Supplementary notes

Diffusion, heat flow, advection, fluxes

Introduction

Advection, diffusion and heat flow may be modeled using partial differential equations (PDEs). We describe some of the ideas and methods in an idealized one dimensional setting. Related but more complicated PDEs are used to model diffusion of chemicals in a cell, movement of pollutants in the atmosphere, heat flow in a computer chip, etc. The PDEs are derived using a combination of *local conservation laws* and *constitutive models*. *Local conservation* means that the quantity being modeled, such as heat flow or chemical diffusion, is not created or destroyed, but moves around in a way described by *fluxes* (also called *currents*). *Constitutive relations* are hypotheses about how fluxes arise from the non-uniform distribution of the quantity itself. Putting these together leads to *advection/diffusion* equations. These notes explain the basic ideas involved.

Time dependent fields, distributed parameter models

A *field*, or a *distributed parameter*, is a function of x and possibly t. For us, t represents a time variable and x is a space variable. A point in physical space is specified by three coordinates, so x should be a vector with three components. However, we will imagine a simpler situation where there is only one space variable. We imagine it as the linear distance down a heat-conducting rod or a fluid “column” such as a long narrow pipe. The local quantity being modeled will by called u. This may represent temperature or concentration of something dissolved such as a die. The temperature or concentration can be different at different places. We write $u(x)$ for the value at location x. The value can be different also at different times. In that case, $u(x,t)$ represents the temperature or concentration at location x at time t. The state of the “system” (the thing or situation being modeled) at time t is determined by the function $u(x,t)$, as a function of x. Such a state “variable” is called a *field* or *distributed parameter*. Physics has many “fields”, including gravitational, electric, magnetic, temperature, etc, “Distributed” means that the temperature (or concentration or whatever) field values are distributed throughout space and time, with a possibly different value at each x and t. The alternative to “distributed” is “lumped”, where the temperature or concentration distribution is represented by a finite number of values, such as the averages over small regions or the temperatures measured at some measuring stations.

We use various notations to describe a distributed parameter field. For example, we might write $u(t)$ to represent the temperature distribution at the specific time t. This distribution is a function of x but this dependence is not explicitly represented in the notation. We might write $u(\cdot,t)$ to indicate that for each time t there is a function of a space variable, but this variable is represented by a dot not a name, Thus, $u(t)$ and $u(\cdot,t)$ might mean basically the same thing. We might write $u(x,\cdot)$ for the function that is the temperature at a specific point x as a function of t.

Fluxes and local conservation

Suppose $u(x,t)$ represents the *density* of something. Density is related to amount by integration[1] We write $Q(a,b,t)$ for the amount of "stuff" in the interval $[a,b]$ at time t. It is found by integrating the density over that interval

$$Q(a,b,t) = \int_a^b u(x,t)\, dx \; . \tag{1}$$

You can think of this as the definition of density; it's the amount of stuff per unit length.[2] If the density is not constant, which means $u(x,t)$ is not constant as a function of x, you divide the interval $[a,b]$ into small Δx length segments. The amount of stuff in a segment near x is $u(x,t)\Delta x$ (the density at x multiplied by the length of the small piece). The total amount of stuff in $[a,b]$ is the sum of these local amounts, which gives the integral (1).

Local conservation means that $Q(a,b,t)$ changes when t changes only by stuff crossing the endpoints a and b. The *flux* (also called *current*) at a at time t, which we call $F(a,t)$ is the rate at which stuff flows across a from left right. Local conservation is the statement that for every a and b,

$$\frac{d}{dt} Q(a,.b,t) = -F(b,t) + F(a,t) \; . \tag{2}$$

The first term on the right, indicates that a positive flux at b draws stuff out of the interval $[a,b]$ which makes Q decrease. Fluxes can be negative, so $F(b,t) < 0$ indicates that stuff is flowing from right to left into $[a,b]$ across b. The second term on the right has a plus sign because $F(a,t) > 0$ means that stuff is entering $[a,b]$ over that a endpoint.

Suppose that the whole "domain" is a larger interval $[0,L]$. The density u and quantity Q are defined only inside the domain, which means $0 \leq x \leq L$ and $0 \leq a \leq b \leq L$. The opposite of local conservation is "global" conservation. This would mean that $Q(0,L,t)$ is independent of t. Global conservation would allow stuff to "teleport" (jump?) from one part of the domain to another without crossing an internal boundary at a or b. For example, suppose $a < b < c < d$ and consider the separated intervals $[a.b]$ and $[c,d]$. Global conservation would allow stuff to go directly from $[a.b]$ to $[c,d]$ without crossing the endpoints b or c.

To illustrate local conservation, consider neighboring intervals $[a,b]$ and $[b,c]$. Then (this is a consequence of (1)) $Q(a,c,t) = Q(a,b,t) + Q(b,c,t)$. If $F(a,t) = 0$ and $F(c,t) = 0$, then stuff flows from $[a,b]$ to $[b,c]$ without changing $Q(a,c,t)$.

$$\begin{aligned}
\frac{d}{dt} Q(a,b,t) &= -F(b,t) \\
\frac{d}{dt} Q(b,c,t) &= \quad F(b,t) \\
\frac{d}{dt} Q(a,c,t) &= \frac{d}{dt} Q(a,b,t) + \frac{d}{dt} Q(b,c,t) = -F(c,t) + F(a,t) = 0 \; .
\end{aligned}$$

Local conservation describes how stuff moves around within the domain without being created or destroyed.

Partial derivative notation

There are many ways to denote partial derivatives. Discussions of partial differential equations tend to use all of them at the same time. The most common ones are as a fraction, like ordinary derivatives,

[1] You might have seen this in a probability class. The probability that a random X lands in an interval $[a,b]$ is the integral of the probability density over the interval.

[2] There is a distinction between *volumetric* density and *specific* density. The definition here is volumetric, though "volume" in one dimension is really length.

then using the partial derivative sign but with the fraction bar removed, then using a subscript:

$$\frac{\partial u(x,t)}{\partial x} = \partial_x u(x,t) = u_x(x,t) \; .$$

We think of ∂_x as the partial derivative *operator* that "operates" on a function by taking the partial derivative. Thus, writing $\partial_x u$ may be read as applying the partial derivative operator to the function u. This is supposed to be analogous to writing A as the operator that is applied to a vector using matrix/vector multiplication.

Constitutive relations

Constitutive relations are physical models that say how fluxes arise from local densities or density differences. *Advection* models a situation where the "stuff" is being carried along by a fluid (liquid or gas) that is is dissolved in. Here, the flux (called *advective flux*) is

$$F_a(x,t) = v\, u(x,t) \; . \tag{3}$$

Here, v is the velocity of the fluid and $u(x,t)$ is the density of stuff (die or pollutant) dissolved in the fluid. The rate of flow across x is the density multiplied by the fluid velocity.

Diffusion is movement of stuff that happens even when the medium it is in is not moving. A *diffusive flux* is

$$F_d(x,t) = -D\, \partial_x u(x,t) \; . \tag{4}$$

The *diffusion coefficient*, D is supposed to be positive. The minus sign models the idea that stuff flows from high to low density. To see this another way, suppose $u(x,t)$ has a maximum inside $[a,b]$ so $u_x(a,t) > 0$ and $u_x(b,t) < 0$. Then $F_d(a,t) < 0$ and $F_d(b,t) > 0$, which says that stuff is leaving the interval at both ends. Stuff is moving from where concentration is large to where concentration is smaller.

Heat flow involves two assumptions about the nature of heat. The model of heat flow involves the local temperature, which is called $u(x,t)$ and the local heat energy density, which is called $e(x,t)$. One assumption is that the local heat energy density is proportional to the temperature. This relationship involves *specific heat*, which is the amount of energy it takes to increase the temperature a small amount. Rather than discussing specific heat more, let's just call the constant of proportionality L, so

$$e(x,t) = Lu(x,t) \; .$$

The other assumption is that the flux of heat energy is proportional to the temperature gradient. Again, to emphasize that we're skipping important physics, let's write K for this proportionality constant. That makes the energy flux equal to

$$F_e(x,t) = -K\partial_x u(x,t) \; .$$

The constant K should be positive to make heat flow from high temperature to low temperature. The constant L should be positive so the heat energy increases when the temperature increases. We write $E(a,b,t)$ for the heat energy in $[a,b]$

$$E(a,b,t) = \int_a^b e(x,t)\, dx \; .$$

The local conservation of heat energy (not local conservation of temperature, which is not true) is

$$\frac{d}{dt}E(a,b,t) = -F_e(b,t) + F_e(a,t) \; .$$

Advection/diffusion/heat equations

The modeling assumptions above can be combined to give partial differential equations that describe the dynamics of temperature or concentration fields. For the concentration field, we have (the derivative "goes under" the integral)

$$\partial_t \int_a^b u(x,t)\,dx = \int_a^b u_t(x,t)\,dx \ .$$

The flux formula then gives

$$\int_a^b u_t(x,t)\,dx = -F(b,t) + F(a,t) \ .$$

The fundamental theorem of calculus says the different on the right is the integral of the derivative

$$F(b,t) - F(a,t) = \int_a^b F_x(x,t)\,dx \ .$$

Therefore

$$\int_a^b u_t(x,t)\,dx = -\int_a^b F_x(x,t)\,dx \ . \tag{5}$$

This equality is true for any interval $[a,b]$. The final step is to argue that an integral identity that holds for any interval implies a "pointwise" identity that holds for eery x and t:

$$u_t(x,t) = -F_x(x,t) \ . \tag{6}$$

To see this, suppose u_t and F_x are continuous. If the pointwise equality (6) does not hold at some (x,t), suppose that $u_t > -F_x$ there. Then $u_t > -F_x$ also in a small interval about x so the integral version (5) would not be true either. Finally, if there is only diffusion, Fick's law (4) applied to (6) gives

$$u_t = Du_{xx} \ . \tag{7}$$

This is "the" diffusion equation. You could model a situation with both advection and diffusion by making a total flux that is the sum of an advective and diffusive flux: $F(x,t) = F_a(x,t) + F_d(x,r)$. This would lead to an *advection diffusion* equation.

To model heat flow, we use (using ∂_x instead of a subscript for the x derivative because F already has a subscript)

$$\int_a^b e_t(x,t)\,dt = -\int_a^b \partial_x F_e(x,t)\,dx \ .$$

and then substitute the physical assumptions to get a PDE for u. The left side is

$$\int_a^b Lu_t(x,t)\,dx \ .$$

The right side is

$$\int_a^b \partial_x \left(Ku_x(x,t) \right) dx = \int_a^b Ku_{xx}(x,t)\,dx \ .$$

Setting the integrands equal pointwise as before gives

$$Lu_t = Ku_{xx} \ .$$

This leads to the *heat equation*

$$u_t = \frac{K}{L} u_{xx} \ .$$

The constant K/L is supposed to be positive. The textbook writes this positive constant by α^2. This derives the book version of the heat equation

$$u_t = \alpha^2 u_{xx} \ . \tag{8}$$

The equations (7) and (8) are the same (except for the names of the coefficients D or α^2). We call them diffusion or heat depending on what is being modeled.

Boundary conditions

The equations (7) or (8) say what happens inside the domain $[0, L]$ but they do not say what happens at the ends. What happens at the ends $x = 0$ and $x = L$ is modeled by *boundary conditions*.

Dirichlet boundary conditions specify the value of u at the ends. It is usually possible to set this value to zero by normalization. Thus

$$u(0,t) = 0 \ \ , \quad u(L,t) = 0 \quad \text{for all } t \ . \tag{9}$$

Neumann boundary conditions specify that there is no flux at the ends. This translates to

$$u_x(0,t) = 0 \ \ , \quad u_x(L,t) = 0 \quad \text{for all } t \ . \tag{10}$$

This would be appropriate for heat flow if the flow were in a one dimensional rod with insulated ends.