MATH 053/126 - Partial Differential Equations

Carter Kruse, November 16, 2023

Homework 6

Instructions/Notation

Please show all steps to get your answers. Specify the problems you discussed with other students (including names).

The starred problems are recommended but not required for undergraduate/non-math major graduate students and required for all math major graduate students.

Notation

- $\mathbb{R}$: The set of all real numbers.
- $\mathbb{R}^+$: The set all positive real numbers $\{x \in \mathbb{R} | x > 0\}$.
- $\frac{\partial u}{\partial t} = \partial_t u = u_t$
- $\frac{\partial u}{\partial x} = \partial_x u = u_x$

Questions

Question 1

We consider an ordinary differential operator L in $(-1, 1)$

$$Lu = -\frac{d}{dx}\left(\left(1 - x^2\right)\frac{du}{dx}\right) + 2u$$

where $u(x) = 1$ for $x = \pm 1$. Assume that there exist eigenfunctions of L that are not constant functions. Show that eigenvalues of L are positive.

In this problem, we are given an ordinary differential operator L in $(-1, 1)$

$$Lu = -\frac{d}{dx}\left(\left(1 - x^2\right)\frac{du}{dx}\right) + 2u$$

where $u(x) = 1$ for $x = \pm 1$, which represents the boundary conditions. Now, let us assume that there exist eigenfunctions of L that are not constant functions. That is, there exist non-constant $u(x)$ such that $Lu = \lambda u$ for eigenvalues λ.

The goal is to demonstrate that the eigenvalues λ of the ordinary differential operator L are positive. To do so, let us multiply the expression $Lu = \lambda u$ by u and integrate over the domain $(-1, 1)$, as follows:

$$\int_{-1}^{1} uLu\,dx = \lambda \int_{-1}^{1} u^2\,dx$$

This may be represented as follows:

$$\int_{-1}^{1} u\left(-\frac{d}{dx}\left(\left(1 - x^2\right)\frac{du}{dx}\right) + 2u\right)dx = \lambda \int_{-1}^{1} u^2\,dx$$

Now, if we apply integration by parts $\left(\int u\,dv = uv - \int v\,du\right)$ to the left-hand side of the equation (given the boundary conditions), we have

$$\int_{-1}^{1} \left(1 - x^2\right) \frac{d^2u}{dx^2} + 2u^2\,dx = \lambda \int_{-1}^{1} u^2\,dx$$

Now, given that $\left(1 - x^2\right) \geq 0$ and $\frac{d^2u}{dx^2} \geq 0$ on $(-1, 1)$, we know that the value of the left-hand side of the equation is strictly positive. Considering $\int_{-1}^{1} u^2\,dx$ is positive as well, this implies that $\lambda > 0$. Therefore, the eigenvalues of L corresponding to non-constant eigenfunctions must be positive. $\square$

Question 2

Show that the eigenfunctions corresponding to different/distinct eigenvalues are orthogonal.

Let L be a symmetric operator on a function space H. We later demonstrate that the ordinary differential operator L in $(-1, 1)$ is symmetric.

To demonstrate that eigenfunctions corresponding to different/distinct eigenvalues (of a symmetric operator L on a function space H) are orthogonal, consider u and v as the eigenfunctions of L corresponding to distinct eigenvalues λ and μ, respectively.

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

Here L is considered to be a symmetric operator on a function space H, so for any $f, g \in H$,

$$\langle Lf, g\rangle = \langle f, Lg\rangle$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product on H. In this case, let $f = u$ and $g = v$, so

$$\langle Lu, v\rangle = \langle u, Lv\rangle$$

If we substitute the eigenvalue expressions for Lu and Lv, respectively, this results in

$$\langle \lambda u, v\rangle = \langle u, \mu v\rangle$$

Now, we may pull λ and μ out of the inner products, since they are constants, which yields

$$\lambda\langle u, v\rangle = \mu\langle u, v\rangle$$

Since $\lambda \neq \mu$ (as the eigenvalues are distinct), we have $\langle u, v\rangle = 0$, indicating that the eigenfunctions u and v corresponding to distinct eigenvalues of the symmetric operator L (on a function space H) must be orthogonal. $\square$

To demonstrate that the ordinary differential operator L in $(-1, 1)$

$$Lu = -\frac{d}{dx}\left(\left(1 - x^2\right)\frac{du}{dx}\right) + 2u$$

where $u(x) = 1$ for $x = \pm 1$ is symmetric, let us consider that it takes the following form

$$\lambda m(x)\, u = -\nabla \cdot (p(x)\, \nabla u) + q(x)\, u$$

where $u = 0$ on ∂D, where $D = (-1, 1)$ and $p(x) > 0$ and $m(x) > 0$. To rigorously proof that L is symmetric over the function space H, consider the following.

Let $u, v \in H$, where H is a suitable function space over $D = (-1, 1)$. That is H is the space of smooth functions that satisfy the boundary conditions $u(\pm 1) = 1$. We define the inner product on H as

$$\langle u, v\rangle = \int_{-1}^{1} uv\, dx$$

To show that L is symmetric, we must show that $\langle Lu, v\rangle = \langle u, Lv\rangle$ for all $u, v \in H$.

$$\langle Lu, v\rangle = \int_{-1}^{1} \left(-\frac{d}{dx}\left(\left(1 - x^2\right)\frac{du}{dx}\right) + 2u\right) v\, dx$$

Now, if we apply integration by parts $\left(\int u\,dv = uv - \int v\,du\right)$ to the right-hand side of the equation (given the boundary conditions), we have

$$\langle Lu, v\rangle = \int_{-1}^{1} \left(1 - x^2\right) \frac{du}{dx}\frac{dv}{dx} + 2uv\,dx$$

This process demonstrates the symmetric of the ordinary differential operator L, as

$$\langle Lu, v\rangle = \int_{-1}^{1} \left(-\frac{d}{dx}\left(\left(1 - x^2\right)\frac{dv}{dx}\right) + 2v\right) u\,dx = \langle u, Lv\rangle$$

Thus, L is symmetric over the function space H. $\square$

Question 3

Prove that

$$\delta\left(a^2 - r^2\right) = \delta\left(a - r\right)/2a$$

for $a > 0$ and $r > 0$.

The expressions $\delta\left(a^2 - r^2\right)$ and $\delta\left(a - r\right)$ are composition of the delta function with the functions

$$w = \phi\left(r\right) = a^2 - r^2 \qquad z = \phi\left(r\right) = a - r$$

The inverse of these functions (for $r > 0$) are

$$r = \phi^{-1}\left(w\right) = \sqrt{a^2 - w} \qquad r = \phi^{-1}\left(z\right) = a - z$$

The following is according to the textbook.

If f is a distribution and $w = \phi\left(r\right)$ is a function with $\phi'\left(r\right) \neq 0$, the meaning of the distribution $g\left(r\right) = f\left[\phi\left(r\right)\right]$ is

$$\left(g, \phi\right) = \left(f, \frac{\phi\left[\phi^{-1}\left(w\right)\right]}{\left|\phi'\left[\phi^{-1}\left(w\right)\right]\right|}\right)$$

for all test functions ϕ. This definition is motivated by the change of variables $w = \phi\left(r\right)$ for ordinary integrals

$$\int f\left[\phi\left(r\right)\right]\phi\left(r\right)\,dr = \int f\left(w\right)\phi\left(r\right)\frac{dw}{\left|\phi'\left(r\right)\right|}$$

Thus, let us calculate the following, respectively

$$\phi'\left(r\right) = -2r = -2\left(\sqrt{a^2 - w}\right) \qquad \phi'\left(r\right) = -1$$

For any test function $\phi\left(r\right)$, we have the following (for $a > 0$), by the definition provided,

$$\left(\delta\left(a^2 - r^2\right), \phi\right) = \left(\delta\left(w\right), \frac{\phi\left(\sqrt{a^2 - w}\right)}{2\sqrt{a^2 - w}}\right) = \frac{\phi\left(a\right)}{2a}$$

$$\left(\delta\left(a - r\right), \phi\right) = \left(\delta\left(z\right), \frac{\phi\left(\sqrt{a - z}\right)}{1}\right) = \phi\left(a\right)$$

This implies that (for all ϕ),

$$\left(\delta\left(a^2 - r^2\right), \phi\right) = \left(\frac{1}{2a}\delta\left(a - r\right), \phi\right)$$

which serves to complete the proof. □

Alternatively, consider the following. We are going to start from $\delta\left(a^2 - r^2\right)$. Let ϕ be an arbitrary test function.

$$\left(\delta\left(a^2 - r^2\right), \phi\left(r\right)\right) = \int_{\mathbb{R}} \delta\left(a^2 - r^2\right)\phi\left(r\right)\,dr$$

Now, we substitute $w = a^2 - r^2$, which results in the following:

$$\begin{aligned}\left(\delta\left(a^2 - r^2\right), \phi\left(r\right)\right) &= \int_{\mathbb{R}} \delta\left(w\right)\phi\left(\sqrt{a^2 - w}\right)\frac{1}{2\left(\sqrt{a^2 - w}\right)}\,dr \\ &= \left(\delta\left(w\right), \frac{\phi\left(\sqrt{a^2 - w}\right)}{2\left(\sqrt{a^2 - w}\right)}\right)\end{aligned}$$

Now, with $a > 0$ and $w = 0$, this is equivalent to

$$\begin{aligned}\left(\delta\left(a^2 - r^2\right), \phi(r)\right) &= \frac{\phi(a)}{2a} \\ &= \frac{1}{2a}\left(\delta(w), \phi(a - w)\right) \\ &= \int_{\mathbb{R}} \frac{1}{2a}\delta(w)\,\phi(a - w)\; dw\end{aligned}$$

Now, we substitute $r = a - w$, which results in the following:

$$\begin{aligned}\left(\delta\left(a^2 - r^2\right), \phi(r)\right) &= \int_{\mathbb{R}} \frac{1}{2a}\delta(a - r)\,\phi(r)\; dr \\ &= \left(\frac{1}{2a}\delta(a - r), \phi(r)\right)\end{aligned}$$

This gives the following

$$\left(\delta\left(a^2 - r^2\right), \phi(r)\right) = \left(\frac{1}{2a}\delta(a - r), \phi(r)\right)$$

which in terms of the distribution, indicates

$$\delta\left(a^2 - r^2\right) = \delta(a - r)/2a$$

This completes the proof. $\square$

Question 4

Consider the equation

$$u_{tt} - c^2 \Delta u + mu = 0 \qquad (t, x) \in \mathbb{R} \times \mathbb{R}^d$$

where $m > 0$ is a constant. *(FYI, this is a relativistic wave equation related to the Schrodinger equation.)* Find the form of energy that is constant in time. Remember to show that energy is constant by taking derivatives.

The energy associated with the wave equation $u_{tt} - c^2 \Delta u + mu = 0 \qquad (t, x) \in \mathbb{R} \times \mathbb{R}^d$ where $m > 0$ is a constant is given by the following integral over the space of the energy density:

$$E(t) = \int_{\mathbb{R}^d} \left[\frac{1}{2} u_t^2 + \frac{1}{2} c^2 |\nabla u|^2 + \frac{1}{2} m u^2 \right] dx$$

This form of energy is constant in time, which is demonstrated by taking the time derivative, as follows:

$$\frac{dE}{dt} = \int_{\mathbb{R}^d} \left[u_t u_{tt} + c^2 \nabla u \cdot \nabla u_t + m u u_t \right] dx$$

Now, let us consider the middle term, which we may use integration by parts to simplify, as follows:

$$\int_{\mathbb{R}^d} \nabla u \cdot \nabla u_t \, dx = \int_{\partial \mathbb{R}^d} u_t \frac{du}{dn} ds - \int_{\mathbb{R}^d} u_t \Delta u \, dx$$

We may assume that the boundary term $\int_{\partial \mathbb{R}^d} u_t \frac{du}{dn} ds$ vanishes, as $u_t, \frac{du}{dn} \to 0$. Thus, we have

$$\int_{\mathbb{R}^d} \nabla u \cdot \nabla u_t \, dx = - \int_{\mathbb{R}^d} u_t \Delta u \, dx$$

Now we have the time derivative written as follows:

$$\frac{dE}{dt} = \int_{\mathbb{R}^d} \left[u_t u_{tt} + c^2 u_t \Delta u + m u u_t \right] dx$$

which may be expressed as

$$\frac{dE}{dt} = \int_{\mathbb{R}^d} \left[u_t \left(u_{tt} + c^2 \Delta u + m u \right) \right] dx$$

Given that the wave equation $u_{tt} + c^2 \Delta u + mu = 0$, we are able to demonstrate that

$$\frac{dE}{dt} = \int_{\mathbb{R}^d} \left[u_t (0) \right] dx = 0$$

Since the time derivative of the energy is zero, the energy $E(t)$ is constant in time. $\square$

Question 5

Solve $u_{xx} + u_{yy} = 1$ in $r = \sqrt{x^2 + y^2} < a$ with $u(x, y)$ vanishing on $r = a$. Hint: Does your solution depend on the angle? Use the Laplacian in the polar coordinate.

To solve the PDE $u_{xx} + u_{yy} = 1$ in $r = \sqrt{x^2 + y^2} < a$ with $u(x, y)$ vanishing on $r = a$, let us consider that the data (the right-hand side) and boundary conditions) are all radially symmetric, so it makes sense to aim for a solution that is radially symmetric. Thus, the solution does *not* depend on the angle, so u is a function of the radius r exclusively.

The Laplacian in polar coordinates ($x = r\cos(\theta)$, $y = r\sin(\theta)$) is given by

$$\Delta = \nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}$$

In this case, we may omit the θ term, since we assume u only depends on r, that is the solution does not depend on the angle.

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r}$$

Thus, we have

$$u_{rr} + \frac{1}{r}u_r = 1 \qquad 0 \leq r \leq 1 \text{ and } u = 0 \text{ for } r = a$$

To solve the differential equation, let us observe that

$$(ru_r)_r = ru_{rr} + u_r \longrightarrow \frac{1}{r}(ru_r)_r = u_{rr} + \frac{1}{r}u_r$$

Thus, we may express the following

$$(ru_r)_r = r$$

To solve the ordinary differential equation, we may simply use integration, as follows:

$$\begin{aligned}(ru_r)_r = r &\Rightarrow ru_r = \frac{1}{2}r^2 + C_1 \\ &\Rightarrow u_r = \frac{1}{2}r + C_1\left(\frac{1}{r}\right) \\ &\Rightarrow u = \frac{1}{4}r^2 + C_1 \ln r + C_2\end{aligned}$$

Since u must be continuous on the entire unit disk ($u(0) < \infty$), $C_1 = 0$. The boundary condition $u(a) = 0$ forces $C_2 = -\frac{a^2}{4}$. Thus, our solution is

$$u(r)\frac{1}{4}\left(r^2 - a^2\right)$$

Converting back to Cartesian coordinates, this is equivalent to

$$u(x, y) = \frac{1}{4}\left(x^2 + y^2 - a^2\right)$$

By the uniqueness theorem for the Dirichlet problem, this much be the only solution. The solution does not append on the angle θ, as we assumed. □

Question 6

The following problem is called the advection-diffusion equation

$$u_t + u_x = u_{xx} \qquad x \in \mathbb{R}$$

where $u(0,x) = \phi(x) \in L^2$. The advection-diffusion equation models a wide range of applications, which is a simple version of the Fokker-Planck equation. Discuss how to solve the problem.

The advection-diffusion equation $u_t + u_x = u_{xx}, x \in \mathbb{R}$ where $u(0,x) = \phi(x) \in L^2$ models the evolution of a quantity $u(t,x)$ that is both diffusing and being advected (transported) in space. In this expression $u(t,x)$ is the unknown function, t is time, and x is the spatial variable.

The term u_t represents the rate of change in time, u_x represents advection or transport in the x direction and u_{xx} represents diffusion.

This equation may be solved using various methods, depending on the characteristics of the problem, boundary conditions, and initial conditions.

Method Of Characteristics
The method of characteristics is often used for first-order linear partial differential equations (PDEs). The idea is to find curves along with the PDE becomes an ordinary differential equation (ODE).

Separation Of Variables
If the initial and boundary conditions are appropriate, separation of variables can be applied. This method assumes $u(t,x) = T(t)X(x)$, and substitutes into the PDE to obtain two separate ODEs. To solve the problem, we solve each ODE and combine the solutions to get the general solution.

Fourier Transform
If the problem has periodic boundary conditions or the initial condition is well-behaved, you may use the Fourier transform to simplify the problem. Apply the Fourier transform to both sides of the PDE and solve the resulting ODE in the frequency domain.

Finite Difference Methods
Numerical (non-analytical) methods such as finite differences methods may be employed to discretize the spatial and temporal derivatives. For example, you may use an explicit or implicit finite difference scheme to approximate the derivatives.

Green's Function Method
If you are dealing with a linear, homogeneous problem with suitable boundary conditions, Green's function method may be applied. You construct the Green's function for the problem and use it to find the solution.

Characteristics & Homogeneous Solution
Rewrite the equation as $u_t + u_x - u_{xx} = 0$ and find the homogeneous solution. Use characteristics to find the particular solution.

The choice of method depends on the specific characteristics of the problem, the form of the initial and boundary conditions, and the desired level of accuracy (for numerical methods).

Further Information

Since this is a linear partial differential equation, we may use methods like separation of variables, Fourier transforms, and Green's functions to try to find analytical solutions. For simple initial conditions $\phi(x)$, we may be able to find a closed form solution. For example, if $\phi(x) = \delta(x)$, the solution is a Gaussian spreading out and moving over time.

For more complex initial conditions, we may have to solve numerically. Common numerical methods include finite differences, finite elements, and spectral methods, which discretize the equation in space and time. The stability and convergence of numerical methods is important. Since there is both diffusion and advection, and implicit method may be preferred over explicit methods.

Boundary conditions need to be specified on a bounded domain, either Dirichlet, Neumann, or periodic. The type of boundary condition impacts the behavior of the solution. Further, the relative size of the diffusion and advection parameters affects the behavior. If diffusion dominates, the solution smoothes out quickly. If advection dominates, the solution propagates without smoothing much.

In summary, analytical solutions are possible in simple cases, while numerical methods are needed for more complex problems. The parameters control the diffusion-advection trade-off and hence the qualitative solution behavior.

The following provides an example of solving the advection-diffusion equation

$$u_t + u_x = u_{xx} \qquad x \in \mathbb{R}$$

where $u(0,x) = \phi(x) \in L^2$, specifically using Fourier transforms. *(The following was discussed alongside Tunmay Gerg.)*

Consider the Fourier transform applied to the advection-diffusion equation, where the Fourier transform is given as

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

This results in the following

$$\widehat{u_t} + \widehat{u_x} = \widehat{u_{xx}} \longrightarrow \partial_t u(k,t) + iku(k,t) = -k^2 u(k,t)$$

This indicates, via simple algebra, that

$$\partial_t u(k,t) = -\left(k^2 + ik\right) u(k,t)$$

That is,

$$u(k,t) = g(k)\, e^{-\left(k^2+ik\right)t}$$

where g is a function of k. This implies the following, which is the Fourier transform with $t = 0$:

$$u(k,0) = g(k) = \int u(x,0)\, e^{-ikx}\, dx = \widehat{\phi(k)}$$

Now, we have

$$u(k,t) = \phi\,\hat{(}k) e^{-\left(k^2+ik\right)t}$$

Let $g(x)$ be such that $\widehat{g(k)} = e^{-\left(k^2+ik\right)t}$, so that the following convolution holds

$$\widehat{g * \phi} = \widehat{g(k)}\widehat{\phi(k)}$$

This implies

$$g * \phi = \overline{\widehat{g(k)}\widehat{\phi(k)}} = \overline{u(k,t)} = u(x,t)$$

Thus, we have the following calculation for $g(x,t)$:

$$\begin{aligned} g(x,t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\left(k^2+ik\right)t} e^{ikx}\, dk \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\left(tk^2+i(t-x)k\right)}\, dk \\ &= \frac{1}{2\pi}\sqrt{\frac{\pi}{t}} e^{\frac{[i(t-x)]^2}{4t}} \\ &= \frac{1}{\sqrt{4\pi t}} e^{\frac{-(x-t)^2}{4t}} \end{aligned}$$

This implies that

$$u(x,t) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{4\pi t}} e^{\frac{-(x-y-t)^2}{4t}} \phi(y)\, dt$$

Further proof is required to show that $u(x,t)$ does solve the partial differential equation and the boundary conditions. For $t \to 0$, the boundary conditions are satisfied, as

$$\frac{1}{\sqrt{4\pi t}} e^{\frac{-(x-y-t)^2}{4t}} \to \delta(x-y)$$

To check that $u(x,t)$ solves the PDE, we need to check that

$$\frac{1}{\sqrt{4\pi t}} e^{\frac{-(x-y-t)^2}{4t}}$$

solves the PDE, which is tedious to check.

Question 7

Derive the Fokker-Planck equation of the random walk. Hint

The following is according to UPenn.

From Random Walk to Diffusion

A stochastic process continuous in both space and time.

This time we will give a (first) derivation of the Fokker-Plank equation, which governs the evolution of the probability density function of a random variable-valued function $X(t)$ that satisfies a "first-order stochastic differential equation." The variable evolves according to what in the literature is called a Wiener process, named after Norbert Wiener, who developed a great deal of the theory of stochastic processes during the first half of this century.

We shall begin with a generalized version of random walk (an "unbalanced" version where the probability of moving in one direction is not necessarily the same as that of moving in the other). Then we shall let both the space and time increments shrink in such a way that we can make sense of the limit as both Δx and Δt approach 0.

Generalized Random Walk

We consider a generalized one-dimensional random walk (on the integers) beginning at $X = 0$. We fix a parameter n, and let p be the probability of taking a step to the right, so that $q = 1 - p$ is the probability of taking a step to the left, at any (integer-valued) time t. In symbols, this is

$$p = \text{Prob}\left(X\left(n+1\right) = x+1 | X\left(n\right) = x\right) \qquad q = \text{Prob}\left(X\left(n+1\right) = x-1 | X\left(n\right) = x\right)$$

Since we are given that $X(0) = 0$, we know that the probability function of $X(n)$ will be essentially binomially distributed (with parameters n and p, except that at time n, x will go from $-n$ to $+n$, skipping every other integer). We will set $v_{x,n} = \text{Prob}\left(X\left(n\right) = x\right)$. Then

$$v_{x,n} = \binom{n}{\frac{1}{2}\left(n+x\right)} p^{\frac{1}{2}(n+x)} q^{\frac{1}{2}(n-x)}$$

for $x = -n, -\left(n-2\right), \ldots, \left(n-2\right), n$ and $v_{x,n} = 0$ otherwise. For use later, we calculate the expected value and variance of $X(n)$.

$$E\left(X_n\right) = \sum_{x=-n}^{n} x v_{x,n} = \sum_{k=0}^{n} \left(2k-n\right) v_{x,2k-n} = \sum_{k=0}^{n} \left(2k-n\right) \binom{n}{k} p^k q^{n-k} = n\left(p-q\right)$$

$$\text{Var}\left(X_n\right) = \sum_{k=0}^{n} \left(2k-n\right)^2 v_{x,2k-n} = \sum_{k=0}^{n} \left(2k-n^2\right) \binom{n}{k} p^k q^{n-k} = 4npq$$

Note that we could have thought of $X(n)$ as the sum of n *independent* binomially distributed random variables, each having expected value $p - q$ and variance $4pq$. The same results follow from this observation.

Now, we let the space and time steps be different from 1. First, we will take time steps of length Δt. This will have little effect on our formulas, except everywhere we will have to note that the number n of time steps needed to get from time 0 to time t is

$$n = \frac{t}{\Delta t}$$

For X, we change our notation so that each step to the left or right has length Δx instead of length 1. This has the effect of scaling $X(t)$ be the factor Δx. Multiplying a random variable by a constant has the effect of multiplying its expected value by the same constant, and its variance by the square of the constant. Therefore, the mean and variance of the total displacement in time t are as follows:

$$E(X(t)) = \frac{t}{\Delta t}(p-q)\Delta x \qquad \text{Var}(X(t)) = 4pqt\frac{(\Delta x)^2}{\Delta t}$$

We want both of these quantities to remain finite for all fixed t, since we do not want our process to have a positive probability of zooming off to infinity in finite time, and we also want $\text{Var}(X(t))$ to be non-zero. Otherwise, the process will be completely deterministic.

We need to examine what we have control over in order to achieve these ends. First off, we do not expect to let Δx and Δt to go to zero in any arbitrary manner. In fact, it seems clear that in order for the variance to be finite and non-zero, we should insist that Δt be roughly the same order of magnitude as $(\Delta x)^2$. This is because we do not seem to have control over anything else in the expression for the variance.

It is reasonable that p and q could vary with Δx, but we probably do not want their product to approach zero (that would force our process to move always to the left or always to the right), and it cannot approach infinity (since the maximum value of pq is $\frac{1}{4}$). So we are forced to make the assumption that $\frac{(\Delta x)^2}{\Delta t}$ remain bounded as they both go to zero. In fact, we will go one step further. We will insist that

$$\frac{(\Delta x)^2}{\Delta t} = 2D$$

for some (necessarily positive) constant D. Now, we are in a bind with the expected value. Under the assumption, the expected value will go to infinity as $\Delta x \to 0$, and the only parameter we have to control it is $p-q$. We need $p-q$ to have the same order of magnitude as Δx in order to keep $E(X(t))$ bounded. Since $p = q$ when $p = q = \frac{1}{2}$, we will make the assumption that

$$p = \frac{1}{2} + \frac{c}{2D}\Delta X$$

for some real (positive, negative, or zero) constant c, which entails

$$q = \frac{1}{2} - \frac{c}{2D}\Delta X$$

The reason for the choice of parameters will be apparent. D is called the diffusion coefficient, and c is called the drift coefficient. Note that as Δx and Δt approach zero, the values $v_{n,k} = \text{Prob}(X(n\Delta t) = k\Delta x) = f(n\Delta t, k\Delta x)$ specify values of the function f at more and more points in the plane. In fact, the points at which f is defined become more and more dense in the plane as Δx and Δt approach zero. This means that given any point in the plane, we can find a sequence of points at which such an f is defined, that approach our given point as Δx and Δt approach zero.

We would like to be able to assert the existence of the limiting function, and that is expressed how the probability of a continuous random variable evolves with time. We will not quite do this, but assuming the existence of the limiting function, we will derive a partial differential equation that it must satisfy. Then, we shall use a probabilistic argument to obtain a candidate limiting function. Finally, we will check that the candidate limiting function is a solution of the partial differential equation. This is not quite a proof of the result we would like, but it touches on the important steps in the argument.

We begin with the derivation of the partial differential equation. To do this, we note that for the original random walk, it was true that $v_{x,n}$ satisfies the following difference equation:

$$v_{k,n+1} = pv_{k-1,n} + qv_{k+1,n}$$

This is just a specification of the matrix of the (infinite discrete) Markov process that defines the random walk. When we pass into the Δx, Δt realm, and use $f(x,t)$ notation instead of $v_{k\Delta x,n\Delta t}$, this equation becomes

$$f(x,t+\Delta t) = pf(x-\Delta x,t) + qf(x+\Delta x,t)$$

Next, we need Taylor's expansion for a function of two variables. As we do the expansion, we must keep in mind that we are going to let $\Delta t \to 0$, and as we do this, we must account for all terms of order up to and including that of Δt. Thus, we must account for both Δx and $(\Delta x)^2$ terms (but $(\Delta x)^3$ and $\Delta t \Delta x$ both go to zero faster than Δt). We expand each of the three expressions in the above equation around the point (x,t), keeping terms up to order Δt and $(\Delta x)^2$.

$$f(x,t+\Delta t) = f(x,t) + \Delta(t)\frac{\partial f}{\partial t}(x,t) + \cdots$$

$$f(x+\Delta x,t) = f(x,t) + \Delta(x)\frac{\partial f}{\partial x}(x,t) + \frac{(\Delta x)^2}{2}\frac{\partial^2 f}{\partial x^2}(x,t) + \cdots$$

$$f(x-\Delta x,t) = f(x,t) - \Delta(x)\frac{\partial f}{\partial x} + \frac{(\Delta x)^2}{2}\frac{\partial^2 f}{\partial x^2}(x,t) + \cdots$$

We substitute these three equations, recalling that $p+q=1$, and get

$$\Delta t\frac{\partial f}{\partial t} = (q-p)\Delta x\frac{\partial f}{\partial x} + \frac{(\Delta x)^2}{2}\frac{\partial^2 f}{\partial x^2}$$

where all the partial derivatives are evaluated at (x,t). Finally, we divide through by Δt, and recall that $(\Delta x)^2/\Delta t = 2D$ and $p-q = \Delta x c/D$ to arrive at the equation

$$\frac{\partial f}{\partial t} = -2c\frac{\partial f}{\partial x} + D\frac{\partial^2 f}{\partial x^2}$$

The last partial differential equation is called the *Fokker-Plank equation.* To find a solution of the Fokker-Plank equation, we use a probabilistic argument to find the limit of the random walk distribution as Δx and Δt approach 0. The key step in this argument uses the fact that the binomial distribution with parameters n and p approaches the normal distribution with expected value np and variance npq as n approaches infinity. In terms of our assumptions about Δx, Δt, p, and q, we see that at time t, the binomial distribution has mean

$$t(p-q)\frac{\Delta x}{\Delta t} = 2ct$$

and variance

$$4pqt\frac{(\Delta x)^2}{\Delta t}$$

which approaches $2Dt$ as $\Delta t \to 0$. Therefore, we expect $X(t)$ in the limit to be normally distributed with mean $2ct$ and variance $2Dt$. Thus, the probability that $X(t) < x$ is

$$\text{Prob}(X(t) < x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{x-2ct}{\sqrt{2Dt}}} e^{-\frac{1}{2}\lambda^2}\,d\lambda$$

Therefore, the pdf of $X(t)$ is the derivative of this with respect to x, which yields

$$f(x,t) = \frac{1}{2\sqrt{\pi Dt}}e^{-\frac{x-2ct}{4Dt}}$$

The following provides an overview of the method outlined. The Fokker-Planck equation describes the evolution of the probability density function of a stochastic process, such as a random walk. For simplicity, let's consider a one-dimensional random walk. The basic idea is to consider the probability distribution function $P(x,t)$, which gives the probability of finding the walker at position x at time t.

In a random walk, at each time step, the walker can move to the right or left with certain probabilities. Let $p_+(x,t)$ be the probability of moving to the right, and $p_-(x,t)$ be the probability of moving to the left. The probability of staying at the same position is $p_0(x,t)$. The Fokker-Planck equation is derived by considering the continuity equaiton for the probability distribution function.

The probability flux, $J(x,t)$, represents the flow of probability in space and time. It is given by

$$J(x,t) = p_-(x,t) - p_+(x,t)$$

The change in probability density in a small time interval Δt is given by

$$\Delta P(x,t) = P(x,t+\Delta t) - P(x,t)$$

According to the continuity equation, this change is related to the probability flux, as follows:

$$\Delta P(x,t) = -\frac{\partial J(x,t)}{\partial x}\Delta x \Delta t$$

Now, we can express the probabilities $p_\pm$ in terms of the probability distribution function P using Taylor expansions:

$$p_+(x,t) = \frac{1}{2}\left[1 - \mu(x,t)\Delta x + \frac{1}{2}\sigma^2(x,t)\Delta t\right]$$

$$p_-(x,t) = \frac{1}{2}\left[1 + \mu(x,t)\Delta x + \frac{1}{2}\sigma^2(x,t)\Delta t\right]$$

$$p_0(x,t) = 1 - \sigma^2(x,t)\Delta t$$

Here, $\mu(x,t)$ is the drift coefficient and $\sigma^2(x,t)$ is the diffusion coefficient. Substitute these expressions into the probability flux $J(x,t)$ and then into the continuity equation. Take the limit as Δx and Δt approach zero, and re-arrange terms to obtain the Fokker-Planck equation.

The result is a partial differential equation that describes how the probability distribution function evolves in time, given the drift and diffusion coefficients of the random walk. The general form of the Fokker-Planck equation is

$$\frac{\partial P(x,t)}{\partial t} = -\frac{\partial}{\partial x}\left[\mu(x,t)P(x,t)\right] + \frac{\partial^2}{\partial x^2}\left[\frac{1}{2}\sigma^2(x,t)P(x,t)\right]$$

Question 8

Read the paper "Investigations on the Theory of the Brownian Movement" by Albert Einstein (1906), and make a summary of the paper (max one page, please type). What is the connection between the Brownian motion and the random walk? Also, compare the diffusion equation and the Fokker-Planck equation.

The paper "Investigations on the Theory of the Brownian Movement" by Albert Einstein (1906) may be found here or here. The following is a summary of the key points from the paper.

In the 1906 paper "Investigations on the Theory of Brownian Movement", Albert Einstein provided a theoretical explanation of Brownian motion - the random, irregular motion of (microscopic) particles suspended in a fluid. Brownian motion had been observed decades earlier by botanist Robert Brown, who noticed the jittery motion of pollen particles suspended in water through a microscope, though there was no satisfactory explanation for this phenomenon at the time.

Einstein approached this problem from the perspective of the kinetic theory of heat/gases, modelling a particle suspended in a fluid as being randomly bombarded by atoms and molecules of the surrounding medium. In this way, Einstein showed that the erratic motion of a Brownian particle is the result of collisions with particles in the surrounding fluid.

By applying statistical methods, Einstein developed a mathematical model that describes the statistical distribution of particle displacements over time. In the paper, an ensemble of particles starting at an initial position is used to derive a probability distribution function describing the likelihood of finding a particle at a given location at time t. In this sense, Einstein derived a quantitative prediction for the diffusion of Brownian particles (given by the probability distribution) according to statistical analysis and the laws of thermodynamics.

A key result from the theory is that the mean square displacement of a Brownian particle is related to the diffusion coefficient, viscosity of the fluid, absolute temperature, and time elapsed, particularly that the mean square displacement should increase linearly with time when observed over sufficiently long durations. This is expressed mathematically in Einstein's diffusion equation for Brownian motion and provides a foundation for analyzing particle diffusion based on kinetic theory.

Einstein further predicted that different sized particles should exhibit different characteristic diffusion coefficients when suspended in the same fluid.

A remarkable aspect of Einstein's work on Brownian motion is the direct connection between macroscopic observables and the microscopic world of atoms and molecules. The framework developed laid the foundation for later developments in statistical mechanics and revolutionized our understanding between macroscopic phenomena and the microscopic world of particles.

Einstein's theoretical predictions were later confirmed experimentally by Jean Perrin in 1908 - 1909. Perrin's careful measurements of the motions of emulsion particles validated Einstein's model and provided empirical estimates of the Avogadro constant and atom sizes.

Now, we aim to consider the connection between Brownian motion and the random walk, which is topical to this class on partial differential equations. The diffusion equation that Einstein derived for Brownian motion is mathematically equivalent to the equation describing a random walk.

That is, Brownian motion may be modelled as a random walk. The random impacts of surrounding particles/molecules on a Brownian particle lead the particle to "randomly walk" (take a random path) through space, i.e. a path that exhibits no overall preferred direction.

The discrete random steps in a random walk process lead to diffusion over time that is described by the same mathematical formalism, therefore the statistics of Brownian motion follow a random walk process. In this sense, Brownian motion is often considered a continuous version of random walk, as it describes the random movement of particles over time.

The diffusion equation describes the time evolution of the probability distribution of a particle's position in Brownian motion, relating the rate of change of the distribution to the diffusion coefficient.

Random Walk

- In a random walk, a particle moves step by step in random directions.
- At each step, the particle makes a random displacement, and the direction of the displacement is typically determined by a random process.
- The classic example is a one-dimensional random walk, where the particle can move either to the left or right with equal probability at each step.

Brownian Motion

- Brownian motion is a continuous version of a random walk, where the particle's position changes continuously over time. *Discrete time steps become infinitesimally small.*
- It involved the random, erratic motion of particles in a fluid due to the collision with surrounding molecules.
- Brownian motion is characterized by continuous, stochastic fluctuations in the position of the particle.

The Fokker-Planck equation is a generalization of Einstein's diffusion equation that describes the time evolution of the probability density function for a particle undergoing Brownian motion. It includes terms for drift and (position-dependent) diffusion coefficients that model external forces acting on the particle, in addition to the random collisions that result in basic Brownian motion.

In this sense, the diffusion equation only models the probability distribution of position, while the Fokker-Planck equation includes velocity dynamics and distributions, providing a more detailed probabilistic description of particle transport in a fluid. The diffusion equation is a simple case of the Fokker-Planck equation without drift or variable diffusivity, so when simplified to only model position distribution, the Fokker-Planck equation reduces to the diffusion equation.

Both the Fokker-Plank and diffusion equation rely on a stochastic differential equation description of the underlying random process, though the former provides a more general description of phase-space probability distributions.

In summary, Brownian motion follows a random walk, described by the diffusion equation for position probabilities. The Fokker-Planck equation generalizes this to phase-space distributions, reducing to the simpler diffusion equation without drift or variable diffusivity.

Diffusion Equation

- The diffusion equation describes the time evolution of the probability density function (pdf) of particles undergoing random motion.
- In one dimension, the diffusion equation is often written as $\frac{\partial P}{\partial t} = D\frac{\partial^2 P}{\partial x^2}$, where P is the probability density, t is time, x is position, and D is the diffusion coefficient.
- This equation describes how the probability distribution spreads over time, representing the process of diffusion.

Fokker-Planck Equation

- The Fokker-Planck equation is a more general form of the diffusion equation and is used to describe the evolution of probability distributions in systems undergoing stochastic processes.
- It is often written as $\frac{\partial P}{\partial t} = -\frac{\partial(J)}{\partial x}$ where J is the probability current.
- The Fokker-Planck equation accounts for both deterministic and stochastic influences on the system.

In summary, Einstein showed that the erratic Brownian motion of microscopic particles may be modelled mathematically using statistical dynamics and the theory of random walks and diffusion. The derived diffusion equation forms the basis for connecting stochastic Brownian motion to macroscopic transport phenomena like particle diffusion. The diffusion equation is a specific form of the Fokker-Planck equation, which provides a broader framework for analyzing a variety of stochastic processes beyond simple diffusion.

Question 9

For smooth $u(x) : [0,1] \to \mathbb{R}$, find the (Frechet) derivative of

$$\int_0^1 (u_x)^2 \, dx$$

with respect to u where $u(0) = u(1) = 0$.

We are given a smooth function $u(x) : [0,1] \to \mathbb{R}$ such that $u(0) = u(1) = 0$. The aim is to find the Frechet derivative (with respect to u) of the functional

$$J[u] = \int_0^1 (u_x)^2 \, dx$$

To do so, consider the following review of calculus, for a function $f(x) : \mathbb{R} \to \mathbb{R}$

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

This may be expressed as the following, which indicates that the derivative is a linear approximation of the function.

$$O\left(h^2\right) + hf'(x) = f(x+h) - f(x)$$

In this case, let us extend the calculus definition to account for the functional, so that $f(x) \to J[u]$, $x \to u(x)$ and $h \to h(x)$, where $h(0) = h(1) = 0$. That is

$$O\left(h^2\right) + hJ'[u] = J[u+h] - J[u]$$

To determine the value of this expression, we compute the following:

$$J[u+h] = \int_0^1 (u_x)^2 + 2u_x h_x + (h_x)^2 \, dx$$

$$J[u+h] = \int_0^1 (u_x)^2 \, dx$$

Thus, we find that

$$O\left(h^2\right) + hJ'[u] = \int_0^1 (u_x)^2 + 2u_x h_x + (h_x)^2 \, dx - \int_0^1 (u_x)^2 \, dx$$

When we consider the $O\left(h^2\right)$ term, we are able to eliminate the $(h_x)^2$ term from the integral. Further, we take the difference between the integrals, as follows:

$$hJ'[u] = \int_0^1 2u_x h_x \, dx$$

Now, we may apply integration by parts, which results in

$$hJ'[u] = 2u_x h \Big|_0^1 - \int_0^1 2u_{xx} h \, dx$$

As we know that the function $h(x)$ vanishes on the boundary, that is $h(0) = h(1) = 0$, this simplifies to the following:

$$hJ'[u] = -2 \int_0^1 u_{xx} h \, dx$$

This serves as the (Frechet) derivative of $\int_0^1 (u_x)^2 \, dx$ with respect to u where $u(0) = u(1) = 0$, and $u(x) : [0,1] \to \mathbb{R}$ is smooth. $\square$

The Frechet derivative of a functional may be expressed generally as follows. Given the functional

$$J[u] = \int |\nabla u|^2 \, dx$$

with $u = 0$ on $\partial\Omega$, the (Frechet) derivative is

$$hJ'[u] = h\frac{dJ}{du} = -2\int \Delta u \, h \, dx$$

In this case, we consider that $h = 0$ on $\partial\Omega$. To find the minimum state of the system, we aim to find where the integrand vanishes (i.e. the derivative is zero).

Question 10

Find a condition for the above integral to have an extreme (max/min) value. Assume that everything is smooth.

To find a condition for the above integral to have an extreme (max/min) value, let us start with the functional

$$J[u] = \int_0^1 (u_x)^2 \, dx$$

The extreme values occur when the first variation $J'[u]$ vanishes for all variations of h. Let us assume that all functions are smooth. From the previous problem, we have that

$$hJ'[u] = -2 \int_0^1 u_{xx} h \, dx$$

which may be generalized as

$$hJ'[u] = h\frac{dJ}{du} = -2 \int \Delta u \, h \, dx$$

For this to vanish for any h (and including the boundary conditions), we must have

$$\Delta u = u_{xx} = 0 \qquad u(0) = u(1) = 0$$

Solving this expression will determine the condition for the function $J[u]$ to be extremized. This is a simple problem, which has the following solution:

$$u(x) = C_1 x + C_2$$

Given the boundary conditions, we know that $C_1 = C_2 = 0$, and so, the solution is

$$u(x) = 0$$

Alternatively, to find the conditions for the functional

$$J[u] = \int_0^1 (u_x)^2 \, dx$$

to have an extreme (max/min) value, we can use the Euler-Lagrange equation. The Euler-Lagrange equation for a functional of the form

$$J[u] = \int_a^b F(x, u, u_x) \, dx$$

is given by

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

In this case, the function is $J[u] = \int_0^1 (u_x)^2 \, dx$ and $F = (u_x)^2$. Therefore

$$\frac{d}{dx}\left(\frac{\partial \left((u_x)^2\right)}{\partial u_x}\right) - \frac{\partial \left((u_x)^2\right)}{\partial u} = 0$$

If we compute this, it results in the following

$$\frac{d}{dx}(2u_x) - 0 = 0$$

which implies that $u_{xx} = 0$, as we had before. The second-order ordinary differential equation for $u(x)$ may be solved as

$$u(x) = C_1 x + C_2$$

Given the boundary conditions, we know that $C_1 = C_2 = 0$, and so, the solution is

$$u(x) = 0$$

This is the only solution satisfying the boundary conditions. Thus, the function $u(x) = 0$ is a critical/extreme point for the functional $J[u]$. In particular, $u(x) = 0$ is the solution that minimizes the function $J[u]$. $\square$

The arguments shown above are relatively intuitive, according to the definition of the problem. Given that the functional

$$J[u] = \int_0^1 (u_x)^2 \, dx$$

includes the integrand $(u_x)^2$, we know that it cannot take on a negative value. Thus, we deduce that the minimum value of the functional $J[u]$ must be zero.

For the functional $J[u]$ to be zero for any given h, we must have that $\Delta u = u_{xx} = 0$, along with the boundary conditions. This gives us the solution. $\square$

Acknowledgements

In working on this homework assignment, independent solutions were discussed with Tunmay Gerg, specifically Questions 6, 7, and 9.