and that $p(a) = p(b) = 0$. Thus, L is non-Hermitian for such solutions.

2.3.4 Classifying Singular Points

For some second order ODE:

$$y''(x) + P(x)y'(x) + Q(x)y(x) = 0$$

If, at some $x = x_0$, $P(x_0)$ or $Q(x_0)$ diverges, then x_0 is called a *regular singular point*. We also must require that $(x - x_0)P(x)$ and $(x - x_0)^2Q(x)$ remain finite, as $x \to x_0$. If, however, $P(x)$ diverges

faster than $1/(x-x_0)$, or $Q(x)$ faster than $1/(x-x_0)^2$, then x_0 is an *irregular singularity.* So, suppose we had:

$$y'' + \frac{x^2}{x-1}y' + \frac{x}{x-1}y = 0$$

Then $x = 1$ is a regular singular point. We see this by noting that both:

$$(x-1)\frac{x^2}{x-1} \qquad (x-1)\frac{x}{x-1}$$

Are finite (the factors just trivially cancel out).

Considering:

$$y'' + \frac{x}{(x-1)^2}y' + \frac{x}{(x-1)^2}y = 0$$

Then $x = 1$ is an irregular singularity, as neither:

$$(x-1)\frac{x}{(x-1)^2} \qquad (x-1)\frac{x}{(x-1)^2}$$

Are finite.

2.4 Series Solutions

2.4.1 Example: Quantum Harmonic Oscillator

Consider the Schrodinger equation for the 1D QM harmonic oscillator:

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x) + \frac{1}{2}m\omega^2x^2\psi(x) = E\psi(x)$$

Now, the solutions to such an equation are square-integrable functions. That is:

$$\int_{-\infty}^{+\infty} |\psi^2(x)|\, dx < \infty$$

And, the way we expressed such a function before, was to say:

$$\psi(x) \in \mathcal{L}^2(\mathbb{R})$$

Now, to solve the equation, let us use the following variables, to make everything look simpler:

$$b \equiv \left(\frac{\hbar^2}{m\omega}\right)^{1/2} \qquad E_\omega \equiv \hbar\omega$$

Then, we have that:

$$z \equiv \frac{x}{b} \quad E = \lambda E_\omega$$

Under these variables, our equation looks like:

$$\frac{d^2\psi(z)}{dx^2} + (2\lambda - z^2)\psi(z) = 0$$

Let us initialy asymptotically analyse this. That is, let $z \to \infty$. So, we are able to ignore the λ term. That is:

$$\frac{d^2\psi(z)}{dx^2} - z^2\psi(z) = 0$$

From which we can guess a solution of the form:

$$\psi(z) \approx e^{\pm\frac{z^2}{2}}$$

We must, however, take only the negative sign, as we must have that the wavefunction decays at infinity. So, let us suppose that the total solution is of the following form:

$$\psi(z) = f(z)e^{-\frac{z^2}{2}}$$

Where $f(z)$ can be neglected as a rising-power. Putting this into our modified Schrodinger equation easily gets us to:

$$f'' - 2zf' + (2\lambda - 1)f = 0$$

Now, let us show that this is of SL form:

$$-\frac{d}{dx}\left(p(x)\frac{d}{dx}y(x)\right) + q(x)y(x) - \lambda\rho(x)y(x) = 0$$

So, following our previous discussion, $\alpha(x)$, which is the coefficient of f', allows us to find $p(x)$; thus:

$$p(x) = \exp\int^x \alpha dx'$$

We see that 'our α' is just:

$$\alpha = -2z \quad \Rightarrow \quad p(z) = \exp\int^z (-2z')dz' = e^{-2z^2}$$

So, to bring our equation to SL form, we multiply our equation by $p(z)$, to give:

$$e^{-2z^2}f'' - 2ze^{-2z^2}zf' + (2\lambda - 1)e^{-2z^2}f = 0$$

Now, this is just (done by inspection):

$$[e^{-2z^2}f']' + (2\lambda - 1)e^{-2z^2}f = 0$$

Which is of SL-form; where $z = \pm\infty$ are regular singularities.

To solve the equation:

$$f'' - 2zf' + (2\lambda - 1)f = 0$$

Let us use a series solution. That is, take:

$$f(z) = \sum_{n=0}^{\infty} c_n z^n$$

So, substituting this in, gives:

$$\sum_{n=0}^{\infty} c_n n(n-1)z^{n-2} - 2\sum_{n=0}^{\infty} znz^{n-1} + (2\lambda - 1)\sum_{n=0}^{\infty} c_n z^n = 0$$

That is:

$$\sum_{n=0}^{\infty} c_n n(n-1)z^{n-2} - 2\sum_{n=0}^{\infty} nz^n + (2\lambda - 1)\sum_{n=0}^{\infty} c_n z^n = 0$$

Now, let us note that, for the first expression, the first two terms will be zero; if $n = 0, 1$, then we have 0 both times. So; let us write:

$$\sum_{n=2}^{\infty} c_n n(n-1)z^{n-2} - 2\sum_{n=0}^{\infty} nz^n + (2\lambda - 1)\sum_{n=0}^{\infty} c_n z^n = 0$$

Let us then define $m \equiv n - 2$; and use that as a summation index (replacing the latter two as well: it makes no difference). Hence:

$$\sum_{m=0}^{\infty} c_{m+2}(m+2)(m+1)z^m - 2\sum_{m=0}^{\infty} mz^m + (2\lambda - 1)\sum_{m=0}^{\infty} c_m z^m = 0$$

That is just:

$$\sum_{m=0}^{\infty} [c_{m+2}(m+2)(m+1) + c_m(-2m + 2\lambda - 1)]z^m = 0$$

So, by linear independence, we have that the coefficient of z^m is zero. Hence:

$$c_{m+2} = \frac{c_m(2m - (2\lambda - 1))}{(m+2)(m+1)}$$

So, we have a recursion relationship to find the coefficients in the solution:

$$c_0 \to c_2 \to c_4 \to \dots$$

Now, thus must terminate, as $f(z)$ is well-behaved. Let us have that the series terminates at some $m = \ell$. That is, so that $c_{\ell+2} = 0$. So that:

$$2\lambda - 1 = 2\ell$$

Hence, we have that:

$$E_{ell} = (\ell + \tfrac{1}{2})\hbar\omega$$

And that:

$$f_\ell = H_\ell(z)$$

The Hermite polynomials.:

$$H_0(z) = 1 \quad H_1(z) = z \quad H_2(z) = 1 - z^2$$

Hence, we have that our total wavefunction has solution:

$$\psi(x) = H_n(\tfrac{x}{b})e^{-\frac{1}{2}(\frac{x}{b})^2}$$

2.5 Legendre Polynomials

Consider Legendre's equation, put into SL form:

$$[(1-x^2)y'(x)]' + \lambda y(x) = 0 \tag{2.13}$$

It is, obviously, singular at $x = \pm 1$. Also, we typically have that $x = \cos\theta$. Then, this range is an obvious choice to make: $x \in [-1, 1]$. So, let us solve using a power series solution, about the origin:

$$y(x) = \sum_{n=0}^{\infty} c_n x^n$$

Then, it is not too hard to get to the recursive relationship:

$$c_{m+2} = \frac{m(m+1)-\lambda}{(m+1)(m+2)} c_m$$

Now, for the series to terminate, we require that $m(m+1) = \lambda$, for some m. Hence, we have a polynomial if $\lambda = n(n+1), n \in \mathbb{N}$. The solutions are called Legendre polynomials $P_n(x)$; and satisfy the orthogonality relation:

$$\int_{-1}^{1} P_n(x) P_m(x) dx = 0 \quad n \neq m$$

2.5.1 Generating Function

For the Legendre polynomials, we suppose that a function will generate the polynomials, then we will check that it does:

$$f(x,t) = \sum_{n=0}^{\infty} t^n P_n(x) \tag{2.14}$$

$$= \frac{1}{\sqrt{1-2tx+t^2}} \tag{2.15}$$

Now, let us find both the x- and t-derivatives of this expression:

$$\frac{d}{dx}: \quad \sum_{n=0}^{\infty} t^n P_n'(x) = t(1-2tx+t^2)^{-3/2} \tag{2.16}$$

$$\frac{d}{dt}: \quad \sum_{n=0}^{\infty} n t^{n-1} P_n(x) = (x-t)(1-2tx+t^2)^{-3/2} \tag{2.17}$$

Now, let us express the first in the following way:

$$\begin{aligned} t(1-2tx+t^2)^{-3/2} &= t(1-2tx+t^2)^{-1}(1-2tx+t^2)^{-1/2} \\ &= t(1-2tx+t^2)^{-1} \sum_n t^n P_n(x) \end{aligned}$$

Hence:

$$\sum_n t^n P_n'(x) = t(1-2tx+t^2)^{-1} \sum_n t^n P_n(x)$$

That is:

$$t\sum_n t^n P_n(x) = (1 - 2tx + t^2)\sum_n t^n P'_n(x)$$

Which is just:

$$\sum_n t^{n+1} P_n(x) = \sum_n (t^n - 2xt^{n+1} + t^{n+2}) P'_n(x)$$

Solve solve this (by equating coefficients of like powers of t), we note the following:

$$\sum_n (t^n - 2xt^{n+1} + t^{n+2}) P'_n(x) = \sum_n t^n P'_n(x) - 2x\sum_n t^{n+1} P'_n(x) + \sum_n t^{n+2} P'_n(x)$$

Where, in each sum above, n starts from zero. Now, we may relabel each summation variable thus:

$$\Rightarrow \quad t^{n+1} P'_{n+1}(x) - 2xt^{n+1} P'_n(x) + t^{n+1} P'_{n-1}(x)$$

Hence, now we equate powers of t, giving:

$$P_n(x) = P'_{n+1}(x) - 2xP'_n(x) + P'_{n-1}(x) \tag{2.18}$$

Now, let us go back a bit. If we mutiply (2.16) by $(x - t)$, and (2.17) by t, then they should be equal. Hence:

$$\sum_n (x - t) t^n P'_n(x) = \sum_n n t^n P(x)$$

That is:

$$xt^n P'_n - t^{n+1} P'_n - nt^n P_n = 0$$

Again, shifting the summation variable, to be able to equate coefficients:

$$xt^n P'_n - t^n P'_{n-1} - nt^n P_n = 0$$

Hence:

$$xP'_n - P'_{n-1} = nP_n \tag{2.19}$$

Now, let us write down (2.18) and (2.19), together:

$$\begin{aligned} P_n &= P'_{n+1} - 2xP'_n + P'_{n-1} \\ nP_n &= xP'_n - P'_{n-1} \end{aligned}$$

From these, eliminate P'_{n-1}, giving:

$$(n + 1)P_n = P'_{n+1} - xP'_n$$

Now, in the above, let $n \to n + 1$:

$$nP_{n-1} = P'_n - xP'_{n-1}$$

Now, add (2.19) times x to the above; giving:

$$x^2 P'_n - xP'_{n-1} + nP_{n-1} = nxP_n + P'_n - xP'_{n-1}$$

Tidying up:

$$(1-x^2)P_n' = n(P_{n-1} - xP_n)$$

Differentiate both sides, with respect to x:

$$[(1-x^2)P_n']' = n(P_{n-1}' - P_n - xP_n')$$

Now, using (2.19) again: $xP_n' - P_{n-1}' = nP_n$, on the RHS

$$[(1-x^2)P_n']' = -n(n+1)P_n$$

Hence, we have shown:

$$[(1-x^2)P_m'(x)]' = -m(m+1)P_m(x)$$

Hence completing the proof.

Let us compute $f(1,t)$:

$$\begin{aligned} f(1,t) &= \sum_{n=0}^{\infty} P_n(1)t^n \\ &= \frac{1}{(1-2tx+t^2)^{1/2}} \\ &= \frac{1}{1-t} \\ &= \sum_n t^n \end{aligned}$$

Hence, we see that $P_n(1) = 1$.

2.5.2 Application: Expand $\frac{1}{|r_1 - r_2|}$

We often need to expand:

$$\frac{1}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|}$$

Now, we can do so in terms of orthogonal functions, in the angle between the two vectors. Let us assume $r_1 > r_2$. So:

$$\begin{aligned} \frac{1}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} &= \frac{1}{\sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos\theta}} \\ &= \frac{1}{r_1} \frac{1}{\sqrt{1 + (\frac{r_2}{r_1})^2 - 2\frac{r_2}{r_1}\cos\theta}} \end{aligned}$$

Now, we identify the expression with the generating function of the Legendre polynomials, as:

$$t = \frac{r_2}{r_1} \qquad x = \cos\theta$$

Hence, we have:

$$\frac{1}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} = \frac{1}{r_1} \sum_{n=0}^{\infty} \left(\frac{r_2}{r_1}\right)^n P_n(\cos\theta) \tag{2.20}$$

For $r_2 > r_1$. We can obviously rearrange the indices if the reverse is true.

2.5.3 Normalisation

So, we wish to compute:

$$\int_{-1}^{1}\left(\sum_n t^n P_n(x)\right)^2 dx$$

So, we have:

$$\begin{aligned}\int_{-1}^{1}\left(\sum_n t^n P_n(x)\right)^2 dx &= \int_{-1}^{1}\frac{1}{1-2tx+t^2}dx\\ &= \frac{1}{t}\ln\left(\frac{1+t}{1-t}\right)\\ &= \frac{1}{t}\sum_{m=1}\left\{\frac{1}{m}t^m-\frac{1}{m}(-t)^m\right\}\\ &= \sum_{m=0}^{\infty}\frac{2t^m}{2m+1}\end{aligned}$$

Hence, going back, we have:

$$\sum_n t^{2n}\int_{-1}^{1}P_n^2(x)dx=\sum_{m=0}^{\infty}\frac{2t^m}{2m+1}$$

That is, we have found the normalisation of the Legendre polynomials:

$$\int_{-1}^{1}P_n^2(x)dx=\frac{2}{2n+1} \tag{2.21}$$

2.6 Gram-Schmidt for Functions

Let us just state the following, then we shall discuss it:

$$y_n=\tilde{y}_n-\sum_{i=0}^{n-1}\hat{y}_i\int\hat{y}_i^*\tilde{y}_n\rho dx \tag{2.22}$$

So, let us say that the set $\{\hat{y}_i\}_{i=0}^{n-1}$ are normalised. The weight function is $\rho(x)$. States can be normalised thus:

$$\hat{y}_n=\frac{y_n}{\left|\int y_n^* y_n\rho dx\right|^{1/2}}$$

So, if we have an non-orthogonal eigenstate $\tilde{y}_n$ (call it our trial), and we wish to make it orthogonal to all states before it; that is, with the set $\{\hat{y}_i\}_{i=0}^{n-1}$, which are orthonormal; we must follow the above iterative procedure. The overlap integral will pick out the "amount" by which the trial eigenstate overlaps the orthogonal set. This overlap is then subtracted. The sum will sweep over all every previous orthogonal eigenstate, subtracting out each overlap. The resulting eigenstate y_n is then, of course, not normalised; but that is easily done, by dividing by the normalisation integral.

As an example to see the iterative process, consider that we have $\hat{y}_0, \hat{y}_1$ as being orthonormal eigenstates. We have some trial eigenstate $\tilde{y}_2$, but it is not orthogonal. Let us find an orthogonal eigenstate y_2:

$$y_2 = \tilde{y}_2 - \hat{y}_0 \int \hat{y}_0^* \tilde{y}_2 \rho dx - \hat{y}_1 \int \hat{y}_1^* \tilde{y}_2 \rho dx$$

And then to normalise the now orthogonal trial eigenstate:

$$\hat{y}_2 = \frac{y_2}{\left| \int y_2^* y_2 \rho dx \right|^{1/2}}$$

And we have hence constructed an orthonormal eigenstate $\hat{y}_2$, from a non-orthogonal, un-normalised trial $\tilde{y}_2$.

3 Green Functions

If we wish to solve a differential equation of the form:

$$(L_x u)(x) = f(x)$$

Plus boundary conditions; where L_x is a linear hermitian operator. Then, to do so, we can always write the solution to the DE, $u(x)$ as:

$$u(x) = \int G(x, x')f(x')dx' \tag{3.1}$$

Where the *Green function* is defined by:

$$L_x G(x, x') = \delta(x - x') \tag{3.2}$$

We shall now supress the x-subscript on the operator: it is clear that it only operates on x, not x'.

To prove the above statement isnt too hard:

$$\begin{aligned} (Lu)(x) &= f(x) \\ \Rightarrow L\int G(x, x')f(x')dx' &= \int LG(x, x')f(x')dx' \\ &= \int \delta(x - x')f(x')dx' \\ &= f(x) \end{aligned}$$

Hence proven. In the first line, we used the linearity of the operator, to be able to bring it inside the integral (which is over x' anyway). Then we used the definitions of the Green function and delta function.

Let us proceed by example.

3.0.1 Example: Electrostatics

Now, from Gauss' law $\nabla \cdot \boldsymbol{E} = \rho/\varepsilon_0$ and $\boldsymbol{E} = -\nabla\Phi$, we can easily derive Poissons equation:

$$\nabla^2\Phi(\boldsymbol{x}) = -\frac{\rho(\boldsymbol{x})}{\varepsilon_0}$$

Note, this is strictly for a static charge distribution, else we would have the $\boldsymbol{A}$ term as well. We have the boundary condition that $\Phi(\boldsymbol{x}) \to 0$ as $|\boldsymbol{x}| \to \infty$.
Now, for a point charge q, at $\boldsymbol{x}'$, we have a charge distribution which is just a delta function:

$$\rho(\boldsymbol{x}) = q\delta(\boldsymbol{x} - \boldsymbol{x}')$$

Note, strictly, we should have written δ^3, but that is understood as its argument has 3 variables. Now, we know that the resulting Poisson equation has the following solution:

$$\nabla^2\Phi(\boldsymbol{x}) = -\frac{q}{\varepsilon_0}\delta(\boldsymbol{x} - \boldsymbol{x}') \quad \Rightarrow \quad \Phi(\boldsymbol{x}) = \frac{q}{4\pi\varepsilon_0}\frac{1}{|\boldsymbol{x} - \boldsymbol{x}'|}$$

Now, upon comparison of the above formalism for Green function, $Lu = f$ is just the Poisson equation. That is, the operation L is just the ∇^2 operator. Hence, from $LG = \delta(x - x')$, we see that the Green function we want, is that satisfying:

$$\nabla^2 G(\boldsymbol{x}, \boldsymbol{x}') = \delta(\boldsymbol{x} - \boldsymbol{x}')$$

We have the form of G, from the above expression for the point charge. That is:

$$G(\boldsymbol{x}, \boldsymbol{x}') = -\frac{1}{4\pi} \frac{1}{|\boldsymbol{x} - \boldsymbol{x}'|}$$

Hence, we have that the solution to Poissons equation is given by:

$$\Phi(\boldsymbol{x}) = \int G(\boldsymbol{x}, \boldsymbol{x}') \left(-\frac{\rho(\boldsymbol{x}')}{\varepsilon_0} \right) d^3x'$$

That is:

$$\Phi(\boldsymbol{x}) = \frac{1}{4\pi\varepsilon_0} \int \frac{\rho(\boldsymbol{x}')}{|\boldsymbol{x} - \boldsymbol{x}'|} d^3x'$$

Which is something we already knew, but we have derived it using Green functions; and in the process, have identified a Green function.

We shall now consider two methods of finding the Green function:

3.1 The Eigenstate Method

Suppose we have the eigenvalue equation:

$$Lu_n(x) = \lambda_n u_n(x)$$

Also suppose that we know the eigenvalues and eigenstates of the operator; and that the eigenstates u_n form a complete orthonormal set. Then, the Green function is given by:

$$G(x, x') = \sum_n \frac{1}{\lambda_n} u_n(x) u_n^*(x') \tag{3.3}$$

Where $\lambda_n \neq 0$. Let us initially show that this Green function works:

$$\begin{aligned} LG(x, x') &= \sum_n \frac{1}{\lambda_n} [Lu_n(x)] u_n^*(x') \\ &= \sum_n u_n(x) u_n^*(x') \\ &= \delta(x - x') \end{aligned}$$

Hence proven. We have used the standard completeness relation.

Now, suppose $\lambda_n = 0$. Let us assume that $\lambda_0 = 0$. Also, by completeness, we have that we can expand any function in terms of eigenstates:

$$u(x) = \sum_{n=0}^{\infty} c_n u_n(x) \qquad f(x) = \sum_n d_n u_n(x)$$

So, let us use these (expanded) functions in the differential equation $Lu = f$; immediately noting that $Lu_n = \lambda_n u_n$:

$$\sum_n \lambda_n c_n u_n(x) = \sum_n d_n u_n(x)$$

That is, obviously:

$$\lambda_n c_n = d_n$$

As a quick aside, note:

$$f(x) = \sum_n d_n u_n \quad \Rightarrow \quad (u_m, f) = \sum_n d_n (u_m, u_n) = \sum_n d_n \delta_{nm} = d_m$$

Hence, we see that:

$$\lambda_n c_n = (u_n, f) = d_n$$

Thus, note that for $\lambda_0 = 0$, we have:

$$(u_0, f) = 0$$

That is, f has no components parallel to the zero mode. That is, f is orthogonal to the zero mode. Hence, we also have that $d_0 = 0$. So, let us expand $u(x)$ again, and inserting our expression $\lambda_n c_n = d_n$:

$$\begin{aligned} u(x) &= \sum_{n=0}^{\infty} c_n u_n(x) \\ &= c_0 u_0(x) + \sum_{n=1}^{\infty} \frac{d_n}{\lambda_n} u_n(x) \\ &= c_0 u_0(x) + \sum_{n=1} (u_n, f) \frac{1}{\lambda_n} u_n(x) \end{aligned}$$

Let us remember that:

$$(u_n, f) = \int u_n^*(x') f(x') dx'$$

And also that:

$$u(x) = \int G(x, x') f(x') dx'$$

Hence, upon inspection of:

$$u(x) = c_0 u_0(x) + \sum_{n=1} \int u_n^*(x') f(x') dx' \frac{1}{\lambda_n} u_n(x)$$

We see that:

$$G(x, x') = c u_0(x) + \sum_{n=1}^{\infty} \frac{1}{\lambda_n} u_n(x) u_n^*(x')$$

CHECK THE CONJUGATED TERM!!

Where:

$$c = \frac{c_0}{\int f(x) dx}$$

3.1.1 Example: Vibrating String

Suppose we have a string, length L, with fixed end points. Consider that it is driven. Assume that we have small displacements.
So, for a 'small bit' of the string, we have a tension $-T(u'(x))$ and $T(u'(x+dx))$ at either end. So, for that small bit of string:

$$\rho dx\ddot{u} = T\frac{du}{dx}(x+dx) - T\frac{du}{dx}(x) - Fdx$$

This results in (putting in the form of the driving force):

$$\rho\frac{\partial^2 u}{\partial t^2} = T\frac{\partial^2 u}{\partial x^2} - F(x)\sin\omega t$$

Now, the system will have a transient phase, which will die away, leaving a steady state $u(x,t) = v(x)\sin\omega t$. Hence:

$$\frac{d^2v}{dx^2} + \frac{\rho\omega^2}{T}v = \frac{F(x)}{T}$$

Which we compress to:

$$\frac{d^2v}{dx^2} + k^2 v = f(x)$$

With boundary conditions that $v(0) = v(L) = 0$.
So, to use the eigenstate method, we must first solve:

$$v_n'' + k^2 v_n = \lambda_n v_n \quad \Rightarrow \quad v_n'' = (\lambda_n - k^2)v_n$$

Let us write this as:

$$v_n'' = -k_n^2 v_n \qquad -k_n^2 \equiv \lambda_n - k^2$$

Which we know the solution to:

$$v_n = \sqrt{\tfrac{2}{L}}\sin k_n x$$

Where $k_n = \frac{n\pi}{L}$, to satisfy the boundary condition. Hence, we see that:

$$\lambda_n = k^2 - k_n^2 \quad n = 1, 2, \ldots$$

Now, we know that from the eigenvalue equation $Lu_n = \lambda_n u_n$, we have the Green function:

$$G(x,x') = \sum_n \frac{1}{\lambda_n} u_n(x) u_n^*(x')$$

With solutions to the differential equation $Lu = f$ being given by:

$$u(x) = \int G(x,x')f(x')dx'$$

Hence, for our problem, we have the Green function:

$$G(x,x') = \sum_{n=1} \frac{2}{L}\frac{\sin k_n x \sin k_n x'}{k^2 - k_n^2}$$

And thus solution:

$$v(x) = \int_0^L G(x, x') f(x') dx'$$

That is:

$$v(x) = \sum_n \frac{2}{L} \int_0^L \frac{\sin k_n x \sin k_n x'}{k^2 - k_n^2} f(x') dx'$$

Notation for which we can compress via:

$$v(x) = \sum_n \sqrt{\frac{2}{L}} d_n \frac{\sin k_n x}{k^2 - k_n^2} \qquad d_n = \int_0^L \sqrt{\frac{2}{L}} \sin k_n x' f(x') dx'$$

3.2 The Continuity Method: Strictly 1D

Here, we solve the differential equation:

$$\left[\frac{d^2}{dx^2} + \alpha(x) \frac{d}{dx} + \beta(x) \right] G(x, x') = \delta(x - x')$$

With some boundary conditions.

So, to do so; first, one solves for $x < x'$ and $x > x'$ separately. Note, this is:

$$LG(x, x') = 0$$

And the integration constants will only be functions of x'.
Then, we fix the integration constants by matching at $x = x'$. We find that G is continuous at $x = x'$, but $\frac{dG}{dx}$ is not, and is discontinuous (depending on the factor infront of $\frac{d^2}{dx^2}$) by one.

Remember, once one has the Green function, one may solve the original differential equation $Lu = f$.

3.2.1 Example: Vibrating String

Let us use this method to do the same problem as before. So; we have the following differential equation:

$$\frac{d^2 v}{dx^2} + k^2 v = f(x)$$

Along with the boundary condition that $v(0) = v(L) = 0$. So, we have:

$$\frac{d^2 G(x, x')}{dx^2} + k^2 G(x, x') = \delta(x - x')$$

And the condition:

$$G(0, x') = G(L, x') = 0$$

Note, we also have $0 \leq x' \leq L$.

So, for $x < x'$:

$$\frac{d^2G(x,x')}{dx^2} + k^2G(x,x') = 0 \qquad G(0,x') = 0$$

Thus having solution:

$$G(x,x') = A(x')\sin kx$$

And, for $x > x'$:

$$\frac{d^2G(x,x')}{dx^2} + k^2G(x,x') = 0 \qquad G(L,x') = 0$$

Having solution:

$$G(x,x') = B(x')\sin(k(L-x))$$

Note, each equation has a different boundary condition, and that is because each excludes the other.

Now, let us match up the Green functions at $x = x'$:

$$A(x')\sin kx' = B(x')\sin k(L-x')$$

To find the discontinuity, let us integrate the original equation, very close to x':

$$\int_{x'-\epsilon}^{x'+\epsilon}\left(\frac{d^2}{dx^2}+k^2\right)G(x,x')dx = \int_{x'-\epsilon}^{x'+\epsilon}\delta(x-x')dx = 1$$

Where we note the RHS by the very definition of the delta-function. Integrating the LHS by parts:

$$\left[\frac{dG}{dx}\right]_{x'-\epsilon}^{x'+\epsilon} + k^2G(x,x')2\epsilon = 1$$

From which we see that:

$$\frac{d}{dx}G(x'+\epsilon,x') - \frac{d}{dx}G(x'-\epsilon,x') = 1$$

As $\epsilon \to 0$. Being careful about which function to use, we get:

$$-B(x')k\cos k(L-x') - A(x')\cos kx' = 1$$

So, from which we can find:

$$A(x') = -\frac{\sin k(L-x')}{k\sin kL} \qquad B(x') = -\frac{\sin kx'}{k\sin kL}$$

Hence, the Green function is:

$$G(x,x') = -\frac{1}{k\sin kL}\begin{cases}\sin kx\sin k(L-x') & x < x'\\ \sin kx'\sin k(L-x) & x > x'\end{cases}$$

Notice the symmetry under exchange of x and x'. This is a common feature of such Green functions.

3.3 Quantum Mechanical Scattering

Here, we shall use Green functions to find the radial distribution of scattered intensity, from a beam of particles which is incident upon some target; if we view the scattered beam from a distance which is a lot larger than the dimensions of the scattering target - which is nice & easy to fulfill, experimentally.

So, we say that some incident beam of particles is incident upon a scattering potential $V(r)$. We require $rV(r) \to 0$ as $r \to \infty$. We shall consider a time-independant approach. Thus, the incident wavefunction is of the form:

$$\phi_0(\boldsymbol{r}) = e^{ikz}$$

If the beam is travelling along the z-direction. It then must satisfy the TISE:

$$-\frac{\hbar^2}{2m}\nabla^2\phi_0(\boldsymbol{r}) + V(\boldsymbol{r})\phi_0(\boldsymbol{r}) = E\phi_0(\boldsymbol{r}) \qquad E = \frac{\hbar^2 k^2}{2m}$$

We can rearrange this, giving:

$$(\nabla^2 + k^2)\phi_0(\boldsymbol{r}) = \frac{2m}{\hbar^2}V(\boldsymbol{r})\phi_0(\boldsymbol{r})$$

Let us define some function as the RHS; thus:

$$(\nabla^2 + k^2)\phi_0(\boldsymbol{r}) = \rho(\boldsymbol{r}) \qquad \rho(\boldsymbol{r}) \equiv \frac{2m}{\hbar^2}V(\boldsymbol{r})\phi_0(\boldsymbol{r})$$

Outgoing spherical waves are of the form:

$$f(\theta,\phi)\frac{e^{ikr}}{r}$$

With the scattering amplitude $f(\theta,\phi)$ being the subject of what we are trying to find. So, let us write the total wavefunction as:

$$\phi(\boldsymbol{r}) = e^{ikz} + f(\theta,\phi)\frac{e^{ikr}}{r}$$

Under the condition that $r \to \infty$. Let this wavefunction tend to some $\chi(\boldsymbol{r})$ as r goes large. That is:

$$\phi(\boldsymbol{r}) \to \chi(\boldsymbol{r}) \qquad \chi(\boldsymbol{r}) \to f(\theta,\phi)\frac{e^{ikr}}{r} \qquad r \to \infty$$

Hence, our 'Schrodinger equation' becomes:

$$(\nabla^2 + k^2)\chi(\boldsymbol{r}) = \rho(\boldsymbol{r})$$

We shall solve this using Green functions methods:

$$(\nabla^2 + k^2)G(\boldsymbol{r},\boldsymbol{r}') = \delta(\boldsymbol{r}-\boldsymbol{r}')$$

Now, if $k = 0$, then we already know the Green function:

$$\nabla^2 G(\boldsymbol{r},\boldsymbol{r}') = \delta(\boldsymbol{r}-\boldsymbol{r}') \quad \Rightarrow \quad G(\boldsymbol{r},\boldsymbol{r}') = -\frac{1}{4\pi}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}$$

Infact, we can use 'translation invariance' to shift our origin to $\boldsymbol{r}'$, so that the above is just:

$$\nabla^2 G_0(\boldsymbol{r}) = \delta(\boldsymbol{r}) \quad \Rightarrow \quad G_0(\boldsymbol{r}) = -\frac{1}{4\pi}\frac{1}{r}$$

We also can verify that the following works, for $k \to 0$:

$$(\nabla^2 + k^2)\frac{e^{ikr}}{r} = 0 \quad r \neq 0$$

That is then:

$$G(\boldsymbol{r}, \boldsymbol{r}') = -\frac{1}{4\pi}\frac{e^{ik|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}$$

Hence, from Green function theory, we can write down $\chi(\boldsymbol{r})$:

$$\chi(\boldsymbol{r}) = \int G(\boldsymbol{r}, \boldsymbol{r}')\rho(\boldsymbol{r}')d^3r'$$

And hence:

$$\phi(\boldsymbol{r}) = e^{ikz} + \int G(\boldsymbol{r}, \boldsymbol{r}')\rho(\boldsymbol{r}')d^3r'$$

Remembering that:

$$\rho(\boldsymbol{r}) \equiv \frac{2m}{\hbar^2}V(\boldsymbol{r})\phi(\boldsymbol{r})$$

Thus:

$$\phi(\boldsymbol{r}) = e^{ikz} + \frac{2m}{\hbar^2}\int G(\boldsymbol{r}, \boldsymbol{r}')V(\boldsymbol{r}')\phi(\boldsymbol{r}')d^3r' \tag{3.4}$$

This is, infact, known as the *Lipman-Schwinger* equation.

3.3.1 The Born Approximation

To solve the above LS-equation, we put it *into itself*, for the $\phi(\boldsymbol{r}')$ term in the integral. That is:

$$\begin{aligned}\phi(\boldsymbol{r}) &= e^{ikz} + \frac{2m}{\hbar^2}\int G(\boldsymbol{r}, \boldsymbol{r}')V(\boldsymbol{r}')e^{ikz'}d^3r' \\ &+ \frac{2m}{\hbar^2}\int\int G(\boldsymbol{r}, \boldsymbol{r}')V(\boldsymbol{r}')G(\boldsymbol{r}', \boldsymbol{r}'')V(\boldsymbol{r}'')\phi(\boldsymbol{r}'')d^3r'd^3r'' + \ldots\end{aligned}$$

For weak potentials, we can ignore the second integral, so we truncate at the first order of potential, giving us:

$$\phi(\boldsymbol{r}) = e^{ikz} + \frac{2m}{\hbar^2}\int G(\boldsymbol{r}, \boldsymbol{r}')V(\boldsymbol{r}')e^{ikz'}d^3r'$$

If we put our form of the Green function in:

$$\phi(\boldsymbol{r}) = e^{ikz} - \frac{m}{2\pi\hbar^2}\int \frac{e^{ik|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}V(\boldsymbol{r}')e^{ikz'}d^3r'$$

Now, to find $f(\theta, \phi)$, we need only know what happens for r large. In this limit, we may write:

$$|\boldsymbol{r} - \boldsymbol{r}'| \approx r\left(1 - \frac{\boldsymbol{r}\cdot\boldsymbol{r}'}{r^2}\right)$$

Also, from the definition, we have:

$$f(\theta,\phi)\frac{e^{ikr}}{r} = -\frac{m}{2\pi\hbar^2}\int \frac{e^{ik|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}V(\boldsymbol{r}')e^{ikz'}d^3r'$$

Using the following definitions:

$$\boldsymbol{k} = k\hat{\boldsymbol{z}} \quad kz = \boldsymbol{k}\cdot\boldsymbol{r} \quad \boldsymbol{k}' = k\frac{\boldsymbol{r}}{r}$$

We end up with the *Born approximation*:

$$f(\theta,\phi) = -\frac{m}{2\pi\hbar^2}\int V(\boldsymbol{r})e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}}d^3r \tag{3.5}$$

Notice that it is the Fourier transform of the scattering potential.

3.4 Time Dependent Wave Equation

In electrodynamics, we have relations such as:

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\Phi = 0$$

We shall use the *D'Alembertian* symbol to make this more compact:

$$\Box \equiv \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2$$

Then, we will have equations like:

$$\Box\Phi = \frac{\rho(\boldsymbol{r},t)}{\varepsilon_0} \qquad \Box\boldsymbol{A} = \mu_0 J(\boldsymbol{r},t)$$

So, it seems natural to discuss such equations for arbitrary sources. Then, we are led to the Green function for the D'Alembertian operator:

$$\Box G(\boldsymbol{r},t;\boldsymbol{r}',t') = \delta(\boldsymbol{r}-\boldsymbol{r}')\delta(t-t') \tag{3.6}$$

And we use the boundary conditions that potentials decay when we are very far away from sources, and cause precedes effect.:

$$G(\boldsymbol{r},t;\boldsymbol{r}',t') \to 0, r\to\infty \qquad G(\boldsymbol{r},t;\boldsymbol{r}',t') = 0, t < t'$$

We shall 'Galilean invariance' to solve. That is, we shift our origin so that we are solving:

$$\Box G(\boldsymbol{r},t) = \delta(\boldsymbol{r})\delta(t) \tag{3.7}$$

We shall proceed by solving for Green functions via Fourier transform.

3.4.1 Solution for Green functions by Fourier Transforms

First, we define the Fourier transform of G:

$$\tilde{G}(\boldsymbol{k},\omega) = \int d^3r \int dt G(\boldsymbol{r},t)e^{-i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} \tag{3.8}$$

$$G(\boldsymbol{r},t) = \int \frac{d^3k}{(2\pi)^3}\int \frac{d\omega}{2\pi}\tilde{G}(\boldsymbol{k},\omega)e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} \tag{3.9}$$

Also, the delta function:

$$\delta(\boldsymbol{r})\delta(t) = \int \frac{d^3k}{(2\pi)^3}\int \frac{d\omega}{2\pi}e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} \tag{3.10}$$

Now, we substitute (3.9) and (3.10) into (3.7):

$$\int \frac{d^3k}{(2\pi)^3}\int \frac{d\omega}{2\pi}\tilde{G}(\boldsymbol{k},\omega)\Box e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} = \int \frac{d^3k}{(2\pi)^3}\int \frac{d\omega}{2\pi}e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}$$

Now:

$$\Box e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} = \left(k^2-\frac{\omega^2}{c^2}\right)e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}$$

Hence, using this:

$$\int \frac{d^3k}{(2\pi)^3}\int \frac{d\omega}{2\pi}\tilde{G}(\boldsymbol{k},\omega)\left(k^2-\frac{\omega^2}{c^2}\right)e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} = \int \frac{d^3k}{(2\pi)^3}\int \frac{d\omega}{2\pi}e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}$$

Now, as the integrals are the same, we equate the integrands:

$$\tilde{G}(\boldsymbol{k},\omega)\left(k^2-\frac{\omega^2}{c^2}\right) = 1$$

That is:

$$\tilde{G}(\boldsymbol{k},\omega) = \frac{-c^2}{\omega^2-c^2k^2}$$

Hence, we can use this in (3.9):

$$G(\boldsymbol{r},t) = \int \frac{d^3k}{(2\pi)^3}\int \frac{d\omega}{2\pi}\left(\frac{-c^2}{\omega^2-c^2k^2}\right)e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} \tag{3.11}$$

Now, let us look at the k-integral. It is of the form:

$$I = \int e^{i\boldsymbol{k}\cdot\boldsymbol{r}}f(k^2)d^3k$$

To do this integral, we choose the k_3 axis to be parallel to $\boldsymbol{r}$. So, the integral:

$$I = \int\int\int e^{ikr\cos\theta}f(k^2)k^2\sin\theta dk d\theta d\phi$$

Now, the $d\phi$ just gives a factor of 2π. The next stage is to make a substitution:

$$x \equiv kr\cos\theta \quad \Rightarrow \quad dx = -kr\sin\theta d\theta$$

Hence, this gives:

$$\begin{aligned} I &= 2\pi \int k^2 f(k^2) dk \int_{\theta=0}^{\pi} e^{ix} \frac{1}{-kr} dx \\ &= 2\pi \int k^2 dk f(k^2) \left[\frac{1}{-ikr} e^{ikr\cos\theta} \right]_0^{\pi} \\ &= 2\pi \int k^2 dk f(k^2) \frac{1}{-ikr} \left[e^{-ikr} - e^{ikr} \right] \\ &= \frac{2\pi}{ir} \int_0^{\infty} k \left[e^{ikr} - e^{-ikr} \right] f(k^2) dk \end{aligned}$$

Now, finally, we do a standard-trick with exponentials:

$$I = \frac{2\pi}{ir} \int_{-\infty}^{\infty} k e^{ikr} f(k^2) dk$$

So, putting this into (3.11):

$$G(\boldsymbol{r}, t) = -\frac{c^2}{(2\pi)^2 ir} \int_{-\infty}^{\infty} k dk \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{1}{\omega^2 - c^2 k^2} e^{i(kr - \omega t)}$$

So, now for the ω integral:

$$I = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{1}{\omega^2 - c^2 k^2} e^{-i\omega t}$$

Now, notice, the integral, as it sweeps over the real line, will sweep past two singularities: $\omega = \pm ck$. So to do the integral, we use contour integration. The residues are:

$$R_+ = \frac{e^{-ickt}}{2ck} \qquad R_- = \frac{e^{ickt}}{-2ck}$$

And the integral is:

$$I = \frac{1}{2\pi} 2\pi i (R_+ + R_-)$$

That is:

$$I = -\frac{i}{2ck} (e^{-ickt} - e^{ickt})$$

So, putting this back in:

$$G(\boldsymbol{r}, t) = \frac{c}{(2\pi)^2 2r} \int_{-\infty}^{\infty} e^{ik(r-ct)} - e^{ik(r+ct)} dk$$

Throughout having made use of $t > 0$. The above integral is just a delta-function:

$$G(\boldsymbol{r}, t) = \frac{c}{4\pi r} [\delta(r - ct) - \delta(r + ct)]$$

The second of which we ignore, due to our causality requirement. So:

$$G(\boldsymbol{r}, t) = \frac{c}{4\pi r} \delta(r - ct)$$

Let reinstate our primed basis:

$$G(\boldsymbol{r}, t; \boldsymbol{r}', t') = \frac{c}{4\pi |\boldsymbol{r} - \boldsymbol{r}'|} \delta(|\boldsymbol{r} - \boldsymbol{r}'| - c(t - t')) \tag{3.12}$$

So, we see that we only have a non-zero solution when events lie within the backward light cone:

$$|\boldsymbol{r} - \boldsymbol{r}'| = c(t - t')$$

3.5 Reducing the Dimension

Suppose we have a Green function for the following operator, in 3+1 (3 spatial, 1 temporal) coordinates:

$$\Box_{3+1} = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2$$

We have just seen that its Green function is:

$$G^{3+1}(\boldsymbol{r}, t; \boldsymbol{r}', t') = \frac{c}{4\pi|\boldsymbol{r} - \boldsymbol{r}'|}\delta(|\boldsymbol{r} - \boldsymbol{r}'| - c(t - t')) \qquad t > t'$$

And zero for $t < t'$. Now, what is the Green function G^{2+1} for the operator in only 2 spatial dimensions:

$$\Box_{2+1} = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2}$$

Now, recall that the original operator/Green function satisfied the following:

$$\Box_{3+1} G^{3+1}(\boldsymbol{r}, t; \boldsymbol{r}', t') = \delta(x - x')\delta(y - y')\delta(z - z')\delta(t - t')$$

Let us integrate both sides of this equation, with respect to z':

$$\int dz' \Box_{3+1} G^{3+1}(\boldsymbol{r}, t; \boldsymbol{r}', t') = \int dz' \delta(x - x')\delta(y - y')\delta(z - z')\delta(t - t')$$

The RHS integral is trivial; and on the LHS we can take the operator out of the integral:

$$\Box_{3+1} \int dz' G^{3+1}(\boldsymbol{r}, t; \boldsymbol{r}', t') = \delta(x - x')\delta(y - y')\delta(t - t')$$

Now notice: on the LHS, if we are integrating over z', then the resulting expression will be completely independent of z'. Hence, the operator effectively looses a dimension to become $\Box_{2+1}$. Thus, we see that:

$$G^{2+1}(x, y, t; x', y', t') = \int dz' G^{3+1}(\boldsymbol{r}, t; \boldsymbol{r}', t')$$

And, for our particular Green function, we can do this:

$$\int dz' G^{3+1}(\boldsymbol{r}, t; \boldsymbol{r}', t') = \int dz' \frac{c}{4\pi|\boldsymbol{r} - \boldsymbol{r}'|}\delta(|\boldsymbol{r} - \boldsymbol{r}'| - c(t - t'))$$

To proceed, notice that the zeros of the delta function (with respect to the integration variable z') occurs at:

$$z' = z \pm \sqrt{c^2(t - t')^2 - (x - x')^2 - (y - y')^2}$$

And that we write a delta function a function as:

$$\delta(f(x)) = \sum_i \frac{\delta(x - x_i)}{\frac{df}{dx}(x_i)}$$

Where $f(x_i) = 0$; hence, doing carefully results in:

$$G^{2+1}(\boldsymbol{x} - \boldsymbol{x}'; t - t') = \frac{c}{2\pi}\frac{1}{\sqrt{c^2(t - t')^2 - |\boldsymbol{x} - \boldsymbol{x}'|^2}}$$

4 Variational Calculus

4.1 Functionals

Now, a *function* is a rule that takes number, and produces a number. That is, its argument is a number:

$$f : \mathbb{R} \mapsto \mathbb{R}$$

A *functional* is something that takes a function, and gives number. That is, its argument is a function:

$$I : (\mathbb{R} \mapsto \mathbb{R}) \mapsto \mathbb{R}$$

An example of such a functional is:

$$I[y] = \int_0^\pi y(x)dx$$

So that $I[\sin(x)] = 2, I[\cos(x)] = 0$ etc.

Notice that we use square brackets around the argument of a functional, where its only explicit argument is y, and not x.

4.2 Stationary Points

For a function, a stationary point is some $\boldsymbol{x}$ where $f(\boldsymbol{x})$ dosent change under a small variation in $\boldsymbol{x}$:

$$\delta f \approx \frac{\partial f}{\partial \boldsymbol{x}} \cdot \delta \boldsymbol{x} = 0 \qquad \nabla f = 0$$

A functional has a stationary point, for $y(x)$, when $I[y]$ dosent change under a small change in its argument. That is, a change:

$$y(x) \to y(x) + \epsilon(x)$$

That is, to first order in $\epsilon(x)$.

Now, let us consider a specific type of functional:

$$I[y] = \int_a^b F(y(x), y'(x), x)dx \tag{4.1}$$

So that the functional is the integral over some function, which takes arguments of $y(x), y'(x)$ and x. We will assume that there is some boundary condition for $y(x)$ at $x = a, b$; and that:

$$\epsilon(a) = \epsilon(b) = 0$$

So, the variation in the functional is:

$$\delta I[y] = I[y + \epsilon] - I[y]$$

That is:

$$\delta I[y] = \int_a^b F(y(x) + \epsilon(x), y'(x) + \epsilon'(x), x)dx - \int_a^b F(y(x), y'(x), x)dx$$

Which we Taylor expand to:

$$\delta I[y] = \int_a^b \left(\epsilon \frac{\partial F}{\partial y} + \epsilon' \frac{\partial F}{\partial y'} \right) dx$$

Where we have implicitly implied that F, y and ϵ:

$$F \equiv F(y(x), y'(x), x) \qquad y \equiv y(x) \qquad \epsilon \equiv \epsilon(x)$$

We then integrate by parts:

$$\begin{aligned}
\delta I[y] &= \int_a^b \left(\epsilon \frac{\partial F}{\partial y} + \epsilon' \frac{\partial F}{\partial y'} \right) dx \\
&= \int_a^b \epsilon \frac{\partial F}{\partial y} dx + \int_a^b \frac{d\epsilon}{dx} \frac{\partial F}{\partial y'} dx \\
&= \int_a^b \epsilon \frac{\partial F}{\partial y} dx + \left(\left[\epsilon \frac{\partial F}{\partial y'} \right]_a^b - \int_a^b \epsilon \frac{d}{dx} \frac{\partial F}{\partial y'} dx \right) \\
&= \left[\epsilon \frac{\partial F}{\partial y'} \right]_a^b + \int_a^b \epsilon \left(\frac{\partial F}{\partial y} - \frac{d}{dx} \frac{\partial F}{\partial y'} \right) dx \\
&= \int_a^b \epsilon \left(\frac{\partial F}{\partial y} - \frac{d}{dx} \frac{\partial F}{\partial y'} \right) dx
\end{aligned}$$

We have noted that $\epsilon(a) = \epsilon(b) = 0$ in the second to last line. We obviously require this to be zero for a 'stationary point' (i.e one with little/no variation for adding on some small function). Also, as the perturbation function is arbitrary, as well as small, we hence deduce that the quantity in brackets is zero. Therefore, we arrive at the Euler-Lagrange equation:

$$\frac{\partial F}{\partial y} - \frac{d}{dx} \frac{\partial F}{\partial y'} = 0 \qquad F(y, y', x), y'(x), y(x) \tag{4.2}$$

We use the functional derivative notation:

$$\frac{\delta I[y]}{\delta y}$$

4.3 Special Cases

4.3.1 No y-dependance

Suppose that the function F is independant of y. That is, $F(y', x)$. Hence, we see that the EL equation becomes:

$$\frac{d}{dx} \frac{\partial F(y', x)}{\partial y'} = 0 \tag{4.3}$$

Which obviously implies that:

$$\frac{\partial F(y', x)}{\partial y'} = const \tag{4.4}$$

This is known as the *first integral*, as it is arrived at by integrating once. Let us proceed via example.

4.3.2 Example: Distance Between Two Points

Let us show that the distance between two points (in 2D) is along a straight line.

Let us assume some path $(x, y(x))$, which goes from $(a, y(a))$ to $(b, y(b))$. So, some small segment of the path has length:

$$ds^2 = dx^2 + dy^2$$

That is:

$$\begin{aligned} ds &= \sqrt{dx^2 + dy^2} \\ &= dx\sqrt{1 + \left(\frac{dy}{dx}\right)^2} \\ &= dx\sqrt{1 + y'^2} \qquad y' \equiv \tfrac{dy}{dx} \end{aligned}$$

Hence, a result we will use regularly:

$$ds = dx\sqrt{1 + y'^2} \tag{4.5}$$

Hence, the total length of the path is just the sum of all the little segments of the path: the integral:

$$\begin{aligned} L &= \int_a^b ds \\ &= \int_a^b dx\sqrt{1 + y'^2} \end{aligned}$$

Therefore, we see that we have a functional of the type where F does not depend on y explicitly (infact, it does not depend on x either). Hence, we use:

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

With c being the constant in question. That is, we have a way of minimising the path length between two points. Hence, differentiating F, wrt y':

$$\frac{y'}{\sqrt{1 + y'^2}} = c$$

Hence:

$$y'^2 = c^2(1 + y'^2)$$

Which can be trivially rearanged to:

$$y' = \frac{c}{\sqrt{1 - c^2}}$$

Now, everything on the RHS is a constant. Hence, as y' is the gradient of a line, we see that the gradient of the shortest line connecting two points is a constant. Which is a straight line.

4.3.3 No x-dependance

Let us consider another 'special case': Where F does not depend on x explicitly. That is, $F(y(x), y'(x))$. Now, writing the EL equation:

$$\frac{\partial F}{\partial y} = \frac{d}{dx}\frac{\partial F}{\partial y'}$$

Now, if $F(y, y')$, then its differential is (by the chain rule):

$$dF = \frac{\partial F}{\partial y}dy + \frac{\partial F}{\partial y'}dy'$$

That is, dividing through by dx:

$$\frac{dF}{dx} = \frac{\partial F}{\partial y}\frac{dy}{dx} + \frac{\partial F}{\partial y'}\frac{dy'}{dx}$$

Which we can obviously write as:

$$\frac{dF}{dx} = \frac{\partial F}{\partial y}y' + \frac{\partial F}{\partial y'}y''$$

Now, if we look back at the EL equation, we see that we have an expression for the middle term. Thus:

$$\frac{dF}{dx} = \frac{d}{dx}\left(\frac{\partial F}{\partial y'}y'\right) + \frac{\partial F}{\partial y'}y''$$

Rewriting the end term:

$$\frac{dF}{dx} = \frac{d}{dx}\left(\frac{\partial F}{\partial y'}y'\right) + \frac{\partial F}{\partial y'}\frac{d}{dx}y'$$

Which we can see is actually exactly the same as the following (if the product rule is carried out on the following):

$$\frac{dF}{dx} = \frac{d}{dx}\left(\frac{\partial F}{\partial y'}y'\right)$$

Therefore:

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

Hence, the first integral in this case is:

$$F - \frac{\partial F}{\partial y'}y' = c \tag{4.6}$$

Where c is a constant.

4.3.4 Example: Snell's Law

Let us derive Snell's law, using our techniques. Fermat's principle of geometrical optics is that light travels along the path that takes least time, or has shortest optical path length.

Consider the situtation where we have some interface between two mediums, having refractive indices n_1, n_2. The angle light makes to the horizontal is ϕ_1, ϕ_2; and to the normal θ_1, θ_2. Let us consider the travel time along some arbitrary path $(x, y(x))$.

So, the light travel time is the sum over little-bits of distance, divided by the speed the light travels:

$$T[y] = \int_a^b \frac{ds}{v(x,y)}$$

We know that the speed of light in a medium is given by $v = c/n$. Hence:

$$T[y] = \frac{1}{c}\int_a^b ds\, n(x,y)$$

We shall use the line-element previously derived, and restrict the argument of the refractive index:

$$T[y] = \frac{1}{c}\int_a^b n(y)\sqrt{1+y'^2}dx$$

Therefore, we see that our function is independant of x. Hence, we use:

$$F - \frac{\partial F}{\partial y'}y' = d \qquad F(y,y') = n(y)\sqrt{1+y'^2}$$

Where d is the constant. Hence:

$$n(y)\sqrt{1+y'^2} - n(y)\frac{y'}{\sqrt{1+y'^2}}y' = d$$

Bringing under a common denominator:

$$\frac{n(y)}{\sqrt{1+y'^2}} = d \quad \Rightarrow \qquad \sqrt{1+y'^2} = \frac{n}{d}$$

Therefore:

$$y'^2 = 1 - \frac{d^2}{n^2(y)}$$

And, if $n(y)$ is constant, then the whole RHS is a constant. Therefore y' is a constant. And therefore we get a straight line.
Now, from the geometry of the system, we see that $\tan\phi = y'$. Therefore:

$$\begin{aligned}\sqrt{1+y'^2} &= \sqrt{1+\tan^2\phi} \\ &= \sqrt{1+\frac{\sin^2\phi}{\cos^2\phi}} \\ &= \sqrt{\frac{\cos^2\phi+\sin^2\phi}{\cos^2\phi}} \\ &= \frac{1}{\cos\phi}\end{aligned}$$

Therefore, we see that:

$$n\cos\phi = const$$

Therefore:

$$n_1\cos\phi_1 = n_2\cos\phi_2$$

And, as $\cos\phi_i = \sin\theta_i$, this easily becomes Snell's law:

$$n_1\sin\theta_1 = n_2\sin\theta_2$$

4.3.5 Example: Brachistochrone

The brachistochrone is the curve along which a particle slides the quickest, between two points A and B, under the influence of gravity.

Think about this as finding the shape of a piece of wire, where a bead is sliding down the wire. The two ends of the wire are fixed, and the problem is to find the optimum shape which the wire must take, for which the bead takes the shortest time to slide down.

Choose the end points $A(0,0)$ and $B(b, y(b))$; with a curve $(x, y(x))$. Then, the travel time is:

$$\begin{aligned} T[y] &= \int_0^b dt \\ &= \int_0^b \frac{dt}{ds} ds \\ &= \int_0^b \frac{ds}{v} \qquad v = \tfrac{ds}{dt} \end{aligned}$$

Now, by energy conservation:

$$\tfrac{1}{2}mv^2 = mgy \qquad \Rightarrow \qquad v = \sqrt{2gy}$$

Hence, using the line-element previously derived:

$$T[y] = \frac{1}{\sqrt{2g}} \int_0^b \frac{1}{\sqrt{y}} \sqrt{1 + y'^2} dx$$

Let us ignore the constants, and write down the associated functional:

$$I[y] = \int_0^b \frac{1}{\sqrt{y}} \sqrt{1 + y'^2} dx$$

Therefore, we see that the function is of a type which is independent of x. Hence, we use (4.6). Giving:

$$\sqrt{\frac{1 + y'^2}{y}} - \frac{1}{\sqrt{y}} \frac{y'}{\sqrt{1 + y'^2}} y' = c$$

Where c is a constant. Bringing under a common denominator:

$$\frac{\sqrt{y}}{y\sqrt{1 + y'^2}} = c$$

That is:

$$\frac{1}{\sqrt{y}\sqrt{1 + y'^2}} = c$$

Let us jig this around, and relable the constant:

$$y(1 + y'^2) = 2R$$

This is infact the equation of a *cycloid*. Now, let us write:

$$y' \equiv \cot(\tfrac{\phi}{2})$$

So that:

$$1 + y'^2 = 1 + \frac{\cos^2(\frac{\phi}{2})}{\sin^2(\frac{\phi}{2})} = \frac{1}{\sin^2(\frac{\phi}{2})}$$

And therefore:

$$y(1 + y'^2) = 2R \qquad \Rightarrow \qquad y(\phi) = 2R\sin^2(\tfrac{\phi}{2})$$

And, using the trig identity $2\sin^2 x = 1 - \cos 2x$, this becomes:

$$y(\phi) = R(1 - \cos\phi)$$

Thus, we have a parameterisation of y in terms of ϕ. Now, just note that $y' = \frac{dy}{dx}$. Now, we need $x(\phi)$. To find this, consider:

$$\begin{aligned}
\frac{dx}{d\phi} &= \frac{dx}{dy}\frac{dy}{d\phi} \\
&= \frac{1}{dy/dx}\frac{dy}{d\phi} \\
&= \frac{1}{\cot(\frac{\phi}{2})} 2R\sin(\tfrac{\phi}{2})\cos(\tfrac{\phi}{2}) \\
&= 2R\sin^2(\tfrac{\phi}{2}) = y(\phi)
\end{aligned}$$

Therefore:

$$x = \int y d\phi = R(\phi - \sin\phi)$$

Hence, we have some parameterisations:

$$x(\phi) = R(\phi - \sin\phi) \qquad y(\phi) = R(1 - \cos\phi)$$

4.4 Variable Endpoints

So, we have, from before:

$$I[y] = \int_a^b F(y(x), y'(x), x)dx$$

With $y(a)$ and $y(b)$ assumed fixed. Now, let us assume that $y(b)$ is free. We shall also have that $\epsilon(a) = 0$, as before, but that nothing is known about $\epsilon(b)$. So, along the same lines as before:

$$\begin{aligned}
\delta I[y] &= \int_a^b \left(\epsilon \frac{\partial F}{\partial y} + \epsilon' \frac{\partial F}{\partial y'}\right) dx \\
&= \int_a^b \epsilon \frac{\partial F}{\partial y} dx + \int_a^b \frac{d\epsilon}{dx}\frac{\partial F}{\partial y'} dx \\
&= \int_a^b \epsilon \frac{\partial F}{\partial y} dx + \left(\left[\epsilon \frac{\partial F}{\partial y'}\right]_a^b - \int_a^b \epsilon \frac{d}{dx}\frac{\partial F}{\partial y'} dx\right) \\
&= \left[\epsilon \frac{\partial F}{\partial y'}\right]_a^b + \int_a^b \epsilon \left(\frac{\partial F}{\partial y} - \frac{d}{dx}\frac{\partial F}{\partial y'}\right) dx \\
&= \int_a^b \epsilon \left(\frac{\partial F}{\partial y} - \frac{d}{dx}\frac{\partial F}{\partial y'}\right) dx + \epsilon(b) \left.\frac{\partial F}{\partial y'}\right|_{x=b}
\end{aligned}$$

Where we now have an 'extra' term, as $\epsilon(b) \neq 0$. Now, we set $\delta I = 0$; and due to the arbitrary nature of the perturbation, we have two equations:

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\frac{\partial F}{\partial y'} = 0 \qquad \left.\frac{\partial F}{\partial y'}\right|_{x=b}$$

So, the EL equation, as before, as well as a new constraint on the derivative of the functional, at the endpoint.

4.4.1 Example: Brachistochrone

We shall redo this problem, with a bead moving along a frictionless wire, from $(0, 0)$ to (x_f, y_f), where x_f is fixed, and y_f is free. Let us find y_f for the shortest fall time.

As before, the functional is:

$$F(x, y') = (1 + y'^2)^{1/2} y^{-1/2}$$

Hence, the EL equation gives:

$$\left[y(1 + y'^2)\right]^{1/2} = k$$

And that gave us that $x = R(\phi - \sin\phi)$ and $y = R(1 - \cos\phi)$. So, our varied endpoint means that to minimise the path, we also have:

$$\left.\frac{\partial F}{\partial y'}\right|_{x=x_f} = \left.\frac{y'}{\sqrt{y(1 + y'^2)}}\right|_{x=x_f} = 0$$

That is:

$$y'(x_f) = \sin\phi_f = 0$$

And hence $\phi_f = \pi$. Hence, as $x = R(\phi - \sin\phi)$, we have that $R = x_f/\pi$, and that $y_f = 2x_f/\pi$.

Therefore, we have solved the Brachistochrone, for the case where the start point is fixed, and that the horizontal distance is fixed. We have then computed the optimum depth for a path connecting the two points, and the shape of that path (the cycloid).

4.5 More Than One Function

Here, we have:

$$I\left[\{y_i\}\right] = \int_a^b F(\{y_i\}, \{y_i'\}, x)dx \qquad i = 1, 2, \ldots, N$$

That is a functional of more than one function. Let us fix the endpoints: $y_i(x)$ is fixed at a, b. The EL equation results in:

$$\frac{\delta I\{y_i\}}{\delta y_i} = 0$$

Which is the same as (doing the expansion & integration again):

$$\frac{\partial F}{\partial y_i} - \frac{d}{dx}\frac{\partial F}{\partial y_i'} = 0 \qquad F(\{y_i\}, \{y_i'\}, x)$$

That is, N EL equations. This is readily used:

4.5.1 Hamilton's Principle of Least Action

In classical mechanics, suppose we have n point particles, with positions $q_i(t)$. Then, the Lagrangian is:

$$L(\{q_i\}, \{\dot{q_i}\}, t) = T(q_i, \dot{q}_i) - V(q_i, t)$$

Then, let us define the action:

$$S = \int_{t_i}^{t_f} L(q_i, \dot{q}_i, t)dt \tag{4.7}$$

With fixed endpoints. The principle of least action is that $\delta S = 0$. Clearly, this corresponds to the 'EL' equation, with a trivial reassignment of letters. This gives the Lagrange equations:

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = 0 \tag{4.8}$$

An example. The Lagrangian for a mass on a spring:

$$L = T - V = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2$$

Clearly here we have $i = 1$ only, and that $q_1 = x$. Hence:

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = -kx - m\ddot{x} = 0$$

Which is trivially:

$$\ddot{x} = -\frac{k}{m}x$$

The first integrals link to conservation laws:

Energy Conservation: $L(q_i, \dot{q}_i)$ That is, when the Lagrangian (i.e. the potential part) is independent of time. From:

$$I[y] = \int_a^b F(y, y')dx \qquad \Rightarrow \qquad F - y'\frac{\partial F}{\partial y'} = const$$

Thus, generalising:

$$L - \sum_i \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} = const$$

If $L = T - V$, and if T is quadratic in $\dot{q}_i$, then:

$$\sum \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} = 2T$$

Hence:

$$T - V - 2T = const \equiv -E$$

Hence, the total energy is a constant.

Momentum Conservation: $L(\dot{q}_i, t)$ Here, we have that the Lagrangian is independant of q_i. It corresponds to the case where F is independent of y. Hence, we have:

$$\frac{\partial L}{\partial \dot{q}_i} = const \equiv p_i$$

Consider and example. The Lagrangian for a relativistic, charged particle in an external electric & magnetic fields:

$$L(\boldsymbol{x}, \dot{\boldsymbol{x}}, t) = -mc^2(1 - \dot{x}/c^2)^{1/2} + q\boldsymbol{A} \cdot \dot{\boldsymbol{x}} - q\phi$$

We get:

$$\frac{d}{dt}\frac{m\dot{x}_i}{(1 - \dot{x}_i^2/c^2)^{1/2}} = q(\nabla_i \boldsymbol{A}) \cdot \boldsymbol{x} - q\partial_t A_i - q\nabla_i \phi$$

Which is equivalent to, if we carefully consider the expression:

$$\frac{d}{dt}\boldsymbol{p} = q(\boldsymbol{v} \times \boldsymbol{B}) + q\boldsymbol{E}$$

4.6 Field Theory

Suppose that we consider more dimensions, so that:

$$x = x^1, x^2, \ldots x^N \qquad \phi(x^1, \ldots x^N) = \phi(x)$$

Then, the generalisation of the action is:

$$\mathcal{S} = \int_\Omega \mathcal{L}(\phi(x), \partial_\mu \phi(x), x)\, d\tau \tag{4.9}$$

Where:

$$d\tau = dx^1 dx^2 \ldots dx^N \qquad \partial_\mu \equiv \frac{\partial}{\partial x^\mu} \equiv \phi_{,\mu}$$

Where we have introduced some notation for the covariant differential operator. We have that Ω is some volume in the N dimensional space. The boundary of the surface has the property that the small variation is zero on it:

$$\epsilon(x)|_{\partial\Omega} = 0$$

Where $\phi(x) \to \phi(x) + \epsilon(x)$. Now, just to see how we motivate our expressions, consider what we had before, when $y(x)$, a function of one variable only. The action was:

$$S = \int_a^b \left(\epsilon \frac{\partial F}{\partial y} + \epsilon' \frac{\partial F}{\partial y'} \right) dx$$

Which actually meant:

$$S = \int_a^b \left(\epsilon \frac{\partial F}{\partial y} + \frac{d\epsilon}{dx} \frac{\partial F}{\partial (dy/dx)} \right) dx$$

But now we have the case that 'x' is only one of many variables 'x_i'. So, we may be able to see that to generalise, we have:

$$\begin{aligned} \delta\mathcal{S} &= \int_\Omega \left(\epsilon(x) \frac{\partial \mathcal{L}}{\partial \phi} + \sum_{\mu=1}^{N} \frac{\partial \epsilon}{\partial x^\mu} \frac{\partial \mathcal{L}}{\partial(\partial\phi/\partial x^\mu)} \right) d\tau \\ &= \int_\Omega \left(\epsilon \frac{\partial \mathcal{L}}{\partial \phi} + \sum_\mu \epsilon_{,\mu} \frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} \right) d\tau \end{aligned}$$

Where we have used our previously defined notation:

$$\psi_{,\mu} \equiv \frac{\partial \psi}{\partial x^\mu}$$

To minimise the action, with the requirement that the variation is zero on the surface, we will end up with the analogue to the EL equation:

$$\frac{\partial \mathcal{L}}{\partial \phi} - \sum \frac{\partial}{\partial x^\mu} \frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} = 0 \tag{4.10}$$

Let us "derive" this by analogy. Consider the previous Lagrangian:

$$L(\{\dot{q}_i\}, \{q_i\}, t) \qquad \dot{q}_i = \frac{dq_i}{dt}$$

The action was minimised with the Euler-Lagrange equation:

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = 0$$

Now, this Lagrangian was a function of coordinates $\{q_i\}$, and only one parameter t; the distinction between coordinates & parameters is that coordinates vary according to the parameterisation of the parameters (for example, a set of quantities which change in time). If the number of parameters now changes; so that we have some set of parameters $\{p_i\}$, then we must change our EL equation

accordingly. Where we had $\dot{q}_i$ before, we must now differentiate q_i with respect to each parameter. So, our new Lagrangian looks more like:

$$L(\{q_i\},\{\partial_j q_i\},\{p_i\}) \qquad \partial_j q_i \equiv \frac{\partial q_i}{\partial p^j} \equiv q_{i,j}$$

So that our EL equation will now look like:

$$\frac{\partial L}{\partial q_i} - \sum_j \frac{\partial}{\partial p^j}\frac{\partial L}{\partial q_{i,j}} = 0$$

We take account of all the paramters involved. As an example of such a set of coordinates & paramters are the coordinates (x, y) which vary both in polar angle & time (ϕ, t). That is, $(x(\phi,t), y(\phi,t))$. That is:

$$\{q_i\} = (x, y) \qquad \{p_i\} = (\phi, t)$$

In which case we would have the two EL equations:

$$\begin{aligned}\frac{\partial L}{\partial x} - \frac{\partial}{\partial \phi}\frac{\partial L}{\partial x_{,\phi}} - \frac{\partial}{\partial t}\frac{\partial L}{\partial x_{,t}} &= 0\\ \frac{\partial L}{\partial y} - \frac{\partial}{\partial \phi}\frac{\partial L}{\partial y_{,\phi}} - \frac{\partial}{\partial t}\frac{\partial L}{\partial y_{,t}} &= 0\end{aligned}$$

It could also be the case that a single coordinate (maybe a scalar field quantity) varies in 3 spatial and one chronological direction, in which case $\phi(x, y, z, t)$; $\{q_i\} = \phi$ and $\{p_i\} = (x, y, z, t)$. One should be very careful with notation! Althogh not a derivation, it lets us see how the EL equation comes about, by analogy. Let us revert back to the above EL-equation, where we have $\{q_i\} = \phi$ (i.e. a single coordinate):

$$\frac{\partial \mathcal{L}}{\partial \phi} - \sum \frac{\partial}{\partial x^\mu}\frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} = 0$$

Consider a 1+1 example (one space, one time): (x, t), so that the scalar field is just $\phi(x, t)$. That is, we have some quantity which varies along the real line (say) & in time. For example, a short blast of heat at one end of a copper rod; the end closest to the heat source will be warmest, and the temperature will decrease in time, after the source has been removed. Then, we would have as our scalar field & parameters $T(x, t)$.
So, the above EL equation expands to:

$$\frac{\partial \mathcal{L}}{\partial \phi} - \frac{\partial}{\partial x}\frac{\partial \mathcal{L}}{\partial \phi_{,x}} - \frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \phi_{,t}} = 0$$

Denoting:

$$\phi_{,x} \equiv \frac{\partial \phi}{\partial x} = \phi' \qquad \phi_{,t} \equiv \frac{\partial \phi}{\partial t} = \dot{\phi}$$

We have:

$$\frac{\partial \mathcal{L}}{\partial \phi} - \frac{\partial}{\partial x}\frac{\partial \mathcal{L}}{\partial \phi'} - \frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot{\phi}} = 0$$

Then, the action takes the form:

$$\mathcal{S} = \int\int \mathcal{L}(\phi, \phi', \dot{\phi}, x, t)\, dx\, dt$$

In this way, we can see that $\int dx\mathcal{L}$ is the Lagrangian; hence the expression $\mathcal{L}$ is the Lagrange density.

So, let us consider a stretched string, with fixed endpoints. It will have line element:

$$dl = \sqrt{1 + y'^2}dx$$

That is, a string along the x-axis, being deformed into the y-axis. So, for small deformations, we can trivially expand this to:

$$dl = \left(1 + \frac{y'^2}{2}\right) dx$$

Its not hard to see that a 'small bit of kinetic energy' has the form:

$$dK = \frac{1}{2}\rho dx\dot{y}^2$$

Where the density of the string remains constant. Similarly, for a tension T, a small bit of potential is just the tension multiplied by the extension. Thus:

$$dV = T(dl - dx) = T\frac{1}{2}y'^2dx$$

Hence, the Lagrangian is:

$$L = \int dK - dV = \int dx(\tfrac{1}{2}\rho\dot{y}^2 - \tfrac{1}{2}y'^2T)$$

Hence, it is clear that the Lagrange density is:

$$\mathcal{L} = \tfrac{1}{2}\rho\dot{y}^2 - \tfrac{1}{2}y'^2T$$

And hence the action:

$$\mathcal{S} = \int\int dxdt(\tfrac{1}{2}\rho\dot{y}^2 - \tfrac{1}{2}y'^2T)$$

And thus, to minimise the action:

$$\frac{\partial\mathcal{L}}{\partial\phi} - \frac{\partial}{\partial x}\frac{\partial\mathcal{L}}{\partial\phi'} - \frac{\partial}{\partial t}\frac{\partial\mathcal{L}}{\partial\dot{\phi}} = 0 \qquad \Rightarrow \qquad Ty'' - \rho\ddot{y} = 0$$

Where we have identified ϕ with y. Hence, we have that the string undergoes a motion described by $Ty'' = \rho\ddot{y}$, which will minimise the action. This is just the standard wave equation:

$$\frac{\partial^2 y}{\partial x^2} = \frac{\rho}{T}\frac{\partial^2 y}{\partial t^2}$$

Let us consider a more complicated example.

Let us discuss the motion of a freely moving elastic band, length $2\pi l$. Let us analyse its motion when it is almost circular.

Here, we shall use ϕ as an angular coordinate, so that $(x(\phi,t), y(\phi,t))$; $\phi \in [0, 2\pi]$. Then, similarly:

$$dV = \kappa\tfrac{1}{2}\left(\sqrt{x'^2 + y'^2} - l\right)^2 d\phi$$

And:

$$dT = \rho l d\phi \tfrac{1}{2}(\dot{x}^2 + \dot{y}^2)$$

Hence:

$$\mathcal{S} = \int_0^T dt \int_0^{2\pi} d\phi \left[\frac{\rho l}{2}(\dot{x}^2 + \dot{y}^2) - \kappa \tfrac{1}{2} \left(\sqrt{x'^2 + y'^2} - l \right)^2 \right]$$

So, we have a Lagrange density:

$$\mathcal{L} = \frac{\rho l}{2}(\dot{x}^2 + \dot{y}^2) - \kappa \tfrac{1}{2} \left(\sqrt{x'^2 + y'^2} - l \right)^2$$

The EL equation we must satisfy, to minimise the action, is (where ϕ is not that in this problem):

$$\frac{\partial \mathcal{L}}{\partial \phi} - \sum \frac{\partial}{\partial x^\mu} \frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} = 0$$

We must identify what these terms are. The $x^\mu = (\phi, t)$; and the $\phi = x, y$. So that the EL equations are the two:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x} - \frac{\partial}{\partial \phi}\frac{\partial \mathcal{L}}{\partial x'} - \frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot{x}} &= 0 \\ \frac{\partial \mathcal{L}}{\partial y} - \frac{\partial}{\partial \phi}\frac{\partial \mathcal{L}}{\partial y'} - \frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot{y}} &= 0 \qquad x' = \frac{\partial x}{\partial \phi} \quad \dot{x} = \frac{\partial x}{\partial t} \end{aligned}$$

Notice that the first term in each will both be zero. So, the equations gives (and these are the equations of motion):

$$\begin{aligned} -\kappa \frac{\partial}{\partial \phi} \left(\frac{x'}{\sqrt{x'^2 + y'^2}} \left[\sqrt{x'^2 + y'^2} - l \right] \right) - \rho l \ddot{x} = 0 \\ -\kappa \frac{\partial}{\partial \phi} \left(\frac{y'}{\sqrt{x'^2 + y'^2}} \left[\sqrt{x'^2 + y'^2} - l \right] \right) - \rho l \ddot{y} = 0 \end{aligned}$$

Then, under the approximation that the string is nearly circular, we write:

$$x(\phi) = (l + \lambda(\phi)) \cos(\phi + \psi(\phi)) \qquad y(\phi) = (l + \lambda(\phi)) \sin(\phi + \psi(\phi))$$

And it is possible (!!) to simplfy the equations of motion a little, using this.

4.7 Constrained Variational Problems

Let us consider what happens if we also have constraints on a minimisation problem.

4.7.1 Lagrange's Method of Undetermined Multipliers

If we wish to minimise a function $f(\boldsymbol{x})$, subject to a set of constraints $g_k(\boldsymbol{x}) = 0$, where $k = 1, 2, \ldots N$; then we can minimise the extended function:

$$F(\boldsymbol{x}, \lambda_1, \ldots, \lambda_N) = f(\boldsymbol{x}) - \lambda_1 g_1(\boldsymbol{x}) - \ldots - \lambda_N g_N(\boldsymbol{x})$$

With respect to $\boldsymbol{x}$ and λ_i. That is:

$$\frac{\partial F}{\partial \boldsymbol{x}} = 0 \qquad \frac{\partial F}{\partial \lambda_i} = g_i(\boldsymbol{x}) = 0$$

For example, suppose we have the extended function $F(x,y) = f(x,y) - \lambda g(x,y)$. Then we have:

$$\begin{aligned} \frac{\partial F}{\partial x} &= \frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = 0 \\ \frac{\partial F}{\partial y} &= \frac{\partial f}{\partial y} - \lambda \frac{\partial g}{\partial y} = 0 \\ \frac{\partial F}{\partial \lambda} &= g(x,y) = 0 \end{aligned}$$

Let us consider another way of thinking about this:

Consider a function $f(x,y,z)$. Then, to extremize the function, we must have:

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz = 0$$

That is, we require:

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = \frac{\partial f}{\partial z} = 0$$

By linear independence. Now consider that there is a set of constraints which are:

$$\varphi(x,y,z) = 0$$

That is:

$$d\varphi = \frac{\partial \varphi}{\partial x}dx + \frac{\partial \varphi}{\partial y}dy + \frac{\partial \varphi}{\partial z}dz = 0$$

So that:

$$df + \lambda d\varphi = \left(\frac{\partial f}{\partial x_i} + \lambda \frac{\partial \varphi}{\partial x_i} \right) dx_i = 0$$

So that we have:

$$\frac{\partial f}{\partial x_i} + \lambda \frac{\partial \varphi}{\partial x_i} = 0$$

When this set of 3 equations is satisfied, we have extremized f subject to the constraint φ.

Example Consider finding the stationary points of $f(x,y) = xy$, on the circle $x^2 + y^2 = 1$. Then, we have:

$$g(x,y) = x^2 + y^2 - 1$$

And we look for stationary points of the extended function:

$$F = xy - \lambda(x^2 + y^2 - 1)$$

And the stationary points are found via:

$$\begin{aligned}\frac{\partial F}{\partial x} &= y - 2\lambda x = 0 \\ \frac{\partial F}{\partial y} &= x - 2\lambda y = 0 \\ \frac{\partial F}{\partial \lambda} &= -(x^2 + y^2 - 1) = 0\end{aligned}$$

So, notice that we can solve these equations. The first two combine to give:

$$\lambda = \pm\frac{1}{2}$$

And the last equation, with the above equation, gives:

$$(x, y) = (\pm 1/\sqrt{2}, \pm 1/\sqrt{2})$$

Thus the positions of the minima found.

Let us now generalise this to functionals:

4.7.2 Generalisation to Functionals

We wish to look for the minimum of a functional $I[y]$, subject to the condition $J[y] = C$, where C is some constant. Again, we form the "extended functional":

$$K[y, \lambda] = I[y] - \lambda[J[y] - C]$$

And minimise $K[y]$.

Again, we shall proceed via example.

Example Let us find the curve of fixed length ℓ, with maximum area. This is infact called the "isoperimetric problem".

Consider some curve, with the origin somewhere inside the curve. We shall use polar coordinates (r, θ). The area inside the whole curve is just:

$$A[r] = \int_0^{2\pi} \frac{1}{2} r^2 d\theta$$

And the curve has length (using the standard line-element):

$$L[r] = \int_0^{2\pi} d\ell = \int_0^{2\pi} \sqrt{r^2 d\theta^2 + dr^2}$$

Which is:

$$L[r] = \int_0^{2\pi} \sqrt{r^2 + r'^2} d\theta \qquad r' \equiv \frac{dr}{d\theta}$$

Now, the constraint is that $L[r] = \ell$:

$$L[r] = \int_0^{2\pi} \sqrt{r^2 + r'^2} d\theta = \ell$$

So, the extended functional:

$$K[r] = A[r] - \lambda(L[r] - \ell)$$

Which is:

$$K[r] = \int_0^{2\pi} \frac{1}{2} r^2 d\theta - \lambda \left(\int_0^{2\pi} \sqrt{r^2 + r'^2} d\theta - \ell \right)$$

That is:

$$K[r] = \int_0^{2\pi} \left(\frac{1}{2} r^2 - \sqrt{r^2 + r'^2} \right) d\theta + \lambda \ell$$

So that to minimise the functional, we use the following first integral:

$$F - r' \frac{\partial F}{\partial r'} = k$$

This gives:

$$\frac{1}{2} r^2 - \lambda \frac{r'^2}{\sqrt{r^2 + r'^2}} = k$$

This is, unfortunately, very hard to solve. Although, we can immediately see one solution: $\lambda = 0$ gives r is a constant: a circle.

4.7.3 Example: The Catenary

This is another example.

What is the equilibrium curve of a flexible chain, of length ℓ density ρ, when we hang it between two points A & B.

So, we wish to minimise the gravitational potential energy:

$$E[y] = \int_A^B \rho g y \sqrt{1 + y'^2} dx$$

Subject to the constraint:

$$L[y] = \int_A^B \sqrt{1 + y'^2} dx = \ell$$

Where we have used the line element:

$$d\ell = \sqrt{1 + y'^2} dx$$

Throughout. So, we have extended functional:

$$K[y] = E[y] - \lambda(L[y] - \ell)$$

Giving:

$$K[y] = \int_A^B (y\rho g - \lambda)\sqrt{1 + y'^2}dx + \lambda\ell$$

We use the same first integral, giving:

$$\frac{\rho g y - \lambda}{\sqrt{1 + y'^2}} = C$$

We solve this by some substitution:

$$u \equiv y - \frac{\lambda}{g\rho} \qquad \alpha \equiv \frac{C}{g\rho}$$

Then, substituting this in, we get:

$$u = \alpha\sqrt{1 + u'^2}$$

Rearranging:

$$\alpha u' = \sqrt{u^2 - \alpha^2}$$

Which we can separate:

$$\int \frac{\alpha}{\sqrt{u^2 - \alpha^2}}du = \int dx$$

This has solution (by substitution, or looking up the standard integral):

$$u = \alpha \cosh\left(\frac{x - x_0}{\alpha}\right)$$

So that:

$$y = \frac{\lambda}{g\rho} + \alpha \cosh\left(\frac{x - x_0}{\alpha}\right)$$

THen the three constants α, x_0 & λ are found by the condition that $y(A)$ and $y(B)$ are known, as is:

$$\ell = \int \left(1 + \alpha^2 \sinh^2(x - x_0)\alpha\right)^{1/2} dx$$

4.7.4 Eigenvalue Problems

Consider the eigenvalue problem:

$$Lu = \lambda\rho u$$

Where L is an Hermitian operator, ρ is real and positive; and $x \in [a, b]$. Now, consider the functional:

$$I[u] = \int_a^b u^* L u dx$$

And let the u be normalised, by the functional:

$$N[u] = \int \rho u^* u dx = 1$$

Now, we shall show that by minimising I, we get the eigenvalue equation back out.

So, let us have the extended functional:

$$J[u] = I[u] - \lambda(N[u] - 1)$$

That is:

$$J[u] = \int u^* L u dx - \lambda \left(\int \rho u^* u dx - 1 \right)$$

Varying the functional:

$$\delta J = \int \delta u^* L u + u^* L \delta u dx - \lambda \int \rho(\delta u^* u + u^* \delta u) dx = 0$$

That is, collecting terms:

$$\delta J = \int \delta u^* (Lu - \lambda \rho u) + \delta u (Lu - \lambda \rho u)^* dx = 0$$

Now, we can vary u & u^*, rather than their real and imgainary parts. Hence:

$$Lu = \lambda \rho u$$

Thus recovered our eigenvalue problem.

4.7.5 Rayleigh-Ritz Method

Suppose that we have a constrained functional:

$$I[\psi] = \int \psi^* L \psi dx \qquad \int \psi^* \psi dx = 1$$

Where the constraint is that ψ is normalised. Then, $I[\psi]$ has a minimum at ψ_0.
Also suppose that we have some unconstrained functional:

$$K[\phi] = \frac{\int \phi^* L \phi dx}{\int \phi^* \phi dx}$$

And that K has a minimum at ϕ_0.
The Rayleigh-Ritz theorem is that $\phi_0 = \psi_0$. Infact, this seems obvious, as they are the same thing. Let us prove it.

Let us compute $K[\psi_0]$, at the minimum of I:

$$K[\psi_0] = \frac{\int \psi_0^* L \psi_0 dx}{\int \psi_0^* \psi_0 dx} = I[\psi_0]$$

Let us also compute $I[\Phi_0]$:

$$I[\Phi_0] = \int \Phi_0^* L \Phi_0 dx$$

Where we have that Φ_0 is related back to the minimum of I via:

$$\Phi_0 = \frac{\phi_0}{\sqrt{N}} = \frac{\phi_0}{\int \phi_0^* \phi_0 dx}$$

Then:

$$\begin{aligned} I[\Phi_0] &= \int \Phi_0^* L \Phi_0 dx \\ &= \frac{1}{N} \int \phi_0^* L \phi_0 dx \\ &= \frac{\int \phi_0^* L \phi_0 dx}{\int \phi_0^* \phi_0 dx} \\ &= K[\phi_0] \end{aligned}$$

We also note that:

$$I[\Phi_0] \geq I[\phi_0]$$

And we have equality if $\Phi_0 = \phi_0$; i.e. if the minima of I and K are the same. We also thus have:

$$K[\phi_0] \geq I[\psi_0]$$

So that the unconstrained minimum is always bigger than the constrained minimum, unless the unconstrained is the same as the constrained.
This is perhaps seen more clearly in the context of quantum mechanics: suppose that there exists an exact ground state eigenstate of a system. However, the exact form of the ground state is unknown. Then, we can compute the expectation value of the Hamiltonian, using some trial wavefunction. The expectation value is then minimised with respect to some variational parameter. Then, the trial ground state eigenstate will always yeild a higher ground state energy than the exact. The lowest possible trial ground state energy will occur when the exact ground state wavefunction is used as a trial:

$$\frac{\langle \psi_t | \hat{H} | \psi_t \rangle}{\langle \psi_t | \psi_t \rangle} \geq E_0$$

Where ψ_t is some trial wavefunction, and E_0 the exact ground state energy. We will get equality when, and only when, the exact ground state wavefunction has been used.

Let us find an approximation for the quartic anharmonic oscillator ground state. The Hamiltonian:

$$\hat{H} = -\frac{1}{2}\frac{d^2}{dx^2} + \frac{1}{2}x^2 + \frac{\lambda}{2}x^4$$

And the trial wavefunction is of the form:

$$\psi(x) = e^{-\alpha x^2/2}$$

Let us first write the normalisation integral:

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} dx = \sqrt{\frac{\pi}{\alpha}}$$

So, we compute:

$$K[\alpha] = \frac{\int \psi \hat{H} \psi dx}{\sqrt{\pi/\alpha}}$$

Which gives (after using some standard integrals):

$$K[\alpha] = \frac{1}{4}(\alpha + 1/\alpha + 3\lambda/2\alpha^2)$$

We then minimise this, with respect to α, that is:

$$\frac{\partial K}{\partial \alpha} = 0$$

This equation is then hard to solve!