

Class Notes: EAS 471, Atmospheric Modelling

John D. Wilson

January 3, 2007

Contents

I	Numerical Methods in Fluid Mechanics	1
1	Introduction to EAS 471 “Atmospheric Modelling”	2
1.1	History of Numerical Weather Prediction	3
2	General Character of Flow Equations	8
2.1	Classification of terms	10
2.2	Discretization	11
2.2.1	Simplest “Finite Difference” Methodology	11
2.2.2	The Laplacian Operator	13
2.3	The Lagrangian (Material) Derivative	15
2.4	Classification of Equations	18
2.5	Terminology with respect to errors	21
3	Linear stability analysis of numerical schemes	23
3.1	Intuitive Stability Analysis	24
3.2	Von Neumann (or Fourier) Linear Stability Analysis	28
3.2.1	Background	28
3.3	Von Neumann Method	29
3.3.1	VN analysis of the advection equation	31
3.4	Exercises	34
3.5	Computational molecules for the 1-d heat equation	34
3.6	Computational molecules for 1-d advection equation	36
3.7	Numerical (“false”) diffusion	37
4	Aliasing and Non-Linear Computational Instability	39
4.1	Aliasing	39
4.2	Non-linear Computational Instability	43
4.2.1	Getting around NLCI	46

5	Action of the Laplacian as a spatially-selective filter	48
5.1	Analysis of the continuous case (no discretization)	48
5.2	Analysis of the discretized problem	49
6	Alternative Discretization Methods	53
6.1	Control Volume Method	53
6.2	Bubnov-Galerkin Methods: the Spectral and Finite Element methods	59
6.2.1	Finite Element Method	60
6.2.2	Spectral Method	60
6.2.3	Spectral Method applied to Non-linear Advection Equa- tion in 1-d	62
II	Model Dynamics versus Model Physics	66
7	Resolved and Unresolved Scales of Motion	67
7.1	Averaging the Governing Equations	69
7.1.1	Example: Averaging the Vapour Conservation Equation	70
8	Modelling the “free” atmosphere (“dynamics”)	75
9	Modelling the Atmospheric Boundary-Layer (“physics”)	86
9.1	Role of the ABL in NWP	88
9.2	Estimating the unresolved fluxes using K-theory	89
9.2.1	Monin-Obukhov Similarity Theory of the ASL	92
9.3	A vintage treatment of the ABL (old Canadian Regional Finite Element Model)	95
9.4	Boundary treatment for a model with lowest level in the ASL	99
10	Modelling cumulus (“physics”)	103
10.1	Convective adjustment schemes	104
10.2	Cloud-model schemes for cumulus parametrization	107
11	Lagrangian approach to dispersion	114
11.1	Taylor’s Lagrangian Theory	116
11.1.1	Comparison with Classical Eddy-Diffusion Solution . .	119
11.1.2	Asymptotic results	122

11.1.3	What is important about G.I.Taylor's result and when is K-theory valid?	122
11.2	Modern single-particle Lagrangian stochastic models	123
11.2.1	Zeroth- and first-order LS models	124
11.2.2	Time step for (1st order) Lagrangian stochastic simulations	125
11.2.3	Thomson's well-mixed constraint on first-order LS models	126
11.2.4	Unique 1-d LS Model for Gaussian Inhomogeneous Turbulence	127
11.2.5	Reflection at Boundaries	130
11.2.6	LS models applicable in disturbed flows	131
12	Parameterizing radiative flux convergence	132
12.1	Definition of "Intensity" of radiative transfer in a given waveband	134
12.2	Radiative Transfer Equation	135
	Case of a non-scattering but emitting medium	138
	Case where $J = 0$	139
12.3	Interactions of radiation and matter	139
12.4	Scattering Function	142
12.4.1	Rayleigh Scattering	143
	Exercises	145
12.4.2	Mie Scattering	145
12.5	Two-stream model for diffuse solar radiation	146
12.6	Two-stream model for snowpack (spectral) albedo	151
III	Appendices	155
13	Review of classical atmospheric dynamics	156
13.0.1	Expression of the governing equations using a pressure axis	158
13.1	Filtering the dynamical equations to eliminate sound and gravity waves	159
13.1.1	Sound Waves	159
13.1.2	Gravity Waves	160
13.2	Scale Analysis	161

<i>CONTENTS</i>	iv
13.3 Filtered dynamical equations	165
13.3.1 Quasi-geostrophic system of equations	168
13.3.2 Barotropic vorticity equation	169
IV References (etc).	171
References	173

Part I

**Numerical Methods in Fluid
Mechanics**

Chapter 1

Introduction to EAS 471 “Atmospheric Modelling”

The course title is quite general, and under such a heading one could cover modelling radiative energy transfer (perhaps further specialised to radiant energy available with a given layer of a forest to sustain transpiration heating of the airstream and biomass, or to the inversion of satellite radiation measurements to derive atmospheric profiles of temperature and humidity), or the prediction of flow around a windbreak or an isolated tree (micrometeorological problems), or the modelling of air quality downstream from a refinery.

Because many or even most students taking this course will be final year undergraduates likely to become weather forecasters, the course will concentrate on atmospheric modelling as applied to weather forecasting, ie. we will deal almost exclusively with the large scales of atmospheric motion. However much of what must be known in the context of numerical weather predictions carries over directly to modelling of atmospheric flow on different scales, to

modelling of other aspects of the atmosphere, and to modelling of the ocean.

Having established that we will generally confine our attention to NWP, and ignoring the history of man’s developing skill in prediction of, say, aerodynamics (eg. Prandtl’s boundary-layer theory, which is relevant to the treatment of the lower boundary condition of a weather model), let’s look quickly at the history of NWP.

1.1 History of Numerical Weather Prediction

The behaviour of the atmosphere and the ocean is governed by a set of conservation equations. There are many interrelated variables (eg. temperature, density, pressure, humidity, three wind components), each of which depends on 4 independent variables (x, y, z, t) or (more naturally, on the global scale) latitude, longitude, altitude and time. Even our supercomputers are only solving approximations to these equations. Can a person do better? Can a person solve these equations in his/her head? Can a person, by looking at the 850, 700, 500 mb analyses (and whatever other tool he/she wants, but not a machine prog) beat the supercomputer with a prediction for 48 hours ahead? At best, sometimes: maybe!

We need to be a little clearer. Our modern models predict the fields of pressure and wind and humidity on a coarse scale (values predicted are spatial averages that do not reveal local features that may impact on local weather). Predictions of “weather” (local rainfall, etc.) are derived by other means from the model product (model output statistics, perhaps). Because the numerical

models almost certainly do not resolve local features (eg. a range of hills) a human forecaster may perform better locally due to local knowledge (eg. predicting cloud on the upslope of a range not resolved by the model). But on average, forecasting over large regions, a person can NOT do better than the machine. Of course not. The computer simply has computational power that far exceeds (for the type of calculation suited to weather prediction) that of any human. Why even try? We have a better way to forecast: find a suitable approximate form of the governing equations, and solve those approximately on the computer. The place for applying meteorological knowledge now, the way of the present and the future, is to improve the models, - not to try to second guess them by standing before a map and mentally assessing vorticity and thickness advection. The computer knows all about that!

Now, a brief history. Monin () identifies the following major steps - needless to say Newton’s development of mechanics and calculus, and the invention and improvement of barometers and thermometers, were necessary precursors.

- U. Leverrier, director of the Paris Astronomical Observatory, organises a weather service under contract from the French government, becoming, on 19 Feb, 1855, the first person to compile a weather map for the same day.
- V. Bjerknes publishes in 1904 “The problem of weather forecasting, considered from the point of view of mathematics and mechanics.”

Weather prediction recognised as an initial-value problem for the hydrodynamic equations of a baroclinic fluid.

- J. Bjerknes in 1917-1919 introduced the concepts of atmospheric fronts and of cyclone formation due to wave instability on the frontal interfaces. These concepts became the basis for modern synoptic methods of short-range weather prediction, but “up to now” (1972) “have had little significance in numerical forecasting... the spatial grids used in numerical forecasting...do not allow consideration of narrow zones with large hydrodynamic field gradients...In order to take such zones into account in numerical forecasts, it is necessary to develop other methods.”
- L. Richardson (1922) publishes “Weather prediction by numerical process.” His forecast was for the Nuremburg-Augsburg area, 20 May 1910, and used the equations of horizontal motion + continuity equation (ie. essentially a primitive equation method). The forecast was unsatisfactory (predicted pressure changes of order 70 mb in 3 hours, an order of magnitude too large), due to:
 1. incomplete initial data, especially aloft (Initialization and data assimilation; objective analysis)
 2. computational instability: poor finite-difference schemes (some schemes, as we will see, have the property of amplifying small errors). (Discretization methods, numerical integration, smoothing-

filtering-boundary effects)

3. equations used allowed amplification of sound and gravity waves, in part initiated by errors in initial conditions. Thus solutions contained a slowly-varying part, the large-scale motion systems, and a system of fast-moving waves which are for purposes of large-scale prediction “noise.” (use of Primitive Equations vs. reduced equations)

Richardson estimated a team of 64000 people would be needed to keep up with the weather on a global basis. (See “Richardson’s Barotropic Forecast: A Reappraisal,” Lynch, 1992; Bull. Am. Meteorol. Soc., Vol. 73).

- Kibel (1940) proposed the quasi-geostrophic expansion, which results in the rejection (filtering out) of sound and gravity waves.
- Development of high-speed computers (many factors, many people involved). ENIAC installed at Princeton University. Charney and Obukhov (circa 1945-49) systematically introduced the geostrophic and hydrostatic assumptions. First numerical forecast by Charney, Fjortoft, and von Neumann (1950), using “equivalent barotropic model,” a development by Rossby and a special case of the Kibel-Obukhov-Charney derivations. Predicted geopotential height of a single constant-pressure surface (near 500 mb), thus 500 mb winds. No prediction of vertical motion, thus no prediction of what we have come to call “weather.”

CHAPTER 1. INTRODUCTION TO EAS 471 "ATMOSPHERIC MODELLING"7

Multi-level models quickly followed.

Sources:

Haltiner and Williams, p 108. Hess, pp 311-317. Holton, pp 173-175.

Monin, pp 1-3. Petterssen, pp 371-373.

Chapter 2

General Character of Flow Equations

Rather than commencing our study with, say, the Navier-Stokes equations which you may have encountered in another course, we will begin with a *generalised conservation equation*¹ which expresses conservation in coordinate-independent form:

$$\frac{\partial}{\partial t} (\rho \phi) + \nabla \cdot (\rho \vec{u} \phi) = \nabla \cdot (\rho K \nabla \phi) + Q \quad (2.1)$$

The generic dependent variable “ ϕ ” and could be a concentration (eg. specific humidity q , so that the absolute humidity or vapour density $\rho_v = \rho q$); a temperature; or a velocity component. K is the diffusivity [$\text{m}^2 \text{s}^{-1}$]; the term it occurs in may arise from the effects of molecular motion, or it may parameterize turbulent diffusion. Q is the volumetric source term (local rate

¹If you want to get a sense of the derivation and meaning of this equation, you could turn to Holton (2004), who derives specifically the conservation equation for air mass (the “continuity equation”. That equation differs a little from our eqn (2.1), because “air” is neither created nor destroyed in situ, so there is no Q term; and because air does not diffuse in air (no diffusive flux).

of production or destruction of “ ϕ ”).

\vec{u} is the velocity vector, so $\vec{u} \phi$ is the (vector) convective flux density of ϕ . On the other hand $-K\nabla\phi$ is the “diffusive” flux density. As general notation for a vector flux density of ϕ let us write \vec{F} and for its Cartesian components F_i which is a compact way of referring to (F_x, F_y, F_z) .

Now this may well be very unfamiliar, so let’s choose Cartesian coordinates, in which case the velocity vector is (in shorthand)

$$\vec{u} \equiv (u, v, w) \quad (2.2)$$

(where u, v, w are the three velocity components) or more pedantically

$$\vec{u} \equiv \hat{i} u + \hat{j} v + \hat{k} w \quad (2.3)$$

where $\hat{i}, \hat{j}, \hat{k}$ are the unit vectors in the three coordinate directions x, y, z . In the Cartesian coordinate system the “grad operator” (∇) has the representation

$$\nabla \equiv \hat{i} \frac{\partial}{\partial x} + \hat{j} \frac{\partial}{\partial y} + \hat{k} \frac{\partial}{\partial z} \quad (2.4)$$

and we should note it can operate on a scalar “directly” (eg. ∇T is the vector gradient in temperature) and on a vector “indirectly” by means of an “inner product” or “dot product”, eg. the velocity divergence

$$\begin{aligned} \nabla \bullet \vec{u} &\equiv \left(\hat{i} \frac{\partial}{\partial x} + \hat{j} \frac{\partial}{\partial y} + \hat{k} \frac{\partial}{\partial z} \right) \bullet (\hat{i} u + \hat{j} v + \hat{k} w) \\ &= \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} \end{aligned} \quad (2.5)$$

(an alternative notation for the velocity divergence is $\text{div} \vec{u}$, so $\text{div} \equiv \nabla \bullet$).

2.1 Classification of terms

The terms of this (or any flow equation) can be classified as:

- *Storage Terms*: The local time rate of change of the “content” of ϕ , $\partial\phi/\partial t$.
- *Transport Terms*: Terms of form $\partial()/\partial s$ where s is any space coordinate. Such a term is a “transport term” because, upon integration with respect to a coordinate (eg. x) from X_1 to X_2 , it reduces to a difference between the influx at X_1 and efflux at X_2 . Such terms cause redistribution of within the flow domain, but no destruction/creation. Sometimes a term which does not appear to be a transport term can, by manipulation of the equation, be transformed to an explicit transport term. Particularly in the context of numerical modelling, it is advantageous to cast advection terms in flux form. This transformation involves the continuity equation. When flow equations are cast in “flux form” it is easier to derive discretisation procedures which conserve properties which should (by inspection of the corresponding differential equations and b/conds) be conserved.
- *Source/sink Terms*: All terms not of the above two categories are production or destruction terms

If we set $\phi = 1$ and $Q = 0$ (no production of “air”), eqn (2.1) reduces² to the continuity equation,

$$\frac{\partial \rho}{\partial t} + \nabla \bullet (\rho \vec{u}) = 0 \quad (2.6)$$

Furthermore if the density ρ is constant (incompressible flow) the continuity equation reduces to $\nabla \bullet \vec{u} = 0$ and the generalized conservation equation (2.1), which was given in “flux” form, may be written in “advection” form as:

$$\frac{\partial \phi}{\partial t} + \vec{u} \bullet \nabla \phi = \nabla \bullet (K \nabla \phi) + \frac{Q}{\rho} \quad (2.7)$$

Note that the presence of a source term Q in general renders a conservation equation “inhomogeneous” (a linear ordinary differential equation has the form $a_0(x) \frac{d^n y}{dx^n} + a_1(x) \frac{d^{n-1} y}{dx^{n-1}} + \dots + a_{n-1}(x) \frac{dy}{dx} + a_n(x) y = Q(x)$, and it is a homogeneous o.d.e. if $Q(x) = 0$). Typically in fluids problems the dependent variable is “forced” either by an inhomogeneity (source) in its conservation equation, or by the nature of the boundary- or initial-conditions (eg. specification of an inward flux across an upwind boundary on an axis having the ‘1-way’ property defined later).

2.2 Discretization

2.2.1 Simplest “Finite Difference” Methodology

Real flow problems involve variables which are spatially continuous (and have spatially continuous derivatives). But in seeking a numerical solution to a

²Since ϕ is a constant, $\nabla \phi = 0$

flow problem, we cannot retain continuous information on the distribution of the dependent variable: knowledge of the values ϕ at discrete points in space and time must suffice. (Aside: if I choose to represent a variable as a wave such as a sine wave, I can know its value at all points; this is the “spectral” approach to discretization, and in this case one adopts a finite set of waves, which when superimposed, represent approximately, the desired spatial variation. More on this later.)

But then how do we approximate a differential equation (which after all expresses our fundamental knowledge, our conservation principle)? A derivative $\partial\phi/\partial x$ is the limit (as $\Delta x \rightarrow 0$) of the ratio $\Delta\phi/\Delta x$. We cannot shrink $\Delta x \rightarrow 0$ if we only know ϕ at discrete points along the x axis.

Patankar (pp26-31) gives an excellent discussion of “discretization,” and Ames(pp 15-19) more thoroughly discusses the particular discretization method covered here. Note that however we “do it,” the outcome will be a set of “difference equations” or “neighbour equations,” which relate the value ϕ_C of our variable at gridpoint C and time t to its neighbours in space and time.

What are usually called “Finite-Difference” methods for obtaining the discretized equations (ie. the neighbour equations) are obtained by approximating derivatives with a truncated Taylor series

$$\begin{aligned}\phi(x + \Delta x) &= \phi(x) + \left(\frac{\partial\phi}{\partial x}\right)_x \Delta x + \frac{1}{2!} \left(\frac{\partial^2\phi}{\partial x^2}\right)_x \Delta x^2 + \frac{1}{3!} \left(\frac{\partial^3\phi}{\partial x^3}\right)_x \Delta x^3 + \dots \\ \phi(x - \Delta x) &= \phi(x) - \left(\frac{\partial\phi}{\partial x}\right)_x \Delta x + \frac{1}{2!} \left(\frac{\partial^2\phi}{\partial x^2}\right)_x \Delta x^2 - \frac{1}{3!} \left(\frac{\partial^3\phi}{\partial x^3}\right)_x \Delta x^3 + \dots\end{aligned}\tag{2.8}$$

We can approximate $\partial\phi/\partial x$ at x as:

$$\begin{aligned}\frac{\partial\phi}{\partial x} &= \frac{\phi(x + \Delta x) - \phi(x)}{\Delta x} + O[\Delta x] \\ \frac{\partial\phi}{\partial x} &= \frac{\phi(x) - \phi(x - \Delta x)}{\Delta x} + O[\Delta x] \\ \frac{\partial\phi}{\partial x} &= \frac{\phi(x + \Delta x) - \phi(x - \Delta x)}{2\Delta x} + O[\Delta x^2]\end{aligned}\quad (2.9)$$

which are called respectively the *forward*, *backward*, and *central* differences; these approximations are sometimes termed “computational molecules” for the first derivative. The central difference is appealing because its error of approximation is of higher order thus (for small Δx) is smaller.

Similarly, the curvature $\partial^2\phi/\partial x^2$ at x may be approximated:

$$\frac{\partial^2\phi}{\partial x^2} = \frac{\phi(x + \Delta x) + \phi(x - \Delta x) - 2\phi(x)}{\Delta x^2} + O[\Delta x^2]\quad (2.10)$$

Ames (1977, p 17) pictures these (and more complex) computational molecules.

2.2.2 The Laplacian Operator

In a Cartesian coordinate system the “Laplacian operator” (or “diffusion” operator) has the representation

$$\nabla^2 \equiv \frac{\partial^2}{\partial x_j \partial x_j}\quad (2.11)$$

where the Einstein summation convention is implied; expanding the summation,

$$\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\quad (2.12)$$

and (eg.) in 1-dimension, $\nabla^2 = \partial^2/\partial x^2$.

I now want to show that the Laplacian operator is a smoothing operator. First, let us note from eqn (2.10) that its simplest computational molecule can be said to compare the “central” value with its neighbours. If the central value is “too large” relative to the average of its neighbours, the term will be negative, and vice versa.

Consider the action of the Laplacian in the 1-d heat equation³

$$\frac{\partial T}{\partial t} = K \frac{\partial^2 T}{\partial x^2} \quad (2.13)$$

where the right-hand side can be seen as the divergence (in 1-d) of a conductive heat flux:

$$\frac{\partial T}{\partial t} = - \frac{\partial}{\partial x} \left(- K \frac{\partial T}{\partial x} \right) \quad (2.14)$$

We could imagine this is governing heat flow in an infinite rod (temperature being uniform on each slice across the x -axis). If at $t = 0$ the temperature distribution in the rod is sinusoidal, we will consider the distribution of the flux, the consequent flux divergence, and the resulting tendency in the temperature. We will see the temperature is “smoothed out.”

Suppose we represent the temperature at discrete points (labelled I) along the x -axis, these points being separated by a constant gridlength Δx . Our other independent variable is t , which we also discretize, on intervals Δt such that $t^n = n \Delta t$. Then T_I^n is the temperature at the I^{th} gridpoint at time n .

³We could equally well call this a 1-d “diffusion equation”. One of the computing assignments will be to solve this equation numerically, using two or more different discretization schemes involving different choices for the necessary computational molecules.

A possible discretisation of the equation is:

$$\frac{T_I^{n+1} - T_I^n}{\Delta t} = K \frac{T_{I+1}^n + T_{I-1}^n - 2 T_I^n}{\Delta x^2} \quad (2.15)$$

which on rearrangement gives the convenient formula:

$$T_I^{n+1} = T_I^n + \frac{K \Delta t}{\Delta x^2} (T_{I+1}^n + T_{I-1}^n - 2 T_I^n) \quad (2.16)$$

This is called an “explicit” algorithm, because its form is “single unknown = combination of knowns”. It is now clear that the diffusion (∇^2) term is a smoother; the new value of T at a given point is bigger (smaller) than the preceding value depending on whether that preceding value is smaller (bigger) than the average of its neighbours. Thus perhaps we could say that, in this context anyway, the ∇^2 operator is a sort of “envy operator.”

2.3 The Lagrangian (Material) Derivative

Consider the trajectory (see Fig. 2.1) of a particular “fluid particle” (a “fluid particle” or “fluid element” is a volume so small as to be considered a point, yet contains vast numbers of molecules). If $X_i(t)$ labels the position of this particle⁴ then its velocity is $U_i = \frac{dX_i}{dt} = (U, V, W)$. X_i is not a field variable, unlike the the coordinate x_i , ie. it makes no sense to write $X_i = X_i(x_i)$, for a given particle cannot be at more than one position at a given time.

Let $q = q(x_i, t)$ be any continuum property of the element. Then the fluid element that is at $X_i(t)$ at time t “carries” $q(X_i(t), t)$ and by the chain rule

⁴Note: We shall generally use capitals to distinguish a “Lagrangian” (ie. particle-following or flow-following) property.

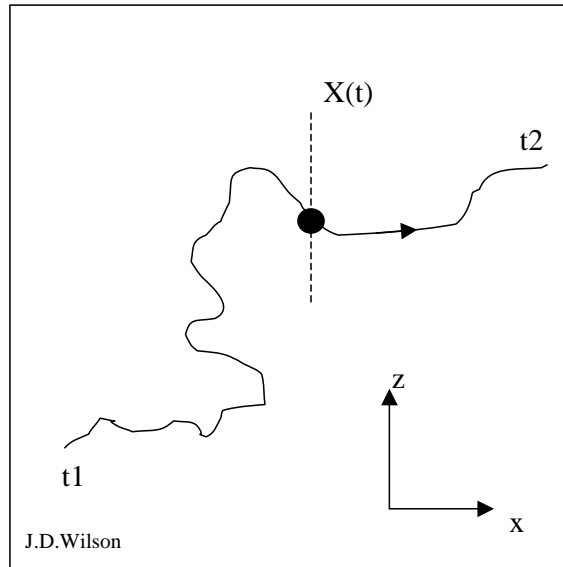


Figure 2.1: A fluid element (particle) trajectory, showing the particle at time t when its position on the x -axis is X .

of differentiation, the change in $q = q(X_i(t), t)$ over interval dt is

$$dq = \left(\frac{\partial q}{\partial t} \right)_{x_i=X_i(t)} dt + \left(\frac{\partial q}{\partial X_i} \right)_t dX_i \quad (2.17)$$

where I have added subscripts to emphasize which variables are held constant during the partial differentiation, and (of course) the repeated subscript i implies summation. We may make the substitution $dX_i = U_i dt$.

How do we interpret $\left(\frac{\partial q}{\partial X_i} \right)_t$? If t is to be held fixed, then how can we have a dX_i , ie. a change in particle position? Well, what we want is the

spatial gradient of q at time t at the location then occupied by our particle, that is, at $x_i = X_i(t)$. So it seems to me we can write

$$\left(\frac{\partial q}{\partial X_i}\right)_t \equiv \left(\frac{\partial q}{\partial x_i}\right)_t \quad (2.18)$$

Upon introducing this simplification and dividing by dt we have:

$$\frac{dq}{dt} = \frac{\partial q}{\partial t} + U_i \frac{\partial q}{\partial x_i} = \frac{\partial q}{\partial t} + \vec{u} \bullet \nabla q \quad (2.19)$$

which expresses the “total” (or “Lagrangian” or “material”) time derivative d/dt as the sum of the “local tendency” $\partial/\partial t$ and “advection”. I have dropped the subscripts on the partial derivatives, for compactness⁵.

Now having introduced the Lagrangian derivative we now have the choice to express our flow equations in a simpler form, eg. the continuity equation (2.6) may be re-written by exploiting the fact that $\nabla \bullet (\rho \vec{u}) = \vec{u} \bullet \nabla \rho + \rho \nabla \bullet \vec{u}$, and by combining terms we have

$$\frac{d\rho}{dt} = -\rho \nabla \bullet \vec{u} \quad (2.20)$$

or

$$\frac{1}{\rho} \frac{d\rho}{dt} = -\nabla \bullet \vec{u} \quad (2.21)$$

The fractional rate of change of density, following a fluid element, is equal to (minus) the velocity divergence, for which reason $\nabla \bullet \vec{u}$ is also given the name: “volumetric dilatation rate”.

⁵I encourage you to think carefully about the meanings of all the steps performed here, and the meanings of all the symbols.

2.4 Classification of Equations

We will take a very informal and schematic approach to this: for more detail see Ames (1977, pp3-5), Patankar (1980, pp 20-22), or texts in mathematics of pde's. The character of a p.d.e. is effectively dictated by the highest order derivatives in the equation, which in fluid mechanics are usually second order spatial derivatives, ie. curvature terms. Recall that a curvature term (eg. on the x axis, $\partial^2\phi/\partial x^2$) couples the dependent variable ϕ to its neighbourhood on both sides, so it implies two-way influence, or two-way spatial connectivity.

“Equilibrium” problems (Ames, p3) are “problems of steady state in which the equilibrium field of ϕ in a domain D is to be determined by solving the differential equation within D subject to conditions on ϕ on the boundary.” These are called “jury problems,” because “the entire solution is passed on by a jury requiring satisfaction of all the boundary conditions and all the internal requirements.” The differential equation will be “elliptic” in the spatial variables, ie. will contain the term $\nabla^2\phi$, ie. involve the Laplacian of the dependent variable⁶.

A prototypical elliptic problem is $\nabla^2\phi = 0$ with $\phi = 1$ on the walls. We saw above that the computational molecule for the Laplacian operator ($\nabla^2\phi$) “looks in all directions” and says “I must look like the average of my neighbours: we must be uniform.” The Laplacian is a smoother in space

⁶Classification of the governing equations logically belongs before one embarks into discretized equations; but it is helpful to understand the action of ∇^2 before addressing the classification.

(however many space dimensions are present). A bit of reasoning should convince you that the solution to this problem must be $\phi = 1$ everywhere.

Virtually all real problems involving mass, momentum, or energy conservation are governed by differential equations that contain diffusion, ie. equations which contain the Laplacian term. Why? Most such problems admit fluxes (transport) due to molecular motion, ie. diffusion, conduction and viscous drag. The prototypical form of these molecular fluxes is $-K \nabla \phi$, where K [$\text{m}^2 \text{s}^{-1}$] is the (appropriate) diffusivity. For reasons that should be apparent from our discussion of the generalized conservation equation, it is the “divergence ” $\nabla \bullet (K \nabla \phi) = K \nabla^2 \phi$ of such fluxes that appears in the governing equations (here for simplicity I assumed the diffusivity K is constant). So most problems will have a diffusion term operating along each spatial coordinate. But often the diffusive transport term along a particular coordinate is tiny compared to a convective flux⁷. In such a case that coordinate direction is, practically, a “one-way coordinate” (see below), and the diffusion term is insignificant in relation to convection. In this situation conditions downstream of a given point do not affect its tendency. A consequence is that no downwind boundary condition is needed. In view of this is useful to extend our concept of the “ellipticity” of a problem, by assigning that character axis by axis... for due to approximations a problem may

⁷Let’s compare advection of water vapour by a $U \sim 20 \text{ m s}^{-1}$ wind with molecular diffusion. The diffusivity of water vapour in air is about $D_w \sim 2 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$. If the inhomogeneity of the water vapour field can be expressed as $\Delta \rho_v L^{-1}$ where L is a length scale, then the ratio of the convective to diffusive fluxes (called the Peclet number) is $(U \Delta \rho_v)/(D_w \Delta \rho_v L^{-2}) = UL^2/D_w$, a huge number.

end up having “elliptic” character on a reduced number of spatial dimensions. Thus we may have ellipticity on only one space axis (eg. our equation retains $\partial^2\phi/\partial x^2$ but is first-order on the other space-axes).

In “marching” or “propagation” problems, “the solution marches out from the initial state guided and modified in transit by the side boundary conditions.”

In time-dependent jury problems, the value of ϕ evolves from some initial value subject to boundary constraints (the equations used in weather prediction are of this form). We have seen that the simplest “computational molecule” for the ∇^2 operator at a given grid point “C” compares the value of ϕ_C with the average of its values at the neighbouring points. In a time-dependent jury problem, an excess in the value of ϕ_C relative to the average of its neighbours will cause a tendency for ϕ_C to be reduced at the subsequent time, ie. the ∇^2 operator is a “smoothing” operator. It is associated with diffusion of the property ϕ , ie. the ∇^2 terms normally appear in the equations as the expression of the physical process of diffusion (either molecular diffusion, or, in the case of some models of turbulent convective transfer, turbulent diffusion).

A “one-way (two-way) coordinate” is such that conditions at a given location in that coordinate are influenced by changes in conditions on only one side (both sides) of that location. Assuming time travel to be impossible, time is presumably a one-way coordinate. Very commonly, equations involving fluid motion which is close to being parallel to a wall (or the ground) are

approximated in such a way that the direction of the flow becomes a one-way coordinate: such problems are solved efficiently by “marching” downstream from a given line along which the solution is known.

2.5 Terminology with respect to errors

Truncation error (Haltiner & Williams, 1980, p 120; Haltiner, 1971, p105; Holton, 2004, p460) “is defined to be the difference between the difference equation and the differential equation.” Thus we can see that if we employ the forward difference for our advection term, for example, our truncation error is of order Δx . The difference equation is said to be *consistent* or *compatible* with the differential equation, provided the truncation error vanishes as the grid intervals and the timestep are reduced indefinitely in size.

Suppose we could obtain an exact solution ϕ_* to the neighbour equations (this is sometimes possible), and our numerical solution is ϕ_{Num} . The exact solution to the differential equation is ϕ . Then:

$$\text{Discretization error} = \phi - \phi_*$$

$$\text{Stability error} = \phi_* - \phi_{Num}$$

A difference solution ϕ_{Num} is said to be “convergent” if it approaches the true solution ϕ as $\Delta t, \Delta x \rightarrow 0$.

Lax Equivalence Theorem: If a difference equation is consistent with the differential equation it represents then stability is the necessary and sufficient condition for convergence. (JW’s comment: presumably there must also be a consistency condition on the numerical initial and boundary conditions too.)

Chapter 3

Linear stability analysis of numerical schemes

Sometimes an apparently reasonable discretization of the governing equations and apparently reasonable choice of grid-lengths and timestep, can lead to an “explosion” of the (computed) solution having nothing to do with the reality of the flow.

The linked equations solved in a realistic weather model are (in the case of those equations expressing conservation of momentum) non-linear, and are solved on a non-uniform grid. However quite simple considerations of the “linearised building blocks” of those more-complex equations can suggest necessary conditions for the stability of a numerical procedure. Methods of formal linear stability analysis are well developed. We will look only at one method, the Von Neumann method. But before doing so, let us note that even simpler thinking can suggest the arising of stability problems.

3.1 Intuitive Stability Analysis

Here and henceforth, let superscript “ n ” imply time $t = n\Delta t$, and indices I, J denote $x = I\Delta x, y = J\Delta y$ (etc).

Example 1: A discretization of the 1-d storage+diffusion equation (heat equation) on a uniform grid, using a forward time difference and the simplest computational molecule for the spatial curvature (Truncation Error of $O(\Delta t + \Delta x^2)$) is:

$$\frac{\phi_I^{n+1} - \phi_I^n}{\Delta t} = \frac{K}{\Delta x^2} [\phi_{I+1}^n + \phi_{I-1}^n - 2\phi_I^n] \quad (3.1)$$

Now assume that at time n the solution is a “two-gridpoint oscillation in space,” ie., $\phi_I^n = (-1)^I$, where (please note) I shan’t bother to carry the “Num” subscript that reminds us this is a numerical solution.

Now our intuition should tell us that at any later time $n+1, n+2, \dots$ the amplitude (of this, or any other wave present at time n) should be *smaller*.

Why? Well, here are the pertinent facts.

- this is a problem in which there are no volumetric sources of ϕ
- and it is a linear problem, so there can be no wave-wave interactions that might increase amplitudes of given fourier components of the solution (this point may be obscure, for now)
- more specifically, it contains only the storage term $\partial\phi/\partial t$ and *diffusion*, and the diffusion term, as we have seen, acts so as to smooth (or damp) the solution progressively as time increases.

Then in view of this intuition, we are entitled to say that if $\phi_I^{n+1} > 1$, not only is the wave is growing, but in particular this is unphysical behaviour - ie. there is must be some artificial (numerical) instability.

Taking the case where I is ODD so $\phi_I^n = -1$, a condition for stability is:

$$\frac{K\Delta t}{\Delta x^2} \leq \frac{1}{2} \quad (3.2)$$

The dimensionless ratio $K\Delta t/\Delta x^2$, which can be thought of as being the reciprocal of a grid Reynolds number $\Delta x (\Delta x/\Delta t)/K$ where K is the effective viscosity, is sometimes called the diffusion number, and the above equation the “diffusion limit.” The conditional instability of this numerical method for the heat equation is confirmed by formal methods.

Example 2: A discretization of the 1-d advection equation using centred time and space differences (truncation error of order $O(\Delta t^2 + \Delta x^2)$) is:

$$\frac{\phi_I^{n+1} - \phi_I^{n-1}}{2\Delta t} = -U \frac{\phi_{I+1}^n - \phi_{I-1}^n}{2\Delta x} \quad (3.3)$$

This is a “leapfrogging scheme,” so-called because of the occurrence of ϕ at three times $n - 1, n, n + 1$. In one of our assignments we will solve this problem numerically.

Now the exact solution for $\phi(x, t)$ is

$$\phi(x, t) = \phi(x - Ut, 0) = \phi_0(x - Ut) \quad (3.4)$$

where $\phi_0(x) = \phi(x, 0)$ is the initial field of ϕ . Recall that we know this is the solution, since the differential equation tells us ϕ is constant following the

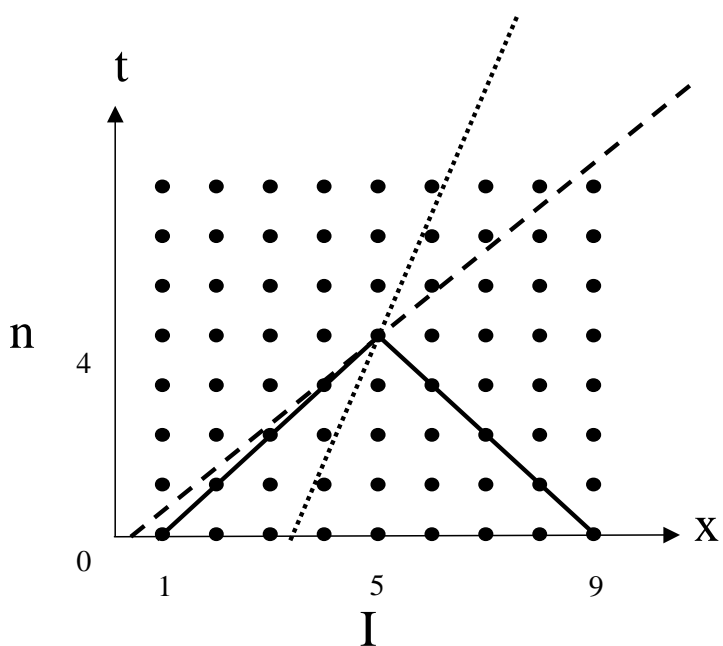


Figure 3.1: Cone of influence; from Haltiner and Williams p121

motion. Thus if we were to have an exact numerical solution on the grid it would be expressible as something like

$$\phi_I^n = \phi(I\Delta x, n\Delta t) = \phi_0(I\Delta x - U n \Delta t) \quad (3.5)$$

Figure (3.1) is a time-space ($t-x$) diagram which shows how the “domain of influence” upon ϕ_I^n expands out in space and backward in time from the point $n = 4, I = 5$, for an arbitrary choice of $\Delta x, \Delta t$. The slope of the cone of influence is $\Delta x/\Delta t$, which is a velocity defined by our choices of $\Delta x, \Delta t$. Also shown are curves $x = Ut$ drawn upstream from the point of interest, ϕ_5^4 , for two possible values of the advection velocity U . The long-dashed line lies outside the domain of influence, hence in that case there is no way for the true solution $\phi_0(x_I - Un\Delta t)$ to influence the numerical solution: the numerical solution cannot be correct. Conversely, the dotted curve, for a smaller U , lies within the domain of influence, hence the correct solution is able to influence the numerical solution. Generalising, the choice of grid spacing and timestep is acceptable provided the path $x = Ut$ lies within the domain of influence. This requires that the Courant number C obey the restriction:

$$C \equiv \frac{U \Delta t}{\Delta x} \leq 1 \quad (3.6)$$

This is called the “Courant-Friedrichs-Levy” (CFL) condition, which was set out in 1928; again, the CFL condition may be derived formally. According to Lindzen and Fox-Rabinowitz (1989; *Monthly Weather Review*, Vol. 117), the CFL condition “calls for a time step that is usually much smaller than the time scale associated with the dominant spatial scale of the phenomena

being described. The point is that as long as we resolve the smaller scales for which the CFL condition is violated, these scales will eventually be excited by non-linearity or even roundoff error.” Also according to Lindzen and Fox-Rabinowitz, “a similar consistency requirement exists between vertical and horizontal resolution.... ; ... excessive horizontal resolution could resolve modes whose vertical wavelength might be too small to be resolved with the existing vertical resolution, and this situation could lead to spatial instability.”

3.2 Von Neumann (or Fourier) Linear Stability Analysis

This is a simple method of stability analysis which is applicable only to finite difference schemes on a uniform grid approximating linear equations. The analysis can be applied to a linearised approximation to a non-linear equation to determine a necessary condition for stability. No account is taken of the influence of boundary conditions.

3.2.1 Background

If $\phi_1, \phi_2, \dots, \phi_N$ are solutions to a linear homogeneous differential equation $L(\phi) = 0$, (where $L()$ is a differential operator¹) then so too is the linear combination $\phi = c_1\phi_1 + c_2\phi_2 + \dots + c_N\phi_N$, where the c 's are constants (Churchill, 1969, p 25). Thus it is very common to represent a function $\phi(x)$, the solution

¹For example in one space dimension x we can write $L \equiv a + b\frac{\partial}{\partial x} + c\frac{\partial^2}{\partial x^2} + \dots$

to such an equation on a finite or infinite interval, by a “series representation” - a superposition of a (possibly infinite) set of special functions. If the special functions are sines/cosines (or exponentials $e^{jk_x x}$ where $j = \sqrt{-1}$ and k_x is the wavenumber) we have the “Fourier series.” Dirichlet (1829) established the general conditions on a function sufficient to ensure the convergence of its Fourier series to the values of the function.

When the solution is a function of several independent variables (say x, y, z, t) it is possible in many cases to assume the variables are separable, viz.

$$\phi(x, y, z, t) = \phi_t(t) \phi_x(x) \phi_y(y) \phi_z(z) \quad (3.7)$$

In such cases typically the series solution is composed of (or can be expressed as) a superposition of “modes:”

$$\exp \left[j \left(\omega t + \vec{k} \cdot \vec{x} \right) \right] \quad (3.8)$$

where ω is the angular frequency (which in general has both real and imaginary parts), and \vec{k} (or k_i) is the vector-wavenumber (which we assume to be purely real). If ω and \vec{k} may vary continuously, the superposition is of the form of an integral ($d\omega dk_x dk_y dk_z$); otherwise in the form of a summation over discrete values of ω, k_j .

3.3 Von Neumann Method

In the VN method we obtain an exact solution ϕ^* to the discretization equation which is approximating the differential equation (obviously therefore the

method will only work for sufficiently simple equations), and deduce conditions required to ensure ϕ^* is bounded on the time axis.

The x, y, z axes in effect extend from $-\infty$ to ∞ , and along these axes we have a discrete representation of our solution. Had we the inclination and time, we should dive into “representation” of a function on a discrete axis (gridlength Δx). We would find, for example, that the maximum meaningful wavenumber for the x axis is $k_{x(MX)} = 1/(2\Delta x)$, and that waves of shorter wavelength (larger wavenumber) cannot “exist” (or be represented) with the given grid spacing. We would also get bogged down in worrying about the conditions under which a series representation is possible: does the axis extend over all time? If not (say in the case of the t -axis), is there periodicity along that axis?

Rather than being so methodical, we will assume that the variables separate, and that the essence of our exact solution to the discretized equation is

$$\phi^*(I\Delta x, n\Delta t) = e^{j(\omega n \Delta t + k I \Delta x)} \quad (3.9)$$

with obvious generalisation to more than one spatial dimension (the grid interval Δx and timestep Δt are both positive). This wave has arbitrary frequency and wavenumber², and from it any general solution can be composed by linear combination. If this wave can be shown to have finite amplitude as

²By mysterious convention, we shall assume the frequency may be complex, but that the wavenumber is real. Recall (Churchill and Brown, 1990) that if $z = x + jy$ is a *non-zero* complex number then it can be represented in polar form as $z = r e^{j\theta}$ where $r = \sqrt{x^2 + y^2}$ is the magnitude of z and $\theta = \arctan(y/x)$ is called the ‘argument’ of z (sometimes $\arg z$).

t becomes large, the discretization equation is considered stable when applied to the given uniform grid. The condition for stability is therefore:

$$|e^{j\omega\Delta t}| \leq 1 \quad (3.10)$$

which guarantees that $|\exp(j\omega n\Delta t)| \leq 1$.

Let $\omega = \omega_R + j\omega_I$ where ω_R, ω_I are both real. Then,

$$e^{j\omega\Delta t} = e^{j\omega_R\Delta t} e^{-\omega_I\Delta t} \quad (3.11)$$

Now for any real θ ,

$$\begin{aligned} |e^{j\theta}| &= |\cos\theta + j\sin\theta| \\ &= \sqrt{\cos^2\theta + \sin^2\theta} \\ &= 1 \end{aligned} \quad (3.12)$$

so

$$|e^{j\omega\Delta t}| = e^{-\omega_I\Delta t} \quad (3.13)$$

For stability, then, we require that $\omega_I \geq 0$. If ω is a purely real number, neutral stability is implied.

3.3.1 VN analysis of the advection equation

Let's take the case of central differencing in time and space:

$$\frac{\phi_I^{n+1} - \phi_I^{n-1}}{2\Delta t} + U \frac{\phi_{I+1}^n - \phi_{I-1}^n}{2\Delta x} = 0 \quad (3.14)$$

Substituting a single mode $\exp[j(\omega n \Delta t + k I \Delta x)]$ of arbitrary frequency and wavenumber (ω, k) we have

$$\begin{aligned} & e^{j(\omega(n+1)\Delta t + k I \Delta x)} - e^{j(\omega(n-1)\Delta t + k I \Delta x)} \\ = & -\frac{U\Delta t}{\Delta x} \left(e^{j(\omega n \Delta t + k(I+1)\Delta x)} - e^{j(\omega n \Delta t + k(I-1)\Delta x)} \right) \end{aligned} \quad (3.15)$$

Now we divide through by the common factor $\exp[j(\omega n \Delta t + k I \Delta x)]$ and the difference equation may be written:

$$e^{j\omega\Delta t} - e^{-j\omega\Delta t} = -\frac{U\Delta t}{\Delta x} (e^{jk\Delta x} - e^{-jk\Delta x}) \quad (3.16)$$

Of course we never forget³ that since for any real θ

$$\begin{aligned} e^{j\theta} & \equiv \cos\theta + j\sin\theta \\ e^{-j\theta} & \equiv \cos\theta - j\sin\theta \end{aligned} \quad (3.17)$$

by sums and differences

$$\begin{aligned} e^{j\theta} + e^{-j\theta} & \equiv 2\cos\theta \\ e^{j\theta} - e^{-j\theta} & \equiv 2j\sin\theta \end{aligned} \quad (3.18)$$

so (provided wavenumber k is real) we can readily extract a quadratic equation for $r = e^{j\omega\Delta t}$, namely

$$r^2 + r \left[2j \frac{U\Delta t}{\Delta x} \sin(k\Delta x) \right] - 1 = 0 \quad (3.19)$$

³This is ‘Euler’s formula’ and (see Churchill and Brown, 1990) it is to be regarded as *defining* the symbol $e^{j\theta}$ or $\exp(j\theta)$.

The two solutions for r are:

$$\begin{aligned} r_+ &= -j \frac{U\Delta t}{\Delta x} \sin(k\Delta x) + \sqrt{1 - \left[\frac{U\Delta t}{\Delta x} \sin(k\Delta x) \right]^2} \\ r_- &= -j \frac{U\Delta t}{\Delta x} \sin(k\Delta x) - \sqrt{1 - \left[\frac{U\Delta t}{\Delta x} \sin(k\Delta x) \right]^2} \end{aligned} \quad (3.20)$$

Now remember that letting $\omega = \omega_R + j\omega_I$ (where ω_R, ω_I are both real) we have $|r| = |\exp(-\omega_I\Delta t)|$ and our condition for stability is that

$$|r| = |e^{-\omega_I\Delta t}| \leq 1 \quad (3.21)$$

which implies our scheme is stable provided $|r_+| \leq 1$ and $|r_-| \leq 1$. Define

$$\alpha = 1 - \left[\frac{U\Delta t}{\Delta x} \sin(k\Delta x) \right]^2 \quad (3.22)$$

such that our two solutions are

$$r_{\pm} = -j \sqrt{1 - \alpha} \pm \sqrt{\alpha} \quad (3.23)$$

Case 1

If, for all k , we have $1 \geq \alpha \geq 0$ then both $\alpha^{1/2}$ and $(1 - \alpha)^{1/2}$ are real. It follows that:

$$|r_{\pm}| = 1 - \alpha + \alpha = 1 \quad (3.24)$$

and we have neutral stability. Since $\sin(k\Delta x)$ can be as large as 1, this neutral stability is only ensured if $U\Delta t/\Delta x \leq 1$, the CFL condition.

Case 2

However if $\alpha < 0$, ie.

$$1 - \left(\frac{U\Delta t}{\Delta x}\right)^2 \sin^2(k\Delta x) < 0 \quad (3.25)$$

(which can only be true generally if $U\Delta t/\Delta x > 1$, ie. if the CFL condition is violated), then $|r_{\pm}|$ may exceed unity: numerical instability.

3.4 Exercises

Perform a Von Neumann stability analysis of one of the following discretizations of the advection equation:

$$\begin{aligned} \frac{\phi_I^{n+1} - \phi_I^n}{\Delta t} &= -U \frac{\phi_{I+1}^n - \phi_{I-1}^n}{2\Delta x} \\ \frac{\phi_I^{n+1} - \phi_I^n}{\Delta t} &= -U \frac{\phi_I^n - \phi_{I-1}^n}{\Delta x} \end{aligned} \quad (3.26)$$

and of one of the following discretizations of the heat equation

$$\begin{aligned} \frac{\phi_I^{n+1} - \phi_I^n}{\Delta t} &= K \frac{\phi_{I+1}^n + \phi_{I-1}^n - 2\phi_I^n}{\Delta x^2} \\ \frac{\phi_I^{n+1} - \phi_I^{n-1}}{2\Delta t} &= K \frac{\phi_{I+1}^n + \phi_{I-1}^n - 2\phi_I^n}{\Delta x^2} \end{aligned} \quad (3.27)$$

3.5 Computational molecules for the 1-d heat equation

A numerical approximation to the diffusion equation or heat equation (eqn 2.13) using a central time difference and the standard computational molecule

for the diffusion term (Richardson's method) is unconditionally unstable (can you prove this using Von Neumann stability analysis?).

In the *Dufort-Frankel* method we modify the standard molecule for the curvature (ie. for a component of the Laplacian operator) by replacing $-2 T_I^n$ with $(-T_I^{n+1} - T_I^{n-1})$ we obtain an unconditionally stable, three time level method for the heat equation

$$\frac{T_I^{n+1} - T_I^{n-1}}{2\Delta t} = K \frac{T_{i+1}^n + T_{i-1}^n - T_I^{n+1} - T_I^{n-1}}{\delta x^2} \quad (3.28)$$

which is explicit, because only T_I^{n+1} appears as an unknown.

In the *Crank-Nicholson* discretization one takes a forward difference for the time derivative and a linear combination of two molecules for the diffusion term:

$$\begin{aligned} \frac{T_I^{n+1} - T_I^n}{\Delta t} &= \lambda K \frac{T_{i+1}^{n+1} + T_{i-1}^{n+1} - 2 T_I^{n+1}}{\delta x^2} \\ &+ (1 - \lambda) K \frac{T_{i+1}^n + T_{i-1}^n - 2 T_I^n}{\delta x^2} \end{aligned} \quad (3.29)$$

If $\lambda = 0$, we have an explicit formula which is conditionally unstable (stability criterion is ??). If $\lambda = 1/2$, we have the scheme of Crank & Nicholson, which is unconditionally stable. This introduces the notion of an implicit scheme: we see that at the new (unknown) time level $(n + 1)$ we do not have T_I^{n+1} in isolation, but the neighbouring values at $I + 1, I - 1$ as well - ie. the difference equation for the unknown T_I^{n+1} contains two other unknowns as well. Therefore with this formula we cannot, from the earlier-time solution, step forward to obtain the solution at time $n+1$ by a simple, explicit formula.

The problem is closed only in the sense that the set of difference equations will contain as many equations as there are unknowns. In this case, solution will necessitate the use of an iterative technique or, more directly, solution of a “tridiagonal” matrix problem - the solution at $(I, n + 1)$ is linked to the solution at $(I - 1, n + 1)$ and $(i + 1, n + 1)$. More on this later.

3.6 Computational molecules for 1-d advection equation

The 1-d advection equation is:

$$\frac{\partial \phi}{\partial t} + U \frac{\partial \phi}{\partial x} = 0 \quad (3.30)$$

The *Euler method*

$$\frac{\phi_I^{n+1} - \phi_I^n}{\Delta t} = -U \frac{\phi_{I+1}^n - \phi_{I-1}^n}{2\Delta x} \quad (3.31)$$

uses a forward difference in time, and a central difference in space. It is explicit, and unconditionally unstable.

The *upstream differencing* method uses a forward difference in time, and an upwind difference in space. It is explicit. One expects a damped or neutral solution provided

$$0 \leq \frac{U \, dt}{dx} \leq 1 \quad (3.32)$$

The *trapezoidal implicit* scheme uses a forward difference in time, while for the space derivative the average of the central-difference for two times is

taken,

$$\frac{\phi_I^{n+1} - \phi_I^n}{\Delta t} + \frac{U}{2} \left(\frac{\phi_{I+1}^{n+1} - \phi_{I-1}^{n+1}}{2\Delta x} + \frac{\phi_{I+1}^n - \phi_{I-1}^n}{2\Delta x} \right) = 0 \quad (3.33)$$

This scheme is implicit, and unconditionally stable.

The *Euler backward* (or *Matsuno*) scheme is an explicit trial-step method

$$\begin{aligned} \phi_I^{*,n+1} &= \phi_I^n - \frac{U\Delta t}{2\Delta x} (\phi_{I+1}^n - \phi_{I-1}^n) \\ \phi_I^{n+1} &= \phi_I^n - \frac{U\Delta t}{2\Delta x} (\phi_{I+1}^{*,n+1} - \phi_{I-1}^{*,n+1}) \end{aligned} \quad (3.34)$$

Here $\phi_I^{*,n+1}$ is a first guess for ϕ_I^{n+1} , which is used in the corrected second step. This scheme is stable so long as $|U\Delta t/\Delta x| \leq 1$. Another explicit 2-step scheme that is subject to the CFL condition is the Lax-Wendroff method (see Haltiner and Williams, p149).

3.7 Numerical (“false”) diffusion

Suppose we have a pure-advection problem: say, the uniform advection along a periodic x-axis of a sin wave:

$$\frac{\partial \phi}{\partial t} + U \frac{\partial \phi}{\partial x} = 0 \quad (3.35)$$

on $-L \leq x \leq L$, with $U = \text{constant}$. Let $\phi(x, 0) = \sin(kx)$, with $k = 2n\pi/L$.

We know the exact solution to this problem. There is no “diffusion.” Whatever initial “form” we have for the initial wave, ie. no matter how “sharp,” our system retains that initial form, or “sharpness.”

However a numerical solution to this problem might NOT have that property of preserving sharpness. In that case, there would be said to arise “false,” or “numerical” diffusion.

Suppose for our discretization of advection we use an upstream difference:

$$\left(\frac{\partial\phi}{\partial x}\right)_I^n = \frac{\phi_I^n - \phi_{I-1}^n}{\Delta x} \quad (3.36)$$

Since one may write a Taylor series expansion

$$\phi_{I-1}^n = \phi_I^n - \Delta x \left(\frac{\partial\phi}{\partial x}\right)_I^n + \frac{\Delta x^2}{2} \left(\frac{\partial^2\phi}{\partial x^2}\right)_I^n + \dots \quad (3.37)$$

it is true by rearrangement that

$$U \frac{\phi_I^n - \phi_{I-1}^n}{\Delta x} = U \left(\frac{\partial\phi}{\partial x}\right)_I^n - \frac{U \Delta x}{2} \left(\frac{\partial^2\phi}{\partial x^2}\right)_I^n + \dots \quad (3.38)$$

Then, by using this upstream difference we might say the equation we are really solving is

$$\frac{\partial\phi}{\partial t} + U \frac{\partial\phi}{\partial x} = \frac{U \Delta x}{2} \frac{\partial^2\phi}{\partial x^2} \quad (3.39)$$

which (by virtue of the non-zero term on the r.h.s.) contains diffusion, which has arisen due to truncation error. See Patankar (p105) for discussion.

Chapter 4

Aliasing and Non-Linear Computational Instability

4.1 Aliasing

A grid with spacing Δx can represent only waves having wavelength

$$\lambda \geq \lambda_{min} = 2\Delta x \quad (4.1)$$

or in terms of wavenumber $k = 2\pi/\lambda$

$$k \leq k_{max} = \pi/\Delta x \quad (4.2)$$

To see that this is so, multiply onto a grid $x(I) = I\Delta x$ the wave $\phi = \sin[k_{max}(1 + \epsilon)x]$ where $0 < \epsilon \ll 1$ (this curve is plotted on Figure 4.1).

The values of ϕ at the grid points are¹

$$\phi(I) = \sin \left[\frac{\pi}{\Delta x} (1 + \epsilon) I \Delta x \right] = \sin [\pi I + \pi I \epsilon] = (-1)^I \sin [\pi I \epsilon] \quad (4.3)$$

Table (4.1) lists the values of the function (c.f. Fig 4.1), and in comparison,

¹ $\sin(A + B) = \sin A \cos B + \cos A \sin B$, and let $A = \pi I \dots$

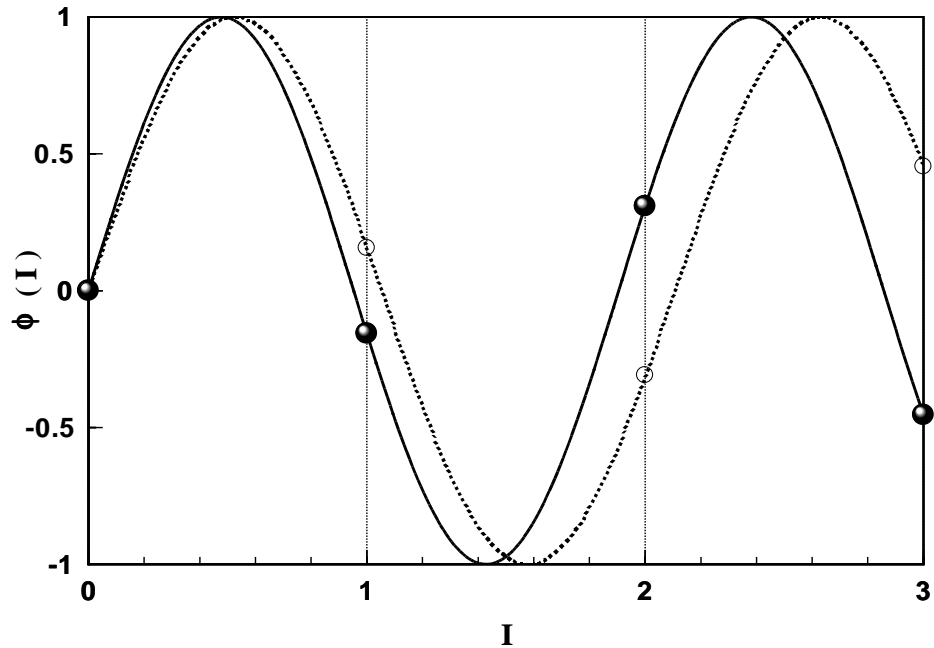


Figure 4.1: *Illustrating the function $\phi = \sin[\pi(1 + 0.05)I]$ (solid curve) and its representation on the discrete grid (symbols \bullet). The dashed line is the curve $\sin[\pi(1 - 0.05)I]$ and symbols (\circ) show its representation. All information along the curve(s) is lost - hidden - and so effectively, to within a sign change, our ' k_{max}^+ ' curve has the same representation as the (slower) ' k_{max}^- ' curve.*

the values of $\phi^* = \sin[k_{max}(1 - \epsilon)x]$. Since $\epsilon \ll 1$ the factor $\sin(\pi\epsilon I)$ varies slowly: but our $\phi(I)$ oscillates with I ; it is a two-gridpoint wave. The wave which “actually” has wavenumber $k_{max}(1 + \epsilon)$ and so is larger than k_{max} has

been *aliased* as a longer wave with wavenumber $k_{max}(1 - \epsilon)$. In general, a wave at $k = k_{max} + \delta k$ (where $\delta k < k_{max}$) is represented on the discrete grid as a wave of lower wavenumber $k^* = 2k_{max} - k = k_{max} - \delta k$.

The phenomenon of ‘aliasing’ is most often encountered in the subject of the “spectral analysis” of a random or quasi-random signal, eg. a time series $u = u(t)$ or a spatial series $g = g(x)$, wherein we seek to compute or interpret the ‘power spectral density function’ (or loosely, the ‘variance spectrum’, or simply ‘spectrum’), defined to be that function of frequency (or wavenumber) that, when integrated over the entire range in frequency (or wavenumber), gives the variance $\sigma_u^2 \equiv \overline{u'^2}$ of the signal, viz.

$$\begin{aligned}\overline{u'^2} &= \int_0^\infty S_u(f) df \\ \overline{g'^2} &= \int_0^\infty S_g(k) dk\end{aligned}\tag{4.4}$$

Now, when we seek to give a statistical description of any time or space series we have many options. Most basically and familiarly, we might list the mean, variance (etc.), these being ‘moments’ of the probability density function (or ‘pdf’), eg.

$$\overline{u'^2} \equiv \int_{-\infty}^\infty (u - \bar{u}) p(u) du\tag{4.5}$$

Less familiarly, we might state the spectrum, or the autocovariance function (these two constitute a ‘Fourier transform pair’). Different information is conveyed by the different options. The key point of relevance here is that the power spectral density can be said (loosely) to convey the *amplitudes*

Table 4.1: Example of the aliasing of an exact function $\phi(x)$ when represented on a discrete grid $x(I) = I\Delta x$, for which the finite interval Δx entails that the maximum representable wavenumber will be $k_{max} = \pi/\Delta x$. Here the function represented on the grid is $\phi = \sin [k_{max}(1 + \epsilon)x]$ with $\epsilon = 0.05$.

I	$\pi\epsilon I$	$\phi(I)$	$\phi^*(I)$
0	0	0	0
1	0.05 π	-0.156	0.156
2	0.10 π	+0.309	-0.309
3	0.15 π	-0.454	+0.454
4	0.20 π	+0.588	-0.588
5	0.25 π	-0.707	+0.707
6	0.30 π	+0.809	-0.809
7	0.35 π	-0.891	+0.891
8	0.40 π	+0.951	-0.951
9	0.45 π	-0.988	+0.988
10	$\pi/2$	+1.000	-1.000
11	0.55 π	-0.988	+0.988
12	0.60 π	+0.951	-0.951
13	0.65 π	-0.891	+0.891

of the various Fourier modes needed to represent the signal, and therefore information on the relative *phases* of the modes is lost, or is not ‘contained’ in the spectral description.

Thus returning to our two waves $k_{max}(1 \pm \epsilon)$ we saw that their representations on the discrete grid differ only by a sign change, ie. a phase shift of π radians. In the context of a spectrum, then, there is no distinction between these two modes. Thus we say the mode $k_{max}(1 + \epsilon)$ has been ‘aliased’ to (or falsely represented as) the mode $k_{max}(1 - \epsilon)$.

4.2 Non-linear Computational Instability

When a finite differencing method is applied to the non-linear advection equation

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = 0 \quad (4.6)$$

the CFL restriction, that $u\Delta t/\Delta x \leq 1$, is a necessary but not a sufficient condition to ensure computational stability.

Why? Let us assume that at time n our discrete solution on the grid $x = I\Delta x$ is a wave of wavenumber k and amplitude U^n ,

$$u_I^n = U^n e^{j k (I\Delta x)} \quad (4.7)$$

Then if we use a central difference for $\partial u/\partial x$

$$u \frac{\partial u}{\partial x} = (U^n)^2 e^{j k I \Delta x} \frac{e^{j k (I+1) \Delta x} - e^{j k (I-1) \Delta x}}{2\Delta x} \quad (4.8)$$

and this is readily simplified to

$$u \frac{\partial u}{\partial x} = \frac{j (U^n)^2}{\Delta x} \sin(k\Delta x) e^{j 2k I \Delta x} \quad (4.9)$$

Now if we use a forward difference in time (results would be similar using a leapfrog scheme) we have:

$$u_I^{n+1} = u_I^n + \frac{\partial u}{\partial t} \Delta t = u_I^n - u \frac{\partial u}{\partial x} \Delta t \quad (4.10)$$

which yields:

$$u_I^{n+1} = U^n e^{j k I \Delta x} - j \frac{\Delta t}{\Delta x} (U^n)^2 \sin(k\Delta x) [\cos(2k I \Delta x) + j \sin(2k I \Delta x)] \quad (4.11)$$

So u_I^{n+1} has a real component

$$U^n \left[\frac{\Delta t}{\Delta x} U^n \sin(k\Delta x) \right] \sin(2k I \Delta x) \quad (4.12)$$

which is a wave at double the starting wavenumber. To be sure, it may only have small amplitude

$$U^n \left[\frac{\Delta t}{\Delta x} U^n \sin(k\Delta x) \right] \quad (4.13)$$

since we presumably impose the CFL condition. But every step will accentuate the problem.

What's happening? A wave-wave interaction in the non-linear term has lead to the production of a shorter wave (higher wavenumber). In our special case it is in fact an interaction of waves of the same wavenumber, but you

can appreciate that if u_I^n is some more complex superposition of waves, say (adding only minimal complexity for starters)

$$u_I^n = U_1^n \sin(k_1 I \Delta x) + U_2^n \sin(k_2 I \Delta x) \quad (4.14)$$

then $u \partial u / \partial x$ will produce interactions of form

$$\sin(k_1 I \Delta x) \cos(k_2 I \Delta x) = \frac{1}{2} [\sin((k_1 + k_2) I \Delta x) + \sin((k_1 - k_2) I \Delta x)] \quad (4.15)$$

Here, starting with waves k_1 and k_2 , we have gathered a longer wave $k_1 - k_2$ and a shorter wave $k_1 + k_2$. For a more-general superposition as the starting point, $u \partial u / \partial x$ will contain many wave cross-products.

We are gathering power at higher wavenumber (steepening gradients). Now this is physically proper: that is what should be happening (in the real meteorological equations other terms would moderate this behaviour). But what if, as this progresses, we generate power at wavenumbers too high to be represented on the grid?, ie. at $k > \pi / \Delta x$. For example, if we start with our “pure” wave k taking the value $k = k_{max} (1 + \epsilon) / 2$ we come out one timestep later with wave

$$\sin(2k I \Delta x) = \sin(\pi I + \pi \epsilon I) = (-1)^I \sin(\pi \epsilon I) \quad (4.16)$$

which due to aliasing, instead of being represented as it should be, a wave shorter than k_{max} , has appeared as a wave just long enough to be represented on the grid.

NLCI of a finite difference scheme, then, is the following: the non-linear term(s), such as $u \partial u / \partial x$, cause (and correctly so!) wave-wave interactions

that produce shorter wave components; the problem is that when these are too short for the grid, they are aliased, so that the finite difference method falsely accumulates energy in the highest wavenumbers representable on the grid.

4.2.1 Getting around NLCI

Possibilities are:

1. Use a finite difference scheme that damps short waves. This is sometimes achieved by including a bogus smoothing (diffusion) term ∇^2 . Unless this fictitiously added term has a typical magnitude that is small compared to the real terms, one has thereby changed the dynamics of the flow model. In turbulent flow such a diffusion term is present anyway, the expression of a turbulent diffusion process; and that may be sufficient to control the gradient-sharpening action of the non-linear advection term.
2. Use a spectral method of discretization, rather than finite differencing. Here variation along the axes (say x) is restricted to be of known, analytical form (polynomial or trigonometric curves) and, in a sense, there is no grid - discretisation takes place in wavenumber space rather than physical space. Thus if we include only the truncated set of wavenumbers k_1, k_2, \dots, k_N (where k_N is the shortest wave included) then the non-linear interaction of k_{N-1} and k_N to produce $k_{SUM} = k_{N-1} + k_N > k_N$

is not problematical, because k_{SUM} lies outside the truncated set of included waves.

Chapter 5

Action of the Laplacian as a spatially-selective filter

Let us re-examine the smoothing action of the Laplacian operator in the 1-d heat equation (our prototypical ‘diffusion’ equation)

$$\frac{\partial T}{\partial t} = K \frac{\partial^2 T}{\partial x^2} \quad (5.1)$$

5.1 Analysis of the continuous case (no discretization)

The temperature $T(x, t)$ and its continuous Fourier transform $\hat{T}(k, t)$ are related as:

$$\begin{aligned} \hat{T}(k, t) &= \int_{-\infty}^{\infty} T(x, t) e^{-jkx} dx \\ T(x, t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{T}(k, t) e^{jkx} dk \end{aligned} \quad (5.2)$$

and from the second equation

$$\frac{\partial^2 T}{\partial x^2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} (jk)^2 \tilde{T}(k, t) e^{jkx} dk \quad (5.3)$$

so that, substituting

$$\int_{-\infty}^{\infty} \frac{\partial \tilde{T}}{\partial t} e^{jkx} dk = -K \int_{-\infty}^{\infty} k^2 \tilde{T}(k, t) e^{jkx} dk \quad (5.4)$$

or

$$\frac{\partial \tilde{T}(k, t)}{\partial t} = -K k^2 \tilde{T}(k, t) \quad (5.5)$$

(check units: does $k^2 K$ have units s^{-1} ?). Thus

$$\frac{\partial \ln \tilde{T}(k, t)}{\partial t} = -k^2 K \quad (5.6)$$

and we can solve to obtain the general solution

$$\int_0^t \frac{\partial \ln \tilde{T}(k, t)}{\partial t} dt = -k^2 K t \quad (5.7)$$

or

$$\tilde{T}(k, t) = \tilde{T}(k, 0) e^{-k^2 K t} \quad (5.8)$$

Well that's a lot of symbols! What does this say? The Fourier mode \tilde{T} whose wavenumber is k is *damped exponentially* as time increases, and the strength of the damping increases as $k^2 K$ (where K is the diffusivity and k is the wavenumber). The 'sharpest' waves suffer the fastest damping. Thus the ∇^2 operator is *scale selective*.

5.2 Analysis of the discretized problem

We'll discretize on a uniform grid (intervals $\Delta x, \Delta t$) using the usual $O(\Delta x^2)$ estimate for the curvature. With a forward difference in time we have

$$T_I^{n+1} = T_I^n + \frac{K \Delta t}{\Delta x^2} (T_{I+1}^n + T_{I-1}^n - 2 T_I^n) \quad (5.9)$$

To simplify, let's use the notation $\gamma = K\Delta t/\Delta x^2$. Then collecting terms,

$$T_I^{n+1} = \gamma T_{I-1}^n + (1 - 2\gamma) T_I^n + \gamma T_{I+1}^n \quad (5.10)$$

Now this is pretty, but how to interpret it? Let's recall that probably, one would choose $\gamma < \frac{1}{2}$ although we are entitled to remain agnostic on that. The basic point is that we have a *digital filter* here, in that the new value T_I^{n+1} at gridpoint I is determined by the three prior values T_J^n , ($J = I - 1, I, I + 1$) which are given the respective weighting factors $\gamma, 1 - 2\gamma, \gamma$. For example if $\gamma = 1/4$ we have a $\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$ filter¹, a particular type of a 'block average'. Just to clarify a little, we can say that we are dealing with a 'recursive filter', in that having computed T_I^{n+1} we then pop it into the right hand side in place of T_I^n , and proceed to compute T_I^{n+2} . It is also pertinent to note (without proof) that the symmetry of the weights about the centre point I has the consequence that the *phase* of any of the modes composing our $T(x, t)$ does not evolve in time.

So what is the 'spectral' consequence of this? Suppose that

$$T_I^n = A^n e^{jkI\Delta x} \quad (5.11)$$

Substituting into eqn (5.10), two skips and a jump later you have that the amplitude gain

$$\frac{T_I^{n+1}}{T_I^n} = 1 - 2\gamma [1 - \cos(k\Delta x)] \quad (5.12)$$

¹Note that things get bizarre if $\gamma > \frac{1}{2}$... the prior value T_I^n at I has a negative influence on T_I^{n+1} .

and the power gain is the square on this. Recall that our formula is only relevant for $k \leq \pi/\Delta x$. Figure (5.1) plots the power gain as a function of $k\Delta x$ out to $k\Delta x = \pi$. Note that the discretized Laplacian for sufficiently small diffusion number γ is a low pass filter (no damping of long waves, maximal damping of the sharpest waves), but note also the warning sign that as our diffusion number γ creeps up towards the stability limit ($\gamma = 1/2$) the power gain at $k\Delta x \rightarrow \pi$ begins to creep up.

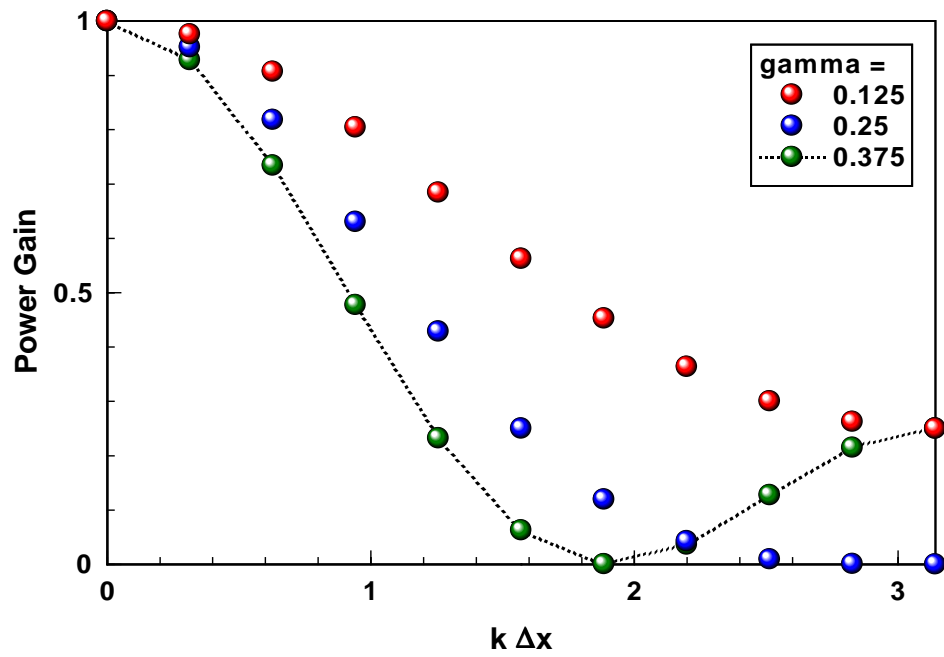


Figure 5.1: Illustrating the impact of the Laplacian operator in the heat equation, in terms of its impact (over a single timestep) on the power of any given mode (wavenumber k) of the solution. The single step power gain is plotted for $\gamma = K\Delta t/\Delta x^2 = \frac{1}{8}, \frac{1}{4}, \frac{3}{8}$, where K is the diffusivity and Δx is the spatial grid interval.

Chapter 6

Alternative Discretization Methods

Our problems and exercises have been with difference (discretisation/neighbour) equations obtained by the finite difference method. In the simulations of incompressible flow often seen in the engineering literature, *Control Volume* methods are increasingly popular. The present generation of medium range forecast models typically use finite differences on the time and vertical axes, but sometimes use a *Spectral scheme* in the horizontal. Thus, you should have some understanding of the more complex discretisation methods. According to Ames, all of these alternative schemes fall under the heading “Methods of Weighted Residuals.”

6.1 Control Volume Method

Patankar calls the control-volume method a “pre-calculus expression of conservation.” Space is divided into adjoining boxes, or control volumes (cv).

The control volumes are imaginary - they are invisible to the flow. The governing equations are (analytically) integrated for application to a control volume: the result is a formulation in terms of (i) fluxes (convective and diffusive) across the cv walls (ii) production/destruction within the cv (iii) changes in storage within the cv.

I'll illustrate by looking at a diffusion problem involving the spread of a passive tracer, concentration c [kg m^{-3}]. We'll consider the flow to be laminar and known (the method generalises easily to calculate dispersion in turbulent flow) and the concentration field to be steady, and uniform along the crosswind (y) axis (stationary, 2-d concentration field).

The governing mass conservation equation is

$$\frac{\partial c}{\partial t} = 0 = -\nabla \cdot (\vec{u} c - D \nabla c) + Q \quad (6.1)$$

where Q [$\text{kg m}^{-3} \text{s}^{-1}$] is the volumetric source/sink term. More explicitly, and assuming we limit to two space dimensions,

$$\frac{\partial}{\partial x} \left(u c - D \frac{\partial c}{\partial x} \right) + \frac{\partial}{\partial z} \left(w c - D \frac{\partial c}{\partial z} \right) = Q \quad (6.2)$$

Now divide up the space into control volumes¹ as in Figure (6.1). Integrating the differential equation for application within such a volume, we

¹For the sake of simplicity we will consider an “unstaggered grid”, ie. a grid at which all dependent variables are co-located at the same gridpoints. This is to be contrasted with the “staggered grid” on which (eg.) pressure gridpoints are offset relative to velocity gridpoints.

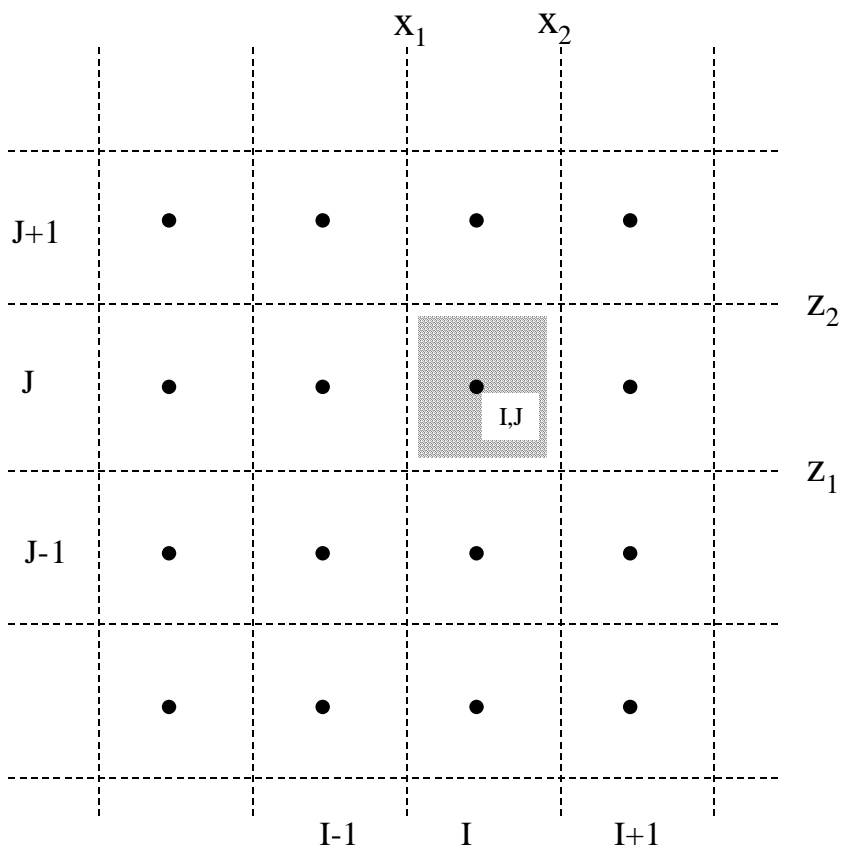


Figure 6.1: Gridpoints on a uniform mesh, and centred within their control volumes. The dashed lines show control volumes faces, eg. x_1, x_2 are the yz -planes defining where faces of the IJ^{th} control volume cut the x -axis.

have

$$\int_{x_1}^{x_2} \int_{z_1}^{z_2} \left[\frac{\partial}{\partial x} \left(u c - D \frac{\partial c}{\partial x} \right) + \frac{\partial}{\partial z} \left(w c - D \frac{\partial c}{\partial z} \right) - Q \right] dx dz = 0 \quad (6.3)$$

which yields

$$\Delta z \left[u c - D \frac{\partial c}{\partial x} \right]_{x_1}^{x_2} + \Delta x \left[w c - D \frac{\partial c}{\partial z} \right]_{z_1}^{z_2} + Q \Delta x \Delta z = 0 \quad (6.4)$$

where $\Delta x, \Delta z$ are the face lengths and $[\cdot]_1^2$ denotes a difference in the argument between location 2 and location 1.

Now we clearly have an equation involving fluxes (convective plus diffusive) across faces - ie. exchanges among control volumes. To progress, we have to express the fluxes across the faces in terms of our resolved values of the concentration.

For simplicity let's make all the grid intervals $\Delta x = \Delta z = \Delta$ equal. Let $c(I, J)$ be the concentration and $u(I, J)$ the alongstream velocity at a gridpoint in the middle of the IJ control volume. Then using an upstream scheme for convection, we could estimate:

$$[u c]_{x_1}^{x_2} = \frac{u(I, J) + u(I + 1, J)}{2} c(I, J) - \frac{u(I - 1, J) + u(I, J)}{2} c(I - 1, J) \quad (6.5)$$

where we have assumed a linear variation of velocity between gridpoints and thus simply averaged. The diffusion term could be written:

$$\left[D \frac{\partial c}{\partial x} \right]_{x_1}^{x_2} = D_{x_2} \frac{c(I, J + 1) - c(I, J)}{\Delta} - D_{x_1} \frac{c(I, J) - c(I, J - 1)}{\Delta} \quad (6.6)$$

In this way we obtain our discretisation equation. Patankar gives a much better interpolation scheme than I have suggested here for the effect of the combination of advection and diffusion across control volume faces.

Example: Calculation of the mean concentration of a gas released continuously from a ground-level line source lying across the wind. Atmosphere assumed neutrally-stratified, wind and turbulence at a given height depending only on two parameters, the friction velocity u_* and the surface roughness length z_0 . The mean wind is

$$u(z) = \frac{u_*}{0.4} \ln \left(\frac{z}{z_0} \right) \quad (6.7)$$

while the mean vertical component (w in the preceding equations) vanishes. We can neglect diffusion along the direction of the mean wind, since advection is overwhelmingly more important; and the (turbulent) diffusivity along the vertical (z) axis is $D = (k_v/S_c)u_*z$ where the von Karman constant $k_v = 0.4$ and the Schmidt number S_c is subject to some uncertainty but appears to be best given the value $S_c \approx 0.6 - 0.7$.

Our discretization equation reduces to

$$\Delta z [u c]_{x_1}^{x_2} = \Delta x \left[D \frac{\partial c}{\partial z} \right]_{z_1}^{z_2} \quad (6.8)$$

(incidentally note our problem is elliptic on the z -axis). Using an upstream difference for the streamwise advection term gives

$$[u c]_{x_1}^{x_2} = u(J) [c(I, J) - c(I - 1, J)] \quad (6.9)$$

(and ensures us a 1-way x -coordinate, ie. a marching problem in x) while for the vertical transport term we can write

$$\begin{aligned} \left[D \frac{\partial c}{\partial z} \right]_{z_1}^{z_2} &= \frac{k_v u_*}{S_c} \left[z_2 \frac{c(I, J+1) - c(I, J)}{\Delta z} - z_1 \frac{c(I, J) - c(I, J-1)}{\Delta z} \right] \\ &= D_n \frac{c(I, J+1) - c(I, J)}{\Delta z} - D_s \frac{c(I, J) - c(I, J-1)}{\Delta z} \quad (6.10) \end{aligned}$$

where $D_n = k_v u_* z_2 / S_c$ is the diffusivity at z_2 (etc). Now if we collect terms we have

$$\begin{aligned} c(I, J) \left[u(J)\Delta z + \frac{\Delta x D_n}{\Delta z} + \frac{\Delta x D_s}{\Delta z} \right] &= c(I, J+1) \frac{\Delta x D_n}{\Delta z} \\ &+ c(I, J-1) \frac{\Delta x D_s}{\Delta z} \\ &+ c(I-1, J) u(J)\Delta z \quad (6.11) \end{aligned}$$

This is clearly a “marching problem” in that if we know our solution at streamwise location $I-1$ then we have a closed set of equations for the column matrix $c(I, J)$, $J = 1, 2, 3, \dots$. It is an implicit-type computational problem, because our unknown at J is linked to its (unknown) neighbours at $J \pm 1$. Thus, when we compute the solution at streamwise location I (already knowing the solution at $I-1$) we have a problem of form:

$$A_J^C c(I, J) = A_J^N c(I, J+1) + A_J^S c(I, J-1) + B_{I,J}, \quad J = 1..J_{mx} \quad (6.12)$$

where the neighbour coefficients (ie. the vectors A_J^C, A_J^N, A_J^S) are (in our case) fixed vectors, though the “what lies behind” vector B_J will be different for the step $I-1 \rightarrow I$ than for the step $I-2 \rightarrow I-1$ (etc).

6.2 Bubnov-Galerkin Methods: the Spectral and Finite Element methods

Suppose in a given physical problem the desired solution $\phi(x)$ must satisfy the differential equation

$$L[\phi(x)] = f(x) \quad (6.13)$$

In the Bubnov-Galerkin methods the solution is expressed as a superposition of linearly independent basis functions $\theta_j(x)$, viz.

$$\phi(x) = \sum_{j=1}^N a_j \theta_j(x) \quad (6.14)$$

Usually N is finite (“truncated set of basis functions”). If the basis functions satisfy the boundary conditions but not the differential equation, this is called a boundary weighted residual method; if they satisfy the differential equation but not the boundary conditions, this is an internal method.

At any point x the error is

$$e(x) = L \left[\sum_{j=1}^N a_j \theta_j(x) \right] - f(x) \quad (6.15)$$

The free coefficients a_j are optimised by forcing the covariance of the error with each of the basis functions to vanish, viz.

$$\int e(x) \theta_j(x) dx = 0, \quad j = 1 \dots N \quad (6.16)$$

Note that the above equation may be interpreted as stating that the “inner product” defined by

$$(e, \theta_j) = \int e(x) \theta_j(x) dx \quad (6.17)$$

must vanish for each j ; and thus that by definition the error function $e(x)$ is orthogonal to each of the basis functions. Probably a more physically-illuminating interpretation is that the integral (or inner product) is a covariance (average along x of the cross-product of $e(x)$ and the basis function $\theta_j(x)$); by forcing that covariance to vanish, we are forcing the error to lie randomly about each of the basis functions (by suitably shaping the basis functions).

The above can be generalised to multi-dimensional problems.

6.2.1 Finite Element Method

Here the basis functions are only locally non-zero: the $\theta_j(x)$ vanish, except in a limited region about each point of interest (grid point). Typically they are chosen to be “tent” functions (also called “chapeau” functions). These functions, eg. along a single space dimension $\theta_j(x)$, are continuous functions of their argument (here x), vanishing except where

$$\begin{aligned}\theta_j(x) &= \frac{x - x_{j-1}}{\Delta x}, & x_{j-1} \leq x \leq x_j \\ \theta_j(x) &= \frac{x_{j+1} - x_j}{\Delta x}, & x_j \leq x \leq x_{j+1}\end{aligned}\tag{6.18}$$

6.2.2 Spectral Method

The θ_j are chosen to be orthogonal functions defined on the entire x axis. In plane geometry they will often be harmonic functions $e^{j(kx+ly)}$. In modern NWP applications they are usually spherical harmonics, with the time dependence of the solution residing in the coefficients (and finite differencing

in the vertical). The usual orthogonal basis functions are

$$Y_{m,n}(\mu, \lambda) = P_{m,n}(\mu) e^{j m \lambda} \quad (6.19)$$

where λ is the longitude and $\mu = \sin \phi$ where ϕ is the latitude. Clearly the $e^{j m \lambda}$ implies sines and cosines to represent the zonal structure. The $P_{m,n}$ represent the meridional structure; they are the associated Legendre functions of the 1st kind. For example a model using pressure p as the vertical coordinate would represent the (height-dependent) velocity as:

$$u(\phi, \lambda, p, t) = \sum_m \sum_n u_{m,n}(p, t) Y_{m,n}(\mu, \lambda) \quad (6.20)$$

The essential difference between spectral and grid point models is that the meteorological fields are represented by a finite sum of wave components of different amplitude and wavelength, rather than by values on a regularly-spaced mesh. One of the attractive consequences of this is that no differencing is required (in the horizontal directions); horizontal derivatives can be evaluated analytically, because the dependent variables are represented by continuous functions. In particular, finite difference approximations to the non-linear advection terms are not needed, so the spectral model does not suffer from NLCI (the generation of small scale noise). However this does not mean a spectral model has no truncation error. The truncation error of a spectral model is related to the choice of J , the maximum wavenumber which determines the smallest wave represented by the model. If J is small, only large waves are allowed. In nature, wave-wave interactions generate ever-shorter waves, so a forecast made with a model allowing only a very

few long wave components would soon depart from reality. Typical spectral models of the 80's carried about 80 waves; a grid model of corresponding resolution would require 2 gridpoints per each shortest wavelength, and so would have a grid spacing of about $1/2 \cdot 360/80 = 2.25^\circ$ (about 250 km). During integration of a spectral model, transformations are made back to physical space (as opposed to wavenumber space); the spherical grid array that results is called a Gaussian grid. It is easier to calculate the physics in physical space.

6.2.3 Spectral Method applied to Non-linear Advection Equation in 1-d

We want to solve

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = 0 \quad (6.21)$$

on $-L \leq x \leq L$ subject to $u(-L, t) = u(L, t)$.

Choose as basis functions

$$\theta_n = e^{j k_n x} \quad (6.22)$$

where $k_n = n\pi/L$ and $n = -N, -N + 1, \dots, -1, 0, 1, \dots, N - 1, N$. These basis functions automatically force the solution to satisfy the boundary conditions.

The inner product will for present purposes be

$$(f, g) = \int_{-L}^L f(x, t) g^*(x, t) dx \quad (6.23)$$

where the * denotes the complex conjugate. The solution is written:

$$u(x, t) = \sum_{-N}^N u_n(t) e^{j n\pi x/L} \quad (6.24)$$

where $u_n(t)$ is the coefficient of the n th wave, in general a complex number (the u_n carry the time-dependence of $u(x,t)$). Substituting into the advection equation, the error function is:

$$e(x) = \sum_{n=-N}^N \frac{du_n}{dt} e^{j n \pi x/L} + \sum_{n=-N}^N u_n(t) e^{j n \pi x/L} \sum_{m=-N}^N \frac{j m \pi}{L} u_m(t) e^{j m \pi x/L} \quad (6.25)$$

To determine the optimal coefficients $u_n(t)$, multiply the error through by $e^{-j p \pi x/L}$ and integrate over the range of x , requiring

$$\int_{-L}^L e(x) e^{-j p \pi x/L} dx = 0 \quad p = -N, \dots, N \quad (6.26)$$

Then

$$\begin{aligned} 0 &= \sum_{n=-N}^N \frac{du_n}{dt} \int_{-L}^L e^{-j p \pi x/L} e^{j n \pi x/L} dx \\ &+ \sum_{n=-N}^N u_n(t) \sum_{m=-N}^N \frac{j m \pi}{L} u_m(t) \int_{-L}^L e^{j n \pi x/L} e^{j m \pi x/L} e^{-j p \pi x/L} dx \end{aligned} \quad (6.27)$$

But (Spiegel, Advanced Mathematics, p187)

$$\begin{aligned} &\int_{-L}^L e^{j m \pi x/L} e^{-j p \pi x/L} dx \\ &= \int_{-L}^L \left(\cos \frac{m \pi x}{L} + j \sin \frac{m \pi x}{L} \right) \left(\cos \frac{p \pi x}{L} - j \sin \frac{p \pi x}{L} \right) dx \\ &= 2 L \delta_{m,p} \end{aligned} \quad (6.28)$$

so we have

$$\frac{du_p}{dt} = - \sum_{n=-N}^N \sum_{m=-N}^N \frac{j m \pi}{2L^2} u_n(t) u_m(t) R_{m,n}^p \quad (6.29)$$

This is a set of $2N + 1$ coupled ordinary differential equations for the advancement in time of the complex coefficients $u_p(t)$, $p = -N \dots N$. The interaction coefficients

$$R_{m,n}^p = \int_{-L}^L e^{j(m+n)\pi x/L} e^{-jp\pi x/L} dx = 2L \delta_{m+n,p} \quad (6.30)$$

vanish unless $m + n = p$, in which case waves m, n are interacting to produce a wavenumber $p = m + n$ component. They may be calculated in advance and stored for all the interactions which will be retained. Splitting the coefficients into their real and imaginary parts,

$$u_n(t) = u_n^R(t) + j u_n^I(t) \quad (6.31)$$

it is straightforward to show that

$$\begin{aligned} \frac{du_p^R}{dt} &= \sum_{n=-N}^N \sum_{m=-N}^N \frac{m\pi}{L} [u_n^I(t) u_m^R(t) + u_m^I(t) u_n^R(t)] \delta_{m+n,p} \\ \frac{du_p^I}{dt} &= \sum_{n=-N}^N \sum_{m=-N}^N \frac{m\pi}{L} [u_n^I(t) u_m^I(t) - u_n^R(t) u_m^R(t)] \delta_{m+n,p} \end{aligned} \quad (6.32)$$

Procedure: We choose a value for N , say $N = 10$ (21 waves). The initial velocity field must now be represented by the decomposition:

$$u(x, 0) = \sum_{n=-10}^{10} u_n(t) e^{jn\pi x/L} \quad (6.33)$$

Suppose we initialize with the special case of a single long cosine wave corresponding to $n = 1$, ie. $u(x, 0) = \cos(\pi x/L)$. Then only the coefficient u_1 is non-zero at commencement of the integration (ie. $u_1(0) \neq 0$), with $u_1^R(0) = 1$ and $u_1^I(0) = 0$. But after the first timestep, at time Δt , we will have non-zero

$u_2(t)$ due to the interaction $1 + 1 \rightarrow 2$. At $2\Delta t$ we have non-zero $u_3(2\Delta t)$ due to the interaction $1 + 2 \rightarrow 3$. Soon all coefficients will be non-zero, after which the further steepening of the gradients implied by the physics is lost due to the retaining of only the limited set of waves.

The non-linear advection equation is a pathological example. When the spectral procedure is applied to weather prediction, it is assumed that the additional terms in the governing equations in any case prohibit the attainment of indefinitely strong gradients.

Part II

Model Dynamics versus Model Physics

Chapter 7

Resolved and Unresolved Scales of Motion

The atmosphere is a continuum. The smallest eddies are very much larger than the molecular mean free path, so one can define each fluid “point” as being actually a small volume containing many molecules and having well-defined pressure, temperature, etc. So, in the atmospheric continuum we clearly never have complete (every-point, every-instant) knowledge of its state. Therefore, in any description of the atmosphere (whether it be concerned with global meteorology, the meso-scale, or micro-meteorology) it is helpful to introduce the notion of “resolved” and “unresolved” scales of motion. We split any variable, say the vertical velocity w , into resolved and unresolved parts, thus:

$$w = \bar{w} + w' \tag{7.1}$$

where $w(x, y, z, t)$ is the total instantaneous vertical velocity, $\bar{w}(x, y, z, t)$ is the resolved motion¹, and w' is the deviation from the resolved motion, the “unresolved motion” or “fluctuation.” We then declare that we are concerned to predict/describe the “resolved fields,” $\bar{u}, \bar{v}, \bar{w}, \bar{p}, \bar{T}, \bar{\rho}_v$. But it is an inescapable fact that the unresolved field usually has a non-negligible influence on the evolution of the resolved field.

Now what might this “resolved” field \bar{w} be? In what sense is \bar{w} a “filtered” field? As a first example, let us suppose we are concerned with the “synoptic scale,” ie. working with dependent variables which have been averaged in the horizontal plane (or on a pressure surface) over distances (X, Y) large enough to average out microscale and mesoscale variations: eg.

$$\bar{w} = \frac{1}{XY} \int_{-X/2}^{X/2} \int_{-Y/2}^{Y/2} w(x, y, z) dx dy \quad (7.2)$$

Note that we could and probably should have also averaged in the vertical, since we cannot resolve all detail along any space dimension.

This averaging operation smooths out the sharp updrafts and downdrafts about a range of hills, or the updrafts about a sharp lake-land or ocean-land boundary, or the intense updrafts within individual cumulus clouds: and we are left with a spatially-averaged vertical velocity \bar{w} which is small. Note that \bar{w} remains a function of x, y , but it is a much smoother function of x, y than w . Detail (unresolved structure) has been “filtered” out.

Now, does the existence of an *unresolved* field w' have any effect on the

¹Common alternative notations for the resolved field are the upper case, W and the angle-bracket $\langle w \rangle$.

evolution of \bar{w} ? Yes. For example, a population of cumulus, though each too small to be a resolved feature (cloud width $L \ll X, Y$), collectively causes very efficient vertical heat and vapour transport, that results in a vertical heat flux by the “unresolved flow” that may even exceed the resolved heat flux $\bar{Q}_H = \rho c_p \bar{w} \bar{T}$.

We have established that in meteorology/climatology and in oceanography, no matter what the scale under consideration, one should always commence by settling the issue of scales to be resolved. That done, we need to find “evolution equations” (prognostic or governing equations) for the resolved field \bar{u} etc., ie. we want differential equations for $\partial\bar{u}/\partial t$, etc. And we must show, formally and quantitatively, the way in which the unresolved field affects the evolution of the resolved field. This is straightforward. We simply average the governing equations (a procedure pioneered by O. Reynolds).

7.1 Averaging the Governing Equations

Let f and g be arbitrary flow variables. Then:

$$\begin{aligned} f &= \bar{f} + f' \\ g &= \bar{g} + g' \end{aligned} \tag{7.3}$$

We would like the averaging process to satisfy the following four conditions:

$$\begin{aligned}
 \overline{f + g} &= \bar{f} + \bar{g} \\
 \overline{\alpha f} &= \alpha \bar{f} \\
 \overline{\frac{\partial f}{\partial s}} &= \frac{\partial \bar{f}}{\partial s} \\
 \overline{\bar{f} g} &= \bar{f} \bar{g}
 \end{aligned} \tag{7.4}$$

where α is (any) constant and $s = x_1, x_2, x_3$ or t .

If these four requirements are met, then in addition the following properties hold true (prove for yourselves):

$$\begin{aligned}
 \overline{\bar{f}} &= \bar{f} \\
 \overline{\bar{f}'} &= 0 \\
 \overline{\bar{f} \bar{g}} &= \bar{f} \bar{g} \\
 \overline{\bar{f} g'} &= 0
 \end{aligned} \tag{7.5}$$

We will average the equations as if our averaging operation was ideal, but in reality neither a space- nor a time-average exactly satisfies rules 3, 4. Anthes (1977) is more careful about this.

7.1.1 Example: Averaging the Vapour Conservation Equation

Conservation of water vapour is expressed by:

$$\frac{\partial \rho_v}{\partial t} = -\nabla \cdot \vec{F}_v + Q_v \tag{7.6}$$

where \vec{F}_v is the (total) vector flux density of water vapour.

Noting that averaging commutes with differentiation, we have

$$\overline{\frac{\partial \rho_v}{\partial t}} = \frac{\partial \bar{\rho}_v}{\partial t} \quad (7.7)$$

and

$$\overline{\nabla \bullet \vec{F}_v} = \nabla \bullet \overline{\vec{F}_v} \quad (7.8)$$

So the average form of the vapour conservation equation is:

$$\frac{\partial \bar{\rho}_v}{\partial t} = -\nabla \bullet \overline{\vec{F}_v} + \bar{Q}_v \quad (7.9)$$

We can safely neglect molecular diffusion of water vapour, so the mean flux density $\overline{\vec{F}_v}$ reduces to (using Cartesian coordinates):

$$\overline{\vec{F}_v} = (\overline{u \rho_v}, \overline{v \rho_v}, \overline{w \rho_v}) \quad (7.10)$$

Now substitute $u = \bar{u} + u'$ (etc.) and multiply to get

$$\begin{aligned} \overline{F_{vx}} &= \overline{u \rho_v} \\ &= \overline{(\bar{u} + u') (\bar{\rho}_v + \rho'_v)} \\ &= \bar{u} \bar{\rho}_v + \overline{u' \rho'_v} \end{aligned} \quad (7.11)$$

and so on.

We will temporarily neglect the source term Q_v (though it can easily be put back in, and needs to be accounted for, evaporation/condensation

with respect to any liquid water present). Then, the averaged form of the equation, ie. the evolution equation for $\bar{\rho}_v$ is:

$$\begin{aligned} \frac{\partial \bar{\rho}_v}{\partial t} = & - \frac{\partial}{\partial x} \bar{u} \bar{\rho}_v - \frac{\partial}{\partial y} \bar{v} \bar{\rho}_v - \frac{\partial}{\partial z} \bar{w} \bar{\rho}_v \\ & - \frac{\partial}{\partial x} \overline{u' \rho'_v} - \frac{\partial}{\partial y} \overline{v' \rho'_v} - \frac{\partial}{\partial z} \overline{w' \rho'_v} \end{aligned} \quad (7.12)$$

The first three terms on the rhs involve the resolved variables (and their gradients) and are simply “resolved advection,” or, advection by the resolved flow, expressed in transport form. It is the second three terms that are novel. They are new unknowns, and represent the divergence of fluxes of water vapour carried by the unresolved flow. That is to say, $\overline{w' \rho'_v}$ is a convective vertical flux density of water vapour, carried by the unresolved flow, and so on. And if $\overline{w' \rho'_v}$ changes with height, there results an influence on the resolved humidity field $\bar{\rho}_v$. Whereas the governing equation for ρ_v is “closed” if we regard the velocity field and Q_v as known (1 equation in 1 unknown, ρ_v), the equation we have derived for the evolution of the resolved humidity is unclosed, for even if we know $\bar{u}, \bar{v}, \bar{w}, \overline{Q_v}$, there appear in our equation in addition to $\bar{\rho}_v$ the unknown unresolved fluxes. This proliferation of unknowns is called “the closure problem.”

So, the form of the influence of unresolved flow upon the resolved flow is that there arise fluxes carried by the unresolved flow, and the divergences of those fluxes appear as a “forcing” in the evolution equation for the resolved flow. It is quite common in numerical weather prediction to make an artificial

separation into “dynamics” and “physics” in the following manner:

$$\frac{\partial \bar{\rho}_v}{\partial t} = \left(\frac{\partial \bar{\rho}_v}{\partial t} \right)_{dyn} + \left(\frac{\partial \bar{\rho}_v}{\partial t} \right)_{phys} \quad (7.13)$$

where

$$\begin{aligned} \left(\frac{\partial \bar{\rho}_v}{\partial t} \right)_{dyn} &= -\frac{\partial}{\partial x} \bar{u} \bar{\rho}_v - \frac{\partial}{\partial y} \bar{v} \bar{\rho}_v - \frac{\partial}{\partial z} \bar{w} \bar{\rho}_v \\ \left(\frac{\partial \bar{\rho}_v}{\partial t} \right)_{phys} &= -\frac{\partial}{\partial x} \overline{u' \rho'_v} - \frac{\partial}{\partial y} \overline{v' \rho'_v} - \frac{\partial}{\partial z} \overline{w' \rho'_v} + Q_v \end{aligned} \quad (7.14)$$

By the same method one may show that there arise unresolved horizontal and vertical heat fluxes $\rho c_p (\overline{u' T'}, \overline{v' T'}, \overline{w' T'})$, whose divergences appear in the evolution equation for mean temperature \bar{T} , and so the tendency in \bar{T} is likewise split into dynamics and physics, with

$$\left(\frac{\partial \bar{T}}{\partial t} \right)_{phys} = -\frac{\partial}{\partial x} \overline{u' T'} - \frac{\partial}{\partial y} \overline{v' T'} - \frac{\partial}{\partial z} \overline{w' T'} + Q_T \quad (7.15)$$

where the source term Q_T includes any latent heat addition/removal and the divergence of the radiative heat flux

$$Q_T = -L Q_v - \nabla \bullet \vec{R}^* \quad (7.16)$$

where L is the latent heat of vapourisation/sublimation. Note that if Q_v is positive, vapour is being created by evaporation, which consumes thermodynamic energy and makes a negative contribution to Q_T .

Completing this, the “physics” term in the \bar{u} -momentum equation, often simply called “friction,” is:

$$\left(\frac{\partial \bar{u}}{\partial t} \right)_{phys} = -\frac{\partial \overline{u' u'}}{\partial x} - \frac{\partial \overline{v' u'}}{\partial y} - \frac{\partial \overline{w' u'}}{\partial z} \quad (7.17)$$

where $\tau_{xz}/\rho = \overline{u'w'}$ is a momentum flux (per unit mass) carried by the unresolved flow, sometimes called a “Reynolds stress.”

We have now covered the decomposition into resolved and unresolved scales, which should be the starting point of virtually any scientific enterprise in meteorology or oceanography. We have seen that the fluxes carried by the unresolved motion feed back on the resolved fields of motion and temperature and humidity, and that the evolution equation for the resolved field is unclosed, until we specify (by some means) the magnitudes of the unresolved fluxes (the closure problem).

Chapter 8

Modelling the “free” atmosphere (“dynamics”)

In Numerical Weather Prediction one speaks of Primitive Equation models, Filtered models, Balanced models, Quasi-Geostrophic models... What does all this mean? Because most students will have studied dynamical meteorology, I shall assume it is unnecessary to go over this topic in detail, particularly since the numerical solution of a modern dynamical weather model is a task beyond our resources (though not our capacities) in the context of our course. Good surveys are given by Holton (2004), and by Haltiner and Williams (1980); but we must note that these writers only passingly allude to the whole conundrum of the influence of unresolved scales, occasionally sticking in a term called ‘F’ (for friction), then dropping it again as hastily as possible. The notion of unresolved topography, likewise, doesn’t get much of a hearing (Wilson, 2002, covers these complications).

Overview

For more details, please turn to the Appendix.

Drop unresolved scales. Having averaged the ‘raw’ equations that express conservation of momentum, mass and energy, we now extract from the total tendencies in the filtered (resolved) field the “dynamic” tendencies. If we drop the explicit recognition of the fact that the resolved variables are averages (\bar{u} , $\bar{p} \rightarrow u$, p , etc.) then the resulting equations *look* like the raw equations - but we know they have a different meaning, since part of the tendency is not included. Anyway, if we do all that, we have the equations that appear in textbooks (eg. Holton) on global meteorology.

Limit the scale of application. The resulting equations of motion contain terms which are not, under normally-expected conditions, all of the same order of magnitude. This is readily ascertained by performing a “scale analysis,” in which one assigns to each of the variables a scale typical of the values expected under the expected conditions. Thus, a scale analysis directed towards the synoptic scale, might assume vertical velocity w to have typical scale $0.01 - 0.1 \text{ m s}^{-1}$ while horizontal components have a scale of about 10 m s^{-1} : but a cloud model or a model of flow about a building or windbreak would probably assign equal scales to horizontal and vertical velocity. This is essentially because the large-scale motion is quasi-horizontal, confined to be so by the underlying surface.

The key results of a scale analysis for synoptic scale meteorology and

numerical weather prediction (Holton, 2004, p35-38; Haltiner & Williams, 1980, pp62-66) are that on the synoptic scale and at mid-latitudes, where the Rossby number

$$R_o = \frac{V}{fL} \quad (8.1)$$

(ratio of typical acceleration to the Coriolis acceleration, where V is a scale for the horizontal velocity and L is the length scale for synoptic scale motion) is small, the horizontal accelerations ($du/dt, dv/dt$) are small terms resulting from small differences between the large pressure gradient and Coriolis terms, while the vertical acceleration (dw/dt) is a small term resulting from small imbalances in large pressure gradient and gravity terms: we say that on the synoptic scale, there is approximate hydrostatic and Geostrophic balance.

Eliminate the possibility of spurious types of motion. The inaccuracy of Richardson’s pioneering numerical forecast is believed to have been partially caused by a growth of spurious (non-meteorological) waves, triggered by errors in his initial data, which obscured the “real” developments. “Filtered” models “are characterised by the absence of sound and gravity waves and a balance between the wind and pressure fields” (Haltiner & Williams, 1980, p208). They are not necessarily based on divergence and vorticity equations, though usually so in practise, since that approach most easily allows filtering out of gravity waves.

Sound Waves

Sound waves are longitudinal waves (particle oscillation is parallel to the direction of propagation), and horizontally propagating sound waves are called “Lamb waves.” There are no meteorological phenomena in which sound waves play a significant role.

If p and ρ are determined by observation and used to initialise a full vertical momentum equation very large spurious accelerations can result. Any convergence in the convective mass flux density $w\rho$ which arises from these spurious velocities can drive erroneous density changes (mass readjustment), as is indicated by the continuity equation

$$\frac{\partial \rho}{\partial t} = - \frac{\partial}{\partial z} \rho w + \dots \quad (8.2)$$

Since we can consider the accelerations to be driven by the pressure force, without the moderation of friction, in effect we have vertically-propagating sound waves.

Replacement of the vertical momentum equation with the hydrostatic equation is sufficient to filter out vertically propagating sound waves, because in that case p is (apparently, ie. in so far as the model is concerned) determined only by the total overlying mass of air, and cannot be perturbed by local mass flux convergences. Lamb waves carry very little energy in the atmosphere, but are important in NWP - they place a restriction on the maximum timestep for numerical integration of the hydrostatic equations (Haltiner & Williams, 1980, p35). They have maximum amplitude at

ground (Holton p176) and may be filtered out by imposing

$$\omega_0 = \left(\frac{dp}{dt} \right)_{gnd} = 0 \quad (8.3)$$

Gravity Waves

Gravity waves are lateral waves. Internal gravity waves (buoyancy waves) play a significant meteorological role at small scales, eg. lee waves (orographic waves). But what “is” a gravity wave?

Let $\theta_e(z)$ be the mean potential temperature profile of the atmosphere (subscript ‘e’ for environment), and let the instantaneous potential temperature of a fluid parcel be $\theta = \theta_e + \theta'$ where θ' is the local and instantaneous deviation of the potential temperature of a fluid parcel from the local mean. The coupled equations

$$\begin{aligned} \frac{dw}{dt} &= g \frac{\theta'}{\theta_e} \\ \frac{d\theta'}{dt} &= -w \frac{\partial \theta_e}{\partial z} \end{aligned} \quad (8.4)$$

(which are in Lagrangian form: d/dt is the derivative following the parcel) describe a system in which

- a moving parcel experiences a vertical acceleration due to the buoyancy force (ie., the momentum balance contains inertia and buoyancy only; most of the gravity has been “cancelled out” by the hydrostatic pressure gradient, leaving only the modified (or “reduced”) gravity $g\theta'/\theta_e$, and

- as the parcel moves it expands or contracts adiabatically, and thus its potential temperature is constant. However relative to its environment its fluctuation temperature θ' is rising or falling at a rate that depends on its ascent rate w .

If we differentiate the first equation to obtain d^2w/dt^2 and close using the second equation for $d\theta'/dt$, we have a wave equation

$$\frac{d^2w}{dt^2} = -w \frac{g}{\theta_e} \frac{\partial \theta_e}{\partial z} = -\frac{w}{N^2} \quad (8.5)$$

where N is the buoyancy frequency (commonly called the Brunt-Vaisala frequency). Since the meteorological equations contain the terms from which we derived this result, it should be no surprise that gravity waves occur in the atmosphere.

If we think of an imaginary plane surface distorting up and down in response to internal gravity waves like a (surface gravity) wave on water, it is intuitive that propagation involves:

$$\frac{\partial}{\partial t} (\nabla \cdot \vec{V}) \neq 0 \quad (8.6)$$

where \vec{V} is the horizontal velocity vector - we need a time-varying confluence/diffluence of the horizontal wind (but not necessarily any convergence/divergence $\nabla \cdot \vec{U}$ of the full vector velocity). Setting

$$\frac{\partial}{\partial t} (\nabla \cdot \vec{V}) = 0 \quad (8.7)$$

is the minimum simplification required to filter out gravity waves (Holton p177). This term does not appear explicitly in the momentum equations,

but it does appear in the divergence equation which we encounter in using a vorticity-streamfunction formulation.

Introduce a convenient vertical coordinate. Under the assumption (or approximation) that the atmospheric pressure distribution is hydrostatic, the momentum equations reduce to the pair

$$\begin{aligned}\frac{d\vec{V}_H}{dt} &= -\frac{1}{\rho} \nabla_H p - 2\vec{\Omega} \times \vec{u} \\ 0 &= -\frac{1}{\rho} \frac{\partial p}{\partial z} - g\end{aligned}\tag{8.8}$$

where \vec{V}_H is the horizontal wind vector, and ∇_H is the horizontal gradient operator ($= (\partial/\partial x, \partial/\partial y)$ in Cartesian coordinates). The hydrostatic equation gives us a 1:1 map between z and p , and in classical dynamics it is usual to use p as our vertical coordinate. And in fact it is even more convenient to use $\sigma = p/p_0(x, y)$, where p_0 is surface pressure, for now the lower boundary of the atmosphere is the flat surface $\sigma = 1$.

Equations in the pressure coordinate. Holton (p176) gives a set of prediction equations for the resolved flow (physics terms not included) under the minimum simplification necessary to filter out sound waves:

$$\begin{aligned}\left(\frac{\partial}{\partial t} + \vec{V} \cdot \nabla_p\right) \vec{V} + \omega \frac{\partial \vec{V}}{\partial p} + f \hat{k} \times \vec{V} &= -\nabla_p \Phi \\ \nabla_p \cdot \vec{V} + \frac{\partial \omega}{\partial p} &= 0 \\ \left(\frac{\partial}{\partial t} + \vec{V} \cdot \nabla_p\right) \frac{\partial \Phi}{\partial p} + \sigma \omega &= 0\end{aligned}\tag{8.9}$$

where \vec{V} is the horizontal velocity¹; Φ is the “geopotential”

$$\Phi = \int_0^z g dz \quad (8.10)$$

$\omega = dp/dt$ is the vertical velocity in pressure (isobaric) coordinates; the grad operator ∇_p is the horizontal gradient operator (applied with p held constant: see Holton p54); and σ is a static stability parameter

$$\sigma = -\frac{1}{\rho} \frac{\partial \ln \theta}{\partial p} \quad (8.11)$$

where θ is the potential temperature.

From this set of equations, which will sustain gravity waves, one may derive by differentiation the vorticity equation²

$$\begin{aligned} \frac{\partial \zeta}{\partial t} + \vec{V} \cdot \nabla (\zeta + f) + \omega \frac{\partial \zeta}{\partial p} \\ = -(\zeta + f) \nabla \cdot \vec{V} + \tilde{k} \cdot \left(\frac{\partial \vec{V}}{\partial p} \times \nabla \omega \right) \end{aligned} \quad (8.12)$$

and the divergence equation

$$\begin{aligned} \frac{\partial}{\partial t} (\nabla \cdot \vec{V}) = & -\nabla^2 \left(\Phi + \frac{\vec{V} \cdot \vec{V}}{2} \right) \\ & - \nabla \cdot (\tilde{k} \times \vec{V} (\zeta + f)) \\ & - \omega \frac{\partial}{\partial p} (\nabla \cdot \vec{V}) - \frac{\partial \vec{V}}{\partial p} \cdot \nabla \omega \end{aligned} \quad (8.13)$$

where

$$\zeta = \tilde{k} \cdot \nabla \times \vec{V} \quad (8.14)$$

¹Note that the Coriolis term may be written $f \hat{k} \times \vec{V} = fu \tilde{j} - fv \tilde{i}$, where \tilde{i} is the unit vector pointing along the east-west direction.

²I shall drop the subscript ‘p’ on ∇ , but recall $\nabla \equiv \left[\left(\frac{\partial}{\partial x} \right)_{y,p,t}, \left(\frac{\partial}{\partial y} \right)_{x,p,t} \right]$.

is the vertical component of the relative vorticity. These equations can be used to replace the two horizontal momentum equations. One advantage of this step is that if the LHS of the divergence equation is set to zero, time-dependent gravity waves are eliminated.

By the Helmholtz theorem, one may split the horizontal velocity into the sum of a nondivergent and irrotational part:

$$\vec{V} = \vec{V}_\psi + \vec{V}_e \quad (8.15)$$

where the nondivergent part satisfies (by definition)

$$\nabla \bullet \vec{V}_\psi = 0 \quad (8.16)$$

and may be expressed in terms of the gradient of a streamfunction ψ :

$$\vec{V}_\psi = \tilde{k} \times \nabla \psi = \left(-\frac{\partial \psi}{\partial y}, \frac{\partial \psi}{\partial x} \right) \quad (8.17)$$

With this decomposition, and after a scale analysis for midlatitude synoptic scale motion - which shows, among other things, that

$$|\vec{V}_\psi| \gg |\vec{V}_e| \quad (8.18)$$

- one can obtain from the divergence equation the non-linear balance equation:

$$\nabla^2 \left(\Phi + \frac{1}{2} (\nabla \psi \bullet \nabla \psi) \right) = \nabla \bullet ((f + \nabla^2 \psi) \nabla \psi) \quad (8.19)$$

which relates the non-divergent part (ie. essentially the horizontal wind) of the wind field via the streamfunction to the “mass” field (the Geopotential:

since we are using the hydrostatic equation, the height of a pressure surface depends only on the density field). The linear balance equation is a further simplification:

$$\nabla^2 \Phi = \nabla \bullet (f \nabla \psi) \quad (8.20)$$

Quasi-geostrophic system of equations

Key attributes: Valid over a restricted range of latitude, with horizontal wind \vec{V} assumed geostrophic (with $f = f_0$) except where its divergence $D = \nabla_p \bullet \vec{V}$ appears in the vorticity equation.

By performing the scale analysis for mid-latitude synoptic scale motion on the governing equation for the vertical component of the vorticity expressed in isobaric coordinates, using the β -plane approximation $f = f_0 + \beta y$, Holton (p180) obtains for synoptic scale motion:

$$\frac{\partial}{\partial t} (\nabla^2 \psi) = -\vec{V}_\psi \bullet \nabla (\nabla^2 \psi + f) + f_0 \frac{\partial \omega}{\partial p} \quad (8.21)$$

The streamfunction is now related to the geopotential by

$$\psi = \frac{\Phi}{f_0} \quad (8.22)$$

so the horizontal velocity is now simply the geostrophic wind:

$$\vec{V}_\psi = \frac{\tilde{k} \times \nabla \Phi}{f_0} \quad (8.23)$$

and the streamfunction is related to the geostrophic (relative) vorticity by the Poisson equation:

$$\nabla^2 \psi = \zeta \quad (8.24)$$

Barotropic vorticity equation

Even by 1912 it had been established (or suggested) by Dines that confluence $\nabla_p \bullet \vec{V} < 0$ in the lower troposphere (ie. the cross-isobar flow into a surface storm) is largely compensated by diffluence in the upper troposphere (and vice versa), such that

$$|\nabla \bullet \vec{V}_p| \gg \frac{1}{p_0} \int_{p=0}^{p_0} \nabla \bullet \vec{V} dp \quad (8.25)$$

This implies that at some midtropospheric altitude $\nabla_p \bullet \vec{V}$ changes sign, ie. passes through zero.

Assuming, then, that there is a *level of non-divergence*, say at 500 mb, we have at that level $\partial\omega/\partial p = 0$, and the Quasi-Geostrophic Vorticity equation reduces to:

$$\frac{\partial}{\partial t} (\zeta) = -\vec{V}_\psi \bullet \nabla (\zeta + f) \quad (8.26)$$

which may be written in shorthand as:

$$\frac{d}{dt} (\zeta + f) = 0 \quad (8.27)$$

(time derivative of the absolute vorticity, following the motion on an isobaric surface, must vanish).

According to Holton, this is “a useful approximate forecast equation.” It gives the evolution of the flow at a single level, the level of non-divergence.

In some years, one of the computational assignments of this course has been to compute the vorticity field over western Canada and the eastern

*CHAPTER 8. MODELLING THE “FREE” ATMOSPHERE (“DYNAMICS”)*86

Pacific from a 500 mb chart, then perform a forecast by solving this equation (the point being, numerical weather forecasting, albeit at a long since surpassed level, is well within our capabilities with a PC).

Chapter 9

Modelling the Atmospheric Boundary-Layer (“physics”)

We have seen that it is useful (and proper) to explicitly define, and find evolution equations for, the resolved state of the atmosphere $\bar{u}, \bar{v}, \bar{w}, \bar{p}, \bar{T}, \bar{\rho}_v, \dots$. In Numerical Weather Prediction, the average can be considered a spatial average over a domain XY of order 100 km x 100 km. We saw that the evolution equations $\partial \bar{\rho}_v / \partial t$, etc., contain the divergences of convective fluxes (of heat and vapour and momentum) “carried by” the unresolved motion. Concentrating on the vertical fluxes (which are in any case those we have best prospect of quantifying), the key unresolved fluxes are:

- heat: $\rho c_p \overline{w'T'}$
- vapour: $\overline{w'\rho'_v}$
- momentum: $\overline{w'u'}, \overline{w'v'}, \overline{w'w'}$

For example, Cumulus convection over a synoptic scale area, causes large vertical heat and vapour fluxes at unresolved scales. The effects of such

processes (eg. the drying and warming of the atmosphere over a region where precipitation is reaching the ground), then, will need to be “parameterised” and is an aspect of the “model physics” or “grid point computations”.

A weather model has, obviously, boundaries in the vertical. Across the lower boundary, there are vertical fluxes into the model domain, and the heat and vapour thus introduced is carried up and redistributed within the atmosphere largely by motions that are unresolved in an NWP model, whose spatial averaging masks the eddies seen near ground. More specifically, the “free atmosphere,” the region where the “dynamics” terms in the equations for the resolved variables are considered dominant, is coupled to the ground through the turbulent Atmospheric Boundary Layer (ABL) (or Planetary Boundary Layer (PBL), which feeds up heat and moisture, as well as transferring drag. These influences (friction, surface heating and evaporation) have an importance to a synoptic scale forecast that increases with the range of the forecast. It can readily be shown, for example, that over 24 hours surface evapotranspiration (E , $\text{kg m}^{-2} \text{s}^{-1}$) can easily increase the moisture content of the PBL by 25% - and this could be an important moisture supply for (eg.) development of a tropical cyclone. Not only that: the surface heat flux (over very wet surfaces the virtual heat flux) determines the PBL stability, and thus convection. So surface fluxes are important w.r.t. changes in storage of heat and moisture (significant especially for longer-term forecasts), and w.r.t. stability (significant both in the short and longer range forecasts). It is believed therefore that even short range (12 hr) forecasts will be im-

proved as surface parameterisation improves. When the lowest grid level of a weather model lies within the atmospheric surface layer (ASL), a shallow layer at the base of the PBL having a depth of order 50 m, the surface fluxes can be computed using the “Monin-Obukhov similarity theory” of the ASL. However in many models (especially older ones) the lowest level is far above the surface layer, and to calculate the boundary fluxes the entire PBL must be parameterised.

“Model physics” is the categorical name given to algorithms of a numerical weather model that calculate the above-mentioned and other mechanisms/processes.

9.1 Role of the ABL in NWP

Near ground, essentially all the vertical transport is performed by the unresolved field: there are strong vertical gradients in temperature, etc., and turbulent eddies transport heat (etc.) down these gradients. An NWP model needs to estimate these vertical fluxes and their divergence, or get the part of the tendency in \bar{T} etc. due to ground forcing wrong. I don’t want to get into this in detail; the turbulent ABL is the subject of the graduate course EAS 572.

We would like our NWP model to correctly calculate the redistribution of heat and vapour by the unresolved eddies (the turbulence) near ground. The turbulent layer has a diurnally, seasonally, and geographically varying depth $h(t)$. Of course, any model which does not contain grid points at heights be-

low $h(t)$ cannot in any way model the evolution of $h(t)$ and the redistribution of heat and moisture by the turbulence. So it is highly desirable to provide sufficient levels near ground to at least resolve the vertical structure of the PBL in a way that is not incompatible with our very coarse knowledge on the horizontal scale. It is not obvious that anything short of a large number (say 50) of vertical grid points will suffice to model the PBL, in which the flow complexity is such that simple K-theory closure (see below) is not really adequate (there can exist, for example, counter gradient fluxes near the outer edge of the convective PBL or “CBL” - ie. mean heat flux is from heights where it is in the mean cooler to where it is in the mean warmer!). Therefore an entire-PBL model will be complex and very time-consuming to compute, and there are competing demands for computer power - eg., to reduce the grid lateral length. In any case (Haltiner & Williams, 1980) for large scale NWP models, the exchange of heat and moisture between the PBL and the free atmosphere above may be more important than details within the PBL. Vertical transfer above the PBL is largely accomplished by cumulus convection and gravity waves (another form of unresolved field), and by the large scale (resolved) vertical velocities.

9.2 Estimating the unresolved fluxes using K-theory

Although there are synoptic scale horizontal gradients in our resolved meteorological fields \bar{u} , \bar{T} , etc., we might postulate some sort of “locally-uniformity”

within the boundary-layer, such that the resolved vertical structure implies some “equilibrium” pattern to the vertical unresolved fluxes, and the horizontal unresolved fluxes are negligible. This permits a 1-dimensional (vertical variation only) treatment of the ABL.

Suppose, then, we concentrate on estimating only the vertical unresolved fluxes,

$$\left(\frac{\partial \bar{T}}{\partial t}\right)_{phys} = -\frac{\partial \overline{w'T'}}{\partial z} \quad (9.1)$$

(etc). The vertical eddy heat flux $Q_H = \rho c_p \overline{w'T'}$ is unknown. If we attempt to derive a rigorous equation for it from the Navier-Stokes equations and the thermodynamic equation, the outcome is

$$\begin{aligned} \frac{\partial \overline{w'T'}}{\partial t} = 0 = & -\frac{\overline{w'^2}}{w'^2} \frac{\partial \bar{T}}{\partial z} - \frac{\partial \overline{w'w'T'}}{\partial z} - \frac{1}{\rho_0} \frac{\overline{\partial p'}}{\partial z} + \frac{g}{T_0} \frac{\overline{w'^2}}{w'^2} \\ & + \nu \overline{T' \nabla^2 w'} + \kappa \overline{w' \nabla^2 T'} \end{aligned} \quad (9.2)$$

The terms ρ_0, T_0 are reference density and temperature, and may be considered to be mean values across the ABL. It is important to note that all terms involving horizontal heterogeneity have been neglected, ie. eqn (9.2) assumes $\partial \overline{w'T'}/\partial x = \partial \overline{w'T'}/\partial y = 0$; the proper interpretation here is that we are assuming the time tendency in statistics of the unresolved flow, and all terms involving horizontal derivatives, are small compared to other terms.

This is a complicated equation which has only deepened our problem, for now we have further unknown terms (the “closure problem”). However the equation at least offers us some insight, and can certainly be used diagnostically in the context of measurements. The first term on the r.h.s. is a

production term: if there is unresolved motion then (by definition) $\overline{w'^2} \neq 0$, and if there is also a temperature gradient, then the motion acts to create a (mean) unresolved flow of heat - the sign of the flux depending on the sign of the temperature gradient. However the fourth term on the r.h.s. shows that the mere existence of temperature variance, in the presence of a gravity field, *always* acts to result in upward transport of heat (temperature variance vanishes only in a neutral atmosphere).

There is no rigorous way around the closure problem. Ultimately an assumption must be made. Here I will describe only the most basic closure assumption, namely the assumption of a linear flux-gradient relationship (most commonly called K-theory or first-order closure or gradient-diffusion closure). The model is certainly wrong in general, but one must go to much trouble to improve upon it.

In K-theory it is assumed that the mean gradient in a given “driving force” (such as temperature or humidity) drives a corresponding turbulent (more generally, unresolved) flux (of heat or moisture). The K-theory closure for the vertical sensible (as distinct from latent; defined below) heat flux density is

$$Q_H = -\rho c_p K_H(z) \left(\frac{\partial \bar{T}}{\partial z} - \gamma_d \right) \quad (9.3)$$

where $K_H(z)$, $\text{m}^2 \text{s}^{-1}$ is the “eddy diffusivity for heat” and the subtraction of the adiabatic lapse rate $\gamma_d = -g/c_p$ ensures zero heat flux in the neutral

(adiabatically mixed) state of stratification. Similarly the moisture flux is

$$E(z) = \overline{w'\rho'_v} = -K_v \frac{\partial \bar{\rho}_v}{\partial z} \quad (9.4)$$

where K_v , $\text{m}^2 \text{s}^{-1}$ is the eddy diffusivity for water vapour.

Note that these flux models are simply an analogy to molecular diffusion (conduction), but with vastly greater diffusivities which are properties of the turbulent flow rather than of the fluid. Experiments have approximately determined the values of the K 's very near ground over the range in atmospheric stratification; those values are themselves affected by stratification, windspeed, and the depth of the PBL.

9.2.1 Monin-Obukhov Similarity Theory of the ASL

The Atmospheric Surface Layer (ASL) is the lowest 50 m (or so) of the ABL, where proximity to ground limits vertical motion and mixing, and so vertical gradients in mean windspeed, temperature, humidity (etc.) are very large. This is an important layer of the atmosphere - we live in it! Let us call its depth h_{ASL} and accept that the precise definition of this depth is vague.

Monin and Obukhov provided an empirical theory of the ASL in its simplest, horizontally-uniform state, by hypothesizing that that state is determined by the (roughly, height-independent¹) turbulent fluxes of momentum and heat across it. Specifically Monin and Obukhov postulated that the most important “global” (or “external” or “governing”) scales are the fric-

¹In the sense that, say, $|\overline{w'T'}(h_{ASL}) - \overline{w'T'}(0)| \leq 0.1 |\overline{w'T'}(0)|$, etc. Thus we may call the hhASL the “constant flux layer”.

tion velocity $u_* = \sqrt{-\overline{u'w'}}$ and the kinematic heat flux $Q_H/\rho c_p = \overline{w'T'}$ (from the latter a turbulent temperature scale $T_* = -\overline{w'T'}/u_*$ is derived). Monin-Obukhov similarity theory (‘MOST’) specifically pertains to the hh-ASL at heights far enough above the surface that no length scales pertaining to the surface are relevant, thus only height z enters the picture; and since buoyancy effects hinge on existence of gravity, let’s just say that a scale g/T_0 is also needed.

From the available parameters, one may form a new length scale, the Monin-Obukhov length

$$L_{MO} = \frac{-u_*^3}{k_v \frac{g}{T_0} \frac{Q_H(0)}{\rho_0 c_p}} \quad (9.5)$$

where the von-Karman constant $k_v = 0.4$ is included purely by convention. If the ASL is neutral, by definition $|L_{MO}| = \infty$, while if the ASL is unstable then $L < 0$ (and vice versa). If $|L| \lesssim 2$ m, we have such light winds and/or strong heat fluxes that the state of the ASL is ill-defined (and difficult to even measure). Otherwise, MOST has been found to give a good description. Its specific predictions are (for example)

$$\begin{aligned} \frac{k_v z}{u_*} \frac{\partial \bar{u}}{\partial z} &= \phi_{mtm} \left(\frac{z}{L_{MO}} \right) \\ \frac{k_{vh} z}{T_*} \frac{\partial \bar{T}}{\partial z} &= \phi_h \left(\frac{z}{L_{MO}} \right) \\ \frac{k_{vw} z}{r_*} \frac{\partial \bar{\rho}_v}{\partial z} &= \phi_w \left(\frac{z}{L_{MO}} \right) \\ \frac{k_{vc} z}{c_*} \frac{\partial \bar{c}}{\partial z} &= \phi_c \left(\frac{z}{L_{MO}} \right) \end{aligned} \quad (9.6)$$

where the functions on the r.h.s. are universal empirical “MO” functions² and the k_v ’s are defined such that in each case $\phi(0) = 1$ (the normalizing lengthscales on the l.h.s. could equally have been L_{MO} , but then different functions would result on the rhs). The scales for water vapour and for a generic gas (concentration \bar{c}) have been formed from their respective fluxes, viz. $r_* = -\overline{w'\rho'_v}/u_*$, $c_* = -\overline{w'c'}/u_*$.

The consensus view at present from the available experiments is that $k_{vw} = k_{vh} = k_v = 0.4$, while (as you discovered in Lab 3) the best atmospheric dispersion experiments indicate $k_{vc} > k_v$, ie. $S_c = k_v/k_{vc} < 1$. Taking the stable case, experiments show $\phi_{mtm} \approx \phi_h \approx \phi_v \approx 1 + \beta z/L_{MO}$ with $\beta \approx 5$. The corresponding wind profile is the “log-linear” profile

$$\bar{u} = \frac{u_*}{k_v} \left[\ln \frac{z}{z_0} + \beta \frac{z - z_0}{L_{MO}} \right] \quad (9.7)$$

where z_0 is a constant of integration, defined to be the location on the z -axis where mean windspeed (apparently) vanishes... the “surface roughness length”. While MO theory does not hold down as low as $z \sim z_0$, in practise we apply it unless we want to go to the trouble of exploiting a more detailed theory.

Now the MOST does not explicitly hinge on introducing the flux-gradient closure, but we can readily see that it provides us formulae for the diffusivities. By re-arranging eqns(9.8) it is easy to show that the K ’s for momentum,

²The ratio z/L_{MO} approximately equals the ratio of the rates of buoyant and shear production of turbulent kinetic energy (more on this below).

heat, water vapour and a (general) gaseous additive are

$$K_{mtm} = \frac{k_v u_* z}{\phi_{mtm}}$$

$$K_h = \frac{k_v u_* z}{\phi_h} \quad (9.8)$$

$$K_w = \frac{k_v u_* z}{\phi_w} \quad (9.9)$$

$$K_c = \frac{k_{vc} u_* z}{\phi_c} \quad (9.10)$$

where we shall often simplify to $K_{tmt} = K$ in notation for the eddy viscosity.

9.3 A vintage treatment of the ABL (old Canadian Regional Finite Element Model)

The K 's were specified as multiples of an eddy viscosity formulated as

$$K = a \lambda \sqrt{k} \quad (9.11)$$

where a is a constant (near 1), λ is a height-dependent “mixing length” that relates to the size of the unresolved eddies, and k is the kinetic energy (per unit mass) in the unresolved motion, defined by

$$k = \frac{1}{2} (\overline{u'^2} + \overline{v'^2} + \overline{w'^2}) \quad (9.12)$$

Since \sqrt{k} is a velocity, we can regard it as a measure of the strength of the unresolved (turbulent) velocities, and λ as a measure of the distance along which an eddy persists. The “exchange coefficient” K , measuring the efficiency of transport by the unresolved motions, depends on both these factors.

An exact evolution equation for k can be derived from the Navier-Stokes equations, by steps we will not cover. In a simplified form in which advective terms have been dropped (which is appropriate in the context of “grid point computations”) the k -equation is:

$$\frac{\partial k(z, t)}{\partial t} = P_S + P_B + T_k - \epsilon \quad (9.13)$$

Here $P = P_S + P_B$ is the rate of production of unresolved kinetic energy (KE), and

$$T_k = - \frac{\partial}{\partial z} \overline{w' \frac{1}{2} (u'^2 + v'^2 + w'^2)} = \frac{\partial}{\partial z} \left(K_k \frac{\partial k}{\partial z} \right) \quad (9.14)$$

is the “turbulent transport” term, modelled (in the closed form on the right) as a gradient-diffusion process with diffusivity $K_k \propto K$. The the KE production rates may be written

$$P_S = - \overline{u'w'} \frac{\partial \bar{u}}{\partial z} = K \left(\frac{\partial \bar{u}}{\partial z} \right)^2 \quad (9.15)$$

$$P_B = \frac{g}{T_0} \overline{w'T'} = - \frac{g}{T_0} \frac{K}{P_r} \left(\frac{\partial \bar{T}}{\partial z} - \gamma_d \right) \quad (9.16)$$

where (again) the closed form exploits the gradient-diffusion approximation for the unresolved fluxes (the turbulent Prandtl number $P_r = K/K_h$ is the ratio of the eddy diffusivities for momentum and heat, so $K_h = K/P_r$), and (in addition to assuming horizontal uniformity) it has been assumed the coordinate axes are aligned with the local mean wind, so that $\bar{v} = 0$. Note that the buoyancy part can be a source or a sink for KE in the unresolved scales, according to the sign of the unresolved heat flux. In unstable stratification

the heat flux is positive and buoyant production is a source for KE. The opposite is true at night, explaining why the surface layer may become very quiescent at night, unless shear production is maintained large by a large free-atmosphere wind at the top of the ABL. It is helpful to note that to within a good approximation

$$\frac{P_B}{P_S} = \frac{z}{L_{MO}} \quad (9.17)$$

which gives us a useful interpretation of the MO length (depth of the layer of shear-dominated turbulence). It is also worthwhile to mention a novel interpretation of the shear production and turbulent transport terms in eqn(9.13) due to K.G. Mcnaughton. Let the symbol F_k represent the mean vertical flux of kinetic energy, ie.

$$F_k = \overline{w \frac{1}{2} u_i u_i} \quad (9.18)$$

(summation over i applies, of course). Then assuming horizontal homogeneity, one may easily show by Reynolds decomposition that

$$F_k = \bar{u} \overline{u'w'} + \overline{w' \frac{1}{2} u_i' u_i'} \quad (9.19)$$

and since we know that (in the hhASL) $\overline{u'w'} = -u_*^2$ is negative, this seems prototypically to be a downward (negative) flux of KE. Then the sum $P_S + T_k$ of (what is conventionally called) “shear production” and turbulent transport is seen to be equivalent to the vertical divergence of this downward flux of kinetic energy.

Returning to eqn (9.13), ϵ is called the turbulent kinetic energy (TKE) dissipation rate, and quantifies the rate of conversion of KE to heat due to

working by viscous forces. It was modelled in the RFE model (in accordance with widespread custom) as

$$\epsilon = \frac{k^{3/2}}{\lambda_e} \quad (9.20)$$

where λ_e is called the equilibrium length scale (discussed below). This formula can be seen to be dimensionally correct, and says that the more energy present, the faster its rate of conversion to heat. The other term in the evolution equation for k is called “turbulent transport,” and actually the diffusion form shown for it is just a model, but a nice model because it results in the k -eqn containing a diffusion operator - and consequently nice numerical behaviour, filtering of short waves, etc.

The mixing lengths were related by

$$\frac{d\lambda}{dt} = \frac{\lambda_e - \lambda}{\tau} \quad (9.21)$$

where τ is some timescale. This simply says that the lengthscale λ is always evolving towards an equilibrium value, λ_e , the latter being a specified continuous function of height (measured relative to ABL depth, ie. z/h) and a discontinuous (categorical) function of stratification (unstable or stable).

I shan’t go into the details of how the RFE modelled the evolution of ABL depth, $h(t)$ the scheme was very simplistic, but note that Delage (1974) used a scheme similar that described above to compute the time evolution of the windspeed, temperature and TKE profiles in a nocturnal boundary-layer, while Wilson et al. (1999) used a similar scheme to compute the spatial variation of mean wind and turbulence across forest belts and cutblocks in

the Boreal forest.

Now we consider the treatment of boundary conditions at $z = 0$.

9.4 Boundary treatment for a model with lowest level in the ASL

“Interactive feedback from the biosphere is probably as important as feedbacks from deep cumulus and from the ocean. The level of complication needed to model this mechanism is, however, a decision that atmospheric modellers must make.” - Hua-Lu Pan, NMC Washington (Monthly Weather Review, Vol. 118, 1990).

Suppose a (typical) modern weather model has its lowest grid level at z_a , nominally the “anemometer height,” which lies with the atmospheric surface layer, *across which the vertical fluxes have negligible change with height*. Let u_a , T_a , q_a be the mean windspeed, temperature, and specific humidity at the anemometer height (I have not bothered to use a notation to emphasize these explicitly as spatial averages). Then the flux densities of momentum, sensible, and latent heat ($Q_E = LE$ is the latent heat flux density, L the latent heat of evaporation or sublimation) are modelled by

$$\begin{aligned}\tau &= \rho C_M^2 u_a^2 \\ Q_H &= \rho c_p C_M C_H u_a (T_0 - T_a) \\ Q_E &= \rho L C_M C_V u_a (q_0 - q_a)\end{aligned}\tag{9.22}$$

where T_0 , q_0 are the temperature and specific humidity at a nominal plane,

the surface³. These equations specify the fluxes “at” ground level, and so can be considered the boundary condition. Within the ABL, these fluxes are carried by the unresolved scales. Any convergence of the fluxes, of course, causes evolution of the corresponding resolved field.

In eqns (9.22) C_M, C_H, C_V are dimensionless transfer coefficients for momentum, heat, and vapour, and are defined by the above equations. Over uniform terrain, they are easily determined from Monin-Obukhov similarity theory. The above formulation is implemented in such a manner as to obey the constraint of energy conservation at the interfacial plane:

$$Q^* = Q_H + Q_E + Q_G \quad (9.23)$$

where Q^* is the net radiation and Q_G the soil heat flux density.

T_0 and q_0 are usually advanced in time from their initial values by some variation of the “force-restore” method (see Deardorff, 1978; J. Geophys. Res. 83). A simple variant of this, the forcing method, models the evolution of T_0 for a “ground” of density ρ_s and specific heat capacity c_s as if T_0 is the average temperature of a slab of soil of depth d_T chosen so that there is zero vertical flux at the base of the slab,

$$\rho_s c_s \frac{\partial T_0}{\partial t} = \frac{Q_G}{d_T} = \frac{Q^* - Q_H - Q_E}{d_T} \quad (9.24)$$

In the so-called “bucket model” of surface hydrology, the specific humidity

³This is an appropriate place to emphasize the connectivity of the material universe. Surface temperature and humidity are determined by a balance that connects the atmosphere to soil/plant/water. Sophisticated lower boundary conditions will resolve a plant layer and a soil layer.

difference “driving” evaporation is modelled as:

$$q_0 - q_a = \beta [q_*(T_0) - q_a] \quad (9.25)$$

Here $q_*(T_0)$ denotes the saturation specific humidity at the surface temperature T_0 . The surface dryness factor $\beta = \min(1, W_o/W_k)$, where W_o is the soil moisture fraction (volume of water per unit volume of soil) in a “skin moisture layer” of depth d_w , and W_k is a threshold value for free evaporation. W_o is constrained to satisfy $0 \leq W_o \leq W_{mx}$, where W_{mx} is a “runoff threshold.” The bucket model is attributed to Manabe (1969; MWR 97); at each grid point there is a bucket and evaporation is reduced from a “potential” value by the ratio of soil water in the bucket W_o to a “field capacity” (saturated soil) value W_k .

The force-restore method for W_o is of the form:

$$\frac{\partial W_o}{\partial t} = \frac{c_1}{\rho_w d_w} (R - E) - \frac{c_2}{\tau} (W_o - W_p) \quad (9.26)$$

Here the first term on the rhs is the “force” term, the excess of the precipitation rate (R) over the evapotranspiration rate E ; ρ_w is the density of water, and c_1 is a proportionality constant. The second term is the “restore” term, modelling recharge (discharge if $W_o > W_p$) of the surface skin moisture layer from (to) deep soil moisture (deep soil has moisture fraction W_p). This model for the evolution of T_0, q_0 is physically defensible (terms used have unambiguous meaning and could be measured) for an ocean or a large lake or a uniform bare soil. But even for a spatially uniform cover of vegetation, there is really no well-defined plane to which can be ascribed a temperature

and humidity which drive the interaction with the atmosphere: so there is an element of artificiality in the adoption of the force-restore method (with the aid of bogus volumetric heat capacities, etc.) to predict the future values of the specific humidity and temperature at a merely nominal plane. Adding to this the spatial variation of the terrain and cover at sub-grid scale, it is clear that the formulation can be regarded as little more than qualitative. Note however that it encapsulates in the simplest way the known dominant feedbacks, and has some very nice attributes: the surface fluxes are constrained to sum to the net radiation (energy is conserved), soil moisture will never be negative, and so on. The simple surface model is unlikely to “do” anything disastrous.

CMC/RPN/MSM weather and climate models now include a simple vegetation canopy layer (within which even complex processes like stem drip are modelled). One can appreciate that with increasingly complex surface models, there is an explosion in the number of surface parameters that require to be specified (soil density, soil specific heat, field capacity, leaf area index, vegetation height, stomatal conductance and its sensitivity to light) and properties that require to be initialised (eg. deep soil moisture) at each grid point.

Chapter 10

Modelling cumulus (“physics”)

Cloud circulations are a scale of motion unresolved by synoptic and even mesoscale models, and are typified by vertical velocities far exceeding the synoptic scale (ie. spatially-averaged) values. Convection in its most intense and effective form is associated with convective clouds (Cumulus), which may be individual, or organised into groups with a scale on the order of the scale of synoptic systems. According to Anthes (1977) “the vertical release and redistribution of energy by small-scale (horizontal dimensions 1-10 km) convective cells exert a significant influence on the evolution of the mesoscale and perhaps even synoptic scale systems,” and “the most widely-studied example of the influence of small-scale convection on a larger scale system is the cooperation between the cumulus clouds and the tropical cyclone.” According also to Anthes “there is also evidence that condensation heating plays a role in extratropical cyclones.” Clouds, provided precipitation reaches the ground, heat and dry their total column of the atmosphere (water out, latent heat left behind). If, in the mean, condensation exceeds evaporation at

any level, there will be heating and drying at that level. Overall, condensation exceeds evaporation aloft, but with the reverse at low levels.

Indirect evidence for the role of convection is that although during day-time steady heat input at ground permits maintenance of an unstable lapse rate despite the very rapid vertical heat transport by convection (whose “effort” is to redistribute heat and weaken the lapse rate), away from the ground persistent unstable lapse rates are not characteristic of the atmosphere - they tend to be removed by the enhanced heat transfer accompanying convection. According to Simard and Girard (CMC Vol IV #4), use of Primitive Equation models in NWP has made parametrization of the role of convection essential: for otherwise the lapse rate may become unstable during a numerical forecast, and intense (and false) synoptic scale vertical velocities can develop and ruin the large-scale forecast. Apparently the first schemes were “convective adjustment schemes” (see below) such as Manabe et al. (1965; MWR 93, p769).

Again from Anthes (1977), “Considerable effort has been directed towards representing the net effect of the many short-lived convective cells (unresolved convective motions) on the environment (model-resolved variables).” Cumulus parametrization schemes fall into two main classes:

10.1 Convective adjustment schemes

The virtual temperature $T_v = T(1 + 0.61 q)$ is defined as the temperature of hypothetical dry air having the same pressure and density as a sample

of moist air. When we consider buoyancy forces, then, it is the virtual temperature that is important¹ (in most cases the difference between T and T_v is small). Thus where we speak of lapse rates below, they are strictly lapse rates of the virtual temperature. Let $\Gamma = -\partial T_v/\partial z$ be the actual lapse rate, and let Γ_d, Γ_m be the dry and moist adiabatic lapse rates (benchmarks).

Suppose the model has computed the temperature and moisture fields for the end of the timestep, neglecting any possible phase-change and unstable lapse rates. The convective adjustment scheme the adjusts the temperature profile to neutral stability and the moisture profile to saturation or sub-saturation, by a procedure that conserves energy and water. The algorithm pays no heed to the mechanisms and scales of “convective adjustment” but simply acts in the assurance that some sort of convective process would, in the real world, have carried out these adjustments.

Dry Convection

Consider a model layer $z_B - z_T$ that is unsaturated ($q < q_*$) but unstable ($\Gamma > \Gamma_d$). Adjust the virtual temperature (but lets drop the subscript v) temperature by $\delta T(z)$ throughout the layer to obtain a revised neutral lapse rate $\Gamma = \Gamma_d$ subject to the constraint

$$\int_{z_B}^{z_T} \rho c_p \delta T(z) dz = 0 \quad (10.1)$$

¹Iribarne & Godson (Atmospheric Thermodynamics) show that the vertical acceleration of a parcel displaced vertically from its initial position is $\frac{d^2 z}{dt^2} = g \frac{T'_v - T_v}{T_v}$ where the ' denotes properties of the parcel, the unprimed the properties of the environment.

Non-convective large scale condensation

Prior to the application of the adjustment scheme the model layer is super-saturated (specific humidity $q > q_*$) but stable ($\Gamma < \Gamma_d$).

Adjust both T and q by amounts $\delta T(z)$, $\delta q(z)$ subject to the constraints

$$\begin{aligned} -L \rho \delta q(z) &= \rho c_p \delta T(z), \quad (\delta T > 0) \\ q(z) + \delta q(z) &= q_*(T + \delta T, p) \end{aligned} \quad (10.2)$$

Here there is no vertical energy transport (thus “non-convective” local condensation) so energy is locally exchanged between latent and sensible form.

Moist convection

The layer is super-saturated ($q > q_*$) and unstable ($\Gamma > \Gamma_m$). Readjust to obtain (saturated) neutrality $\Gamma = \Gamma_m$ such that:

$$\begin{aligned} - \int_{z_B}^{z^T} L \rho \delta q(z) dz &= \int_{z_B}^{z^T} \rho c_p \delta T(z) dz \\ q(z) + \delta q(z) &= q_*(T + \delta T, p) \end{aligned} \quad (10.3)$$

This is solved numerically by successive approximations. When condensation occurs, the resulting precipitation is

$$P \text{ [kg m}^{-2}\text{]} = - \int_{z_B}^{z^T} \rho \delta q(z) dz \quad (10.4)$$

and the latent heat is released instantly to the layer.

By stabilizing one layer the stability of adjacent layers is modified, and though one might think the process would spread over a large number of

layers and thus simulate deep convection, in practise that does not seem to be the case. It has been common for modellers to mess with the humidity threshold for condensation in order to tweak this scheme. Convective adjustment remains an algorithm in many modern models, surprisingly, even models that boast other cloud-influence algorithms. The “Manabe scheme” was the first widely used convective adjustment procedure; a more recent convective adjustment scheme is that of Betts and Miller (1986).

10.2 Cloud-model schemes for cumulus parametrization

Cloud model schemes consider the effect of subgrid scale penetrative convection (or shallow convection, ie. stratus, etc., that is, forms of cloud that are increasingly important with modern high-vertical resolution models) on the large scale by a model of the clouds themselves. “Once the cloud distribution and properties (temp, humidity, momentum) are determined, their effect on the environment may be computed as additional ”eddy terms” (convergences of fluxes carried by unresolved scales) in the larger-scale equations.” (Anthes, 1977).

We’ll cover the classic and simplest deep-Cumulus parametrization scheme of Kuo (1965) which was later improved by Kuo (1974), Anthes (1977), Emanuel (1991). Reviews are given by Emmanuel and Raymond (1992, 1993). Kuo’s scheme (“almost as easy to apply as Manabe’s convective adjustment, but superior” - Simard and Girard, CMC IV #4) was developed

in the context of modelling tropical cyclones, but he states it is “in no way limited to the hurricane problem investigated here.” Global weather models sometimes boast a “modified” Kuo scheme. The concept of the scheme is that at the end of each timestep, having stepped forward to predict provisional fields of temperature \tilde{T} and specific humidity \tilde{q} without allowing for the effects of Cumuli, we correct these temperatures and humidities. We do so, however, in a manner a little more skillfull than that called above “convective adjustment.” If μ is the fractional area of sky that is covered by deep cumulus clouds, then the temperature at level z after the dissolution of the cloud (mixing) will be

$$T(z) = \tilde{T}(z) + \mu \left(T_c(z) - \tilde{T}(z) \right) \quad (10.5)$$

where \tilde{T} is the environmental temperature prior to this mixing-in of the cloud air and $T_c(z)$ is the temperature in the cloud.

It is a necessity of course that all quantities that must be evaluated in the parameterization should be derivable from the large scale (ie. model resolved) variables. In describing his scheme, Kuo states “we shall derive formulas that express the latent heat released by the deep cumulus purely in terms of parameters of large scale quantities.” Kuo assumes that “deep cumulus convective motions bring moist surface air directly to high levels.”

Cloud Occurrence

It is assumed that cumulus convection always occurs² in “deep” (quantify?) layers having the potential to support convection: specifically, layers of conditionally (or unconditionally) unstable stratification over areas of mean low-level convergence.

Cloud Location and Properties

Cloud base (z_B) is presumed to be the lifting condensation level (LCL) of surface air (in the CMC Spectral Model, the LCL is computed by assuming a parcel from the surface arrives at the top of boundary layer carrying the height-average properties of the boundary-layer, ie. boundary-layer mean temperature and humidity). The in-cloud vertical distribution of temperature $T_c(z)$ and specific humidity $q_c = q_*(T_c)$ are presumed moist adiabatic (the profiles of T_c and q_c for the presumed cloud are therefore readily calculated). Cloud top (z_T) occurs where this moist adiabat from the lifting condensation level crosses the environmental “sounding” (presumably the latter is estimated from the model predictions prior to application of the scheme). Complete mixing (cloud dissolution) of the cloud, level-by-level, with the environment is assumed.

²Cloud occurrence is quantified in the parameterisation scheme by a bulk dimensionless indicator μ for which a value $\mu = 0$ would indicate no convection.

Cloud Duration

“The cumulus clouds exist only momentarily,” (Kuo), and “dissolve by mixing with the environmental air (also called by Kuo the ”mean surroundings”) at the same level, so that heat and moisture carried up by the cloud air are imparted to the environmental air.” The assumption of instantaneous dissipation is justified by the fact that the longevity of individual clouds is short relative to the timescale of the evolution of the large-scale weather.

Moisture Supply

The total rate [$\text{kg m}^{-2} \text{s}^{-1}$] at which water vapour becomes available to the vertical column above unit ground area is

$$M_t = - \int_0^\infty \nabla_H \cdot (\rho q \vec{V}_H) dz + E_0 \quad (10.6)$$

where E_0 is the surface evaporation rate (the subscript ‘H’ indicates the horizontal components only are retained). Kuo speaks of “control of the water vapour supply (for tropical storms) by the low level mean flow field,” presumably meaning that both vapour density and the flow convergence are numerically largest near the ground/ocean (maximum cross-isobar flow). When we multiply M_t by a time interval Δt (the model timestep) we have an amount $M_t \Delta t$ of vapour (per unit ground area) available to make (and therefore “used” by the model to make) cloud columns from environmental air.

Moisture Needed

In order to form a cloud column spanning ($z_B - z_T$) it is assumed we raise the temperature from environmental temperature T_e (in practise given by the model output \tilde{T} before application of the cloud parameterisation scheme) to T_c (known, see above) by condensing water vapour (it is also assumed that all this condensed water is precipitated out); the amount of water vapour needed per unit ground area is easily calculated as

$$W_1 [\text{kg m}^{-2}] = \frac{\rho c_p}{L} \int_{z_B}^{z_T} (T_c(z) - \tilde{T}(z)) dz \quad (10.7)$$

and is to be drawn from the “accession flux” M_t . In addition to this “condensing part,” there is a non-condensing “humidification part,” that raises the humidity of the cloud column ($z_B - z_T$) to saturation,

$$W_2 [\text{kg m}^{-2}] = \int_{z_B}^{z_T} \rho (q_c - \tilde{q}) dz \quad (10.8)$$

This is also readily calculated, and likewise to be drawn from the accession flux

Fractional sky cover

The dimensionless ratio

$$\mu = \frac{M_t \Delta t}{W_1 + W_2} \quad (10.9)$$

is the ratio of the amount of vapour available (the supply) to the amount of vapour needed for cloud formation, over timestep Δt . If the timestep is sufficiently small, we can ensure $\mu < 1$ (moisture supply $M_t \Delta t$ too small

relative to the required moisture for the cloud column); then μ can be considered as “the fractional area of the sky that is covered by newly formed cumulus cloud as a result of the accession of moisture by advection and by evaporation from below.” And μ can be used in the above-suggested manner to correct the forecast. However Kuo suggested that μ be interpreted only “somewhat figuratively” as the fractional cloud cover; and he argued that even if Δt is sufficiently large that μ exceeds 1, the correction procedure is still valid.

Thus if \tilde{T}, \tilde{q} are the forecast temperature and specific humidity for the end of the timestep *without* allowing for the effects of cumuli, then the cumulus-corrected temperature and humidity are:

$$\begin{aligned} T &= \tilde{T} + \mu (T_c - \tilde{T}) \\ q &= \tilde{q} + \mu (q_c - \tilde{q}) \end{aligned} \quad (10.10)$$

Precipitation

The precipitation is that part of the moisture used in warming the air from \tilde{T} to T , so the mass of precipitation falling out of the cloud layer on unit area over the timestep is

$$P \text{ [kg m}^{-2}\text{]} = \frac{\mu}{L} \int_{z_B}^{z_T} \rho c_p (T_c - \tilde{T}) dz \quad (10.11)$$

It was soon noted that this formulation underestimated the convective precipitation rate and therefore atmospheric warming in the tropics, apparently because an excessive fraction of the moisture accession is used to hu-

midify the column. Kuo (1974) suggested a more-realistic partitioning of the moisture accession. Anthes (1977) noted that “imperfectly represented cloud processes may interact with the larger-scale model flow in unrealistic ways that are not permitted with the simpler (convective) adjustment schemes” and presented a “generalised framework for representing the effect of deep, precipitating cumulonimbus clouds on their environment in mesoscale or synoptic scale models. It is not intended to model shallow, non-precipitating clouds.”

Chapter 11

Lagrangian modelling of dispersion due to unresolved velocity field

What is often called “eddy diffusion”, and modelled as such, is in reality a convection process, better called “turbulent convection” or “convection by the unresolved flow”. The word “dispersion” is preferable to “diffusion”, for it encodes one’s recognition that the process is not a diffusion process. We will consider the dispersion of a non-buoyant and non-reactive species whose instantaneous concentration is c . It is evident that in the context of turbulent or unresolved flow, there is prospect of success only for a theory of concentration *statistics*. So we shall be focused on modelling the mean concentration (C or equivalently \bar{c}).

The mass conservation equation is:

$$\frac{\partial c}{\partial t} = -\nabla \bullet (\bar{\mathbf{u}} c - D \nabla c) + Q \quad (11.1)$$

and we already know that if we perform a Reynolds decomposition $c = C + c'$

(etc.) then we will find

$$\frac{\partial C}{\partial t} = - \frac{\partial}{\partial x_i} (U_i C + \overline{u'_i c'}) \quad (11.2)$$

(where I have dropped the molecular diffusion term, and assumed there is no in-situ production). The term $\partial \overline{u'_i c'} / \partial x_i$ is the flux of c carried by the unresolved motion (turbulent or unresolved flux), and its divergence impacts the evolution of the mean (resolved) concentration. The eddy diffusion model introduces a closure assumption, viz.

$$\overline{u'_i c'} = - K_{ij} \frac{\partial C}{\partial x_j} \quad (11.3)$$

There is no a priori reason not to regard the eddy diffusivity as a tensor; however this is rarely found to be necessary, and more typically the eddy-diffusion model (K-theory) is expressed

$$\begin{aligned} \overline{u'c'} &= -K_x \frac{\partial C}{\partial x} \\ \overline{v'c'} &= -K_y \frac{\partial C}{\partial y} \\ \overline{w'c'} &= -K_z \frac{\partial C}{\partial z} \end{aligned} \quad (11.4)$$

where these K's are scalar eddy diffusivities ($[m^2 s^{-1}]$), allowed to be different for different directions of motion.

But the mass conservation equation can also be written

$$\frac{dc}{dt} = D \nabla^2 c + Q \quad (11.5)$$

In general, transport by (true, ie. molecular) diffusion can be neglected relative to convection provided the Peclet number $P_e = \frac{UL}{D} \gg 1$ where U, L

are velocity and length scales of the turbulence. Therefore, provided there are no in-situ sources or sinks of c ,

$$\frac{dc}{dt} = 0 \quad (11.6)$$

This is the basis for the Lagrangian approach to turbulent dispersion. It says that away from any source, concentration is constant along trajectories. If a fluid element is at $t = 0$ “tagged” with concentration C_0 , its concentration will remain constant at $c = C_0$ - or almost so, except for a slow “leakage” through the diffusion term on the rhs.

Consider a point source emitting N particles (or kilograms) per second, and a volume V downstream. Any given particle, say particle- i , may or may not pass through volume V . Let t_i be the time it spends in V (perhaps composed of many intervals of occupancy separated by periods of absence from V). Then if \bar{t} is the average time a particle spends in V , the mean concentration therein is simply:

$$C = \frac{N \bar{t}}{V} \quad (11.7)$$

Thus if we can calculate trajectories, we can calculate mean concentrations.

11.1 Taylor’s Lagrangian Theory

Taylor’s (1921) analytical theory for dispersion in homogeneous turbulence shows up the weakness in the Eulerian approach, and suggests the important factors in constructing a Lagrangian model. Consider the vertical motion

(there is nothing special about our choice of the vertical axis) of a fluid element which is marked purple and released at $z = t = 0$ into an infinite domain of homogeneous and stationary turbulence¹. At subsequent times (t) its vertical coordinate is:

$$Z = \int_0^{t'=t} W(t') dt' \quad (11.8)$$

where $W(t')$ is the particle velocity. This vertical velocity is by definition

$$W(t') = \left(\frac{dZ}{dt} \right)_{t=t'} \quad (11.9)$$

Now we may write

$$\frac{dZ^2}{dt} = 2 Z(t) \frac{dZ}{dt} \quad (11.10)$$

If we consider the release of many particles in such a way that they sample many realisations of the transporting velocity field, we may define an ensemble average value for any property of these fluid elements at any time t since release. For example, $\overline{Z^2}$, “mean square Z ,” denotes the mean value (in an ensemble average sense) of the square of the particle displacement along the z -axis at time t since release. Obviously this somehow relates to spread, and we will define the “spread” as:

$$\sigma_z = \sqrt{\overline{Z^2}} \quad (11.11)$$

In fact σ_z is the standard deviation of the concentration distribution along z .

The particle position (transition) probability density function $p(z, t|0, 0)$, [m^{-1}]

¹In an alternative and broader terminology, the prescription is that *velocity statistics* of the unresolved motion do not vary in space or time.

is defined such that $p(z_1, t|0, 0)dz$ is the probability that the particle released at $z = t = 0$ is at later time t between $z_1 - dz/2$ and $z_1 + dz/2$. Often this (transition) probability density will be called simply the concentration (but note again that it is a density on a line, rather than in a volume). Clearly it must satisfy

$$\int_{-\infty}^{\infty} p(z, t|0, 0) dz = 1 \tag{11.12}$$

since the particle must lie *somewhere* on the z -axis. Now by the usual rules of probability, the mean square particle displacement is

$$\overline{Z^2}(t) = \int_{-\infty}^{\infty} z^2 p(z, t|0, 0) dz \tag{11.13}$$

and σ_z is just the standard deviation of the concentration distribution.

Now we return to derivation of a prediction for σ_z . The ensemble-averaging operation we assume as usual to commute with differentiation or integration.

Then

$$\begin{aligned} \frac{d\overline{Z^2}}{dt} &= \overline{\frac{dZ^2}{dt}} \\ &= \overline{2 Z \frac{dZ}{dt}} \\ &= \overline{2 W(t) \int_0^{t'=t} W(t') dt'} \\ &= 2 \int_0^{t'=t} \overline{W(t) W(t')} dt' \end{aligned} \tag{11.14}$$

Now define a Lagrangian velocity autocorrelation coefficient

$$R(t, t') = \frac{\overline{W(t) W(t')}}{\sigma_w^2} \tag{11.15}$$

which, defining $\tau = t - t'$, may be re-written:

$$R(t', \tau) = \frac{\overline{W(t') W(t' + \tau)}}{\sigma_w^2} \quad (11.16)$$

But statistics of stationary turbulence are independent of any particular “starting time” t' , so that $R(t', \tau) = R(\tau)$ (and in fact it is also true that $R(-\tau) = R(\tau)$). Therefore, noting $d\tau = -dt'$ and $\tau = 0$ when $t' = t$ and $\tau = t$ when $t' = 0$, we have:

$$\frac{d\overline{Z^2}}{dt} = 2 \sigma_w^2 \int_0^{\tau=t} R(\tau) d\tau \quad (11.17)$$

This is Sir Geoffrey Ingram Taylor’s famous result for the rate of increase of mean square particle displacement, which is revealed to depend on the Lagrangian velocity autocorrelation function. Taylor’s solution does not tell us the form of the concentration distribution, but gives us the second moment of that unknown distribution.

Note that we can define a Lagrangian integral time scale from $R(\tau)$ as

$$T_L = \int_0^{\tau=\infty} R(\tau) d\tau \quad (11.18)$$

11.1.1 Comparison with Classical Eddy-Diffusion Solution

Remember we are considering the time-dependent evolution of the vertical position of a particle released at $z = t = 0$, and we aim to determine ensemble mean statistics (average over many trial releases). Now with the eddy diffusion closure the governing mass conservation equation reduces to

$$\frac{\partial C}{\partial t} + \bar{u} \frac{\partial \bar{c}}{\partial x} = \frac{\partial}{\partial x} \left(K_x \frac{\partial C}{\partial x} \right) + \frac{\partial}{\partial y} \left(K_y \frac{\partial C}{\partial y} \right) + \frac{\partial}{\partial z} \left(K_z \frac{\partial C}{\partial z} \right) \quad (11.19)$$

where (of course) we have assumed there is no in-situ production or destruction ($Q = 0$), and neglected molecular transport. Now let us set the mean velocity to zero, and assume the K 's are independent of position - homogeneous turbulence. Furthermore, since we don't care where in x, y the particle is, we can define a new concentration $C_*[m^{-1}]$

$$C_*(z, t) = \int_{-\infty}^{x=\infty} \int_{-\infty}^{y=\infty} C(x, y, z) dx dy \quad (11.20)$$

Thus C_* is the earlier-mentioned probability density $p(z, t|0, 0)$ for particle position, in the sense that $C_*(z_1, t)dz$ is the probability that a particle released at $z = t = 0$ is at time t lying in the range $z_1 - dz/2$ to $z_1 + dz/2$. By performing the double integral in x, y on the mass conservation equation we find C_* satisfies:

$$\frac{\partial C_*}{\partial t} = K_z \frac{\partial^2 C_*}{\partial z^2} \quad (11.21)$$

It is easy to show that the solution to this equation that satisfies

$$C_*(z, 0) = \delta(z - 0) \quad (11.22)$$

(unit release at $t = z = 0$) is the Gaussian (ie. Normal) distribution

$$C_*(z, t) = \frac{1}{\sqrt{2\pi} \sigma_z} e^{-\frac{z^2}{2\sigma_z^2}} \quad (11.23)$$

with standard deviation $\sigma_z(t) = \sqrt{2 K_z t}$.

Observations show that the concentration distribution is more or less Gaussian, for all distances from the source (but remember: this is an unbounded domain of homogeneous turbulence). But the really important point

is that this result states that

$$\frac{\partial \sigma_z^2}{\partial t} = 2 K_z \quad (11.24)$$

while G. I. Taylor's result, which is exact, says

$$\frac{d\overline{Z^2}}{dt} = 2 \sigma_w^2 \int_0^{\tau=t} R(\tau) d\tau \quad (11.25)$$

To reconcile these results we must have:

$$K_z = \sigma_w^2 \int_0^{\tau=t} R(\tau) d\tau \quad (11.26)$$

The eddy diffusivity, which we would have liked to be a property of the turbulence, turns out to be a function of the time since the particle was released: and only when $t \gg T_L$ (in the "far field") does the eddy diffusivity assume a value independent of t , namely the "far field limit" $K_\infty = \sigma_w^2 T_L$.

Thus we conclude the classical diffusion theory, which predicts a plume width σ_z increasing as \sqrt{t} for all t is wrong at short travel times. In truth, σ_z goes linearly with t in the "near field" (when $t \ll T_L$). In the near field, travel time is so short that the particle flies with an essentially constant velocity - that it assumed upon release.

Taylor's result for the time-derivative of $\overline{Z^2}$ may be further integrated to obtain:

$$\overline{z^2} = 2 \sigma_w^2 \int_0^t (t - \tau) R(\tau) d\tau \quad (11.27)$$

If, as is often assumed, the autocorrelation function is exponential with timescale T_L , ie.

$$R(\tau) = e^{-\frac{\tau}{T_L}} \quad (11.28)$$

the result for the spread specialises to:

$$\frac{\overline{z^2}}{2\sigma_w^2} = t T_L - T_L^2 \left[1 - e^{-\frac{t}{T_L}} \right] \quad (11.29)$$

11.1.2 Asymptotic results

The following results are valid for a general autocorrelation function.

Small time limit (near field): If $t \ll T_L$ then $R(\tau) = 1$ throughout the required integral. It follows that:

$$\overline{z^2} = \sigma_w^2 t^2 \quad (11.30)$$

or $\sigma_z = \sigma_w t$. The plume spreads linearly in time for short times after release: that is because for short times the spread is not a diffusion process, but simply an advection at a constant random starting velocity.

Large time limit (far field):

If $t \gg T_L$ then $\sigma_z = \sqrt{2\sigma_w^2 T_L} \sqrt{t}$.

11.1.3 What is important about G.I.Taylor's result and when is K-theory valid?

- If we insist on using an eddy diffusivity K to close the mass conservation equation, we have to accept that, of the totality of particles seen at a point z , a sub-ensemble of particles having been released at time t_1 may have a different diffusivity from a sub-ensemble of particles that were released at t_2 . This is logically absurd, the result of K-theory being fundamentally wrong.

- The statistical evolution of concentration in this simple turbulence is completely governed by two statistical properties of the turbulence: σ_w, T_L .
- The “far field” diffusivity is $K_\infty = \sigma_w^2 T_L$.

We expect K-theory to be useful when the turbulence is “fine-grained” relative to the length scale over which there is curvature in the concentration distribution (Corrsin, 1974).

11.2 Modern single-particle Lagrangian stochastic models

Taylor’s analytical solution is highly restricted in its validity. With the arrival of computers came the possibility of “Monte-Carlo” type Lagrangian simulations that treat dispersion by literally computing the paths of computational particles. The name “Lagrangian stochastic model” has emerged as the most common term for what is also known as the trajectory-simulation, random-flight, random-walk, or (occasionally) Markov-chain method.

We shall consider the modelling of dispersion in a flow whose velocity statistics are known, and horizontally-uniform. For the time being we will concentrate on motion along a single dimension (z). Evidently if we can generate a sequence for the vertical velocity W of a “particle” we can generate a set of trajectories.

These are not real trajectories. Then we must ask, what properties of

real trajectories, ie. what statistical properties of the Lagrangian velocity W , is it necessary to mimic correctly? Surely, from previous considerations, σ_w and T_L . The first LS models were heuristic, but due to the contribution of Thomson (1987) we now have rigorous criteria.

11.2.1 Zeroth- and first-order LS models

The “zeroth-order” LS model for motion along a single axis (say z , the vertical), is a random walk in position (the drunkard’s walk, or Random Displacement Model, RDM). Over each time increment dt the increment dZ in particle position is given by:

$$dZ = a dt + b d\xi \quad (11.31)$$

where a, b are deterministic coefficients, and $d\xi$ is drawn randomly from a Gaussian distribution with mean zero, and (by convention) variance dt . This model is easily shown to be equivalent to the treatment of turbulent convection as a diffusion process, and for stationary turbulence

$$\begin{aligned} a &= \frac{\partial K}{\partial z} \\ b^2 &= 2 K \end{aligned} \quad (11.32)$$

The RDM model is fundamentally wrong, and manifests as such close to a source.

In the Random Displacement model, X_i is treated as Markovian, ie. correlation of particle velocity from one timestep to the next is ignored: thus the problem that the model is invalid for travel times short compared to

the typical velocity correlation timescale. This deficiency is remedied in the “first-order” LS model, wherein the Markovian state variable is (X_i, U_i) , and the velocity evolves in time according to a “generalised Langevin equation” (Thomson, 1987),

$$dU_i = a_i dt + \sqrt{C_0 \epsilon} d\xi_i \quad (11.33)$$

where dU_i are the increments over timestep dt in the three components of Lagrangian velocity, ϵ is the rate of dissipation of the kinetic energy of the unresolved flow (TKE), and C_0 is a dimensionless and (supposedly) universal constant (prescription of the universal constant C_0 in an LS model is equivalent to the prescription of the turbulent Schmidt number S_c in a K-theory model). The 1st- order LS model correctly predicts the rate of dispersion even in the near field of a source, where travel time t is not large w.r.t. T_L , in contradistinction to the RDM and to Eulerian models.

11.2.2 Time step for (1st order) Lagrangian stochastic simulations

Integral (outer) scales measure typical spatial and temporal persistence of the unresolved (turbulent) velocity. In stationary, homogeneous turbulence, the Lagrangian integral timescale can be determined from the velocity variance (σ_w^2) and the TKE dissipation rate (ϵ) as (Tennekes, 1979)

$$T_L = \frac{2 \sigma_w^2}{C_0 \epsilon} \quad (11.34)$$

In a simulation we should ensure the timestep $dt \ll T_L$

11.2.3 Thomson's well-mixed constraint on first-order LS models

Suppose we would like to determine the coefficient a in the generalized Langevin equation for vertical motion

$$dW = a dt + b d\xi \tag{11.35}$$

where the coefficient b is known ($b = \sqrt{C_0 \epsilon}$).

Let us define $p(z, w, t)$ as the density function (in $z - w$ space) for particle location. For example, for release at $z = t = 0$ with a random velocity from the Eulerian pdf $g_a(w)$,

$$p(z, w, 0) = \delta(z - 0) g_a(w) \tag{11.36}$$

Now, being a statistic, $p(z, w, t)$ undergoes a deterministic evolution. Over discrete intervals Δt , it evolves according to the ‘‘Chapman-Kolmogorov’’ equation

$$p(z, w, t + \Delta t) = \int_{z_0} \int_{w_0} p(z, w, t + \Delta t | z_0, w_0, t) p(z_0, w_0, t) dz_0 dw_0 \tag{11.37}$$

which expresses conservation of probability (at any time after release, the particle must exist somewhere with some velocity). It is possible to derive from the CK equation a differential equation for the evolution of the density function. Provided the random forcing ($d\xi$) is Gaussian (which it must be, on other grounds), the differential equation corresponding to the stochastic

model (eqn 11.35) is the “Fokker-Plank equation,”

$$\frac{\partial p(z, w, t)}{\partial t} = - \frac{\partial}{\partial z} (w p) - \frac{\partial}{\partial w} (a p) + \frac{1}{2} \frac{\partial^2}{\partial w^2} (b^2 p) \quad (11.38)$$

The rhs is the divergence of a probability flux (in $z - w$ space): $w p$ is the flux along the z -axis, while $a p - \frac{1}{2} \partial(b^2 p)/\partial w$ is the flux along the w -axis.

Now, we can apply our LS model to any initial distribution of tracer. Suppose we apply it to the particular case where the tracer is initially well-mixed in the flow,

$$p(z, w, 0) \propto \rho(z) g_a(z, w) \quad (11.39)$$

where I have implicitly assumed stationarity of the flow, since the Eulerian velocity pdf g_a lacks time-dependence. Assume the fluid density ρ is constant. Then, $p(z, w, 0) \propto g_a(z, w)$ and, since the particles are initially well-mixed, they must remain so, ie. $p(z, w, t) \propto g_a(z, w)$. Hence, g_a must be a stationary solution to the FP equation,

$$\frac{\partial}{\partial w} (a g_a) = - \frac{\partial(w g_a)}{\partial z} + \frac{1}{2} \frac{\partial^2 (b^2 g_a)}{\partial w^2} \quad (11.40)$$

Since b is known (see above), this result determines the conditional mean acceleration a , for a specified Eulerian velocity pdf! This is a very powerful result, constraining the Lagrangian stochastic model.

11.2.4 Unique 1-d LS Model for Gaussian Inhomogeneous Turbulence

It is usual to treat surface-layer turbulence in the atmosphere as Gaussian, but vertically-inhomogeneous. And except very close to ground (or within a

canopy; a case we exclude here), the mean velocity is large w.r.t. the typical fluctuations (u'): thus dispersion in the streamwise direction is frequently ignored, or at least, u' is treated as independent of w . This leads to a focus on 1-component models, for the particle's vertical velocity, W .

The Eulerian velocity pdf, if assumed Gaussian, is

$$g_a(z, w) = \frac{1}{\sqrt{2\pi} \sigma_w(z)} \exp\left(-\frac{w^2}{2\sigma_w^2}\right) \quad (11.41)$$

It follows that the (unique) well-mixed, 1-component (here W) model for stationary Gaussian turbulence (Thomson, 1987) is

$$\begin{aligned} a &= -\frac{C_0 \epsilon(z)}{2\sigma_w^2(z)} W + \frac{1}{2} \frac{\partial \sigma_w^2}{\partial z} \left(\frac{w^2}{\sigma_w^2} + 1 \right) \\ b &= \sqrt{C_0 \epsilon(z)} \end{aligned} \quad (11.42)$$

The influence of atmospheric stratification enters through the turbulence parameters and the mean velocity profile $\bar{u}(z)$.

In inhomogeneous turbulence a Lagrangian timescale $T_L(z)$ can still be defined through eqn (11.34), but can no longer be interpreted as an integral time scale. In the neutral surface layer, where $\sigma_w \approx 1.3 u_*$, inhomogeneity arises solely through the TKE dissipation rate.

It can be shown that the model compared in detail with the Project Prairie Grass field observations by Wilson et al. (1981; WTK) is the discrete-time version of eqn (11.42). Using a time scale parameterisation equivalent to $C_0 = 3.1$, which value accords with latest estimates from simpler flows, WTK showed this LS model to be in excellent agreement with the observations.

They also obtained good agreement with unstable and stable PPG runs, using the parameterisations (which are consistent with $C_0 = 3.1$ in neutral conditions):

$$\begin{aligned} \frac{2 \sigma_w^2}{C_0 \epsilon} &= T_L(z) = \frac{0.5 z}{\sigma_w} \left(1 - 6 \frac{z}{L}\right)^{\frac{1}{4}}, L < 0 \\ \frac{2 \sigma_w^2}{C_0 \epsilon} &= T_L(z) = \frac{0.5 z}{\sigma_w} \left(1 + 5 \frac{z}{L}\right)^{-1}, L > 0 \end{aligned} \quad (11.43)$$

where L is the Monin-Obukhov length. These formulae can be used to parameterise the compound variable $C_0\epsilon$ in (modern) LS models of the equilibrium atmospheric surface layer.

We have focused here on vertical dispersion, and ignored the important question of the impact of the fluctuating alongwind component u' , which is correlated with the vertical velocity, and of the lateral turbulence component. The inclusion of u' and/or v' presents a difficult problem. There is firstly the difficulty that as yet we do not have a unique well-mixed 2- or 3-dimensional LS model, even if we are prepared to accept Gaussian velocity pdfs.

The second difficulty of including U, V , a difficulty not peculiar to LS modelling, is that u' and v' contain energy at much lower frequencies than w - wind direction and strength change on a continuous range of scales, from the almost-microscopic (Kolmogorov) scale, up through scales of order 10 min (cloud/meso scale), to hourly and daily scales (changing large scale pressure patterns). Therefore no matter what our choice of averaging time, there is always the likelihood of an irregular and possibly multi-modal distribution of material along the horizontal direction(s). In general the (mean) mass

distribution in x, y cannot be sharply predicted unless we can forecast the mesoscale fluctuations in wind direction. Practical means of accomplishing this involve prediction of the mesoscale wind variations with a prognostic mesoscale wind-field model and coupling these to a LS model for the turbulence.

11.2.5 Reflection at Boundaries

Except in cases that permit analytic solution, an LS model will be implemented with finite timestep Δt , and it may be necessary to tack onto the model an algorithm that assures particles remain in the computational domain (not necessarily only because the timestep is finite). The best available criterion remains the w.m.c., whose satisfaction may be checked simply by calculating (numerically) the evolution of an initially well-mixed tracer distribution.

For example for surface-layer simulations we may choose to reflect trajectories at height $z_{refl} \geq z_0$, where typically there is little loss of accuracy with $z_{refl} \sim 10z_0$. Suppose that at the end of a timestep a particle has moved to a position $Z < z_{refl}$. The reflection strategy is simply to correct the position and velocity as if the particle, like a billiard ball, had bounced off an apparent wall at z_{refl} , viz.

$$\begin{aligned} Z &\leftarrow 2z_{refl} - Z \\ W &\leftarrow -W \end{aligned} \tag{11.44}$$

11.2.6 LS models applicable in disturbed flows

Throughout this section we have assumed the flow and turbulence to be horizontally-uniform, so that the joint pdf for the Eulerian velocity fluctuations $g_a(u', v', w'; z)$ is a function only of the height z . In a disturbed flow, $g_a = g_a(\vec{u}', \vec{x})$.

Provided an analytical description of g_a is given, it remains possible to derive a well-mixed LS model for the disturbed flow - and this is a *huge* advance over the Eulerian description, where any such attempt would be desperately heuristic.

Suppose that g_a remained Gaussian. Then the turbulence would be described by the velocity standard deviations ($\sigma_u = \sigma_u(x, y, z)$, etc.) and covariances like ($\overline{u'w'} = \overline{u'w'}(x, y, z)$). The LS model will involve not only vertical derivatives of flow statistics, but also the horizontal derivatives, such as $\partial\sigma_u/\partial x$, $\partial\overline{u'w'}/\partial y$ and so on.

Of course, such an LS model demands that all these flow statistics be provided - it must be “fed” this information. It is no mean feat to supply such data. Measurements are out of the question. The only possibility, if it must be done, is a judicious combination of a sufficient number of measurements in the disturbed flow that one can prove (establish the correctness of) a disturbed flow model, which can then be used interpolatively and extrapolatively to “procure” all the needed derivatives, specifically, in the case where g_a is presumed Gaussian, all components of $\partial\overline{u_i}/\partial x_j$ and $\partial/\partial x_k \overline{u'_i u'_j}$.

Chapter 12

Parameterizing radiative flux convergence

In the mid-latitude cloudless summer atmosphere, direct atmospheric heating due to solar absorption is of order $+2 K \text{ day}^{-1}$, while the long wave cooling rate is of order $-2 K \text{ day}^{-1}$. This would hardly recommend parameterizing direct radiative heating in the context of short range weather forecasting, however, much stronger local heating/cooling rates may occur in cloudy layers. And in the context of climate models, it is essential to develop confidence that “cloud feedbacks” have been included with fidelity: according to Mitchell (2004) “Most of the range in (modelled) climate sensitivity is associated with differences in cloud feedback”, and the overall cloud feedback ranges from being moderately negative to being strongly positive, among the various climate models.

Difficulties with radiative transfer are principally computational. There exist several very different computational techniques proven to give a good description of radiative scattering and absorption. However since in princi-

ple for greatest accuracy a spectral description is needed (ie. one calculates the radiation-matter interactions in detail for each of a larger number of very narrow wavebands,- so called “line by line” methods), the most rigorous approaches are too cumbersome to be employed in weather/climate models. Widely differing approaches to computing radiation energy transfer stem from differences in the system under study, in objectives, and in computational resources. As what is (perhaps) the most surprising of models, sometimes a *diffusion* approximation of the radiative transport equation is used, a satisfactory approach provided the medium is highly scattering but only very weakly absorbing.

Stephens (1992) makes the following statement, which reminds us of the general nature of the so-called “physics” packages of modern models: “The object of any parametrization of cloud-radiation interactions for use in models of the large scale circulation of the atmosphere is to provide a simple, accurate and fast method of calculating the radiative flux distribution within the atmosphere preferably based on the appropriate physics of the problem at hand.”

Present parametrizations typically employ “band” models, ie. it is assumed that one may with suitable degree of accuracy define mean optical coefficients for scattering and absorption over a selected set of spectral intervals. Due to the manifold possible interactions of photons of various frequencies with molecules or particles/droplets in the atmosphere, it is necessary to employ finer subdivisions of the spectrum than simply the familiar di-

chotomy: shortwave (solar, about $0.1 - 4 \mu\text{m}$) versus longwave (terrestrial, about $4 - 100 \mu\text{m}$).

12.1 Definition of “Intensity” of radiative transfer in a given waveband

Radiative transfer is omnidirectional - photons are flying about in any and every direction. The fundamental descriptor of the radiation field (see Fig. 12.1) is the intensity $I(\mathbf{x}, \mathbf{s})$, with units [$\text{J s}^{-1} \text{m}^{-2} \text{steradian}^{-1}$], defined more specifically to be the intensity of energy transfer at position \mathbf{x} into a cone of unit solid angle about the direction \mathbf{s} (in a spectral treatment, the “spectral” or “monochromatic” intensity I_λ would have units [$\text{J s}^{-1} \text{m}^{-2} \mu\text{m steradian}^{-1}$]). For example, I might be the diffuse solar intensity; or the longwave intensity. (See also Liou, 2002; p4, Fig. 1.3)

The direction \mathbf{s} is characterised by a unit vector (\mathbf{s}), and relative to Cartesian axes (see Fig. 12.1) aligned with the local zenith (unit vector \hat{k}), \mathbf{s} has components

$$\mathbf{s} = \hat{i} \cos \theta \sin \phi + \hat{j} \sin \theta \sin \phi + \hat{k} \cos \phi \quad (12.1)$$

The *net* energy flux density¹ across a unit of area whose normal is oriented

¹By convention (eg. see Liou, 2002, p6) a radiative flux has units [W] while more commonly we encounter the “radiant flux density” [W m^{-2}], an example of which is the solar constant.

along the arbitrary unit vector \mathbf{d} is given by²

$$F_{\mathbf{d}}(\mathbf{x}) = \int_{4\pi} I(\mathbf{x}, \mathbf{s}) (\mathbf{d} \bullet \mathbf{s}) d\omega \quad (12.2)$$

where $d\omega = \sin \phi d\theta d\phi$. Thus for example the *net* vertical radiant energy flux density (due to all photons within the waveband in question, ie. which contribute to the intensity I) is:

$$F_z(\mathbf{x}) = \int_{\theta=0}^{2\pi} \int_{\phi=0}^{\pi} I(\mathbf{x}, \theta, \phi) \cos \phi \sin \phi d\theta d\phi \quad (12.3)$$

while

$$F_z \uparrow(\mathbf{x}) = \int_{\theta=0}^{2\pi} \int_{\phi=0}^{\pi/2} I(\mathbf{x}, \theta, \phi) \cos \phi \sin \phi d\theta d\phi \quad (12.4)$$

is the hemispheric (1-way, upward) flux, in Liou's terminology, the "irradiance" from the lower into the upper hemisphere.

Similar expressions obtain for the net fluxes in the orthogonal directions F_x, F_y , and the heating rate (again, due to photons of this waveband) is given by

$$h [\text{W m}^{-3}] = -\nabla \bullet \mathbf{F} = -\frac{\partial F_x}{\partial x} - \frac{\partial F_y}{\partial y} - \frac{\partial F_z}{\partial z} \quad (12.5)$$

So if we can compute the intensity $I(\mathbf{x}, \mathbf{s})$, we can derive all the quantities of interest.

12.2 Radiative Transfer Equation

When radiation of incident intensity $I(\mathbf{s})$ traverses a finite distance of medium, we may distinguish two modes of change:

²Note that unit vectors don't carry units: thus $\mathbf{d} \bullet \mathbf{s}$ does not, as would appear on first sight, contribute an unwanted $[\text{m}^2]$ to make us irritable and rob us of sleep.

- extinction, due to absorption and/or scattering (away from direction \mathbf{s}), which decreases intensity, and is *modeled* as: $dI = -e I da$ (ie. assumption of linearity)
- emission and/or scattering *into* the direction \mathbf{s} , which increases intensity: $dI = +e J da$

Here J is the “source function,” and (to repeat) may be non-zero due to scattering and/or due to thermal emission; e is the extinction coefficient; and da measures the “amount of matter” per unit area encountered over the path across which dI occurs. If $d\ell$ is the increment in distance³ and ρ is the mass density, then if we define $da = \rho d\ell$ [kg m^{-2}], then e [$\text{m}^2 \text{kg}^{-1}$] is called the “mass” extinction coefficient (or mass extinction cross section; eg. Liou, p27, whose k_λ is a monochromatic mass extinction coefficient); while if $da = d\ell$, e is called the “volume” extinction coefficient.

Now $e da$ is dimensionless, and it can be defined to be equal to the increment in optical path⁴, $d\tau_o \equiv e da$. Combining the above definitions we

³There is a subtlety here: as observers we will focus on a certain volume of the atmosphere, perhaps a slab, and assign a depth dz or a slant distance (eg. $dz/\cos\phi_0$, where ϕ_0 is solar zenith angle, zero if the sun is at the zenith) across the slab; but an electromagnetic wave (or photon) traversing our slab may experience multiple reflection and so, in a sense, take a much longer path. Thus it seems to me we have to think of our $d\ell$ as having to do with our mapping, not to do with distances travelled by photons.

⁴It is important to bear in mind that, in general, a distinction is made between an optical “path” and an optical “depth”. The terminology is a minefield... here for example we have a subscript o on our increment in optical path, and we drop the subscript to denote the projection of $d\tau_o$ onto the local vertical; but other authors reverse the convention.

obtain the radiative transfer equation⁵ (c.f. Liou, 2002, eqn1.4.5):

$$\frac{dI(\mathbf{x}, \mathbf{s})}{d\tau_o} = -I(\mathbf{x}, \mathbf{s}) + J(\mathbf{x}, \mathbf{s}) \quad (12.6)$$

This is a continuum equation, and it makes sense if the increment in optical path is large relative to the sizes of molecules, aerosols, cloud droplets, and anything else interacting with our radiation beam (viz., hailstones, ducks, and any farm animals remaining aloft since having been uplifted by the latest tornado).

Case of a plane-parallel atmosphere

In some circumstances it may be appropriate to consider that $I(\mathbf{x}, \mathbf{s})$ varies only with z , thus $I = I(z, \mathbf{s})$, and as a surrogate for z we may use the optical depth τ defined by Fig. (12.2). Then writing $\mu = \cos \phi$ where ϕ denotes the departure of the photon path from the local zenith, the RTE becomes⁶:

$$\mu \frac{dI(\tau, \mathbf{s})}{d\tau} = -I(\tau, \mathbf{s}) + J(\tau, \mathbf{s}) \quad (12.7)$$

Now the largest radiative flux convergences in the atmosphere will occur in cloudy or hazy atmospheres, and it is common to introduce the further simplification/assumption that the intensity $I(\tau, \mathbf{s})$ depends on the zenith angle, but not the azimuth; ie. that the dependence of I on direction \mathbf{s} , in general entailing dependence on two angles, reduces to a dependence on the

⁵The spectral equivalent of this equation is called the Schwarzschild equation.

⁶To add to the confusion where plenty is already available, depending on whether they take $d\tau/dz$ as positive or negative, different authors will have different signs for the terms in I . For example Liou takes $\tau = \int_z^\infty e(z')\rho(z')dz'$, calling this the “normal optical thickness (or depth)”, and for which obviously $d\tau/dz < 0$.

emergent zenith angle, or μ : $I = I(\tau, \phi)$. For example, suppose I is the diffuse solar intensity. The simplification under discussion is valid if, when looking up at the sky at any angle relative to the zenith, we see the same intensity no matter which compass direction we face.

Case of a non-scattering but emitting medium

This (for example) is the case of longwave radiation ($\lambda \sim 4 - 100 \mu\text{m}$) in an ideally particulate-free atmosphere (since λ is vastly greater than molecular diameter, and there are no other particles to interact with). The source function, by restriction entirely due to emission, is the Planck function.

But since the atmospheric gases are selective absorbers and emitters, a spectral description is essential, so we have to revert to

$$\mu \frac{dI_\lambda(\tau, \phi)}{d\tau} = -I_\lambda(\tau, \phi) + B_\lambda(\tau) \quad (12.8)$$

where the Planck function - with units of spectral intensity - is

$$B_\lambda(T) = \frac{2hc^2}{\lambda^5 (e^{hc/k_B\lambda T} - 1)} \quad (12.9)$$

($h = 6.626 \times 10^{-34} \text{ J s}^{-1}$ Planck's constant; $k_B = 1.3806 \times 10^{-23} \text{ J K}^{-1}$ Boltzmann's constant; c speed of light). Note that the Planck function is isotropic (a black body radiates with equal intensity into every equal element of solid angle making up the hemisphere it faces), so, for example, the Stefan-Boltzmann law for the emittance ($[\text{W m}^{-2}]$) of a black body is

$$\Phi = \pi \int B_\lambda(T) d\lambda = \sigma T^4 \quad (12.10)$$

where the π (rather than 2π , the solid angle subtended by a hemisphere) pops magically out from an integration of the projection factor $(\mathbf{s} \cdot \hat{k})$ over the 2π steradians of the hemisphere,

$$\int_{\theta=0}^{2\pi} \int_{\phi=0}^{\pi/2} \cos \phi \sin \phi d\theta d\phi = \pi \quad (12.11)$$

In the context of multi-spectral satellite remote sensing using an radiation bands in the atmospheric window, eqn (12.8) is the basis for inversion to deduce the atmospheric temperature profile $T(z)$.

Case where $J = 0$

Since the atmospheric gases emit a negligible amount of energy in the solar waveband, if we neglect multiple scattering (which first knocks a photon out of the path \mathbf{s} of the beam, then kicks it back onto another path parallel to and thus indistinguishable from \mathbf{s}) the RTE (eqn 12.7) for a plane-parallel atmosphere reduces to

$$\mu \frac{dI(\tau, \theta, \phi)}{d\tau} = -I(\tau, \theta, \phi) \quad (12.12)$$

(here the assumption of azimuthal symmetry no longer applies) and we have the familiar exponential attenuation law.

12.3 Interactions of radiation and matter

Since extinction (loss of intensity) may be the result either of absorption (complete removal of photons) or of photon path-deflection (scattering), one

splits the extinction coefficient as⁷:

$$e = k + s \quad (12.13)$$

where k is the absorption coefficient, and s the scattering coefficient. Dividing by e ,

$$1 = \frac{k}{e} + \frac{s}{e} \quad (12.14)$$

where $\omega_o \equiv s/e$ is the scattered fraction of extinguished intensity, known as the “single scattering albedo”, and $(1 - \omega_o) \equiv k/e$ is the absorbed fraction.

Note that when one considers solar radiation in a real atmosphere composed of molecules and particulates (aerosols, cloud droplets, etc.), then “Since scattering by particles and absorption by gases occur simultaneously, the exact amount of absorber along the light path cannot be known” (Fouquart and Bonnel, 1980; *Contrib. Atmos. Phys.*, Vol. 53).

Simple scattering

No interaction with any form of internal energy. (The internal energy of any large scale material system consists primarily of contributions from translational, rotational, and vibrational energy, ie. “all those energies associated with random motions and relative positions of the small parts within the system”. King, *Thermophysics*)

⁷There is much variation in terminology; for example Liou (2002) uses k for the extinction whereas I am using it for the absorption coefficient.

Scattering

The interaction produces a change in internal energy, but the excited state has a short lifetime and thus re-emission occurs with negligible loss to translational energy (coherent scattering corresponds to single-step re-emission, incoherent scattering to multi-step).

Multiple scattering

In general a scattering volume (finite volume of the continuum) will contain many scatterers; thus light scattered off one particle may be scattered again by another, and so on. Thus a photon that enters a scattering volume with direction \mathbf{s} may potentially exit that volume with the same direction of propagation after an arbitrary number of scattering interaction... all that complexity has to be integrated into the extinction coefficient, and our RTE is a superficial (macroscopic) treatment.

Absorption

The interaction produces a change in internal energy, and before re-emission of a photon, a molecular or atomic collision(s) occur(s), - entailing non-radiative transitions (“deactivation”).

Absorption and thermal emission in the atmosphere are considered isotropic processes (the one the reverse of the other) - ie. a unit mass of air has the same probability of absorbing a photon, regardless of the direction it came from; and thermal emission from unit mass is evenly distributed with re-

spect to direction). However scattering preserves a memory of the direction of incidence, ie. is not isotropic, and so a formalism is required.

12.4 Scattering Function

Define the scattering function⁸ $p(\theta, \phi; \theta', \phi')$ such that $p(\theta, \phi; \theta', \phi') d\omega/4\pi$ is the probability that a photon incident at angle (θ', ϕ') will be scattered into solid angle $d\omega$ centred on emergent direction (θ, ϕ) . A scattered photon (by definition of “scattering”) must emerge at some angle: thus we have the normalization property of p ,

$$\int_{\theta=0}^{2\pi} \int_{\phi=0}^{\pi} p(\theta, \phi; \theta', \phi') \frac{\sin \phi d\phi d\theta}{4\pi} = 1 \quad (12.15)$$

For “isotropic scattering”, by definition $p(\theta, \phi; \theta', \phi') \equiv 1$.

As an example, suppose we take $\theta' = \phi' = 0$ and assume that the scattering function has azimuthal symmetry: then the forward-scattered fraction would be

$$f = \int_{\theta=0}^{2\pi} \int_{\phi=0}^{\frac{\pi}{2}} p(\phi) \frac{\sin \phi d\phi d\theta}{4\pi} \quad (12.16)$$

Evaluation of the scattering function is a problem in physics; various exact analytical solutions are available for the interaction of an incident plane electromagnetic wave with a single particle. Of course, in application to finite volumes of the atmosphere the scattering volume will contain many

⁸The “phase function” $P(\cos \Theta)$ defined by Liou (2002, p90) differs only notationally from our p ... Liou’s Θ is the angular deflection of the scattered path away from the incident direction, and the fact that his notation is $P(\cos \Theta)$ rather than $P(\Theta)$ is of infinitesimally small (but negative) significance.

types of particles, so the bulk scattering function must represent the global consequence of the ensemble of scatterers on incident E/M waves of the given character (single wavelength λ or waveband $\lambda_1 \rightarrow \lambda_2$).

12.4.1 Rayleigh Scattering

Rayleigh scattering is the scattering of unpolarized light from very small particles (radius $r \ll \lambda$, the wavelength of the radiation). Thus the scattering of sunlight by the molecules of the atmospheric gases is Rayleigh scattering.

The scattering probability in this case is highly dependent on wavelength, varying as λ^{-4} . Thus the “Rayleigh optical depth” of the earth’s atmosphere (clear air, absorption neglected, extinction entirely through scattering) is strongly dependent on the wavelength in question (see Table 12.1). In particular, the optical depth for blue light is much greater than for red light, implying much stronger scattering of blue than red light, out of the solar beam (and thus available to strike our eye as diffuse solar radiation, ie. as a component of the sky colour).

For Rayleigh scattering of incident unpolarized radiation, the scattering function (here written s.t. emergent angles are measured w.r.t. the incident direction) has azimuthal symmetry, and is given by⁹:

$$p(\phi) = \frac{3}{4} (1 + \cos^2 \phi) \quad (12.17)$$

The scattering function gives the distribution of the angle of scattering, for photons that *are* scattered - it does *not* give the probability of scattering.

⁹Liou calls this $P(\cos \Theta)$, ie. his scattering deflection angle Θ is our ϕ .

Table 12.1: Rayleigh optical thickness of earth's atmosphere at sea level, under clear skies.

λ [μm]	τ^R
0.25	2.74
0.30	1.25
0.35	0.650
0.40	0.373
0.45 (“blue”)	0.229
0.50	0.149
0.55	0.101
0.60	0.0708
0.65	0.0512
0.70 (“red”)	0.0379
0.75	0.0287
0.80	0.0221
0.85	0.0173
0.90	0.0138
0.95	0.0111
1.00	0.0090

This is a bit confusing; what is the difference between a photon's not being scattered, and its being forward-scattered onto exactly the same path as it had before the scattering interaction?

In any case, it is interesting to note that the Rayleigh scattering function is not wavelength-dependent, and so exerts no influence on the red/blue character of the atmosphere... it is (rather) the wavelength dependence of the probability of scattering that controls sky colour.

Exercises

1. Show from eqn (12.16) that the forward integral of the Rayleigh scattering function is $f = 1/2$
2. Show that the Rayleigh scattering function has maxima $p = 3/2$ at $\phi = 0^\circ, 180^\circ$ and minima $p = 3/4$ at $\phi = 90^\circ, 270^\circ$

12.4.2 Mie Scattering

The atmosphere always contains dust, haze, and other types of particles (cloud and rain drops) in addition to the molecular size particles of the atmospheric gases, and their radiative effects must be considered. Scattering by spherical particles of radius r about equal to or greater than the wavelength λ is given by the Mie theory, which however is too complicated to be discussed here.

12.5 Two-stream model for diffuse solar radiation

We have seen that at the most fundamental level, radiative transfer is quantified by the intensity I . We have seen, too, how the heating rate due to radiative divergence is calculated from I . Provided we can supply in sufficient detail the material properties k, e (etc.) and realistically model scattering (of photons of any wavelength) by any and all forms of material in the atmosphere, there is no fundamental barrier to arbitrarily accurate computations of radiation distribution in the atmosphere. Unfortunately such a rigorous and general treatment is beyond available computing resources in the context of global climate modelling, and may well always be; and furthermore specifying the distribution of particulate matter and the scattering and absorption coefficients is no small task. Simplified models are a necessity, and Stephens (1992) considers that in the context of parameterizing the interaction of radiation and cloud it is important to distinguish

- Inherent optical properties; functions of cloud microphysics that may not be able to be supplied by large scale models
- Apparent optical properties; properties that are apparent to the medium and come about through illumination of the cloud by radiation

It is common to resort to some variant of the following “two-stream” (upward, downward) model¹⁰, and we will here visualize that model in terms of its application to the profile of diffuse solar radiation (formulae of the same nature apply for longwave radiation).

We begin by defining $F^+(\tau)$ as the upward and $F^-(\tau)$ the downward stream of radiation; these are both flux densities, having units $[\text{W m}^{-2}]$, and the net diffuse solar radiative energy flux density (defined positive for net energy transport towards earth) is

$$F^* = F^- - F^+ \quad (12.18)$$

Now, for either component of the flux we may write:

$$dF = -F d\tau = -F (k/e + s/e) d\tau \quad (12.19)$$

Let us (as earlier) define $s/e = \omega_o$, the “single-scattering albedo,” ie. the scattered fraction of the extinguished radiation; then $1 - \omega_o$ is the absorbed fraction of extinction. Since we now admit of only two directions of transfer,

¹⁰According to Lenoble (1985; Radiative Transfer in Scattering and Absorbing Atmospheres: Standard Computational Procedures), this model *cannot* be derived rigorously from the RTE. Thus the two-stream model must be regarded as heuristic, though presumably it can be ‘calibrated’ by adjusting the parameter inputs (viz., single scattering albedo; forward and backward scattering probabilities) in reference to measurements or to a fully-directional model. Stephens (1992), noting that all current radiation schemes for large scale models are based on the “plane parallel version of the radiative transfer equation” (ie. two-stream approximation) states that such a procedure “may produce arbitrarily large errors due to the subgrid scale effects of cloud morphology” (ie. lateral radiative exchange due to cloud boundaries). Be that as it may, many authors adopt this description as a starting point in the analysis of atmospheric radiation. Liou (2002) gives the two stream model in at least two forms; that of his eqns (6.5.8), and in the form of his generalized two-stream approximation, eqns(6.5.27).

our description of the directionality of scattering reduces to the specification of

- b , the backscattered fraction of the diffuse radiation
- b_o , the up-scattered fraction of the scattered part of the solar beam
- $f_o (= 1 - b_o)$, the down-scattered fraction of (the scattering from) the solar beam

and we require empirical data or a suitable physical model of the appropriate scattering process to evaluate these parameters. (Note that the form of the equations to follow indicates that the net effect of *multiple* scattering is handled by means of these adjustable parameters.)

Let $S(\tau)$ [W m^{-2}] be the beam strength measured on a horizontal surface: then $S(0) = 1370\mu_o$ where $\mu_o = \cos(\phi_o)$, ϕ_o being the solar zenith angle. Now, over a positive (downward) increment $d\tau$, the change in F^- is made up of loss by absorption and scattering; gain by forward scattering from itself; gain by backward scattering of F^+ ; and gain by downward scattering of the beam. Thus

$$dF^- = -d\tau F^- + (1 - b) \omega_o d\tau F^- + b \omega_o d\tau F^+ + (1 - b_o) \omega_o d\tau \frac{S(\tau)}{\mu_o}$$

and a similar result holds for the change in F^+ . Rearranging¹¹

$$\begin{aligned}\mu_1 \frac{dF^-}{d\tau} &= F^+ (\omega_o b) - F^- (1 - \omega_o + b\omega_o) + \omega_o(1 - b_o) \frac{S(\tau)}{\mu_o} \\ \mu_1 \frac{dF^+}{d\tau} &= F^+ (1 - \omega_o + b\omega_o) - F^- (\omega_o b) - \omega_o b_o \frac{S(\tau)}{\mu_o}\end{aligned}\quad (12.20)$$

where

$$\begin{aligned}\frac{dS}{d\tau} &= -\frac{S}{\mu_o} \\ S(\tau) &= S(0) \exp\left(-\frac{\tau}{\mu_o}\right)\end{aligned}\quad (12.21)$$

The boundary conditions are:

$$\begin{aligned}S(0) &= 1370\mu_o \\ F^-(0) &= 0 \\ F^+(\tau_{gnd}) &= \alpha (F^-(\tau_{gnd}) + S(\tau_{gnd})\mu_o)\end{aligned}\quad (12.22)$$

where α is the albedo of the ground (the second of these equations states that there is no incoming diffuse radiation at the top of the atmosphere). Analytical solution is probably possible, and numerical solution is easy. The link between increments in τ and the corresponding increments in pressure (or decrements in height) is straightforward,

$$\frac{d\tau}{dp} = -\frac{d\tau}{dz} \left(-\frac{\partial p}{\partial z}\right)^{-1} = \frac{-1}{\rho g} \frac{d\tau}{dz}\quad (12.23)$$

¹¹The projection factor μ_1 “allows (approximately) for the mean obliquity of the rays to the vertical direction” (Paltridge and Platt, “Radiative processes in Meteorology and Climatology”, p74) but is commonly suppressed with corresponding readjustment of the coefficients of all other terms (eg. Liou, p108). Schuster (1905; “Radiation through a foggy atmosphere”, *Astrophys. J.*, Vol. 21, 1-22) is credited with first use of a two-stream formulation.

Now if e is the mass extinction coefficient, $d\tau = e \rho dz$, thus

$$\frac{d\tau}{dp} = -\frac{e}{g} \quad (12.24)$$

To progress to specific predictions of diffuse radiation in (say) cloudy skies, we need specifically to know the profile of the extinction coefficient, and the other parameters that have been introduced. This connects us with the cloud microphysics. According to Stephens (1992), a cloud of depth h has optical depth (for the shortwave band):

$$\tau = \int_0^h dz \int_0^\infty n(r) 2\pi r^2 dr \quad (12.25)$$

where $n(r)$ [m^{-4}] is the droplet size distribution. The radiation budget (and thus the climate) predicted by large scale models is extremely sensitive to assumptions about the cloud microphysics.

According to Stephens (1992), the two-stream approach is extensively used in large scale models to parameterise solar radiation, and also, with the simplification that scattering is completely ignored ($\omega_o = 0, b = 0$) for longwave radiation. Apparently for some types of cloud (esp. high, cold cirrus) the latter simplification is not defensible. As regards spectral resolution, “It is customary to divide both the solar and the infrared spectrum into a number of spectral bands and sum over these bands to provide the required broadband solar and infrared fluxes. A number of factors complicate the selection of these bands: the extremely high spectral variability of molecular absorption; the superimposed slower spectral variability of droplet and crystal scattering and absorption; and, the slower spectral variation of the

sources of flux (spectral solar flux, Planck black body function). In addition to these factors, the desire to include selected (greenhouse) gases that absorb in discrete spectral and narrow intervals (eg. CH₄) complicates the choice of spectral resolution.”

12.6 Two-stream model for snowpack (spectral) albedo

This is a subject of obvious relevance to climate modelling. Measured snow albedos exceed 0.8 in the range $0.3 \leq \lambda \leq 0.7 \mu\text{m}$, but drop steeply with increasing λ in the near IR ($0.8 - 1.5 \mu\text{m}$) and remain low at larger λ (though with local maxima at $\lambda = 1.1, 1.3, 1.8, 2.2 \mu\text{m}$). As snow ages, irrespective of melting, the snow grains increase in size and sphericity (reducing the total “surface free energy” by reducing their surface-to-volume ratio) and albedo is reduced at all wavelengths (Wiscombe and Warren, 1981; *J. Atmos. Sci.*, Vol. 37, 2712-2733; WW81).

Dunkle and Bevan (1956; *J. Meteorology*, Vol. 13, 212-216) adopted the two-stream model with tunable absorption and scattering parameters. They assumed the snow/icepack is of finite depth, homogeneous and of constant optical properties, and that the “radiant energy within the medium is perfectly diffuse at any depth” (ie. no beam radiation: such a model of course cannot reproduce the salient observation that snowpack albedo increases with increasing solar zenith angle); the absorption and scattering coefficients are considered independent of depth, and the scattering is isotropic (equal prob-

abilities of forward and backward scattering). Later authors relaxed the assumption of isotropic scattering, including Bohren and Barkstrom (1974) who treated scattering from individual snow grains (assumed spherical), and properly treated the strongly preferential forward scattering; BB74 showed that it is refraction through snow grains, not reflection from them, that is mainly responsible for snow albedo.

WW81 simplified the model of BB74 and gave a model that computes the spectral albedo at any wavelength in the solar spectrum, as a function of effective snow grain size (the only adjustable parameter), zenith angle, ratio of diffuse to beam incident solar radiation, snow layer depth and albedo of the underlying surface. The WW81 model is considered to be based entirely on measurable inputs.

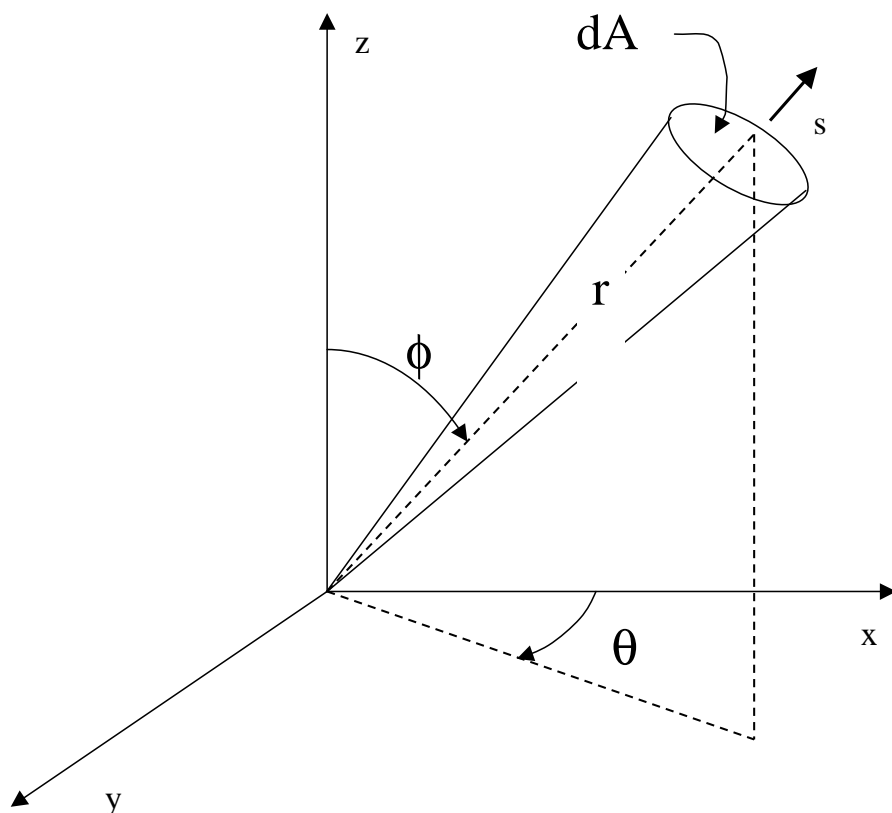


Figure 12.1: Definitional diagram for the radiative intensity I in direction \mathbf{s} . As defined here, the intensity I is a broadband property (though a spectral intensity I can be defined similarly), and has units $[\text{J s}^{-1} \text{m}^{-2} \text{steradian}^{-1}]$. The solid angle defined by the cone is $d\omega = dA/r^2$, while the element of area is $dA = (r \sin \phi d\theta) r d\phi$. (Note that Liou, 2002 gives a slightly more complex definition, in which the element of area at the entry point to the cone has its normal parallel to the z -axis rather than parallel to \mathbf{s} . Also, Liou's terminology for the angles is exchanged, ie. ϕ and θ exchange roles.)

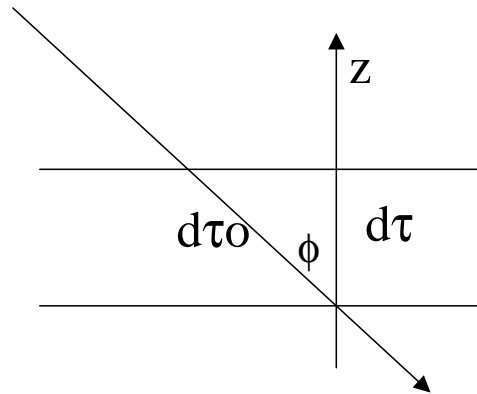


Figure 12.2: Replacing the increment of vertical distance with the projection onto the vertical axis of the increment in optical path: $d\tau = d\tau_o \cos \phi$. By integration, we may use $\tau = \tau(z)$ as the vertical coordinate. It is conventional to define $\tau = 0$ at the top of the atmosphere ($z = \infty, p = 0$). The value of τ at sea-level or ground depends on the waveband in question, and on the condition of the overlying atmosphere. Considering for the moment solar radiation in a clear sky, values at sea level are strongly wavelength-dependent, but in the range of about 1 - 0.1. By contrast, in the presence of a thick layer of cloud, τ may be as large as 50 or more.

Part III
Appendices

Chapter 13

Review of classical atmospheric dynamics

The following discussion is brief, because derivation of the equations is not a part of this course (although it is intended that you will learn something of the general character of flow equations and the meaning and origin of particular kinds of terms such as diffusion or advection terms). Note first that the word ‘filtering’, as used below, implies the neglect of some of the terms that appear in our dynamic tendency equations (ie. it is a filtering that is *additional* to that applied to define the resolved scales).

For completeness, we will write down the complete set of “raw” equations, that we would next “filter” as outlined above. The momentum equations (Navier-Stokes equations) for the rotating earth are

$$\frac{d\vec{u}}{dt} = -\frac{1}{\rho} \nabla p - 2\vec{\Omega} \times \vec{u} + \vec{g} + \vec{F} \quad (13.1)$$

(Haltiner, 1971, p3, attaches the subscript ‘3’, eg. \vec{V}_3 , ∇_3 to emphasize this is the 3-d velocity - for later he uses \vec{V} and \vec{V}_p for the horizontal velocity

respectively on constant elevation and constant pressure surfaces). In eqn (13.1) viscous momentum transfer has been neglected, the gravity vector \vec{g} absorbs also the centripetal force (a small correction), and \vec{F} is any other force. Regarding this latter, which in view of the preceding discussion should rightly include the influence of unresolved scales of motion (that is, friction: the divergence of the unresolved or turbulent momentum flux vector), in the context of the dynamics of the so-called free atmosphere, it is neglected.

The continuity equation (conservation of “air”) is

$$\frac{\partial \rho}{\partial t} = - \nabla \bullet (\rho \vec{u}) \quad (13.2)$$

or equivalently

$$\frac{d\rho}{dt} = - \rho \nabla \bullet \vec{u} \quad (13.3)$$

Conservation of water vapour, neglecting molecular diffusion (since it is overwhelmed by convection) can be expressed:

$$\frac{d\rho_v}{dt} = - \rho_v \nabla \bullet \vec{u} + Q_v \quad (13.4)$$

where ρ_v is the absolute humidity (vapour density) and Q_v is a source/sink term due to condensation/evaporation. This is just the Lagrangian form of the vapour conservation equation we exploited above in our example of filtering.

Conservation of thermodynamic energy (the first law of thermodynamics) can be written

$$c_p T \frac{d \ln \theta}{dt} = c_v \frac{dT}{dt} - \frac{1}{\rho} \frac{dp}{dt} = Q_T \quad (13.5)$$

where θ is the potential temperature and Q_T is a source/sink term, the rate of heat addition/removal by, say, radiative divergence.

Suppose now we have filtered this equation set, and we now extract from the total tendencies in the filtered (resolved) field the “dynamic” tendencies. Suppose we drop the explicit recognition of the fact that the resolved variables are averages... the resulting equations look like the raw equations, but we know they have a different meaning, since part of the tendency is not included. Anyway, if we do that, we have the equations that appear in textbooks (eg. Holton) on global meteorology.

13.0.1 Expression of the governing equations using a pressure axis

Under the assumption that the atmospheric pressure distribution is hydrostatic, the momentum equations reduce to the pair

$$\begin{aligned}\frac{d\vec{V}_H}{dt} &= -\frac{1}{\rho}\nabla_H p - 2\vec{\Omega} \times \vec{u} + \vec{F}_H \\ 0 &= -\frac{1}{\rho}\frac{\partial p}{\partial z} - g\end{aligned}\tag{13.6}$$

where \vec{V}_H is the horizontal wind vector, and ∇_H is the horizontal gradient operator ($= (\partial/\partial x, \partial/\partial y)$ in Cartesian coordinates).

The hydrostatic equation gives us a 1:1 map between z and p , so that we could if we wished use p as our vertical coordinate. But taking an even more general stance, suppose $\eta(x, y, z, t)$ was to be our vertical coordinate. Let $F = F(x, y, z)$ be an *arbitrary* field, which we now re-write as $F =$

$F(x, y, \eta(x, y, z))$. It is easy to show that:

$$\begin{aligned} \left(\frac{\partial F}{\partial x}\right)_{yz} &= \left(\frac{\partial F}{\partial x}\right)_{y\eta} + \left(\frac{\partial F}{\partial \eta}\right)_{xy} \left(\frac{\partial \eta}{\partial x}\right)_{yz} \\ \left(\frac{\partial F}{\partial y}\right)_{xz} &= \left(\frac{\partial F}{\partial y}\right)_{x\eta} + \left(\frac{\partial F}{\partial \eta}\right)_{xy} \left(\frac{\partial \eta}{\partial y}\right)_{xz} \\ \left(\frac{\partial F}{\partial z}\right)_{xy} &= \left(\frac{\partial F}{\partial \eta}\right)_{xy} \left(\frac{\partial \eta}{\partial z}\right)_{xy} \end{aligned} \quad (13.7)$$

and so if $\eta = p$, then

$$\left(\frac{\partial}{\partial z}\right)_{xy} = -\rho g \left(\frac{\partial}{\partial p}\right)_{xy} \quad (13.8)$$

13.1 Filtering the dynamical equations to eliminate sound and gravity waves

The inaccuracy of Richardson’s pioneering numerical forecast is believed to have been partially caused by a growth of spurious (non-meteorological) waves, triggered by errors in his initial data, which obscured the “real” developments. “Filtered” models “are characterised by the absence of sound and gravity waves and a balance between the wind and pressure fields” (Haltiner & Williams, 1980, p208). They are not necessarily based on divergence and vorticity equations, though usually so in practise, since that approach most easily allows filtering out of gravity waves.

13.1.1 Sound Waves

Sound waves are longitudinal waves: particle oscillation parallel to direction of propagation. When propagating horizontally, sound waves are called

“Lamb waves.” There are no meteorological phenomena in which sound waves play a significant role.

13.1.2 Gravity Waves

Gravity waves are lateral waves. Internal gravity waves (buoyancy waves) play a significant meteorological role at small scales, eg. lee waves (orographic waves). But what “is” a gravity wave?

Let $\theta_e(z)$ be the mean potential temperature profile of the atmosphere (subscript ‘e’ for environment), and let the instantaneous potential temperature of a fluid parcel be $\theta = \theta_e + \theta'$ where θ' is the local and instantaneous deviation of the potential temperature of a fluid parcel from the local mean.

The coupled equations

$$\begin{aligned}\frac{dw}{dt} &= g \frac{\theta'}{\theta_e} \\ \frac{d\theta'}{dt} &= -w \frac{\partial \theta_e}{\partial z}\end{aligned}\tag{13.9}$$

(which are in Lagrangian form: d/dt is the derivative following the parcel) describe a system in which

- a moving parcel experiences a vertical acceleration due to the buoyancy force (ie., the momentum balance contains inertia and buoyancy only; most of the gravity has been “cancelled out” by the hydrostatic pressure gradient, leaving only the modified (or “reduced”) gravity $g\theta'/\theta_e$, and
- as the parcel moves it expands or contracts adiabatically, and thus its potential temperature is constant. However relative to its environment

its fluctuation temperature θ' is rising or falling at a rate that depends on its ascent rate w .

If we differentiate the first equation to obtain d^2w/dt^2 and close using the second equation for $d\theta'/dt$, we have a wave equation

$$\frac{d^2w}{dt^2} = -w \frac{g}{\theta_e} \frac{\partial \theta_e}{\partial z} = -\frac{w}{N^2} \quad (13.10)$$

where N is the buoyancy frequency (commonly called the Brunt-Vaisala frequency). Since the meteorological equations contain the terms from which we derived this result, it should be no surprise that gravity waves occur in the atmosphere.

13.2 Scale Analysis

The full equations of motion contain terms which are not, under normally-expected conditions, all of the same order of magnitude. This is readily ascertained by performing a “scale analysis,” in which, nb., one assigns to each of the variables a scale typical of the values expected under the expected conditions. Thus, a scale analysis directed towards the synoptic scale, might assume vertical velocity w to have typical scale $0.01-0.1 \text{ m s}^{-1}$ while horizontal components have a scale of about 10 m s^{-1} : but a cloud model or a model of flow about a building or windbreak would probably assign equal scales to horizontal and vertical velocity. This is essentially because the large-scale motion is quasi-horizontal, confined to be so by the underlying surface.

Remember, we consider “synoptic scale” to mean that we are working with dependent variables which (though it is not usually explicitly recognised in the symbols used in the texts) have been averaged in the horizontal plane (or on a pressure surface) over distances large enough to average out microscale and mesoscale variations: eg.

$$\bar{w}(x, y) = \frac{1}{XY} \int_{-X/2}^{X/2} \int_{-Y/2}^{Y/2} w(x, y) dx dy \quad (13.11)$$

Thus what Holton calls vertical velocity (w) in the dynamical equations for NWP is really our \bar{w} .

Methodical scale analysis is considered by some (eg., Haltiner & Williams, 1980, p53) to have been introduced into large-scale meteorology by Charney (1948): but according to Atkinson (Dynamical Meteorology: An Introductory Selection) the method of scale analysis and the essence of its implications was largely laid out by Jeffreys (1922). The key results for synoptic scale meteorology and numerical weather prediction (Holton, 2004, p35-38; Haltiner & Williams, 1980, pp62-66) are that on the synoptic scale and at mid-latitudes, where the Rossby number

$$R_o = \frac{V}{fL} \quad (13.12)$$

(ratio of typical acceleration to the Coriolis acceleration, where V is a scale for the horizontal velocity and L is the length scale for synoptic scale motion) is small, the horizontal accelerations ($du/dt, dv/dt$) are small terms resulting from small differences between the large pressure gradient and Coriolis terms, while the vertical acceleration (dw/dt) is a small term resulting from

small imbalances in large pressure gradient and gravity terms. There is approximate hydrostatic and Geostrophic balance.

The vertical acceleration is governed by a small imbalance between two large terms, namely vertical pressure gradient and gravitational acceleration

$$\frac{dw}{dt} = - \frac{1}{\rho} \frac{\partial p}{\partial z} - g$$

$O(10) \qquad O(10)$ (13.13)

On the synoptic scale the vertical acceleration is very small. But if p and ρ are determined by observation and used to initialise a full vertical momentum equation very large spurious accelerations can result. Any convergence in the convective mass flux density $w\rho$ which arises from these spurious velocities can drive erroneous density changes (mass readjustment), as is indicated by the continuity equation

$$\frac{\partial \rho}{\partial t} = - \frac{\partial}{\partial z} \rho w + \dots$$

(13.14)

Since we can consider the accelerations to be driven by the pressure force, without the moderation of friction, in effect we have vertically-propagating sound waves (speed $c = \sqrt{(c_p/c_v) R T} \approx 20\sqrt{T}$).

Obviously the raw flow equations admit sound waves (if they did not, there could be no such thing as a sound wave). The above argues that initiation of sound waves at initialisation is possible. We would need to examine the coupled set of discretised governing equations (with their corresponding grid intervals $(\Delta x, \Delta y, \Delta z, \Delta t)$) to determine whether a model based on

the raw equations is vulnerable to a growth of such waves once present, and that is beyond our scope. Let us take it as established that amplification of sound waves can be a problem.

Sound waves are associated with density changes. As may be seen by writing the continuity equation in the form

$$\frac{1}{\rho} \frac{d\rho}{dt} = - \nabla \bullet \vec{u} \quad (13.15)$$

a specification that the velocity field be non-divergent assures a constant density (along trajectories) and would eliminate all sound waves. Though this is the usual practise in analysis of small scale motion in the atmospheric boundary-layer layer, it cannot be justified for synoptic scale motion.

Replacement of the vertical momentum equation with the hydrostatic equation is sufficient to filter out vertically propogating soundwaves, because in that case p is (apparently, ie. in so far as the model is concerned) determined only by the total overlying mass of air, and cannot be perturbed by local mass flux convergences. Lamb waves (horizontally-propogating sound) carry very little energy in the atmosphere, but are important in NWP - they place a restriction on the maximum timestep for numerical integration of the hydrostatic equations (Haltiner & Williams, 1980, p35). They have maximum amplitude at ground (Holton p176) and may be filtered out by imposing

$$\omega_0 = \left(\frac{dp}{dt} \right)_{gnd} = 0 \quad (13.16)$$

Errors in initial (horizontal) winds and pressure will similarly cause large

horizontal accelerations, exciting both sound and gravity waves. If we think of an imaginary plane surface distorting up and down in response to internal gravity waves like a (surface gravity) wave on water, it is intuitive that propagation involves:

$$\frac{\partial}{\partial t} (\nabla \cdot \vec{V}) \neq 0 \quad (13.17)$$

where \vec{V} is the horizontal velocity vector - we need a time-varying confluence/diffluence of the horizontal wind (but not necessarily any convergence/divergence $\nabla \cdot \vec{U}$ of the full vector velocity). Setting

$$\frac{\partial}{\partial t} (\nabla \cdot \vec{V}) = 0 \quad (13.18)$$

is the minimum simplification required to filter out gravity waves (Holton p177). This term does not appear explicitly in the momentum equations, but it does appear in the divergence equation which we encounter in using a vorticity-streamfunction formulation.

13.3 Filtered dynamical equations

Holton (p176) gives a set of prediction equations for the resolved flow (physics terms not included) under the minimum simplification necessary to filter out sound waves:

$$\begin{aligned} \left(\frac{\partial}{\partial t} + \vec{V} \cdot \nabla_p \right) \vec{V} + \omega \frac{\partial \vec{V}}{\partial p} + f \hat{k} \times \vec{V} &= - \nabla_p \Phi \\ \nabla_p \cdot \vec{V} + \frac{\partial \omega}{\partial p} &= 0 \\ \left(\frac{\partial}{\partial t} + \vec{V} \cdot \nabla_p \right) \frac{\partial \Phi}{\partial p} + \sigma \omega &= 0 \end{aligned} \quad (13.19)$$

where \vec{V} is the horizontal velocity¹; Φ is the “geopotential”

$$\Phi = \int_0^z g \, dz \quad (13.20)$$

$\omega = dp/dt$ is the vertical velocity in pressure (isobaric) coordinates; the grad operator ∇_p is the horizontal gradient operator (applied with p held constant: see Holton p54); and σ is a static stability parameter

$$\sigma = -\frac{1}{\rho} \frac{\partial \ln \theta}{\partial p} \quad (13.21)$$

where θ is the potential temperature.

From this set of equations, which will sustain gravity waves, one may derive by differentiation the vorticity equation

$$\begin{aligned} \frac{\partial \zeta}{\partial t} + \vec{V} \cdot \nabla (\zeta + f) + \omega \frac{\partial \zeta}{\partial p} \\ = -(\zeta + f) \nabla \cdot \vec{V} + \tilde{k} \cdot \left(\frac{\partial \vec{V}}{\partial p} \times \nabla \omega \right) \end{aligned} \quad (13.22)$$

and the divergence equation

$$\begin{aligned} \frac{\partial}{\partial t} (\nabla \cdot \vec{V}) = & -\nabla^2 \left(\Phi + \frac{\vec{V} \cdot \vec{V}}{2} \right) \\ & -\nabla \cdot \left(\tilde{k} \times \vec{V} (\zeta + f) \right) \\ & -\omega \frac{\partial}{\partial p} (\nabla \cdot \vec{V}) - \frac{\partial \vec{V}}{\partial p} \cdot \nabla \omega \end{aligned} \quad (13.23)$$

where

$$\zeta = \tilde{k} \cdot \nabla \times \vec{V} \quad (13.24)$$

¹Note that the Coriolis term may be written $f \hat{k} \times \vec{V} = fu \tilde{j} - fv \tilde{i}$, where \tilde{i} is the unit vector pointing along the east-west direction.

is the vertical component of the relative vorticity. These equations can be used to replace the two horizontal momentum equations. One advantage of this step is that if the LHS of the divergence equation is set to zero, time-dependent gravity waves are eliminated.

By the Helmholtz theorem, one may split the horizontal velocity into the sum of a nondivergent and irrotational part:

$$\vec{V} = \vec{V}_\psi + \vec{V}_e \quad (13.25)$$

where the nondivergent part satisfies (by definition)

$$\nabla \bullet \vec{V}_\psi = 0 \quad (13.26)$$

and may be expressed in terms of the gradient of a streamfunction ψ :

$$\vec{V}_\psi = \tilde{k} \times \nabla \psi = \left(-\frac{\partial \psi}{\partial y}, \frac{\partial \psi}{\partial x} \right) \quad (13.27)$$

With this decomposition, and after a scale analysis for midlatitude synoptic scale motion - which shows, among other things, that

$$|\vec{V}_\psi| \gg |\vec{V}_e| \quad (13.28)$$

- one can obtain from the divergence equation the non-linear balance equation:

$$\nabla^2 \left(\Phi + \frac{1}{2} (\nabla \psi \bullet \nabla \psi) \right) = \nabla \bullet \left((f + \nabla^2 \psi) \nabla \psi \right) \quad (13.29)$$

which relates the non-divergent part (ie. essentially the horizontal wind) of the wind field via the streamfunction to the “mass” field (the Geopotential:

since we are using the hydrostatic equation, the height of a pressure surface depends only on the density field). The linear balance equation is a further simplification:

$$\nabla^2 \Phi = \nabla \cdot (f \nabla \psi) \quad (13.30)$$

13.3.1 Quasi-geostrophic system of equations

Key attributes: Valid over a restricted range of latitude, with horizontal wind V assumed geostrophic (with $f = f_0$) except where its divergence $D = \nabla_p \cdot \vec{V}$ appears in the vorticity equation.

By performing the scale analysis for mid-latitude synoptic scale motion on the governing equation for the vertical component of the vorticity expressed in isobaric coordinates, using the β -plane approximation $f = f_0 + \beta y$, Holton (p180) obtains for synoptic scale motion:

$$\frac{\partial}{\partial t} (\nabla^2 \psi) = - \vec{V}_\psi \cdot \nabla (\nabla^2 \psi + f) + f_0 \frac{\partial \omega}{\partial p} \quad (13.31)$$

The streamfunction is now related to the geopotential by

$$\psi = \frac{\Phi}{f_0} \quad (13.32)$$

so the horizontal velocity is now simply the geostrophic wind:

$$\vec{V}_\psi = \frac{\tilde{k} \times \nabla \Phi}{f_0} \quad (13.33)$$

and the streamfunction is related to the geostrophic (relative) vorticity by the Poisson equation:

$$\nabla^2 \psi = \zeta \quad (13.34)$$

The hydrostatic thermodynamic energy equation is:

$$\frac{\partial}{\partial t} \left(\frac{\partial \psi}{\partial p} \right) = - \vec{V}_\psi \cdot \nabla \left(\frac{\partial \psi}{\partial p} \right) - \frac{\sigma}{f_0} \omega \quad (13.35)$$

According to Holton (p130 but note this was published 1979) this “quasi-geostrophic system,” equations () and others derived from them (eg. the diagnostic “omega-equation” for the vertical velocity, obtained by eliminating the time derivatives from the equations(,)), constitute the core of modern dynamic meteorology, a simple “noise-immune” system. However this model must be modified if low latitudes are to be included (because at low latitude the streamfunction cannot be determined from the geostrophic vorticity field). According to HW (p69), normally when more accuracy is required than can be obtained with the quasi-geostrophic equations, the primitive equations are used.

13.3.2 Barotropic vorticity equation

Even by 1912 it had been established (or suggested) by Dines that confluence $\nabla_p \cdot \vec{V} < 0$ in the lower troposphere is largely compensated by diffluence in the upper troposphere (and vice versa), such that

$$|\nabla \cdot \vec{V}_p| \gg \frac{1}{p_0} \int_{p=0}^{p_0} \nabla \cdot \vec{V} dp \quad (13.36)$$

This implies that at some midtropospheric altitude $\nabla_p \cdot \vec{V}$ changes sign, ie. passes through zero.

Assuming, then, that there is a *level of non-divergence*, say at 500 mb, we have at that level $\partial\omega/\partial p = 0$, and the Quasi-Geostrophic Vorticity equation

reduces to:

$$\frac{\partial}{\partial t} (\nabla^2 \psi) = - \vec{V}_\psi \bullet \nabla (\nabla^2 \psi + f) \quad (13.37)$$

which may be written in shorthand as:

$$\frac{d_h}{dt} (\zeta + f) = 0 \quad (13.38)$$

According to Holton, this is “a useful approximate forecast equation.” It gives the evolution of the flow at a single level, the level of non-divergence.

Part IV
References (etc).

Author

These notes prepared by Dr. John D. Wilson (Department of Earth & Atmospheric Sciences, University of Alberta, Edmonton, Canada) for use by students enrolled in *EAS471* (“Atmospheric Modelling”).

References

- Ames, W.F. 1977. *Numerical Methods for Partial Differential Equations*. Academic Press.
- Anthes, R.A. 1977. A Cumulus Parameterization Scheme Utilizing a One-Dimensional Model. *Monthly Weather Review*, **105**, 270–286.
- Churchill, R.V. 1969. *Fourier Series and Boundary Value Problems (2nd Edition)*. McGraw-Hill.
- Churchill, R.V., & Brown, J.W. 1990. *Complex Variables and Applications (5th Edition)*. McGraw-Hill.
- Corrsin, S. 1974. Limitations of Gradient Transport Models. *Adv. Geophys.*, **18A**, 25–60.
- Delage, Y. 1974. A Numerical Study of the Nocturnal Atmospheric Boundary Layer. *Quart. J. R. Met. Soc.*, **100**, 351–364.
- Emmanuel, K.A. 1991. A Scheme for Representing Cumulus Convection in Large-Scale Models. *J. Atmos. Sci.*, **48**, 2313–2335.

- Emmanuel, K.A., & Raymond, D.J. 1992. Report from a Workshop on Cumulus Parameterization. *Bull. American Meteorol. Soc.*, **73**, 318–325.
- Emmanuel, K.A., & Raymond, D.J. 1993. *The Representation of Cumulus Convection in Numerical Models*. Meteorological Monographs, vol. 24, no. 46. American Meteorol. Soc.
- Haltiner, G.J. 1971. *Numerical Weather Prediction*. J. Wiley.
- Haltiner, G.J., & Williams, R.T. 1980. *Numerical Prediction and Dynamic Meteorology*. J. Wiley.
- Holton, J.R. 2004. *An Introduction to Dynamic Meteorology (Fourth Edition)*. Elsevier.
- Kuo, H.L. 1965. On Formation and Intensification of Tropical Cyclones Through Latent Heat Release by Cumulus Convection. *J. Atmos. Sci.*, **22**, 40–63.
- Kuo, H.L. 1974. Further Studies of the Parameterization of the Influence of Cumulus Convection on Large-Scale Flow. *J. Atmos. Sci.*, **31**, 1232–1240.
- Liou, K.N. 2002. *An Introduction to Atmospheric Radiation*. 2 edn. Academic Press.
- Mitchell, J.F.B. 2004. Can We Believe Predictions of Climate Change? *Q. J. R. Meteorol. Soc.*, **130**, 2341–2360.

- Patankar, S.V. 1980. *Numerical Heat Transfer and Fluid Flow*. Hemisphere Publ. (McGraw-Hill Book Co.), New York (ISBN 0-07-0487404).
- Stephens, G.L. 1992. *Parameterization of Subgrid Cloud Radiative Transport Processes for Atmospheric Prediction Models: Present, Past and Future*. J. Wiley. Pages 203–230.
- Taylor, G.I. 1921. Diffusion by Continuous Movements. *Proc. London Math. Soc., Ser. 2*, **A20**, 196–211.
- Tennekes, H. 1979. The Exponential Lagrangian Correlation Function and Turbulent Diffusion in the Inertial Subrange. *Atmos. Env.*, **13**, 1565–1567.
- Thomson, D.J. 1987. Criteria for the selection of stochastic models of particle trajectories in turbulent flows. *J. Fluid Mech.*, **180**, 529–556.
- Wilson, J.D. 2002. Representing Drag on Unresolved Terrain as a Distributed Momentum Sink. *J. Atmos. Sci.*, **59**, 1629–1637.
- Wilson, J.D., & Flesch, T.K. 1999. Wind and Tree Sway in Forest Cutblocks, III: A Windflow Model to Diagnose Spatial Variation. *Agric. Forest Meteo.*, **93**, 259–282.
- Wilson, J.D., Thurtell, G.W., & Kidd, G.E. 1981. Numerical Simulation of Particle Trajectories in Inhomogeneous Turbulence. III. Comparison of Predictions with Experimental Data for the Atmospheric Surface-Layer. *Boundary-Layer Meteorol.*, **21**, 443–463.