PDE methods for Pricing Derivative Securities

Diane Wilcox

University of Cape Town

.

2005

Contents

Part I

Mathematical Theory: Parabolic PDE

Chapter 1

Introduction to Partial Differential Equations and Diffusion processes

Many phenomena are described by *functions whose values at a given point* (for example, in time) depends on *values at neighbouring points.* The equation which determines such a function usually contains derivatives of the function in order to capture information about *rates of change of the function with respect to the underlying variable(s) or parameter(s).* Such equations are referred to as *differential equations* .

1.1 Some general definitions

Formally, an **ordinary differential equation** (ODE) is an equation involving an unknown function u of one independent variable x and derivatives of that function. By appropriate rearrangement of terms, such an equation may be written as :

$$F(x, u, \frac{du}{dx}, \frac{d^2u}{dx^2}, \frac{d^3u}{dx^3}, \ldots) = 0. \tag{1.1}$$

The **order** of the equation is the order of the highest derivative and the equation is said to be **linear** if u and all it's derivatives are of the 1st degree (there are no terms of the form u^3 or $(\frac{du}{dx})^2$, etc ...) and there are no products of u and it's derivatives.

A general nth order linear ODE may be written as :

$$a_0(x)u(x) + a_1(x)\frac{du}{dx} + \ldots + a_n(x)\frac{d^nu}{dx^n} = f(x), \tag{1.2}$$

where $a_i(x),\ i = 1, \ldots n$ are known functions of x. If $f(x) = 0$, then equation (1.2) is said to be **homogeneous**.

A **partial differential equation** (PDE) is an equation involving an unknown function u of two or more one independent variables and partial derivatives of that function. By appropriate rearrangement of terms, such an equation in two independent variables x and y may be written as

$$F(x, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial^2 u}{\partial x^2}, \frac{\partial^2 u}{\partial x \partial y}, \frac{\partial^2 u}{\partial y^2}, \ldots, \frac{\partial^k u}{\partial x^r \partial y^s}, \ldots) = 0, \tag{1.3}$$

where $k = r + s$.

As before, the **order** of the equation is the order of the highest derivative and the equation is said to be **linear** if u and all it's derivatives are of the 1st degree (i.e. there are no terms of the form u^2 or $(\frac{\partial^2 u}{\partial x^2})^3$, etc...) and there are no products of u and it's derivatives.

A 1st order linear PDE may be written as :

$$a(x,y)\frac{\partial u}{\partial x} + b(x,y)\frac{\partial u}{\partial y} + c(x,y)u = g(x,y). \tag{1.4}$$

The general form of a 2nd order linear PDE is given by

$$au_{xx} + 2bu_{xy} + cu_{yy} + du_x + eu_y + fu = g, \tag{1.5}$$

where a, b, c, d, e and f are known functions of x and y and $u_x := \frac{\partial u}{\partial x}$, $u_{xx} := \frac{\partial^2 u}{\partial x^2}$ and $u_{xy} := \frac{\partial^2 u}{\partial x \partial y}$, etc ...

If $g = 0$ in equation (1.4) or equation (1.5) then the equation is said to be **homogeneous**.

1.2 Examples

The best known PDE in the mathematics of financial markets is, of course, the Black-Scholes equation for valuing options:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0,$$

where S denotes stock price, t denotes the time, and V, the value of the option, is a function of S and t. The equation itself is quite a general object and one cannot speak of a solution without reference to additional constraints, such as possible values which the option may attain when exercised. These conditions specify the valuation problem at hand.

Historically, three 2nd-order linear PDE have been of fundamental interest because they exhibit distinct behaviour and have led to the clarification of general theories and methods:

Let $u(x) : \mathbb{R}^2 \to \mathbb{R}$ be a function of two variables.

- The **Laplace equation** originated out of problems in potential theory in physics :

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

- The **Wave Equation** describes the amplitudes $u(x,t)$ of vibrating membranes and other wave functions :

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 u}{\partial t^2}.$$

- The **Heat Equation** describes the temperature distribution $u(x,t)$ in the plane when heat is allowed to flow from warm areas to cool ones :

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial t}.$$

The above examples are models of problems with 2 independent variables. More generally we may let $u(x) : \mathbb{R}^n \to \mathbb{R}$ be a function of $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ and the above examples may be formulated naturally in higher dimensions. Since methods of solution often generalise to higher dimensions, it is often sufficient to understand examples for $n = 1, 2$ or 3. However, it is sometimes the case that a particular method for solving is not suited to higher dimensions.

1.3 Classification of 2nd order linear PDE

The behavior of known solutions of the classic 2nd order PDE's indicates that the presence of the higher order terms is significant. Thus, for a general theory of 2nd order PDE's, we consider an equivalent form of equation (1.5):

$$a(x,y)\frac{\partial^2 u}{\partial x^2} + 2b(x,y)\frac{\partial^2 u}{\partial x \partial y} + c(x,y)\frac{\partial^2 u}{\partial y^2} = F(x,y,u,u_x,u_y). \tag{1.6}$$

The **discriminant**[1] of a 2nd order linear PDE is defined as :

$$D := ac - b^2. \tag{1.9}$$

The PDE is said to be

$$\begin{aligned} \textit{elliptic} \quad & if \quad D > 0, \\ \textit{parabolic} \quad & if \quad D = 0, \\ \textit{hyperbolic} \quad & if \quad D < 0. \end{aligned}$$

N.B.: If a, b or c are function of the independent variables, then the discriminant varies with the values of these variables.

By this classification,

the *heat equation is parabolic* everywhere ($a = 1, b = 0$ and $c = 0$),

the *Laplace equation is elliptic* everywhere and

the *wave equation is hyperbolic* everywhere.

[1]Alternatively, if we consider the general form (1.5):

$$a(x,y)\frac{\partial^2 u}{\partial x^2} + b(x,y)\frac{\partial^2 u}{\partial x \partial y} + c(x,y)\frac{\partial^2 u}{\partial y^2} = F(x,y,u,u_x,u_y).. \tag{1.7}$$

then the discriminant is equivalently defined as :

$$D := 4ac - b^2. \tag{1.8}$$

Sometimes, it defined $D := b^2 - 4ac$. In this case, the inequalities for the classification of PDE as elliptic or hyperbolic must be reversed.

Examples 1.3.1. *For the PDE,* $2xu_{xx} - u_{tt} = 0$*, the discriminant is* $D = -2x$*. Thus, the equation is*

$$\begin{aligned} &elliptic \quad if \quad x < 0, \\ &parabolic \quad if \quad x = 0, \\ &hyperbolic \quad if \quad x > 0. \end{aligned}$$

We will see that the Black-Scholes equation is also a parabolic equation. In fact, by a suitable change in variables it is possible to transform the BS equation into one which has the form of the heat equation.

The flow of heat through a solid, and the continuous Brownian motion of particles in a liquid are examples of **diffusion processes**. These processes can be modelled by the heat equation. The *forward Kolmogorov equation*, also known as the *Fokker-Planck* equation, is another parabolic equation derived in the modelling diffusion processes :

$$\frac{\partial u}{\partial t} = -(c\frac{\partial u}{\partial x}) + \frac{1}{2}D\frac{\partial^2 u}{\partial x^2},$$

where c and D are constants.

1.4 Function spaces and operators

A **Vector space** V is a non-empty set of vectors $x, y, z, \ldots$ endowed two operations referred to as *addition* and *scalar multiplication* (i.e. multiplication by a scalar quantity) such that the following axioms are satisfied:

Let $\mathbb{K}$ denote $\mathbb{R}$ or $\mathbb{C}$ (scalar field). For any x, y and $z \in V$ and $\alpha, \beta \in \mathbb{K}$ we have

	Property	Interpretation
V1	$x + y$ is a vector in V	V is *closed* under addition
V2	$x + y = y + x$	Addition is *commutative*
V3	$(x + y) + z = z + (y + z)$	Addition is *associative*
V4	There exists a unique vector $0 \in V$	V has an *additive identity*
	such that $x + 0 = 0 + x = x$	
V5	For every $x \in V$ there exists a unique vector	For every $x \in V$ there exists an
	$-x \in V$ such that $x + (-x) = 0 = (-x) + x$	*additive inverse*
V6	$\alpha x \in V$	V is *closed* under scalar multiplication
V7	$\alpha(x + y) = \alpha x + \alpha y$	*Distributivity I*
V8	$(\alpha + \beta)x = \alpha x + \beta x$	*Distributivity II*
V9	$\alpha(\beta x) = (\alpha\beta)x$	Multiplication is *associative*
V10	If $\alpha = 1$, then $\alpha x = x$	*Scaling by unity*

Remarks : Because of this linear structure, vector spaces are often referred to as *linear spaces.* To establish that a set is a vector space, one need only verify that it is closed under addition and scalar multiplication.

Examples 1.4.1.

1. *The spaces $\mathbb{R}^n$ or $\mathbb{C}^n$ of n-tuples of real and complex numbers, respectively, are vector spaces under usual addition and scalar multiplication.*

2. *The set $M_n(\mathbb{R})$ of $n \times n$ matrices with real entries, is a vector space under matrix addition.*

3. *The set P_n of polynomials of degree at most n is a vector space under the usual addition and scalar multiplication of functions*[2].

4. *The set $C([a,b]) := \{f : [a,b] \to \mathbb{R} \mid f$ is continuous on $[a,b]$ $\}$ is a vector space under the usual addition and scalar multiplication of functions.*

5. *If $\Omega \subset \mathbb{R}^n$ then the set*

$$C(\Omega) := \{\ u : \Omega \to \mathbb{R} \mid u \ \text{ is continuous on } \Omega\ \}$$

of real-valued functions of n variables is a vector space. Similarly, for $m \geq 1$,

$$C^m(\Omega) := \{u : \Omega \to \mathbb{R} \mid u \text{ and its partial derivatives to order } m \text{ are continuous on } \Omega\}$$

of real-valued functions of n variables is a vector space.

Definition 1.4.2. *Given vector spaces V and W, an* **operator** *$T : V \to W$ is a mapping from V into W which assigns a vector $w = Tv \in W$ for each vector v in the domain $D(T)$ of T.*

Examples 1.4.3.

1. *An $r \times s$ matrix is an operator from $\mathbb{R}^r$ into $\mathbb{R}^s$.*

2. *Let $\Omega \subset \mathbb{R}$. Then $D : C^1(\Omega) \to C(\Omega)$ given by $Df := \frac{df}{dx}$ is a* differential operator.

3. *The operator $L : C^n(\Omega) \to C(\Omega)$ given by*

$$Lu := \frac{\partial^\alpha u}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}}, \quad \sum_{i=1}^{n} \alpha_1 = \alpha$$

is a partial differential operator.

Definition 1.4.4. *A* **linear operator** *is an operator $L : V \to W$ which satisfies*

$$L(\alpha x + \beta y) = \alpha L(x) + \beta L(y)$$

for all $x, y \in V$ and scalars α and β.

Examples 1.4.5. *The operators in Examples 1.4.3 are all linear operators.*

[2] $(f+g)(x) = f(x) + g(x)$ and $(\alpha f)(x) = \alpha(f(x))$

The following linear operators are basic to the study of PDE :

1)	Laplace operator	L $= \Delta := \frac{\partial^2}{\partial x_1^2} + \ldots + \frac{\partial^2}{\partial x_n^2}$
2)	Diffusion operator	L$= \frac{\partial}{\partial t} - \Delta$
3)	D'Alembert operator	L$= \frac{\partial^2}{\partial t^2} - \Delta$

The above operators are referred to as elliptic, parabolic and hyperbolic operators, respectively, by the classification of the associated equation $Lu = f$. By a suitable change of variables, an elliptic operator L may be transformed into the Laplacian in an alternative co-ordinate system. Similarly, parabolic and hyperbolic operators may be transformed into the diffusion and d'Alembert operators, respectively.

1.5 Linear DE and the principle of superposition

Recall that the operation of differentiation acts as a *linear transformation* on the set of continuous real-valued functions which have 1st derivatives. In other words, if f and g are continuous real-valued functions which have 1st derivatives, and α and β are scalars (i.e. $\alpha,\ \beta \in \mathbb{R}$ are constants) then

$$\frac{d}{dx}(\alpha f + \beta g) = \alpha\frac{d}{dx}f \ + \ \beta\frac{d}{dx}g.$$

The same is true for partial differentiation. Thus, linear partial differential equations may be written in the form $Lu = f$, where u is unknown and f is given function, and in general, a linear partial differential operator, L, transforms a function u of $x = (x_1, x_2, \ldots x_n) \in \mathbb{R}^n$, into another function $L(u)$ given by:

$$\begin{aligned} L(u) \quad &= \quad a(x)u \ + \ b_1(x)\frac{\partial u}{\partial x_1} \ + \ b_2(x)\frac{\partial u}{\partial x_2} \ + \ \ldots \ + \ b_n(x)\frac{\partial u}{\partial x_n} \\ &\qquad + c_{11}(x)\frac{\partial^2 u}{\partial x_1^2} \ + \ c_{12}(x)\frac{\partial^2 u}{\partial x_1 \partial x_2} \ + \ \ldots + c_{nn}(x)\frac{\partial^2 u}{\partial x_n^2} \ + \ \ldots \\ &= \quad a(x)u \ + \ \sum_{i=1}^{n} b_i(x)\frac{\partial u}{\partial x_i} \ + \ \sum_{i,j=1}^{n} c_{ij}(x)\frac{\partial^2 u}{\partial x_i \partial x_j} \ + \ldots, \end{aligned}$$

where the dots indicate higher-order terms, but it is understood that the sum contains only finitely many terms. Clearly we may write L independently of the function u as follows:

$$L := \ a(x) \ + \ \sum_{i=1}^{n} b_i(x)\frac{\partial}{\partial x_i} \ + \ \sum_{i,j=1}^{n} c_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} \ + \ldots$$

If $u_1, \ldots u_k$ are functions which are *sufficiently smooth,* i.e. $u_1, \ldots u_k$ have enough derivatives for $L(u_i)$ to be defined for $1 \le i \le k$, and if $c_1, \ldots c_k$ are scalars (constants) then

$$L(c_1 u_1 + \ldots + c_k u_k) \ = \ c_1 L(u_1) + \ldots + c_k L(u_k).$$

For *infinite* linear combinations such that the series

$$\sum_{k=1}^{\infty} c_k u_k \qquad \textit{and} \qquad \sum_{k=1}^{\infty} c_k L(u_k)$$

both converge, we have

$$L\,[\sum_{k=1}^{\infty} c_k u_k\,] \;=\; \sum_{k=1}^{\infty} c_k L(u_k).$$

For integrals, we generalise as follows: suppose $u(x,\lambda)$ is a function of $x \in \mathbb{R}^n$ and of the parameter λ, where $a < \lambda < b$, and suppose $g(\lambda)$ is an integrable function of λ on (a,b), then

$$L\,[\int_a^b g(\lambda)u(x\lambda)d\lambda\,] \;=\; \int_a^b g(\lambda)\,L\,[u(x\lambda)]\,d\lambda.$$

Here we may think of $\int_a^b g(\lambda)u(x\lambda)d\lambda$ as the continuous form of the series $\sum_{k=1}^{\infty} c_k u_k$ in which the coefficients c_k are replaced with $g(\lambda)$ in the "continuous" sum.

The **principle of superposition** states :

> *If u_i satisfy the a linear homogenous equation, then an arbitrary linear combination $c_1u_1 + \ldots + c_ku_k$ satisfies the same homogeneous equation.*

Symbolically, we may write this as follows :

$$L(u_k) = 0 \;\; for \;\; all \;\; k \quad \Rightarrow \quad L[\sum_{k=1}^{\infty} c_k u_k\,] = \sum_{k=1}^{\infty} c_k L(u_k) = 0.$$

For the integral (continuous sum) version , the principle of superposition states:

$$L[u(x,\lambda)] = 0 \;\; for \;\; a < \lambda < b \quad \Rightarrow \quad L\,[\int_a^b g(\lambda)u(x,\lambda)d\lambda\,] \;=\; \int_a^b g(\lambda)L[u(x,\lambda)]d\lambda \;=\; 0.$$

1.6 Types of problems within the solution of a PDE

(1) *Boundary Value Problems (BVP)*

e.g. the *Dirichelet problem* on a domain $\Omega \subset \mathbb{R}^n$, where Ω is open and connected[3], refers to the problem of finding a function u which satisfies the Laplace equation within Ω, i.e.

$$\Delta u(x) = 0 \text{ for } x \in \Omega$$

and which obeys the additional condition that

$$u(x) = f \text{ for } x \in \partial\Omega.$$

where $\partial\Omega$ denotes the boundary of Ω and f is a known function defined on $\partial\Omega$.

(2) *Initial (and Final) Value Problems (IVP)*

e.g. the heat equation specified as:

$$\begin{aligned} u_t - u_{xx} &= 0 \quad\;\; for \;\; -\infty < x < \infty, \;\; t > 0 \\ u(x,0) &= f(x) \quad for \;\; -\infty < x < \infty, \;\; t = 0, \end{aligned}$$

[3]for now we may take these conditions to simply mean that Ω is a nice and unbroken subset of $\mathbb{R}^n$

where f denotes some initial temperature distribution and the value of $u(x,t)$ evolves with time.

e.g. the Black-Scholes PDE for which the payoff function is given. Here the payoff function is a *final condition* and depends on S_T, the price of the stock at some expiry (final) time T.

$$\begin{aligned} \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV &= 0 \qquad for \;\; 0 \le t < T \\ V(S_T, T) &= payoff \;\; at \;\; t = T, \end{aligned}$$

where S denotes stock price, t denotes the time, T is the expiry date and V, the value of the option, is a function of S and t. The BS-PDE can be transformed in to an initial value problem by letting $\tau := T - t$. The final condition $V(S_T, T)$ is now an initial condition in the variable τ, i.e. $V(S_\tau, 0) = payoff$ at $\tau = 0$. The PDE is referred to as a *backward PDE*:

(3) *Eigenvalue Value Problems (EVP)*

Definition 1.6.1. *Let $L : V \to W$ be a linear operator. A non-zero vector $v \in V$ is said to be an* **eigenvector** *for L if there exists a scalar λ such that $Lv = \lambda v$. If V is a function space, then an eigenvector is often referred to as an* **eigenfunction**.

Consider the following ODE[4]

$$\begin{aligned} v'' &= \lambda v, \quad 0 < x < \pi \\ v(0) = v(\pi) &= = 0, \end{aligned}$$

where $v : \mathbb{R} \to \mathbb{R}$ and $v' := \frac{dv}{dx}$. Clearly $v = 0$ is a solution. We are interested in non-zero solutions, i.e. we want to find the *eigenfunctions* $v_n(x)$ and their corresponding *eigenvalues* for the problem $Lv = \lambda v$, where $L = \frac{d^2}{dx^2}$.

1.7 Random walk derivation of a diffusion equation

Let $(\Omega, \mathbb{P})$ denote a probability space. A discrete valued random variable (r.v.), x, defined on all of Ω assumes values x_i with probability p_i ($\sum_i p_i = 1$). Its expectation and variance are

$$\begin{aligned} \mathbb{E}(x) &:= \sum_i x_i p_i, \quad \text{denoted } \langle x \rangle \\ \text{var}(x) &:= \left\langle x - \langle x \rangle^2 \right\rangle = \langle x^2 \rangle - \langle x \rangle^2. \end{aligned}$$

Consider an unrestricted 1-D random walk: A particle at the origin executes random steps $+\delta$ (right) or $-\delta$ (left). Let x_i denote the r.v. which assumes the values $+\delta$ or $-\delta$ at the i^{th} step; assume that each step is independent of the others. By construction, the x_i are independent, identically distributed (i.i.d.) r.v. The position of the particle after the n^{th} step is given by

$$X_n = \sum_{i=1}^{n} x_i.$$

[4] such problems arise in the method of *separation of variables* and *Sturm-Liouville problems* which we will meet in subsequent chapters.

Let

$$\begin{aligned}\mathbb{P}(x_i = +\delta) &= p \\ \mathbb{P}(x_i = -\delta) &= q,\end{aligned}$$

where $p + q = 1$. It follows that

$$\begin{aligned}\mathbb{E}(X_n) &= \mathbb{E}(\sum_{i=1}^{n} x_i) = \sum_{i=1}^{n} \mathbb{E}(x_i) = (p-q)\delta n && (1.10) \\ \mathbb{E}(x_i^2) &= (+\delta)^2 p + (-\delta)^2 q = \delta^2 \\ \mathrm{var}(X_n) &= \mathrm{var}(\sum_{i=1}^{n} x_i) = \sum_{i=1}^{n} \mathrm{var}(x_i) = \sum_{i=1}^{n} [\langle x_i^2\rangle - \langle x_i\rangle^2] \\ &= \sum_{i=1}^{n} [\delta^2 - (p-q)^2\delta^2] = 4pq\delta^2 n. && (1.11)\end{aligned}$$

For the continuous case, consider the following mathematical model for physical Brownian motion. Experimentally, the average displacement of the particle per unit time is c, and the variance of the observed displacement is $D > 0$.

Suppose the particle executes r jumps per unit time. Then from (1.10) and (1.11) we have

$$\begin{aligned}(p-q)r\delta &\approx c && (1.12) \\ 4pq\delta^2 r &\approx D. && (1.13)\end{aligned}$$

The motion of particles appear continuous, so we examine behaviour under the limit $\delta \to 0$ and $r \to 0$, constrained by the conditions that c and D remain fixed. If $p \neq q$ and

$$\lim_{\delta\to 0}(p-q) \neq 0 \neq \lim_{r\to\infty}(p-q),$$

then as $\delta \to 0,\ r \to \infty$ jointly, we have

$$\delta r \longrightarrow \frac{c}{p-q}.$$

Hence, as $\delta \to 0,\ r \to \infty$ jointly,

$$4pq\delta^2 r \longrightarrow \left(\frac{4pqc}{p-q}\right)\delta \longrightarrow \infty.$$

But, from (1.13),

$$4pq\delta^2 r \longrightarrow D \neq 0.$$

Thus, $\lim(p-q) \to 0$, with $p + q = 1$, so that $\lim p = \lim q = \frac{1}{2}$. Now $p = q = \frac{1}{2}$ implies that $c = 0$ and $D = 0$, which in turn implies that there is no variance and the process is deterministic. Thus, (1.10) and (1.11) hold if we have

$$\begin{aligned}p &= \frac{1}{2}(1+b\delta) \\ q &= \frac{1}{2}(1-b\delta),\end{aligned}$$

where b is a constant such that $0 \leq p,\ q \leq 1$ and $p+q=1$. In this case we have $p-q=b\delta$. It follows that as $\delta \to 0,\ r \to \infty$ jointly, $\delta^2 r \to D,\ p \to \frac{1}{2},\ q \to \frac{1}{2}$ and $b = \frac{c}{D}$.

Now there are r steps per unit time if and only if 1 step of length δ occurs in time $\tau := \frac{1}{r}$ units.

Returning to the motion of our particle starting at $x=0,\ t=0$, we are interested in the probability of the particle being at position x at time t after n steps, i.e. we wish to know

$$v(x,t) := \mathbb{P}(X_n = x,\ n\tau = t).$$

For the continuum limit, we construct difference equations for $v(x,t)$ and show that it converges to the a PDE. We have (the master equation) :

$$v(x,t+\tau) = p\,v(x-\delta,t) + q\,v(x+\delta,t). \tag{1.14}$$

Expanding the Taylor series

$$\begin{aligned} v(x,t+\tau) &= v(x,t) + \tau v_t(x,t) + O(\tau^2) \\ v(x\pm\delta,t) &= v(x,t) \pm \delta v_x(x,t) + \frac{1}{2}\delta^2 v_{xx}(x,t) + O(\delta^3), \end{aligned} \tag{1.15}$$

where $O(y^k)$ means that $\lim\limits_{y\to 0} \frac{O(y^k)}{y^k} < \infty$[5]. Substituting (1.15) into (1.14), we obtain

$$v_t(x,t) = \left((q-p)\frac{\delta}{\tau}\right) v_x(x,t) + \frac{1}{2}(\frac{\delta^2}{\tau})v_{xx}(x,t) + O(\tau) + O(\frac{\delta^3}{\tau}). \tag{1.16}$$

As $\delta \to 0$ and $\tau \to 0$, equation (1.16) converges to

$$v_t = -cv_x + \frac{1}{2}Dv_{xx},$$

and $v(x,t)$ may interpreted as the probability density function for the continuous r.v. x at time t.

1.8 Heat flux derivation of a diffusion equation

Suppose the problem is to model heat flowing through a medium, e.g. a gas or a liquid, located in 3-D space.

Let $x = (x_1, x_2, x_3)$ denote the co-ordinates of a point in space. We want to find the temperature distribution in the medium, $u(x,t)$. The value of $u(x,t)$ will depend on

[5]i.e. the error due to ignoring higher order terms is smaller in absolute value than some constant times y^k if y is close enough to 0.

- the presence of heat in the body and
- the flow of heat into or out of the body through the surface.

There are two laws (= equations) which model the situation :

1. The *balance* or *conservation* equation for the conservation of energy.
2. The *constitutive equation* which describes how heat flows in the medium.

Assuming there are no other types of energy in the situation, (1) can be stated as follows:

$$\begin{array}{ccc} \textit{rate of change} & & \textit{heat generated by sources within the body} \\ \textit{of} & = & + \\ \textit{thermal energy in body} & & \textit{flow of heat into body from outside.} \end{array}$$

Using the notation :

$$\begin{array}{rl} x = & (x_1, x_2, x_3) = \text{position in space} \\ t = & \text{time} \\ \triangle = & \text{small volume element } (= \triangle x_1 \triangle x_2 \triangle x_3 \text{)} \\ R = & \text{region occupied by medium} \\ S = & \text{surface of medium} \\ c = & \text{heat capacity (= amount of heat generated per unit mass and per unit rise} \\ & \text{in temperature)} \\ \rho = & \text{mass density} \\ u = & \text{temperature} \\ f(x,t) = & \text{energy generated inside the body per unit volume and per unit time} \\ q(x,t) = & \text{(normal component of) heat flux through surface (which may be negative)} \end{array}$$

and assuming that properties of the medium are identical throughout the medium (the medium is said to be *homogeneous* in this case), we have :

$$\begin{array}{rcl} \text{total thermal energy in region} & = & \int\limits_R c(x)\rho(x)u(x)dV \\ & (\approx & \sum\limits_{\triangle} c.(\rho.\triangle).u \text{) ,} \end{array}$$

$$\text{total energy from source} \quad = \quad \int\limits_R f(x,t)dV,$$

and

$$\begin{array}{rcl} \text{total flux through surface} & = & -\oint\limits_S q(x,t).ndA \\ & = & -\int\limits_R div\, q.dV, \end{array}$$

where n is the outward normal vector to S and the last equality follows from the divergence theorem. Thus, from the balance equation, we have :

$$\frac{d}{dt}\int\limits_R cpu\, dV = \int\limits_R f\, dV - \int\limits_R div\, q\, dV \tag{1.17}$$

or

$$\int_R (cp\frac{du}{dt} - f + div\, q)dV = 0.$$

Thus, since the volume is arbitary and the integrand is *smooth*,

$$cp\frac{du}{dt} - f + div\, q = 0.$$

The *constitutive equation* is given by *Fourier's law* :

$$\text{heat flux} \;=\; q \;=\; -k\nabla u,$$

where k is referred to as the *thermal conductivity function* and depends on the material of the medium[6]. Thus (1.17) becomes

$$\frac{\partial u}{\partial t} - div(k\nabla u) = Q, \qquad Q = \frac{f}{cp}, \tag{1.18}$$

which is referred to as the *heat equation.*

To solve this equation we need information about u on S and u at time $t = 0$.

The following properties may be deduced from this equation :

1. heat energy is conserved
2. heat flows in a manner that distributes evenly throughout the medium (from hot to cold and faster when there is greater initial temperature differences).

1.9 On mathematical modelling and Solving PDE

Very broadly, the process of solving a problem by mathematical methods involves three components, namely:

(1) *formulating the problem mathematically,*

(2) *solving the mathematical problem, and*

(3) *interpretting the solution.*

Once a model has been formulated, one needs to know :

---

[6]Here

- $k < 0$ implies heat flows from hot to cold
- $\nabla u = 0$ implies temperature is constant and there is no heat flow
- $\nabla u =$ large implies q is large, i.e. the greater the temp. difference, the greater the flow.

- *does a solution exists?*

This is sort of question is referred to as the *existence problem.* For example, there are some mathematical formulations of minimisation problems which do not have solutions. In this case, the problem probably does not lie with the math, but with the formulation. [7] We will be fortunate to work with problems for which solutions are known to exist. The simplest way to prove existence is to *construct a solution* that meets all the conditions originally imposed. One also is interested:

- *is the solution unique ?*

- *is the solution stable ?* [8]

These questions lead to the following definition:

Definition 1.9.1. *A solution to a PDE is said to* **depend continuously on data** *if small changes in the data (boundary or initial conditions) produce small changes in the solution. A problem is said to be* **well-posed** *if*

(1) a solution to the problem exists,

(2) the solution is unique, and

(3) the solution depends continuously on data

In general, the number of initial / boundary conditions which are required to solve a PDE uniquely, depends on several factors - in general, it is complicated to determine the appropriate form for data needed for a solution. The notion of well-posedness was proposed by Hadamard to serve as set of a *guidelines* for formulating a problem and determine which data are necessary for a solution to be found [9]. The last condition is important for many applications - it is preferred that the solution changes only a little when conditions which specify the problem change a little. More can be said on the well-posedness of a problem with the aid of deeper analysis of geometry of the classification of 2nd order PDE's - we will not investigate that here.

Another question is :

- *what method was used to construct the solution ?*

This is particularly relevant when numerical methods are used to solve the problem: one wants to know how good an approximation is and is concerned about the error in the solution, how long an iteration takes to converge, etc... Rigour is always important when there may be exceptional cases to a seemingly logical argument. Later we will see that when using a series expansion to represent a solution, one must be sure that the series converges!

---

[7] Problems of proving the existence of solutions have been important in the development of mathematics. The Dirichelet problem, for example, had a significant early role in the development of functional analysis.

[8] Here we are asking whether the solution changes significantly if a part of the input information / data is changed only very slightly.

[9] When the problem is set in a homogeneous medium, when the shape of the boundary is "nice" and when the boundary data is simple, then solutions to PDE exist. In particular, there are always solutions for PDE in this course!

And another is:

- how differentiable is the solution found ?

This refers to the problem if *regularity* of the solution. Depending on the model, it may be required that the solution be very regular, say $k-$times continuously differentiable. Such problems may be really hard to solve and proofs would have to include verifications that solutions are smooth enough. An alternative strategy is to consider the *existence* and *smoothness* problems separately. The idea is to consider the PDE for a very wide class of functions and to find a *weak solution*. Since not much is being expected of such a solution, it is anticipated that questions on existence, uniqueness and stability are easier to solve. Once it is establised that the problem is well-posed for a broad class of generalised solutions, the question of regularity of solutions may be addressed - for some problems, it may be that weak solutions are smooth enough to qualify as a meaningful solution[10].

In applications, we are not interested in general solutions - we usually seek some specific solution to the given problem, where the unique solution is specified by additional data, i.e. boundary conditions. In the case where one of the variables is a time parameter, we may have some initial or final condition for the problem.

In the case of the diffusion equation, which contains a time dependence, we have will find it appropriate to specify the function u at initial time $t = 0$, i.e. the initial value of the density distribution is given as part of the problem . For the wave equation, which also contains time dependency, both $u(x, t)$ and $u_t(x, t)$ are usually specified, since the problem usually involves a description of the vibrations of the system as a result of some initial impulse.

When the values of x are constrained to a bounded or partially bounded region, then $u(x, t)$ and/or $u_x(x, t)$ (or a linear combination of these) must be given on the boundaries for all $t > 0$. Even when the region is not constrained, there are usually explicit or implicit conditions for the behaviour of the solution at infinity.

Boundary value problems for parabolic and hyperbolic equations may not well-posed. Solutions to these equations evolve in time, and their behaviour at later times is given by previous states. Thus, a boundary value problem which arbitrarily specifies the values of the solution at two or more distinct times, is usually *not* reasonable.

Examples 1.9.2.

Consider the hyperbolic equation $u_{xy} = 0$ *on the square* $[0, 1] \times [0, 1]$. *Since* $u_{xy} = 0$, *it follows that* u_x *is constant as a function of the parameter* y, *and in particular,* $u_x(x, 0) = u_x(x, 1)$. *Now if* $u_x(x, 0) = f'(x)$ *and* $u_x(x, 1) = g'(x)$ *are given by the boundary conditions* $u(x, 0) = f(x)$ *and* $u(x, 1) = g(x)$, *then the problem cannot be solved for arbitrary* f *and* g; *a solution exists only if* $f'(x) = g'(x)$.

[10]We will see more of this notion when we consider Green's functions and solutions to PDE

Chapter 2

Fourier series methods

2.1 Fourier's solution to the heat equation - method of separation of variables

The problem which Joseph Fourier (1768 - 1830) considered in his paper *On the Propagation of Heat in Solid Bodies*[1] : was

find a solution for the diffusion of heat in a 1-dimensional rod of length l

given by the equation with boundary conditions and initial condition:

$$\begin{aligned} \frac{\partial u}{\partial t} - k\frac{\partial^2 u}{\partial x^2} &= 0, && (2.1)\\ u(0,t) &= u(l,t) = 0 \quad for \quad t > 0,, && (2.2)\\ u(x,0) &= f(x) \quad for \quad 0 < x < l. && (2.3) \end{aligned}$$

His solution can be summarised as follows:

Let $u(x,t) = \phi(x)\varphi(t)$. Then, substituting into equation (2.1), we get

$$\frac{\phi_{xx}(x)}{\phi(x)} = \frac{k\varphi_t(t)}{\varphi(t)}. \tag{2.4}$$

Since equation (2.4) holds for all x such that $0 \leq x \leq l$, and for all $t > 0$, the LHS and RHS must be equal to some constant, say $-\lambda$. Thus, (2.4) is given by:

$$\begin{aligned} \phi_{xx}(x) + \lambda\phi(x) &= 0, \quad and && (2.5)\\ \varphi_t(x) + \lambda k\varphi(t) &= 0. && (2.6) \end{aligned}$$

Now the boundary conditions give at (2.2) imply that $\phi(0) = \phi(l) = 0$. Thus, the solution to the eigenvalue problem (2.5) can be written:

$$\phi(x) = c_n . \sin(\frac{n\pi x}{l}), \tag{2.7}$$

[1] His work on the topic began in 1804 and was completed in 1807. Under review by Lagrange, Laplace, Monge and Lacroix, the work was highly regarded but criticized, mainly for the controversial trigonometric expansions for the representation functions. In this regard, his arguments were considered incomplete and not general.

with corresponding eigenvalues

$$\lambda_n = (\frac{n\pi}{l})^2.$$

The solution to equation (2.6) is then given by

$$\varphi(t) = e^{-\lambda_n kt}. \tag{2.8}$$

Thus, combining (2.7) and (2.8), it follows that we have solutions

$$u_n(x,t) = c_n \sin(\frac{n\pi x}{l}) e^{-(\frac{n\pi}{l})^2 kt}\ , \quad n \in \mathbb{N}. \tag{2.9}$$

Since (2.1) is a linear equation, it follows from (2.9) that our general solution can be written

$$\begin{aligned} u(x,t) &= \sum_{n=1}^{\infty} u_n(x,t) \\ &= \sum_{n=1}^{\infty} c_n \sin(\frac{n\pi x}{l}) e^{-(\frac{n\pi}{l})^2 kt} \end{aligned} \tag{2.10}$$

with the initial condition that $u(x,0) = f(x)$. The problem is solved with :
provided

$$f(x) = \sum_{n=1}^{\infty} c_n \sin(\frac{n\pi x}{l}). \tag{2.11}$$

This last statement leads to the question: *when can a function be represented as a trigonometric series?*

A rigorous answer for pointwise representation was first given by Dirichet in 1829. The contemporary form of a much more powerful theorem on the uniform convergence of a series representation for square-integrable functions was developed only after the theory of Hilbert spaces was fleshed out.

2.2 Trigonometric Series and pointwise convergence

We continue with the question : *which functions admit trigonometric representations of the form (2.11)?* Since these are sums of periodic functions, a natural place to begin is with the class of periodic functions (since we asking about the convergence of a series made up of periodic functions).

Definition 2.2.1. *A function f defined on the real line is said to be* **periodic** *with period P if $f(x) = f(x + P)$ for all $x \in \mathbb{R}$.*

Recall that the sine function satisfies $\sin(x) = -\sin(-x)$ and that the cosine function satisfies $\cos(x) = \cos(-x)$ and we have the following general definitions:

Definition 2.2.2. *A function f defined on an interval $[a, b]$ is said to be* **even** *if it satisfies $f(x) = f(-x)$ for all $x \in [a, b]$, it is said to be* **odd** *if it satisfies $f(x) = -f(-x)$ for all $x \in [a, b]$.*

We note that the definition is vacuous when $a \geq 0$ or when $b \leq 0$.

Any function can be written as the sum of an odd function, $f = f_0 + f_1$, where f_0 is an the function and f_1 is even function defined as follows:

$$f_0(x) \ := \ \frac{1}{2}(\ f(x) - f(-x)\ ) \quad \textit{and} \quad f_1(x) \ := \ \frac{1}{2}(\ f(x) + f(-x)\ ).$$

We could rephrase our question: *which functions on the interval $[-\pi, \pi]$ admit trigonometric representations of the form* :

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} [\, a_n \cos(nx) + b_n \sin(nx)\,] \quad ? \tag{2.12}$$

> The restriction to functions which are 2π periodic on the interval $[-\pi, \pi]$ is just a mathematical convenience. The results which are developed hold equivalently for functions of period P by making suitable adjustments (by rescaling shifting of co-ordinate axes) for the problem considered. For an arbitrary function f on a bounded interval $[a, b]$, one may pass to its *periodic extension*[2] - any mathematical conclusions for $\tilde{f}$ are valid for the restriction of $\tilde{f}$ to f defined on $[a, b]$.

Hence, we will proceed to tackle the problem: *which functions on the interval $[-l, l]$ admit trigonometric representations of the form* :

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} [\, a_n \cos(\frac{n\pi x}{l}) + b_n \sin(\frac{n\pi x}{l})\,] \quad ? \tag{2.13}$$

Equivalently, we could develop the theory for functions on in the interval $[0, l]$.

[2]The periodic extension of a function f defined on $[a, b]$ is given by

$$\tilde{f} \ = \begin{cases} f(x) & \text{for } x \in [a,b], \\ f(x - k(b-a)) & \text{for } x > b,\ x = y + k(b-a) \text{ for some } y \in [a,b] \text{ and } k \in \mathbb{N}^+ \\ f(x + k(b-a)) & \text{for } x < a,\ x = y - k(b-a) \text{ for some } y \in [a,b] \text{ and } k \in \mathbb{N}^+ \end{cases}$$

Clearly $\tilde{f}$ is $(b-a)$-periodic and defined on the entire $\mathbb{R}$

Orthogonality of sines and cosines and calculating coefficients for Fourier series

For $m, n \in \mathbb{N}$, we have

$$\int_{-l}^{l} \sin(\frac{n\pi x}{l}) \cos(\frac{m\pi x}{l}) dx = \int_{0}^{l} \sin(\frac{n\pi x}{l}) \cos(\frac{m\pi x}{l}) dx = 0,$$

and

$$\frac{1}{l}\int_{-l}^{l} \sin(\frac{m\pi x}{l}) \sin(\frac{n\pi x}{l}) dx = \frac{2}{l}\int_{0}^{l} \sin(\frac{m\pi x}{l}) \sin(\frac{n\pi x}{l}) dx = \begin{cases} 0 & if \;\; m \neq n, \\ 1 & if \;\; m = n \neq 0, \end{cases}$$

and finally

$$\frac{1}{l}\int_{-l}^{l} \cos(\frac{m\pi x}{l}) \cos(\frac{n\pi x}{l}) dx = \frac{2}{l}\int_{0}^{l} \cos(\frac{m\pi x}{l}) \cos(\frac{n\pi x}{l}) dx = \begin{cases} 0 & if \;\; m \neq n, \\ 1 & if \;\; m = n \neq 0, \\ 2 & if \;\; m = n = 0. \end{cases}$$

Now suppose

$$f(x) = \sum_{n=1}^{\infty} b_n \sin(\frac{n\pi x}{l}). \tag{2.14}$$

Then multiplying the LHS and RHS of (2.14) by $\frac{1}{l}\sin(\frac{k\pi x}{l})$ and integrating over $[-l, l]$, we have

$$\begin{aligned} \frac{1}{l}\int_{-l}^{l} f(x) \sin(\frac{k\pi x}{l}) dx &= \frac{1}{l}\int_{-l}^{l} \sum_{n=1}^{\infty} b_n \sin(\frac{n\pi x}{l}) \sin(\frac{k\pi x}{l}) dx \\ &= \sum_{n=1}^{\infty} \frac{b_n}{l}\int_{-l}^{l} \sin(\frac{n\pi x}{l}) \sin(\frac{k\pi x}{l}) dx \\ &= c_k \end{aligned}$$

Exercises 2.2.3. *Derive the formula for coefficients for a function defined on* $[0, l]$.

Definition 2.2.4. *If f is $2l$-periodic and Riemann integrable on $[-l, l]$, then the coefficients a_k and b_k are given by*

$$a_k := \frac{1}{l}\int_{-l}^{l} f(x)\cos(\frac{k\pi x}{l})dx$$

$$b_k := \frac{1}{l}\int_{-l}^{l} f(x)\sin(\frac{k\pi x}{l})dx$$

are called the **Fourier coefficients** *of f and the corresponding series,*

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} [\, a_n \cos(\frac{n\pi x}{l}) + b_n \sin(\frac{n\pi x}{l})\,]$$

is called the **Fourier series** *of f.*

NB : We have NOT shown or assumed that we have equality (2.13). We have merely given a name to the series expansion appearing on the RHS of (2.13).

Exercises 2.2.5.

1. *Consider Definition 2.2.4. Give the corresponding definition if f is l-periodic and Riemann integrable on $[0, l]$.*
2. *If f is periodic with period P then show that the following integral is independent of a:*

$$\int_{a}^{a+P} f(x)dx$$

3. *Show that*

$$\int_{-a}^{a} f(x)dx = \begin{cases} 0 & if \quad f \quad is \quad odd \\ 2\int_0^a f(x)dx & if \quad f \quad is \quad even \end{cases}$$

4. *Let f be the 2π periodic function given by $f(x) = |x|$ for $-\pi \le x \le \pi$. Sketch the function and find its Fourier series.*

Complex Fourier series

We recall the properties of the complex exponential function and how it is related to the cosine and sine functions of a variable $x \in \mathbb{R}$:

$$\begin{aligned} e^{ix} &= \cos x + i \sin x \\ \cos x &= \frac{1}{2}(e^{ix} + e^{-ix}) \\ \sin x &= \frac{1}{2i}(e^{ix} - e^{-ix}) \end{aligned}$$

The advantages of working with the sine and cosine functions are that they are odd and even functions, respectively. The advantage of the working with the exponential function is that we have

$$\begin{aligned} (e^{ix})' &= \frac{d}{dx}(e^{ix}) = ie^{ix}, \quad and \\ e^{i(x_1+x_2)} &= e^{ix_1} + e^{ix_2}. \end{aligned}$$

Using the complex exponential function we can rewrite equation (2.13) by :

$$f(x) = \sum_{-\infty}^{\infty} c_n e^{inx/l} \tag{2.15}$$

where

$$\begin{aligned} c_0 &= \frac{a_0}{2}, \\ c_n &= \frac{1}{2}(a_n - ib_n), \quad and \\ c_{-n} &= \frac{1}{2}(a_n + ib_n) \end{aligned}$$

for $n \in \mathbb{N}$. Equivalently we may start with equation (2.15) and derive equation (2.13). The odd and even properties of sine and cosine are used when n is negative in equation (2.15)), and the coefficients are given by

$$\begin{aligned} a_0 &= 2c_0, \\ a_n &= c_n - c_{-n}, \quad and \\ b_n &= i(c_n - c_{-n}), \end{aligned}$$

for $n \in \mathbb{N}$.

Assuming that a function f defined on $[-l, l]$ is given by the series representation of (2.15) (or equivalently (2.13), then we may find any coefficients c_m (or, equivalently, a_m and b_m) by multiplying the LHS and RHS of (2.15) by $e^{-imx/l}$, integrating from $-l$ to l, and then solving for c_m.

Exercises 2.2.6. *Show that*

$$\frac{1}{2l}\int_{-l}^{l} e^{i(m-n)\pi x/l} dx = \begin{cases} 0 & if \quad m \neq n \\ 1 & if \quad m = n \end{cases} \tag{2.16}$$

Definition 2.2.7 (Fourier series and coefficients II). *If f is $2l$-periodic and Riemann integrable on $[-l, l]$, then the coefficients c_n given by*

$$c_n = \frac{1}{2l}\int_{-l}^{l} f(x)e^{-in\pi x/l}dx. \tag{2.17}$$

are called the **Fourier coefficients** *of f and the corresponding series,*

$$\sum_{-\infty}^{\infty} c_n e^{in\pi x/l},$$

is called the **Fourier series** *of f.*

Proposition 2.2.8. *If f is integrable on $[-l, l]$, then its complex Fourier coefficients satisfy*

$$\frac{1}{2l}\int_{-l}^{l}\sum_{-N}^{N} c_k\overline{f(x)}e^{ikx}dx = \sum_{-N}^{N}|c_k|^2 = \frac{1}{2l}\int_{-l}^{l}\sum_{-N}^{N}\overline{c_k}f(x)e^{ikx}dx.$$

PROOF
We have

$$\begin{aligned}
\frac{1}{2l}\int_{-l}^{l}\sum_{-N}^{N} c_k\overline{f(x)}e^{ikx}dx &= \sum_{-N}^{N} c_k\frac{1}{2l}\int_{-l}^{l}\overline{f(x)}e^{ikx}dx \\
&= \sum_{-N}^{N} c_k\frac{1}{2l}\int_{-l}^{l}\overline{f(x)e^{-ikx}}dx \\
&= \sum_{-N}^{N} c_k\overline{\frac{1}{2l}\int_{-l}^{l} f(x)e^{-ikx}dx} \\
&= \sum_{-N}^{N} c_k\overline{c_k}.
\end{aligned}$$

The second equality follows similarly.

Theorem 2.2.9 (Bessel's inequality). *If f is a $2l-$ periodic function and is Riemann integrable on $[-l,l]$, then the Fourier coefficients c_n satisfy*

$$\sum_{-\infty}^{\infty} |c_n|^2 \;\le\; \frac{1}{2l}\int_{-l}^{l} |f(x)|^2 dx$$

PROOF
We give a proof for the case when $l = \pi$. The general case follows from the obvious substitution. Recall that $|z|^2 = z\bar{z}$ for any complex number $z \in \mathbb{C}$. Thus,

$$\begin{aligned} & \left| f(x) - \sum_{-N}^{N} c_n e^{inx} \right|^2 \\ = \;& \left(f(x) - \sum_{-N}^{N} c_n e^{inx}\right) \left(\overline{f(x)} - \sum_{-N}^{N} \overline{c_n} e^{-inx}\right) \\ = \;& |f(x)|^2 \;-\; \sum_{-N}^{N} \left[c_n \overline{f(x)} e^{inx} \;+\; \overline{c_n} f(x) e^{-inx}\right] \;+\; \sum_{m,n=-N}^{N} c_m \overline{c_n} e^{-i(m-n)x} \end{aligned}$$

Dividing the LHS and RHS by 2π, integrating from $-\pi$ to π and applying the definitions of the Fourier coefficients c_n and the orthogonality of $\{e^{inx}\}_{n\in\mathbb{N}}$ we get :

$$\begin{aligned} 0 \;\le\;& \frac{1}{2\pi}\int_{-\pi}^{\pi} \left| f(x) - \sum_{-N}^{N} c_n e^{inx} \right|^2 dx \\ =\;& \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2 dx \;-\; \sum_{-N}^{N} \left[c_n\overline{c_n} + \overline{c_n} c_n\right] \;+\; \sum_{-N}^{N} c_n \overline{c_n} \\ =\;& \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2 dx \;-\; \sum_{-N}^{N} |c_n|^2. \end{aligned}$$

Letting $N \to \infty$, the result follows.

$\diamond$

Corollary 2.2.10. *The Fourier coefficients a_n, b_n and c_n all tend to zero as $n \to \infty$ (and as $n \to -\infty$ in the case of c_n).*

Definition 2.2.11. *Suppose $-\infty < a < b < \infty$. A function f defined on the closed interval $[a, b]$ is said to be* **piecewise continuous** *if*

(i) f *is continuous on $[a, b]$ except at finitely many point $x_1, \ldots, x_n$*

(ii) *at each of the points $x_1, \ldots, x_n$, the left and right limits of f exist, i.e.*

$f(x_{i-}) := \lim\limits_{x \to x_i^-} f(x)$, *and*

$f(x_{i+}) := \lim\limits_{x \to x_i^+} f(x)$ *exist at each of the points $x_1, \ldots, x_n$,*

Note that if $x_i = a$ (or $x_i = b$) for some $i \in 1, \ldots n$, then only the right (respectively, left) limit is expected to exist.

Definition 2.2.12. *A function f defined on the closed interval $[a, b]$ is said to be* **piecewise smooth** *if f and its first derivative f' are each piecewise continuous on $[a, b]$. $f : \mathbb{R} \to \mathbb{R}$ is said to be* **piecewise continuous** *or* **piecewise smooth** *if f satisfies this property on very bounded interval $[a, b]$.*

The set of points where f' is discontinuous will include the set of points where f is discontinuous, and may include other points as well.

On pointwise and uniform convergence

The distinction between pointwise and uniform convergence is especially important when one considers, for example, a sequence of functions f_n and the sequence

$$\int_a^b f_n(x)dx.$$

If $f_n \overset{pointwise}{\longrightarrow} f$ then it is **not** necessarily the case that $\int_a^b f_n(x)dx \longrightarrow \int_a^b f(x)dx$. However, we do have

$$f_n \overset{uniformly}{\longrightarrow} f \quad \Rightarrow \quad \int_a^b f_n(x)dx \longrightarrow \int_a^b f(x)dx.$$

Notation: As usual, L_p, $p \geq 1$ denote the classes of functions whose p^{th} powers are Lebesgue integrable.

Definition 2.2.13 (Pointwise, uniform and root mean square / L_2 convergence). *Consider $f_n : \Omega \to \mathbb{R}$ and $f_n : \Omega \to \mathbb{R}$ where $\Omega \subset \mathbb{R}$.*

- *f_n converges pointwise to f on Ω if for any $\epsilon > 0$ and $x \in \Omega$ there exists N such that for $n > N$ we have $|f_n(x) - f(x)| < \epsilon$.*
- *f_n converges uniformly to f on Ω if for any $\epsilon > 0$ there exists N such that for $n > N$ we have $|f_n(x) - f(x)| < \epsilon$ for all $x \in \Omega$.*
- *f_n converges in mean to f on Ω if for any $\epsilon > 0$ there exists N such that for $n > N$ we have $||f_n(x) - f(x)||_1 < \epsilon$, i.e. $\int_\Omega |f_n(x) - f(x)| d\mu \to 0$.*
- *f_n converges to f in root-mean-squared on Ω if for any $\epsilon > 0$ there exists N such that for $n > N$ we have $||f_n(x) - f(x)||_2 < \epsilon$, i.e. $\int_\Omega |f_n(x) - f(x)|^2 d\mu \to 0$.*

Theorem 2.2.14 (Pointwise convergence of Fourier series).
Suppose f is defined on $[-l, l]$ and

(i) f *is square-integrable on* $[-l, l]$*, i.e.* $\int_{-l}^{l} |f|^2 dx < \infty$,

(ii) f *is piecewise smooth on* $[-l, l]$,

(iii) f *is extended outside* $[-l, l]$ *by periodic extension.*

Then for every $x \in [-l, l]$

$$\lim_{N\to\infty} \frac{a_0}{2} + \sum_{n=1}^{N} [\, a_n \cos(\frac{n\pi x}{l}) + b_n \sin(\frac{n\pi x}{l})\,] \;= \frac{1}{2}([\; f(x_-) + f(x_+)\;],$$

where a_n *and* b_n *are the Fourier coefficients of* f.
In particular, if f *is continuous at* $x \in [-l, l]$*, then*

$$\lim_{N\to\infty} \frac{a_0}{2} + \sum_{n=1}^{N} [\, a_n \cos(\frac{n\pi x}{l}) + b_n \sin(\frac{n\pi x}{l})\,] \;= f(x).$$

Proof
See Appendix

Corollary 2.2.15 (Pointwise convergence of Fourier series - complex form).
Suppose f is defined on $[-l, l]$ and

(i) f *is square-integrable on* $[-l, l]$*, i.e.* $\int_{-l}^{l} |f|^2 dx < \infty$,

(ii) f *is piecewise smooth on* $[-l, l]$,

(iii) f *is extended outside* $[-l, l]$ *by periodic extension.*

Then for every $x \in [-l, l]$

$$\lim_{N\to\infty} \sum_{-N}^{N} c_n e^{in\pi x/l} \;= \frac{1}{2}([\; f(x_-) + f(x_+)\;],$$

where c *are the Fourier coefficients of* f.
In particular, if f *is continuous* $x \in [-l, l]$*, then*

$$\lim_{N\to\infty} \sum_{-N}^{N} c_n e^{in\pi x/l} \;= f(x).$$

Differentiation of Fourier series

Theorem 2.2.16. *Suppose f is a $2l-$periodic function and is piecewise smooth on $\mathbb{R}$, and a_n, b_n and c_n are the Fourier coefficients f. Let a_n', b_n' and c_n' denote the Fourier coefficients of f'. Then*

$$a_n' = \tfrac{n\pi}{l}\, b_n, \qquad b_n' = -\tfrac{n\pi}{l}\, a_n, \qquad c_n' = \tfrac{in\pi}{l}\, c_n.$$

Theorem 2.2.17. *Suppose f is a $2l-$periodic function and that f' satisfy the condition of Theorem 2.2.14 (or equivalently 2.2.15). If*

$$\sum_{-\infty}^{\infty} c_n e^{inx/l} = \tfrac{a_0}{2} + \sum_{n=1}^{\infty} \left[\, a_n \cos(\tfrac{n\pi x}{l}) + b_n \sin(\tfrac{n\pi x}{l}) \,\right]$$

is the Fourier series of $f(x)$, then $f'(x)$ has the derived series

$$\sum_{-\infty}^{\infty} \tfrac{c_n in\pi}{l} e^{inx/l} = \tfrac{a_0}{2} + \sum_{n=1}^{\infty} \left[\, \tfrac{b_n n\pi}{l} \cos(\tfrac{n\pi x}{l}) - \tfrac{a_n n\pi}{l} \sin(\tfrac{n\pi x}{l}) \,\right]$$

for all x such that $f'(x)$ exists.

2.3 Bases for infinite-dimensional vectors spaces

In the rest of this chapter we glimpse into the framework which gives us L_2 convergence of the fourier series of a function.

If X is a $n-$dimensional vector space then we can find basis $\Psi := \{\psi_i\}_{i=1}^n$ for X such that any element in X may be written as a linear combination of elements of Ψ. When X has the additional structure of an inner product then we can find bases with additional properties which are very useful to work with. Recall that the **dot product** or **inner product** of two vectors $u = (u_1, u_2, u_3)$ and $v = (v_1, v_2, v_3)$ in $\mathbb{R}^3$ is a real-valued function from $\mathbb{R}^3 \times \mathbb{R}^3$ into $\mathbb{R}$ defined by

$$u \cdot v \ := \ u_1v_1 + u_2v_2 + u_3v_3$$

or equivalently by

$$u \cdot v \ := \ ||u||\,||v||\cos\theta,$$

where θ is the angle between u and v. Clearly, if u and v are non-zero vectors then $u \cdot v = 0$ if and only if $\theta = 90^\circ$. Hence, we say two vectors are **perpendicular** or **orthogonal** if and only if their inner product is zero. These definitions are generalised to higher dimensions as follows. The inner product of two vectors $u = (u_1, u_2, \ldots, u_n)$ and $v = (v_1, v_2, \ldots, v_n)$ in $\mathbb{R}^n$ was defined by

$$u \cdot v \ := \ u_1v_1 + u_2v_2 + \ \ldots \ u_nv_n,$$

and u and v are said to be orthogonal, denoted $u \perp v$, if and only if $u \cdot v = 0$. The function $\_ \cdot \_ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ satisfies the axioms given in the next definition.

Definition 2.3.1 (inner product axioms). *Let X be a vector space. An inner product on X is a function which maps $X \times X$ into $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, written $(\_ , \_) : X \times X \to \mathbb{K}$, such that for $u, v, w \in X$ and scalars $\alpha, \beta \in \mathbb{K}$:*

[c1] $(u, v) = \overline{(v, u)}$
[c2] $(\alpha u + \beta v, w) = \alpha(u, w) + \beta(v, w)$
[c3] $(u, u) \geq 0$; $(u, u) = 0$ *if and only if* $u = 0$

We write $(X, (\_ , \_)\ )$ to indicate that a vector space X is endowed with an inner product structure.

Remarks
(1) $(u, u) \in \mathbb{R}$ for any $u \in (X, (\_ , \_)\ )$ since $(u, u) = \overline{(u, u)}$.
(2) $(\alpha u, v) = \alpha(u, v)$ while $(u, \alpha v) = \overline{\alpha}(u, v)$.

Examples 2.3.2. *It is can be verified that the following examples satisfy the inner product axioms.*

1. *The vector space $\mathbb{C}^n$ (or $\mathbb{R}^n$) with* $(u, v) := u_1\overline{v_1} + u_2\overline{v_2} + \ldots u_n\overline{v_n}$.
2. *The vector spaces l_2 of 2−summable sequences of complex (or real) numbers with* $(u, v) := u_1\overline{v_1} + u_2\overline{v_2} + \ldots . = \sum_{i=1}^{\infty} u_i\overline{v_i}$ [3].
3. *The vector spaces $L_2(\Omega)$ of square-integrable functions with* $(f, g) := \int_\Omega f(x)\overline{g(x)}\, dx$.
4. *The vector space $C([a, b])$ with* $(f, g) := \int_\Omega f(x)\overline{g(x)}\, dx$.

Definition 2.3.3 (orthogonality). *Suppose $v, w \in (X, (\_ , \_)\ )$. If $(v, w) = 0$, then v and w are said to be* **orthogonal**, *and we write $v \perp w$. A vector v is said to be orthogonal to a set W is $v \perp w$ for every $w \in W$ and the sets V and W are said to be orthogonal if $v \perp w$ for every $v \in V$ and $w \in W$.*

If X is an inner product space and W is a subspace of X as a vector space, then W inherits the inner product structure defined on X and is an inner product space in itself.

Definition 2.3.4 (norm axioms). *Let X be a vector space. A norm on X is a function which maps $X \times X$ into $\mathbb{R}$ written $||\_|| : X \to \mathbb{R}$, such that for $u, v \in X$ and scalars $\alpha \in \mathbb{R}$ we have:*

[n1] $||u|| \geq 0$; $||u|| = 0$ *if and only if* $u = 0$
[n2] $||\alpha u|| = |\alpha|\, ||u||$
[n3] $||u + v|| \leq ||u|| + ||v||$

We write $(X, ||\_||\ )$ to indicate that the vector space X is endowed with a norm function.

Clearly the norm function is a generalisation of the modulus function for $\mathbb{R}$ and $\mathbb{C}$ [4]. The norm of the difference between two vectors is naturally interpreted as the distance between the two elements in the vector space[5].

[3] The property that the sequences in l_2 are 2−summable is enough to ensure the convergence of the summation series given for (u, v).

[4] It is well-known that the modulus functions $\mathbb{R}$ and $\mathbb{C}$ on satisfies the conditions [n1], [n2] and [n3].

[5] More formally, a norm function on a vector space induces another function d from $X \times X$ into $\mathbb{R}$ given by $d(x, y) := ||x - y||$. The function d satisfies the axioms of a *metric* . The converse is not generally true, i.e. not every metric defined on a vector space is induced by a norm function

Examples 2.3.5. *It is can be verified that the following examples satisfy the norm axioms. Minkowski's inequality is needed to verify [n3] for (3), (4) and (5).*

1. *The vector space* $\mathbb{R}^n$ *(or* $\mathbb{C}^n$*) with* $||u||_2 := \sqrt{u_1^2 + u_2^2 + \dots u_n^2}$.
2. *The vector space* $\mathbb{R}^n$ *(or* $\mathbb{C}^n$*) with* $||u||_\infty := \max\{\ |u_i|\ \}$.
3. *The vector spaces* l_p, $1 \leq p < \infty$, *of* $p-$*summable sequences with* $||x||_p := (\sum\limits_{i=1}^{\infty} |x_i|^p\ )^{\frac{1}{p}}$.
4. *The vector spaces* $L_p(\Omega)$, $1 \leq p < \infty$, *of* $p-$*integrable functions with* $||f||_p := (\int\limits_\Omega |f(x)|^p\ )^{\frac{1}{p}}$.
5. *The vector space* $C([a,b])$ *with* $||f||_2 := (\int\limits_\Omega |f(x)|^2\ )^{\frac{1}{2}}$.
6. *The vector space* $C([a,b])$ *with* $||f||_\infty := \sup\limits_{x\in K} |f(x)|$.

Proposition 2.3.6 (Norm generated by an inner product). *Suppose* $(X,\ (\_,\_))$ *is an inner product space. Then the function* $||\_||$ *defined by* $||u|| := (u,u)^{\frac{1}{2}}$ *is a norm function.*

Examples 2.3.7.

1. *The norm for* L_2 *is precisely the norm induced by the inner product for* L_2 *given in Examples 2.3.2, i.e.*

$$||f||_2 \ = \ \int\limits_\Omega |f(x)|^2\,dx \ = \ \sqrt{(f,f)} \ = \ (\ \int\limits_\Omega f(x)\overline{f(x)}\,dx\ )^{\frac{1}{2}}$$

2. *As in (1) the norm for the sequence space* l_2 *is precisely the norm induced by the inner product given for* l_2 *(see Examples 2.3.2).*

We have shown that any inner product space is a normed space. The converse is not always true. To determine whether the norm function on a vector space is in fact given by some inner product structure we have the following rule which must be satisfied by any norm induced by an inner product.

Proposition 2.3.8 (Parallelogram law). *If* $(X,\ (\_,\_))$ *is an inner product space, then*

$$||u+v||^2 + ||u-v||^2 = 2(\ ||u||^2 + ||v||^2\ ) \tag{2.18}$$

Definition 2.3.9. *A set of vectors* $\{\psi_i\}_{i\in\mathbb{N}}$ *in an inner product space* $(X,(\_,\_),)$ *is said to be orthogonal if the elements are mutually orthogonal, i.e. if* $(\psi_i,\psi_j) = 0$ *for all* $i \neq j$*. The set is said to be* **orthonormal** *if it is orthogonal and* $||\psi_i|| = \sqrt{(\psi,\psi)} = 1$ *for each* $i \in \mathbb{N}$.

Suppose $\Psi = \{\psi_i\}_{i=1}^n$ is an orthonormal set in an $n-$dimensional inner product space $(X,(\_,\_),)$. Then Ψ is a basis[6] for X. To see this, suppose

$$a_1\psi_1 + a_2\psi_2 + \dots + a_n\psi_n = 0 \tag{2.19}$$

[6]see Glossary of Linear Algebra terms

Then taking the inner product of both sides of (2.19) with ψ_1 yields $a_1 = 0$. Proceeding in the same way for each $\psi_i,\ i = 1$ to n, proves that $a_i = 0$ for each $i = 1$ to n. Thus, Ψ is a linearly independent set. Since Ψ is a basis for X, any $x \in X$ may be written

$$x = x_1\psi_1 + x_2\psi_2 + \ldots + x_n\psi_n, \tag{2.20}$$

where $x_i,\ i = 1$ to n are constants. As before, we take the inner product of both sides of (2.20) with ψ_1. This time we find that $x_1 = (x, \psi_1)$. Proceeding in the same way for each $\psi_i,\ i = 1$ to n, yields $x_i = (x, \psi_i)$. Thus, we may write

$$x = \sum_{i=1}^{n} (x, \psi_i)\psi_i.$$

Another computation is the following. Suppose $x, y \in X$. Then

$$(x, y) \ = \ \left(\sum_{i=1}^{n} x_i\psi_i, \sum_{j=1}^{n} y_j\psi_j \right) \ = \ \sum_{i=1}^{n}\sum_{j=1}^{n} x_i\,\overline{y_j}\,(\psi_i, \psi_j) = \sum_{i=1}^{n} x_i\,\overline{y_i}$$

It is always possible to construct an *orthonormal* basis from an arbitrary basis for an inner product space[7].

[7]The **Gram-Schmidt orthonormalisation** procedure is an algorithm for constructing an orthonormal basis $\Psi = \{\psi_i\}_{i=1}^n$ from any given basis $U = \{u_i\}_{i=1}^n$ of $n-$ dimensional inner product space.

To illustrate, let $U = \{u_i\}_{i=1}^3$ be a basis for $\mathbb{R}^3$. To derive Ψ, let

$$\psi_1 = \frac{u_1}{||u_1||}.$$

Next, project u_2 onto the subspace which is orthogonal to ψ_1 via the projection P_2 defined by $P_2x = x - (x, \psi_1)\psi_1$. To get ψ_2, normalise the resulting vector, i.e. let

$$\psi_2 = \frac{Pu_2}{||Pu_2||}.$$

Lastly, project u_3 onto the subspace which is orthogonal to the subspace spanned by ψ_1 and ψ_2 via the projection P_3 defined by $P_3x = x - (\ (x, \psi_1)\psi_1 + (x, \psi_2)\psi_2\ )$. To get ψ_3, normalise the resulting vector , i.e. let

$$\psi_3 = \frac{Pu_3}{||Pu_3||}.$$

In general, given a basis $U = \{u_i\}_{i=1}^n$, derive an orthonormal basis $\Psi = \{\psi_i\}_{i=1}^n$ by letting

$$\psi_i = \frac{P_iu_i}{||P_iu_i||},$$

where the projections P_i, are defined by

$$\begin{aligned} P_1u &= u \\ P_iu &= u - \sum_{k=1}^{i-1} (u, \psi_i)\psi_i, \quad i = 2, 3 \ldots, n. \end{aligned}$$

In an $n-$dimensional vector space, we can represent an element x as a linear combination of vectors in a basis $\{\psi_i\}_{i=1}^n$, i.e. x can be written as finite sum as in (2.20). However, we have encountered function spaces for which it is not possible to find a finite set $\{\psi_i\}_{i=1}^n$ which spans the whole set. In order to extend the notion of a basis to function spaces, such as $L_2(\Omega)$, we have to trade in *equality* (to a finite linear combination) for *approximation* by a finite linear combination of elements in a *countable infinite* set. This is the goal which underlies the content of the rest of this section.

Definition 2.3.10. *Let X be an inner product space and let $\Psi = \{\psi_i\}_{i=1}^{\infty}$ be an orthonormal set. Then Ψ is said to be* **maximal** *if there is no nonzero element in X which is orthogonal to Ψ, i.e. Ψ is maximal if and only if it is impossible to find another vector in H which is orthogonal to all $\{\psi_i\}_{i=1}^{\infty}$.*

A maximal orthonormal set in a Hilbert space H is called an **orthonormal basis** *for H*[8]*.*

Theorem 2.3.11 (The Fourier Series Theorem).

Let H be a Hilbert space and let $\Psi = \{\psi_i\}_{i=1}^{\infty}$ be an orthonormal set. Then every $x \in H$ can be represented as

$$x = \sum_{i=1}^{\infty} (x, \psi_i)\psi_i$$

if and only if Ψ is a maximal orthonormal set[9]*.*

2.4 Sturm-Liouville problems

Definition 2.4.1. *A* **Sturm-Liouville** *operator L is a linear operator defined which satisfies :*

$$Lu = \frac{1}{w}(-(pu')' + qu)$$

on an interval $[a, b]$, where u is a twice differentiable function, $u' = \frac{du}{dx}$ and p, p', q and w are continuous $\mathbb{R}-$valued functions which satisfy

$$\begin{aligned} p(x) &> 0, \\ q(x) &\geq 0, \quad \text{and} \\ w(x) &\geq 0 \quad \text{on} \quad [a, b]. \end{aligned}$$

Let B_1 and B_2 denote boundary operators, i.e. linear operators which specify the boundary values of a continuous function, defined by

---

We first note that the Gram-Schmidt procedure may be generalised to function spaces. Suppose $U = \{u_i\}_{i=1}^{\infty}$ is a linearly independent set in an function space with an inner product structure. Derive an orthonormal set $\Psi = \{\psi_i\}_{i=1}^{\infty}$ by letting

$$\psi_i = \frac{P_i u_i}{||P_i u_i||},$$

where the projections P_i, are defined inductively by

$$\begin{aligned} P_1 u &= u \\ P_i u &= u - \sum_{k=1}^{i-1} (u, \psi_i)\psi_i, \quad i = 2, 3 \ldots. \end{aligned}$$

[8]The reason for this definition becomes clear after the Fourier series theorem.

[9]It is precisely this equivalence which motivates the definition that such a set Ψ is called an orthonormal *basis*, as defined earlier.

$$
\begin{aligned}
B_1 u &= \alpha_1 u(a) - \beta_1 u'(a) \quad \textit{and} \\
B_2 u &= \alpha_2 u(a) - \beta_2 u'(b),
\end{aligned}
$$

such that the constants α_i *and* β_i *satisfy*

$$
\alpha_i \geq 0, \quad \beta_i \geq 0 \quad \textit{and} \quad \alpha_i + \beta_i > 0.
$$

A **regular Sturm-Liouville problem** *is an eigenvalue problem of the form*

$$
\begin{aligned}
Lu(x) &= \lambda u(x) \quad x \in (a,b) \\
B_1 u = 0 \quad \textit{and} \quad & B_2 u = 0.
\end{aligned} \tag{2.21}
$$

Remarks The equation $Lu = \lambda u$ can be rewritten $-(pu')' + qu = \lambda w u$, and it is usually encountered in this form.

The conditions (bounded domain for u, constraints on p, q, w, α_i and β_i) for the above definition may not all hold for a given problem, in which case the problem is referred to as a *singular Sturm-Liouville problem.*

Problem (2.21) is considered on the space $L_2([a,b])$ endowed with the inner product

$$
(f,g) := \int_a^b f(x)\overline{g(x)}w(x)\,dx.
$$

The function w is referred to as a *weighting function* and L_2 with this inner product is called a *weighted inner product space.*

Clearly L is not defined for all functions in $L_2([a,b])$. The domain of L is a proper subspace:

$$
D(L) := \{\ f \in L_2([a,b]) \cap C^2([a,b]) \mid B_1 f = B_2 f = 0\ \}.
$$

It can be shown that $D(L)$ is dense in $L_2([a,b])$.

Examples 2.4.2.
The EVP encountered in the problem

$$
\begin{aligned}
u_t &= k u_{xx} \quad x \in (0,l) \\
u(0,t) = u(l,t) &= 0 \quad t > 0 \\
u(x,0) &= f(x) \quad t = 0, \quad x \in (0,l)
\end{aligned}
$$

is a Sturm-Liouville problem (verify).

Theorem 2.4.3. *The set of eigenfunctions* $\{\psi_i\}_{i=1}^{\infty}$ *of a Sturm-Liouville are a maximal orthonormal set for* $L_2([a,b])$*.*

Examples 2.4.4.

Remark: Bessels functions constitute another type of maximal orthonormal family; there are several others.

2.5 Applications to Finance

Eigenfunction expansions have been applied in several ways to problems related to derivative securities, for example :

- *Applications of Eigenfunction Expansions in Continuous-Time Finance* by Alan L. Lewis, Mathematical Finance **8**, 349 - 383 (1998)
- *Exotic Spectra* by V. Linetsky, Risk, April 2002, and *Shadow Interest* by V. Gorovoi and V. Linetsky, Risk, December 2003; several more papers are downloadable from the website http://users.iems.nwu.edu/ linetsky/
- *Black-Scholes goes hypergeometric*, by C. Albanese, G. Compolieti, P. Carr and A. Lipton, Risk, December 2001

Orthonormal **wavelet bases** are special types of orthonormal bases which have been used extensively in *signal processing and the analysis of financial market data.* The mathematical theory of wavelets is an extension of Fourier series analysis.

On the theoretic side of things, Fourier series and wavelet bases are employed in constructions of Brownian motion (hence, they are used to show that there does in fact exist mathematical objects which satisfy the axioms which define Brownian motion).

Chapter 3

Integral transform methods

3.1 A heuristic motivation for fourier integral transform

Suppose f has a fourier series representation on $[-l, l]$, i.e.

$$\begin{aligned} f(x) &= \frac{1}{2l} \sum_{-\infty}^{\infty} C_{n,l}\ e^{in\pi x/l} \\ \text{where} \quad C_{n,l} &= \int_{-l}^{l} f(y)\ e^{-in\pi y/l}\ dy. \end{aligned}$$

Here we are using '$C_{n,l}$' instead of just 'C_n' in order to include information of the interval under consideration.

Now let $\Delta\xi = \frac{\pi}{l}$, $\xi_n = n\,\Delta\,\xi = \frac{n\pi}{l}$. Then since $\frac{1}{2l} = \frac{1}{2\pi}\ \Delta\xi$, for $x \in [-l, l]$ we have:

$$\begin{aligned} f(x) &= \frac{1}{2\pi} \sum_{-\infty}^{\infty} C_{n,l}\ e^{i\,\xi_n\,x}\ \Delta\xi \\ C_{n,l} &= \int_{-l}^{l} f(y)\ e^{-i\,\xi_n\,y}\ dy. \end{aligned}$$

If $f(x)$ vanishes rapidly as $x \to \pm\infty$, then we have

$$C_{n,l} \approx \int_{-\infty}^{\infty} f(y)\ e^{-i\,\xi_n\,y}\ dy,$$

which we denote by $\tilde{f}(\xi_n)$. Letting $l \longrightarrow \infty$, it follows $\Delta\xi \longrightarrow 0$ and

$$\begin{aligned} f(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{f}(\xi) e^{i\,\xi\,x}\ d\xi \\ \text{where} \quad \tilde{f}(\xi) &= \int_{-\infty}^{\infty} f(y)\ e^{-i\xi y} dy. \end{aligned}$$

Substituting the expression $\tilde{f}(\xi)$ into the integral representation for $f(x)$ we have

$$f(x) \;=\; \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{\infty}^{\infty} f(y)\; e^{-i\xi y}\; dy\; e^{i\xi x}\; d\xi \;=\; \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{\infty}^{\infty} f(y)\; e^{i\xi(x-y)}\; dy\; d\xi.$$

This equality is referred to as the **Fourier integral theorem** or alternatively the **Fourier inversion formula**. From it we may define several integral pairs which, when put together, amount to the same thing. We adopt the following convention[1]:

the **Fourier integral transform** of a function $f(x)$ is defined by :

$$\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)\, e^{-i\xi x}\, dx,$$

and the **Fourier inverse transform** of $\hat{f}(\xi)$ is defined by

$$f(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(\xi)\, e^{i\xi y}\, d\xi.$$

NB: As with Fourier series, it is *not* always the case that

1. the integral transform $\hat{f}(\xi)$ exists,
2. the integral representation is equal to the function $f(x)$.

The proof of pointwise convergence of the integral representation is similar to the proof of pointwise convergence of Fourier series. When f is a rapidly decreasing function (defined in the next chapter) the proof of equality is straight forward (see Yosida). If $f \in L_2(\mathbb{R})$ the proof is more sophisticated and is known as the Plancheral theorem.

[1] An alternative pair could be:

$$\begin{aligned} \hat{f}(\xi) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x)\, e^{i\xi x}\, dx, \quad \text{and} \\ f(y) &= \int_{-\infty}^{\infty} \hat{f}(\xi)\, e^{-i\xi y}\, dy. \end{aligned}$$

Here the normalization factors in front of the integrals have changed and the minus sign in the exponent has been shifted from the transform integral to the inverse transform integral.

3.1.1 Some basic properties

We use the notation $F[f(x)] := \hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} f(x)\, e^{-i\xi x}\, dx$.

The integral $\hat{f}(\xi)$ exists if $f \in L_1(\mathbb{R})$ and

1. $F[\alpha f + \beta g] = \alpha F[f] + \beta F[g]$ (linearity)
2. $F[f(x-a)] = e^{-ia\xi} F[f]$ (shift)
3. $F[f'] = i\xi F[f]$ provided $f' \in L_1$ (the derivative formula)

PROOF of (3)
Suppose $f, f' \in L_1$. Then $\lim\limits_{x \to \pm\infty} f(x) = 0$. Thus,

$$\begin{aligned}
F[f'](\xi) &= \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{-i\xi x}\ f'(x)\ dx \\
&= \frac{1}{\sqrt{2\pi}} \lim_{b\to\infty} \int\limits_{-b}^{b} e^{-i\xi x}\ f'(x)\ dx \\
\text{(integrating by parts)} \quad &= \frac{1}{\sqrt{2\pi}} \lim_{b\to\infty} \left([e^{-i\xi x}\ f(x)]_{-b}^{+b} - \int\limits_{-b}^{b} f(x)\ e^{-i\xi x}\ .(-i\xi)\ dx \right) \\
\left(\text{since } \lim_{b\to\pm\infty} e^{-i\xi b}\ f(b) = 0\right) \quad &= \frac{1}{\sqrt{2\pi}} \lim_{b\to\infty} i\xi \int\limits_{-b}^{b} f(x)\ e^{-i\xi x}\ dx \\
&= i\ \xi\ \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} f(x)\ e^{-i\xi x}\ dx \\
&= i\ \xi\ F[f].
\end{aligned}$$

Examples 3.1.1.

The Fourier transform of e^{-x^2}: $\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{-x^2} \, e^{-i\xi x} \, dx$. Now $x^2 + i\,\xi\,x = (x + \frac{i\xi}{2})^2 + \frac{\xi^2}{4}$. Letting $\kappa = x + \frac{i\xi}{2}$, $d\kappa = dx$, and

$$\hat{f}(\xi) \;=\; \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{-k^2 - \xi^2/4} \; d\kappa \;=\; \frac{1}{\sqrt{2}} \, e^{-\xi^2/4},$$

since[2] $\int\limits_{-\infty}^{\infty} e^{-\kappa^2} \; d\kappa = \sqrt{\pi}$.

3.1.2 Convolution

The convolution $f * g$ of two functions f and g on $\mathbb{R}$ is defined by:

$$(f * g)(x) = \int\limits_{-\infty}^{\infty} f(y)\, g(x-y)\, dy,$$

provided the integral exists.

Various conditions can be given to ensure that the the integral converges:

1. If $f \in L_1(\mathbb{R})$, i.e. if $\int\limits_{-\infty}^{0} |f(x)| \; dx \; < \; \infty$, and if g is uniformly bounded, i.e. if $\exists M \;\, s.t. |g(x)| \; < M \;\forall \;\; x \in \mathbb{R}$, then

$$\int\limits_{\infty}^{\infty} |f(y) g(x-y)| \; dy \;\; \leq M \int\limits_{-\infty}^{\infty} |f(y)| \; dy \;\; < \;\; \infty.$$

2. Similarly, if $g \in L_1(\mathbb{R})$ and if f is uniformly bounded, then

$$\int\limits_{-\infty}^{\infty} |f(y) g(x-y)| \; dy < \infty.$$

---

[2]

$$\begin{aligned} I = \int\limits_{-\infty}^{\infty} e^{-\kappa^2} \; d\kappa \;\; \Rightarrow \;\; I^2 \;\; &= \;\; \int\limits_{-\infty}^{\infty} e^{-x^2} \; dx \int\limits_{-\infty}^{\infty} e^{-y^2} dy \;=\; \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} e^{-(x^2+y^2)} \; dx\, dy \\ &= \int\limits_{0}^{2\pi} \int\limits_{0}^{\infty} e^{-r^2} \; r \; dr \; d\Theta \;\;=\;\; \frac{1}{2} \int\limits_{0}^{2\pi} \int\limits_{0}^{\infty} e^{-u} \; du \, d\Theta \;\;=\;\; \frac{1}{2} \int\limits_{0}^{2\pi} d\Theta \\ &= \;\; \pi. \end{aligned}$$

Similarly $\int\limits_{0}^{\infty} e^{-\kappa^2} \, d\kappa \; = \frac{\pi}{2}$.

3. If f is piecewise continuous and if g is bounded on the same bounded interval $[a,b]$ with g vanishing outside $[a,b]$, i.e. $g(x) = 0 \forall\, x \notin [a,b]$, then

$$\begin{aligned}\int_{-\infty}^{\infty} |f(y)g(x-y)|\, dy &= \int_a^b |f(y)|\, |g(x-y)|\, dy \\ &\leq M \int_a^b |f(y)|\, dy \, < \, \infty.\end{aligned}$$

4. See exercises.

3.2 The Convolution Theorem

$$\begin{aligned}F[(f*g)(x)] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (f*g)\ (x)\ e^{-i\xi x}\ dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(y)\ g(x-y)\ dy\ e^{-i\xi x}\ dx \\ \text{(multiply by } e^{-i\xi y}.e^{i\xi}) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(y)\, e^{-i\xi y}\ g(x-y)\, e^{-i\xi(x-y)}\ dy\ dx \\ \text{(substitute } z = x-y) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(y)\, e^{-i\xi y}\ g(x) e^{-i\xi z}\ dy\ dz \\ &= \sqrt{2\pi}\ \hat{f}(\xi)\ \hat{g}(\xi)\end{aligned}$$

We have proved that $F\ [f*g]\ =\ \sqrt{2\pi}\ \hat{f}\ \hat{g}$. Similarly $F^{-1}\ [\hat{f}\hat{g}]\ =\frac{1}{\sqrt{2\pi}}\ (f*g)\ (x)$.

Before proceeding to apply Fourier transform to solve a diffusion problem, we note:

1. Fourier transforms are applied to functions of a variable x ranging over all $\mathbb{R}$. Usually this is a spatial variable.
2. The calculation of Fourier transform can be technical and summaries of commonly used transform pairs are given in tables. Two transform pairs which occur frequently in the sequel are"

$$\begin{aligned}F\ [e^{-ax^2}] &= \frac{1}{\sqrt{2a}}\ e^{-\xi^2/4a} \\ F^{-1}\ [e^{-a\xi^2}] &= \frac{1}{\sqrt{2a}}\ e^{-x^2/4a},\end{aligned}$$

where a is constant.

3.3 Applying Fourier transforms to solve the IVP

$$u_t - k\, u_{x,x} = 0 \qquad x \in \mathbb{R}\,,\ t > 0 \tag{3.1}$$

$$u(x,0) = f(x). \tag{3.2}$$

Taking Fourier transforms of u_t and u_{xx} we have

$$\begin{aligned}
F\,[u(x,t)] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u(x,t)\, e^{-i\xi x}\, dx \qquad = \hat{u}(\xi,t) \\
F[\frac{\partial}{\partial t}\, u\,(x,t)] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{\partial}{\partial t}\, u(x,t)\, e^{-i\xi x}\, dx = \widehat{\frac{\partial u}{\partial t}}\,(\xi,t) = \hat{u}_t\,(\xi) \\
F[\frac{\partial}{\partial x}\, u(x,t)] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{\partial}{\partial x}\, u(x,t)\, e^{-i\xi x}\, dx = i\xi\, \widehat{u}(\xi,t) \\
F[\frac{\partial^2}{\partial x^2}\, u(x,t)] &= (i\,\xi)^2\, \hat{u}\,(\xi,t) = -\xi^2\, \hat{u}\,(\xi,t).
\end{aligned}$$

Thus, (3.1) and (3.2) become

$$\hat{u}_t(\xi,t) = -k\,\xi^2\, \hat{u}\,(\xi,t) \tag{3.3}$$

$$\hat{u}\,(0,t) = \widehat{f}(\xi). \tag{3.4}$$

For each $\xi \in \mathbb{R}$, (3.3) is the (simplest) ODE in the t-variable (of the form $y' = Ay$), with general solution:

$$\hat{u}\,(\xi,t) = c(\xi)\, e^{-\xi^2 kt}.$$

Since $\widehat{f}(\xi) = \widehat{u}\,(\xi,0) = c\,(\xi)$, it follows that

$$\begin{aligned}
\hat{u}\,(\xi,t) &= \hat{f}(\xi) e^{-\xi^2 kt} \\
&= \hat{f}(\xi)\, \hat{g}(\xi)
\end{aligned}$$

for some $g(x)$ such that $\hat{g}(\xi) = e^{-k\xi^2 t}$. Now $\hat{g}(\xi) = e^{-k\xi^2 t}$ implies that $g(x) = \frac{1}{\sqrt{2kt}} e^{-x^2/4kt}$. Thus, by the convolution theorem,

$$\begin{aligned}
u(x,t) &= F^{-1}\,[\hat{u}\,(\xi,t)] = F^{-1}\,[\hat{f}\hat{g}] \\
&= \frac{1}{\sqrt{2\pi}}\,(f * g)(x) \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(y)\, \frac{1}{\sqrt{2kt}}\, \exp[-\frac{(x-y)^2}{4kt}]\, dy \\
&= \frac{1}{\sqrt{4\pi kt}} \int_{-\infty}^{\infty} f(y)\, \exp[-\frac{(x-y)^2}{4kt}]\, dy.
\end{aligned}$$

3.4 Higher dimensional transforms

Suppose $f \in L_1(\mathbb{R}^n)$. Then

$$F[f](\xi) = \hat{f}(\xi) = (\frac{1}{\sqrt{2\pi}})^n \int_{\mathbb{R}} f(x)\, e^{-i(\xi\cdot x)}\, dx,$$

where $\xi, x \in \mathbb{R}^n$ and $\xi \cdot x = x_1\xi_1 + x_2\xi_2 + ... + x_n\xi_n$, and

$$F^{-1}[\hat{f}] = f(x) = (\frac{1}{\sqrt{2\pi}})^n \int_{\mathbb{R}^n} \hat{f}(\xi)\, e^{i(\xi\cdot x)}\, d\xi.$$

3.4.1 Properties

1. $\widehat{\frac{\partial}{\partial x_k} f}(\xi) = i\,\xi_k \hat{f}\,(\xi).$
2. $\widehat{\frac{\partial}{\partial {x_k}^2} f}(\xi) = -{\xi_k}^2\, \hat{f}\,(\xi).$
3. $\widehat{\nabla^2 f}(\xi) = -({\xi_1}^2 + {\xi_2}^2 + ... + {\xi_n}^2)\, \hat{f}\,(\xi).$

Suppose $u \in L_1(\mathbb{R}^3) \cap C^2(\mathbb{R}^3)$. Then

$$\begin{aligned}
&\widehat{\nabla^2 u}(\xi) \\
=\ & (\frac{1}{\sqrt{2\pi}})^3 \int\int\int \nabla^2 u(x_1, x_2, x_3)\, e^{-i(\xi_1 x_1 + \xi_2 x_2 + \xi_3 x_3)}\, dx_1\, dx_2\, dx_3 \\
=\ & (\frac{1}{\sqrt{2\pi}})^3 \Big(\int\int\int \frac{\partial^2}{\partial x_1^2}\, u\, e^{-i(\xi\cdot x)}\, dx + \int\int\int \frac{\partial^2}{\partial x_2^2}\, u\, e^{-i(\xi\cdot x)}\, dx \\
& \quad + \int\int\int \frac{\partial^2}{\partial x_3^2}\, u\, e^{-i(\xi\cdot x)}\, dx \Big).
\end{aligned}$$

Now

$$\begin{aligned}
& \int\int\int \frac{\partial^2}{\partial x_1^2}\, u(x_1, x_2, x_3)\, e^{-i\xi_1 x_1}\, e^{-i\xi_2 x_2}\, e^{-i\xi_3 x_3}\, dx_1\, dx_2\, dx_3 \\
=\ & \int\int \Big(\int \frac{\partial^2}{\partial x_1^2}\, u\,(x_1, x_2, x_3)\, e^{-i\xi_1 x_1}\, dx_1 \Big)\, e^{-i\xi_2 x_2}\, e^{-i\xi_3 x_3}\, dx_2\, dx_3 \\
=\ & \int\int \Big(-\xi_1^2 \int u\,(x_1, x_2, x_3)\, e^{-i\xi_1 x_1}\, dx_1 \Big)\, e^{-i\xi_2 x_2}\, e^{-i\xi_3 x_3}\, dx_2\, dx_3 \\
=\ & -\xi_1^2 \int\int\int u\, e^{-i(\xi x)}\, dx
\end{aligned}$$

Similarly, for $k = 2, 3$,

$$\begin{aligned}
& \int\int\int \frac{\partial^2}{\partial x_k^2}\, u\,(x_1, x_2, x_3)\, e^{-i(\xi x)}\, dx \\
=\ & -\xi_k^2 \int\int\int u\,(x_1, x_2, x_3)\, e^{-i(\xi x)}\, dx.
\end{aligned}$$

Hence,

$$\begin{aligned}
\widehat{\nabla^2 u}(\xi) &= -({\xi_1}^2 + {\xi_2}^2 + {\xi_3}^2)\, (\frac{1}{\sqrt{2\pi}})^3 \int\int\int u\,(x)\, e^{-i(\xi\cdot x)}\, dx \\
&= -({\xi_1}^2 + {\xi_2}^2 + {\xi_3}^2)\, \hat{u}\,(\xi)
\end{aligned}$$

3.5 Fourier sine and cosine transforms

Recall that the Fourier integral transform was motivated by considering the Fourier series expansion $\sum_{-\infty}^{\infty} C_n\, e^{in\pi x/l}$ of a function $f(x)$ defined the interval $[-l, l]$ with

$$\sum_{-\infty}^{\infty} C_n\, e^{i\,n\,\xi\,x/l} \quad \overset{l\to\infty}{\longrightarrow} \quad \int_{-\infty}^{\infty} \hat{f}(\xi)\, e^{i\,\xi\,x}\, dx$$

Similarly, we can motivate the sine and cosine transforms on the semi-finite interval $[0, \infty)$:

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi x}{l}\right) \quad \overset{l\to\infty}{\longrightarrow} \quad \sqrt{\frac{2}{\pi}} \int_0^{\infty} \tilde{f}(\xi)\cos(\xi x)\, dx$$

$$\sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi x}{l}\right) \quad \overset{l\to\infty}{\longrightarrow} \quad \sqrt{\frac{2}{\pi}} \int_0^{\infty} \tilde{f}(\xi)\sin(\xi x)\, f(x)\, dx$$

The **Fourier cosine transform** of $f(x)$ is defined

$$F_c\,[f] \;=\; \sqrt{\frac{2}{\pi}} \int_0^{\infty} \cos\,(\xi x)\, f(x)\, dx \;=\; \hat{f}_c(\xi)$$

and the Fourier cosine inversion formula is given by

$$f(x) \;=\; F^{-1}[\hat{f}_c] \;=\; \sqrt{\frac{2}{\pi}} \int_0^{\infty} \hat{f}_c(\xi)\, \cos\,(\xi x)\, d\xi.$$

Similarly, the Fourier sine transform of $f(x)$ is defined

$$F_s\,[f] \;=\; \sqrt{\frac{2}{\pi}} \int_0^{\infty} \sin\,(\xi x)\, f(x)\, dx \;=\; \hat{f}_s(\xi),$$

and the inversion formula is given by

$$f(x) \;=\; F^{-1}[\hat{f}_s] \;=\; \sqrt{\frac{2}{\pi}} \int_0^{\infty} \hat{f}_s(\xi)\, \sin\,(\xi x)\, d\xi.$$

3.6 Comments on Integral transforms

When we considered Sturm-Liouville problems we noted that different boundary condition gave rise to different families of eigenfunctions (and hence different bases). We saw that the diffusion problem

$$\begin{aligned} u_t - k u_{xx} &= 0 \\ u(x,0) &= f(x) \\ u(0,t) = u(l,t) &= 0, \end{aligned}$$

was coupled to a Sturm-Louiville problem in the x-variable[3] and gave rise to the orthonormal basis $\{\sin(\frac{n\pi x}{l})\}_{n\in\mathbb{N}}$ for $L_2[0,l]$.

Similarly, different boundary conditions for unbounded or semi-bounded variables make it meaningful to consider different integral transforms. When the spatial variables is unbounded (or semi-bounded, e.g. $x \in [0,\infty)$), then one may try to solve the PDE by using Fourier transforms or Fourier sine or cosine transforms. Fourier sine (cosine) transforms may be useful when the spatial variable is unbounded and the function being transformed is known to be odd (even).

Laplace transforms (reviewed in the next section) are (usually) applied to time dependent problems (hyperbolic and parabolic PDE) where the equation holds for all time $t \in [0,\infty)$. The transform is applied to the time variable. In such cases the method can be used when the spatial variable is bounded ($x \in [0,l]$) or semi-bounded ($x \in [0,\infty)$).

[for more on integral transforms, see Zauderer]

[3]After separating variables with $u(x,t) = \psi(x)\phi(t)$, the function $\psi(x)$ satisfied the condition of a regular SLP on the interval $[a,b] = [0,l]$. The boundary operators

$$\begin{aligned} B_1\,\psi &= \alpha_1\,\psi(a) - \beta_1\,\psi'(a) \;=\; 0 \\ B_2\,\psi &= \alpha_2\,\psi(b) - \beta_2\,\psi'(b) \;=\; 0 \end{aligned}$$

satisfied $\alpha_1 = \alpha_2 = 1$ and $\beta_1 = \beta_2 = 0$.

3.7 Laplace transforms

The Laplace transform can be motivated as a special case of the Fourier transform (see [Zauderer]). Laplace transforms are applicable to ODE and PDE and are introduced in most basic references on ODE and engineering mathematics handbooks.

The **Laplace transform** of a function f is given by:

$$L[f] = \overline{f}(s) := \int_0^\infty e^{-st} f(t)\,dt.$$

It is defined if f is exponential order[4], i. e. if $\exists\, c, M > 0$ and $T > 0$ such that $\forall\, t \geq T$ we have $|f(t)| \leq Me^{ct}$.

The inverse transform is given by

$$L^{-1}[\overline{f}] \;=\; \frac{1}{2\pi i} \int_\gamma e^{st}\,\overline{f}(s)\,ds,$$

where γ is a closed curve in the complex plane $\mathbb{C}$. In general these integrals are difficult to compute and require techniques from complex integration. Traditionally Laplace transforms are looked up in tables (which usually accompany textbook notes).

Properties of Laplace transforms:

Linearity:

$$L(\alpha\, f + \beta\, g) \;=\; \alpha\, L\,[f] + \beta\, L\,[g] \quad \text{for} \;\; \alpha, \beta \in \mathbb{R}.$$

The derivative formula for Laplace transform:

$$L[\frac{\partial^n}{\partial t^n}] \;=\; \overline{f^{(n)}}\;(s) \;=\; s^n\,\overline{f}\,(s) - s^{n-1}.f(0) - \;...\; - f^{n-1}(0)$$

(the proof is easily verified as an exercise).

Taking Laplace transforms of a differential equation (together with application of the derivative formula) often gives rise to interesting ratios of polynomials $p_1(s)/p_2(s)$. Taking inverse Laplace transforms of such expressions may require manipulation via partial fractions (see for example [Haberman, § 12.2]).

[4]This condition ensures that the integral which defines $\overline{f}$ is finite.

Examples 3.7.1.

[Some simple Laplace transforms.]

1. If $f(t) = 1$,

$$\begin{aligned} L[f] &= \int_0^\infty 1\, e^{-st}\, dt \;=\; \lim_{b\to\infty} \int_0^b e^{-st}\, dt \;=\; \lim_{b\to\infty} \frac{e^{-st}}{s} \Big|_0^b \\ &= \lim_{b\to\infty} \frac{-e^{-st}+1}{s} \;=\; \frac{1}{s}. \end{aligned}$$

2. $L\,[t] = \frac{1}{s^2}$.

3. $L\,[e^{-at}] = \frac{1}{s+a}, \qquad s > -a.$

4. $L\,[\sin(at)] = \frac{a}{s^2+a^2}$.

The **error** and **complementary error functions** occur commonly in problems. The error function is defined by:

$$\mathrm{erf}\,(x) \;=\; \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2}\, du.$$

The **complementary error function** is defined by

$$\begin{aligned} \mathrm{erfc}\,(x) &= 1 - \mathrm{erf}\,(x) \;=\; \frac{2}{\sqrt{\pi}} \int_0^\infty e^{-u^2}\, du \;-\; \int_0^x e^{-u^2}\, du \\ &= \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-u^2}\, du. \end{aligned}$$

[insert sketch]

From tables of Laplace transforms,

$$f(x) = \mathrm{erfc}\,(\frac{a}{2\sqrt{t}}) \quad \Longleftrightarrow \quad L[f] = \frac{1}{s}\, e^{-a\sqrt{s}}.$$

Examples 3.7.2.

[An application of Laplace transforms to simple diffusion problems]

$$\begin{aligned} u_t - k u_{xx} &= 0, \qquad x \geq 0, \quad t \geq 0 \\ u(0,t) &= u_0 \qquad t \geq 0 \\ u(x,0) &= f(x). \end{aligned} \tag{3.5}$$

Taking Laplace transforms in the t variable, it follows from the derivative formula that,

$$\begin{aligned} \overline{u_t}\,(x,s) &= s\,\overline{u}\,(x,s) - f(x), \\ L\,[f(x)] &= \frac{f(x)}{s}. \end{aligned}$$

Hence, (3.5) becomes

$$\begin{aligned} s\overline{u}\,(x,s) - \frac{f(x)}{s} &= \overline{u}_{xx}\,(x,s) \\ \overline{u}\,(0,s) &= L[u_0] \;=\; \frac{u_0}{s},, \end{aligned} \tag{3.6}$$

which is an ODE in the x-variable.

For the simplified case when $f(x) = 0$, the solution to (3.6), with auxiliary equation $r^2 - s/k = 0$, is

$$\overline{u}\,(x,s) \;=\; A(s)\,e^{-x\sqrt{s/k}} \;+\; B(s)\,e^{x\sqrt{s/k}}.$$

Now $B(s) = 0$ (otherwise $\overline{u}\,(x,s) \to \infty \, as \, x \to \infty$). Thus, $\overline{u}(0,s) \;=\; A(s) \;=\; \frac{u_0}{s}$ and $\overline{u}(x,s) \;=\; \frac{u_0}{s}\, e^{-x\sqrt{s/k}}$. Taking inverses,

$$\begin{aligned} u(x,t) = L^{-1}\,[\overline{u}(x,s)] &= u_0\; L^{-1}\,[\frac{1}{s}\,e^{-x\sqrt{s/k}}] \\ &= u_0 \text{ erfc } (\frac{x}{2\sqrt{kt}}) \\ &= u_0\;\frac{2}{\sqrt{\pi}}\int_0^x e^{-(y^2/4kt)}\;dy.\frac{d}{dy}(\frac{y}{2\sqrt{kt}}) \\ &= \frac{u_0}{\sqrt{\pi kt}}\int_0^x e^{-(y^2/4kt)}\;dy. \end{aligned}$$

The case when $f(x) \neq 0$ involves much more complicated.

When the problem is given by

$$\begin{aligned} u_t - k u_{xx} &= 0 \qquad 0 < x < l, \; t > 0 \\ u(0,t) &= 0, \;\; u(l,t) \;=\; u_0 \\ u(x,0) &= f(x), \end{aligned}$$

then Laplace transforms may be used as an alternative to the method of separation of variables.

Calculation becomes complicated quite quickly and tidying things up to take inverse transforms may be tricky (c.f. Zauderer Example 5.12 (2nd edition) or Zill, page 584 (2nd edition)).

3.8 Laplace Transforms and Convolution

If one defines

$$(f * g)(t) = \int_0^t f(\tau)\, g(t-\tau\,)d\tau,$$

then the convolution theorem also holds for Laplace transforms:

$$L\,[f * g] \quad = \quad \overline{f}(s)\,\overline{g}(s),$$

and

$$L^{-1}[\overline{f}(s)\ \overline{g}(s)] \quad = \quad (f * g)(t).$$

Examples 3.8.1.

[Another application of Laplace transforms to simple PDE]

$$\begin{aligned} u_t + x\ u_x &= x, \qquad x > 0, t > 0 & (3.7)\\ u(x,0) &= \ x > 0 & (3.8)\\ u(0,t) &= 0\,t > 0 & (3.9) \end{aligned}$$

We may take the Laplace transform with respect to either variable. Transforming (3.7) and applying the derivative formula in the t variable we get

$$\underbrace{s\,\overline{u}\,(x,s) - u\,(x,0)}_{L[u_t]} + \underbrace{x\ \overline{u_x}\,(x,s)}_{L[xu_x]} = \underbrace{\frac{x}{s}}_{L[x]} \qquad (3.10)$$

Since $u\,(x,0) = 0$, (3.10) can be written

$$\overline{u}_x\,(x,s) + \frac{s}{x}\,\overline{u}\,(x,s) = \frac{1}{s}. \qquad (3.11)$$

This is a 1^{st} order linear ODE in the x-variable with integrating factor x^s. Solving (3.11) we get

$$\overline{u}\,(x,s) = \frac{c}{x^s} + \frac{x}{s(s+1)},$$

where c is a constant of integration. Since

$$\overline{u}\,(0,s) = \int_0^\infty u\,(0,t)\,e^{-st}\,dt = 0,$$

it follows that $c = 0$. Thus, $\overline{u}\ (x,s) = \frac{x}{s(s+1)}$ and (from table of transforms) we have

$$\begin{aligned} u\,(x,t) &= x\,L^{-1}\,(\frac{1}{s}) - x\,L^{-1}\,(\frac{1}{s+1})\\ &= x\,(1-e^{-t}) \end{aligned}$$

(which satisfies (3.7) and its given initial and boundary conditions).

Examples 3.8.2.

[An application to ODE]
Solve

$$\begin{aligned} \frac{d^2}{dt^2} + 4y &= 3 \quad \text{subjected to} \\ y(0) &= 1 \\ \frac{dy}{dt}(0) &= 5. \end{aligned}$$

Taking Laplace transform yields

$$s^2\,\overline{y}\,(s) - s - 5 + 4\,\overline{y}(s) = \frac{3}{5}.$$

Thus,

$$\begin{aligned} \overline{y}\,(s) &= (\frac{3}{5} + 5 - 5)\frac{1}{s^2+4} \\ &= \frac{3}{s(s^2+4)} + \frac{s}{s^s+4} - \frac{5}{s^2+4}. \end{aligned}$$

Now, by partial fractions[5]

$$\frac{3}{s(s^2+4)} = \frac{3}{4} - \frac{3}{4}(\frac{5}{s^2+4})$$

Hence,

$$\overline{y}\,(s) = \frac{3}{4} - \frac{3}{4}(\frac{5}{s^2+4}) + \frac{5}{s^2+4} - \frac{5}{s^2+4}.$$

Taking inverse transforms from table we get

$$y(t) = \frac{3}{4} - \frac{1}{4}\cos{(2t)} + \frac{5}{2}\sin{(2t)}.$$

5

$$\frac{3}{s(s^2+4)} = \frac{a}{s} + \frac{bs+c}{s^2+4} = \frac{as^2+4a+bs^2+sc}{s(s^2+4)}$$

implies $as^2 + bs^2 = 0$ which in turn implies $a = -b$, and $4a + sc = 0$ implies $c = 0$ and $= \frac{3}{4}$.

Chapter 4

Fundamental solutions and Greens functions

4.1 Scaling and Similarity solutions

This sections gives a technique for solving the diffusion equation by considering underlying symmetries of equation of without reference to boundary condition or initial values.

The homogeneous diffusion equation,

$$u_t - \nabla^2 u = 0 \tag{4.1}$$

involves 1 derivative in the time variable and 2 derivatives in the spatial variables. Considering the case $u = u(x,t)$ where x is 1 dimensional, it follows that if u is a solution to (4.1), then so is $u(\lambda x, \lambda^2 t)$. To see this let $X = \lambda x,\ T = \lambda^2 t$ then $u_{xx} = u_{XX}\lambda^2$ and $u_t = u_T\ \lambda^2$ and hence (4.1) implies

$$u_T - u_{XX} = 0.$$

Now for $\frac{X}{x} = \sqrt{\frac{T}{t}}$ and $\forall \lambda \neq 0$ we have $\frac{X}{\sqrt{T}} = \frac{x}{\sqrt{t}}$. Thus, our solution seems[1] to depend on $\frac{x}{\sqrt{t}}$, i.e. it seems that $u(x,t) = V(\frac{x}{\sqrt{t}})$. Suppose[2]

$$u(x,t) = \frac{1}{t^\alpha} V\,(\frac{x}{\sqrt{t^\beta}}), \tag{4.2}$$

where $x \in \mathbb{R},\ t > 0,\ \alpha, \beta \in \mathbb{R}$.

Substituting (4.2) into (4.1)we get

$$\alpha t^{-(\alpha+1)}\ \ V(\xi) - \beta t^{-(\alpha+1)}\,\xi\, V\xi\,(\xi) - t^{-(\alpha+2\beta)}\, V_{\xi\xi}(\xi) = 0, \tag{4.3}$$

where $V(\xi) = V(\frac{x}{t^\beta})$. Letting $\beta = \frac{1}{2}$ in (4.3) we get

$$\alpha V(\xi) + \frac{\xi}{2}\, V_\xi(\xi) + V_{\xi\xi}(\xi) = 0. \tag{4.4}$$

Now let $\alpha = \frac{1}{2}$. Then (4.4) becomes

$$\frac{d}{d\xi}\,(\frac{1}{2}\,\xi\, V + V_\xi) = 0$$

[1] we're conjecturing

[2] proceeding with $u(x,t) = v(\frac{x}{\sqrt{t}})$ also works but the given approach is quicker

and hence,

$$\frac{1}{2}\,\xi V + V_\xi = c, \tag{4.5}$$

where c is constant.

Now we note some reasonable assumptions for a diffusion problem:

1. $u(x,t) \longrightarrow 0$ as $x \longrightarrow \pm\infty$, and
2. $u_x(x,t) \longrightarrow 0$ as $x \longrightarrow \pm\infty$,

Thus,

$$\begin{aligned} V(\xi) &\longrightarrow 0 \text{ as } \xi \longrightarrow 0, \\ V_\xi(\xi) &\longrightarrow 0 \text{ as } \xi \longrightarrow 0. \end{aligned}$$

and hence $c = 0$ in (4.5). It follows that

$$V(\xi) = K\,e^{-\xi^2/4}$$

and

$$u(x,t) = K\,t^{-\frac{1}{2}}\,e^{-x^2/4t}$$

Letting $K = \frac{1}{\sqrt{4\pi}}$ we get

$$u(x,t) = \frac{1}{\sqrt{4\pi t}}\,e^{-x^2/4t}$$

and

$$\int_{-\infty}^{\infty} u(x,t)\,dx = 1.$$

$$u(x,t) = \frac{1}{\sqrt{4\pi t}} e^{-x^2/4t}$$

is referred to as the fundamental solution for (4.1). The technique given in this section only works when special symmetries exist in the equation. Once the fundamental solution to the PDE is found, it can be shown that general solutions are built out of these fundamental solutions[3].

[3]For reference on the above calculation see L.C Evans, chapter 2 (for $x \in \mathbb{R}^n$). In Dewynne, Howison & Wilmott, chapter 5, proceed with $u(x,t) = V(\frac{x}{\sqrt{t}})$, but omit detail

4.2 Introduction to Greens function

Case 1 - Series solutions: The series solution to the problem

$$\begin{aligned} u_t &= ku_{xx} \\ u(x,0) &= g(x) \\ u(0,t) = u(l,t) &= 0, \end{aligned}$$

was found to be

$$u(x,t) = \sum_{n=1}^{\infty} a_n \sin\left(\frac{n\pi x}{l}\right) e^{-k(\frac{n\pi}{l})^2 t}.$$

The initial condition

$$u(x_0,t) = g(x_0) = \sum a_n \sin(\frac{n\pi x}{l})$$

implied that

$$a_n = \frac{2}{l}\int_0^l g(x_0)\sin(\frac{n\pi x}{l})dx_0.$$

Here, the variable x_0 is used to emphasise that $g(x_0)$ is given at $t = 0$. Thus,

$$\begin{aligned} u(x,t) &= \sum_{n=1}^{\infty}(\frac{2}{l}\int_0^l g(x_0)\sin(\frac{n\pi x_0}{l})dx_0)\sin(\frac{n\pi x}{l})e^{-k(\frac{n\pi}{l})^2 t} \\ &= \int_0^l g(x_0)(\sum_{n=1}^{\infty}\frac{2}{l}\sin(\frac{n\pi x}{l})\sin(\frac{n\pi x_0}{l})e^{-k(\frac{n\pi}{l})^2 t}dx_0. \end{aligned}$$

Since we are integrating over the entire region $[0,l]$ it follows that **the initial value** $g(x_0) = u(x_0,0)$ **at every position** $x_0 \in [0,l]$ **contributes to the subsequent value of** $u(x,t)$ **at position** x **at time** t. Let

$$G(x,t;x_0,0) := \sum_{n=1}^{\infty}\frac{2}{l}\sin(\frac{n\pi x_0}{l})\sin(\frac{n\pi x}{l})e^{-k(\frac{n\pi}{l})^2 t}.$$

G is referred to as the influence function for the initial condition and expresses that the value of u at position x at time t is due to the value at x_0 at time $t = 0$:

$$u(x,t) = \int_0^l g(x_0)G(x,t;x_0,0)dx_0.$$

More generally, we found the solution to

$$\begin{aligned} u_t &= ku_{xx} + F(x,t) \\ u(x_0,0) &= g(x_0) \\ u(0,t) = u(l,t) &= 0, \end{aligned}$$

by taking series expansions of each of the functions in the PDE :

$$u(x,t) = \sum_{n=1}^{\infty} a_n \sin(\frac{n\pi x}{l}) e^{-k(\frac{n\pi}{l})^2 t} + \int_0^t F_n(\tau) e^{-k(\frac{n\pi}{l})^2(t-\tau)} \sin(\frac{n\pi x}{l})\, d\tau,$$

where

$$a_n = \frac{2}{l}\int_0^l g(x_0)\sin\ (\frac{n\pi x_0}{l})\, dx_0, \text{ and}$$

$$F_n(\tau) = \frac{2}{l}\int_0^l F(y,\tau)\sin(\frac{n\pi y}{l})\, dy.$$

Thus,

$$\begin{aligned} u(x,t) &= \int_0^l g(x_0) \sum_{n=1}^{\infty} \frac{2}{l}\sin(\frac{n\pi x}{l})\sin(\frac{n\pi x_0}{l}) e^{-k(\frac{n\pi}{l})^2 t}\, dx_0 \\ &+ \int_0^t \int_0^l F(y,\tau) \sum_{n=1}^{\infty} \frac{2}{l}\sin(\frac{n\pi x}{l})\sin(\frac{n\pi y}{l}) e^{-k(\frac{n\pi}{l})^2(t-\tau)} dy\, d\tau \end{aligned}$$

Let

$$G(x,t;\ y,\tau) = \sum_{n=1}^{\infty} \frac{2}{l}\ \sin(\frac{n\pi x}{l})\ \sin(\frac{n\pi x_0}{l})\, e^{-k(\frac{n\pi}{l})^2(t-\tau)},$$

we have

$$u(x,t) = \int_0^l g(x_0)\, G(x,t;x_0,0)\ dx_0 + \int_0^t \int_0^l F(y,\tau)\, G(x,t;y,\tau) d\tau\ dy.$$

Case 2 - Integral solutions: For the IVP

$$\begin{aligned} u_t &= k u_{xx}, \quad x \in \mathbb{R}\ , t > 0 \\ u(x,0) &= g(x) \end{aligned} \tag{4.6}$$

we found the solution by taking Fourier transforms:

$$u(x,t) = \frac{1}{\sqrt{4\pi k t}} \int_{-\infty}^{\infty} g(y) \exp(-\frac{(x-y)^2}{4kt})\ dy. \tag{4.7}$$

Similarly,

$$\begin{aligned} u_t &= k u_{xx} + F(x,t) \\ u(x,0) &= g(x) \end{aligned} \tag{4.8}$$

can be solved by taking transforms of each of the terms and solving the resulting ODE problem. When $g(x) \neq 0$ the problem can be solved by reducing it to (4.6) via Duhamel's ppl. In this case the solution is

$$
\begin{aligned}
u(x,t) &= \frac{1}{\sqrt{4\pi kt}} \int_{-\infty}^{-\infty} g(y) \exp\left(-\frac{(x-u)^2}{4kt}\right) dy \\
&+ \int_0^t \frac{1}{\sqrt{4\pi kt(t-\tau)}} \int_{-\infty}^{-\infty} F(y,\tau) \exp\left(-\frac{(x-y)^2}{4k(t-\tau)}\right) dy\, d\tau.
\end{aligned}
$$

Letting

$$
G(x,t;y;\tau) = \frac{1}{\sqrt{4\pi kt(t-\tau)}} \exp\left(-\frac{(x-y)^2}{4k(t-\tau)}\right),
$$

the solutions to (4.6) and (4.8) can be written as

$$
\begin{aligned}
u(x,t) &= \int_{-\infty}^{\infty} g(y)\, G(x,t;y,0)\, dy \quad \text{and} \\
u(x,t) &= \int_{-\infty}^{\infty} g(y)\, G(x,t;y,0)\, dy + \int_0^t \int_{-\infty}^{\infty} F(y,\tau) G(x,t;y,\tau)\, dy\, d\tau,
\end{aligned}
$$

respectively.

The function $G(x,t;y,\tau)$ is referred to as **Green's function**. This name arises in the *method of Green's functions* for constructing solutions. In *our* solution of the diffusion problems, we obtained the Green's function via separation of variables and integral transforms[4]. Green's functions are also referred to as **fundamental solutions** (see Definition 4.5.4).

For the (non-homogeneous) diffusion problems on both bounded and unbounded regions, the Green's function describes how

1. the initial condition at position y at time $t = 0$ contributes to the subsequent value of the solution at position x at a later time time t
2. the value of the driving term at position y at time τ contributes to the subsequent value of the solution at position x at a later time time t.

The Green's function of a diffusion equation is closely to the **transition probability density function** for some random walk. In Chapter 1 we were considered a 1-dimensional random walk motivated by empirical observation of physical Brownian motion. We were interested in the probability $v(x,t)$ of the particle being at position x at time t after n steps given an initial position of $x = 0$ at time $t = 0$. We used 1-timestep transition probabilities and an asymptotic limit argument to derive a parabolic PDE which described $v(x,t)$. Equivalently, our solution v to the PDE gives the transition probability density function $v(x,t;x',t')$ for the (continuous limit) of the random walk, where the function $v(x,t;x',t')$ describes the probability of the particle being at position x at time t given that it was at position x' at time t' $(t > t')$.

[4]For more on the method of Greens functions, see the example Zauderer or Habermann.

4.3 The Dirac Delta function

Let

$$\delta^{\epsilon}\ (x) = \begin{cases} \frac{1}{2\epsilon} & |x| \le \epsilon \\ 0 & |x| > \epsilon \end{cases}$$

Then

1. $\int\limits_{-\infty}^{\infty} \delta^{\epsilon}\ (x) = 1$

2. $$\lim_{\epsilon \to 0}\ \delta^{\epsilon}\ (x) = \begin{cases} 0 & x\ \neq 0 \\ \infty & x = 0 \end{cases}$$

The object $\lim\limits_{\epsilon \to 0} \delta^{\epsilon}\ (x)$ is **not** a function in the consistent mathematical framework of analysis of real or complex valued functions since $'\infty'$ is not a number. However we can make sense of the notion by passing to the limit within the confines of an integral. This leads to the following definition.

Definition 4.3.1 (1^{st} encounter with the Dirac delta function). *The Dirac delta function $\delta(x)$ is defined within an integral as follows:*

$$\int_{-\infty}^{\infty} \delta(x)\, f(x)\, dx := \lim_{\epsilon \to 0} \int_{-\infty}^{\infty} \delta^{\epsilon}(x)\, f(x) dx, \tag{4.9}$$

where f is a continuous function.

Suppose f is a continuous function, then by the mean value theorem[5],

$$\begin{aligned}
\int_{-\infty}^{\infty} \delta(x)\, f(x)\, dx &= \lim_{\epsilon \to 0} \int_{-\infty}^{\infty} f(x)\, \delta^{\epsilon}(x)\, dx \;=\; \lim_{\epsilon \to 0} \int_{-\epsilon}^{\epsilon} f(x)\, \frac{1}{2\epsilon}\, dx \\
&= \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_{-\epsilon}^{\epsilon} f(x)\, dx \;=\; \lim_{\epsilon \to 0} \frac{1}{2\epsilon}\, f(a)[\epsilon - (-\epsilon)] \quad \text{for some } a \in [-\epsilon, \epsilon] \\
&= \lim_{\epsilon \to 0} f(a), \quad a \in [-\epsilon, \epsilon] \\
&= f(0)
\end{aligned}$$

Similarly it can be shown that $f(\xi) = \int\limits_{-\infty}^{\infty} f(x)\, \delta(x - \xi)\, dx$.

[5]The mean-value theorem for integrals: If f and g are continuous on $[a, b]$ and $g(x)$ does not change sign on $[a, b]$ then $\exists\ \xi$ such that $\int\limits_a^b f(x) g(x)\, dx = f(\xi) \int\limits_a^b g(x)\, dx$

4.4 Rapidly Decreasing Functions and Distributions

Formal definitions are included in this section for the sake of a more complete presentation - the reader may skip ahead to Definition 4.4.4 at the end.

Definition 4.4.1. *Let $\Re$ denote the class of function $f \in C^{\infty}(\mathbb{R})$ which satisfy*

$$\sup_{x\in\mathbb{R}} |x^{\beta} \frac{d^{\alpha}}{dx^{\alpha}} f(x)| < \infty \quad \alpha, \beta \in \mathbb{N}.$$

Such functions are referred to as **rapidly decreasing functions** *(at ∞).*

Examples 4.4.2. *Examples of rapidly decreasing functions are*

1. $f(x) = e^{-|x|^2}$
2. *All functions in $C_0^{\infty}(\mathbb{R})$ where*

   $$C_0^{\infty}(\mathbb{R}) = \{f \in\ C^{\infty}(\mathbb{R}) \mid f(x) = 0 \, for \, x \notin [a, b] \text{ for some closed bound interval } [a, b]\}$$

3. *A classic example of a function in $C_0^{\infty}(\mathbb{R})$ is*

$$f\ (x) = \begin{cases} \exp\ (\frac{1}{(|x|^2-1)}) & if\ \ |x| < 1 \\ 0 & otherwise. \end{cases}$$

The definition for rapidly decreasing functions can be extended to functions of many variables as follows:

- Let D^{α} denote the operator $\frac{\partial^{\alpha_1+\alpha_2+\cdots\cdots+\alpha_n}}{\prod_{i=1}^{n} \partial x_i^{\alpha_i}}$ on C^{∞} where $\alpha = (\alpha_1 .....\alpha_n)$.
- The support of a function $f \in C^k(\mathbb{R}^n)$ is the closed subset of $\mathbb{R}^n$ defined by $\text{supp}(f) = \{x \in \mathbb{R}^n \mid f(x) \neq 0\}$. The support of $f : \mathbb{R} \to \mathbb{R}$ is a *compact set* if it is the finite union of closed bounded intervals $[a, b]$. Similarly, a compact set in $\mathbb{R}^n$ is the cartesian product of compact subsets of $\mathbb{R}$. Thus, the set $C_0^{\infty}(\mathbb{R}^n)$ of infinitely differentiable functions with compact support is defined by

  $$C_0^{\infty}(\mathbb{R}^n) = \{f \in C^{\infty}(\mathbb{R}^n) \mid f \text{ has compact support}\}.$$

- The set of rapidly decreasing functions on $\mathbb{R}^n$ is defined by

  $$\Re(\mathbb{R}^n) := \{f \in C_0^{\infty}(\mathbb{R}^n) \mid \sup_{x\in\mathbb{R}} \mid x^{\beta} D^{\alpha} f(x)| < \infty\},$$

  where $\alpha = (\alpha_1, ...\alpha_n)$, $\beta = (\beta_1, ...\beta_n) \in \mathbb{N}^n$ and $x^{\beta} = \prod_{j=1}^{n} x^{\beta j}$.

The 1-variable examples generalise to $\mathbb{R}^n$ in the obvious way.

Definition 4.4.3.
A linear function $T : \Re(\mathbb{R}^n) \to \mathbb{R}$ is called a **tempered distribution**. *The set of all tempered distribution will be denoted $\mathcal{S}$.*

The space of rapidly decreasing functions $\Re(\mathbb{R}^n)$ is contained in a larger space of so-called **test functions** *denoted $\mathcal{D}(\mathbb{R}^n)$. The set $\mathcal{D}(\mathbb{R}^n)$ is just the set C_0^{∞} of infinitely differentiable* *functions with compact support such that C_0^∞ is endowed with a particular topological structure*[6]. **Generalised functions** *or* **distributions**[7], *are the continuous linear functionals on* $\mathcal{D}(\mathbb{R}^)$.

The theory of (tempered) distributions is well-developed[8] and the motivation for their study is manyfold:

- The notion of the limit $\int\limits_{-\infty}^{\infty} f(x)\,\delta(x) = \lim\limits_{\varepsilon\to 0} \int\limits_{-\infty}^{\infty} f(x)\,\delta^{\varepsilon}(x) = 0$ is not yet rigourously defined.
- The physical (practical) intuition of evaluating a function in a distributional sense should be made rigorous.
- Distributions are used for proving existence of fundamental solution to PDE.
- Distributions are used in calculations for constructing solutions to PDE.

Definition 4.4.4 (2^{nd} encounter with the Dirac delta function).
The Dirac delta function is a linear functional on $C(\mathbb{R}^n)$ given as follows:

$$\begin{aligned} \delta_x : C(\mathbb{R}^n) &\rightarrow \mathbb{R} \\ \delta_x(f) &\mapsto \int_{-\infty}^{\infty} f(\xi)\,\delta(x-\xi)d\xi \;=\; f(x). \end{aligned}$$

Hence, $\delta_x \in \mathcal{S}$ since $C(\mathbb{R}^n) \supset \Re(\mathbb{R}^n)$.

[6]This topology is given by the sup norm, $||f||_\infty := \sup\limits_{x\,\varepsilon\,\mathbb{R}} |f(x)|$ together with the family of functions $P_{K,m}(f) = \sup\limits_{\alpha<m, x\in K} |D^\alpha f(x)|$ for each compact set K of $\mathbb{R}^n$. The functions $P_{K,m}(f)$ are referred to as semi-norms.

[7]This special class of functions was introduced by Laurent Schwartz in

[8]See Yosida, for example.

4.5 Distributional Evaluation of Functions and Greens Functions

This section serves to motivate the distributional evaluation of a function (via test functions) and interprets Green's functions in terms of distributions. Examples 4.5.2 and Theorem 4.5.3 may be skipped.

Consider a function f of say a single variable x. Then f may be evaluated by

1. the usual pointwise evaluation of $f(x)$ at every point x.
2. the average value of f over sub-intervals of the domain.
3. the weighted average of f by means of test functions.

The first approach is simple but may not correspond to practical reality. Consider for example attempting to obtain f by the measurement of its value for different values of x (x may be temperature, for example). The short coming lies in the fact that the value of x can generally not be specified exactly. At best it may only be possible to obtain $x \in [a, b]$ and measure corresponding $f(x)$.

The average value $\frac{1}{b-a} \int_a^b f(x)\, d(x)$ offers an alternative but it assumes that f is uniformly distributed on $[a, b]$.

The integral $\int_a^b f(x)\, p(x)\, d(x)$ is the continuous analogue of the weighted average $\sum_{i=1}^{n} f(x)\, p(x)$ where p is a density function containing information on how the value of f varies on $[a, b]$. The integral then gives a valuation of f in terms of the density $P(x)$ which carries information on the occurrence of $x \in [a, b]$, i.e. it is an expected value. This last approach leads to the consideration of *test functions* which may be considered as abstract density functions.

Definition 4.5.1 (Integral Equations). [9]
In general a linear integral equation involves an integral operator

$$L := \int_{\Omega} k(x, y)\, dy \quad \Omega \subset \mathbb{R},$$

[9] [Moiseiwitsch] Integral equations refer to equations in which some unknown function is given within an integral. The general theory of integral equations can be traced back to the pioneering contributions of Volterra and Fredholm round the start of the 20^{th} century. Before then only particular examples were studied. As early as 1782, Laplace used the integral transform:

$$f(x) = \int_0^{\infty} e^{-xs}\, \varphi(s)\, ds \tag{4.10}$$

to solve linear difference & differential equations.

In 1822 Fourier found the reciprocal formulas :

$$f(x) = \frac{2}{l} \int_0^{\infty} \sin_{(\cos)}(x\xi)\, \varphi(\xi)\, dx \tag{4.11}$$

$$\varphi(\xi) = \int_0^{\infty} \sin_{(\cos)}(x\xi)\, \varphi(x)\, dx \tag{4.12}$$

Here (4.5) and (4.6) may be viewed as integral equation f with φ the unknown function to be solved for.

where $k(x,y)$ is referred to the as the kernel[10]*. A linear integral equation for an unknown function f, given a known function g is : $L[f](x) = \int\limits_{\Omega} k(x,y)\, f(y)\, dy = g(x)$.*

Examples 4.5.2. *Famous linear integral equations are given as follows:*

- *The Fredholm equation of the first kind,*

$$\int\limits_a^b k(x,y)\, f(y)\, dy = g(x)$$

  (which includes Laplace's and Fourier's integral).

- *The Fredholm equation of the second kind,*

$$\int\limits_a^b k(x,y)\, f(y)\, dy + g(x) = f(x)$$

  and its corresponding homogenous equation

$$\int\limits_a^b k(x,y)\, f(y)\, dy = f(x).$$

- *The Volterra equation of the first kind*

$$\int\limits_a^x k(x,y)\, f(y)\, dy = g(x)$$

- *The Volterra equation of the second kind is defined analogously, i.e. in the Fredholm equation of second kind replace b with x in the limits of integration.*

Integral equations are a topic of investigation in their own right, but they clearly have close connections with differential equations. Consider the first order ODE

$$y' = f(x,y)$$

where $y = y(x),\ y' = \frac{d}{dx}y$ and f is any continuous real valued function on some rectangle

$$R = \{(x,y) \mid |x - x_0| < a, |y - y_0| < b \text{ for some } x_0, y_0 \in \mathbb{R}\}.$$

Consider the IVP

$$y' = f(x,y), \quad y(x_0) = y_0 \tag{4.13}$$

and the integral equation

$$y = y_0 + \int\limits_{x_0}^x f(t,y)\, dt. \tag{4.14}$$

A solution to (4.13) is a function φ defined on an interval $I \supset \{x_0\}$ which satisfies both conditions, while a solution to (4.14) is a continuous function φ on some interval $I \supset \{x_0\}$ such that $(x, \varphi(x)) \in \mathbb{R}^2$ for all $x \in I$ and $\varphi(x) = y_0 + \int\limits_{x_0}^x f(t, \varphi(t))\, dt$ for all $x \in I$.

[10]This kernel has nothing to do with the *kernel* N(T) (or nullspace) of a linear operator

Theorem 4.5.3.
A function φ is a solution to (4.13) on I if and only if it is a solution to (4.14) on I[11].

Proof
Suppose φ is a solution to (4.13) on I. Then

$$\varphi'(t) = f(t, \varphi(t)) \tag{4.15}$$

on I. Since φ is continuous on I and f is continuous, $F(t) := f(t, \varphi(t))$ is continuous (composition of continuous function). Integrating (4.15) from x_0 to x we have

$$\begin{aligned}\varphi(x) &= \varphi(x_0) + \int_{x_0}^{x} f(t, \varphi(t))\, dt \\ &= y_0 + \int_{x_0}^{x} f(t, \varphi(t))\, dt.\end{aligned}$$

Conversely, suppose φ satisfies (4.14), i.e.

$$\varphi(x) = y_0 + \int_{x_0}^{x} f(t, \varphi(t))\, dt \tag{4.16}$$

Then, differentiating with respect to x and applying Liebnitz's rule[12]

$$\begin{aligned}\frac{d}{dx}\varphi(x) &= \frac{d}{dx}\int_{x_0}^{x} f(t, \varphi(t))\, dt \\ &= f(x, \varphi(x))\end{aligned}$$

for all $x \in I$. From (4.14) it also follows that $\varphi(x_0) = y_0$ and hence φ satisfies (4.13).

◇

[11]The solution to (4.14) can be constructed by successive approximations. This is the crux of the proof of *Peano's existence theorem* for solution to first order ODE.

Equation (4.14) is a Volterra equation of second kind. One can also show that a solution to an IVP for a second order linear equation

$$y''(x) + a_1(x)y'(x) + a_2(x)y(x) = f(x)$$

satisfies a Volterra equation of second kind - though the kernels may be completely different! This line of reasoning enables one to verify the existence of solution to the DE problem at hand.
[12]Liebnitz's rule:

$$\frac{d}{dx}\int_{u(x)}^{v(x)} f(t)\, d(t) = f(v(x))\frac{dv}{dx} - f(u(x))\frac{du}{dx}.$$

More generally, if L is a linear partial differential operator, then the problem of solving an equation of the form $Lu = f$, where f is known and u is unknown, may be transformed to one of finding a solution for some integral equation.

Suppose L is invertible , i.e. L^{-1} exists as a linear operator which satisfies $LL^{-1} = L^{-1}L = I$ where I is the identity operator satisfying $I(f) = f$. Then $Lu = f \Leftrightarrow u = L^{-1}\, f$. Presumably L^{-1}, the inverse of a (partial) differential operator, will be an integral operator. Assuming that L is acts on $u(x)$, a function of x, the $u = L^{-1}\, f$ may be written

$$u(x) = \int_{\Omega} k(x,y)\, f(y)\, dy. \tag{4.17}$$

Applying L to the RHS and LHS of (4.17) it follows from the original DE, $Lu = f$, that

$$[Lu](x) = \int_{\Omega} L(k(x,y))\, f(y)\, dy = f(x). \tag{4.18}$$

Assuming equality of functions within the integral, it follows from (4.13) that we seek $k(x,y)$ such that $L(k(x,y)) = \delta(x-y)$

Definition 4.5.4. *A* **fundamental solution**[13] *of a partial differential operator L is a distribution G such that $LG = \delta$.*

We described the Dirac delta function within the context of a limit of integrals and saw that for a continuous function f

$$\int_{\mathbb{R}} f(x)\, \delta(x-\xi) dx = f(\xi) \tag{4.20}$$

for $\xi \in \mathbb{R}$. Thus, as ξ ranges over the domain of f, the function f can be completely specified.

We also noted that the mapping $T_\xi(f) := f(\xi)$ is a linear functional on a vector space of functions. Furthermore, as f varies over the function space, it is the *set of functions* which defines $\delta_\xi(x) = \delta(x-\xi)$ and gives the more consistent definition of $\delta_\xi(x)$ as the linear functional T_ξ (cf. Definitions 4.3.1 and 4.4.4).

Now suppose Φ is a function space and $\phi \in \Phi$. Provided the integral is well-defined, the mapping $T_\phi : \Phi \to \mathbb{R}$ given by

$$T_\phi(f) = \int f(x)\, \phi(x)\, dx \tag{4.21}$$

[13] The existence of a fundamental solution for every linear PDE with constant coefficients was proved independently by B. Malgrange and L.Ehrenpreiss around 1954 - 1955. A proof due to L.Hormander (who won the Fields medal in 1962) is accessible with some effort (see Yosida).

In 1957 H.Lewy constructed the equation

$$-i\,\frac{\partial u}{\partial x_1} + \frac{\partial u}{\partial x_2} - 2(x_1 + ix_2)\frac{\partial u}{\partial x_3} = f(x_3) \tag{4.19}$$

which has no solution if f if not *analytic* [an analytic function f is a complex function with complex derivative of every point in it's domain (and hence, it has derivatives of all orders) such that it agrees locally with it's Taylor series $\sum_{n=1}^{\infty} \frac{1}{n!}\,(x-a)^n\, f^{(n)}(a)$; analytic functions are also referred to as regular or holomorphic functions].

is a linear functional on Φ. Thus, using inner product notation, we conclude that **a family of functions Φ can describe f by valuation**[14] **in terms of** $\{(f, \varphi) \in \mathbb{R} \mid \phi \in \Phi\}$.

This provides a more easily used definition for generalised functions that is consistent with the abstract rigorous definition in Section 4.4. We can summarise the key properties of as follows:

1. A generalised function f is *defined* with respect to a family of test functions Φ (which vanish at ∞ and are rapidly decreasing). It is a continuous linear functional which can be denoted $(f, \_) : \Phi \to \mathbb{R}$, i.e. for $\varphi \in \Phi$ we have $\phi \to (f, \varphi)$.

2. For φ, $\psi \in \Phi$ and $\alpha, \beta \in \mathbb{R}$ we have
$$(f, \alpha\varphi + \beta\psi) = \alpha\,(f, \varphi) + \beta\,(f, \psi) \tag{4.22}$$

3. Two generalised functions, f and g , are equal if $(f, \varphi) = (g, \varphi)$ for every $\varphi \in \Phi$.

4. The mapping $(f, \_) : \varphi \to \mathbb{R}$ is continuous function with respect to the topology on Φ, i.e. if $\varphi_n(x), n \in \mathbb{N}$ is a set of test functions such that $\varphi_n \to \varphi \in \Phi$ then
$$\lim_{n\to\infty} (f, \varphi_n) = (f, \varphi). \tag{4.23}$$

5. The set of all generalised functions with respect to Φ forms a vector space (linear functionals are special types of linear operators). Furthermore, if a sequence of generalised functions, f_n satisfies
$$\lim_{n\to\infty} (f_n, \varphi) = (f, \varphi) \tag{4.24}$$
for every $\varphi \in \Phi$, then we say f_n converges to f (this type of convergence is referred to as weak convergence).

Thus, our first definition of the delta function (Definition 4.3.1) as the limit of the function δ^ε in the context of an integral is made rigorous in the context of generalised functions and the definition is consistent with our second definition of the delta function (Definition 4.4.4) as the *point-evaluation* linear functional $T_\xi(f) = f(\xi)$. In what follows we assume that a generalised function is concretely represented as an integral
$$(f, \varphi) = \int f(x)\,\varphi(x)\,dx. \tag{4.25}$$

Definition 4.5.5. *The equality* $(f', \varphi) = -(f, \varphi)$ *defines the* **generalised** *or* **distributional derivative** *of a function.*

To motivate the definition, suppose f and f' generalised functions of a single real variable. Then, since $\varphi' \in \Phi$, we have
$$\begin{aligned}(f', \varphi) &= \int_{-\infty}^{\infty} f'(x)\,\varphi(x)\,dx \\ &= F(x)\,\varphi(x)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f(x)\,\varphi'(x)\,dx \\ &= -(f, \varphi)\end{aligned}$$

[14] Riesz's theorem ensures that any linear functional on an inner product space can be represented as as an inner product.

Examples 4.5.6. *Let*

$$H\ (x-\xi) = \begin{cases} 1 & \text{if } x \geq \xi \\ 0 & \text{if } x < \xi \end{cases}$$

Then H is a generalised function given by

$$(H, \varphi) = \int_{-\infty}^{\infty} H(x-\xi)\,\varphi(x)\,dx = \int_{\xi}^{\infty} \varphi(x)\,dx$$

and H' is a generalised function

$$(H', \varphi) \quad = \quad -(H, \varphi) = -\int_{\xi}^{\infty} \varphi'(x)\,dx = -\varphi(x)]_{\xi}^{\infty} = \varphi(\xi).$$

But

$$(\delta_{\xi}, \varphi) \quad = \quad \int_{-\infty}^{\infty} \delta(x-\xi)\,\varphi(x)\,dx = \varphi(\xi).$$

Thus, $H'(x-\xi) = \delta(x-\xi)$.

Definition 4.5.7. *The* **Fourier transform $\hat{f}$ of a generalised function**[15] *f is defined by:*

$$(\hat{f}, \varphi) := (f, \hat{\varphi}),$$

where $\varphi \in \Re(\mathbb{R})$.

To motivate this definition, suppose $\hat{f}$ and $\hat{g}$ exist. Then

$$\begin{aligned}
\int_{-\infty}^{\infty} \hat{f}(\xi)\,g(\xi)\,d(\xi) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)\,e^{-i\xi x}\,dx\,g(\xi)\,d\xi \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x)\,g(\xi)\,e^{-i\xi x}\,dx\,d\xi \\
&= \int_{-\infty}^{\infty} f(x)\,\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(\xi)\,e^{-i\xi x}\,d\xi\,dx \\
&= \int_{-\infty}^{\infty} f(x)\,\hat{g}(x)\,dx.
\end{aligned}$$

[15]If $\varphi \in \Re(\mathbb{R})$ then $\hat{\varphi}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \varphi(x)\,e^{-i\xi x}\,dx$ is defined. In general $\hat{\varphi}$ may not be defined functions in $\mathcal{D}(\mathbb{R})$. In fact this is the reason for working with the class of rapidly decreasing functions and not just functions in $C_0^{\infty}(\mathbb{R})$.

Examples 4.5.8. *The Fourier of the delta $\delta\,(x-y)$ is computed by*

$$\begin{aligned}
(\hat{\delta}(\xi-y),\varphi(\xi)) &= (\delta(x-y),\hat{\delta}(x)) \\
&= \int_{-\infty}^{\infty}\delta(x-y)\,\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-i\xi x}\,\varphi(\xi)d\xi\,dx \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\delta(x-y)\,e^{-i\xi x}\,dx\varphi(\xi)\,d\xi \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-i\xi x}\,\varphi(\xi)\,d\xi \\
&= (\frac{1}{\sqrt{2\pi}}\,e^{-i\xi x},\varphi(\xi)).
\end{aligned}$$

Thus, $\hat{\delta}(\xi-y)=\frac{1}{\sqrt{2\pi}}\,e^{-i\xi y}$.

4.6 Application to Diffusion problems

First we need to define the n-dimensional delta function. The 2-dimensional delta function $\delta(x_1,x_2)$ is defined by $\delta(x_1,x_2)=\delta(x_1)\delta(x_2)$. To see this

$$\begin{aligned}
\int\int\delta(x_1,x_2)\,f(x_1,x_2)\,dx_1\,dx_2 &= \int\int\delta(x_1)\,\delta(x_2)\,f(x_1,x_2)\,dx_1\,dx_2 \\
&= \int\int\delta(x_1)\;f(x_1,x_2)\,dx_1\,\delta(x_2)\,dx_2 \\
&= \int f(0,x_2)\,\delta(x_2)\,dx_2 \\
&= f(0,0)
\end{aligned}$$

as expected. This definition extends to functions of n variables: $\delta(x_1,...,x_2)=\delta(x_1)...\delta(x_n)$.

Recall that if

$$L\,G(x) \;=\; \delta(x)$$

then

$$\begin{aligned}
[Lu](x) &= f(x) \\
&= \int f(y)\,\delta(y-x)\,dy \\
&= \int f(y)\,L\,G(y-x)\,dy,
\end{aligned}$$

which implies

$$u(x) \;=\; L^{-1}\int f(y)\,L\,G\,(y-x)\,dy,$$

provided L is invertible[16]. For the diffusion equation this amounts to solving the problem

$$\frac{\partial G}{\partial t}-k\frac{\partial G}{\partial x^2}=\delta(x-x_0)\,\delta(t-t_0), \tag{4.26}$$

[16] With some more effort it can be shown that $L\,G(x)=\delta(x)\Rightarrow u(x)=\int f(y)\,G(y-x)\,dy$ is a solution to the equation $Lu=f$ even when L is not invertible (see Yosida)

where x, x_0 and $\xi \in \mathbb{R}$, subject to the condition $G(x,t;x_0,t_0) = 0$ for $t < t_0$.

Taking Fourier transforms in (4.26)

$$\frac{\partial \hat{G}}{\partial t} + k\,\xi^2\,\hat{G} = \frac{1}{\sqrt{2\pi}}\;e^{-i\xi x_0}\,\delta(t-t_0). \tag{4.27}$$

This is an ODE in the t variable which is solved by means of the integrating factor $\exp[\int_0^t k\,\xi^2\,dt] = e^{k\xi^2 t}$: multiply the LHS and RHS of (4.26) by integrating factor

$$e^{k\xi^2 t}\,\frac{\partial \hat{G}}{\partial t} + k\xi^2\,e^{k\xi^2 t}\,\hat{G} = \frac{1}{\sqrt{2\pi}}\,e^{-i\xi x_0}\,e^{k\xi^2 t}\,\delta(t-t_0)$$

$\Leftrightarrow$

$$\begin{aligned}
& \frac{d}{dt}\,(e^{k\xi^2 t}\,\hat{G}) &=& \int \frac{1}{\sqrt{2\pi}}\,e^{-i\xi x_0}\,e^{k\xi^2 t}\,\delta(t-t_0)dt \\
\Leftrightarrow\quad & e^{k\xi^2 t}\,\hat{G} &=& \int \frac{1}{\sqrt{2\pi}}\,e^{-\xi x_0}\,e^{k\xi^2 t_0}\,\delta(t-t_0)dt \\
\Rightarrow\quad & e^{k\xi^2 t}\,\hat{G} &=& \frac{1}{\sqrt{2\pi}}\,e^{-\xi x_0}\,e^{k\xi^2 t_0}\,H(t-t_0) + C(\xi) \\
\Rightarrow\quad & \hat{G} &=& \frac{1}{\sqrt{2\pi}}\,e^{-\xi x_0}\,e^{k\xi^2 t_0(t-t_0)}\,H(t-t_0) + C(\xi)\,e^{-k\xi^2 t}
\end{aligned}$$

Since $\hat{G}(\xi,t;x_0,t_0) = 0 = H(t-t_0)$ at $t = 0$ it follows $C(\xi)\,e^{-k\xi^2 t} = 0$, i.e. $C(\xi) = 0$. Thus,

$$\begin{aligned}
G(x,t;x_0,x_0) &= \frac{1}{\sqrt{2\pi}}\int \hat{G}(\hat{\xi},t;x_0,t)\,e^{i\xi x}\,d\xi \\
&= \frac{1}{\sqrt{2\pi}}\int e^{-\xi x_0}\,e^{-k\xi^2(t-t_0)}\,e^{i\xi x}\,d\xi\,(H(t-t_0)) \\
&= \frac{1}{\sqrt{2\pi}}\left[\sqrt{\frac{\pi}{k(t-t_0)}}\exp\left(-\frac{(x-x_0)^2}{4k(t-t_0}\right)\right]H(t-t_0) \\
&= \frac{1}{\sqrt{4\pi k(t-t_0)}}\;\exp\left(-\frac{(x-x_0)^2}{4k(t-t_0)}\right)\,H(t-t_0)
\end{aligned}$$

This method extends easily to higher dimensions: recall the Fourier transform of a function of n-variable

$$F[f](\xi) = \hat{f}(\xi) = (\frac{1}{\sqrt{2\pi}})^n \int\limits_{\mathbb{R}^n} f(x)\, e^{-i(\xi * x)}\, dx \tag{4.28}$$

where $x = (x_1, ..., x_n), \xi = (\xi_1, ..., \xi_n)$

It follows that the Fourier transform of $\delta(x-y) = \delta(x_1 - y_1)....\delta(x_n - y_n)$ is given by

$$\begin{aligned}
(\delta(\xi - y), \varphi(\xi)) &= (\delta(x-y), \varphi(x)) \\
&= \int\limits_{\mathbb{R}^n} \delta(x-y)(\frac{1}{\sqrt{2\pi}})^n \int\limits_{\mathbb{R}^n} e^{-i(\xi * x)}\, \varphi(\xi) d\xi\, dx \\
&= (\frac{1}{\sqrt{2\pi}})^n \int\limits_{\mathbb{R}^n} \int\limits_{\mathbb{R}^n} \delta(x-y)\, e^{-i(\xi * x)}\, dx\, \varphi(\xi)\, d\xi \\
&= (\frac{1}{\sqrt{2\pi}})^n \int\limits_{\mathbb{R}^n} e^{-i(\xi * y)}\, \varphi(\xi)\, d\xi.
\end{aligned}$$

Thus $\hat{\delta}(\xi - y) = (\frac{1}{\sqrt{2\pi}})^n\, e^{-i(\xi - y)}$. Proceeding as before for the case when x is 1-dimensional, 3-dimensional Fourier transform yields

$$G(x_1, x_2, x_3, t;\ \xi_1, \xi_2, \xi_3, t_0) = \frac{H(t - t_0)}{(\sqrt{4\pi k (t - t_0)})^3} \exp \left(\frac{\sum\limits_{i=1}^{3} (x_i \xi_i)^2}{4k(t - t_0)} \right). \tag{4.29}$$

Part II

Application to Derivative Securities

Chapter 5

Black-Scholes PDE and solutions

5.1 The BSM differential equation

The PDE provides a mathematical model for the price of derivative securities under very particular assumptions[1]:

1. The stock price follows geometric Brownian motion with μ and σ constant.
2. Short selling of securities, with full use of proceeds, is permitted.
3. There are no transaction costs or taxes and all securities are perfectly divisible.
4. There are no (riskless) arbitrage opportunities.
5. Trading is continuous.
6. The risk-free interest rate is constant and the same for all maturities.

5.1.1 Derivation of the PDE

Assume a lognormal stock process for stock price S:

$$dS = \mu S dt + \sigma S dz \tag{5.1}$$

Suppose f is the price derivative dependent on S then

$$\begin{aligned} f : \mathbb{R} \times \mathbb{R}^+ &\longrightarrow \mathbb{R} \\ (S, t) &\longrightarrow f(S, t) \end{aligned}$$

If f is twice differentiable then by Itô's lemma

$$df = (\mu S \frac{\partial f}{\partial S} + \frac{\partial f}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2})dt + \frac{\partial f}{\partial S}\sigma S dz \tag{5.2}$$

In particular, the Brownian motion underlying f and S are the same. In discrete form, (5.1) and (5.2) may be written,

[1]see Hull, Chapter 12

$$\Delta S = \mu S \Delta t + \sigma S \Delta z \tag{5.3}$$

$$\Delta f = (\frac{\partial f}{\partial S}\mu S + \frac{\partial f}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2})\Delta t + \frac{\partial f}{\partial S}\sigma S \Delta z \tag{5.4}$$

We want to create risk free portfolio by putting together negatively correlated instruments in order to replicate an option:

- Choose -1 quantity of derivative (short position)
- $+M$ quantity of share

Then the portfolio is equal to

$$\begin{aligned} \Pi &= -f + MS, \\ \Delta\Pi &= -\Delta f + M\Delta S, \\ &= -(\frac{\partial f}{\partial t} + \mu S\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2})\Delta t \ - \sigma S \frac{\partial f}{\partial S}\Delta z + M\mu S d\Delta t + M\sigma S \Delta z \end{aligned}$$

Here $dz =$ is the risky part $\rightsquigarrow$ let $M = \frac{\partial f}{\partial s}$

$$\Delta\Pi = (-\frac{\partial f}{\partial S} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2})\Delta t \tag{5.5}$$

This gives us a riskless portfolio for time t to $t + \Delta t$. What should the return on portfolio be?

Let r denote rate of return of riskless security.

- If r is greater than rate of return on Π then short the portfolio and long the riskless asset.
- If r is less than rate of return on Π then long the portfolio and short the riskless asset.

Thus we may assume that rate of return of portfolio equals the rate of return of riskless security in market.

$$\begin{aligned} \frac{\Delta\Pi}{\Delta t} &= r\Pi, \quad \text{or} \\ \Delta\Pi &= r\Pi\Delta t \end{aligned} \tag{5.6}$$

From (5.5) and (5.6)

$$(-\frac{\partial f}{\partial S} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2})\Delta t = r(-f + \frac{\partial f}{\partial S}S)\Delta t \tag{5.7}$$

i.e.

$$\frac{\partial f}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 f}{\partial S^2} + rS\frac{\partial f}{\partial S} - rf = 0.$$

5.2 Truncated expectation of log-normally distributed random variable

The density function for a normally distributed random variable $\mathbf{X}$ with mean μ_x and variance σ_x^2, i.e $X \sim N(\mu_x, \sigma_x^2)$, is

$$f_x(x) \quad = \quad \frac{1}{\sigma\sqrt{2\pi}} \exp[-\frac{(x-\mu_x)^2}{2\sigma_x^2}].$$

The density function for the random variable $\mathbf{Y} = \mathbf{e}^{\mathbf{X}}$ for $X \sim N(\mu_x, \sigma_x^2)$ is

$$f_y(y) = \begin{cases} \frac{1}{y}\frac{1}{\sigma\sqrt{2\pi}} \exp[-\frac{(\ln y-\mu_x)^2}{2\sigma_x^2}] & \text{if } y > 0 \\ 0 & \text{if } y \leq 0 \end{cases}$$

The truncated expectation[2] of $\mathbf{Y}$ is

$$\begin{aligned} \mathbb{E}(Y|Y>a) &= \int_a^{+\infty} y f_Y(y)dy, \\ &= \int_{\ln a}^{+\infty} e^x f_X(x)dx, \\ &= \int_{\ln a}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp[\frac{2\sigma^2 x - x^2 + 2\mu x - \mu^2}{2\sigma^2}]dx, \\ &= \int_{-\ln a}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp[-\frac{(x-(\sigma^2+\mu))^2 + \sigma^4 + 2\mu\sigma^2}{2\sigma^2}]dx, \\ &= \frac{1}{\sigma\sqrt{2\pi}} \exp[\frac{\sigma^2}{2}+\mu] \int_{\ln a}^{+\infty} \exp[-\frac{(x-(\sigma^2+\mu))^2}{2\sigma^2}]dx. \end{aligned}$$

Let $z = \frac{x-(\sigma^2+\mu)}{\sigma}$. Then $dz = \frac{1}{\sigma}dx$ and

$$\mathbb{E}(Y|Y>a) = \frac{1}{\sqrt{2\pi}} \exp[\frac{\sigma^2}{2}+\mu] \int_{\frac{\ln a-(\sigma^2+\mu)}{\sigma}}^{+\infty} \exp[-\frac{z^2}{2}]dz, \tag{5.8}$$

$$= \exp[\frac{\sigma^2}{2}+\mu] N((\sigma^2+\mu-\ln a)/\sigma), \tag{5.9}$$

where

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} \int_{-x}^{+\infty} e^{\frac{z^2}{2}} dz.$$

When a→ 0 the truncated mean becomes the usual mean of a log-normal variable, i.e. $\mu_x = \exp[\frac{\sigma^2}{2}+\mu]$

[2]or *complementary cumulative distribution*

5.3 Solving the Black-Scholes PDE

The PDE for the derivative value is

$$\frac{\partial f}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 f}{\partial S^2} + rS\frac{\partial f}{\partial S} - rf = 0, \tag{5.10}$$

where $0 \leq S \leq \infty$, $0 \leq t \leq T$ and $f(S,T) =$ payoff (final condition).
(5.10) is a backward equation. Letting $\tau = T - t$, (5.10) becomes

$$\frac{\partial f}{\partial \tau} - \frac{\sigma^2 S^2}{2}\frac{\partial^2 f}{\partial S^2} - rS\frac{\partial f}{\partial S} + rf = 0, \tag{5.11}$$

subject to the initial condition $f(S_\tau, 0) =$ payoff.
More constraints:

- $S = 0 \Rightarrow S$ remains zero $\Rightarrow f(0,\tau) = 0$ and
- $S \to \infty \Rightarrow f > 0$ and f will be exercised.

Now let $y = \ln S$. Then (5.11) becomes

$$\frac{\partial f}{\partial \tau} - \frac{\sigma^2}{2}\frac{\partial^2 f}{\partial y^2} - (r - \frac{\sigma^2}{2})\frac{\partial f}{\partial y} + rf = 0 \tag{5.12}$$

for $-\infty < y < \infty$.
Note that now all the coefficients are now constant!

Letting $\omega(y,\tau) = e^{r\tau} f(y,\tau)$, (5.3) becomes

$$\frac{\partial \omega}{\partial \tau} - \frac{\sigma^2}{2}\frac{\partial^2 \omega}{\partial y^2} - (r - \frac{\sigma^2}{2})\frac{\partial \omega}{\partial y} = 0. \tag{5.13}$$

Note that $f(y,\tau) = e^{-r\tau}\omega(y,\tau)$ is sometimes referred to as a future value of $\omega(y,\tau)$.

Finally, let $x = y + (r - \frac{\sigma^2}{2})\tau$. Then (5.13) becomes

$$\frac{\partial \omega}{\partial \tau} - \frac{1}{2}\sigma^2\frac{\partial^2 \omega}{\partial x^2} = 0. \tag{5.14}$$

The fundamental solution or Greens function to (5.14) is

$$G(x,\tau;\xi,0) = \frac{1}{\sqrt{2\pi\sigma^2\tau}} e^{-\frac{(x-\xi)^2}{2\sigma^2\tau}}. \tag{5.15}$$

Thus,

$$\omega(y,\tau) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{\sigma^2 2\pi\tau}} \exp\left(-\frac{(y + (r - \frac{\sigma^2}{2})\tau - \xi)^2}{2\sigma^2\tau}\right) p(\xi) d\xi$$

where $p(\xi)$ is the initial condition (in τ variable), i.e. it is the payoff of the derivative at time T and ξ is the value of value of $y = \ln S$ at time $\tau = 0$, i.e. $\xi = \ln(S_T)$.

The Black-Scholes PDE holds for any derivative security modelled under the set assumptions. By substituting the changes of variables into (5.15) we obtain

$$V(S,\tau) = e^{-r\tau} \int_{-\infty}^{\infty} G(S,\tau;S_T,0)V(S_T,0)dS_T, \tag{5.16}$$

where

$$G(S,\tau;S_T,0) = \frac{1}{S_T\sqrt{2\pi\tau\sigma^2}} \exp\left(-(\frac{\ln(S/S_T)+(r-\frac{\sigma^2}{2})\tau)^2}{2\sigma^2\tau}\right), \tag{5.17}$$

is the Green's function (or fundamental solution) which quantifies the probabilities for all possible payoffs $V(S_T,0)$ for all possible underlying security prices S_T $(= S_{\tau=0})$ at expiry $\tau = 0$, given the initial underlying security price of S at time $t = 0$ $(\tau = T)$.

5.4 The price of a European vanilla call

The payoff at time $t = T$ for European (vanilla) call option is $V(S_t,T) = \max(S_T - K, 0)$. Thus, using $S_\tau = e^{y_\tau}$, $p(\xi) = \max(e^\xi - K, 0)$ and the fact that $e^\xi \geq K$ iff $\xi \geq \ln K$, it follows that

$$\begin{aligned}
\omega(y,\tau) &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{\sigma^2 2\pi\tau}} \exp\left(-\frac{(y+(r-\frac{\sigma^2}{2})\tau-\xi)^2}{2\sigma^2\tau}\right) \max(e^\xi - K,0)d\xi, \\
&= \int_{\ln K}^{+\infty} \frac{1}{\sqrt{\sigma^2 2\pi\tau}} \exp\left(-\frac{(y+(r-\frac{\sigma^2}{2})\tau-\xi)^2}{2\sigma^2\tau}\right) (e^\xi - K)d\xi.
\end{aligned}$$

With the last expression as a difference between two integrals, we compute the first:

$$\begin{aligned}
&= \int_{\ln K}^{+\infty} e^\xi \frac{1}{\sigma\sqrt{\tau 2\pi}} \exp\left(-\frac{(y+(r-\frac{\sigma^2}{2})\tau-S)^2}{2\sigma^2\tau}\right) d\xi \\
&= \exp\left(\frac{\sigma^2\tau}{2} + y + (r-\frac{\sigma^2}{2})\tau\right) N\left(\frac{\sigma^2\tau + y + (r-\frac{\sigma^2}{2})\tau - \ln K}{\sigma\sqrt{\tau}}\right) \\
&= e^{r\tau} S N\left(\frac{\ln(S/K)+(r+\frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}\right).
\end{aligned}$$

Here, the second last equality follows by identifying:

$$\ln K \equiv \ln a, \quad \sigma \equiv \sigma\sqrt{\tau}, \quad \mu \equiv y + (r-\frac{\sigma^2}{2})\tau, \quad \text{and} \quad x \equiv \xi,$$

in the following integrals

$$\begin{aligned}
\int_{\ln a}^{+\infty} e^x f_x(x)dx &= \int_{\ln a}^{+\infty} e^x \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx, \\
&= \exp\left(\frac{\sigma^2}{2}+\mu\right) N\left(\frac{\sigma^2+\mu-\ln a}{\sigma}\right).
\end{aligned}$$

Similarly, the second integral is found

$$
\begin{aligned}
&= \int_{\ln K}^{+\infty} K \frac{1}{\sigma\sqrt{\tau 2\pi}} \exp\left(-\frac{(y+(r-\frac{\sigma^2}{2})\tau-\xi)^2}{2\sigma^2\tau}\right) d\xi \\
&= KN\left(\frac{y+(r-\frac{\sigma^2}{2})\tau-\ln K}{\sigma\sqrt{\tau}}\right) \\
&= KN\left(\frac{\ln(S/K)+(r-\frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}\right),
\end{aligned}
$$

where the second equality was obtained by identifying

$$
\begin{aligned}
z &\equiv \frac{(y+(r-\frac{\sigma^2}{2})\tau)-\xi}{\sigma\sqrt{\tau}}, \quad dz = \frac{1}{\sigma\sqrt{\tau}} d\xi, \quad \text{and} \\
x &\equiv \frac{y+(r-\frac{\sigma^2}{2})\tau-\ln K}{\sigma\sqrt{\tau}}
\end{aligned}
$$

to get

$$
\begin{aligned}
N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{z^2}{2}} dz &= \frac{1}{\sqrt{2\pi}} \int_{-x}^{+\infty} e^{-\frac{z^2}{2}} dz, \\
&= \frac{1}{\sqrt{2\pi}} \int_{\frac{\ln K - y + (r-\frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}}^{+\infty} e^{-z^2} dz.
\end{aligned}
$$

Letting

- $d_1 = \frac{\ln(S/K)+(r+\frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}$ and
- $d_2 = d_1 - \sigma\sqrt{\tau}$

we have

$$
\begin{aligned}
c(S,\tau) &= e^{-r\tau}\omega(y,\tau), \\
&= e^{-r\tau}(e^{r\tau}SN(d_1)+KN(d_2)), \\
&= SN(d_1) - e^{-r\tau}KN(d_2),
\end{aligned}
$$

Immediate observations and the price put options

1. This formula gives the value of $c(S,\tau)$ where $\tau = T - t =$ is the time left to expiry at time t. When $t = 0 : c(S_\tau, T) = c(S_t, 0) = SN(d_1) + e^{-rT}KN(d_2)$, where

$$
\begin{aligned}
d_1 &= \frac{\ln(S/K)+(r+\frac{\sigma^2}{2})T}{\sigma\sqrt{T}}, \\
d_2 &= d_1 - \sigma\sqrt{T},
\end{aligned}
$$

2. Using put-call parity,

$$\begin{aligned} p(S,\tau) &= c(S,\tau) + Ke^{-r\tau} - S, \\ &= S(N(d_1) - 1) + Ke^{-r\tau}(1 - N(d_2)), \\ &= Ke^{-r\tau}N(-d_2) - SN(-d_1), \end{aligned}$$

3. If $S_0 > K$ and T is regarded as a variable then

$$\lim_{T\to 0} d_1 = \infty \quad \text{and} \quad \lim_{T\to 0} d_2 = \infty$$

so that $N(d_1) = N(d_2) = 1,\ N(-d_1) = N(-d_2) = 0$, which in turn implies that $c = S - K,\ p = 0$ as expected when $S > K$ at $T = 0$. Similarly if $S_0 < K$ then $c = 0, p = K - S_0$ so that $c = 0, p = K - S$ as expected when $S < K$ at $T = 0$.

4. If $S \to \infty$ then c is a.s. going to exercised and $c \approx S - Ke^{-r\tau}$.

5. If $\sigma \to 0$ then $c \approx S - Ke^{-r\tau}$

Exercises 5.4.1.

1. *Plot a surface for numerical values of $V(S,t)$. To do so, pick numerical values for K, r, σ, T, represent $V(s,t)$ as a function of S and t and plot in $3D$ space, for example $r = 0.65,\ K = 100,\ \sigma = 0.8 = 80\%$ (higher σ values leads to a better picture), $T = 1,\ t \in [0,1]$ and $S \in [50, 150]$. It is important to note that the surface obtained does not change; the option value traces out a path on the surface as the value of S traces out its sample path from $t = 0$ to $t = 1$*

2. *Solutions for Binary options: A binary call option is defined by*

$$\begin{aligned} \textit{payoff} &= \begin{cases} 1 & \textit{if } S_T > K \\ 0 & \textit{if } S_T < K \end{cases} \\ &= H(S,K) \end{aligned}$$

Show that

   (a) *the value of a European binary call is $c_B(S,\tau) = e^{-r\tau}N(d_2)$*
   (b) *the value of a European binary put is $p_B(S,\tau) = e^{-r\tau}(1 - N(d_2))$*

Chapter 6

First extensions of the Black-Scholes model

6.1 The model for dividends and applications

6.2 Greeks

6.3 Compound Options

These derivatives are *options on options.* There are 4 main types:

call on call, call on put, put on call and put on put.

NB: For compound options there are *2 strike prices and 2 expiry dates.*

6.3.1 The value of a European call on European call

We consider the case when there are no dividends. For the cases incorporating dividends, the solution can be modified in the obvious ways.

Let T_1 and T_2 be times such that $0 < T_1 < T_2$. There is one underlying (risky) security, denoted S and We use the following notation:

S_t	value of stock at time t
$\tilde{c}(S_{T_1}, T_1)$	value of European call on S with expiry T_2 and strike X_2
$c(S_{T_1}, T_1)$	value of European call on $\tilde{c}$ with expiry T_1 and strike X_1

Remarks

- The holder of the call c will have the right to buy the call option $\tilde{c}$ at time T_1 for the price X_1.
- The option $\tilde{c}$ is an option to purchase the underlying S at time T_2 for price X_2.
- The compound option will be exercised if $\tilde{c}(S_{T_1}, T_1) > X_1$.
- Compound options can be applied to price American options

The Black-Scholes PDE holds for any derivative security modelled under the set assumptions. In particular, it should hold for $c(S,t)$ for $T < T_1$.

For a call on call option $c(S,0)$, the underlying security is $\tilde{c}$ and

$$\begin{aligned} c(\xi,0) &= \max(\tilde{c} - X_1, 0) \\ &= \max(S_{T_1} N(d_1) - X_2 e^{-r(T_2-T_1)} N(d_2) - X_1,\ 0). \end{aligned}$$

If X is the value of S_t for which $\tilde{c} = X_1$ and $\tau = T_1 - t$, it follows from (5.16) and (5.17) that :

$$\begin{aligned} c(S,t) &= e^{-r(T_1-t)} \int_{-\infty}^{\infty} G(S,\tau;\xi,0) c(\xi,0) d\xi, \\ &= e^{-r(T_1-t)} \int_{X}^{\infty} G(S,\tau;\xi,0) \left(S_{T_1} N(d_1) - X_2 e^{-r(T_2-T_1)} N(d_2) - X_1 \right) d\xi, \end{aligned}$$

where

$$\begin{aligned} d_1 &= \frac{\ln(S_{T_1}/X_2) + (r + \frac{\sigma^2}{2})(T_2 - T_1)}{\sigma\sqrt{T_2 - T_1}} \\ d_2 &= d_1 - \sigma\sqrt{T_2 - T_1}. \end{aligned}$$

is determined for the valuation of $\tilde{c}$ at time T_1.

After *lengthy* calculation[1], the value of a call on call can be found to be

$$c(S,t) = SN_2(a_1, b_1; \rho) - X_2 e^{-r(T_2-t)} N_2(a_2, b_2; \rho) - X_1 e^{-r(T_1-t)} N(a_2),$$

where $N_2(x, y, \rho)$ is the standard bivariate normal distribution with correlation coefficient ρ,

$$\begin{aligned} a_1 &= \frac{\ln(S/X) + (r + \frac{\sigma^2}{2})(T_1 - t)}{\sigma\sqrt{T_1 - t}} \\ a_2 &= a_1 - \sigma\sqrt{T_1 - t}, \\ b_1 &= \frac{\ln(S/X_2) + (r + \frac{\sigma^2}{2})(T_2 - t)}{\sigma\sqrt{T_2 - t}} \\ b_2 &= b_1 - \sigma\sqrt{T_2 - t}, \quad \text{and} \\ \rho &= \frac{\text{cov}(Z(T_2) - Z(t), Z(T_1) - Z(t))}{\sqrt{\text{var}(Z(T_2) - Z(t))\text{var}(Z(T_1) - Z(t))}} = \sqrt{\frac{T_1 - t}{T_2 - t}}. \end{aligned}$$

Here ρ is the correlation coefficient between the overlapping Brownian increments of amount $T_1 - t$ and $T_2 - t$. $N_2(a_2, b_2; \rho)$ may be interpreted as the probability that $S_{T_1} > X$ at time T_1 and $S_{T_2} > X_2$ at time T_2, given an asset price of S at time t.

[1] *Mathematical Models of Financial Derivatives* (Springer Financial Mathematics) by Y.-K. Kwok

Chapter 7

Valuing American options

Throughout this chapter C_A and P_A will denote the values of American call and put options, respectively and C_E and P_E will denote their European counterparts. V_A and V_E are used to denote generic American and European options, respectively. Both t and $\tau = T - t$ are used to denote the time parameter, where T is the option's expiry date.

7.1 Bounds and put-call parity

7.1.1 Upper bounds

1. A call option can never cost more than the underlying stock price:

$$\begin{aligned} C_A &\leq S_t, \\ C_E &\leq S_t. \end{aligned}$$

2. Similarly, a put option can never be worth more than its strike price:

$$\begin{aligned} P_A &\leq K, \\ P_E &\leq K. \end{aligned}$$

   Furthermore, for European options $P_E \leq e^{-r\tau} K$ (= present value of strike price).

7.1.2 Lower bounds on European option values

Lower bounds for European options on non-dividend paying stock

1. $C_E \ \geq \ S_t - Ke^{-r\tau}, \quad \tau = T - t$

   ARBITRAGE ARGUMENT : Suppose $C_E(t) + Ke^{-r\tau} < S_t$. Then, short sell S_t, buy the call, bank $S_t - C_E(t) > Ke^{-r\tau}$. Equivalently, $S_t - C_E(t) = Ke^{-r\tau} + Z$. At expiry exercise call and return stock (close-out position) to make a riskless profit $Ze^{r\tau}$.

2. $P_E \; \geq Ke^{-r\tau}S_t$.

PORTFOLIO ARGUMENT : Consider,

$$\begin{array}{lcll} \text{portfolio A} & \equiv & P_E + S_t & \text{(put + share)} \\ \text{portfolio B} & \equiv & Ke^{-r\tau} & \text{(cash amount).} \end{array}$$

At expiry portfolio B is worth K and portfolio A is worth $\max(S_T, K)$ since

$$\begin{array}{lcl} S_T \leq K & \Rightarrow & \text{put is exercised and portfolio A} \equiv \text{ K} \\ S_T > K & \Rightarrow & \text{put is not exercised and portfolio A} \equiv S_t. \end{array}$$

Thus, portfolio A $\geq$ portfolio B, i.e. $P_E - S_T \geq Ke^{-r\tau}$.

Lower bounds for European options on stock paying Discrete Dividends

Deduct the riskless component D, the present value of known dividends, from the stock price and value the option on the risky part of stock price $\tilde{S} = S - D$, treating this as non-dividend paying stock. Then

$$\begin{array}{lcl} C_E & \geq & (S_t - D) - Ke^{-r\tau} \\ P_E & \geq & Ke^{-r\tau} - (S_t - D). \end{array}$$

Lower bounds for European options on stock paying continuous dividend yield

Adjust stock price model, $S_t \rightsquigarrow S_te^{-q\tau}$, and treat the latter price as a non-dividend paying stock. Then

$$\begin{array}{lcl} C_E & \geq & S_te^{-q\tau} - ke^{-r\tau}, \\ P_E & \geq & Ke^{-r\tau} - S_te^{-q\tau}, \end{array}$$

Put-call parity for European options becomes $C_E + Ke^{-r\tau} = P_E + S_te^{-r\tau}$.

7.1.3 Lower bounds and early exercise for American options on non-dividend paying stock

American options have early exercise privilege. Hence, their cost must incorporate some extra cost or an *early exercise premium*, i.e. $V_A \geq V_E$. Four key properties characterise American options. The first one is:

$$V_A \;\geq\; \underbrace{\text{payoff if exercised}}_{=\text{intrinsic value}}$$

The *time value* of V_A is the value of option arising from time left to expiry and is given by

$$\text{time value} = V_A - \text{ intrinsic value} .$$

1. $C_A \geq C_E \geq S_t - Ke^{-r\tau} \geq S_t - K$

   It is never optimal to exercise a call option C_A early. If stock is not going to be held and is in-the-money, then trade the option.

2. $P_A \geq P_E \geq Ke^{-r\tau} - S_t$. It may be the case that $K - S_t \geq P_E \geq Ke^{-r\tau}$ (depending on the stock value), but it is always the case that $P_A \geq K - St$!

   It *may* be optimal to exercise P_A early. If P_A is in-the-m then sell stock and collect time value of K.

7.1.4 Put-call symmetry for American options

We review the case for non-dividend paying stock; the dividend paying case follows analogously. Consider

$$\begin{aligned} \text{portfolio A} &\equiv C_E + K, \\ \text{portfolio B} &\equiv P_A + S_0, \end{aligned}$$

It follows that

TIME t	S_t	PORTFOLIO VALUE	
$0 < t < T$	$S_t < K$	$A = C_E + Ke^{rt}$ $B = K$	$A \geq B$
	$S_t > K$	$A = C_E + Ke^{rt}$ $B = P_A + S_t$	C_E is in-the money P_A is out-the money
at expiry $t = T$	$S_T < K$	$A = Ke^{rT}$ $B = K$	$A \geq B$
	$S_T > K$	$A = (Ke^{rT} - K) + S_T$ $B = S_T$	$A \geq B$

Thus,

$$P_A + S_0 \leq C_E + K \implies S_0 - K \leq C_A - P_A.$$

Now consider

$$\begin{aligned} \text{portfolio A} &\equiv C_A + Ke^{rT}, \\ \text{portfolio B} &\equiv P_E + S_0 \end{aligned}$$

It follows that

TIME t	S_t	PORTFOLIO VALUE	
$0 < t < T$	$S_t < K$	$A = Ke^{-r\tau} + C_A$ $B = P_E + S_t$	C_A is out-the money P_E is in-the money
	$S_t > K$	$A = Ke^{-r\tau} + C_A$ $B = P_E + S_T$	C_A is still not exercised P_E is out-the money
at time $t = T$	$S_T < K$	$A = K$ $B = K$	$C_A = 0$ P_E is exercised
	$S_T > K$	$A = S_T$ $B = S_T$	C_A is exercised $P_E = 0$

Thus,

$$C_A + Ke^{-rT} \;\leq\; P_E + S_0 \;\leq\; P_A + S_0.$$

7.2 American options on discrete dividend paying stock

Let D denote present value of discrete dividend and

$$\begin{aligned} t_d &= \text{dividend date },\\ t_{d-} &= \text{time just before payment},\\ t_{d+} &= \text{time just after payment }. \end{aligned}$$

Assume $r > 0$. The lower bounds for C_A and C_E are

$$\begin{aligned} C_E &\geq S_t - D - Ke^{-r(T-t)}\\ C_A &\geq S_t - K \end{aligned}$$

The decision to exercise early depends on whether $Ke^{-rt}D$ lies on left or right of K.

Away from t_d, there is no dividend payment and C_A behaves like C_E. The only time that C_A may be worth exercising is just before t_d.

If D is sufficiently large then $S_{t_d} - D - Ke^{-r(T-t_d)} \leq S_t - K$. If the inequality holds, then $D \geq K(1 - e^{-r(T-t_d)})$ (which depends on t_d). In this case, C_A should be exercised early if the value of C_E drops below the intrinsic value $S_{t_d} - K$ just before t_d (such an event is possible since C_E is bounded below by $S_{t_d} - D - Ke^{-r(T-t)}$ while $C_A \geq S_{t_d} - K$). Since the value of C_E will drop below $S_{t_d} - K$ if S_{t_d} is high enough we need to determine what is *high enough*?

Now C_A is continuous across dividend date (exercise) and C_A can be valued as C_E away from the dividend date. Thus,

$$\begin{aligned} C_A(S_{t_{d-}}, t_{d-}) &\approx C_A(S_{t_{d+}}, t_{d+}) \qquad \text{with } \epsilon\text{−difference}\\ &= C_E(S_{t_{d+}}, t_{d+})\\ &= C_E(S_{t_{d-}} - D, t_{d-})\\ &= C_E(\tilde{S}_{t_{d-}}, t_{d-}), \end{aligned}$$

where

$$\tilde{S} = \begin{cases} S - De^{-r(t_d - t)} & \text{for } t < t_d\\ S & \text{for } t > t_d \end{cases}$$

is the risky part of the stock value and $C_E(\tilde{S}_{t_{d-}}, t_{d-})$ has time to expiry $T - t_{d-}$.

But $C_A(S_{t_{d-}}, t_{d-}) \approx C_A(S_{t_{d+}}, t_{d+}) \geq S_{t_d} - K$. Thus, S_{t_d} is high enough when

$$C_E(\tilde{S}_{t_{d-}}, t_{d-}) \;=\; S_{t_d} - K. \qquad (7.1)$$

Letting $S^*_{t_d}$ denote the value of S which satisfies equation (7.1), it follows that the value of C_A at t_d is given by

$$C_A = \begin{cases} S_{t_d} - K & \text{if } S_{t_{d-}} \geq S^*_{t_d} \\ C_E(\tilde{S}_{t_{d-}}, t_{d-}) & \text{if } S_{t_{d-}} < S^*_{t_d} \end{cases}$$

This can valued by means of a compound option which is exercised at time t_d with payoff:

$$\text{payoff at } t_d = \begin{cases} S_{t_d} - K & \text{if } S > S^* \text{ (and the call is exercised)} \\ C_E(\tilde{S}_{t_{d-}}, t_{d-}) & \text{if } S \leq S^* \text{ (and the call is held to expiry).} \end{cases}$$

Only the risky part of S is used in the valuation of the option expiring at t_d. This gives us

$$C_A(S, 0) \;=\; e^{-rt_d} \left(\int_{S^*}^{\infty} (\tilde{S}_{t_d} + D - K) G \, d\tilde{S}_{t_d} + \int_0^{S_*} C_E(\tilde{S}_{t_{d-}}, t_{d-}) G \, d\tilde{S}_{t_d} \right).$$

where $\tilde{S}$ is the risky part of the stock price on the dividend date and $\tilde{S}+D-K \;=\; S-K$ is the payoff if call is exercised as t_d and G ... Once S^* has been determined (by Newtons' method, for example), this expression evaluated by formulae for compound options or numerically.

The argument generalises to the case where there are a finite number of discrete dividends. The analogue for put options is much more complicated. In that case, if the put value drops below its intrinsic value, it may be worthwhile to hold onto the stock till after the dividend is paid and the option behaves like a European option for a brief while prior to the dividend date[1].

[1]For some more qualitative discussion on this case, see K.-Y. Kwok.

7.3 The perpetual American put

A perpetual American put option

- ∼ can be exercised at any time, i.e. $\nexists$ expiry date,
- ∼ is time independent and is only a function of the underlying asset
- ∼ has payoff $P_p = \max(K - S, 0)$ if exercised and hence, $P_p \geq \max(K - S, 0)$.

For this problem the Black-Scholes equation becomes:

$$\frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} \quad = \quad rV. \tag{7.2}$$

Equation (7.2) has a general solution $V(S) = AS + BS^{\frac{-2r}{\sigma^2}}$, where A, B are constant. Now option value $V(S) \to 0$ as $S \to \infty$ and hence, $A = 0$.

Next, the put will not be exercised when $S > K$ and its insurance value becomes meaningless if $S \ll K$, (for example if $S = 0$). In the latter case, the put may exercised to collect the time-value of the payoff - waiting for S to drop even lower would diminish the *effective* value of the payoff, which includes time value, if the put is exercised later.

Lastly, the *absence* of a boundary constraint $S \geq S_*$ for the PDE causes the value of solutions to decrease, in which case $P_p = V(s)$ may drop below the intrinsic value.

Let S^* denote the optimal exercise boundary (OEB). Gere S^* is not function of time. Then $V(S^*) = K - S$. Since $V(S^*) = B(S^*)^{\frac{-2r}{\sigma^2}}$, we have $B = (K - S^*)(S^*)^{\frac{2r}{\sigma^2}}$.

Now the OEB should give a value for which $V(S^*)$ is maximal when V is considered as a function of S^*. Thus,

$$\begin{aligned} \frac{\partial}{\partial S^*}V(S, S^*) &= \frac{\partial}{\partial S^*}(K - S^*)(S^*)^{\frac{2r}{\sigma^2}}.(S)^{\frac{-2r}{\sigma^2}} \\ &= 0 \quad \text{when } S^* = \frac{K}{1 + \sigma^2/2r}. \end{aligned}$$

For this S^* we have $B = \frac{\sigma^2}{2}(\frac{K}{1+\sigma^2/2r})^{1+2r/\sigma^2}$ and hence,

$$V(S) \;=\; \frac{\sigma^2}{2r}(\frac{K}{1 + \sigma^2/2r})^{1+2r/\sigma^2}.s^{-2r/\sigma^2}.$$

It is easy to verify that

$$\frac{\partial V}{\partial S}\,|_{S=S^*} \;=\; -1 \;=\; \text{slope of payoff function.}$$

This is the *smooth-pasting condition.*

7.4 The optimal exercise boundary & constraints for V_A

7.4.1 Introduction

We have already seen that the option price V_A cannot drop below its intrinsic value. The price must also satisfy the Black-Scholes PDE[2] and is determined by the boundary conditions which define the domain where the PDE holds. For American options, this boundary is not known in advance.

Consider the American put option. Since $P_A(S,t) \geq K - S$ and for some (S_t, t) we have $P_E(S,t) < K - S$, it follows the domain for the European put is not the same as that of the American put. The boundary conditions for pricing vanilla European options are $0 \leq S < \infty$, $0 \leq t \leq T$, and no restrictions other than final payoffs determine how the option value should behave.

Since boundary conditions for a PDE influence the values of the solutions, there must be some critical value S_t^* at each time $t \in [0,T]$ at which the option should be exercised. As in the case of perpetual options, S_t^* is referred to as the **optimal exercise boundary** (OEB). In this case S_t^* is a function of t.

One could immediately try to guess at the value of S_t^*. For example, boundary conditions for American options could simply be

$$V_A \geq \max(V_E,\ \text{intrinsic value of } V_A)$$

[2]assuming the condition of the Black-Scholes model

It turns out that the behaviour of the option value determines more constraints for the problem. In particular, we have that

1. $V_A \geq$ intrinsic value
2. V_A must be continuous as a function of S; $\frac{\partial V_A}{\partial S}$ must be continuous as a function of S (continuity of delta)
3. $\frac{\partial V}{\partial S} |_{S_t = S_t^*} = \frac{\partial}{\partial S}$ [payoff function] (smooth pasting or high contact condition)
4. $\frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV \leq 0$ (Black-Scholes inequality)

We discuss items (2) - (4) in more detail in the following sections. Clearly these constraints rule out our first naive guess for the boundary S_t^*. Moreover, the boundary conditions cannot be determined prior to solving the PDE. Instead S_t^* *is found as part of its solution* as a **free boundary problem**[3].

We do know that the value of S_t^* depends on t. For the case of put options S_t^* is a increasing function of t and for call options it is an decreasing function of t.

Since the option value $V_A(S,t)$ is determined by the exercise boundary S_t^*, when the option is in-the-money, there is a tradeoff between its insurance value and the time-value of the payoff if the option is exercised. The insurance value may be greater than the value of the payoff if the option is exercised. In this case the option is still within the domain where it satisfies the Black-Scholes PDE and should not be exercised.

7.4.2 Continuity of delta and the smooth-pasting condition

We consider the case for an American put option.

Arbitrage argument

Let $P_A(S,t)$ denote an American put option and let S_t^* denote its exercise boundary. The slope of the payoff function $\max(K-S,0)$, is -1 or 0. In particular, its slope at S_t^* must be $\frac{\partial}{\partial S}(K-S) = -1$. To see this, note that the slope of P_A at $S = S_t^*$ could be

$$\frac{\partial}{\partial S} P(S^*, t) \begin{cases} < -1 \\ > -1 \\ = 1 \end{cases}$$

[3]See ... for the melting ice problem as the free boundary analogue of the standard heat equation problem.

Suppose $\frac{\partial}{\partial S} P \mid_{S_t = S_t^*} < -1$. As S increases across S_t^* we have $P_A(S,t)$ dropping below its intrinsic value which contradicts the lower bound assumption.

If $\frac{\partial}{\partial S} P \mid_{S_t = S_t^*} > -1$, then the value of P near S_t^* is increased by choosing a smaller S_t^*. Thus if $\frac{\partial}{\partial S} P \mid_{S_t = S_t^*} > -1$ then S_t^* is too high. [4] Thus, we are left with

$$\frac{\partial}{\partial S} P \mid_{S_t = S_t^*} = -1 = \frac{\partial}{\partial S}[\text{payoff}].$$

The equality $\frac{\partial}{\partial S} P \mid_{S_t = S_t^*} = \frac{\partial}{\partial S}[\text{payoff}]$ is referred to as the **smooth pasting** or **high contact condition**. This property is nicely illustrated by the valuation of the perpetual American put.

Mathematical argument

Let $f(S,t;b(t))$ denote the solution to the Black-Scholes PDE over domain $\Omega = \{(S,t) \mid S \in (0,b(t)),\ t \in (0,T]\}$, where $b(t)$ is some known boundary. The holder of an American put option[5] will choose an early exercise policy which maximises the value of the contract:

$$P(S,t) = \max_{b(t)} f(S,t;b(t)).$$

For fixed t we can write $F(S,b) := f(S,t;b(t))$ and we may assume that F is differentiable with respect to b for $0 \leq S \leq b$. Along the boundary (when $S = b$) we can write $h(b) = F(b,b)$ so that $h(b) = K - b$ for the case of an American put. Thus, along the boundary we have

$$\frac{dF}{db} = \frac{dh}{db} = \frac{\partial F}{\partial S}(S,b)|_{S=b} + \frac{\partial F}{\partial b}(S,b)|_{S=b}.$$

Letting b^* be the value of b which maximises F, it follows that $\frac{\partial F}{\partial b}(S,b^*) = 0$ and

$$\frac{dh}{db}|_{b=b^*} = \frac{d}{db}(K-b)|_{b=b^*} = -1.$$

Thus,

$$\begin{aligned} \frac{\partial F}{\partial S}|_{S=b^*} &= -1, \quad \text{i.e.} \\ \frac{\partial P}{\partial S}(S^*(t),t) &= -1, \end{aligned}$$

[4] We have not specified S_t^*, we merely know that it exists and find that by varying possible values for S_t^*, the option value increases while still ensuring that the boundary constraints are satisfied.

[5] The argument for a call option is the case presented in [Kwok 4.1.2].

where $S^*(t) = b^*(t)$.

Similarly, it can be shown that

$$\frac{\partial C}{\partial S}(S^*(t), t) \quad = \quad 1.$$

7.4.3 The Black-Scholes inequality

Let $\Pi = V - \Delta S$. Then

$$\begin{aligned} d\Pi - r\Pi\, dt &\equiv \text{excess return on } \Pi \text{ over the risk free rate } r \\ &= \left(\frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV\right) dt \\ &\quad \text{when } \Delta = \frac{\partial V}{\partial S} \\ &= 0 \end{aligned}$$

when the return on the portfolio is the riskfree rate.

For American options, V_A may be exercised sub-optimally (for example if the put is held and not exercised at $S = S_t^*$). Now the holder cannot achieve more than the riskfree rate (if the option was priced correctly), irrespective of his/her exercise strategy. However, the writer of the contract (who has no control over how the contract is exercised) may achieve more than the riskfree rate if the option is *not* exercised optimally. Thus, for the value of an American option we have

$$\frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV \quad \leq \quad 0. \tag{7.3}$$

It follows that we have 2 regions for the value of an American put option:

(a)	$P_A > K - S$	Black-Scholes PDE holds	$S > S_t^*$	option is *alive*
(b)	$P_A \leq K - S$	Black-Scholes inequality holds	$S \leq S_t^*$	option is *dead.*

Exercises 7.4.1. *Substitute the value $P_A = K - S$ into the LHS of (7.3) and determine whether inequality or equality holds.*

7.5 Integral solutions for American options on stock paying continuous dividends

7.5.1 Further properties of the optimal exercise boundary

$C(S,\tau)$ is an increasing function of τ ($C(S,t)$ is an decreasing function of t), i.e. $\frac{\partial C}{\partial \tau} > 0$

Asymptotic behaviour of the optimal exercise boundary close to expiry

In this section we are concerned only with the behaviour of the optimal exercise boundary at and just before expiry. We consider the case for a call option. The case the put option follows similarly and is left as an exercise. The argument which follows uses the Black-Scholes PDE for the continuous dividends case to obtain an expression for $\frac{\partial V}{\partial \tau}$ together with the fact that $\frac{\partial V}{\partial \tau}$ must be positive for an option which is held optimally to expiry.

Recall that the PDE for case of continuous dividends is

$$\frac{\partial V}{\partial \tau} = \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2} + (r-q)S\frac{\partial V}{\partial S} - rV. \tag{7.4}$$

Let $\tau = 0$. When the call option is in-the-money ($S > K$), the option value is $C(S_\tau, 0) = S - K$. If we substitute this value into (7.4) we obtain

$$\frac{\partial C}{\partial \tau}|_{\tau=0} \;=\; \frac{\partial C}{\partial \tau}(S,0) = (r-q)S - r(S-K) = rK - qS, \qquad S > K.$$

Now $\frac{\partial C}{\partial \tau}(S,0) > 0$ for the call to be kept alive close to expiry[6]. Treating $\frac{\partial C}{\partial \tau}(S,0)$ as a function of S, it follows that $\frac{\partial C}{\partial \tau}(S,0) = rK - qS = 0$ when $S = \frac{r}{q}K$. Clearly $rK - qS > 0$ when $S < \frac{r}{q}K$.

Assume $r > q$. Since our discussion concerns the case when the call is in-the-money, the call is kept alive when $K < S < \frac{r}{q}K$ (which is only possible if $r > q$). This is consistent with our intuition that the dividend $qS\triangle t$ earned in the short time $\triangle t$ before expiry is less than the interest $rK\triangle t$ which earned by holding the amount K in the bank and not exercising the option. On the other hand, if $K < \frac{r}{q}K < S$, then the option should be exercised to ensure $\frac{\partial C}{\partial \tau}(S,0) > 0$. Thus, when $r > q$

$$\lim_{\tau\to 0^+} S^*(\tau) = \frac{r}{q}K.$$

[6]otherwise $C(S,\tau)$ must be a decreasing function of τ and will have to drop below $C(S_\tau,0) = S - K$ just before expiry.

This is consistent with our arguments for the case when $q = 0$: in this case $\lim_{\tau \to 0^+} S^*(\tau) = \infty$. Together with the property that $C(S, \tau)$ is an increasing function of τ, the call should never be exercised before expiry!

Assume $r < q$ and hence, $\frac{r}{q}K < K$. If $S > \frac{r}{q}K$ then the contract should be exercised and not held to expiry. On the other hand $rK - qS \geq 0$ when $S \leq \frac{r}{q}K$ and the call will not be exercised. Thus, for call held optimally to expiry, it follows that $S^*(0) \leq K$. Clearly $S^*(0) > K$ for $\tau > 0$. Thus, when $r < q$

$$\lim_{\tau \to 0^+} S^*(\tau) = K.$$

Further analysis gives approximations of $S^*(\tau)$ close to expiry when $r > q$ (see [WHD], Section 6.5). In particular, it has been shown that

$$S^*(\tau) \approx \frac{r}{q}K\left(1 + \xi_0\sqrt{\frac{1}{2}\sigma^2\tau} + ...\right),$$

where $\xi_0 \approx 0.9034$, as $\tau \to 0$.

Exercises 7.5.1. *Show that for an American put option*

$$\lim_{\tau \to 0^+} S^*(\tau) = \begin{cases} K & \textit{when } r > q \\ \frac{r}{q}K. & \textit{when } r < q. \end{cases}$$

7.5.2 Non-homogeneous PDE for American options

We give an argument for an American put options. The discussion for an American call option follows analogously.

In what follows, we consider the value of an American put option at any arbitrary time $\tau \in [0, T]$. We have established that when $S > S^*(\tau)$, the put value must satisfy the Black-Scholes PDE:

$$\frac{\partial V}{\partial \tau} - \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2} - (r-q)S\frac{\partial V}{\partial S} + rV = 0. \tag{7.5}$$

i.e.

I	when $S > S^*(\tau)$	option is *alive*	P satisfies (7.5)
II	when $S \leq S^*(\tau)$	option is *dead*	$P = K - S$

For case II, evaluate the LHS of (7.5) when $P = K - S$:

$$\begin{aligned} & \frac{\partial P}{\partial \tau} - \frac{\sigma^2 S^2}{2}\frac{\partial^2 P}{\partial S^2} - (r-q)S\frac{\partial Vp}{\partial S} + rP \\ = \; & 0 - 0 - (r-q)S(-1) + r(K - S) \\ = \; & rK - qS. \end{aligned}$$

This leads to

$$\frac{\partial P}{\partial \tau} - \frac{\sigma^2 S^2}{2}\frac{\partial^2 P}{\partial S^2} - (r-q)S\frac{\partial P}{\partial S} + rP = \begin{cases} 0 & \text{if } S > S^*(\tau) \\ rK - qS & \text{if } S \leq S^*(\tau) \end{cases} \tag{7.6}$$

Letting $F(S,\tau)$ denote the RHS of (7.6), it follows that the value of an American put option on dividend paying stock satisfies the non-homogenous PDE

$$\frac{\partial P}{\partial \tau} - \frac{\sigma^2 S^2}{2}\frac{\partial^2 P}{\partial S^2} - (r-q)S\frac{\partial P}{\partial S} + rP = F(S,\tau)$$

with initial condition (in τ variable) $P(S_T, 0) = K - S_T$. Its solution is therefore,

$$P(S,\tau) = \int_0^\infty P(S_T,0)G(S,\tau;S_T,0)dS_T + \int_0^\tau \int_0^\infty F(S_\xi, \tau-\xi)G(S,\tau;S_\xi,\xi)dS_\xi\, d\xi, \tag{7.7}$$

where[7]

$$G(S,\tau;S_\xi,\xi) = \frac{e^{-r\xi}}{S_\xi\sqrt{2\pi\xi\sigma^2}} \exp\left(-\frac{\ln(S/S_\xi) + (r-q-\frac{\sigma^2}{2})\xi}{2\sigma^2\xi}\right),$$

Exercises 7.5.2. *Derive a non-homogeneous PDE for an American call option on a stock paying a continuous dividend yield and give an integral solution for its value.*

[7]Here we are absorbing the discounting factor $e^{-r\xi}$ into the Green's function. In keeping with earlier formulation we could have written

$$P(S,\tau) = e^{-r\tau}\int_0^\infty P(S_T,0)G(S,\tau;S_T,0)dS_T + \int_0^\tau e^{-r\xi}\int_0^\infty F(S_\xi,\tau-\xi)G(S,\tau;S_\xi,\xi)dS_\xi\, d\xi,$$

where

$$G(S,\tau;S_\xi,\xi) = \frac{1}{S_\xi\sqrt{2\pi\tau\sigma^2}} \exp\left(-\frac{\ln(S/S_\xi) + (r-q-\frac{\sigma^2}{2})\tau}{2\sigma^2\tau}\right).$$

Either convention is acceptable.

7.5.3 Integral solution for the option value

From equation (7.7) we have

$$\begin{aligned} P(S,\tau) &= \int_0^K (K-S_T)G(S,\tau;S_T,0)dS_T \\ &+ \int_0^{\tau}\int_0^{S*(\tau-\xi)} (rK-qS_\xi)G(S,\tau;S_\xi,\xi)dS_\xi\, d\xi. \end{aligned} \tag{7.8}$$

Mathematical constraints for the solution (7.8) to the free boundary problem are that P and $\frac{\partial P}{\partial S}$ are continuous along $S*(\tau)$. The first term in (7.8) is the price for a European put option, while the second is the early exercise premium. Clearly we have

$$\begin{aligned} P(S,\tau) &= Ke^{-r\tau}N(-d_2) - Se^{-q\tau}N(-d_1) \\ &+ \int_0^{\tau} rKe^{-r\xi}N(-d_{\xi_2}) - qSe^{-q\tau}N(-d_{\xi_1})d\xi, \end{aligned} \tag{7.9}$$

where

$$\begin{aligned} d_1 &= \frac{\ln(\frac{S}{K}) + (r-q+\frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}, \quad d_2 = d_1 - \sigma\sqrt{\tau} \\ d_{\xi_1} &= \frac{\ln(\frac{S}{S^*(\tau-\xi)}) + (r-q+\frac{\sigma^2}{2})\xi}{\sigma\sqrt{\xi}}, \quad d_{\xi_2} = d_{\xi_1} - \sigma\sqrt{\xi}. \end{aligned}$$

Now on the boundary $P(S^*(\tau),\tau) = K - S^*(\tau)$, so applying this to (7.9), we obtain an equation for $S^*(\tau)$:

$$\begin{aligned} K - S^*(\tau) &= Ke^{-r\tau}N(-\hat{d}_2) - S^*(\tau)e^{-q\tau}N(-\hat{d}_1) \\ &+ \int_0^{\tau} rKe^{-r\xi}N(-\hat{d}_{\xi_2}) - qS^*(\tau)e^{-q\tau}N(-\hat{d}_{\xi_1})d\xi, \end{aligned} \tag{7.10}$$

where

$$\begin{aligned} \hat{d}_1 &= \frac{\ln(\frac{S^*(\tau)}{K}) + (r-q+\frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}, \quad \hat{d}_2 = \hat{d}_1 - \sigma\sqrt{\tau} \\ \hat{d}_{\xi_1} &= \frac{\ln(\frac{S^*(\tau)}{S^*(\tau-\xi)}) + (r-q+\frac{\sigma^2}{2})\xi}{\sigma\sqrt{\xi}}, \quad \hat{d}_{\xi_2} = \hat{d}_{\xi_1} - \sigma\sqrt{\xi}. \end{aligned}$$

To solve for $S^*(\tau)$, the future value $S^*(\tau-\xi)$, $0 < \xi \le \tau$ is needed. Hence, the solution can be obtained by starting with knowledge of $S^*(0)$ and computing $S^*(\tau)$ by iterating backwards in time.

Alternatively, the smooth pasting condition $\frac{\partial P}{\partial S}|_{S=S^*(\tau)} = -1$ can be used together with derivatives with respect to S in (7.9):

$$\begin{aligned} 0 &= 1 + \frac{\partial P}{\partial S}|_{S=S^*(\tau)} \\ &= N(\hat{d}_1) - \int_0^\tau \left[\frac{(r-q)e^{-q\xi}}{\sigma\sqrt{2\pi\xi}} \exp(-\frac{\hat{d}_{\xi_1}^2}{2}) + qe^{-q\xi}N(-\hat{d}_{\xi_1}) \right] d\xi. \end{aligned} \tag{7.11}$$

Direct solution of this integrable is not tractable, but numerical solution definitely is [8].

7.6 Linear Complimentarity and Variational Inequality formulations

In this section we formulate the American option value problem in two closely related ways which are both completely free of reference to its optimal exercise boundary. The topics are introduced via a simple obstacle problem.

7.6.1 An obstacle problem

Consider an elastic string stretched over some smooth object whose description (position in space) are known. We assume that the string and its slope are continuous. Clearly either

the string is in contact with the object (in which case its position is known),

or

the string lies above the object and it is stretch straight. .

Let $u(x)$ and $f(x)$ denote the positions of the string and the fixed obstacle, respectively. We may assume that A and B are located at $x = 1$ and $x = -1$ respectively. Furthermore, we assume that $f(\pm 1) \leq 0$, $f(x) > 0$ for some $x \in [-1, 1]$ and that f is twice differentiable with $f'' < 0$ (this ensures that the contact region is unique). We also assume that both u and u' are continuous. Let x_1^* denote the point where the string first touches f and x_2^* the point where the string first lifts off f (passing from left to right).

[8] see Kwok for further details in the computation of (7.11)

The obstacle problem amounts to finding $u(x)$, such that

$$
\begin{array}{lll}
x = -1 & \Rightarrow & u(x) = 0 \\
x \in [-1, x_1^*] & \Rightarrow & u''(x) = 0 \\
x = x_1^* & \Rightarrow & u(x) = f(x) \text{ and } u'(x) = f'(x) \\
x \in [x_1^*, x_2^*] & \Rightarrow & u(x) = f(x) \\
x = x_2^* & \Rightarrow & u(x) = f(x) \text{ and } u'(x) = f'(x) \\
x \in [x_2^*, 1] & \Rightarrow & u''(x) = 0 \\
x = 1 & \Rightarrow & u(x) = 0
\end{array}
$$

More succinctly, the problem can be summarised as

$$
\begin{array}{ll}
\text{either} & u = f \text{ and } u'' = f'' \\
\text{or} & u > f \text{ and } u'' = 0.
\end{array}
$$

7.6.2 Linear Complementarity problems and formulation for an American put

A complimentarity problem can be summarised simply as follows: solve for some unknown parameter which determines V and W, where V and W satisfy

$$VW = 0, \quad V \geq 0, \quad B \geq 0.$$

The obstacle problem of the previous section can be formulated neatly in this way: determine u, where

$$u''(u - f) = 0, \quad u'' \geq 0, \quad u - f \geq 0,$$

subject to

$$u(-1) = u(1) = 0, \qquad u, u' \text{ are continuous.}$$

Formulation for the American put

The valuation of an American option can be transformed into linear complimentarity problem. We use the following change of variables:

$$S = Ke^x, \quad t = T - 2\tau/\sigma^2, \quad k_1 = 2r/\sigma^2,$$

$$P(S, \tau) = Ke^{-\frac{1}{2}(k_1-1)x - \frac{1}{4}(k_1+1)^2\tau} u(x, \tau).$$

Under these transformations, the Black-Scholes PDE becomes

$$u_\tau = u_{xx},$$

the payoff of a put option, $\max(K - S, 0)$, becomes

$$g(x, \tau) = e^{-\frac{1}{4}(k_1+1)^2\tau} \max(e^{-\frac{1}{2}(k_1-1)x} - e^{-\frac{1}{4}(k_1+1)x},\ 0)$$

and the initial condition (in τ variable) corresponds to the condition $u(x, 0) = g(x, 0)$. Furthermore, the constraint that the option value may not drop below its intrinsic value becomes:

$$u(x, \tau) - g(x, \tau) \geq 0.$$

We have that $\lim_{x\to\infty} u(x,\tau) = 0$ and u and u' are continuous. Lastly, the Black-Scholes inequality (which holds since the holder may not exercise the contract optimally) becomes[9]:

$$u_\tau - u_{xx} \geq 0.$$

Putting the constraints together we obtain:

$$(u_\tau - u_{xx})(u(x,\tau) - g(x,\tau)) = 0, \quad u_\tau - u_{xx} \geq 0, \quad u(x,\tau) - g(x,\tau) \geq 0, \tag{7.12}$$

subject to

$$u(x,0) = g(x,0), \quad u, u' \text{ are continuous.}$$

The problem given by (7.12) can be solved numerically by a finite difference scheme[10]. An additional constraint for numerical implementation is needed. The problem must be restricted to a finite interval and thus, the problem will be restricted to $x \in [LB, UB]$, where LB (UB) is some sufficiently small (large) bound on the value of x. Numerically, we need boundary conditions for u at these extremes and can let $u(LB,\tau) = g(LB,\tau)$ and $u(UB,\tau) = 0$ (which have obvious financial interpretations).

Before numerical implementation, we need to know that a solution to the problem exists[11]. In the next section, the problem is transformed into a variational inequality. An advantage of the next approach is that it can be proved (via techniques from *functional analysis*) that a solution to the mathematical problem does in fact exist. Moreover the next approach lends itself to a finite-element solution which is more robust for when the underlying variable is no longer assumed to be continuous (for example if the stock price has jumps).

7.6.3 Variational inequality formulation for an American put

Variational inequalities are described in terms of families of test functions. We return to the obstacle problem for its simplicity - the functions f and u are as before. Let Φ denote the set of all functions ϕ such that

- $\phi(-1) = 0 = \phi(1)$
- $\phi(x) \geq f(x)$ on $[-1, 1]$
- $\phi(x)$ is continuous
- ϕ'(x) is piecewise continuous

Clearly $u \in \Phi$ and for any $\phi \in \Phi$, $\phi - f \geq 0$. Thus, since $u'' \leq 0$, it follows that $-u''(\phi - f) \geq 0$. Hence, for $\phi \in \Phi$,

$$\int_{-1}^{1} -u''(\phi - f) \geq 0. \tag{7.13}$$

This holds, in particular, for $\phi = u$. From the complementarity condition we have more:

$$\int_{-1}^{1} -u''(u - f) = 0. \tag{7.14}$$

[9]Note that the inequality is reversed in the t variable.
[10]see Wilmott, DeWynne and Howison, *Option pricing: Mathematical models and computation.*
[11]Otherwise the output of such an implementation would be garbage even though some answer is obtained.

Thus, subtracting (7.14) from (7.13), we have

$$\int_{-1}^{1} -u''(\phi - u) \geq 0. \tag{7.15}$$

Integrating, we obtain

$$[-u'(\phi - u)]_{-1}^{+1} \ + \int_{-1}^{1} u'(\phi - u)' \geq 0.$$

Since $u(\pm 1) = \phi(\pm 1)$, it follows that

$$\int_{-1}^{1} u'(\phi - u)' \geq 0 \tag{7.16}$$

for all $\phi \in \Phi$. Hence, the variational inequality formulation for the obstacle problem amounts to finding u such that

$$\int_{-1}^{1} u'(\phi - u)' \ \geq \ 0, \quad \forall \ \phi \in \Phi.$$

This problem can be solved numerically by finite-element methods[12].

The variational inequality formulation for the American put

For the American put option, we define Φ to be the set of $\phi(x, \tau)$ such that

- $\phi(x, \tau)$ and $\phi_\tau(x, \tau)$ are continuous and $\phi_x(x, \tau)$ is piecewise continuous
- $\phi(x, \tau) \geq g(x, \tau)$ for all (x, τ).
- $\phi(x, 0) = g(x, 0)$
- $\phi(LB, \tau) = g(LB, \tau)$ and $\phi(UB, \tau) = g(UB, \tau) = 0$

As for the obstacle problem, $u(x, \tau) \in \Phi$ and $\phi(x, \tau) - g(x, \tau) \geq 0$ implies

$$(u_\tau - u_{xx})(\phi(x, \tau) - g(x, \tau)) \ \geq \ 0.$$

Thus, for $\tau \in [0, \frac{\sigma^2}{2}T]$,

$$\int_{LB}^{UB} (u_\tau - u_{xx})(\phi(x, \tau) - g(x, \tau))dx \ \geq \ 0. \tag{7.17}$$

[12] Finite Element methods are based on the same foundations as the mathematical theory of Fourier Series and orthonormal bases in Hilbert Spaces. See Wilmott, DeWynne and Howison for a quick introduction and B.D.Reddy's book for a more in depth but very readable treatment of finite element methods.

From the complementarity formulation for the put, we also have

$$\int_{LB}^{UB} (u_\tau - u_{xx})(u(x,\tau) - g(x,\tau))dx \;=\; 0. \tag{7.18}$$

Subtracting (7.18) from (7.17)

$$\int_{LB}^{UB} (u_\tau - u_{xx})(\phi(x,\tau) - u(x,\tau))dx \;\geq\; 0,$$

which must hold for all $\phi \in \Phi$.

Integrating by parts,

$$\begin{aligned} 0 \;&\leq\; \int_{LB}^{UB} [u_\tau(\phi - u) - u_{xx}(\phi - u)]dx \\ &=\; \int_{LB}^{UB} [u_\tau(\phi - u) + u_x(\phi_x - u_x)]dx - [u_x(\phi - u)]_{LB}^{UB}. \end{aligned}$$

Since $\phi(LB,\tau) = g(LB,\tau)$ and $\phi(UB,\tau) = g(UB,\tau)$ for all $\phi \in \Phi$, the problem is reduced to finding $u(x,\tau) \in \Phi$ such that

$$\int_{LB}^{UB} [u_\tau(\phi - u) + u_x(\phi_x - u_x)]dx \geq 0,$$

for all $\phi \in \Phi$ and $\tau \in [0, \frac{\sigma^2}{2}T]$. Notice that that the integral for the variational problem does not contain reference to g - the payoff constraint g is contained implicitly in the definition of Φ.

When tackling the problem numerically, it is easier to implement the variational problem in the transformed variables and then transform the solution into its option pricing counterpart as a final step.

Exercises 7.6.1. *Derive the variational inequality for the price of an American call option.*

$- - - - - - - - - fin - - - - - - - - -$

Appendix A

Glossary on Linear Algebra terms

[1]

Many definitions are expressed in terms of technical terms which are also included in the glossary. Needless to say, there must be some terms which are fundamental in the sense that they are defined without reference to others.

- **algebraic multiplicity of an eigenvalue:** The algebraic multiplicity of an eigenvalue c of a matrix A is the number of times the factor $t - c$ occurs in the characteristic polynomial of A.
- **basis for a subspace:** A basis for a subspace W is a set of vectors $\{v_1, ..., v_k\}$ in W such that:
  1. $\{v_1, ..., v_k\}$ is linearly independent; and
  2. $\{v_1, ..., v_k\}$ spans W.
- **characteristic polynomial of a matrix:** The characteristic polynomial of a $n \times n$ matrix A is the polynomial in t given by the formula $\det(A - tI)$.
- **column space of a matrix:** The column space of a matrix is the subspace spanned by the columns of the matrix considered as a set of vectors. See also: row space.
- **consistent linear system:** A system of linear equations is consistent if it has at least one solution. See also: inconsistent.
- **defective matrix:** A matrix A is defective if A has an eigenvalue whose geometric multiplicity is less than its algebraic multiplicity.
- **diagonalizable matrix:** A matrix is diagonalizable if it is *similar* to a diagonal matrix.
- **dimension of a subspace:** The dimension of a subspace W is the number of vectors in any basis of W. (If W is the subspace $\{0\}$, we say that its dimension is 0.)

[1]This glossary is attributed to Gene Herman, Grinnell University. See also http://www.math.uic.edu/math310/glossary.html. In this version, definitions for norm (length), inner product (scalar product or dot product), transpose have been added.

- **dot product:** The dot product of two vectors $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$ in $\mathbb{R}^n$ is defined $u \cdot v = u_1v_1 + \ldots + u_nv_n$. The dot product is also denoted $\langle u, v\rangle$ or (u, v).
- **echelon form of a matrix:** A matrix is in row echelon form if:
  1. all rows that consist entirely of zeros are grouped together at the bottom of the matrix; and
  2. the first (counting left to right) nonzero entry in each nonzero row appears in a column to the right of the first nonzero entry in the preceding row (if there is a preceding row).
- **eigenspace of a matrix:** The eigenspace associated with the eigenvalue c of a matrix A is the null space of $A - cI$.
- **eigenvalue of a matrix:** An eigenvalue of a square matrix A is a scalar c such that $Ax = cx$ holds for some nonzero vector x. See also: eigenvector.
- **eigenvector of a matrix:** An eigenvector of a square matrix A is a nonzero vector x such that $Ax = cx$ holds for some scalar c. See also: eigenvalue.
- **elementary matrix:** An elementary matrix is a matrix that is obtained by performing an elementary row operation on an identity matrix:
- **elementary row operations:** The elementary row operations performed on a matrix are:
  1. interchange two rows;
  2. multiply a row by a nonzero scalar
  3. add a constant multiple of one row to another.
- **equivalent linear systems:** Two systems of linear equations in n unknowns are equivalent if they have the same set of solutions.
- **geometric multiplicity of an eigenvalue:** The geometric multiplicity of an eigenvalue c of a matrix A is the dimension of the eigenspace of c.
- **homogeneous linear system:** A system of linear equations $Ax = b$ is homogeneous if $b = 0$.
- **inconsistent linear system:** A system of linear equations is inconsistent if it has no solutions. See also: consistent.
- **inner product:** See dot product
- **inverse of a matrix:** The matrix B is an inverse for the matrix A if $AB = BA = I$.
- **invertible matrix:** A matrix is invertible if it has an inverse.
- **least-squares solution of a linear system:** A least-squares solution to a system of linear equations $Ax = b$ is a vector x that minimizes the length of the vector $Ax - b$.
- **length of a vector:** The length of a vectors $v = (v_1, \ldots, v_n)$ in $\mathbb{R}^n$ is defined $||v|| = \sqrt{v_1^2 + \ldots + v_n^2}$
- **linear combination of vectors:** A vector v is a linear combination of the vectors v1, ..., vk if there exist scalars a1, ..., ak such that v = a1v1 + ... + akvk.

- **linear dependence relation for a set of vectors:** A linear dependence relation for the set of vectors $\{v_1, ..., v_k\}$ is an equation of the form $a_1v_1 + ... + a_kv_k = 0$, such that not all the scalars $\{a_1, ..., a_k\}$ are zero, i.e. at least one a_i is non-zero.
- **linearly dependent set of vectors:** The set of vectors $\{v_1, ..., v_k\}$ is linearly dependent if the equation $a_1v_1 + ... + a_kv_k = 0$ has a solution, where not all the scalars a1, ..., ak are zero (i.e., it is linearly dependent if $\{v_1, ..., v_k\}$ satisfies a linear dependence relation).
- **linearly independent set of vectors:** The set of vectors $\{v_1, ..., v_k\}$ is linearly independent if the only solution to the equation $a_1v_1 + ... + a_kv_k = 0$ is the solution where all the scalars $a_1, ..., a_k$ are zero. (i.e., if $\{v_1, ..., v_k\}$ does not satisfy any linear dependence relation).
- **linear transformation :** A linear transformation from V to W is a function T from V to W such that: $T(u+v) = T(u)+T(v)$ for all vectors $u, v \in V$; and $T(\alpha v) = \alpha T(v)$ for all vectors $v \in V$ and all scalars α.
- **nonsingular matrix:** A square matrix A is nonsingular if the only solution to the equation $Ax = 0$ is $x = 0$. See also: singular.
- **norm of a vector:** See length of vector
- **null space of a matrix:** The null space of a $m \times n$ matrix A is the set of all vectors $x \in \mathbb{R}^n$ such that $Ax = 0$.
- **null space of a linear transformation:** The null space of a linear transformation T is the set of vectors v in its domain such that $T(v) = 0$.
- **nullity of a matrix:** The nullity of a matrix is the dimension of its null space.
- **orthogonal complement of a subspace:** The orthogonal complement of a subspace S of $\mathbb{R}^n$ is the set of all vectors $v \in \mathbb{R}^n$ such that v is orthogonal to every vector in S.
- **orthogonal set of vectors:** A set of vectors in $\mathbb{R}^n$ is orthogonal if the dot product of any two of them is 0.
- **orthogonal matrix:** A matrix A is orthogonal if A is invertible and its inverse equals its transpose; i.e., $A^{-1} = A^T$.
- **orthonormal set of vectors:** A set of vectors in $\mathbb{R}^n$ is orthonormal if it is an orthogonal set and each vector has length 1.
- **range of a linear transformation:** The range of a linear transformation T is the set of all vectors $T(v)$, where v is any vector in its domain.
- **rank of a matrix:** The rank of a matrix A is the number of nonzero rows in the reduced row echelon form of A; i.e., the dimension of the row space of A.
- **rank of a linear transformation:** The rank of a linear transformation (and hence of any matrix regarded as a linear transformation) is the dimension of its range. Note: A theorem tells us that the two definitions of rank of a matrix are equivalent.
- **reduced row echelon form of a matrix:** A matrix is in reduced row echelon form if:

  1. the matrix is in row echelon form;
  2. the first nonzero entry in each nonzero row is the number 1;

3. and the first nonzero entry in each nonzero row is the only nonzero entry in its column.

- **row equivalent matrices:** Two matrices are row equivalent if one can be obtained from the other by a sequence of elementary row operations.
- **row space of a matrix:** The row space of a matrix is the subspace spanned by the rows of the matrix considered as a set of vectors. See also: column space.
- **scalar product:** See dot product
- **similar matrices:** Matrices A and B are similar if there is a square invertible matrix P such that $P^{-1}AP = B$.
- **singular matrix:** A square matrix A is singular if the equation $Ax = 0$ has a nonzero solution for x. See also: nonsingular.
- **span of a set of vectors:** The span of the set of vectors $\{v_1, ..., v_k\}$ is the subspace V consisting of all linear combinations of $v_1, ..., v_k$. One also says that the subspace V is spanned by the set of vectors $\{v_1, ..., v_k\}$ and that this set of vectors spans V.
- **subspace:** A subset W of $\mathbb{R}^n$ is a subspace of $\mathbb{R}^n$ if:

  1. the zero vector, 0 is in W
  2. $x + y$ is in W whenever x and y are in W
  3. αx is in W whenever x is in W and α is any scalar

- **symmetric matrix:** A matrix A is symmetric if it equals its transpose; i.e., $A = A^T$.
- **transpose matrix:** The transpose of a matrix A is obtained by interchanging the rows and columns of A.

Appendix B

On flow, flux, divergence, curl and Green's theorem in 2D

B.1 Basics from vector calculus (calculus of several variables)

Ingredients:

- gradient ∇f
- level curves & normals
- line integrals
- normal & tangential components
- flow & flux
- divergence & curl.

Here the 1^{st} three items are defined for *a scalar field* and the last three for elements of a *vector field*.

A **curve** c in space is a vector valued function e.g $c(t) = (x(t), y(t), z(t))$ in 3-dimensional space or $\mathbb{R}^n$. Limits and derivatives of $r(t)$ with respect to t are computed componentwise.

The **length of a curve** c as t ranges between 2 values is given by Pythagoras theorem:

$$s = \int_a^b ||r'(t)||\, dt$$

Thus, the length from $r(0)$ to $r(t)$ is

$$s = \int_0^t ||r'(t)||\, dt$$

and hence,

$$\frac{ds}{dt} = ||r'(t)||.$$

The **unit tangent vector** is defined:

$$T = \frac{r'(t)}{||r'(t)||} = \frac{dr}{dt} \Big/ \frac{ds}{dt} = \frac{dr}{ds}$$

Curvature is given by

$$K = ||\frac{dT}{ds}|| = ||\frac{dT}{dt} \Big/ \frac{ds}{dt}|| = ||\frac{T'(t)}{r'(t)}||$$

The **unit normal** is defined by:

$$N = \frac{1}{||\frac{dT}{dt}||} \frac{dT}{dt}.$$

Orthogonality between N and T follows from the fact that

$$\begin{aligned} T \cdot T &= ||T|| = 1 \qquad \textit{and} \\ \frac{d}{dt}(T \cdot T) &= T \cdot \frac{dT}{dt} + \frac{dT}{dt} \cdot T = 0, \end{aligned}$$

and hence,

$$T \cdot \frac{dT}{dt} = -\frac{dT}{dt} \cdot T = 0$$

The **gradient operator** ∇ of a vector field is a vector valued function derived from a scalar function. For example, if $f(x, y)$ is a function of x and y then

$$\nabla f = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}).$$

Suppose $c = f(x, y)$ is a level curve of $z = f(x, y)$ and $r(t) = (x(t), y(t))$ is a parameterisation of the curve then:

$$\begin{aligned} 0 = \frac{df}{dt} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} \\ &= \nabla f \cdot \frac{dr}{dt}\ . \end{aligned}$$

This implies that ∇f is perpendicular to $r'(t)$ and to T.

Similarly, if $c = f(x, y, z)$ is a level surface S of $f(x, y, z)$ and $r(t) = (x(t), y(t), z(t))$ is a curve which lies on S, then

$$f(x(t), y(t), z(t)) = c.$$

This implies that

$$\begin{aligned} 0 = \frac{df}{dt} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial t} \\ &= \nabla f \cdot r'(t). \end{aligned}$$

Thus, ∇f is perpendicular to $r'(t)$ and to T. The latter can be written $\nabla f \perp T$.

Consider again a surface S given by $f(x, y, z) = c$. Let $P(x_0, y_0, z_0)$ be a point on the surface where $\nabla f(x_0, y_0, z_0) \neq 0$. The tangent plane at P is the plane through P which is orthogonal to ∇f (and which contains vectors which are tangent to the surface S at $P(x_0, y_0 z_0)$). Thus, if $r = (x, y, z)$ and $r_0 = (x_0, y_0, z_0)$ are the co-ordinates of points in the plane, then $r - r_0$ is a vector within the plane. Thus,

$$\nabla f(x_0, y_0, z_0) \cdot (r - r_0) = 0,$$

i.e.

$$f_x(x_0, y_0, z_0)\,(x - x_0) + f_y(x_0, y_0 z_0)\,(y - y_0) + f_z(x_0, y_0, z_0)\,(z - z_0) = 0,$$

gives an equation for the tangent plane at $P(x_0, y_0 z_0)$. If (x, y, z) is close to (x_0, y_0, z_0) then this may be written

$$f_x(x_0, y_0, z_0)\triangle x_0 + f_y(x_0, y_0, z_0)\triangle y_0 + f_z(x_0, y_0, z_0)\triangle z_0 = 0.$$

Line integrals are integrals of scalar functions along a curve. For example, the integral of $f(x, y, z)$ along $r(t) = (x(t), y(t), z(t))$ is given by:

$$\int_c f(x, y, z)\, ds.$$

Since s can be parameterised as a function of t by means of $ds = ||r('t)||\, dt$, it follows that the integral of f along c can be written

$$\int_c f(x, y, z)\, ds = \int_c f(x, y, z)\, ||r'(t)||\, dt.$$

To illustrate, consider the following example:
[insert sketch]

$$\begin{aligned} c_1 &= (t, t, 0) \\ c_2 &= (1, 1, t) \\ c_3 &= (t, t, t) \end{aligned}$$

for $0 \leq t \leq 1$. Then for $f = x - 3y^2 + z$ we have

$$\begin{aligned} \int_{c_1 \cup c_2} f ds &= \int_{c_1} (t - 3t^2)\sqrt{2}\, dt + \int_{c_2} (1 - 3 + t) dt = -\frac{1}{2}(\sqrt{2} + 3), \quad and \\ \int_{c_3} f ds &= \int_0^1 (t - 3t^2 - t)\sqrt{3}\, dt = 0 \end{aligned}$$

Here $||r'(t)||$ is the length of the curve.

A **vector field** in space is a function which defines a vector at each point in its domain. For example in $\mathbb{R}^3$, $F(x, y, z) = (P(x, y, z), Q(x, y, z), R(x, y, z))$, where P, Q, and R are functions of x, y, and z. The scalar component of F in direction T is $F.T$ and the scalar component of F in direction N is $F.N$. To see this, note that for vectors $\vec{u}$ and $\vec{v}$ we have

$$\vec{u} \cdot \vec{v} = ||u||\,||v||\cos\theta$$

[insert sketch]

The scalar component of $\vec{u}$ along $\vec{v}$ is $d = a.\cos\theta = ||u||.(\frac{\vec{u}.\vec{v}}{||u||\,||v||}) = \vec{u}.(\frac{\vec{v}}{||v||})$. For $F \cdot T$ and $F \cdot N$ we use the fact that T and N are unit vectors (i.e. $||T|| = 1$ and $||N|| = 1$).

The **flow** of a vector field $F = (P, Q)$ in R^2 along a curve c, where c is described by $r(t) = (x(t), y(t))$, is given by:

$$\begin{aligned} \text{flow (or circulation)} &= \int_c F \cdot T \;\; ds \quad \text{(where } T \text{ is tangent to } c\text{)} \\ &= \int_c F \cdot dr \quad (\text{since } T = \frac{dr}{ds}) \\ &= \int_c Pdx + Qdy \quad (\text{since } \frac{dr}{ds} = (\frac{dx}{ds},\ \frac{dy}{ds})\ ). \end{aligned}$$

The **flux** of $F = (P(x,y), Q(x,y))$ across a closed curve in R^2 is given by:

$$\begin{aligned} \text{flux} &= \int_c F \cdot N \; ds \quad \text{(where } N \text{ is the outward normal to } c\text{)} \\ &= \oint Pdy - Qdx, \end{aligned}$$

where $N = (\frac{dy}{ds},\ -\frac{dx}{ds})$, since $T = \frac{dr}{ds} = (\frac{dx}{ds},\ \frac{dy}{ds})$ and $T \perp N$.

Note for flow and flux in R^2, we have used the fact that $r(t) = (x(t), y(t))$ can be reparameterised to $r(s) = (x(s), y(s))$ by means of $S = \int_0^t ||r'(t)||dt$.

The 3D analogue *flow* is natural: let c be given by $r(t) = (x(t), y(t), z(t))$ and $F = (P(x,y,z), Q(x,yz), R(x,y,z))$. Then

$$\int_c Fdr = \int_c Pdx + Qdy + Rdz.$$

In 3D it also makes sense to talk about *flux* through a surface S. But we need surface integrals for this.

The **divergence** of a field at a point is the flux density at that point. To interpret this in $\mathbb{R}^2$ let $F(x,y)i + Q(x,y)j$.
[insert sketch]

The rate at which 'fluid' leaves the rectangle across bottom is

$$F(x,y).(-j)\triangle x = -Q(x,y)\triangle x.$$

We have:
Rate of flux through top $= F(x, y+\triangle y).j\triangle x = Q(x, y+\triangle y)\triangle x$

Rate of flux through left $= F(x,y).-i\triangle y = -P(x,y)\triangle y$
Rate of flux through right $= F(x+\triangle x).i\triangle y = P(x+\triangle x)\triangle y$.
Thus, the flux out left and right

$$= (P(x+\triangle x) - P(x,y))\triangle y \approx (\frac{\partial P}{\partial x}\triangle x)\triangle y,$$

and the flux out top and bottom

$$= (Q(x, y+\triangle y) - Q(x,y))\triangle x \approx (\frac{\partial Q}{\partial y}\triangle y)\triangle x.$$

Thus,

$$\begin{aligned} & \frac{\text{total flux across (rectangle) boundary}}{\text{area (of rectangle)}} \\ = \quad & \frac{(\frac{\partial p}{\partial x} + \frac{\partial Q}{\partial y})\ \triangle x\ \triangle y}{\triangle x \triangle y} \\ = \quad & \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} \\ = \quad & \nabla.F \\ \stackrel{\text{def}}{=} \quad & \text{Divergence, denoted Div F.} \end{aligned}$$

The **circulation density** of a vector field F in $\mathbb{R}^2$ is defined by:
[**insert sketch**]

$$\begin{aligned} & \frac{\text{total flow along boundary (of rectangle)}}{\text{area (of rectangle)}} \\ = \quad & \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \\ \stackrel{\text{def}}{=} \quad & \text{curl F } (= \nabla \times F). \end{aligned}$$

• curl F can be interpreted as the rotation effect in the field where rotation at a point can be thought of as rotation about on axis through the point. [**insert sketch**]

• div F measurers the tendency of a fluid/flow to diverge from a point. **[insert sketch]**

B.2 Green's Theorem Theorem (in $\mathbb{R}^2$)

[for calculating line integrals as area and vice versa)]

Let $F(x, y) = (P(x, y), Q(x, y))$, let C be a simple piece-wise smooth closed curve and let D be the region enclosed by C together with its boundary.
Then

$$\oint Pdy - Qdx = \int\int (\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y})dxdy \tag{B.1}$$

$$\tag{B.2}$$

$$\text{(flux through } c) \qquad \text{(total divergence of } F \text{ over } D) \tag{B.3}$$

and

$$\oint Pdx + Qdy = \int\int (\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})dxdy \tag{B.4}$$

$$\tag{B.5}$$

$$\text{(flow along } c) \qquad \text{(total curl for } F \text{ over } D) \tag{B.6}$$

(B.1) is referred to as the normal form of Green's theorem and can be written

$$\oint F \cdot N ds = \int\int div\, F\, dA.$$

(B.4) is referred to as the tangential form of Green's theorem and can be written

$$\oint F \cdot T\, ds = \int\int curl\, F\, dA.$$

Green's theorem for simple closed piecewise smoothly curve is proved by direct computation of the LH and RH sides of (B.1) and (B.4).

Appendix C

Taylor series, derivative matrices, approximations and Jacobians

C.1 Taylor series for several variables

If $f(x)$ is n-times differentiable and x is close to a then:

$$f(x) \quad \approx \quad f(a) + f'(a)(x-a) + \frac{1}{2!}f''(a)(x-a)^2 + ... + \frac{1}{n!}f^{(n)}(a)(x-a)^n.$$

Suppose $f(x,y)$ has continuous partial derivatives about (a,b). Let $x = a + \alpha b,\ y = b + \alpha k$ and define $g(\alpha) := f(a + \alpha h, b + \alpha k)$. Then

$$\begin{aligned} g(0) &= f(a,b) \quad \text{and} \\ g'(\alpha) &= \frac{\partial g}{\partial x}\frac{\partial x}{\partial \alpha} + \frac{\partial g}{\partial y}\frac{\partial y}{\partial \alpha} \ = \ h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y}. \end{aligned}$$

Thus,

$$g'(0) \quad = \quad hf_x(a,b) + kf_y(a,b).$$

Similarly,

$$\begin{aligned} g''(\alpha) &= h.\frac{\partial}{\partial x}g'(\alpha) + k\frac{\partial}{\partial y}g'(\alpha) \\ &= h^2 f_{xx}(x,y) + 2hkf_{xy}(x,y) + k^2 f_{yy}(x,y), \end{aligned}$$

and

$$g'(0) \quad = \quad h^2 f_{xx}(a,b) + 2hkf_{xy}(a,b) + k^2 f_{yy}(a,b).$$

By computing the Taylor series for g about $\alpha = 0$ we find

$$g(\alpha) = g(0) + g'(0)\alpha + \frac{1}{2!}g''(0)\alpha^2...$$

and hence,

$$
\begin{aligned}
f(x,y) &= f(a,b) + (hf_x(a,b) + kf_y(a,b))\alpha \\
&\quad + \frac{1}{2!}\left(h^2 f_{xx}(a,b) + 2khf_{xy}(a,b) + k^2 f_{yy}(a,b)\right)\alpha^2 \ + \ ... \\
&= f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b) \\
&\quad + \frac{1}{2!}\left(f_{xx}(a,b)(x-a)^2 + f_{xy}(a,b)(x-a)(y-b) + f_{yy}(a,b)(y-b)^2\right) \\
&\quad + \text{ higher order terms.}
\end{aligned}
$$

C.2 Affine approximation, the derivative matrix

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear map. From linear algebra we know T can be written as a matrix.
The set $\{(1,0,0),\ (0,1,0),\ (0,0,1)\}$ is a basis for $\mathbb{R}^3$ and $(x,y,z) = x(1,0,0) + y(0,1,0) + z(0,0,1)$ We may equivalently express this in terms of column vectors.
Suppose T is a transformation such that

$$
T\begin{pmatrix}1\\0\\0\end{pmatrix} = \begin{pmatrix}a\\b\\c\end{pmatrix},\ T\begin{pmatrix}0\\1\\0\end{pmatrix} = \begin{pmatrix}d\\e\\f\end{pmatrix},\ T\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}g\\h\\i\end{pmatrix}.
$$

By linearity of T , if $T\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}u\\w\\v\end{pmatrix}$ *then*

$$
\begin{aligned}
\begin{pmatrix}u\\v\\w\end{pmatrix} &= x\,T\begin{pmatrix}1\\0\\0\end{pmatrix} + y\,T\begin{pmatrix}0\\1\\0\end{pmatrix} + z\begin{pmatrix}0\\0\\1\end{pmatrix} \\
&= x\begin{pmatrix}a\\b\\c\end{pmatrix} + y\begin{pmatrix}d\\e\\f\end{pmatrix} + z\begin{pmatrix}g\\h\\i\end{pmatrix} = \begin{pmatrix}a & d & g\\ b & e & h\\ c & f & i\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}.
\end{aligned}
$$

Thus,

$$
T \equiv \begin{pmatrix}a & d & g\\ b & e & h\\ c & f & i\end{pmatrix}.
$$

If $F : \mathbb{R}^n \Rightarrow \mathbb{R}^n$ is not linear then we cannot do this. However F may be approximated by an affine map, $A =$ *linear map + transition*, as follows :

Let

$$
F\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}f(x,y)\\g(x,y)\end{pmatrix},
$$

where f and g have Taylor expansions about (x_0, y_0). Then for (x,y) close to (x_0, y_0) we have

$$
\begin{aligned}
F\begin{pmatrix}x\\y\end{pmatrix} &= \begin{pmatrix}f(x,y)\\g(x,y)\end{pmatrix} = \begin{pmatrix}f(x_0+h, y_0+k)\\g(x_0+h, y_0+)\end{pmatrix} \\
&= \begin{pmatrix}f(x_0,y_0) + hf_x(x_0,y_0) + kf_y(x_0,y_0) + h.o.t.\\ g(x_0,y_0) + hg_x(x_0,y_0) + kg_y(x_0,y_0) + h.o.t.\end{pmatrix},
\end{aligned}
$$

i.e.,

$$F\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f(x_0, y_0) \\ g(x_0, y_0) \end{pmatrix} + \begin{pmatrix} f_x\ (x_0,\ y_0)\ f_y\ (x_0,\ y_0) \\ g_x\ (x_0,\ y_0)\ g_y\ (x_0,\ y_0) \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix}\ +\ h.o.t.$$

If h and k are small, we may ignore the h. o. t. and use

$$F\begin{pmatrix} x \\ y \end{pmatrix} \approx F\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + F'\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}\begin{pmatrix} h \\ k \end{pmatrix},$$

where

$$F'\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} f_x\ (x_0, y_0)\ f_y\ (x_0, y_0) \\ g_x\ (x_0, y_0)\ g_y\ (x_0, y_0) \end{pmatrix}$$

is called the derivative matrix and the approximation is called the affine approximation for F about $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$.

NOTE: the determinant of $F'\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ is called the Jacobian and is usually denoted by $\frac{\partial(f,g)}{\partial(x,y)}$.

C.3 The Jacobian I - geometry of a transformation

Consider a change of co-ordinates from (u, v) to (x, y). This may be written

$$\begin{pmatrix} x \\ y \end{pmatrix} = F\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} x(u, v) \\ y(u, v) \end{pmatrix},$$

where x and y are both functions of u and v.
Suppose $z = f(x, y) = f(x(u, v), y(u, v))$. Then

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u} \quad and \quad \frac{\partial z}{\partial v} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial v}.$$

We may write this

$$\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{pmatrix}\begin{pmatrix} \frac{\partial z}{\partial x} \\ \frac{\partial z}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{\partial z}{\partial u} \\ \frac{\partial z}{\partial v} \end{pmatrix} \quad or \quad F'\begin{pmatrix} z_x \\ z_y \end{pmatrix} = \begin{pmatrix} z_u \\ z_v \end{pmatrix},$$

where F' is the derivative matrix

$$F'\ := \begin{pmatrix} x_u\ y_u \\ x_v\ y_v \end{pmatrix}.$$

This gives us the *affine approximation*:

$$F\begin{pmatrix} u \\ v \end{pmatrix} \approx \left(F'\begin{pmatrix} u_0 \\ v_0 \end{pmatrix}\right)\begin{pmatrix} u - u_0 \\ v - v_0 \end{pmatrix} + \text{constant}.$$

For the change of co-ordinate to make sense, we require F to be one-one (more than one point cannot be mapped to the same image). If F is a nice (not pathological) one-one transformation then (i.e. for our purposes) the derivative matrix will be invertible too, i.e. the Jacobian $\frac{\partial(f,g)}{\partial(x,y)}$ will be non-zero. Conversely, if F can be affinely approximated on small domains $D_,$ and the derivative matrix F' is invertible on D_i (i.e the Jacobian $\frac{\partial(f,g)}{\partial(x,y)} \neq 0$) then F is invertible on $D = \cup D_i$.

Examples C.3.1.

Consider the transformation $u = x, \ v = 0$ which maps the square

$$[0,1] \times [0,1] = \{ (x,y) \mid x \in [0,1], y \in [0,1] \}$$

to the line segment

$$[0,1] \times \{0\} = \{(u,v) \,| u \in [0,1] \, v = 0\} \subset \mathbb{R}^2.$$

Then,

$$\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} x_u \ y_u \\ x_v \ y_v \end{vmatrix} = \begin{vmatrix} 1 \ 0 \\ 0 \ 0 \end{vmatrix} = 0.$$

C.4 The Jacobian II - the scaling factor in integrals

From linear algebra we have the following:

Let $A \subset \mathbb{R}^n$ be a bounded region (finite volume) and $T : \mathbb{R}^n \Rightarrow \mathbb{R}^n$ be an invertible linear map. Then the following holds[1]:

$$\text{volume of } T(A) \quad = \quad (\text{volume of } A).|detT|$$

or

$$|detT| \quad = \quad \frac{\text{volume of } T(A)}{\text{volume of } A}.$$

[1] The proof can be broken down into 4 steps:

1. If Q denotes the unit 'cube' in $\mathbb{R}^n$ (Q has sides of length 1, and volume= 1) then $vol(T(Q)) = |detT|$. This is proved by writing $T = E_1 \, E_2.....E_n$ where $E_i, 1 \le i \le n$, are elementary matrices, and the facts that

   (a) $det(AB) = detA.detB$, and

   (b) if E is elementary matrix then $vol(E(Q)) = |detE|$.

2. If C is a 'cube' of arbitrary volume in $\mathbb{R}^n$, then since the scaling of the sides of C by T will be equal to the scaling of Q we have

$$\frac{vol(T(C))}{vol(C)} = \frac{vol(T(Q))}{vol(Q)}.$$

   This implies $vol(T(C)) = vol(C).vol(T(Q))$ since $vol(Q) = 1$.

3. If A is a region in $\mathbb{R}^n$ then A can be built up from disjoint cubes C_i of arbitrarily small volume (the C_i $'s$ form a partition of A) and $vol(A) = \sum_i vol(C_i)$.

4. Since T is injective (one-one), if $\{C_i\}$ forms a partition of a set $A \subset \mathbb{R}^n$, then $\{T(C_i)\}$ forms a partition of $T(A)$. Thus,

$$\begin{aligned} vol(T(A)) &= \sum_i vol(T(C_i)) \\ &= \sum_i vol(T(Q)).vol(C_i) \\ &= vol(T(Q)) \sum_i vol(C_i) \\ &= |detT| \times vol(A). \end{aligned}$$

Now let D be a closed bounded set in $\mathbb{R}^n$ and suppose $\varphi : \mathbb{R}^n \Rightarrow \mathbb{R}^n$ is a (nonlinear) differentiable mapping that is one-one within D and such that the derivative matrix $\varphi'(x_1, ..., x_n)$ is invertible. Then, if $E = \varphi(D)$ and $f : E \to \mathbb{R}$ is integrable we have

$$\int\limits_E f(y)\, dV_{E} = \int\limits_D f(\varphi(x))\, |\varphi'(x_1 ....., x_n)|\, dV_{D}.$$

To see this, let $||C_{E}||$ denote the size of the biggest cube in a partition $\{C_{E_i}\}$ of E:

$$\begin{aligned}
\int\limits_E f(y)\, dV_{E} &= \lim_{||C_{E}|| \to 0} \sum_i f(y_i).\, vol(C_{E_i}), \qquad\qquad (y_i \in C_{E_i}) \\
&= \lim_{||C_{D}|| \to 0} \sum_i f(\varphi(x_i)).\, vol\,(\varphi(C_{D_i})) \\
&= \lim_{||C_{D}|| \to 0} \sum_i f(\varphi(x_i)).\, \frac{vol(\varphi(C_{D_i}))}{vol(C_{D_i})}\, .vol(C_{D_i}).
\end{aligned}$$

Now if $vol(C_i)$, for a cube C_i, is small then the action of φ on C_i may be approximated by an affine approximation as follows: For $x \in \mathbb{R}^n$ near $x_0 \in \mathbb{R}^n$ we have

$$\varphi(x) \approx\ \varphi(x_0) + (\varphi'(x_0))\ (x - x_0).$$

Here, $\varphi'(x_0)$ is a linear matrix and the term $\varphi(x_0)$ is equivalent to translation in $\mathbb{R}^n$ (no change in volume).

WLOG, we may consider the effect of φ on a small cube C_i which is centered at the origin. Thus, $x \in C_i$ implies x is close to $x_0 = 0$ and, $\varphi(x) \approx\ \varphi'(x)+$ constant.

It follows that $vol\,(\varphi(C_i)) \approx vol\,(\varphi'(C_i))$ and, hence,

$$\frac{vol(\varphi(C_i))}{vol(C_i)} \approx \frac{vol(\varphi'(C_i))}{vol(C_i)} = |det\varphi'(x_0)|.$$

Thus,

$$\begin{aligned}
\int\limits_E f(y)\, dV_{E} &= \lim_{||C_D|| \to 0} \sum_i f(\varphi(x_i))|det\varphi'(x_0)|\, vol(C_{D_i}) \\
&= \int\limits_D (f \circ \varphi)\ (x)\ |det\ \varphi'(x_0)|\, dV_{D}
\end{aligned}$$

NB: The conditions on φ (stated earlier) are enough to ensure that φ is invertible. From the linearisation above it is clear that φ is invertible if φ' is invertible.

Appendix D

Quick review for solving 1st and 2nd order linear ODE

The General form of a linear Ordinary Differential Equation (ODE) is

$$a_n(x)\,\frac{d^n\,u}{dx^n}+a_{n-1}(x)\,\frac{d^{n-1}\,u}{dx^{n-1}}+....+a_1(x)\,\frac{du}{dx}+a_0(x)\,u=g(x),$$

where $u=u(x)$ is an unknown function.

D.1 First and simplest methods for solving simple 1st order linear ODE

An equation of the form

$$\frac{du}{dx}=\frac{g(x)}{h(u)} \tag{D.1}$$

is separable. It can be written: $h(u)du = g(x)dx$, and can be solved by directly by integrating, i.e finding a solution for

$$\int h(u)\,du=\int g(x)dx$$

gives a solution for (D.1).

A first order linear ODE $a_1(x)\frac{du}{dx}+a_0(x)u=g(x)$ may be written as

$$\frac{du}{dx}+p(x)u=Q(x). \tag{D.2}$$

Using the integrating factor $f(x)=exp\,[\int p(x)\,dx]$, (D.2) may be written

$$f\,\frac{du}{dx}+fp(x)u=fQ(x). \tag{D.3}$$

Now (D.3) holds if and only if

$$\begin{aligned} \frac{d}{dx}(f.u) &= f\,Q(x) \\ \Leftrightarrow \qquad f.u &= \int f.Q(x)\,dx \\ \Leftrightarrow \qquad u &= \frac{1}{f}\int f\,Q(x)dx. \end{aligned}$$

[see also exact differential equations & other methods...]

D.2 On existence and uniqueness

In our first encounters with differential equations we usually assume that solutions exist. In fact many users of mathematics adopt the attitude that finding a solution demonstrate it's existence and more over, that the nature of the model ensures uniqueness.

General warning:

1. Solving a DE on computer (numerically) may be costly (in time and effort) - in such cases it is worth knowing beforehand if a solution exists.
2. Even though it may be argued that the equation describes some real problem and hence, that a solution must exist, it can be debated whether the equation concerned is the correct description of the problem.
3. There may be inconsistencies in the method and reasoning used to "find a solution" - faulty methods or unsound reasoning may fail at a critical time (even if it works other times).

Theorem D.2.1 (Uniqueness Theorem - see for example, Coddington). *Suppose the n^{th} order linear ODE*

$$a_n(x)\, y^n + ... + a_1(x)\, y'(x) + a_0(x)y(x) = f(x)$$

has continuous coefficients on an interval I. Then the solution satisfying the initial condition $y(x_0) = y_0,\ y'(x_0) = y_1, ...$ *and* $y^{n-1}\,(x_0) = y_n - 1$ *is unique for* $x \geq x_0$ *on I.*

In our treatment of series solutions we can ensure our methods are rigorous by appealing to some Hilbert space theory.

D.3 Linear Independence, the Wronskian, fundamental and general solutions

Suppose $\{\varphi_n\}$ is a set of n functions such that each $\varphi_i, 1 \leq\ i \leq n$ is at least $(n-1)$ times differentiable.

Consider the equation

$$C_1\,\varphi_1(x) + ... + C_n\,\varphi_n(x) = 0. \tag{D.4}$$

If $\{\varphi_n\}$ is a linearly independent set on the interval $[a, b]$ then for $x \in [a, b]$ equation (D.4) implies $C_1 = C_2 = ...C_n = 0$[1].

From (D.4) and differentiability of the $\varphi_i' s$ we have

$$\begin{aligned} c_1\,\varphi'(x) \quad &+ \quad ... + \ c_n\,\varphi'(x) &= \ 0 \\ c_1\,\varphi''(x) \quad &+ \quad ... + \ c_n\,\varphi''(x) &= \ 0 \\ c_1\,\varphi^{(n-1)}(x) \quad &+ \quad ... + \ c_n\,\varphi^{(n-1)}(x) &= \ 0 \end{aligned} \tag{D.5}$$

For each $x \in [a, b]$ the system of equation (D.5) yields an $n \times n$ matrix.

The determinant of the matrix is denoted $W(x)$ and is called the **Wronskian.**

Let

$$\varphi_1(x) \ = \ e^x, \ \ \varphi_2(x) = xe^x, \ \ \varphi_3(x) = e^{2x}.$$

Then

$$\begin{aligned} \varphi_1'(x) \ &= \ e^x \ = \ \varphi_1''(x), \\ \varphi_2'(x) \ &= \ xe^x + e^x, \ \ \varphi_2''(x) \ = \ xe^x + 2e^x, \quad \text{and} \\ \varphi_3'(x) \ &= \ 2e^{2x}, \qquad \varphi_3''(x) \ = \ 4e^{2x}. \end{aligned}$$

For this set, (D.5) can be written

$$\begin{pmatrix} e^x & xe^x & e^{2x} \\ e^x & xe^x + e^x & 2e^{2x} \\ e^x & xe^x + 2e^x & 4e^{2x} \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

From linear algebra we have

1. Matrix B is invertible (non-singular) if and only if det $B \neq 0$.
2. If B is not invertible (if B is singular) then the equation $Bu = 0$ has more than one solution (in fact, infinitely many solutions). On the other hand, if B is invertible then $u = 0$ is the only solution for $Bu = 0$.

[1]Suppose $C_i \neq 0$ for some $1 \leq i \leq n$. Since addition is commutative, we may assume $i = 1$. Then (D.4) can be written

$$-C_1\,\varphi_1 = C_2\,\varphi_2 + ... + C_n\,\varphi_n.$$

Since the LHS $\neq 0$ it follows the RHS $\neq 0$ and

$$\varphi_1 = -C_2/C_1\,\varphi_2 + ... + (C_n/C_1)\,\varphi_n,$$

i.e. φ_1 can be written as a linear combination of the other $\varphi_i' s$.

Now if $W(x) \neq 0$ for some $x \in [a,b]$ then equation (D.5) has the unique solution $C_1 = C_2 = ...C_n = 0$.

We have shown $W(x) \neq 0 \Rightarrow$ *set is linearly independent.* However $W(x) = 0$ does not imply a set of functions is linearly dependent. For example, the Wronskian is zero for

$$g_1\ (x) = \begin{cases} 0 & x \geq 0 \\ x^4 & x < 0 \end{cases}$$

$$g_2\ (x) = \begin{cases} x^4 & x \geq 0 \\ 0 & x < 0 \end{cases}$$

However these functions are linearly independent.

Now suppose $\varphi_1(x), ..., \varphi_n(x)$ are solutions of an n^{th} order ODE

$$a_n\, y^{(n)}\,(x) + ... + a_1 y'(x) + a_0 y(x) = 0. \tag{D.6}$$

If $W(\alpha) = 0$ for some $\alpha \in [a,b]$ then we can find $c_1^*, ..., c_n^*$ such that $c_i^* \neq 0$ for some $1 \leq i \leq n$. By the principle of superpositioning

$$f(x) = c_1^*\, \varphi_1(x) + ... + c_n^*\, \varphi_n(x) \tag{D.7}$$

is a solution for the linear ODE (D.6).
Now

$$\begin{aligned} f(\alpha) &= c_1^*\, \varphi_1(\alpha) + ... + c_n^*\, \varphi_n(\alpha) = 0 \\ f^{(n-1)}(\alpha) &= c_1^*\, \varphi_1^{(n-1)}(\alpha) + ... + c_n^*\, \varphi_n^{(n-1)}(\alpha) = 0. \end{aligned}$$

Hence, substituting (D.7) into (D.6) we get $f^{(n)}(\alpha) = 0$. However, $y(x) = 0$ is also a solution for (D.6) and $y(x)$ and $f(x)$ satisfy the same initial conditions on the interval $[a,b]$. By the uniqueness theorem $f(x) = y(x) = 0$. We have shown $W(x) = 0$ *implies that set of solutions for linear the ODE is linearly dependent.*

Theorem D.3.1. *If $\varphi_1(x), ..., \varphi_n(x)$ are solutions for a linear n^{th} order ODE on an interval $[a,b]$ then $\{\varphi_i\}_{i=1}^n$ is a linearly independent set if and only if their Wronskian $W(x) \neq 0$ for all $x \in [a,b]$.*

Theorem D.3.2. *The general n^{th} order linear homogeneous equation*

$$a_n(x) y^{(n)}(x) + ... + a_1(x) y'(x) + a_0(x) y(x) = 0 \tag{D.8}$$

has n linearly independent solutions $\varphi_1, ..., \varphi_n$ on some interval I and it's general solution on I is of the form

$$y(x) = c_1 \varphi_1(x) + ... + c_n \varphi_n(x),$$

where $c_1, ..., c_n$ are arbitrary constants.

PROOF
Show that solution $\varphi_1, ..., \varphi_m$ must be linearly dependent if $m > n$. Clearly $y(x)$ satisfies (D.8). We can show there are at most n linearly independent solutions. To see there are exactly n we turn to the differential operator viewpoint developed in the next section.

D.4 Differential operator methods for linear ODE

Differentiation is a an operation which acts on a differentiable function to produce a new function:

$$\frac{d}{dx}(f(x)) = \frac{df}{dx} = f'(x).$$

Using the notation $D = \frac{d}{dx},\ D^2 = \frac{d^2}{dx^2},\ D^n = \frac{d^n}{dx^n}$, we have $D^n f(x) = f^{(n)}(x)$. Thus, our usual n^{th} order linear homogeneous ODE

$$a_n(x)y^{(n)}(x) + ... + a_1(x)y'(x) + a_0(x)y(x) = 0 \tag{D.9}$$

may be written

$$(a_n D^n + ... + a_1 D_y + a_0)\, y = 0. \tag{D.10}$$

If we let $L := a_n D^n + ... + a_1 D + a_0$ then (D.10) may be written $L[y] = 0$. Since D is a linear operator, if f and g are solutions for (D.10), then $L(\alpha f + \beta g) = 0$.

Consider the equation

$$y''' + 3y'' + 4y' + y = 0. \tag{D.11}$$

This may be written

$$(D^3 + 3D^2 + 4D + 1)y = 0.$$

Thus, if $P(\lambda) = \lambda^3 + 3\lambda^2 + 4\lambda + 1$, then $P(D) = D^3 + 3D^2 + 4D + 1$, and letting $L = P(D)$ (D.11) may be written

$$L[y] \quad = \quad P(D)[y] = 0.$$

<u>Note</u>: $P(\lambda)$ is a polynomial function in λ but $P(D)$ is a (linear) differential operator.

We have

$$\begin{aligned} D[e^{\lambda x}] &= \lambda e^{\lambda x} \\ D^2[e^{\lambda x}] &= \lambda^2 e^{\lambda x} \\ D^n[e^{\lambda x}] &= \lambda^n e^{\lambda x}. \end{aligned}$$

Thus, for a polynomial $P(\lambda) = \lambda^n + a_n - 1\lambda^{n-1} + ... + a_1\lambda + a_0$, we have

$$\begin{aligned} P(D)[e^{\lambda x}] &= (\lambda^n + a_{n-1}\lambda^{n-1} + ... + a_1\lambda + a_0)\, e^{\lambda x} \\ &= P(\lambda)e^{\lambda x}. \end{aligned}$$

Now, if $e^{\lambda x}$ is a solution to the equation $P(D)[y] = 0$, then it follows that $P(\lambda)\, e^{\lambda x} = 0$. Since $e^{\lambda x} > 0$. We must have $P(\lambda) = 0$. The n^{th} degree polynomial will have n roots $\lambda_1, ..., \lambda_n$. Hence, $P(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)...(\lambda - \lambda_n)$ and our ODE may be written $(D - \lambda_1)\,(D - \lambda_2)...(D - \lambda_n)y = 0$. Thus, if λ is a root, then $y(x) = e^{\lambda x}$ satisfies the ODE. If we have repeated roots $\lambda_i = \lambda_j = \gamma$, then we may write

$$P(\lambda) = (\lambda - \gamma)^2\, Q(\lambda),$$

where $Q(\lambda)$ is an $(n-2)$-degree polynomial. Clearly $P(\gamma) = 0$ and

$$\frac{dP}{d\lambda} \quad = \quad 2(\lambda - \gamma)Q(\lambda) + (\lambda - \gamma)^2 Q'(\lambda).$$

Hence, $\frac{dP}{d\lambda}|_{\lambda=\gamma} = 0$.

If we differentiate $P(D)[e^{\lambda x}] = P(\lambda)e^{\lambda x}$ with respect to λ, we get

$$P(D)[xe^{\lambda x}] = P'(\lambda)e^{\lambda x} + P(\lambda)\, xe^{\lambda x}.$$

Thus, if $\lambda = \gamma$ is a repeated root of $P(\lambda) = 0$, then $e^{\gamma x}$ and $xe^{\gamma x}$ are solutions of $P(D)[y] = 0$.

D.5 Solving 2nd order linear ODE

Consider the equation

$$y'' + by' + cy = 0, \tag{D.12}$$

where $y = y(x)$ and $y' = \frac{dy}{dx}$.

Try the solution[2] $y = e^{\lambda x}$.
To find λ, substitute $y = e^{\lambda x}$ into (D.12):

$$\begin{aligned} \lambda^2 e^{\lambda x} + b\,\lambda e^{\lambda x} + c\,e^{\lambda x} &= 0 \\ \Leftrightarrow \qquad (\lambda^2 + b\,\lambda + c)\,e^{\lambda x} &= 0. \end{aligned}$$

Since $e^{\lambda x} > 0$ for all λ and x we must have

$$\lambda^2 + b\,\lambda + c = 0. \tag{D.13}$$

(D.13) is referred to as the *auxilliary equation* and, since it is a quadratic, it must have 2 roots. There are 3 cases to consider, depending on the value of $b^2 - 4c$:

1. If $b^2 - 4c > 0$, then the roots λ_1, λ_2 are real and distinct. In this case $y_1 = e^{\lambda_1 x}$ and $y_2 = e^{\lambda_2 x}$ are solutions.

2. If $b^2 - 4c = 0$ then the roots $\lambda = \lambda_1 = \lambda_2$ are real and equal. In this case $y_1 = e^{\lambda x}$ and $y_2 = x\,e^{\lambda x}$ are solutions (from the previous section; this can be verified by direct substitution).

3. If $b^2 - 4c < 0$ then the roots λ_1, λ_2 are complex conjugates. In this case $y_1 = e^{(\alpha+i\beta)x}$ and $y_2 = e^{(\alpha-i\beta)x}$, are solutions where $\lambda_1 = \alpha + i\beta,\ \lambda_2 = \overline{\lambda_1} = \alpha - i\beta$.

In each case our 2 solutions are linearly independent and the general solution may be written

$$y(x) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}. \tag{D.14}$$

[2]We are looking for a function whose second derivative is a linear combination of the function and it's first derivative. In this case the guess, which is preceding any theory, is the most apparent when $b = 0$ or $c = 0$; a guess of this sort is referred to as an *ansatz*.

In case 3, (D.14) can be written

$$\begin{aligned}
y(x) &= C_1 e^{(\alpha+i\beta)x} + C_2 e^{(\alpha-i\beta)x} \\
&= C_1 e^{\alpha x}\left(\cos\,\alpha\,x + i\sin\beta\,x\right) + C_2 e^{\alpha x}\left(\cos\,\alpha\,x - i\sin\beta\,x\right) \\
&= e^{\alpha x}\left((C_1 + C_2)\cos\alpha\,x + i(C_1 - C_2)\sin\beta\,x\right) \\
&= e^{\alpha x}\left(D_1\cos\alpha\,x + D_2\sin\beta\,x\right) \\
&= D_1 e^{\alpha x}\cos\beta\,x + D_2 e^{\alpha x}\,\sin\beta\,x,
\end{aligned}$$

where $D_1 \;=\; C_1 + C_2$ and $D_2 = i(C_1 - C_2)$.

For this case the solutions are real-valued when D_1 and D_2 are real. The functions

$$\begin{aligned}
y_1 &= e^{\alpha x}\cos\beta\,x \\
y_2 &= e^{\alpha x}\sin\beta x
\end{aligned}$$

are our 2 linearly independent solutions.