Chapter 1

Preliminaries

If you would understand anything, observe its beginning. - Aristotle.

Since the 1980s we have thought of ocean circulation in terms of the "global conveyer belt", in which cold polar waters sink and then circulate around the ocean basins, eventually being warmed in the tropics. But the truth is that this circulation has a typical speed of only a few cm/s, and it is generally accompanied by oscillations many times faster (1m/s is not uncommon), with periods ranging from seconds to months.

The largest oscillations are the majestic mesoscale eddies which spin off strong currents like the Agulhas and the Gulf Stream (figure 1.1). Today, much research is focused on the next size smaller: the submesoscale eddies. Smaller than this are the gravity waves and, at the smallest scale, three-dimensional turbulence.

One must measure for a year or more in order to average out the oscillations and discern the mean "conveyer belt" current. But to think of these oscillations as something we can average away is to fool ourselves, for it is largely the oscillations that govern the conveyer belt. We can't understand one without the other. Today, the majority of attention in physical oceanography is on the oscillations.

One way to understand such a chaotic profusion of motions is to ask what would happen if, at some initial instant, the ocean was calm, with steady, orderly currents. Would the oscillations develop spontaneously? If so, how? That is the essential goal of instability theory.

For example, figure 1.1 shows Gulf Stream eddies simulated in a global-scale numerical model. No human mathematician could solve the equations that describe these intricate motions but, using the methods of linear perturbation theory, we can not only predict their length and time scales but also understand quite a lot about what causes them. The trick is to imagine a fictitious Gulf Stream that is straight and eddy-free then study what happens in the very first few moments - after the current begins to buckle but before it grows so complex as to be mathematically intractable.

In this course we will study instabilities on scales from centimeter to global, driven by mechanisms including gravity, current shear and the Earth's rotation. While the emphasis is on oceanic applications, all of these instabilities are active in the atmosphere and in other geophysical fluid systems, and we will frequently look at examples from these other systems.

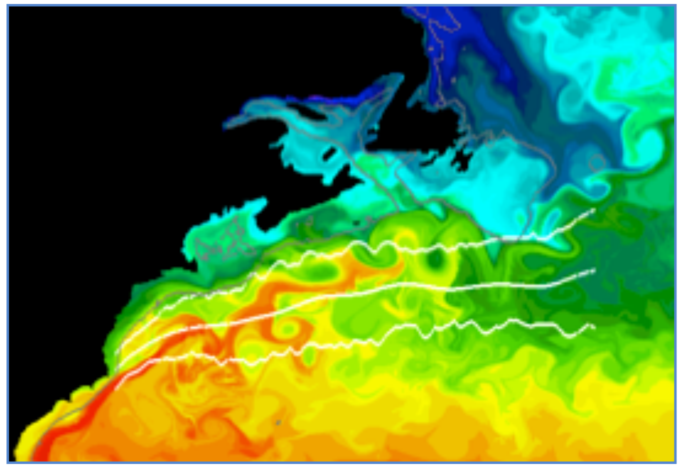

Figure 1.1: Instability of the Gulf Stream in a global ocean model. Colors represent sea surface temperature. (U. of Miami)

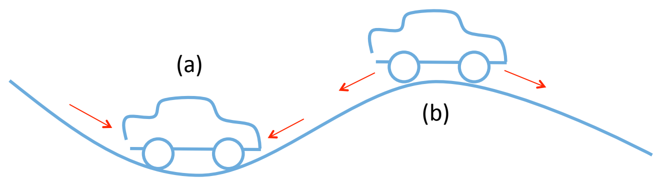

Figure 1.2: No emergency brakes! Where would you park? Arrows show the gravitational force that acts when the car is displaced slightly from equilibrium.

1.1 What is instability?

Suppose that the emergency brake in your car doesn't work, and you have to park somewhere in hilly country. Where can you park so that your car doesn't roll away (figure 1.2)? I hope you would park at point (a), the bottom of a valley. But what about point (b), the top of a hill? In theory, you could park there and the car would not roll away. But you would have to park at *exactly* the right spot, and any little disturbance would cause your car to roll away.

In mathematical terms, we say that both points (a) and (b) are equilibrium states, i.e. states at which the system can remain steady in time. The difference is in what happens when the system is displaced slightly from the equilibrium. If you park at the bottom of the valley (a), and if something then pushes the car slightly to the left or the right, gravity will pull it back toward its original location. The car will rock back and forth and eventually come to rest due to friction. In contrast, if you park at the top of the hill (b) and the car is moved slightly, gravity pulls it further from the equilibrium point. The further the car travels, the steeper the slope and the stronger the pull of gravity. Goodbye car! We say that equilibrium (a) is stable, while (b) is unstable.

The equations that describe geophysical fluid systems are in general far too complicated to solve analytically. One way to get around this problem is to look for equilibria, i.e. solutions that are valid when all time derivatives are set to zero. Flows are often found to be close to such equilibria. For example, the surface of a lake is in equilibrium if it is horizontal. Although this is never exactly true, it is pretty close on average.

Once we have identified an equilibrium state, the next step is to determine its stability. If the equilibrium is stable, disturbances will have the form of oscillations (e.g. the car in figure 1.2a), or waves. Geophysical waves are the topic of the companion course OC681. If the equilibrium is unstable, then small disturbances grow exponentially. Instabilities will be our topic here.

1.1.1 The role of forcing

You might reasonable wonder why we ever see unstable systems in nature. It is much more usual to see systems close to stable equilibria. For example, the surface of a lake is never perfectly horizontal, but it usually stays pretty close, because the horizontal equilibrium state is stable.

But a sufficiently strong wind destabilizes that horizontal equilibrium state, and waves grow as a result.[1] If the waves grow large enough they fall prey to a second kind of instability as the crests roll over and break (convective instability; chapter 2). The surface then returns (on average) to a horizontal state until a new set of waves emerges. Eventually the wind dies down and the horizontal state is once again stable.

The oceans and atmosphere are almost always turbulent, and this cycle of instability is the reason. Forcing by wind, sun, gravity and planetary rotation tends to push the system away from stable equilibrium and towards unstable states. Instability and turbulence (see OC674) then act to relax the system back toward a stable state.

1.2 Goals of the course

In this course we have three main goals.

1. **Mechanisms:** We aim to understand, on an intuitive level, the basic physical processes that generate instability. In the car example, we've seen how motion away from equilibrium alters the effect of gravity (arrows in figure 1.2), resulting in oscillations or instability. Geophysical examples will take a bit more work to understand, but we'll do it.

2. **Rules of thumb:** We would like to be able to predict the stability or instability of a system quickly with minimal math. In the car example, we are able to predict whether an equilibrium point will be stable or unstable without knowing the details of the shape of the hill or valley. All we need to know is whether the equilibrium is a maximum or a minimum of elevation, i.e. whether the curvature at that point is negative or positive.

   We can invent similar rules for most types of geophysical flow instability. These allow us to estimate not only the likelihood of instability, but also the spatial and temporal scales on which it will grow. These can help us identify the particular mechanism through which a geophysical flow is becoming unstable. For example, the Gulf Stream eddies shown in figure 1.1 could be due to different instabilities (which you will learn about later). By comparing their observed length scales, and the time they take to grow, with rules of thumb based on various known instability types, we can take a first guess as to the mechanism.

3. **Numerical solution methods:** Sometimes a rule of thumb is not enough. We want to determine quantitative details of an instability, perhaps in a situation where many physical factors interact. In that case we may have to solve a nontrivial set of differential equations. Many advanced analytical methods are available, but in this course we will focus on numerical methods. Since the 1980s,

[1] The process is similar to shear instability, covered in chapters 3-5.

computers have had the capacity to do something unprecedented: to solve a differential equation whose coefficients are defined using observational measurements. That capability is now in use in the analysis of oceanographic and atmospheric observations. Our discussion of numerical methods will prepare the student to learn this technique.

1.3 Tools

Below are three topics I'll expect you to have some familiarity with. Under each topic is listed one or more things that you should be able to do.

1. **Calculus**:
   - Solve this boundary value problem:
   $$y'' = -y; \quad y(0) = y(\pi) = 0. \tag{1.3.1}$$
   - Derive this Taylor series approximation:
   $$\frac{1}{1+x} \approx 1 - x + x^2, \quad \text{for } |x| \ll 1.$$
   - Understand the meaning of (though not necessarily solve) a partial differential equation, e.g.
   $$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = -\frac{\partial \pi}{\partial x}.$$
2. **Linear algebra**:
   - Compute the eigenvalues of a 2×2 matrix.
3. **Matlab**: Homework will be done using the Matlab programming environment. You don't have to be an expert; you'll learn as you go. But if you've never used Matlab at all it would be worth familiarizing yourself with the basic syntax. Try the following:
   - Write a function, and a script that calls it.
   - Define a matrix and compute its eigenvalues using eig.
   - Make a line plot using plot and an image plot using pcolor.

   Matlab is available free for OSU students:

   http://is.oregonstate.edu/accounts-support/software/software-list/matlab

1.4 Perturbation theory: the parking problem revisited [optional]

We now consider the car-parking example from a more rigorous perspective. This example will illustrate some of the fundamental ideas of perturbation theory. Let ℓ be the distance the car rolls along the hill, measured left-to-right from some arbitrary origin (figure 1.3). The force of gravity is $-mg$, where m is the mass of the car and $g = 9.81$m/s at Earth's surface. The component of gravity in the direction of ℓ is

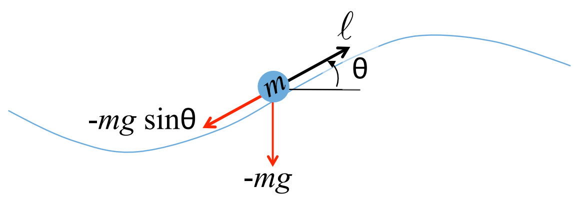

Figure 1.3: Definition diagram for the car-parking example.

$-mg\sin\theta$, where $\theta(\ell)$ is the angle of the road from the horizontal at any point ℓ. Newton's second law is therefore

$$m\frac{d^2\ell}{dt^2} = -mg\sin\theta.$$

Abbreviating $\sin\theta$ as $s(\ell)$ and removing the common factor m, we have

$$\frac{d^2\ell}{dt^2} + gs(\ell) = 0. \tag{1.4.1}$$

A general solution requires specifying the function $s(\ell)$ which, of course, is different for every road. An exact solution would be very difficult since $s(\ell)$ is in general nonlinear. Progress can be made if we identify equilibrium points: values of ℓ at which $s(\ell) = 0$, and concern ourselves only with the behavior close to those points. Equilibria exist when $\theta = 0$, i.e. at the bottom of a valley or the top of a hill (as can be seen by setting all time derivatives to zero in 1.4.1).

Let $\ell = \ell_0$ be an equilibrium point, and seek a solution

$$\ell(t) = \ell_0 + \varepsilon\ell_1(t) + \varepsilon^2\ell_2(t) + \ldots, \tag{1.4.2}$$

where ε is a parameter measuring the amplitude of the disturbance. Because we'll be assuming that ε is small, the early terms in the series are the most important. We now expand the unknown function $s(\ell)$ in a Taylor series about ℓ_0 and substitute (1.4.2) into the result:

$$\begin{aligned} s(\ell) &= s(\ell_0) + s'(\ell_0)(\ell - \ell_0) + \ldots \\ &= s(\ell_0) + \varepsilon s'(\ell_0)\ell_1 + \varepsilon^2 G + \ldots \end{aligned}$$

Here, G stands for some complicated terms whose details don't matter. [2] Now substitute into (1.4.1):

$$\frac{d^2\ell_0}{dt^2} + \varepsilon\frac{d^2\ell_1}{dt^2} + \varepsilon^2\frac{d^2\ell_2}{dt^2} + \ldots + g\ \left[s(\ell_0) + \varepsilon s'(\ell_0)\ell_1 + \varepsilon^2 G + \ldots\right] = 0,$$

or, gathering powers of ε,

$$\frac{d^2\ell_0}{dt^2} + gs(\ell_0) + \varepsilon\left[\frac{d^2\ell_1}{dt^2} + gs'(\ell_0)\ell_1\right] + \varepsilon^2\left[\frac{d^2\ell_2}{dt^2} + gG\right] + \ldots = 0. \tag{1.4.3}$$

Now here is a subtle but important point. Regardless of values of the quantities in square brackets, we can always find one or more values of ε that satisfy (1.4.3). But that's not what we're looking for. What we want is to take the limit $\varepsilon \to 0$, and have the equation be satisfied throughout that limiting process. In other words, the equations has to be true for every value of ε. That can only be true if the coefficients of the powers of ε all vanish individually. [3] That leaves us with an infinite sequence of equations whose solutions are the

[2] Try it if you like. You should get $G = s'(\ell_0)\ell_2 + \frac{1}{2}s''(\ell_0)\ell_1^2$.

[3] As an analogy, consider a quadratic equation $ax^2 + bx + c = 0$. For given values of a, b and c, you can easily find solutions for x using the quadratic formula. But what if the equation has to be satisfied for *every* x? That's only possible if $a = b = c = 0$.

unknown functions ℓ_0, $\ell_1(t)$, $\ell_2(t)$, etc.:

$$\frac{d^2\ell_0}{dt^2} + gs(\ell_0) = 0, \tag{1.4.4}$$

$$\frac{d^2\ell_1}{dt^2} + gs'(\ell_0)\ell_1 = 0, \tag{1.4.5}$$

$$\frac{d^2\ell_2}{dt^2} + gG = 0, \tag{1.4.6}$$

$$\ldots\text{etc.}$$

Now, let's look again at our putative solution (1.4.2). We already know the first term, ℓ_0; it's just the equilibrium position (hill or valley). As a result, the first equation (1.4.4) is satisfied trivially; $s(\ell_0) = 0$ and $d^2\ell_0/dt^2 = 0$.

The next term in (1.4.2), $\varepsilon\ell_1(t)$, is the only one that matters if ε is made sufficiently small. We therefore concern ourselves only with the second equation, (1.4.5), whose solution is ℓ_1. This is a linear ordinary differential equation, and very easy to solve because $s'(\ell_0)$ is a constant. For tidiness we'll abbreviate $s'(\ell_0)$ as s'_0. We'll consider two cases.

If $s'_0 > 0$, we define $gs'_0 = \omega^2$, and the general solution is

$$\ell_1 = A\sin\omega t + B\cos\omega t,$$

where A and B are constants to be determined by the initial conditions. This oscillatory solution describes the car rocking back and forth after being displaced from a stable equilibrium point (a valley).

If $s'_0 < 0$, we define $-gs'_0 = \sigma^2$, and the solution is

$$\ell_1 = Ae^{\sigma t} + Be^{-\sigma t}.$$

As long as $A \neq 0$, the first term grows exponentially and will eventually dominate the solution. This describes the unbounded motion of the car away from an unstable equilibrium, i.e. the top of a hill.

Here are some general features of stability analysis that the car-parking problem illustrates:

1. The equation (1.4.5) is a linear, homogeneous ordinary differential equation. This is always true. In general, though, the coefficients will not be constant, and the solution will be much more difficult. For that reason we will often resort to numerical methods.
2. In the general case, solutions can be oscillatory, exponential, or a combination. Our main interest is in exponential solutions with positive growth rate σ. Often, we can define a simple rule of thumb to tell us which type of solution will be found; here, we only need to know the sign of s'_0.
3. Having solved for ℓ_1, it is possible to substitute the result into (1.4.6) and solve for ℓ_2, and so on to even higher orders. There are people who do this, but we won't.
4. The solution is valid only if the neglected terms in (1.4.2) are indeed negligible. How can we tell if this is true? Most commonly, we regard the smallness of the neglected terms as a *hypothesis* whose validity we test by comparing the solution with reality.

1.5 Numerical solution of a boundary value problem

The basic geophysical flow instabilities are analyzed as solutions of two-point boundary value problems. In his section I'll define this class of problems and introduce a matrix-based finite difference method for solving them.

1.5.1 Defining the problem

Let $f(x)$ be the solution to a second-order ordinary differential equation with independent variable x. Complete specification of f requires two pieces of information in addition to the equation itself. These can be either

- values of f and its first derivative at some initial point which we'll label as zero, i.e. $f(0)$ and $f'(0)$, or
- values of f at two points, say $f(0)$ and $f(L)$.

The first case is called an initial value problem; the second is called a boundary value problem.

A critical difference between these two classes of problem is that the first generally has a solution while the second generally does not. Here's a simple example:

$$f'' = -k^2 f. \tag{1.5.1}$$

The general solution is

$$f = A\cos kx + B\sin kx, \tag{1.5.2}$$

where A and B are constants to be determined. Consider first the initial value problem. Suppose we have initial conditions $f(0) = 0$ and $f'(0) = 1$. The solution is then (1.5.2) with $A = 0$ and $B = 1/k$.

Now, consider the boundary value problem with conditions $f(0) = 0$ and $f(L) = 0$. The first condition is satisfied if $A = 0$, but the second can be satisfied only if $k = \pm n\pi/L$, where n is any integer. These special values of k are called the eigenvalues of the problem, and unless k is equal to one of those eigenvalues, the problem has no solution. We also call this a differential eigenvalue problem. It is analogous to the more familiar matrix eigenvalue problem, as we will now see.

Suppose that

- $\vec{x}$ is a list of possible values of x arranged as a vector;
- $\vec{f}$ and $\vec{f}^{(2)}$ are vectors composed of the corresponding values of f and its second derivative, respectively;
- $\underset{\sim}{D}$ is a matrix such that $\underset{\sim}{D}\vec{f} = \vec{f}^{(2)}$.

We can now write (1.5.3) as

$$\underset{\sim}{D}\vec{f} = -k^2\vec{f}, \tag{1.5.3}$$

which is a standard matrix eigenvalue problem. Because the matrix eigenvalue problem can be easily solved using standard numerical routines (e.g. the Matlab function *eig*), this approach suggests a convenient way to solve the differential eigenvalue problem. But first we have to identify this matrix $\underset{\sim}{D}$ that transforms a vector into its second derivative.

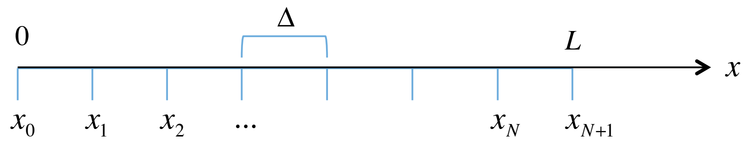

Figure 1.4: Discretizing the x axis.

1.5.2 Discretization: the derivative matrix

We discretize the independent variable x by choosing a vector of values:

$$x_i = i\Delta, \quad \text{where } \; i = 0, 1, 2, \dots, N+1.$$

The first and last values correspond to the boundaries, say $x_0 = 0$ and $x_{N+1} = L$ (figure 1.4). This requires that

$$\Delta = L/(N+1).$$

Note that the x_i are evenly spaced. This restriction is not necessary, but it simplifies the math. We can now discretize the solution f,

$$f_i = f(x_i),$$

and the k^{th} derivative

$$f_i^{(k)} = \left.\frac{d^k f}{dx^k}\right|_{x=x_i}$$

The finite difference approximation to the derivative $f^{(k)}$ is a weighted sum of f_i values at neighboring points. A well-known example is:

$$\boxed{f_i' = \frac{f_{i+1} - f_i}{\Delta},} \tag{1.5.4}$$

which approximates the first derivative to arbitrary accuracy as $\Delta \to 0$. In general

$$f_i^{(k)} \;=\; \sum_{j=j_1}^{j_2} A_j^{(k)} \, f_{i+j}.$$

The range of the summation, $j = j_1, \dots, j_2$, is called the stencil. For example, in (1.5.4), $k = 1$, $j_1 = 0$ and $j_2 = 1$, and the weights are $A_0^{(1)} = -1/\Delta$ and $A_1^{(1)} = 1/\Delta$.

The weights are computed by means of a Taylor series expansion of f about x_i:

$$f_{i+j} \;=\; f(x_i + j\Delta) \;=\; f_i \;+\; j\Delta f_i^{(1)} \;+\; \frac{1}{2}(j\Delta)^2 f_i^{(2)} \;+\; \dots \;+\; \frac{1}{n!}(j\Delta)^n f_i^{(n)}. \tag{1.5.5}$$

For example, suppose we want to approximate the first derivative using the three-point stencil $j = -1, 0, 1$:

$$\tilde{f}_i' = A f_{i-1} + B f_i + C f_{i+1},$$

where the tilde indicates the approximation. Using (1.5.5), we can write

$$\begin{aligned}
A f_{i-1} + B f_i + C f_{i+1} \;&=\; A\left[f_i - \Delta f_i' + \frac{1}{2}\Delta^2 f_i'' - \frac{1}{6}\Delta^3 f_i''' + \dots\right] \;+\; B\left[f_i\right] \;+\; C\left[f_i + \Delta f_i' + \frac{1}{2}\Delta^2 f_i'' + \frac{1}{6}\Delta^3 f_i''' + \dots\right] \\
&=\; (A+B+C) f_i \;+\; (-A+C)\Delta f_i' \;+\; (A+C)\frac{1}{2}\Delta^2 f_i'' + \;(-A+C)\frac{1}{6}\Delta^3 f_i''' \;+\; \dots \\
&=\; f_i'.
\end{aligned} \tag{1.5.6}$$

The final equality expresses our wish that the approximation $\tilde{f}_i'$ equal the true value f_i', a wish that will not be granted. We try to find values for A, B and C so that (1.5.6) is satisfied for all functions f, which requires that the final equality hold separately for the terms multiplying each derivative $f^{(n)}$. The terms multiplying f_i, f_i' and f_i'' (colored blue) give:

$$\begin{aligned} A+B+C &= 0 \\ (-A+C)\Delta &= 1 \\ (A+C) &= 0. \end{aligned}$$

We now have three equations for three unknowns. Since this is all the equations we can satisfy, we can equate only the blue terms in (1.5.6). The solution is:

$$A = -\frac{1}{2\Delta}, \quad B = 0, \quad C = \frac{1}{2\Delta}, \tag{1.5.7}$$

or

$$\boxed{\tilde{f}_i' = \frac{f_{i+1} - f_{i-1}}{2\Delta}} \tag{1.5.8}$$

This is called a centered difference owing to its symmetry.

How accurate is this approximation? Recall that, so solve (1.5.6), we had to ignore the red term. That term is a measure of the error in (1.5.8). Using our computed solution (1.5.7) for $A - C$ in (1.5.6), we have

$$\tilde{f}_i' = \frac{f_{i+1} - f_{i-1}}{2\Delta} = f_i' + \frac{1}{6}\Delta^2 f_i''' + \ldots$$

We can't tell the value of the error term in general because it depends on the function f. What we can do is conclude that, as we shrink the grid spacing Δ to zero for a given f, the error decreases in proportion to Δ^2. We therefore say that the approximation is accurate to second order in Δ.

For comparison, you could derive (1.5.4) in the same way (try it!), and you'd find that the error is proportional to Δ, i.e. (1.5.4) is accurate only to first order. We conclude that (1.5.8) is more accurate than (1.5.4) in the sense that the error decreases more rapidly as $\Delta \to 0$.

We can represent (1.5.8) using a matrix:

$$f_i' = D_{ij}^{(1)} f_j.$$

For example, if $N = 4$, then

$$\underset{\sim}{D} = \frac{1}{2\Delta} \begin{bmatrix} -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 \end{bmatrix}$$

Note, however, that the matrix is not square. It requires values of f at $x_0, x_1, \ldots, x_{N+1}$, but returns the derivative only at $x_1, \ldots, x_N$. For use as part of an eigenvalue problem such as (1.5.3), only a square matrix will do.

One solution is to replace the first and last equations (the top and bottom rows) with expressions that don't depend on f_0 and f_{N+1}. This requires the use of one-sided derivatives, which are derived in the same way as (1.5.8). The simplest choice is to use (1.5.4) for the top row and its counterpart for the bottom row:

$$f_1' = \frac{f_2 - f_1}{\Delta}; \quad f_N' = \frac{f_N - f_{N-1}}{\Delta} \tag{1.5.9}$$

We can represent the result using the matrix

$$\underset{\sim}{D} = \frac{1}{2\Delta}\begin{bmatrix} -2 & 2 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -2 & 2 \end{bmatrix},$$

which we call a derivative matrix Another strategy is to use the top and bottom rows to incorporate boundary conditions, as we discuss in the next section.

1.5.3 Incorporating boundary conditions

If the derivative matrix is intended for use solving a boundary value problem, we can incorporate the boundary conditions into the top and bottom rows instead of (1.5.9). For example, suppose we have the Dirichlet boundary conditions $f(0) = f_0 = 0$ and $f(L) = f_{N+1} = 0$.[4] Then (1.5.8) gives, for $i = 1$ and N,

$$f_1' = \frac{f_2}{2\Delta}; \qquad f_N' = \frac{-f_{N-1}}{2\Delta}, \tag{1.5.10}$$

and the derivative matrix for $N = 4$ becomes

$$\underset{\sim}{D} = \frac{1}{2\Delta}\begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}$$

For higher N, of course, the pattern is repeated through the interior of the matrix.

> **Exercise**: Derive the matrix representing the second derivative at second order, with boundary conditions $f_0 = f_{N+1} = 0$:
>
> $$\underset{\sim}{D}^{(2)} = \frac{1}{\Delta^2}\begin{bmatrix} -2 & 1 & 0 & 0 & \ldots \\ 1 & -2 & 1 & 0 & \ldots \\ 0 & 1 & -2 & 1 & \ldots \\ & & \ddots & & \\ \ldots & 0 & 1 & -2 & 1 \\ \ldots & & & 1 & -2 \end{bmatrix} \tag{1.5.11}$$
>
> Program this matrix in Matlab, with Δ chosen so that $f_0 = f_{N+1} = 0$, and find its eigenvalues. Verify that these correspond to the analytical solution of the differential eigenvalue problem (1.5.3) with the same boundary conditions. Try it with different values of N and see how it affects the accuracy of the results. See if the error decreases in proportion to Δ^2.

[4] In case you don't remember, Dirichlet boundary conditions specify the value of the solution, while Neumann conditions specify the derivative. We'll use both kinds in this course.

1.6 Equations of motion

We'll assume that space is measured by a Cartesian coordinate system $\vec{x} = \{x, y, z\}$, with z directed opposite to gravity. Corresponding unit vectors are $\hat{e}^{(x)}, \hat{e}^{(y)}$ and $\hat{e}^{(z)}$. The velocity is $\vec{u} = D\vec{x}/Dt = \{u, v, w\}$. Here D/Dt represents the material derivative, i.e. the time derivative as measured by an observer moving with the flow:

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + \vec{u} \cdot \vec{\nabla}. \tag{1.6.1}$$

The momentum equation is based on the Boussinnesq approximation, which requires two assumptions:

1. The fluid is incompressible (or nondivergent):

$$\boxed{\vec{\nabla} \cdot \vec{u} = 0.} \tag{1.6.2}$$

   This is also known as the continuity equation.

2. The density remains close to a uniform value ρ_0:

$$\rho = \rho_0 + \rho^*, \quad \text{where } |\rho^*| \ll \rho_0. \tag{1.6.3}$$

We begin with the Navier-Stokes equation for velocities measured in a rotating reference frame:

$$\rho \left[\frac{D\vec{u}}{Dt} - \vec{u} \times f\hat{e}^{(z)} \right] = -\vec{\nabla} p - \rho g \hat{e}^{(z)} + \mu \nabla^2 \vec{u}. \tag{1.6.4}$$

The second term in brackets is the Coriolis acceleration, with $f = f_0 \sin\phi$, where ϕ is the latitude and f_0 is twice the planetary rotation rate (see table 1.1 for values).

We now expand the pressure as

$$p = p_H + p^*,$$

where p_H is in hydrostatic balance with the uniform density ρ_0:

$$\vec{\nabla} p_H = -\rho_0 g \hat{e}^{(z)}.$$

The result is:

$$(\rho_0 + \rho^*) \left[\frac{D\vec{u}}{Dt} - \vec{u} \times f\hat{e}^{(z)} \right] = -\vec{\nabla} p^* - \rho^* g \hat{e}^{(z)} + \mu \nabla^2 \vec{u}.$$

Now, based on (1.6.3), we replace $(\rho_0 + \rho^*)$ on the left-hand side with the constant ρ_0. Finally, we divide through by ρ_0, resulting in

$$\boxed{\frac{D\vec{u}}{Dt} = -\vec{\nabla}\pi + b\hat{e}^{(z)} + \nu \nabla^2 \vec{u} + \vec{u} \times f\hat{e}^{(z)}.} \tag{1.6.5}$$

The accelerations appearing on the right-hand side of (1.6.5) are

- the pressure gradient, with

$$\pi = \frac{p^*}{\rho_0}, \tag{1.6.6}$$

name	symbol	unit	seawater	air
dynamic viscosity	μ	$kg\ m^{-2}s^{-3}$	10^{-3}	1.6×10^{-5}
gravitational acceleration	g	ms^{-2}	9.81	9.81
Coriolis parameter	f_0	s^{-1}	1.458×10^{-4}	1.458×10^{-4}
density	ρ_0	$kg\ m^{-3}$	1027	1.2
kinematic viscosity	ν	m^2s^{-1}	10^{-6}	1.4×10^{-5}
thermal density coefficient	α	K^{-1}	10^{-4}	3×10^{-3}
saline density coefficient	β	psu^{-1}	7×10^{-4}	
thermal diffusivity	κ_T	m^2s^{-1}	1.4×10^{-7}	1.9×10^{-5}
saline diffusivity	κ_S	m^2s^{-1}	10^{-9}	

Table 1.1: Typical terrestrial parameter values to be used in this course. Beware when using these "constants" for general purposes; some of them can vary significantly.

- buoyancy:

$$b = -g\frac{\rho^*}{\rho_0}, \quad \text{and} \tag{1.6.7}$$

- viscosity, with dynamic viscosity μ and kinematic viscosity

$$\nu = \frac{\mu}{\rho_0}. \tag{1.6.8}$$

- The final term is the Coriolis acceleration due to the Earth's rotation. We will omit this term for most of the course, but it will appear in the discussion of baroclinic instability.

The density of seawater is governed by two separate properties, temperature T and salinity S. If the water is close to a uniform state with $T = T_0$ and $S = S_0$, we can use the linearized equation of state:

$$b = \alpha g(T - T_0) - \beta g(S - S_0).$$

Here, α and β are coefficients for thermal and saline buoyancy, taken to be constants. In the absence of sources, T and S obey Fickian diffusion equations with molecular diffusivities κ_T and κ_S:

$$\frac{DT}{Dt} = \kappa_T\nabla^2T; \qquad \frac{DS}{Dt} = \kappa_S\nabla^2S.$$

For most applications, the fact that buoyancy has two components is not important, and we will use a single equation

$$\boxed{\frac{Db}{Dt} = \kappa\nabla^2b.} \tag{1.6.9}$$

Unless otherwise specified, you can think of b as proportional to temperature.

1.6.1 A note on viscosity and turbulence

To the human senses, and in most measurements, a fluid appears as a continuous medium. Although we recognize that a fluid is really made of discrete molecules, the science of fluid mechanics is not concerned with such microscopic details. When we talk about, say, the velocity $\vec{u}$ at a point $\vec{x}$, we really mean an average of $\vec{u}$ over some volume of space, centered on $\vec{x}$, that is tiny but nonetheless large enough to encompass many molecules. With that understanding, we can think of $\vec{u}$ as a continuous function of $\vec{x}$, and therefore employ the powerful tools of calculus to understand the flow.

Although we are not interested in molecular motions *per se*, we must account for the effect they have on the motions that we *are* interested in. That is where viscosity comes in - it represents the frictional effect that molecular interactions exert on the macroscopic motions that we can perceive and measure.[5] The assumption that molecular effects can be represented in this way is called the continuum hypothesis.

In the study of geophysical fluids, the continuum hypothesis is extended to larger scales. We are not only not interested in the motions of individual molecules, but we are also not interested in macroscopic motions smaller than a certain scale. In the study of weather, for example, we do not try to understand every little gust of wind. When we talk about the wind speed at a certain time and place, we usually mean an average that encompasses many wind gusts.

As with molecular motions, though, we must account for the effect the gusts have on the larger-scale motions that we're trying to understand. Exploiting the obvious analogy, the effect of the gusts is usually represented as an "effective" viscosity, often called eddy viscosity or turbulent viscosity. This analogy is highly imperfect. Eddy viscosity is not a property of the fluid but of the flow, and it can vary greatly in space and time in ways we do not understand.

These caveats aside, the eddy viscosity concept is a useful first step towards understanding the effect of small-scale turbulence. The assumption that eddy viscosity is uniform in space and time is, well, better than nothing. In this course, the quantity ν that we call "viscosity" can refer to either molecular or eddy viscosity. Similarly, the diffusivity κ can refer to diffusion either by molecular motions or by small-scale turbulence.

Further reading

An excellent introduction to instability and turbulence in the ocean is

Thorpe, S.A., 2005: *The Turbulent Ocean*, Cambridge University Press.

A more advanced discussion of turbulence theory, with particular attention to eddy viscosity, may be found in chapters 4, 5 and 10 of

Pope, S.B., 2000: *Turbulent flows*, Cambridge University Press.

A derivation of the Navier-Stokes equations (which every student should see at least once) may be found in

Kundu, P.J., I.M. Cohen and D.R. Dowling, 2016: *Fluid Mechanics*, Academic Press.

The Kundu text includes a huge amount of other useful information.

[5] The derivation of the viscous term, leading to the so-called Navier-Stokes equations, is something every student should see at least once. It can be found in any good text on basic fluid mechanics, e.g. (Kundu et al., 2016).