LECTURE 2

Dimensional Analysis, Scaling, and Similarity

1. Systems of units

The numerical value of any quantity in a mathematical model is measured with respect to a system of units (for example, meters in a mechanical model, or dollars in a financial model). The units used to measure a quantity are arbitrary, and a change in the system of units (for example, from meters to feet) cannot change the model.

A crucial property of a quantitative system of units is that the value of a dimensional quantity may be measured as some multiple of a basic unit. Thus, a change in the system of units leads to a rescaling of the quantities it measures, and the ratio of two quantities with the same units does not depend on the particular choice of the system. The independence of a model from the system of units used to measure the quantities that appear in it therefore corresponds to a scale-invariance of the model.

Remark 2.1. Sometimes it is convenient to use a logarithmic scale of units instead of a linear scale (such as the Richter scale for earthquake magnitudes, or the stellar magnitude scale for the brightness of stars) but we can convert this to an underlying linear scale. In other cases, qualitative scales are used (such as the Beaufort wind force scale), but these scales ("leaves rustle" or "umbrella use becomes difficult") are not susceptible to a quantitative analysis (unless they are converted in some way into a measurable linear scale). In any event, we will take connection between changes in a system of units and rescaling as a basic premise.

A *fundamental system of units* is a set of independent units from which all other units in the system can be derived. The notion of independent units can be made precise in terms of the rank of a suitable matrix [**7, 10**] but we won't give the details here.

The choice of fundamental units in a particular class of problems is not unique, but, given a fundamental system of units, any other derived unit may be constructed uniquely as a product of powers of the fundamental units.

Example 2.2. In mechanical problems, a fundamental set of units is mass, length, time, or M, L, T, respectively, for short. With this fundamental system, velocity $V = LT^{-1}$ and force $F = MLT^{-2}$ are derived units. We could instead use, say, force F, length L, and time T as a fundamental system of units, and then mass $M = FL^{-1}T^2$ is a derived unit.

Example 2.3. In problems involving heat flow, we may introduce temperature (measured, for example, in Kelvin) as a fundamental unit. The linearity of temperature is somewhat peculiar: although the 'zeroth law' of thermodynamics ensures that equality of temperature is well defined, it does not say how temperatures can be 'added.' Nevertheless, empirical temperature scales are defined, by convention, to be linear scales between two fixed points, while thermodynamics temperature is an energy, which is additive.

Example 2.4. In problems involving electromagnetism, we may introduce current as a fundamental unit (measured, for example, in Ampères in the SI system) or charge (measured, for example, in electrostatic units in the cgs system). Unfortunately, the officially endorsed SI system is often less convenient for theoretical work than the cgs system, and both systems remain in use.

Not only is the distinction between fundamental and derived units a matter of choice, or convention, the number of fundamental units is also somewhat arbitrary. For example, if dimensional constants are present, we may reduce the number of fundamental units in a given system by setting the dimensional constants equal to fixed dimensionless values.

Example 2.5. In relativistic mechanics, if we use M, L, T as fundamental units, then the speed of light c is a dimensional constant ($c = 3 \times 10^8\,\mathrm{ms}^{-1}$ in SI-units). Instead, we may set $c = 1$ and use M, T (for example) as fundamental units. This means that we measure lengths in terms of the travel-time of light (one nanosecond being a convenient choice for everyday lengths).

2. Scaling

Let $(d_1, d_2, \ldots, d_r)$ denote a fundamental system of units, such as (M, L, T) in mechanics, and a a quantity that is measurable with respect to this system. Then the dimension of a, denoted $[a]$, is given by

$$[a] = d_1^{\alpha_1} d_2^{\alpha_2} \ldots d_r^{\alpha_r} \tag{2.1}$$

for suitable exponents $(\alpha_1, \alpha_2, \ldots, \alpha_r)$.

Suppose that $(a_1, a_2, \ldots, a_n)$ denotes all of the dimensional quantities appearing in a particular model, including parameters, dependent variables, and independent variables. We denote the dimension of a_i by

$$[a_i] = d_1^{\alpha_{1,i}} d_2^{\alpha_{2,i}} \ldots d_r^{\alpha_{r,i}}. \tag{2.2}$$

The invariance of the model under a change in units $d_j \mapsto \lambda_j d_j$ implies that it is invariant under the scaling transformation

$$a_i \to \lambda_1^{\alpha_{1,i}} \lambda_2^{\alpha_{2,i}} \ldots \lambda_r^{\alpha_{r,i}} a_i \qquad i = 1, \ldots, n$$

for any $\lambda_1, \ldots \lambda_r > 0$.

Thus, if

$$a = f(a_1, \ldots, a_n)$$

is any relation between quantities in the model with the dimensions in (2.1) and (2.2), then f must have the scaling property that

$$\lambda_1^{\alpha_1} \lambda_2^{\alpha_2} \ldots \lambda_r^{\alpha_r} f(a_1, \ldots, a_n) = f\left(\lambda_1^{\alpha_{1,1}} \lambda_2^{\alpha_{2,1}} \ldots \lambda_r^{\alpha_{r,1}} a_1, \ldots, \lambda_1^{\alpha_{1,n}} \lambda_2^{\alpha_{2,n}} \ldots \lambda_r^{\alpha_{r,n}} a_n\right).$$

A particular consequence of this invariance is that any two quantities that are equal must have the same dimension (otherwise a change in units would violate the equality). This fact is often useful in finding the dimension of some quantity.

Example 2.6. According to Newton's second law,

$$\text{force} = \text{rate of change of momentum with respect to time.}$$

Thus, if F denotes the dimension of force and P the dimension of momentum, then $F = P/T$. Since $P = MV = ML/T$, we conclude that $F = ML/T^2$ (or mass $\times$ acceleration).

3. Nondimensionalization

Scale-invariance implies that we can reduce the number of quantities appearing in a problem by introducing dimensionless quantities.

Suppose that $(a_1, \dots, a_r)$ are a set of quantities whose dimensions form a fundamental system of units. We denote the remaining quantities in the model by $(b_1, \dots, b_m)$, where $r + m = n$. Then, for suitable exponents $(\beta_{1,i}, \dots, \beta_{r,i})$ determined by the dimensions of $(a_1, \dots, a_r)$ and b_i, the quantity

$$\Pi_i = \frac{b_i}{a_1^{\beta_{1,i}} \dots a_r^{\beta_{r,i}}}$$

is dimensionless, meaning that it is invariant under the scaling transformations induced by changes in units.

A dimensionless parameter Π_i can typically be interpreted as the ratio of two quantities of the same dimension appearing in the problem (such as a ratio of lengths, times, diffusivities, and so on). In studying a problem, it is crucial to know the magnitude of the dimensionless parameters on which it depends, and whether they are small, large, or roughly of the order one.

Any dimensional equation

$$a = f(a_1, \dots, a_r, b_1, \dots, b_m)$$

is, after rescaling, equivalent to the dimensionless equation

$$\Pi = f(1, \dots, 1, \Pi_1, \dots, \Pi_m).$$

Thus, the introduction of dimensionless quantities reduces the number of variables in the problem by the number of fundamental units. This fact is called the 'Buckingham Pi-theorem.' Moreover, any two systems with the same values of dimensionless parameters behave in the same way, up to a rescaling.

4. Fluid mechanics

To illustrate the ideas of dimensional analysis, we describe some applications in fluid mechanics.

Consider the flow of a homogeneous fluid with speed U and length scale L. We restrict our attention to incompressible flows, for which U is much smaller that the speed of sound c_0 in the fluid, meaning that the Mach number

$$M = \frac{U}{c_0}$$

is small. The sound speed in air at standard conditions is $c_0 = 340\,\mathrm{ms}^{-1}$. The incompressibility assumption is typically reasonable when $M \leq 0.2$.

The physical properties of a viscous, incompressible fluid depend upon two dimensional parameters, its mass density ρ_0 and its (dynamic) viscosity μ. The dimension of the density is

$$[\rho_0] = \frac{M}{L^3}.$$

The dimension of the viscosity, which measures the internal friction of the fluid, is given by

$$[\mu] = \frac{M}{LT}. \tag{2.3}$$

To derive this result, we explain how the viscosity arises in the constitutive equation of a Newtonian fluid relating the stress and the strain rate.

4.1. The stress tensor

The stress, or force per unit area, $\vec{t}$ exerted across a surface by fluid on one side of the surface on fluid on the other side is given by

$$\vec{t} = \mathbf{T}\vec{n}$$

where $\mathbf{T}$ is the Cauchy stress tensor and $\vec{n}$ is a unit vector to the surface. It is a fundamental result in continuum mechanics, due to Cauchy, that $\vec{t}$ is a linear function of $\vec{n}$; thus, $\mathbf{T}$ is a second-order tensor [**25**].

The sign of $\vec{n}$ is chosen, by convention, so that if $\vec{n}$ points into fluid on one side A of the surface, and away from fluid on the other side B, then $\mathbf{T}\vec{n}$ is the stress exerted by A on B. A reversal of the sign of $\vec{n}$ gives the equal and opposite stress exerted by B on A.

The stress tensor in a Newtonian fluid has the form

$$\mathbf{T} = -p\mathbf{I} + 2\mu\mathbf{D} \tag{2.4}$$

where p is the fluid pressure, μ is the dynamic viscosity, $\mathbf{I}$ is the identity tensor, and $\mathbf{D}$ is the strain-rate tensor

$$\mathbf{D} = \frac{1}{2}\left(\nabla\vec{u} + \nabla\vec{u}^{\top}\right).$$

Thus, $\mathbf{D}$ is the symmetric part of the velocity gradient $\nabla\vec{u}$.

In components,

$$T_{ij} = -p\delta_{ij} + \mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)$$

where δ_{ij} is the Kronecker-δ,

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

Example 2.7. Newton's original definition of viscosity (1687) was for shear flows. The velocity of a shear flow with strain rate σ is given by

$$\vec{u} = \sigma x_2 \vec{e}_1$$

where $\vec{x} = (x_1, x_2, x_3)$ and $\vec{e}_i$ is the unit vector in the i^{th} direction. The velocity gradient and strain-rate tensors are

$$\nabla\vec{u} = \begin{pmatrix} 0 & \sigma & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \mathbf{D} = \frac{1}{2}\begin{pmatrix} 0 & \sigma & 0 \\ \sigma & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The viscous stress $\vec{t}_v = 2\mu\mathbf{D}\vec{n}$ exerted by the fluid in $x_2 > 0$ on the fluid in $x_2 < 0$ across the surface $x_2 = 0$, with unit normal $\vec{n} = \vec{e}_2$ pointing into the region $x_2 > 0$, is $\vec{t}_v = \sigma\mu\vec{e}_1$. (There is also a normal pressure force $\vec{t}_p = -p\vec{e}_1$.) Thus, the frictional viscous stress exerted by one layer of fluid on another is proportional the strain rate σ and the viscosity μ.

4.2. Viscosity

The dynamic viscosity μ is a constant of proportionality that relates the strain-rate to the viscous stress.

Stress has the dimension of force/area, so

$$[\mathbf{T}] = \frac{ML}{T^2}\frac{1}{L^2} = \frac{M}{LT^2}.$$

The strain-rate has the dimension of a velocity gradient, or velocity/length, so

$$[\mathbf{D}] = \frac{L}{T}\frac{1}{L} = \frac{1}{T}.$$

Since $\mu\mathbf{D}$ has the same dimension as $\mathbf{T}$, we conclude that μ has the dimension in (2.3).

The *kinematic viscosity* ν of the fluid is defined by

$$\nu = \frac{\mu}{\rho_0}.$$

It follows from (2.3) that ν has the dimension of a diffusivity,

$$[\nu] = \frac{L^2}{T}.$$

The kinematic viscosity is a diffusivity of momentum; viscous effects lead to the diffusion of momentum in time T over a length scale of the order $\sqrt{\nu T}$.

The kinematic viscosity of water at standard conditions is approximately $1\,\mathrm{mm}^2/\mathrm{s}$, meaning that viscous effects diffuse fluid momentum in one second over a distance of the order 1 mm. The kinematic viscosity of air at standard conditions is approximately $15\,\mathrm{mm}^2/\mathrm{s}$; it is larger than that of water because of the lower density of air. These values are small on every-day scales. For example, the timescale for viscous diffusion across room of width 10 m is of the order of 6×10^6 s, or about 77 days.

4.3. The Reynolds number

The dimensional parameters that characterize a fluid flow are a typical velocity U and length L, the kinematic viscosity ν, and the fluid density ρ_0. Their dimensions are

$$[U] = \frac{L}{T}, \quad [L] = L, \quad [\nu] = \frac{L^2}{T}, \quad [\rho_0] = \frac{M}{L^3}.$$

We can form a single independent dimensionless parameter from these dimensional parameters, the *Reynolds number*

$$\mathrm{R} = \frac{UL}{\nu}. \tag{2.5}$$

As long as the assumptions of the original incompressible model apply, the behavior of a flow with similar boundary and initial conditions depends only on its Reynolds number.

The inertial term in the Navier-Stokes equation has the order of magnitude

$$\rho_0 \vec{u} \cdot \nabla \vec{u} = O\left(\frac{\rho_0 U^2}{L}\right),$$

while the viscous term has the order of magnitude

$$\mu \Delta \vec{u} = O\left(\frac{\mu U}{L^2}\right).$$

The Reynolds number may therefore be interpreted as a ratio of the magnitudes of the inertial and viscous terms.

The Reynolds number spans a large range of values in naturally occurring flows, from 10^{-20} in the very slow flows of the earth's mantle, to 10^{-5} for the motion of bacteria in a fluid, to 10^6 for air flow past a car traveling at 60 mph, to 10^{10} in some large-scale geophysical flows.

Example 2.8. Consider a sphere of radius L moving through an incompressible fluid with constant speed U. A primary quantity of interest is the total drag force D exerted by the fluid on the sphere. The drag is a function of the parameters on which the problem depends, meaning that

$$D = f(U, L, \rho_0, \nu).$$

The drag D has the dimension of force (ML/T^2), so dimensional analysis implies that

$$D = \rho_0 U^2 L^2 F\left(\frac{UL}{\nu}\right).$$

Thus, the dimensionless drag

$$\frac{D}{\rho_0 U^2 L^2} = F(\mathrm{R}) \tag{2.6}$$

is a function of the Reynolds number (2.5), and dimensional analysis reduces the problem of finding a function f of four variables to finding a function F of one variable.

The function $F(\mathrm{R})$ has a complicated dependence on R which is difficult to determine theoretically, especially for large values of the Reynolds number. Nevertheless, experimental measurements of the drag for a wide variety of values of U, L, ρ_0 and ν agree well with (2.6) (see Figure 1.

4.4. The Navier-Stokes equations

The flow of an incompressible homogeneous fluid with density ρ_0 and viscosity μ is described by the incompressible Navier-Stokes equations,

$$\begin{aligned} &\rho_0 \left(\vec{u}_t + \vec{u} \cdot \nabla \vec{u}\right) + \nabla p = \mu \Delta \vec{u}, \\ &\nabla \cdot \vec{u} = 0. \end{aligned} \tag{2.7}$$

Here, $\vec{u}(\vec{x}, t)$ is the velocity of the fluid, and $p(\vec{x}, t)$ is the pressure. The first equation is conservation of momentum, and the second equation is conservation of volume.

Remark 2.9. It remains an open question whether or not the three-dimensional Navier-Stokes equations, with arbitrary smooth initial data and appropriate boundary conditions, have a unique, smooth solution that is defined for all positive times. This is one of the Clay Institute Millenium Prize Problems.

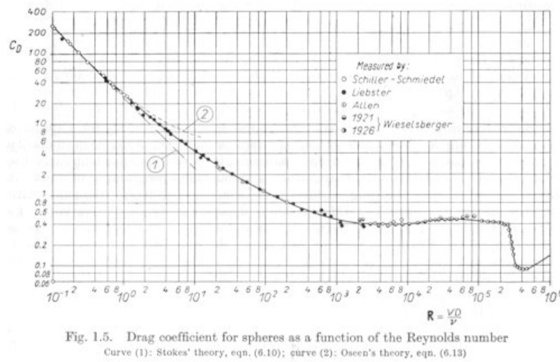

FIGURE 1. Drag on a sphere as a function of Reynold's number (from Schlichting).

Let U, L be a typical velocity scale and length scale of a fluid flow, and define dimensionless variables by

$$\vec{u}^* = \frac{\vec{u}}{U}, \quad p^* = \frac{p}{\rho U^2}, \quad \vec{x}^* = \frac{\vec{x}}{L}, \quad t^* = \frac{Ut}{L}.$$

Using these expressions in (2.7), and dropping the stars on the dimensionless variables, we get

$$\begin{aligned} &\vec{u}_t + \vec{u} \cdot \nabla \vec{u} + \nabla p = \frac{1}{\mathrm{R}} \Delta \vec{u}, \\ &\nabla \cdot \vec{u} = 0, \end{aligned} \tag{2.8}$$

where R is the Reynolds number defined in (2.5).

4.5. Euler equations

The nondimensionalized equation (2.8) suggests that for flows with high Reynolds number, we may neglect the viscous term on the right hand side of the momentum equation, and approximate the Navier-Stokes equation by the *incompressible Euler equations*

$$\begin{aligned} &\vec{u}_t + \vec{u} \cdot \nabla \vec{u} + \nabla p = 0, \\ &\nabla \cdot \vec{u} = 0. \end{aligned}$$

The Euler equations are difficult to analyze because, like the Navier-Stokes equations, they are nonlinear. Moreover, the approximation of the Navier-Stokes equation by the Euler equations is problematic. High-Reynolds number flows develop complicated small-scale structures (for instance, boundary layers and turbulence) and, as a result, it is not always possible to neglect the second-order spatial derivatives $\Delta \vec{u}$ in the viscous term in comparison with the first-order spatial derivatives $\vec{u} \cdot \nabla \vec{u}$ in the inertial term, even though the viscous term is multiplied by a small coefficient.

4.6. Stokes equations

At low Reynolds numbers a different nondimensionalization of the pressure, based on the viscosity rather than the inertia, is appropriate. Using

$$\vec{u}^* = \frac{\vec{u}}{U}, \quad p^* = \frac{p}{\rho U^2}, \quad \vec{x}^* = \frac{\vec{x}}{L}, \quad t^* = \frac{Ut}{L},$$

in (2.7), and dropping the stars on the dimensionless variables, we get

$$\begin{aligned} &\mathrm{R}\left(\vec{u}_t + \vec{u}\cdot\nabla\vec{u}\right) + \nabla p = \Delta\vec{u}, \\ &\nabla\cdot\vec{u} = 0. \end{aligned}$$

Setting $\mathrm{R} = 0$ in these equations, we get the *Stokes equations*,

$$\nabla p = \Delta\vec{u}, \qquad \nabla\cdot\vec{u} = 0. \tag{2.9}$$

These equations provide a good approximation for low Reynolds number flows (although nonuniformities arise in using them on unbounded domains). They are much simpler to analyze than the full Navier-Stokes equations because they are linear.

5. Stokes formula for the drag on a sphere

As an example of the solution of the Stokes equations for low Reynolds number flows, we will derive Stokes' formula (1851) for the drag on a sphere moving at constant velocity through a highly viscous fluid.

It is convenient to retain dimensional variables, so we consider Stokes equations (2.9) in dimensional form

$$\mu\Delta\vec{u} = \nabla p, \qquad \nabla\cdot\vec{u} = 0. \tag{2.10}$$

We note for future use that we can eliminate the pressure from (2.10) by taking the curl of the momentum equation, which gives

$$\Delta\,\mathrm{curl}\,\vec{u} = 0. \tag{2.11}$$

Before considering axisymmetric Stokes flow past a sphere, it is useful to look at the two-dimensional equations. Using Cartesian coordinates with $\vec{x} = (x, y)$ and $\vec{u} = (u, v)$, we may write (2.10) as 3×3 system for (u, v, p):

$$\mu\Delta u = p_x, \qquad \mu\Delta v = p_y, \qquad u_x + v_y = 0.$$

Here, $\Delta = \partial_x^2 + \partial_y^2$ is the two-dimensional Laplacian. In a simply connected region, the incompressibility condition implies that we may introduce a streamfunction $\psi(x, y)$ such that $u = \psi_y$ and $v = -\psi_x$. The momentum equation then becomes

$$\mu\Delta\psi_y = p_x, \qquad \Delta\psi_x = -p_y.$$

The elimination of p by cross-differentiation implies that ψ satisfies the biharmonic equation

$$\Delta^2\psi = 0.$$

Thus, the two-dimensional Stokes equations reduce to the biharmonic equation.

Similar considerations apply to axisymmetric flows, although the details are more complicated. We will therefore give a direct derivation of the solution for flow past a sphere, following Landau and Lifshitz [**34**].

We denote the radius of the sphere by a, and adopt a reference frame moving with the sphere. In this reference frame, the sphere is at rest and the fluid velocity far away from the sphere approaches a constant velocity $\vec{U}$. The pressure also approaches a constant, which we may take to be zero without loss of generality.

The appropriate boundary condition for the flow of a viscous fluid past a solid, impermeable boundary is the no-slip condition that the velocity of the fluid is equal to the velocity of the body. Roughly speaking, this means that a viscous fluid 'sticks' to a solid boundary.

Let $\vec{x}$ denote the position vector from the center of the sphere and $r = |\vec{x}|$ the distance. We want to solve the Stokes equations (2.10) for $\vec{u}(\vec{x})$, $p(\vec{x})$ in the exterior of the sphere $a < r < \infty$, subject to the no-slip condition on the sphere,

$$\vec{u}(\vec{x}) = 0 \qquad \text{at } r = a, \tag{2.12}$$

and the uniform-flow condition at infinity,

$$\vec{u}(\vec{x}) \sim \vec{U}, \quad p(\vec{x}) \to 0 \qquad \text{as } r \to \infty. \tag{2.13}$$

First, we will solve for the velocity. Since $\vec{u}$ is divergence free and the exterior of a sphere is simply connected, we can write it as

$$\vec{u}\,(\vec{x}) = \vec{U} + \operatorname{curl} \vec{A}\,(\vec{x}) \tag{2.14}$$

where $\vec{A}$ is a vector-streamfunction for the deviation of the flow from the uniform flow. We can always choose $\vec{A}$ to be divergence free, and we require that the derivatives of $\vec{A}$ approach zero at infinity so that $\vec{u}$ approaches $\vec{U}$.

We will show that we can obtain the solution by choosing $\vec{A}$ to be of the form

$$\vec{A}\,(\vec{x}) = \nabla f(r) \times \vec{U} \tag{2.15}$$

for a suitable scalar valued function $f(r)$ which we will determine. This form for $\vec{A}$ is dictated by linearity and symmetry considerations: since the Stokes equations are linear, the solution must be linear in the velocity $\vec{U}$; and the solution must be invariant under rotations about the axis parallel to $\vec{U}$ through the center of the sphere, and under rotations of $\vec{U}$.

Using the vector identity

$$\operatorname{curl}\left(f\vec{F}\right) = \nabla f \times \vec{F} + f \operatorname{curl} \vec{F},$$

and the fact that $\vec{U}$ is constant, we may also write (2.15) as

$$\vec{A} = \operatorname{curl}\left(f\vec{U}\right), \tag{2.16}$$

which shows, in particular, that $\nabla \cdot \vec{A} = 0$.

By use of the vector identity

$$\operatorname{curl}\operatorname{curl} \vec{F} = \nabla\left(\nabla \cdot \vec{F}\right) - \Delta \vec{F}, \tag{2.17}$$

and (2.15), we find that

$$\operatorname{curl} \vec{u} = \operatorname{curl}\operatorname{curl} \vec{A} = -\Delta \vec{A} = -\Delta\left(\nabla f \times \vec{U}\right).$$

Using this result in (2.11), we find that

$$\Delta^2\left(\nabla f \times \vec{U}\right) = 0.$$

Since $\vec{U}$ is constant, it follows that

$$\nabla\left(\Delta^2 f\right) \times \vec{U} = 0.$$

Since f depends only on r, this equation implies that $\nabla\left(\Delta^2 f\right) = 0$, so $\Delta^2 f$ is constant. Since the derivatives of $\vec{A}$ decay at infinity, $\Delta^2 f$ must also decay, so the constant is zero, and therefore f satisfies the biharmonic equation

$$\Delta^2 f = 0. \tag{2.18}$$

Writing $g = \Delta f$, which is a function of $r = |\vec{x}|$, and using the expression for the three-dimensional Laplacian in spherical-polar coordinates, we get

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dg}{dr}\right) = 0.$$

Integrating this equation, we get $g(r) = 2b/r + c$ where b, c are constant of integration. Since $\Delta f \to 0$ as $r \to \infty$, we must have $c = 0$, so

$$\Delta f = \frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{df}{dr}\right) = \frac{2b}{r}.$$

Integrating this equation and neglecting an additive constant, which involves no loss of generality because $\vec{A}$ depends only on ∇f, we get

$$f(r) = br + \frac{c}{r} \tag{2.19}$$

where c is another constant of integration.

Using this expression for f in (2.15), then using the result in (2.14), we find that

$$\vec{u}\left(\vec{x}\right) = \vec{U} - \frac{b}{r}\left[\vec{U} + \frac{1}{r^2}\left(\vec{U}\cdot\vec{x}\right)\vec{x}\right] + \frac{c}{r^3}\left[\frac{3}{r^2}\left(\vec{U}\cdot\vec{x}\right)\vec{x} - \vec{U}\right]. \tag{2.20}$$

This velocity field satisfies the boundary condition at infinity (2.13). Imposing the boundary condition (2.12) on the sphere, we get

$$\left(1 - \frac{b}{a} - \frac{c}{a^3}\right)\vec{U} + \frac{1}{a^3}\left(\frac{3c}{a^2} - b\right)\left(\vec{U}\cdot\vec{x}\right)\vec{x} = 0 \qquad \text{when } |\vec{x}| = a.$$

This condition is satisfied only if the coefficient of each term vanishes, which gives

$$b = \frac{3a}{4}, \qquad c = \frac{a^3}{4}.$$

Thus, from (2.19), the solution for f is

$$f(r) = \frac{3ar}{4}\left(1 + \frac{a^2}{3r^2}\right), \tag{2.21}$$

and, from (2.20), the solution for the velocity field is

$$\vec{u}\left(\vec{x}\right) = \left(1 - \frac{3a}{4r} - \frac{a^3}{4r^3}\right)\vec{U} + \frac{1}{r^2}\left(\frac{3a^3}{4r^3} - \frac{3a}{4r}\right)\left(\vec{U}\cdot\vec{x}\right)\vec{x}. \tag{2.22}$$

Remark 2.10. A noteworthy feature of this solution is its slow decay to the uniform flow. The difference

$$\vec{U} - \vec{u}\left(\vec{x}\right) \sim \frac{3a}{4r}\left[\vec{U} + \frac{1}{r^2}\left(\vec{U}\cdot\vec{x}\right)\vec{x}\right]$$

is of the order $1/r$ as $r \to \infty$. This makes the analysis of nondilute suspensions of particles difficult, even with linearity of the Stokes equations, because the hydrodynamic interactions between particles have a long range.

To get the pressure, we first compute $\Delta\vec{u}$. Using (2.16) in (2.14), and applying the vector identity (2.17), we get

$$\begin{aligned}\vec{u} &= \vec{U} + \operatorname{curl}\operatorname{curl}\left(f\vec{U}\right)\\ &= \vec{U} + \nabla\left[\nabla\cdot\left(f\vec{U}\right)\right] - (\Delta f)\,\vec{U}.\end{aligned}$$

Taking the Laplacian of this equation, then using the identity $\nabla\cdot\left(f\vec{U}\right) = \vec{U}\cdot\nabla f$ and the fact that f satisfies the biharmonic equation (2.18), we get

$$\Delta\vec{u} = \nabla\Delta\left(\vec{U}\cdot\nabla f\right).$$

Use of this expression in the momentum equation in (2.10) gives

$$\nabla\left[\mu\Delta\left(\vec{U}\cdot\nabla f\right) - p\right] = 0.$$

It follows that the expression inside the gradient is constant, and from (2.13) the constant is zero. Therefore,

$$p = \mu\Delta\left(\vec{U}\cdot\nabla f\right).$$

Using (2.21) in this equation, we find the explicit expression

$$p = -\left(\frac{3\mu a}{2r^3}\right)\vec{U}\cdot\vec{x}. \tag{2.23}$$

Thus, (2.22) and (2.23) is the solution of (2.10) subject to the boundary conditions (2.13)–(2.12).

A primary quantity of interest is the drag force F exerted by the fluid on the sphere. This force is given by integrating the stress over the surface $\partial\Omega$ of the sphere:

$$\vec{F} = \int_{\partial\Omega}\mathbf{T}\vec{n}\,dS.$$

Here, $\vec{n}$ is the unit outward normal to the sphere, and $\mathbf{T}$ is the Cauchy stress tensor, given from (2.4) by

$$\mathbf{T} = -p\mathbf{I} + \mu\left(\nabla\vec{u} + \nabla\vec{u}^{\top}\right).$$

A direct calculation, whose details we omit, shows that the force is in the direction of $\vec{U}$ with magnitude

$$F = 6\pi\mu a U, \tag{2.24}$$

where U is the magnitude of $\vec{U}$.

This expression for the drag on a spherical particle is found to be in excellent agreement with experimental results if $\mathrm{R} < 0.5$, where

$$\mathrm{R} = \frac{2aU}{\nu}$$

is the Reynolds numbers based on the particle diameter, and $\nu = \mu/\rho_0$ is the kinematic viscosity, as before.

For example, consider a particle of radius a and density ρ_p falling under gravity in a fluid of density ρ_0. At the terminal velocity U, the viscous drag must balance the gravitational buoyancy force, so

$$6\pi\mu a U = \frac{4}{3}\pi a^3\left(\rho_p - \rho_0\right)g$$

where g is the acceleration due to gravity. This gives

$$U = \frac{2a^2 g}{9\nu}\left(\frac{\rho_p}{\rho_0} - 1\right)$$

The corresponding Reynolds number is

$$\mathrm{R} = \frac{4a^3 g}{9\nu^2}\left(\frac{\rho_p}{\rho_0} - 1\right).$$

For a water droplet falling through air [**9**], we have $\rho_p/\rho_0 \approx 780$ and $\nu \approx 15\,\mathrm{mm\,s^{-1}}$. This gives a Reynolds number of approximately $1.5 \times 10^4\, a^3$ where a is measured in mm. Thus, Stokes formula is applicable when $a \leq 0.04\,\mathrm{mm}$, corresponding to droplets in a fine mist.

6. Kolmogorov's 1941 theory of turbulence

> Finally, if we are to list the reasons for studying homogeneous turbulence, we should add that it is a profoundly interesting physical phenomenon which still defies satisfactory mathematical analysis; this is, of course, the most compelling reason.[1]

High-Reynolds number flows typically exhibit an extremely complicated behavior called turbulence. In fact, Reynolds first introduced the Reynolds number in connection with his studies on transition to turbulence in pipe flows in 1895. The analysis and understanding of turbulence remains a fundamental challenge. There is, however, no precise definition of fluid turbulence, and there are many different kinds of turbulent flows, so this challenge is likely to be one with many different parts.

In 1941, Kolmogorov proposed a simple dimensional argument that is one of the basic results about turbulence. To explain his argument, we begin by describing an idealized type of turbulence called homogeneous, isotropic turbulence.

6.1. Homogeneous, isotropic turbulence

Following Batchelor [**8**], let us imagine an infinite extent of fluid in turbulent motion. This means, first, that the fluid velocity depends on a large range of length scales; we denote the smallest length scale (the 'dissipation' length scale) by λ_d and the largest length scale (the 'integral' length scale) by L. And, second, that the fluid motion is apparently random and not reproducible in detail from one experiment to the next.

We therefore adopt a probabilistic description, and suppose that a turbulent flow is described by a probability measure on solutions of the Navier-Stokes equations such that expected values of the fluid variables with respect to the measure agree with appropriate averages of the turbulent flow.

This probabilistic description is sometimes interpreted as follows: we have an 'ensemble' of many different fluid flows — obtained, for example, by repeating the same experiment many different times — and each member of the ensemble corresponds to a flow chosen 'at random' with respect to the probability measure.

A turbulent flow is said to be homogeneous if its expected values are invariant under spatial translations — meaning that, on average, it behaves the same way at each point in space — and isotropic if its expected values are also independent of spatial rotations. Similarly, the flow is stationary if its expected values are invariant

[1] G. K. Batchelor, *The Theory of Homogeneous Turbulence.*

under translations in time. Of course, any particular realization of the flow varies in space and time.

Homogeneous, isotropic, stationary turbulence is rather unphysical. Turbulence is typically generated at boundaries, and the properties of the flow vary with distance from the boundary or other large-scale features of the flow geometry. Moreover, turbulence dissipates energy at a rate which appears to be nonzero even in the limit of infinite Reynolds number. Thus, some sort of forcing (usually at the integral length scale) that adds energy to the fluid is required to maintain stationary turbulence. Nevertheless, appropriate experimental configurations (for example, high-Reynolds number flow downstream of a metal grid) and numerical configurations (for example, direct numerical simulations on a 'box' with periodic boundary conditions and a suitable applied force) provide a good approximation to homogeneous, isotropic turbulence.

6.2. Correlation functions and the energy spectrum

We denote expected values by angular brackets $\langle\cdot\rangle$. In a homogeneous flow the two-point correlation

$$Q = \langle \vec{u}(\vec{x},t)\cdot\vec{u}(\vec{x}+\vec{r},t)\rangle \tag{2.25}$$

is a function of the spatial displacement $\vec{r}$, and independent of $\vec{x}$. In a stationary flow it is independent of t. Furthermore, in an isotropic flow, Q is a function only of the magnitude $r = |\vec{r}|$ of the displacement vector.

Note, however, that even in isotropic flow the general correlation tensor

$$\mathbf{Q}(\vec{r}) = \langle \vec{u}(\vec{x},t)\otimes\vec{u}(\vec{x}+\vec{r},t)\rangle\,,$$

with components $Q_{ij} = \langle u_i u_j\rangle$, depends on the vector $\vec{r}$, not just its magnitude, because a rotation of $\vec{r}$ also induces a rotation of $\vec{u}$.

For isotropic turbulence, one can show [**8**] that the two-point correlation (2.25) has the Fourier representation

$$Q(r) = 2\int_0^\infty \frac{\sin kr}{kr}\,E(k)\,dk$$

where $E(k)$ is a nonnegative function of the wavenumber magnitude k.

In particular, it follows that

$$\frac{1}{2}\langle \vec{u}(\vec{x},t)\cdot\vec{u}(\vec{x},t)\rangle = \int_0^\infty E(k)\,dk. \tag{2.26}$$

Thus, $E(k)$ may be interpreted as the mean kinetic energy density of the turbulent flow as a function of the wavenumber $0 \le k < \infty$.

6.3. The five-thirds law

In fully developed turbulence, there is a wide range of length scales $\lambda_d \ll \lambda \ll L$ that are much greater than the dissipation length scale and much less than the integral length scale. This range is called the inertial range. The corresponding wavenumbers are $k = 2\pi/\lambda$, with dimension

$$[k] = \frac{1}{L}.$$

It appears reasonable to assume that the components of a turbulent flow which vary over length scales in the inertial range do not depend on the viscosity ν of the fluid or on the integral length scale and velocity.

Kolmogorov proposed that, in the inertial range, the flow statistics depend only on the mean rate per unit mass ϵ at which the turbulent flow dissipates energy. It would not make any difference if we used instead the mean rate of energy dissipation per unit volume, since we would have to nondimensionalize this by the fluid density, to get the mean energy dissipation rate per unit mass. The dimension of this rate is is

$$[\epsilon] = \frac{ML^2}{T^2} \cdot \frac{1}{T} \cdot \frac{1}{M} = \frac{L^2}{T^3}.$$

From (2.26), the spectral energy density has dimension

$$[E(k)] = \frac{L^3}{T^2}.$$

If the only quantities on which $E(k)$ depends are the energy dissipation rate ϵ and the wavenumber k itself, then, balancing dimensions, we must have

$$E(k) = C\epsilon^{2/3}k^{-5/3}, \tag{2.27}$$

where C is a dimensionless constant, called the Kolmogorov constant.

Thus, Kolmogorov's 1941 (K41) theory predicts that the energy spectrum of a turbulent flow in the inertial range has a power-law decay as a function of wavenumber with exponent $-5/3$; this is the "five-thirds law."

The spectral result (2.27) was, in fact, first stated by Oboukhov (1941). Kolmogorov gave a closely related result for spatial correlations:

$$\left\langle \left|\vec{u}\left(\vec{x}+\vec{r},t\right) - \vec{u}\left(\vec{x},t\right)\right|^2 \right\rangle = C\epsilon^{2/3}r^{2/3}.$$

This equation suggests that the velocity of a turbulent flow has a 'rough' spatial dependence in the inertial range, similar to that of a non-differentiable Hölder-continuous function with exponent 1/3.

Onsager rediscovered this result in 1945, and in 1949 suggested that turbulent dissipation might be described by solutions of the Euler equation that are not sufficiently smooth to conserve energy [**20**]. The possible relationship of non-smooth, weak solutions of the incompressible Euler equations (which are highly non-unique and can even increase in kinetic energy without some kind of additional admissibility conditions) to turbulent solutions of the Navier-Stokes equations remains unclear.

6.4. The Kolmogorov length scale

The only length scale that can be constructed from the dissipation rate ϵ and the kinematic viscosity ν, called the Kolmogorov length scale, is

$$\eta = \left(\frac{\nu^3}{\epsilon}\right)^{1/4}.$$

The K41 theory implies that the dissipation length scale is of the order η.

If the energy dissipation rate is the same at all length scales, then, neglecting order one factors, we have

$$\epsilon = \frac{U^3}{L}$$

where L, U are the integral length and velocity scales. Denoting by R_L the Reynolds number based on these scales,

$$\mathrm{R}_L = \frac{UL}{\nu},$$

it follows that

$$\frac{L}{\eta} = \mathrm{R}_L^{3/4}.$$

Thus, according to this dimensional analysis, the ratio of the largest (integral) length scale and the smallest (dissipation) length scale grows like $\mathrm{R}_L^{3/4}$ as $\mathrm{R}_L \to \infty$.

In order to resolve the finest length scales of a three-dimensional flow with integral-scale Reynolds number R_L, we therefore need on the order of

$$N_L = \mathrm{R}_L^{9/4}$$

independent degrees of freedom (for example, N_L Fourier coefficients of the velocity components). The rapid growth of N_L with R_L limits the Reynolds numbers that can be attained in direct numerical simulations of turbulent flows.

6.5. Validity of the five-thirds law

Experimental observations, such as those made by Grant, Stewart and Moilliet (1962) in a tidal channel between islands off Vancouver, agree well with the five-thirds law for the energy spectrum, and give $C \approx 1.5$ in (2.27). The results of DNS on periodic 'boxes', using up to 4096^3 grid points, are also in reasonable agreement with this prediction.

Although the energy spectrum predicted by the K41 theory is close to what is observed, there is evidence that it is not exactly correct. This would imply that there is something wrong with its original assumptions.

Kolmogorov and Oboukhov proposed a refinement of Kolmogorov's original theory in 1962. It is, in particular, not clear that the energy dissipation rate ϵ should be assumed constant, since the energy dissipation in a turbulent flow itself varies over multiple length scales in a complicated fashion. This phenomenon, called 'intermittency,' can lead to corrections in the five-thirds law [**23**]. All such turbulence theories, however, depend on some kind of initial assumptions whose validity can only be checked by comparing their predictions with experimental or numerical observations.

6.6. The benefits and drawbacks of dimensional arguments

As the above examples from fluid mechanics illustrate, dimensional arguments can lead to surprisingly powerful results, even without a detailed analysis of the underlying equations. All that is required is a knowledge of the quantities on which the problem being studied depends together with their dimensions. This does mean, however, one has to know the basic laws that govern the problem, and the dimensional constants they involve. Thus, contrary to the way it sometimes appears, dimensional analysis does not give something for nothing; it can only give what is put in from the start.

This fact cuts both ways. Many of the successes of dimensional analysis, such as Kolmogorov's theory of turbulence, are the result of an insight into which dimensional parameters play an crucial role in a problem and which parameters can be ignored. Such insights typical depend upon a great deal of intuition and experience, and they may be difficult to justify or prove.[2]

[2] As Bridgeman ([**10**], p. 5) puts it in his elegant 1922 book on dimensional analysis (well worth reading today): "The untutored savage in the bushes would probably not be able to apply the methods of dimensional analysis to this problem and obtain results that would satisfy us." Hopefully, whatever knowledge may have been lost since then in the area of dimensional analysis has been offset by some gains in cultural sensitivity.

Conversely, it may happen that some dimensional parameters that appear to be so small they can be neglected have a significant effect, in which case scaling laws derived on the basis of dimensional arguments that ignore them are likely to be incorrect.

7. Self-similarity

If a problem depends on more fundamental units than the number of dimensional parameters, then we must use the independent or dependent variables themselves to nondimensionalize the problem. For example, we did this when we used the wavenumber k to nondimensionalize the K41 energy spectrum $E(k)$ in (2.27). In that case, we obtain self-similar solutions that are invariant under the scaling transformations induced by a change in the system of units. For example, in a time-dependent problem the spatial profile of a solution at one instant of time might be a rescaling of the spatial profile at any other time.

These self-similar solutions are often among the few solutions of nonlinear equations that can be obtained analytically, and they can provide valuable insight into the behavior of general solutions. For example, the long-time asymptotics of solutions, or the behavior of solutions at singularities, may be given by suitable self-similar solutions.

As a first example, we use dimensional arguments to find the Green's function of the heat equation.

7.1. The heat equation

Consider the following IVP for the Green's function of the heat equation in $\mathbb{R}^d$:

$$\begin{aligned} u_t &= \nu \Delta u, \\ u(x,0) &= E\delta(x). \end{aligned}$$

Here δ is the delta-function, representing a unit point source at the origin. Formally, we have

$$\int_{\mathbb{R}^d} \delta(x)\, dx = 1, \qquad \delta(x) = 0 \quad \text{for } x \neq 0.$$

The dimensioned parameters in this problem are the diffusivity ν and the energy E of the point source. The only length and times scales are those that come from the independent variables (x,t), so the solution is self-similar.

We have $[u] = \theta$, where θ denotes a unit of temperature. Furthermore, since

$$\int_{\mathbb{R}^d} u(x,0)\, dx = E,$$

we have $[E] = \theta L^d$. The rotational invariance of the Laplacian, and the uniqueness of the solution, implies that the solution must be spherically symmetric. Dimensional analysis then gives

$$u(x,t) = \frac{E}{(\nu t)^{d/2}} f\left(\frac{|x|}{\sqrt{\nu t}}\right).$$

Using this expression for $u(x,t)$ in the PDE, we get an ODE for $f(\xi)$,

$$f'' + \left(\frac{\xi}{2} + \frac{d-1}{\xi}\right) f' + \frac{d}{2} f = 0.$$

We can rewrite this equation as a first-order ODE for $f' + \frac{\xi}{2} f$,

$$\left(f' + \frac{\xi}{2} f\right)' + \frac{d-1}{\xi}\left(f' + \frac{\xi}{2} f\right) = 0.$$

Solving this equation, we get

$$f' + \frac{\xi}{2} f = \frac{b}{\xi^{d-1}},$$

where b is a constant of integration. Solving for f, we get

$$f(\xi) = ae^{-\xi^2/4} + be^{-\xi^2/4} \int \frac{e^{-\xi^2}}{\xi^{d-1}}\, d\xi,$$

where a s another constant of integration. In order for f to be integrable, we must set $b = 0$. Then

$$u(x,t) = \frac{aE}{(\nu t)^{d/2}} \exp\left(-\frac{|x|^2}{4\nu t}\right).$$

Imposing the requirement that

$$\int_{\mathbb{R}^d} u(x,t)\, dx = E,$$

and using the standard integral

$$\int_{\mathbb{R}^d} \exp\left(-\frac{|x|^2}{2c}\right)\, dx = (2\pi c)^{d/2},$$

we find that $a = (4\pi)^{-d/2}$, and

$$u(x,t) = \frac{E}{(4\pi\nu t)^{d/2}} \exp\left(-\frac{|x|^2}{4\nu t}\right).$$

8. The porous medium equation

In this section, we will further illustrate the use of self-similar solutions by describing a problem for point-source solutions of the porous medium equation, taken from Barenblatt [**7**]. This solution is a one-dimensional version of the radially symmetric self-similar solution of the porous medium equation

$$u_t = \nabla \cdot (u \nabla u)$$

found by Zeldovich and Kompaneets (1950) and Barenblatt (1952).

We consider the flow under gravity of an incompressible fluid in a porous medium, such as groundwater in a rock formation. We suppose that the porous medium sits above a horizontal impermeable stratum, and, to simplify the discussion, that the flow is two-dimensional. It is straightforward to treat three-dimensional flows in a similar way.

Let x and z denote horizontal and vertical spatial coordinates, respectively, where the impermeable stratum is located at $z = 0$. Suppose that the porous medium is saturated with fluid for $0 \le z \le h(x,t)$ and dry for $z > h(x,t)$. If the wetting front $z = h(x,t)$ has small slope, we may use a quasi-one dimensional approximation in which we neglect the z-velocity components of the fluid and average x-velocity components with respect to z.

The volume of fluid (per unit length in the transverse y-direction) in $a \leq x \leq b$ is given by

$$\int_a^b nh(x,t)\,dx$$

where n is the porosity of the medium. That is, n is the ratio of the open volume in the medium that can be occupied by fluid to the total volume. Typical values of n are 0.3–0.7 for clay, and 0.01, or less, for dense crystalline rocks. We will assume that n is constant, in which case it will cancel from the equations.

Let $u(x,t)$ denote the depth-averaged x-component of the fluid velocity. Conservation of volume for an incompressible fluid implies that for any x-interval $[a,b]$

$$\frac{d}{dt}\int_a^b nh(x,t)\,dx = -\left[nhu\right]_a^b.$$

In differential form, we get

$$h_t = -\left(hu\right)_x \tag{2.28}$$

For slow flows, we can assume that the pressure p in the fluid is equal to the hydrostatic pressure

$$p = \rho_0 g\left(h - z\right).$$

It follows that the total pressure 'head', defined by

$$\frac{p}{\rho_0 g} + z,$$

is independent of z and equal to $h(x,t)$.

According to Darcy's law, the volume-flux (or velocity) of a fluid in a porous medium is proportional to the gradient of the pressure head, meaning that

$$u = -kh_x, \tag{2.29}$$

where k is the permeability, or hydraulic conductivity, of the porous medium.

Remark 2.11. Henri Darcy was a French water works engineer. He published his law in 1856 after conducting experiments on the flow of water through columns of sand, which he carried out in the course of investigating fountains in Dijon.

The permeability k in (2.29) has the dimension of $L^2/(HT)$, where H is the dimension of the head h. Since we measure h in terms of vertical height, k has the dimension of velocity. Typical values of k for consolidated rocks range from 10^{-9} m/day for unfractured metamorphic rocks, to 10^3 m/day for karstic limestone.

Using (2.29) in (2.28), we find that $h(x,t)$ satisfies the *porous medium equation*

$$h_t = k\left(hh_x\right)_x. \tag{2.30}$$

We may alternatively write (2.30) as

$$h_t = \frac{1}{2}k\left(h^2\right)_{xx}.$$

This equation was first considered by Boussinesq (1904).

The porous medium equation is an example of a degenerate diffusion equation. It has a nonlinear diffusivity equal to kh which vanishes when $h = 0$. As we will see, this has the interesting consequence that (2.30) has solutions (corresponding to wetting fronts) that propagate into a region with $h = 0$ at finite speed — behavior one would expect of a wave equation, but not at first sight of a diffusion equation.

8.1. A point source solution

Consider a solution $h(x,t)$ of the porous medium equation (2.30) that approaches an initial point source:

$$h(x,t) \to I\delta(x), \qquad t \to 0^+,$$

where $\delta(x)$ denotes the Dirac delta 'function.' Explicitly, this means that we require

$$\begin{aligned} & h(x,t) \to 0 \qquad \text{as } t \to 0^+ \text{ if } x \neq 0, \\ & \lim_{t\to 0^+} \int_{-\infty}^{\infty} h(x,t)\, dx = I. \end{aligned} \tag{2.31}$$

The delta-function is a distribution, rather than a function. We will not discuss distribution theory here (see [**44**] for an introduction and [**27**] for a detailed account). Instead, we will define the delta-function formally as a 'function' δ with the properties that

$$\begin{aligned} & \delta(x) = 0 \qquad \text{for } x \neq 0, \\ & \int_{-\infty}^{\infty} f(x)\delta(x)\, dx = f(0) \end{aligned}$$

for any continuous function f.

The solution of the porous medium with the initial data (2.31) describes the development of of wetting front due to an instantaneous 'flooding' at the origin by a volume of water I. It provides the long time asymptotic behavior of solutions of the porous medium equation with a concentrated non-point source initial condition $h(x,t) = h_0(x)$ where h_0 is a compactly supported function with integral I.

The dimensional parameters at our disposal in solving this problem are k and I. A fundamental system of units is L, T, H where H is a unit for the pressure head. Since we measure the pressure head h in units of length, it is reasonable to ask why we should use different units for h and x. The explanation is that the units of vertical length used to measure the head play a different role in the model than the units used to measure horizontal lengths, and we should be able to rescale x and z independently.

Equating the dimension of different terms in (2.30), we find that

$$[k] = \frac{L^2}{HT}, \qquad [I] = LH.$$

Since we assume that the initial data is a point source, which does not define a length scale, there are no other parameters in the problem.

Two parameters k, I are not sufficient to nondimensionalize a problem with three fundamental units. Thus, we must also use one of the variables to do so. Using t, we get

$$\left[(kIt)^{1/3}\right] = L, \qquad [t] = T, \qquad \left[\frac{I^{2/3}}{(kt)^{1/3}}\right] = H$$

Dimensional analysis then implies that

$$h(x,t) = \frac{I^{2/3}}{(kt)^{1/3}} F\left(\frac{x}{(kIt)^{1/3}}\right)$$

where $F(\xi)$ is a dimensionless function.

Using this similarity form in (2.30), we find that $F(\xi)$ satisfies the ODE

$$-\frac{1}{3}\left(\xi F' + F\right) = \left(FF'\right)'. \tag{2.32}$$

Furthermore, (2.31) implies that

$$F(\xi) \to 0 \qquad \text{as } |\xi| \to \infty, \tag{2.33}$$

$$\int_{-\infty}^{\infty} F(\xi)\, d\xi = 1. \tag{2.34}$$

Integrating (2.32), we get

$$-\frac{1}{3}\xi F + C = FF' \tag{2.35}$$

where C is a constant of integration.

The condition (2.33) implies that $C = 0$. It then follows from (2.35) that either $F = 0$, or

$$F' = -\frac{1}{3}\xi,$$

which implies that

$$F(\xi) = \frac{1}{6}\left(a^2 - \xi^2\right)$$

where a is a constant of integration.

In order to get a solution that is continuous and approaches zero as $|\xi| \to \infty$, we choose

$$F(\xi) = \begin{cases} \left(a^2 - \xi^2\right)/6 & \text{if } |\xi| < a, \\ 0 & \text{if } |\xi| \geq a. \end{cases}$$

The condition (2.34) then implies that

$$a = \left(\frac{9}{2}\right)^{1/3}.$$

Thus, the solution of (2.30)–(2.31) is given by

$$h(x,t) = \frac{I^{2/3}}{6\,(kt)^{1/3}}\left[\left(\frac{9}{2}\right)^{2/3} - \frac{x^2}{(kIt)^{2/3}}\right] \qquad \text{if } |x| < (9kIt/2)^{1/3}$$

with $h(x,t) = 0$ otherwise.

This solution represents a saturated region of finite width which spreads out at finite speed. The solution is not a classical solution of (2.30) since its derivative h_x has a jump discontinuity at $x = \pm(9kIt/2)^{1/3}$. It can be understood as a weak solution in an appropriate sense, but we will not discuss the details here.

The fact that the solution has length scale proportional to $t^{1/3}$ after time t could have been predicted in advance by dimensional analysis, since $L = (kIt)^{1/3}$ is the only horizontal length scale in the problem. The numerical factors, and the fact that the solution has compact support, depend upon the detailed analytical properties of the porous medium equation equation; they could not be shewn by dimensional analysis.

8.2. A pedestrian derivation

Let us consider an alternative method for finding the point source solution that does not require dimensional analysis, but is less systematic.

First, we remove the constants in (2.30) and (2.31) by rescaling the variables. Defining

$$u(x, \bar{t}) = \frac{1}{I} h(x, t), \qquad \bar{t} = kIt,$$

and dropping the bars on $\bar{t}$, we find that $u(x, t)$ satisfies

$$u_t = (uu_x)_x. \tag{2.36}$$

The initial condition (2.31) becomes

$$\begin{aligned} &u(x, t) \to 0 \qquad \text{as } t \to 0^+ \text{ if } x \neq 0, \\ &\lim_{t \to 0^+} \int_{-\infty}^{\infty} u(x, t)\, dx = 1. \end{aligned} \tag{2.37}$$

We seek a similarity solution of (2.36)–(2.37) of the form

$$u(x, t) = \frac{1}{t^m} f\left(\frac{x}{t^n}\right) \tag{2.38}$$

for some exponents m, n. In order to obtain such a solution, the PDE for $u(x, t)$ must reduce to an ODE for $f(\xi)$. As we will see, this is the case provided that m, n are chosen appropriately.

Remark 2.12. Equation (2.38) is a typical form of a self-similar solution that is invariant under scaling transformations, whether or not they are derived from a change in units. Dimensional analysis of this problem allowed us to deduce that the solution is self-similar. Here, we simply seek a self-similar solution of the form (2.38) and hope that it works.

Defining the similarity variable

$$\xi = \frac{x}{t^n}$$

and using a prime to denote the derivative with respect to ξ, we find that

$$\begin{aligned} u_t &= -\frac{1}{t^{m+1}} \left(mf + n\xi f'\right) \\ (uu_x)_x &= \frac{1}{t^{2m+2n}} \left(ff'\right)'. \end{aligned} \tag{2.39}$$

In order for (2.36) to be consistent with (2.38), the powers of t in (2.39) must agree, which implies that

$$m + 2n = 1. \tag{2.40}$$

In that case, $f(\xi)$ satisfies the ODE

$$\left(ff'\right)' + n\xi f' + mf = 0.$$

Thus, equation (2.36) has a one-parameter family of self-similar solutions. The ODE for similarity solutions is easy to integrate when $m = n$, but it is not as simple to solve when $n \neq m$.

To determine the value of m, n for the point source problem, we compute that, for solutions of the form (2.38),

$$\int_{-\infty}^{\infty} u(x,t)\,dx = t^{n-m}\int_{-\infty}^{\infty} f\left(\xi\right)\,d\xi.$$

Thus, to get a nonzero, finite limit as $t \to 0^+$ in (2.37), we must take $m = n$, and then (2.40) implies that $m = n = 1/3$. We therefore recover the same solution as before.

8.3. Scaling invariance

Let us consider the scaling invariances of the porous medium equation (2.36) in more detail.

We consider a rescaling of the independent and dependent variables given by

$$\tilde{x} = \alpha x, \qquad \tilde{t} = \beta t, \qquad \tilde{u} = \mu u \tag{2.41}$$

where α, β, μ are positive constants. Writing u in terms of $\tilde{u}$ in (2.36) and using the transformation of derivatives

$$\partial_x = \alpha\partial_{\tilde{x}}, \qquad \partial_t = \beta\partial_{\tilde{t}},$$

we find that $\tilde{u}\left(\tilde{x},\tilde{t}\right)$ satisfies the PDE

$$\tilde{u}_{\tilde{t}} = \frac{\alpha^2}{\beta\mu}\left(\tilde{u}\tilde{u}_{\tilde{x}}\right)_{\tilde{x}}.$$

Thus, the rescaling (2.41) leaves (2.36) invariant if $\alpha^2 = \beta\mu$.

To reformulate this invariance in a more geometric way, let $E = \mathbb{R}^2 \times \mathbb{R}$ be the space with coordinates (x,t,u). For $\alpha, \beta > 0$ define the transformation

$$g\left(\alpha,\beta\right): E \to E, \qquad g(\alpha,\beta) : (x,t,u) \mapsto \left(\alpha x, \beta t, \frac{\alpha^2}{\beta}u\right). \tag{2.42}$$

Then

$$G = \{g(\alpha,\beta) : \alpha,\beta > 0\} \tag{2.43}$$

forms a two-dimensional Lie group of transformations of E:

$$g(1,1) = I, \qquad g^{-1}\left(\alpha,\beta\right) = g\left(\frac{1}{\alpha},\frac{1}{\beta}\right),$$
$$g\left(\alpha_1,\beta_1\right)g\left(\alpha_2,\beta_2\right) = g\left(\alpha_1\alpha_2,\beta_1\beta_2\right)$$

where I denotes the identity transformation.

The group G is commutative (in general, Lie groups and symmetry groups are not commutative) and is generated by the transformations

$$(x,t,u) \mapsto \left(\alpha x, t, \alpha^2 u\right), \qquad (x,t,u) \mapsto \left(x, \beta t, \frac{1}{\beta}u\right). \tag{2.44}$$

Abusing notation, we use the same symbol to denote the coordinate u and the function $u(x,t)$. Then the action of $g(\alpha,\beta)$ in (2.42) on $u(x,t)$ is given by

$$u(x,t) \mapsto \frac{\alpha^2}{\beta}u\left(\frac{x}{\alpha},\frac{t}{\beta}\right). \tag{2.45}$$

This map transforms solutions of (2.36) into solutions. Thus, the group G is a symmetry group of (2.36), which consist of the symmetries that arise from dimensional analysis and the invariance of (2.36) under rescalings of its units.

8.4. Similarity solutions

In general, a solution of an equation is mapped to a different solution by elements of a symmetry group. A *similarity solution* is a solution that is mapped to itself by a nontrivial subgroup of symmetries. In other words, it is a fixed point of the subgroup. Let us consider the case of similarity solutions of one-parameter subgroups of scaling transformations for the porous medium equation; we will show that these are the self-similar solutions considered above.

The one-parameter subgroups of G in (2.43) are given by

$$H_n = \{g(\beta^n, \beta) : \beta > 0\} \qquad \text{for } -\infty < n < \infty,$$

and $\{g(\alpha, 1) : \alpha > 0\}$. From (2.45), a function $u(x,t)$ is invariant under H_n if

$$u(x,t) = \beta^{2n-1} u\left(\frac{x}{\beta^n}, \frac{t}{\beta}\right)$$

for every $\beta > 0$. Choosing $\beta = t$, we conclude that $u(x,t)$ has the form (2.38) with $m = 1 - 2n$. Thus, we recover the self-similar solutions considered previously.

8.5. Translational invariance

A transformation of the space E of dependent and independent variables into itself is called a *point transformation*. The group G in (2.43) does not include all the point transformations that leave (2.36) invariant. In addition to the scaling transformations (2.44), the space-time translations

$$(x,t,u) \mapsto (x - \delta, t, u), \qquad (x,t,u) \mapsto (x, t - \varepsilon, u), \tag{2.46}$$

where $-\infty < \delta, \varepsilon < \infty$, also leave (2.36) invariant, because the terms in the equation do not depend explicitly on (x,t).

As we will show in Section 9.8, the transformations (2.44) and (2.46) generate the full group of point symmetries of (2.36). Thus, the porous medium equation does not have any point symmetries beyond the obvious scaling and translational invariances. This is not always the case, however. Many equations have point symmetries that would be difficult to find without using the theory of Lie algebras.

Remark 2.13. The one-dimensional subgroups of the two-dimensional group of space-time translations are given by

$$(x,t,u) \mapsto (x - c\varepsilon, t - \varepsilon, u),$$

where c is a fixed constant (and also the space translations $(x,t,u) \mapsto (x - \varepsilon, t, u)$). The similarity solutions that are invariant under this subgroup are the traveling wave solutions

$$u(x,t) = f(x - ct).$$

9. Continuous symmetries of differential equations

Dimensional analysis leads to scaling invariances of a differential equation. As we have seen in the case of the porous medium equation, these invariances form a continuous group, or Lie group, of symmetries of the differential equation.

The theory of Lie groups and Lie algebras provides a systematic method to compute all continuous point symmetries of a given differential equation; in fact, this is why Lie first introduced the the theory of Lie groups and Lie algebras.

Lie groups and algebras arise in many other contexts. In particular, as a result of the advent of quantum mechanics in the early 20^{th}-century, where symmetry considerations are crucial, Lie groups and Lie algebras have become a central part of mathematical physics.

We will begin by describing some basic ideas about Lie groups of transformations and their associated Lie algebras. Then we will describe their application to the computation of symmetry groups of differential equations. See Olver [**40, 41**], whose presentation we follow, for a full account.

9.1. Lie groups and Lie algebras

A manifold of dimension d is a space that is locally diffeomorphic to $\mathbb{R}^d$, although its global topology may be different (think of a sphere, for example). This means that the elements of the manifold may, locally, be smoothly parametrized by d coordinates, say $\left(\varepsilon^1, \varepsilon^2, \dots, \varepsilon^d\right) \in \mathbb{R}^d$. A Lie group is a space that is both a manifold and a group, such that the group operations (composition and inversion) are smooth functions.

Lie groups almost always arise in applications as transformation groups acting on some space. Here, we are interested in Lie groups of symmetries of a differential equation that act as point transformations on the space whose coordinates are the independent and dependent variables of the differential equation.

The key idea we want to explain first is this: the Lie algebra of a Lie group of transformations is represented by the vector fields whose flows are the elements of the Lie Group. As a result, elements of the Lie algebra are often referred to as 'infinitesimal generators' of elements of the Lie group.

Consider a Lie group G acting on a vector space E. In other words, each $g \in G$ is a map $g : E \to E$. Often, one considers Lie groups of linear maps, which are a subgroup of the general linear group $GL(E)$, but we do not assume linearity here.

Suppose that $E = \mathbb{R}^n$, and write the coordinates of $x \in E$ as $\left(x^1, x^2, \dots, x^n\right)$. We denote the unit vectors in the coordinate directions by

$$\partial_{x^1}, \quad \partial_{x^2}, \quad \dots, \quad \partial_{x^n}.$$

That is, we identify vectors with their directional derivatives.

Consider a vector field

$$\vec{v}(x) = \xi^i(x)\partial_{x^i},$$

where we use the summation convention in which we sum over repeated upper and lower indices. The associated flow is a one-parameter group of transformations obtained by solving the system of ODEs

$$\frac{dx^i}{d\varepsilon} = \xi^i\left(x^1, x^2, \dots, x^n\right) \qquad \text{for } 1 \le i \le n.$$

Explicitly, if $x(\varepsilon)$ is a solution of this ODE, then the flow $g(\varepsilon) : x(0) \mapsto x(\varepsilon)$ maps the initial data at $\varepsilon = 0$ to the solution at 'time' ε.

We denote the flow $g(\varepsilon)$ generated by the vector field $\vec{v}$ by

$$g(\varepsilon) = e^{\varepsilon\vec{v}}.$$

Conversely, given a flow $g(\varepsilon)$, we can recover the vector field that generates it from

$$\vec{v}(x) = \left.\frac{d}{d\varepsilon}g(\varepsilon)\cdot x\right|_{\varepsilon=0}.$$

That is, $\vec{v}(x)$ is the tangent, or velocity, vector of the solution curve through x.

Example 2.14. A linear vector field has the form

$$\vec{v}(x) = a^i_j x^j \partial_{x^i}.$$

Its flow is given by the usual exponential $e^{\varepsilon\vec{v}} = e^{\varepsilon A}$ where the linear transformation A has matrix $\left(a^i_j\right)$.

The flow $e^{\varepsilon\vec{v}}$ of a smooth linear vector field $\vec{v}$ is defined for all $-\infty < \varepsilon < \infty$. The flow of a nonlinear vector field may exists only for sufficiently small values of ε, which may depend on the initial data. In that case we get a local Lie group of flows. Since we only use local considerations here, we will ignore this complication.

9.2. The Lie bracket

In general, a Lie algebra $\mathfrak{g}$ is a vector space with a bilinear, skew-symmetric bracket operation

$$[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$$

that satisfies the Jacobi identity

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0.$$

The Lie bracket of vector fields $\vec{v}$, $\vec{w}$ is defined by their commutator

$$[\vec{v}, \vec{w}] = \vec{v}\vec{w} - \vec{w}\vec{v},$$

where the vector fields are understood as differential operators. Explicitly, if

$$\vec{v} = \xi^i \partial_{x^i}, \qquad \vec{w} = \eta^i \partial_{x^i},$$

then

$$[\vec{v}, \vec{w}] = \left(\xi^j \frac{\partial \eta^i}{\partial x^j} - \eta^j \frac{\partial \xi^i}{\partial x^j}\right) \partial_{x^i}.$$

The Lie bracket of vector fields measures the non-commutativity of the corresponding flows:

$$[\vec{v}, \vec{w}]\,(x) = \frac{1}{2}\,\frac{d^2}{d\varepsilon^2}\left(e^{\varepsilon\vec{v}} e^{\varepsilon\vec{w}} e^{-\varepsilon\vec{v}} e^{-\varepsilon\vec{w}}\right) x\Big|_{\varepsilon=0}.$$

One can show that the Lie bracket of any two vector field that generate elements of a Lie group of transformations also generates an element of the Lie group. Thus, the infinitesimal generators of the Lie group form a Lie algebra.

9.3. Transformations of the plane

As simple, but useful, examples of Lie transformation groups and their associated Lie algebras, let us consider some transformations of the plane.

The rotations of the plane $g(\varepsilon) : \mathbb{R}^2 \to \mathbb{R}^2$ are given by

$$g(\varepsilon) : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \cos\varepsilon & -\sin\varepsilon \\ \sin\varepsilon & \cos\varepsilon \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \qquad \text{where } \varepsilon \in \mathbb{T}.$$

These transformations form a representation of the one-dimensional Lie group $SO(2)$ on $\mathbb{R}^2$. They are the flow of the ODE

$$\frac{d}{d\varepsilon}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix}.$$

The vector field on the right hand side of this equation may be written as

$$\vec{v}(x, y) = -y\partial_x + x\partial_y,$$

and thus

$$g(\varepsilon) = e^{\varepsilon \vec{v}}.$$

The Lie algebra $\mathfrak{so}(2)$ of $SO(2)$ consists of the vector fields

$$\{-\varepsilon y \partial_x + \varepsilon x \partial_y : \varepsilon \in \mathbb{R}\}.$$

The translations of the plane in the direction (a, b)

$$(x, y) \mapsto (x - \varepsilon a, y - \varepsilon b)$$

are generated by the constant vector field

$$a\partial_x + b\partial_y$$

The rotations and translations together form the orientation-preserving Euclidean group of the plane, denoted by $E^+(2)$. The full Euclidean group $E(2)$ is generated by rotations, translations, and reflections.

The Euclidean group is not commutative since translations and rotations do not commute. As a result, the corresponding Lie algebra $\mathfrak{e}(2)$ is not trivial. For example, if $\vec{v} = \partial_x$ is an infinitesimal generator of translations in the x-direction, and $\vec{w} = -y\partial_x + x\partial_y$ is an infinitesimal generator of rotations, then $[\vec{v}, \vec{w}] = \partial_y$ is the infinitesimal generator of translations in the y-direction.

The scaling transformations

$$(x, y) \mapsto (e^{\varepsilon r} x, e^{\varepsilon s} y)$$

are generated by the vector field

$$rx\partial_x + sy\partial_y.$$

Together with the translations and rotations, the scaling transformations generate the conformal group of angle preserving transformations of the plane.

Finally, as a nonlinear example, consider the vector field

$$\vec{v}(x, y) = x^2\partial_x - y^2\partial_y.$$

This generates the local flow

$$(x, y) \mapsto \left(\frac{x}{1 - \varepsilon x}, \frac{y}{1 + \varepsilon y}\right).$$

9.4. Transformations of function

Next, we want to consider the action of point transformations on functions.

Suppose that $f : \mathbb{R}^n \to \mathbb{R}$. We denote the coordinates of the independent variables by $x = \left(x^1, x^2, \dots, x^n\right) \in \mathbb{R}^n$, and the coordinate of the dependent variable by $u \in \mathbb{R}$. We assume that f is scalar-valued only to simplify the notation; it is straightforward to generalize the discussion to vector-valued functions $f : \mathbb{R}^n \to \mathbb{R}^m$.

Let $E = \mathbb{R}^n \times \mathbb{R}$ be the space with coordinates $\left(x^1, \dots, x^n, u\right)$. Then the graph Γ_f of f is the subset of E given by

$$\Gamma_f = \{(x, u) \in E : u = f(x)\}.$$

Consider a local one-parameter group of point transformations $g(\varepsilon) : E \to E$ on the space of independent and dependent variables. These transformations induce a local transformation of functions, which we denote in the same way,

$$g(\varepsilon) : f \mapsto g(\varepsilon) \cdot f$$

that maps the graph of f to the graph of $g(\varepsilon) \cdot f$. The global image of the graph of f under $g(\varepsilon)$ need not be a graph; it is, however, locally a graph (that is, in a sufficiently small neighborhood of a point $x \in \mathbb{R}^n$ and for small enough values of ε, when $g(\varepsilon)$ is sufficiently close to the identity).

To express the relationship between f and $\tilde{f} = g \cdot f$ explicitly, we write g as

$$g(\varepsilon) : (x, u) \mapsto (\tilde{x}, \tilde{u}), \qquad \tilde{x} = \tilde{X}(x, u, \varepsilon), \quad \tilde{u} = \tilde{U}(x, u, \varepsilon).$$

Then, since

$$g(\varepsilon) : \{(x, u) : u = f(x)\} \mapsto \{(\tilde{x}, \tilde{u}) : \tilde{u} = \tilde{f}(\tilde{x}, \varepsilon)\},$$

we have

$$\tilde{U}\left(x, f(x), \varepsilon\right) = \tilde{f}\left(\tilde{X}(x, f(x), \varepsilon), \varepsilon\right).$$

This is, in general, a complicated implicit equation for $\tilde{f}$ in terms of f.

Example 2.15. Suppose that $f : \mathbb{R} \to \mathbb{R}$ is the square function $f : x \mapsto x^2$ and $g(\varepsilon) : \mathbb{R} \times \mathbb{R} \to \mathbb{R} \times \mathbb{R}$ is the rotation

$$g(\varepsilon) \cdot (x, u) = ((\cos\varepsilon)x - (\sin\varepsilon)u, (\sin\varepsilon)x + (\cos\varepsilon)u).$$

If $u = f(x)$, then

$$\tilde{x} = (\cos\varepsilon)x - (\sin\varepsilon)x^2, \qquad \tilde{u} = (\sin\varepsilon)x + (\cos\varepsilon)x^2.$$

Solving the first equation for x in terms of $\tilde{x}$, then using the second equation to express $\tilde{u}$ in terms of $\tilde{x}$, we find that $\tilde{u} = \tilde{f}\left(\tilde{x}, \varepsilon\right)$ where

$$\tilde{f}\left(\tilde{x}, \varepsilon\right) = \frac{2\tilde{x}^2 \cos\varepsilon + 2\tilde{x}\tan\varepsilon}{1 - 2\tilde{x}\sin\varepsilon + \sqrt{1 - 4\tilde{x}\sin\varepsilon / \cos^2\varepsilon}}$$

Thus, the image of the function $x \mapsto x^2$ under the rotation $g(\varepsilon)$ is the function

$$x \mapsto \frac{2x^2 \cos\varepsilon + 2x\tan\varepsilon}{1 - 2x\sin\varepsilon + \sqrt{1 - 4x\sin\varepsilon / \cos^2\varepsilon}}.$$

Note that this reduces to $x \mapsto x^2$ if $\varepsilon = 0$, and that the transformed function is only defined locally if $\varepsilon \neq 0$.

9.5. Prolongation of transformations

In order to obtain the symmetries of a differential equation, we use a geometric formulation of how the derivatives of a function transform under point transformations.

To do this, we introduce a space $E^{(k)}$, called the k^{th} *jet space*, whose coordinates are the independent variables, the dependent variable, and the derivatives of the dependent variable of order less than or equal to k.

We will use multi-index notation for partial derivatives. A multi-index α is an n-tuple

$$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$$

where each $\alpha_i = 0, 1, 2, \ldots$ is a nonnegative integer. The α-partial derivative of a function $f : \mathbb{R}^n \to \mathbb{R}$ is

$$\partial^\alpha f = \partial_{x^1}^{\alpha_1} \partial_{x^2}^{\alpha_2} \ldots \partial_{x^n}^{\alpha_n} f.$$

This partial derivative has order $|\alpha|$ where

$$|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n.$$

We define $E^{(k)}$ to be the space with coordinates $(x, u, \partial^\alpha u)$ where α runs over all multi-indices with $1 \leq |\alpha| \leq k$. When convenient, we will use alternative notations for the partial-derivative coordinates, such as u_{x^i} for $\partial_{x^i} u$ and u_α for $\partial^\alpha u$.

Example 2.16. Written out explicitly, the coordinates on the first-order jet space $E^{(1)}$ are $\left(x^1, x^2, \dots, x^n, u, u_{x^1}, u_{x^2}, \dots, u_{x^n}\right)$. Thus, $E^{(1)}$ has dimension $(2n+1)$.

Example 2.17. For functions $u = f(x, y)$ of two independent variables, the second-order jet space $E^{(2)}$ has coordinates $(x, y, u, u_x, u_y, u_{xx}, u_{xy}, u_{yy})$.

Point transformations induce a map of functions to functions, and therefore they induce maps of the derivatives of functions and of the jet spaces.

Specifically, suppose that $g(\varepsilon) : E \to E$ is a point transformation. We extend, or prolong $g(\varepsilon)$, to a transformation

$$\mathbf{pr}^{(k)} g(\varepsilon) : E^{(k)} \to E^{(k)}$$

in the following way. Given a point $(x, u, \partial^\alpha u) \in E^{(k)}$, pick a function $f : \mathbb{R}^n \to \mathbb{R}$ whose value at x is u and whose derivatives at x are $\partial^\alpha u$, meaning that

$$f(x) = u, \qquad \partial^\alpha f(x) = \partial^\alpha u \qquad \text{for } 1 \leq |\alpha| \leq k.$$

For example, we could choose f to be a polynomial of degree k.

Suppose that $g(\varepsilon) \cdot (x, u) = (\tilde{x}, \tilde{u})$ is the image of $(x, u) \in E$ under $g(\varepsilon)$ and $\tilde{f} = g(\varepsilon) \cdot f$ is the image of the function f. We define the image of the jet-coordinates by

$$\widetilde{\partial^\alpha u} = \tilde{\partial}^\alpha \tilde{f}\,(\tilde{x})\,.$$

That is, they are the values of the derivatives of the transformed function $\tilde{f}\,(\tilde{x})$. One can show that these values do not depend on a particular choice of the function f, so this gives a well-defined map $\mathbf{pr}^{(k)} g(\varepsilon)$ on $E^{(k)}$ such that

$$\mathbf{pr}^{(k)} g(\varepsilon) : (x, u, \partial^\alpha u) \mapsto \left(\tilde{x}, \tilde{u}, \widetilde{\partial^\alpha u}\right).$$

9.6. Prolongation of vector fields

Suppose that $g(\varepsilon) : E \to E$ is generated by the vector field

$$\vec{v}(x, u) = \xi^i(x, u)\partial_{x^i} + \varphi(x, u)\partial_u. \tag{2.47}$$

Then, writing the coordinates of $E^{(k)}$ as (x, u, u_α), the prolonged transformation

$$\mathbf{pr}^{(k)} g(\varepsilon) : E^{(k)} \to E^{(k)}$$

is generated by a vector field $\mathbf{pr}^{(k)} \vec{v}$ on $E^{(k)}$. This prolonged vector field has the form

$$\mathbf{pr}^{(k)} \vec{v} = \xi^i \partial_{x^i} + \varphi \partial_u + \sum_{|\alpha|=1}^{k} \varphi^\alpha \partial_{u_\alpha},$$

where the φ^α are suitable coefficient functions, which are determined by $\vec{v}$.

The prolongation formula expresses the coefficients φ^α of the prolonged vector field in terms of the coefficients ξ^i, φ of the original vector field. We will state the result here without proof (see [**40**] for a derivation).

To write the prolongation formula in a compact form — see (2.49) below — we define the total derivative $D_{x^i} F : E^{(k)} \to \mathbb{R}$ of a function $F : E^{(k)} \to \mathbb{R}$ with respect to an independent variable x^i by

$$D_{x^i} F = \partial_{x^i} F + \sum_{|\alpha|=0}^{k} u_{\alpha, i} \partial_{u_\alpha} F.$$

Here, we use the notation

$$u_{\alpha,i} = \partial_{x^i}\partial^\alpha u$$

to denote the coordinate of the corresponding derivative. That is, $u_{\alpha,i} = u_\beta$ where $\beta_i = \alpha_i + 1$ and $\beta_j = \alpha_j$ for $j \neq i$.

In other words, the total derivative $D_{x^i}F$ of F with respect to x^i is what we would obtain by differentiating F with respect to x^i *after* the coordinates u, u_α have been evaluated at a function of x and its derivatives.

If $\alpha = (\alpha_1, \ldots, \alpha_n)$ is a multi-index, we define the α-total derivative by

$$D^\alpha = D_{x^1}^{\alpha_1} D_{x^2}^{\alpha_2} \ldots D_{x^n}^{\alpha_n}.$$

Total derivatives commute, so the order in which we take them does not matter.

Finally, we define the characteristic $Q : E^{(1)} \to \mathbb{R}$ of the vector field (2.47) by

$$Q\left(x, u, \partial u\right) = \varphi(x, u) - \xi^i(x, u) u_{x^i},$$

where the summation convention is understood, and $\partial u = (u_{x^1}, \ldots, u_{x^n})$ is the first-derivative coordinate.

Then the k^{th}-prolongation of the vector field (2.47) is given by

$$\mathbf{pr}^{(k)}\vec{v} = \xi^i \partial_{x^i} + \varphi \partial_u + \sum_{|\alpha|=1}^{k} \varphi^\alpha \partial_{u_\alpha}, \tag{2.48}$$

$$\varphi^\alpha = D^\alpha Q + \xi^i u_{\alpha,i}. \tag{2.49}$$

This is the main result needed for the algebraic computation of symmetries of a differential equation. See Olver [**40**] for the prolongation formula for systems.

9.7. Invariance of a differential equation

A k^{th} order differential equation for a real-valued function $u(x)$ may be written as

$$F\left(x, u, \partial^\alpha u\right) = 0$$

where $F : E^{(k)} \to \mathbb{R}$ and $1 \leq |\alpha| \leq k$. Here, we abuse notation and use the same symbols for the coordinates u, $\partial^\alpha u$ and the functions $u(x)$, $\partial^\alpha u(x)$.

A local point transformation $g(\varepsilon) : E \to E$ is a symmetry of the differential equation if $g(\varepsilon) \cdot u$ is a solution whenever u is a solution. This means that, for all ε in the neighborhood of 0 for which $g(\varepsilon)$ is defined, we have

$$F\left(\mathbf{pr}^{(k)} g(\varepsilon) \cdot (x, u, \partial^\alpha u)\right) = F\left(x, u, \partial^\alpha u\right) \qquad \text{on } F\left(x, u, \partial^\alpha u\right) = 0. \tag{2.50}$$

Suppose that $g(\varepsilon) = e^{\varepsilon \vec{v}}$. Then, differentiating (2.50) with respect to ε and setting $\varepsilon = 0$, we conclude that

$$\mathbf{pr}^{(k)}\vec{v} \cdot F\left(x, u, \partial^\alpha u\right) = 0 \qquad \text{on } F\left(x, u, \partial^\alpha u\right) = 0, \tag{2.51}$$

where $\mathbf{pr}^{(k)}\vec{v}$ acts on F by differentiation. Conversely, if F satisfies the 'maximal rank' condition $DF \neq 0$ on $F = 0$, which rules out degenerate ways of the equation such as $F^2 = 0$, we may integrate the infinitesimal invariance condition (2.51) to obtain (2.50).

The condition (2.51) is called the *determining equation* for the infinitesimal symmetries of the differential equation. It is typically a large, over-determined system of equations for $\xi^i(x, u)$, $\varphi(x, u)$ and their derivatives, which is straightforward (though tedious) to solve.

Thus, in summary, to compute the point symmetries of a differential equation

$$F(x, u, \partial^\alpha u) = 0$$

we use the prolongation formula (2.48)–(2.49) to write down the infinitesimal invariance condition (2.51), solve the resulting equations for $\xi^i(x, u)$ and $\varphi(x, u)$, then integrate the vector fields (2.47) to obtain the symmetries $g = e^{\vec{v}}$.

9.8. Porous medium equation

Let us return to the porous medium equation (2.36).

The space E of independent and dependent variables has coordinates (x, t, u). We may write the equation as

$$F(u, u_x, u_t, u_{xx}) = 0 \tag{2.52}$$

where $F : E^{(2)} \to \mathbb{R}$ is given by

$$F(u, u_x, u_t, u_{xx}) = -u_t + u_x^2 + u u_{xx}.$$

A vector field $\vec{v}$ on E is given by

$$\vec{v}(x, t, u) = \xi(x, t, u)\partial_x + \tau(x, t, u)\partial_t + \varphi(x, t, u)\partial_u. \tag{2.53}$$

From (2.48), the second prolongation of $\vec{v}$ has the form

$$\mathbf{pr}^{(2)}\vec{v} = \xi\partial_x + \tau\partial_t + \varphi\partial_u + \varphi^x\partial_{u_x} + \varphi^t\partial_{u_t} + \varphi^{xx}\partial_{u_{xx}} + \varphi^{xt}\partial_{u_{xt}} + \varphi^{tt}\partial_{u_{tt}}.$$

The infinitesimal invariance condition (2.51) applied to (2.52) gives

$$-\varphi^t + 2u_x\varphi^x + \varphi u_{xx} + u\varphi^{xx} = 0 \qquad \text{on } u_t = u_x^2 + u u_{xx}. \tag{2.54}$$

From (2.49), we have

$$\begin{aligned}
\varphi^x &= D_x Q + \xi u_{xx} + \tau u_{xt},\\
\varphi^t &= D_t Q + \xi u_{xt} + \tau u_{tt},\\
\varphi^{xx} &= D_x^2 Q + \xi u_{xxx} + \tau u_{xxt},\\
\varphi^{xt} &= D_x D_t Q + \xi u_{xxt} + \tau u_{xtt},\\
\varphi^{tt} &= D_t^2 Q + \xi u_{xtt} + \tau u_{ttt},
\end{aligned}$$

where the characteristic Q of (2.53) is given by

$$Q(x, t, u, u_t, u_x) = \varphi(x, t, u) - \xi(x, t, u)\, u_x - \tau(x, t, u)\, u_t,$$

and the total derivatives D_x, D_t of a function $f(x, t, u, u_x, u_t)$ are given by

$$\begin{aligned}
D_x f &= f_x + u_x f_u + u_{xx} f_{u_x} + u_{xt} f_{u_t},\\
D_t f &= f_t + u_t f_u + u_{xt} f_{u_x} + u_{tt} f_{u_t}.
\end{aligned}$$

Expanding the total derivatives of Q, we find that

$$\begin{aligned}
\varphi^x &= \varphi_x + (\varphi_u - \xi_x)\, u_x - \tau_x u_t - \xi_u u_x^2 - \tau_u u_x u_t,\\
\varphi^t &= \varphi_t - \xi_t u_x + (\varphi_u - \tau_t)\, u_t - \xi_u u_x u_t - \tau_u u_t^2,\\
\varphi^{xx} &= \varphi_{xx} + (2\varphi_{xu} - \xi_{xx})\, u_x - \tau_{xx} u_t + (\varphi_{uu} - 2\xi_{xu})\, u_x^2\\
&\quad - 2\tau_{xt} u_x u_t - \xi_{uu} u_x^3 - \tau_{uu} u_x^2 u_t + (\varphi_u - 2\xi_x)\, u_{xx}\\
&\quad - 2\tau_x u_{xt} - 3\xi_u u_x u_{xx} - \tau_u u_t u_{xx} - 2\tau_u u_x u_{xt}.
\end{aligned}$$

We use these expressions in (2.54), replace u_t by $uu_{xx}+u_x^2$ in the result, and equate coefficients of the terms that involve different combinations of the spatial derivatives of u to zero.

The highest derivative terms are those proportional to u_{xt} and $u_x u_{xt}$. Their coefficients are proportional to τ_x and τ_u, respectively, so we conclude that $\tau_x = 0$ and $\tau_u = 0$, which implies that $\tau = \tau(t)$ depends only on t.

The remaining terms that involve second-order spatial derivatives are a term proportional to $u_x u_{xx}$, with coefficient ξ_u, so $\xi_u = 0$ and $\xi = \xi(x,t)$, and a term proportional to u_{xx}. Equating the coefficient of the latter term to zero, we find that

$$\varphi = (2\xi_x - \tau_t)\, u. \tag{2.55}$$

Thus, φ is a linear function of u.

The terms that are left are either proportional to u_x^2, u_x, or involve no derivatives of u. Equating to zero the coefficients of these terms to zero, we get

$$\begin{aligned}
&\tau_t - 2\xi_x + \varphi_u + u\varphi_{uu} = 0,\\
&\xi_t - u\xi_{xx} + 2\varphi_x + 2u\varphi_{xu} = 0,\\
&\varphi_t - u\varphi_{xx} = 0.
\end{aligned}$$

The first equation is satisfied by any φ of the form (2.55). The second equation is satisfied if $\xi_t = 0$ and $\xi_{xx} = 0$, which implies that

$$\xi = \varepsilon_1 + \varepsilon_3 x$$

for arbitrary constants ε_1, ε_3. The third equation holds if $\tau_{tt} = 0$, which implies that

$$\tau = \varepsilon_2 + \varepsilon_4 t$$

for arbitrary constants ε_2, ε_3. Equation (2.55) then gives

$$\varphi = (2\varepsilon_3 - \varepsilon_4)\, u$$

Thus, the general form of an infinitesimal generator of a point symmetry of (2.36) is

$$\vec{v}(x,t,u) = (\varepsilon_1 + \varepsilon_3 x)\,\partial_x + (\varepsilon_2 + \varepsilon_4 t)\,\partial_t + (2\varepsilon_3 - \varepsilon_4)\, u\partial_u.$$

We may write this as

$$\vec{v} = \sum_{i=1}^{4} \varepsilon_i \vec{v}_i$$

where the vector fields $\vec{v}_i$ are given by

$$\vec{v}_1 = \partial_x, \qquad \vec{v}_2 = \partial_t \tag{2.56}$$

$$\vec{v}_3 = x\partial_x + 2u\partial_u \qquad \vec{v}_4 = t\partial_t - u\partial_u \tag{2.57}$$

The vector fields $\vec{v}_1$, $\vec{v}_2$ in (2.56) generate the space and time translations

$$(x,t,u) \mapsto (x+\varepsilon_1, t, u), \qquad (x,t,u) \mapsto (x, t+\varepsilon_2, u),$$

respectively. The vector fields $\vec{v}_3$, $\vec{v}_4$ in (2.57) generate the scaling transformations

$$(x,t,u) \mapsto \left(e^{\varepsilon_3} x, t, e^{2\varepsilon_3} u\right) \qquad (x,t,u) \mapsto \left(x, e^{\varepsilon_4} t, e^{-\varepsilon_4} u\right).$$

These are the same as (2.44) with

$$\alpha = e^{\varepsilon_3}, \qquad \beta = e^{\varepsilon_4}.$$

Thus the full point symmetry group of the porous medium equation is four dimensional, and is generated by space and time translations and the two scaling transformations that arise from dimensional analysis.

This result is, perhaps, a little disappointing, since we did not find any new symmetries, although it is comforting to know that there are no other point symmetries to be found. For other equations, however, we can get symmetries that are not at all obvious.

Example 2.18. Consider the one-dimensional heat equation

$$u_t = u_{xx}. \tag{2.58}$$

The determining equation for infinitesimal symmetries is

$$\varphi^t = \varphi^{xx} \qquad \text{on } u_t = u_{xx}.$$

Solving this equation and integrating the resulting vector fields, we find that the point symmetry group of (2.58) is generated by the following transformations [**40**]:

$$\begin{aligned}
&u(x,t) \mapsto u(x-\alpha, t),\\
&u(x,t) \mapsto u(x, t-\beta),\\
&u(x,t) \mapsto \gamma u(x,t),\\
&u(x,t) \mapsto u(\delta x, \delta^2 t),\\
&u(x,t) \mapsto e^{-\epsilon x+\epsilon^2 t} u(x-2\epsilon t, t),\\
&u(x,t) \mapsto \frac{1}{\sqrt{1+4\eta t}} \exp\left[\frac{-\eta x^2}{1+4\eta t}\right] u\left(\frac{x}{1+4\eta t}, \frac{t}{1+4\eta t}\right),\\
&u(x,t) \mapsto u(x,t) + v(x,t),
\end{aligned}$$

where $(\alpha, \dots, \eta)$ are constants, and $v(x,t)$ is an arbitrary solution of the heat equation. The scaling symmetries involving γ and δ can be deduced by dimensional arguments, but the symmetries involving ϵ and η cannot.

As these examples illustrate, given a differential equation it is, in principle, straightforward (but lengthy) to write out the conditions that a vector field generates a point symmetry of the equation, solve these conditions, and integrate the resulting infinitesimal generators to obtain the Lie group of continuous point symmetries of the equation. There are a number of symbolic algebra packages that will do this automatically.

Finally, we note that point symmetries are not the only kind of symmetry one can consider. It is possible to define 'generalized' (also called 'nonclassical' or 'higher') symmetries on infinite-dimensional jet spaces (see [**40**], for an introduction). These are of particular interest in connection with completely integrable equations, such as the Korteweg-de Vries (KdV) equation, which possess 'hidden' symmetries that are not revealed by their point symmetries