Atmospheric and Oceanic Fluid Dynamics

Fundamentals and Large-Scale Circulation

Geoffrey K. Vallis

CAMBRIDGE

Contents

An asterisk indicates more advanced material that may be omitted on a first reading. A dagger indicates material that is still a topic of research or that is not settled.

Preface *page* **xix**
Notation xxiv

Part I FUNDAMENTALS OF GEOPHYSICAL FLUID DYNAMICS 1

1 Equations of Motion 3
1.1 Time Derivatives for Fluids 3
1.1.1 Field and material viewpoints 3
1.1.2 The material derivative of a fluid property 4
1.1.3 Material derivative of a volume 6
1.2 The Mass Continuity Equation 7
1.2.1 An Eulerian derivation 7
1.2.2 Mass continuity via the material derivative 9
1.2.3 A general continuity equation 11
1.3 The Momentum Equation 11
1.3.1 Advection 12
1.3.2 The pressure force 12
1.3.3 Viscosity and diffusion 13
1.3.4 Hydrostatic balance 13
1.4 The Equation of State 14
1.5 Thermodynamic Relations 16
1.5.1 A few fundamentals 16
1.5.2 Various thermodynamic relations 18
1.6 Thermodynamic Equations for Fluids 22

- 1.6.1 Thermodynamic equation for an ideal gas 23
- 1.6.2 * Thermodynamic equation for liquids 26
- 1.7 * More Thermodynamics of Liquids 31
  - 1.7.1 Potential temperature, potential density and entropy 31
  - 1.7.2 * Thermodynamic properties of seawater 33
- 1.8 Sound Waves 37
- 1.9 Compressible and Incompressible Flow 38
  - 1.9.1 Constant density fluids 38
  - 1.9.2 Incompressible flows 39
- 1.10 The Energy Budget 40
  - 1.10.1 Constant density fluid 40
  - 1.10.2 Variable density fluids 42
  - 1.10.3 Viscous effects 43
- 1.11 An Introduction to Non-Dimensionalization and Scaling 43
  - 1.11.1 The Reynolds number 44

2 Effects of Rotation and Stratification 51

- 2.1 Equations in a Rotating Frame 51
  - 2.1.1 Rate of change of a vector 52
  - 2.1.2 Velocity and acceleration in a rotating frame 53
  - 2.1.3 Momentum equation in a rotating frame 54
  - 2.1.4 Mass and tracer conservation in a rotating frame 54
- 2.2 Equations of Motion in Spherical Coordinates 55
  - 2.2.1 * The centrifugal force and spherical coordinates 55
  - 2.2.2 Some identities in spherical coordinates 57
  - 2.2.3 Equations of motion 60
  - 2.2.4 The primitive equations 61
  - 2.2.5 Primitive equations in vector form 62
  - 2.2.6 The vector invariant form of the momentum equation 63
  - 2.2.7 Angular momentum 64
- 2.3 Cartesian Approximations: The Tangent Plane 66
  - 2.3.1 The f-plane 66
  - 2.3.2 The beta-plane approximation 67
- 2.4 The Boussinesq Approximation 67
  - 2.4.1 Variation of density in the ocean 68
  - 2.4.2 The Boussinesq equations 68
  - 2.4.3 Energetics of the Boussinesq system 72
- 2.5 The Anelastic Approximation 73
  - 2.5.1 Preliminaries 73
  - 2.5.2 The momentum equation 74
  - 2.5.3 Mass conservation 75
  - 2.5.4 Thermodynamic equation 76
  - 2.5.5 * Energetics of the anelastic equations 76
- 2.6 Changing Vertical Coordinate 77
  - 2.6.1 General relations 77
  - 2.6.2 Pressure coordinates 78

2.6.3 Log-pressure coordinates 80
2.7 Scaling for Hydrostatic Balance 80
2.7.1 Preliminaries 80
2.7.2 Scaling and the aspect ratio 81
2.7.3 * Effects of stratification on hydrostatic balance 82
2.7.4 Hydrostasy in the ocean and atmosphere 84
2.8 Geostrophic and Thermal Wind Balance 85
2.8.1 The Rossby number 85
2.8.2 Geostrophic balance 86
2.8.3 Taylor–Proudman effect 88
2.8.4 Thermal wind balance 89
2.8.5 * Effects of rotation on hydrostatic balance 91
2.9 Static Instability and the Parcel Method 91
2.9.1 A simple special case: a density-conserving fluid 92
2.9.2 The general case: using potential density 93
2.9.3 Lapse rates in dry and moist atmospheres 95
2.10 Gravity Waves 98
2.10.1 Gravity waves and convection in a Boussinesq fluid 98
2.11 * Acoustic-Gravity Waves in an Ideal Gas 100
2.11.1 Interpretation 101
2.12 The Ekman Layer 104
2.12.1 Equations of motion and scaling 105
2.12.2 Integral properties of the Ekman layer 107
2.12.3 Explicit solutions. I: a bottom boundary layer 109
2.12.4 Explicit solutions. II: the upper ocean 112
2.12.5 Observations of the Ekman layer 113
2.12.6 * Frictional parameterization of the Ekman layer 114

3 Shallow Water Systems and Isentropic Coordinates 123
3.1 Dynamics of a Single, Shallow Layer 123
3.1.1 Momentum equations 124
3.1.2 Mass continuity equation 125
3.1.3 A rigid lid 127
3.1.4 Stretching and the vertical velocity 128
3.1.5 Analogy with compressible flow 129
3.2 Reduced Gravity Equations 129
3.2.1 Pressure gradient in the active layer 130
3.3 Multi-Layer Shallow Water Equations 131
3.3.1 Reduced-gravity multi-layer equation 133
3.4 Geostrophic Balance and Thermal wind 134
3.5 Form Drag 135
3.6 Conservation Properties of Shallow Water Systems 136
3.6.1 Potential vorticity: a material invariant 136
3.6.2 Energy conservation: an integral invariant 139
3.7 Shallow Water Waves 140
3.7.1 Non-rotating shallow water waves 140

3.7.2 Rotating shallow water (Poincaré) waves 141
3.7.3 Kelvin waves 143
3.8 Geostrophic Adjustment 144
3.8.1 Non-rotating flow 145
3.8.2 Rotating flow 146
3.8.3 * Energetics of adjustment 148
3.8.4 * General initial conditions 149
3.8.5 A variational perspective 151
3.9 Isentropic Coordinates 152
3.9.1 A hydrostatic Boussinesq fluid 152
3.9.2 A hydrostatic ideal gas 153
3.9.3 Analogy to shallow water equations 154
3.10 Available Potential Energy 155
3.10.1 A Boussinesq fluid 156
3.10.2 An ideal gas 158
3.10.3 Use, interpretation, and the atmosphere and ocean 159

4 Vorticity and Potential Vorticity 163
4.1 Vorticity and Circulation 163
4.1.1 Preliminaries 163
4.1.2 Simple axisymmetric examples 164
4.2 The Vorticity Equation 165
4.2.1 Two-dimensional flow 167
4.3 Vorticity and Circulation Theorems 168
4.3.1 The ‘frozen-in’ property of vorticity 168
4.3.2 Kelvin’s circulation theorem 171
4.3.3 Baroclinic flow and the solenoidal term 173
4.3.4 Circulation in a rotating frame 173
4.3.5 The circulation theorem for hydrostatic flow 174
4.4 Vorticity Equation in a Rotating Frame 175
4.4.1 The circulation theorem and vortex tilting 175
4.4.2 The vertical component of the vorticity equation 176
4.5 Potential Vorticity Conservation 178
4.5.1 PV conservation from the circulation theorem 178
4.5.2 PV conservation from the frozen-in property 180
4.5.3 PV conservation — an algebraic derivation 182
4.5.4 Effects of salinity and moisture 183
4.5.5 Effects of rotation, and summary remarks 183
4.6 * Potential Vorticity in the Shallow Water System 184
4.6.1 Using Kelvin’s theorem 184
4.6.2 Using an appropriate scalar field 185
4.7 Potential Vorticity in Approximate, Stratified Models 186
4.7.1 The Boussinesq equations 186
4.7.2 The hydrostatic equations 187
4.7.3 Potential vorticity on isentropic surfaces 187
4.8 * The Impermeability of Isentropes to Potential Vorticity 188

4.8.1 Interpretation and application 190

5 Simplified Equations for Ocean and Atmosphere 197

5.1 Geostrophic Scaling 198
5.1.1 Scaling in the shallow water equations 198
5.1.2 Geostrophic scaling in the stratified equations 200
5.2 The Planetary-Geostrophic Equations 203
5.2.1 Using the shallow water equations 203
5.2.2 Planetary-geostrophic equations for stratified flow 205
5.3 The Shallow Water Quasi-Geostrophic Equations 207
5.3.1 Single-layer shallow water quasi-geostrophic equations 207
5.3.2 Two-layer and multi-layer quasi-geostrophic systems 211
5.3.3 † Non-asymptotic and intermediate models 214
5.4 The Continuously Stratified Quasi-Geostrophic System 215
5.4.1 Scaling and assumptions 215
5.4.2 Asymptotics 216
5.4.3 Buoyancy advection at the surface 219
5.4.4 Quasi-geostrophy in pressure coordinates 220
5.4.5 The two-level quasi-geostrophic system 221
5.5 * Quasi-geostrophy and Ertel Potential Vorticity 224
5.5.1 * Using height coordinates 224
5.5.2 Using isentropic coordinates 225
5.6 * Energetics of Quasi-Geostrophy 226
5.6.1 Conversion between APE and KE 227
5.6.2 Energetics of two-layer flows 228
5.6.3 Enstrophy conservation 229
5.7 Rossby Waves 229
5.7.1 Waves in a single layer 229
5.7.2 Rossby waves in two layers 232
5.8 * Rossby Waves in Stratified Quasi-Geostrophic Flow 234
5.8.1 Setting up the problem 234
5.8.2 Wave motion 235
Appendix: Wave Kinematics, Group Velocity and Phase Speed 236
5.A.1 Kinematics and definitions 236
5.A.2 Wave propagation 237
5.A.3 Meaning of group velocity 239

Part II INSTABILITIES, WAVE–MEAN FLOW INTERACTION AND TURBULENCE 245

6 Barotropic and Baroclinic Instability 247
6.1 Kelvin–Helmholtz Instability 248
6.2 Instability of Parallel Shear Flow 250
6.2.1 Piecewise linear flows 251

- 6.2.2 Kelvin–Helmholtz instability, revisited 253
- 6.2.3 Edge waves 254
- 6.2.4 Interacting edge waves producing instability 254
- 6.3 Necessary Conditions for Instability 258
  - 6.3.1 Rayleigh's criterion 258
  - 6.3.2 Fjørtoft's criterion 260
- 6.4 Baroclinic Instability 261
  - 6.4.1 A physical picture 261
  - 6.4.2 Linearized quasi-geostrophic equations 263
  - 6.4.3 Necessary conditions for baroclinic instability 264
- 6.5 The Eady Problem 265
  - 6.5.1 The linearized problem 266
  - 6.5.2 Atmospheric and oceanic parameters 268
- 6.6 Two-Layer Baroclinic Instability 271
  - 6.6.1 Posing the problem 271
  - 6.6.2 The solution 272
- 6.7 An Informal View of the Mechanism of Baroclinic Instability 277
  - 6.7.1 The two-layer model 278
  - 6.7.2 Interacting edge waves in the Eady problem 280
- 6.8 * The Energetics of Linear Baroclinic Instability 282
- 6.9 * Beta, Shear and Stratification in a Continuous Model 284
  - 6.9.1 Scaling arguments for growth rates, scales and depth 284
  - 6.9.2 Some numerical calculations 287

7 Wave–Mean Flow Interaction 295

- 7.1 Quasi-geostrophic Preliminaries 296
  - 7.1.1 Potential vorticity flux in the linear equations 297
- 7.2 The Eliassen–Palm Flux 298
  - 7.2.1 The Eliassen–Palm relation 299
  - 7.2.2 The group velocity property 300
  - 7.2.3 * The orthogonality of modes 302
- 7.3 The Transformed Eulerian Mean 304
  - 7.3.1 Quasi-geostrophic form 304
  - 7.3.2 The TEM in isentropic coordinates 306
  - 7.3.3 Residual and thickness-weighted circulation 307
  - 7.3.4 * The TEM in the primitive equations 309
- 7.4 The Non-acceleration Result 314
  - 7.4.1 A derivation from the potential vorticity equation 314
  - 7.4.2 Using TEM to given the non-acceleration result 316
  - 7.4.3 The EP flux and form drag 317
- 7.5 Influence of Eddies on the Mean Flow in the Eady Problem 319
  - 7.5.1 Formulation 319
  - 7.5.2 Solution 321
  - 7.5.3 The two-level problem 324
- 7.6 * Necessary Conditions for Instability 324
  - 7.6.1 Stability conditions from pseudomomentum conservation 325

7.6.2 Inclusion of boundary terms 325
7.7 * Necessary Conditions for Instability: Use of Pseudoenergy 327
7.7.1 Two-dimensional flow 327
7.7.2 * Stratified quasi-geostrophic flow 330
7.7.3 * Applications to baroclinic instability 331

8 Basic Theory of Incompressible Turbulence 337
8.1 The Fundamental Problem of Turbulence 338
8.1.1 The closure problem 338
8.1.2 Triad interactions in turbulence 339
8.2 The Kolmogorov Theory 341
8.2.1 The physical picture 341
8.2.2 Inertial-range theory 342
8.2.3 * Another expression of the inertial-range scaling argument 348
8.2.4 A final note on our assumptions 349
8.3 Two-Dimensional Turbulence 349
8.3.1 Energy and enstrophy transfer 351
8.3.2 Inertial ranges in two-dimensional turbulence 354
8.3.3 † More about the phenomenology 358
8.3.4 Numerical illustrations 360
8.4 Predictability of Turbulence 361
8.4.1 Low-dimensional chaos and unpredictability 362
8.4.2 * Predictability of a turbulent flow 363
8.4.3 Implications and weather predictability 365
8.5 * Spectra of Passive Tracers 366
8.5.1 Examples of tracer spectra 368

9 Geostrophic Turbulence and Baroclinic Eddies 377
9.1 Effects of Differential Rotation 377
9.1.1 The wave–turbulence cross-over 378
9.1.2 Generation of zonal flows and jets 380
9.1.3 † Joint effect of β and friction 382
9.2 Stratified Geostrophic Turbulence 384
9.2.1 An analogue to two-dimensional flow 384
9.2.2 Two-layer geostrophic turbulence 385
9.2.3 Phenomenology of two-layer turbulence 388
9.3 † A Scaling Theory for Geostrophic Turbulence 391
9.3.1 Preliminaries 392
9.3.2 Scaling properties 393
9.3.3 The halting scale and the β-effect 395
9.4 † Phenomenology of Baroclinic Eddies in the Atmosphere and Ocean 395
9.4.1 The magnitude and scale of baroclinic eddies 396
9.4.2 Baroclinic eddies and their lifecycle in the atmosphere 397
9.4.3 Baroclinic eddies and their lifecycle in the ocean 400

10 Turbulent Diffusion and Eddy Transport 407
10.1 Diffusive Transport 408
  10.1.1 An explicit example 409
10.2 Turbulent Diffusion 409
  10.2.1 Simple theory 409
  10.2.2 * An anisotropic generalization 413
  10.2.3 Discussion 415
10.3 Two-Particle Diffusivity 415
  10.3.1 Large particle separation 416
  10.3.2 Separation within the inertial range 417
10.4 Mixing Length Theory 419
  10.4.1 Requirements for turbulent diffusion 421
  10.4.2 A macroscopic perspective 422
10.5 Homogenization of a Scalar that is Advected and Diffused 423
  10.5.1 Non-existence of extrema 423
  10.5.2 Homogenization in two-dimensional flow 424
10.6 † Transport by Baroclinic Eddies 425
  10.6.1 Symmetric and antisymmetric diffusivity tensors 426
  10.6.2 * Diffusion with the symmetric tensor 426
  10.6.3 * Skew diffusion 427
  10.6.4 The story so far 429
10.7 † Eddy Diffusion in the Atmosphere and Ocean 430
  10.7.1 Preliminaries 430
  10.7.2 Magnitude of the eddy diffusivity 430
  10.7.3 * Structure: the symmetric transport tensor 432
  10.7.4 * Structure: the antisymmetric transport tensor 435
  10.7.5 Examples 437
10.8 † Thickness Diffusion 440
  10.8.1 Equations of motion 440
  10.8.2 Diffusive thickness transport 442
10.9 † Eddy Transport and the Transformed Eulerian Mean 443
  10.9.1 Potential vorticity diffusion 443

Part III LARGE-SCALE ATMOSPHERIC CIRCULATION 449

11 The Overturning Circulation: Hadley and Ferrel Cells 451
11.1 Basic Features of the Atmosphere 452
  11.1.1 The radiative equilibrium distribution 452
  11.1.2 Observed wind and temperature fields 453
  11.1.3 Meridional overturning circulation 456
  11.1.4 Summary 457
11.2 A Steady Model of the Hadley Cell 457
  11.2.1 Assumptions 457
  11.2.2 Dynamics 458

11.2.3 Thermodynamics 460
11.2.4 Zonal wind 462
11.2.5 Properties of solution 463
11.2.6 Strength of the circulation 464
11.2.7 † Effects of moisture 465
11.2.8 The radiative equilibrium solution 466
11.3 A Shallow Water Model of the Hadley Cell 468
11.3.1 Momentum balance 468
11.3.2 Thermodynamic balance 468
11.4 † Asymmetry Around the Equator 469
11.5 Eddies, Viscosity and the Hadley Cell 473
11.5.1 Qualitative considerations 473
11.5.2 An idealized eddy-driven model 474
11.6 The Hadley Cell: Summary and Numerical Solutions 477
11.7 The Ferrel Cell 480

12 Zonally Averaged Mid-Latitude Atmospheric Circulation 485
12.1 Surface Westerlies and the Maintenance of a Barotropic Jet 486
12.1.1 Observations and motivation 486
12.1.2 The mechanism of jet production 487
12.1.3 A numerical example 495
12.2 Layered Models of the Mid-latitude Circulation 496
12.2.1 A single-layer model 497
12.2.2 A two-layer model 503
12.2.3 Dynamics of the two-layer model 507
12.3 † Eddy Fluxes and an Example of a Closed Model 513
12.3.1 Equations for a closed model 513
12.3.2 * Eddy fluxes and necessary conditions for instability 514
12.4 A Stratified Model and the Real Atmosphere 516
12.4.1 Potential vorticity and its fluxes 516
12.4.2 Overturning circulation 522
12.5 † The Tropopause and the Stratification of the Atmosphere 522
12.5.1 A radiative–convective model 526
12.5.2 Radiative and dynamical constraints 528
12.6 † Baroclinic eddies and Potential Vorticity Transport 529
12.6.1 A linear argument 530
12.6.2 Mixing potential vorticity and baroclinic adjustment 530
12.6.3 Diffusive transport of potential vorticity 532
12.7 † Extratropical Convection and the Ventilated Troposphere 534
Appendix: TEM for the Primitive Equations in Spherical Coordinates 536

13 Planetary Waves and the Stratosphere 541
13.1 Forced and Stationary Rossby Waves 542
13.1.1 A simple one-layer case 542
13.1.2 Application to Earth's atmosphere 543

13.1.3 * One-dimensional Rossby wave trains 545
13.1.4 The adequacy of linear theory 548
13.2 * Meridional Propagation and Dispersion 549
13.2.1 Ray tracing 549
13.2.2 Rossby waves and Rossby rays 550
13.2.3 Application to an idealized atmosphere 553
13.3 * Vertical Propagation of Rossby Waves in a Stratified Medium 554
13.3.1 Model formulation 554
13.3.2 Model solution 555
13.3.3 Properties of the solution 559
13.4 * Effects of Thermal Forcing 560
13.4.1 Thermodynamic balances 561
13.4.2 Properties of the solution 562
13.4.3 Numerical solutions 563
13.5 Stratospheric Dynamics 566
13.5.1 A descriptive overview 566
13.5.2 † Dynamics of the overturning circulation 569
13.5.3 † The polar vortex and the quasi-horizontal circulation 575

Part IV LARGE-SCALE OCEANIC CIRCULATION 581

14 Wind-Driven Gyres 583

14.1 The Depth Integrated Wind-Driven Circulation 585
14.1.1 The Stommel model 586
14.1.2 Alternative formulations 587
14.1.3 Approximate solution of Stommel model 589
14.2 Using Viscosity Instead of Drag 593
14.3 Zonal Boundary Layers 597
14.4 * The Nonlinear Problem 599
14.4.1 A perturbative approach 600
14.4.2 A numerical approach 600
14.5 * Inertial Solutions 601
14.5.1 Roles of friction and inertia 603
14.5.2 Attempting an inertial western boundary solution 604
14.5.3 A fully inertial approach: the Fofonoff model 606
14.6 Topographic Effects on Western Boundary Currents 608
14.6.1 Homogeneous model 608
14.6.2 Advective dynamics 609
14.6.3 Bottom pressure stress and form drag 611
14.7 * Vertical Structure of the Wind-Driven Circulation 613
14.7.1 A two-layer quasi-geostrophic Model 613
14.7.2 The functional relationship between ψ and q 616
14.8 * A Model with Continuous Stratification 619
14.8.1 Depth of the wind's influence 619
14.8.2 The complete solution 620

15 **The Buoyancy-Driven Ocean Circulation** **627**
15.1 Sideways Convection 629
15.1.1 Two-dimensional convection 630
15.1.2 † Phenomenology of the overturning circulation 633
15.2 The Maintenance of Sideways Convection 634
15.2.1 The energy budget 635
15.2.2 Conditions for maintaining a thermally-driven circulation 635
15.2.3 Surface fluxes and non-turbulent flow at small diffusivities 637
15.2.4 The importance of mechanical forcing 639
15.3 Simple Box Models 640
15.3.1 A two-box model 640
15.3.2 * More boxes 644
15.4 A Laboratory Model of the Abyssal Circulation 646
15.4.1 Set-up of the laboratory model 646
15.4.2 Dynamics of flow in the tank 647
15.5 A Model for Oceanic Abyssal Flow 650
15.5.1 Completing the solution 652
15.5.2 Application to the ocean 653
15.5.3 A two-hemisphere model 655
15.6 * A Shallow Water Model of the Abyssal Flow 656
15.6.1 Potential vorticity and poleward interior flow 657
15.6.2 The solution 658
15.7 Scaling for the Buoyancy-Driven Circulation 659
15.7.1 Summary remarks on the Stommel–Arons model 661

16 **The Wind- and Buoyancy-Driven Ocean Circulation** **667**
16.1 The Main Thermocline: an Introduction 667
16.1.1 A simple kinematic model 668
16.2 Scaling and Simple Dynamics of the Main Thermocline 670
16.2.1 An advective scale 671
16.2.2 A diffusive scale 672
16.2.3 Summary of the physical picture 673
16.3 The Internal Thermocline 674
16.3.1 The M equation 674
16.3.2 * Boundary-layer analysis 676
16.4 The Ventilated Thermocline 681
16.4.1 A reduced gravity, single-layer model 682
16.4.2 A two-layer model 683
16.4.3 The shadow zone 686
16.4.4 † The western pool 688
16.5 † A Model of Deep Wind-Driven Overturning 691
16.5.1 A single-hemisphere model 693
16.5.2 A cross-equatorial wind-driven deep circulation 697
16.6 † Flow in a Channel and the Antarctic Circumpolar Current 700
16.6.1 Steady and eddying flow 701

16.6.2 Vertically integrated momentum balance 702
16.6.3 Form drag and baroclinic eddies 703
16.6.4 † An idealized adiabatic model 708
16.6.5 Form stress and Ekman stress at the ocean bottom 709
16.6.6 Differences between gyres and channels 710
Appendix: Miscellaneous Relationships in a Layered Model 710
16.A.1 Hydrostatic balance 711
16.A.2 Geostrophic and thermal wind balance 711
16.A.3 Explicit cases 712

References 717

Index 738

We must be ignorant of much, if we would know anything.
Cardinal John Newman (1801–1890).

Preface

THIS IS A BOOK on the fluid dynamics of the atmosphere and ocean, with an emphasis on the fundamentals and on the large-scale circulation, the latter meaning flows from the scale of the first deformation radius (a few tens of kilometres in the ocean, several hundred kilometres in the atmosphere) to the global scale. The book is primarily a textbook; it is designed to be accessible to students and could be used as a text for graduate courses. It may be also useful as an introduction to the field for scientists in other areas and as a reference for researchers in the field, and some aspects of the book have the flavour of a research monograph.

Atmospheric and oceanic fluid dynamics (AOFD) is fascinating field, and simultaneously both pure and applied. It is a pure field because it is intimately tied to some of the most fundamental and unsolved problems in fluid dynamics — problems in turbulence and wave–mean flow interaction, problems in chaos and predictability, and problems in the general circulation itself. Yet it is applied because the climate and weather so profoundly affect the human condition, and so a great deal of effort goes into making predictions — indeed the practice of weather forecasting is a remarkable example of a successful applied science, in spite of the natural limitations to predictability that are now reasonably well understood. The field is plainly important, for we live in the atmosphere and the ocean covers about two-thirds of the Earth. It is also very broad, encompassing such diverse topics as the general circulation, gyres, boundary layers, waves, convection and turbulence. My goal in this book is present a coherent selection of these topics, concentrating on the foundations but without shying away from the boundaries of active areas of research — for a book that limits itself to what is absolutely settled would, I think, be rather dry, a quality best reserved for martinis and humour.

AOFD is closely related to the field of *geophysical fluid dynamics* (GFD). The latter can be, depending on one's point of view, both a larger and a smaller field than the former. It is larger because GFD, in its broadest meaning, includes not just the fluid dynamics of the Earth's atmosphere and ocean, but also the fluid dynamics of such things as the Earth's interior, volcanoes, lava flows and planetary atmospheres; it is the fluid mechanics of all

Part I

FUNDAMENTALS OF GEOPHYSICAL FLUID DYNAMICS

If a body is moving in any direction, there is a force, arising from the Earth's rotation, which always deflects it to the right in the northern hemisphere, and to the left in the southern.

William Ferrel, *The influence of the Earth's rotation upon the relative motion of bodies near its surface,* 1858.

CHAPTER
TWO

Effects of Rotation and Stratification

THE ATMOSPHERE AND OCEAN are shallow layers of fluid on a sphere in that their thickness or depth is much less than their horizontal extent. Furthermore, their motion is strongly influenced by two effects: rotation and stratification, the latter meaning that there is a mean vertical gradient of (potential) density that is often large compared with the horizontal gradient. Here we consider how the equations of motion are affected by these effects. First, we consider some elementary effects of rotation on a fluid and derive the Coriolis and centrifugal forces, and then we write down the equations of motion appropriate for motion on a sphere. Then we discuss some approximations to the equations of motion that are appropriate for large-scale flow in the ocean and atmosphere, in particular the hydrostatic and geostrophic approximations. Following this we discuss gravity waves, a particular kind of wave motion that is enabled by the presence of stratification, and finally we talk about how rotation leads to the production of certain types of boundary layers — Ekman layers — in rotating fluids.

2.1 THE EQUATIONS OF MOTION IN A ROTATING FRAME OF REFERENCE

Newton's second law of motion, that the acceleration on a body is proportional to the imposed force divided by the body's mass, applies in so-called inertial frames of reference. The Earth rotates with a period of almost 24 hours (23h 56m) relative to the distant stars, and the surface of the Earth manifestly is not, in that sense, an inertial frame. Nevertheless, because the surface of the Earth is moving (in fact at speeds of up to a few hundreds of metres per second) it is very convenient to describe the flow relative to the Earth's surface, rather than in some inertial frame. This necessitates recasting the equations into a form that is appropriate for a rotating frame of reference, and that is the subject of this section.

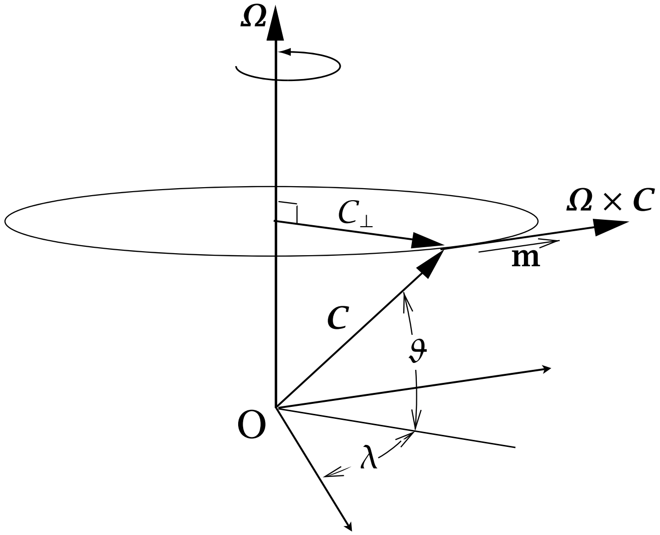


Fig. 2.1 A vector C rotating at an angular velocity Ω. It appears to be a constant vector in the rotating frame, whereas in the inertial frame it evolves according to $(\mathrm{d}C/\mathrm{d}t)_I = \Omega \times C$.

2.1.1 Rate of change of a vector

Consider first a vector $\boldsymbol{C}$ of constant length rotating relative to an inertial frame at a constant angular velocity $\boldsymbol{\Omega}$. Then, in a frame rotating with that same angular velocity it appears stationary and constant. If in a small interval of time δt the vector $\boldsymbol{C}$ rotates through a small angle $\delta\lambda$ then the change in $\boldsymbol{C}$, as perceived in the inertial frame, is given by (see Fig. 2.1)

$$\delta \boldsymbol{C} = |\boldsymbol{C}| \cos\vartheta \, \delta\lambda \, \boldsymbol{m}, \tag{2.1}$$

where the vector $\boldsymbol{m}$ is the unit vector in the direction of change of $\boldsymbol{C}$, which is perpendicular to both $\boldsymbol{C}$ and $\boldsymbol{\Omega}$. But the rate of change of the angle λ is just, by definition, the angular velocity so that $\delta\lambda = |\boldsymbol{\Omega}|\delta t$ and

$$\delta \boldsymbol{C} = |\boldsymbol{C}||\boldsymbol{\Omega}| \sin\hat{\vartheta} \, \boldsymbol{m} \, \delta t = \boldsymbol{\Omega} \times \boldsymbol{C} \, \delta t. \tag{2.2}$$

using the definition of the vector cross product, where $\hat{\vartheta} = (\pi/2 - \vartheta)$ is the angle between $\boldsymbol{\Omega}$ and $\boldsymbol{C}$. Thus

$$\left(\frac{\mathrm{d}\boldsymbol{C}}{\mathrm{d}t}\right)_I = \boldsymbol{\Omega} \times \boldsymbol{C}, \tag{2.3}$$

where the left-hand side is the rate of change of $\boldsymbol{C}$ as perceived in the inertial frame.

Now consider a vector $\boldsymbol{B}$ that changes in the inertial frame. In a small time δt the change in $\boldsymbol{B}$ as seen in the rotating frame is related to the change seen in the inertial frame by

$$(\delta \boldsymbol{B})_I = (\delta \boldsymbol{B})_R + (\delta \boldsymbol{B})_{rot}, \tag{2.4}$$

where the terms are, respectively, the change seen in the inertial frame, the change due to the vector itself changing as measured in the rotating frame, and the change due to the rotation. Using (2.2) $(\delta \boldsymbol{B})_{rot} = \boldsymbol{\Omega} \times \boldsymbol{B} \, \delta t$, and so the rates of change of the vector $\boldsymbol{B}$ in the inertial and rotating frames are related by

$$\boxed{\left(\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}\right)_I = \left(\frac{\mathrm{d}\boldsymbol{B}}{\mathrm{d}t}\right)_R + \boldsymbol{\Omega} \times \boldsymbol{B}}\,. \tag{2.5}$$

This relation applies to a vector $\boldsymbol{B}$ that, as measured at any one time, is the same in both inertial and rotating frames.

2.1.2 Velocity and acceleration in a rotating frame

The velocity of a body is not measured to be the same in the inertial and rotating frames, so care must be taken when applying (2.5) to velocity. First apply (2.5) to $\boldsymbol{r}$, the position of a particle to obtain

$$\left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\right)_I = \left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\right)_R + \boldsymbol{\Omega} \times \boldsymbol{r} \tag{2.6}$$

or

$$\boldsymbol{v}_I = \boldsymbol{v}_R + \boldsymbol{\Omega} \times \boldsymbol{r}. \tag{2.7}$$

We refer to $\boldsymbol{v}_R$ and $\boldsymbol{v}_I$ as the relative and inertial velocity, respectively, and (2.7) relates the two. Apply (2.5) again, this time to the velocity $\boldsymbol{v}_R$ to give

$$\left(\frac{\mathrm{d}\boldsymbol{v}_R}{\mathrm{d}t}\right)_I = \left(\frac{\mathrm{d}\boldsymbol{v}_R}{\mathrm{d}t}\right)_R + \boldsymbol{\Omega} \times \boldsymbol{v}_R, \tag{2.8}$$

or, using (2.7)

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{v}_I - \boldsymbol{\Omega} \times \boldsymbol{r})\right)_I = \left(\frac{\mathrm{d}\boldsymbol{v}_R}{\mathrm{d}t}\right)_R + \boldsymbol{\Omega} \times \boldsymbol{v}_R, \tag{2.9}$$

or

$$\left(\frac{\mathrm{d}\boldsymbol{v}_I}{\mathrm{d}t}\right)_I = \left(\frac{\mathrm{d}\boldsymbol{v}_R}{\mathrm{d}t}\right)_R + \boldsymbol{\Omega} \times \boldsymbol{v}_R + \frac{\mathrm{d}\boldsymbol{\Omega}}{\mathrm{d}t} \times \boldsymbol{r} + \boldsymbol{\Omega} \times \left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\right)_I. \tag{2.10}$$

Then, noting that

$$\left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\right)_I = \left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\right)_R + \boldsymbol{\Omega} \times \boldsymbol{r} = (\boldsymbol{v}_R + \boldsymbol{\Omega} \times \boldsymbol{r}), \tag{2.11}$$

and assuming that the rate of rotation is constant, (2.10) becomes

$$\left(\frac{\mathrm{d}\boldsymbol{v}_R}{\mathrm{d}t}\right)_R = \left(\frac{\mathrm{d}\boldsymbol{v}_I}{\mathrm{d}t}\right)_I - 2\boldsymbol{\Omega} \times \boldsymbol{v}_R - \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \boldsymbol{r}). \tag{2.12}$$

This equation may be interpreted as follows. The term on the left-hand side is the rate of change of the relative velocity as measured in the rotating frame. The first term on the right-hand side is the rate of change of the inertial velocity as measured in the inertial frame (or, loosely, the inertial acceleration). Thus, by Newton's second law, it is equal to the force on a fluid parcel divided by its mass. The second and third terms on the right-hand side (including the minus signs) are the *Coriolis force* and the *centrifugal force* per unit mass. Neither of these are true forces — they may be thought of as quasi-forces (i.e., 'as if' forces); that is, when a body is observed from a rotating frame it seems to behave as if unseen forces are present that affect its motion. If (2.12) is written, as is common, with the terms $+2\boldsymbol{\Omega} \times \boldsymbol{v}_r$ and $+\boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \boldsymbol{r})$ on the left-hand side then these terms should be referred to as the Coriolis and centrifugal *accelerations*.[1]

Centrifugal force

If $\boldsymbol{r}_\perp$ is the perpendicular distance from the axis of rotation (see Fig. 2.1 and substitute $\boldsymbol{r}$ for $\boldsymbol{C}$), then, because $\boldsymbol{\Omega}$ is perpendicular to $\boldsymbol{r}_\perp$, $\boldsymbol{\Omega} \times \boldsymbol{r} = \boldsymbol{\Omega} \times \boldsymbol{r}_\perp$. Then, using the vector identity $\boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \boldsymbol{r}_\perp) = (\boldsymbol{\Omega} \cdot \boldsymbol{r}_\perp)\boldsymbol{\Omega} - (\boldsymbol{\Omega} \cdot \boldsymbol{\Omega})\boldsymbol{r}_\perp$ and noting that the first term is zero, we see that the centrifugal force per unit mass is just given by

$$\boldsymbol{F}_{ce} = -\boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \boldsymbol{r}) = \Omega^2 \boldsymbol{r}_\perp. \tag{2.13}$$

This may usefully be written as the gradient of a scalar potential,

$$\boldsymbol{F}_{ce} = -\nabla \Phi_{ce}. \tag{2.14}$$

where $\Phi_{ce} = -(\Omega^2 r_\perp^2)/2 = -(\boldsymbol{\Omega} \times \boldsymbol{r}_\perp)^2/2$.

Coriolis force

The Coriolis force per unit mass is:

$$\boldsymbol{F}_{Co} = -2\boldsymbol{\Omega} \times \boldsymbol{v}_R. \tag{2.15}$$

It plays a central role in much of geophysical fluid dynamics and will be considered extensively later on. For now, we just note three basic properties.

(i) There is no Coriolis force on bodies that are stationary in the rotating frame.

(ii) The Coriolis force acts to deflect moving bodies at right angles to their direction of travel.

(iii) The Coriolis force does no work on a body because it is perpendicular to the velocity, and so $\boldsymbol{v}_R \cdot (\boldsymbol{\Omega} \times \boldsymbol{v}_R) = 0$.

2.1.3 Momentum equation in a rotating frame

Since (2.12) simply relates the accelerations of a particle in the inertial and rotating frames, then in the rotating frame of reference the momentum equation may be written

$$\frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} + 2\boldsymbol{\Omega} \times \boldsymbol{v} = -\frac{1}{\rho}\nabla p - \nabla \Phi. \tag{2.16}$$

We have dropped the subscript R; henceforth, unless ambiguity is present, all velocities without a subscript will be considered to be relative to the rotating frame.

2.1.4 Mass and tracer conservation in a rotating frame

Let ϕ be a scalar field that, in the inertial frame, obeys

$$\frac{\mathrm{D}\phi}{\mathrm{D}t} + \phi \nabla \cdot \boldsymbol{v}_I = 0. \tag{2.17}$$

Now, observers in both the rotating and inertial frame measure the same value of ϕ. Further, $\mathrm{D}\phi/\mathrm{D}t$ is simply the rate of change of ϕ associated with a material parcel, and therefore is reference frame invariant. Thus,

$$\left(\frac{\mathrm{D}\phi}{\mathrm{D}t}\right)_R = \left(\frac{\mathrm{D}\phi}{\mathrm{D}t}\right)_I, \tag{2.18}$$

where $(\mathrm{D}\phi/\mathrm{D}t)_R = (\partial\phi/\partial t)_R + \boldsymbol{v}_R \cdot \nabla\phi$ and $(\mathrm{D}\phi/\mathrm{D}t)_I = (\partial\phi/\partial t)_I + \boldsymbol{v}_I \cdot \nabla\phi$ and the local temporal derivatives $(\partial\phi/\partial t)_R$ and $(\partial\phi/\partial t)_I$ are evaluated at fixed locations in the rotating and inertial frames, respectively.

Further, since $\boldsymbol{v} = \boldsymbol{v}_I - \boldsymbol{\Omega} \times \boldsymbol{r}$, we have that

$$\nabla \cdot \boldsymbol{v}_I = \nabla \cdot (\boldsymbol{v}_I - \boldsymbol{\Omega} \times \boldsymbol{r}) = \nabla \cdot \boldsymbol{v}_R \tag{2.19}$$

since $\nabla \cdot (\boldsymbol{\Omega} \times \boldsymbol{r}) = 0$. Thus, using (2.18) and (2.19), (2.17) is equivalent to

$$\frac{\mathrm{D}\phi}{\mathrm{D}t} + \phi\nabla \cdot \boldsymbol{v} = 0, \tag{2.20}$$

where all observables are measured in the *rotating* frame. Thus, the equation for the evolution of a scalar whose measured value is the same in rotating and inertial frames is unaltered by the presence of rotation. In particular, the mass conservation equation is unaltered by the presence of rotation.

Although we have taken (2.18) as true a priori, the individual components of the material derivative differ in the rotating and inertial frames. In particular

$$\left(\frac{\partial\phi}{\partial t}\right)_I = \left(\frac{\partial\phi}{\partial t}\right)_R - (\boldsymbol{\Omega} \times \boldsymbol{r}) \cdot \nabla\phi \tag{2.21}$$

because $\boldsymbol{\Omega} \times \boldsymbol{r}$ is the velocity, in the inertial frame, of a uniformly rotating body. Similarly,

$$\boldsymbol{v}_I \cdot \nabla\phi = (\boldsymbol{v}_R + \boldsymbol{\Omega} \times \boldsymbol{r}) \cdot \nabla\phi. \tag{2.22}$$

Adding the last two equations reprises and confirms (2.18).

2.2 EQUATIONS OF MOTION IN SPHERICAL COORDINATES

The Earth is very nearly spherical and it might appear obvious that we should cast our equations in spherical coordinates. Although this does turn out to be true, the presence of a centrifugal force causes some complications which we must first discuss. The reader who is willing ab initio to treat the Earth as a perfect sphere and to neglect the horizontal component of the centrifugal force may skip the next section.

2.2.1 * The centrifugal force and spherical coordinates

The centrifugal force is a potential force, like gravity, and so we may therefore define an 'effective gravity' equal to the sum of the true, or Newtonian, gravity and the centrifugal force. The Newtonian gravitational force is directed approximately toward the centre of the Earth, with small deviations due mainly to the Earth's oblateness. The line of action of the effective gravity will in general differ slightly from this, and therefore have a component in the 'horizontal' plane, that is the plane perpendicular to the radial direction. The magnitude of the centrifugal force is $\Omega^2 r_\perp$, and so the effective gravity is given by

$$\boldsymbol{g} \equiv \boldsymbol{g}_{eff} = \boldsymbol{g}_{grav} + \Omega^2 \boldsymbol{r}_\perp, \tag{2.23}$$

where $\boldsymbol{g}_{grav}$ is the Newtonian gravitational force due to the gravitational attraction of the Earth and $\boldsymbol{r}_\perp$ is normal to the rotation vector (in the direction $\boldsymbol{C}$ in Fig. 2.2), with $r_\perp =$

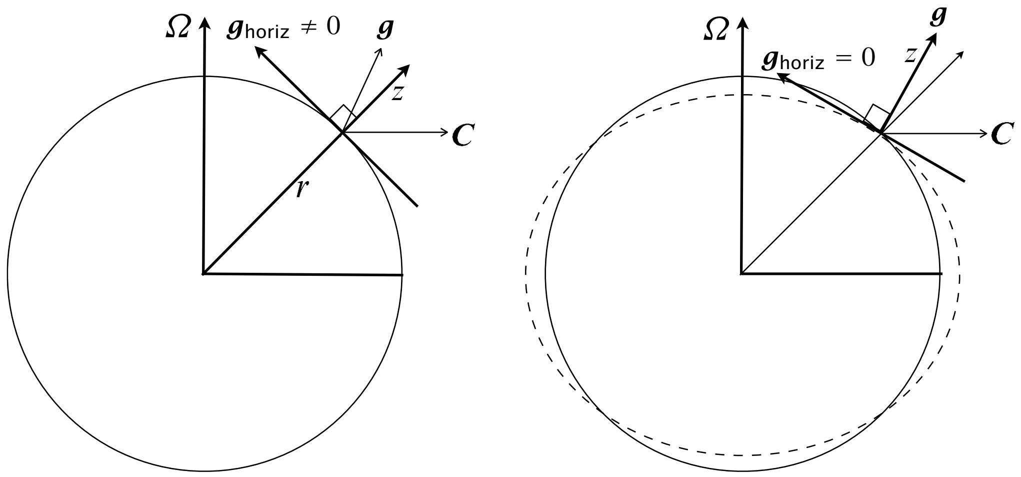


Fig. 2.2 Left: directions of forces and coordinates in true spherical geometry. $\boldsymbol{g}$ is the effective gravity (including the centrifugal force, $\boldsymbol{C}$) and its horizontal component is evidently non-zero. Right: a modified coordinate system, in which the vertical direction is defined by the direction of $\boldsymbol{g}$, and so the horizontal component of $\boldsymbol{g}$ is identically zero. The dashed line schematically indicates a surface of constant geopotential. The differences between the direction of $\boldsymbol{g}$ and the direction of the radial coordinate, and between the sphere and the geopotential surface, are much exaggerated and in reality are similar to the thickness of the lines themselves.

$r\cos\vartheta$. Both gravity and centrifugal force are potential forces and therefore we may define the *geopotential, Φ*, such that

$$\boldsymbol{g} = -\nabla\Phi. \tag{2.24}$$

Surfaces of constant Φ are not quite spherical because $r_\perp$, and hence the centrifugal force, vary with latitude (Fig. 2.2); this has certain ramifications, as we now discuss.

The components of the centrifugal force parallel and perpendicular to the radial direction are $\Omega^2 r\cos^2\vartheta$ and $\Omega^2 r\cos\vartheta\sin\vartheta$. Newtonian gravity is much larger than either of these, and at the Earth's surface the ratio of centrifugal to gravitational terms is approximately, and no more than,

$$\alpha \approx \frac{\Omega^2 a}{g} \approx \frac{(7.27\times10^{-5})^2\times6.4\times10^6}{10} \approx 3\times10^{-3}. \tag{2.25}$$

(Note that at the equator and pole the horizontal component of the centrifugal force is zero and the effective gravity is aligned with Newtonian gravity.) The angle between $\boldsymbol{g}$ and the line to the centre of the Earth is given by a similar expression and so is also small, typically around 3×10^{-3} radians. However, the horizontal component of the centrifugal force is still large compared to the Coriolis force, their ratio in mid-latitudes being given by

$$\frac{\text{horizontal centrifugal force}}{\text{Coriolis force}} \approx \frac{\Omega^2 a\cos\vartheta\sin\vartheta}{2\Omega u} \approx \frac{\Omega a}{4|u|} \approx 10, \tag{2.26}$$

using $u = 10\,\mathrm{m\,s^{-1}}$. The centrifugal term therefore dominates over the Coriolis term, and is largely balanced by a pressure gradient force. Thus, if we adhered to true spherical coordinates, both the horizontal and radial components of the momentum equation would be dominated by a static balance between a pressure gradient and gravity or centrifugal terms. Although in principle there is nothing wrong with writing the equations this way, it obscures the dynamical balances involving the Coriolis force and pressure that determine the large-scale horizontal flow.

A way around this problem is to use the direction of the geopotential force to *define* the vertical direction, and then for all geometric purposes to regard the surfaces of constant Φ as if they were true spheres.[2] The horizontal component of effective gravity is then identically zero, and we have traded a potentially large dynamical error for a very small geometric error. In fact, over time, the Earth has developed an equatorial bulge to compensate for and neutralize the centrifugal force, so that the effective gravity does act in a direction virtually normal to the Earth's surface; that is, the surface of the Earth is an oblate spheroid of nearly constant geopotential. The geopotential Φ is then a function of the vertical coordinate alone, and for many purposes we can just take $\Phi = gz$; that is, the direction normal to geopotential surfaces, the local vertical, is, in this approximation, taken to be the direction of increasing r in spherical coordinates. It is because the oblateness is very small (the polar diameter is about 12 714 km, whereas the equatorial diameter is about 12 756 km) that using spherical coordinates is a very accurate way to map the spheroid, and if the angle between effective gravity and a natural direction of the coordinate system were not small then more heroic measures would be called for.

If the solid Earth did not bulge at the equator, the *behaviour* of the atmosphere and ocean would differ significantly from that of the present system. For example, the surface of the ocean is nearly a geopotential surface, and if the solid Earth were exactly spherical then the ocean would perforce become much deeper at low latitudes and the ocean basins would dry out completely at high latitudes. We could still choose to use the spherical coordinate system discussed above to describe the dynamics, but the shape of the surface of the solid Earth would have to be represented by a topography, with the topographic height increasing monotonically polewards nearly everywhere.

2.2.2 Some identities in spherical coordinates

The location of a point is given by the coordinates (λ, ϑ, r) where λ is the angular distance eastwards (i.e., longitude), ϑ is angular distance polewards (i.e., latitude) and r is the radial distance from the centre of the Earth — see Fig. 2.3. (In some other fields of study co-latitude is used as a spherical coordinate.) If a is the radius of the Earth, then we also define $z = r - a$. At a given location we may also define the Cartesian increments $(\delta x, \delta y, \delta z) = (r\cos\vartheta\,\delta\lambda, r\,\delta\vartheta, \delta r)$.

For a scalar quantity ϕ the material derivative in spherical coordinates is

$$\frac{\mathrm{D}\phi}{\mathrm{D}t} = \frac{\partial\phi}{\partial t} + \frac{u}{r\cos\vartheta}\frac{\partial\phi}{\partial\lambda} + \frac{v}{r}\frac{\partial\phi}{\partial\vartheta} + w\frac{\partial\phi}{\partial r}, \tag{2.27}$$

where the velocity components corresponding to the coordinates (λ, ϑ, r) are

$$(u, v, w) \equiv \left(r\cos\vartheta\frac{\mathrm{D}\lambda}{\mathrm{D}t}, r\frac{\mathrm{D}\vartheta}{\mathrm{D}t}, \frac{\mathrm{D}r}{\mathrm{D}t}\right). \tag{2.28}$$

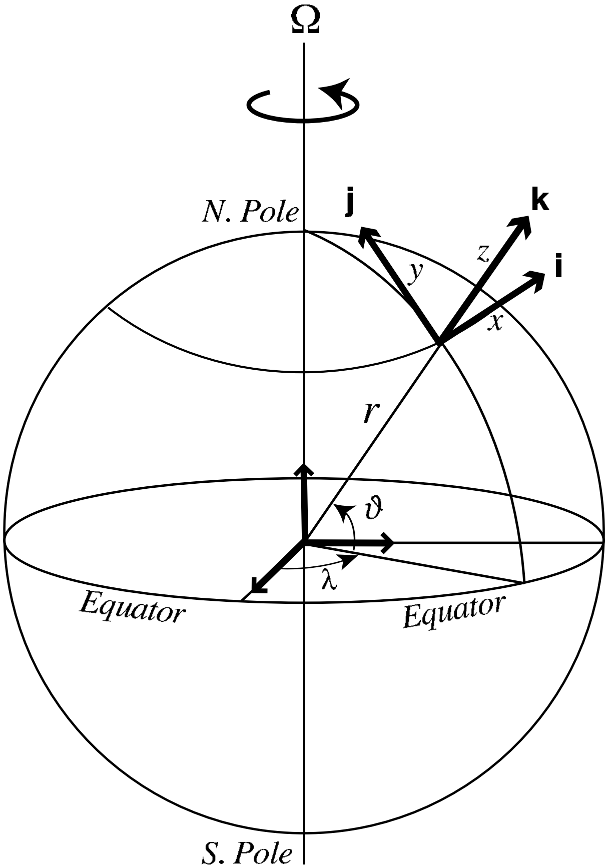


Fig. 2.3 The spherical coordinate system. The orthogonal unit vectors **i**, **j** and **k** point in the direction of increasing longitude λ, latitude ϑ, and altitude z. Locally, one may apply a Cartesian system with variables x, y and z measuring distances along **i**, **j** and **k**.

That is, u is the zonal velocity, v is the meridional velocity and w is the vertical velocity. If we define $(\mathbf{i},\mathbf{j},\mathbf{k})$ to be the unit vectors in the direction of increasing (λ,ϑ,r) then

$$\boldsymbol{v} = \mathbf{i}u + \mathbf{j}v + \mathbf{k}w. \tag{2.29}$$

Note also that $\mathrm{D}r/\mathrm{D}t = \mathrm{D}z/\mathrm{D}t$.

The divergence of a vector $\boldsymbol{B} = \mathbf{i}B^\lambda + \mathbf{j}B^\vartheta + \mathbf{k}B^r$ is

$$\nabla\cdot\boldsymbol{B} = \frac{1}{\cos\vartheta}\left[\frac{1}{r}\frac{\partial B^\lambda}{\partial\lambda} + \frac{1}{r}\frac{\partial}{\partial\vartheta}(B^\vartheta\cos\vartheta) + \frac{\cos\vartheta}{r^2}\frac{\partial}{\partial r}(r^2B^r)\right]. \tag{2.30}$$

The vector gradient of a scalar is:

$$\nabla\phi = \mathbf{i}\frac{1}{r\cos\vartheta}\frac{\partial\phi}{\partial\lambda} + \mathbf{j}\frac{1}{r}\frac{\partial\phi}{\partial\vartheta} + \mathbf{k}\frac{\partial\phi}{\partial r}. \tag{2.31}$$

The Laplacian of a scalar is:

$$\nabla^2\phi \equiv \nabla\cdot\nabla\phi = \frac{1}{r^2\cos\vartheta}\left[\frac{1}{\cos\vartheta}\frac{\partial^2\phi}{\partial\lambda^2} + \frac{\partial}{\partial\vartheta}\left(\cos\vartheta\frac{\partial\phi}{\partial\vartheta}\right) + \cos\vartheta\frac{\partial}{\partial r}\left(r^2\frac{\partial\phi}{\partial r}\right)\right]. \tag{2.32}$$

The curl of a vector is:

$$\operatorname{curl}\boldsymbol{B} = \nabla\times\boldsymbol{B} = \frac{1}{r^2\cos\vartheta}\begin{vmatrix} \mathbf{i}\,r\cos\vartheta & \mathbf{j}\,r & \mathbf{k} \\ \partial/\partial\lambda & \partial/\partial\vartheta & \partial/\partial r \\ B^\lambda r\cos\vartheta & B^\vartheta r & B^r \end{vmatrix}. \tag{2.33}$$

The vector Laplacian $\nabla^2\boldsymbol{B}$ (used for example when calculating viscous terms in the momentum equation) may be obtained from the vector identity:

$$\nabla^2\boldsymbol{B} = \nabla(\nabla\cdot\boldsymbol{B}) - \nabla\times(\nabla\times\boldsymbol{B}). \tag{2.34}$$

Only in Cartesian coordinates does this take the simple form:

$$\nabla^2\boldsymbol{B} = \frac{\partial^2\boldsymbol{B}}{\partial x^2} + \frac{\partial^2\boldsymbol{B}}{\partial y^2} + \frac{\partial^2\boldsymbol{B}}{\partial z^2}. \tag{2.35}$$

The expansion in spherical coordinates is, to most eyes, rather uninformative.

Rate of change of unit vectors

In spherical coordinates the defining unit vectors are **i**, the unit vector pointing eastwards, parallel to a line of latitude; **j** is the unit vector pointing polewards, parallel to a meridian; and **k**, the unit vector pointing radially outward. The directions of these vectors change with location, and in fact this is the case in nearly all coordinate systems, with the notable exception of the Cartesian one, and thus their material derivative is not zero. One way to evaluate this is to consider geometrically how the coordinate axes change with position. Another way, and the way that we shall proceed, is to first obtain the effective rotation rate $\boldsymbol{\Omega}_{flow}$, relative to the Earth, of a unit vector as it moves with the flow, and then apply (2.3). Specifically, let the fluid velocity be $\boldsymbol{v} = (u, v, w)$. The meridional component, v, produces a displacement $r\delta\vartheta = v\delta t$, and this gives rise a local effective vector rotation rate around the local zonal axis of $-(v/r)\mathbf{i}$, the minus sign arising because a displacement in the direction of the north pole is produced by negative rotational displacement around the **i** axis. Similarly, the zonal component, u, produces a displacement $\delta\lambda r\cos\vartheta = u\delta t$ and so an effective rotation rate, about the Earth's rotation axis, of $u/(r\cos\vartheta)$. Now, a rotation around the Earth's rotation axis may be written as (see Fig. 2.4)

$$\boldsymbol{\Omega} = \Omega(\mathbf{j}\cos\vartheta + \mathbf{k}\sin\vartheta). \tag{2.36}$$

If the scalar rotation rate is not Ω but is $u/(r\cos\vartheta)$, then the vector rotation rate is

$$\frac{u}{r\cos\vartheta}(\mathbf{j}\cos\vartheta + \mathbf{k}\sin\vartheta) = \mathbf{j}\frac{u}{r} + \mathbf{k}\frac{u\tan\vartheta}{r}. \tag{2.37}$$

Thus, the total rotation rate of a vector that moves with the flow is

$$\boldsymbol{\Omega}_{flow} = -\mathbf{i}\frac{v}{r} + \mathbf{j}\frac{u}{r} + \mathbf{k}\frac{u\tan\vartheta}{r}. \tag{2.38}$$

Applying (2.3) to (2.38), we find

$$\frac{\mathrm{D}\mathbf{i}}{\mathrm{D}t} = \boldsymbol{\Omega}_{flow}\times\mathbf{i} = \frac{u}{r\cos\vartheta}(\mathbf{j}\sin\vartheta - \mathbf{k}\cos\vartheta), \tag{2.39a}$$

$$\frac{\mathrm{D}\mathbf{j}}{\mathrm{D}t} = \boldsymbol{\Omega}_{flow}\times\mathbf{j} = -\mathbf{i}\frac{u}{r}\tan\vartheta - \mathbf{k}\frac{v}{r}, \tag{2.39b}$$

$$\frac{\mathrm{D}\mathbf{k}}{\mathrm{D}t} = \boldsymbol{\Omega}_{flow}\times\mathbf{k} = \mathbf{i}\frac{u}{r} + \mathbf{j}\frac{v}{r}. \tag{2.39c}$$

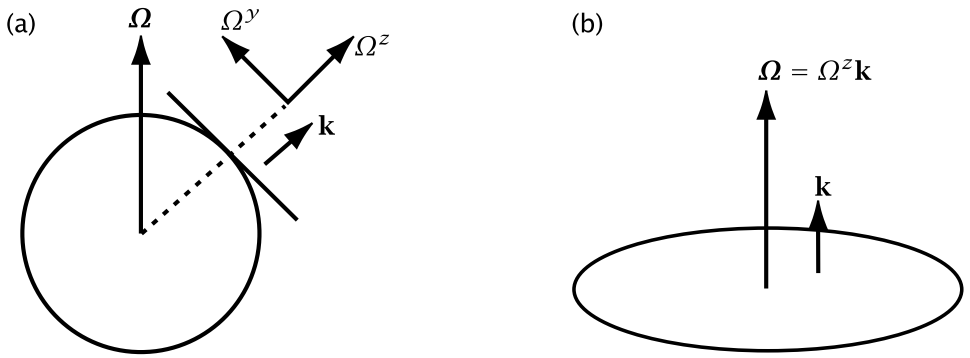


Fig. 2.4 (a) On the sphere the rotation vector $\boldsymbol{\Omega}$ can be decomposed into two components, one in the local vertical and one in the local horizontal, pointing toward the pole. That is, $\boldsymbol{\Omega} = \Omega_y\mathbf{j} + \Omega_z\mathbf{k}$ where $\Omega_y = \Omega\cos\vartheta$ and $\Omega_z = \Omega\sin\vartheta$. In geophysical fluid dynamics, the rotation vector in the local vertical is often the more important component in the horizontal momentum equations. On a rotating disk, (b), the rotation vector $\boldsymbol{\Omega}$ is parallel to the local vertical $\mathbf{k}$.

2.2.3 Equations of motion

Mass Conservation and Thermodynamic Equation

The mass conservation equation, (1.36a), expanded in spherical co-odinates, is

$$\begin{aligned}\frac{\partial\rho}{\partial t} &+ \frac{u}{r\cos\vartheta}\frac{\partial\rho}{\partial\lambda} + \frac{v}{r}\frac{\partial\rho}{\partial\vartheta} + w\frac{\partial\rho}{\partial r}\\ &+ \frac{\rho}{r\cos\vartheta}\left[\frac{\partial u}{\partial\lambda} + \frac{\partial}{\partial\vartheta}(v\cos\vartheta) + \frac{1}{r}\frac{\partial}{\partial r}(wr^2\cos\vartheta)\right] = 0.\end{aligned} \tag{2.40}$$

Equivalently, using the form (1.36b), this is

$$\frac{\partial\rho}{\partial t} + \frac{1}{r\cos\vartheta}\frac{\partial(u\rho)}{\partial\lambda} + \frac{1}{r\cos\vartheta}\frac{\partial}{\partial\vartheta}(v\rho\cos\vartheta) + \frac{1}{r^2}\frac{\partial}{\partial r}(r^2 w\rho) = 0. \tag{2.41}$$

The thermodynamic equation, (1.108), is a tracer advection equation. Thus, using (2.27), its (adiabatic) spherical coordinate form is

$$\frac{\mathrm{D}\theta}{\mathrm{D}t} = \frac{\partial\theta}{\partial t} + \frac{u}{r\cos\vartheta}\frac{\partial\theta}{\partial\lambda} + \frac{v}{r}\frac{\partial\theta}{\partial\vartheta} + w\frac{\partial\theta}{\partial r} = 0, \tag{2.42}$$

and similarly for tracers such as water vapour or salt.

Momentum Equation

Recall that the inviscid momentum equation is:

$$\frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} + 2\boldsymbol{\Omega}\times\boldsymbol{v} = -\frac{1}{\rho}\nabla p - \nabla\Phi, \tag{2.43}$$

where Φ is the geopotential. In spherical coordinates the directions of the coordinate axes change with position and so the component expansion of (2.43) is

$$\frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} = \frac{\mathrm{D}u}{\mathrm{D}t}\mathbf{i} + \frac{\mathrm{D}v}{\mathrm{D}t}\mathbf{j} + \frac{\mathrm{D}w}{\mathrm{D}t}\mathbf{k} + u\frac{\mathrm{D}\mathbf{i}}{\mathrm{D}t} + v\frac{\mathrm{D}\mathbf{j}}{\mathrm{D}t} + w\frac{\mathrm{D}\mathbf{k}}{\mathrm{D}t} \tag{2.44a}$$

$$= \frac{\mathrm{D}u}{\mathrm{D}t}\mathbf{i} + \frac{\mathrm{D}v}{\mathrm{D}t}\mathbf{j} + \frac{\mathrm{D}w}{\mathrm{D}t}\mathbf{k} + \boldsymbol{\Omega}_{flow} \times \boldsymbol{v}, \tag{2.44b}$$

using (2.39). Using either (2.44a) and the expressions for the rates of change of the unit vectors given in (2.39), or (2.44b) and the expression for $\boldsymbol{\Omega}_{flow}$ given in (2.38), (2.44) becomes

$$\begin{aligned}\frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} = \mathbf{i}\left(\frac{\mathrm{D}u}{\mathrm{D}t} - \frac{uv\tan\vartheta}{r} + \frac{uw}{r}\right) + \mathbf{j}\left(\frac{\mathrm{D}v}{\mathrm{D}t} + \frac{u^2\tan\vartheta}{r} + \frac{vw}{r}\right)\\ + \mathbf{k}\left(\frac{\mathrm{D}w}{\mathrm{D}t} - \frac{u^2+v^2}{r}\right).\end{aligned} \tag{2.45}$$

Using the definition of a vector cross product the Coriolis term is:

$$\begin{aligned}2\boldsymbol{\Omega}\times\boldsymbol{v} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 0 & 2\Omega\cos\vartheta & 2\Omega\sin\vartheta \\ u & v & w \end{vmatrix}\\ &= \mathbf{i}\,(2\Omega w\cos\vartheta - 2\Omega v\sin\vartheta) + \mathbf{j}\,2\Omega u\sin\vartheta - \mathbf{k}\,2\Omega u\cos\vartheta.\end{aligned} \tag{2.46}$$

Using (2.45) and (2.46), and the gradient operator given by (2.31), the momentum equation (2.43) becomes:

$$\frac{\mathrm{D}u}{\mathrm{D}t} - \left(2\Omega + \frac{u}{r\cos\vartheta}\right)(v\sin\vartheta - w\cos\vartheta) = -\frac{1}{\rho r\cos\vartheta}\frac{\partial p}{\partial\lambda}, \tag{2.47a}$$

$$\frac{\mathrm{D}v}{\mathrm{D}t} + \frac{wv}{r} + \left(2\Omega + \frac{u}{r\cos\vartheta}\right)u\sin\vartheta = -\frac{1}{\rho r}\frac{\partial p}{\partial\vartheta}, \tag{2.47b}$$

$$\frac{\mathrm{D}w}{\mathrm{D}t} - \frac{u^2+v^2}{r} - 2\Omega u\cos\vartheta = -\frac{1}{\rho}\frac{\partial p}{\partial r} - g. \tag{2.47c}$$

The terms involving Ω are called Coriolis terms, and the quadratic terms on the left-hand sides involving $1/r$ are often called metric terms.

2.2.4 The primitive equations

The so-called *primitive equations* of motion are simplifications of the above equations frequently used in atmospheric and oceanic modelling.[3] Three related approximations are involved.

(i) *The hydrostatic approximation.* In the vertical momentum equation the gravitational term is assumed to be balanced by the pressure gradient term, so that

$$\frac{\partial p}{\partial z} = -\rho g. \tag{2.48}$$

The advection of vertical velocity, the Coriolis terms, and the metric term $(u^2+v^2)/r$ are all neglected.

(ii) *The shallow-fluid approximation.* We write $r = a + z$ where the constant a is the radius of the Earth and z increases in the radial direction. The coordinate r is then replaced by a except where it used as the differentiating argument. Thus, for example,

$$\frac{1}{r^2}\frac{\partial(r^2w)}{\partial r} \to \frac{\partial w}{\partial z}. \tag{2.49}$$

(iii) The traditional approximation. Coriolis terms in the horizontal momentum equations involving the vertical velocity, and the still smaller metric terms uw/r and vw/r, are neglected.

The second and third of these approximations should be taken, or not, together, the underlying reason being that they both relate to the presumed small aspect ratio of the motion, so the approximations succeed or fail together. If we make one approximation but not the other then we are being asymptotically inconsistent, and angular momentum and energy conservation are not assured (see section 2.2.7 and problem 2.13). The hydrostatic approximation also depends on the small aspect ratio of the flow, but in a slightly different way. For large-scale flow in the terrestrial atmosphere and ocean all three approximations are in fact all very accurate approximations. We defer a more complete treatment until section 2.7, in part because a treatment of the hydrostatic approximation is done most easily in the context of the Boussinesq equations, derived in section 2.4.

Making these approximations, the momentum equations become

$$\frac{\mathrm{D}u}{\mathrm{D}t} - 2\Omega \sin\vartheta v - \frac{uv}{a}\tan\vartheta = -\frac{1}{a\rho\cos\vartheta}\frac{\partial p}{\partial \lambda}, \tag{2.50a}$$

$$\frac{\mathrm{D}v}{\mathrm{D}t} + 2\Omega \sin\vartheta u + \frac{u^2\tan\vartheta}{a} = -\frac{1}{\rho a}\frac{\partial p}{\partial \vartheta}, \tag{2.50b}$$

$$0 = -\frac{1}{\rho}\frac{\partial p}{\partial z} - g, \tag{2.50c}$$

where

$$\frac{\mathrm{D}}{\mathrm{D}t} = \left(\frac{\partial}{\partial t} + \frac{u}{a\cos\vartheta}\frac{\partial}{\partial\lambda} + \frac{v}{a}\frac{\partial}{\partial\vartheta} + w\frac{\partial}{\partial z}\right). \tag{2.51}$$

We note the ubiquity of the factor $2\Omega\sin\vartheta$, and take the opportunity to define the *Coriolis parameter*, $f \equiv 2\Omega\sin\vartheta$.

The corresponding mass conservation equation for a shallow fluid layer is:

$$\begin{aligned}\frac{\partial\rho}{\partial t} &+ \frac{u}{a\cos\vartheta}\frac{\partial\rho}{\partial\lambda} + \frac{v}{a}\frac{\partial\rho}{\partial\vartheta} + w\frac{\partial\rho}{\partial z}\\ &+ \rho\left[\frac{1}{a\cos\vartheta}\frac{\partial u}{\partial\lambda} + \frac{1}{a\cos\vartheta}\frac{\partial}{\partial\vartheta}(v\cos\vartheta) + \frac{\partial w}{\partial z}\right] = 0,\end{aligned} \tag{2.52}$$

or equivalently,

$$\frac{\partial\rho}{\partial t} + \frac{1}{a\cos\vartheta}\frac{\partial(u\rho)}{\partial\lambda} + \frac{1}{a\cos\vartheta}\frac{\partial}{\partial\vartheta}(v\rho\cos\vartheta) + \frac{\partial(w\rho)}{\partial z} = 0. \tag{2.53}$$

2.2.5 Primitive equations in vector form

The primitive equations may be written in a compact vector form provided we make a slight reinterpretation of the material derivative of the coordinate axes. Let $\boldsymbol{u} = u\mathbf{i} + v\mathbf{j} + 0\mathbf{k}$ be the horizontal velocity. The primitive equations (2.50a) and (2.50b) may be written as

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f}\times\boldsymbol{u} = -\frac{1}{\rho}\nabla_z p, \tag{2.54}$$

where $\boldsymbol{f} = f\mathbf{k} = 2\Omega \sin\vartheta \mathbf{k}$ and $\nabla_z p = [(a\cos\vartheta)^{-1}\partial p/\partial\lambda, a^{-1}\partial p/\partial\vartheta]$, the gradient operator at constant z. In (2.54) the material derivative of the horizontal velocity is given by

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} = \mathbf{i}\frac{\mathrm{D}u}{\mathrm{D}t} + \mathbf{j}\frac{\mathrm{D}v}{\mathrm{D}t} + u\frac{\mathrm{D}\mathbf{i}}{\mathrm{D}t} + v\frac{\mathrm{D}\mathbf{j}}{\mathrm{D}t}, \tag{2.55}$$

where instead of (2.39) we have

$$\frac{\mathrm{D}\mathbf{i}}{\mathrm{D}t} = \widetilde{\boldsymbol{\Omega}}_{flow} \times \mathbf{i} = \mathbf{j}\frac{u\tan\vartheta}{a}, \tag{2.56a}$$

$$\frac{\mathrm{D}\mathbf{j}}{\mathrm{D}t} = \widetilde{\boldsymbol{\Omega}}_{flow} \times \mathbf{j} = -\mathbf{i}\frac{u\tan\vartheta}{a}, \tag{2.56b}$$

where $\widetilde{\boldsymbol{\Omega}}_{flow} = \mathbf{k}u\tan\vartheta/a$ [which is the vertical component of (2.38), with r replaced by a.]. The advection of the horizontal wind $\boldsymbol{u}$ is still by the three-dimensional velocity $\boldsymbol{v}$. The vertical momentum equation is the hydrostatic equation, (2.50c), and the mass conservation equation is

$$\frac{\mathrm{D}\rho}{\mathrm{D}t} + \rho\nabla\cdot\boldsymbol{v} = 0 \qquad \text{or} \qquad \frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\boldsymbol{v}) = 0, \tag{2.57}$$

where $\mathrm{D}/\mathrm{D}t$ on a scalar is given by (2.51), and the second expression is written out in full in (2.53).

2.2.6 The vector invariant form of the momentum equation

The 'vector invariant' form of the momentum equation is so-called because it appears to take the same form in all coordinate systems — there is no advective derivative of the coordinate system to worry about. With the aid of the identity $(\boldsymbol{v}\cdot\nabla)\boldsymbol{v} = -\boldsymbol{v}\times\boldsymbol{\omega} + \nabla(\boldsymbol{v}^2/2)$, where $\boldsymbol{\omega} \equiv \nabla\times\boldsymbol{v}$ is the relative vorticity, the three-dimensional momentum equation may be written:

$$\frac{\partial\boldsymbol{v}}{\partial t} + (2\boldsymbol{\Omega} + \boldsymbol{\omega})\times\boldsymbol{v} = -\frac{1}{\rho}\nabla p - \frac{1}{2}\nabla\boldsymbol{v}^2 + \boldsymbol{g}. \tag{2.58}$$

In spherical coordinates the relative vorticity is given by:

$$\begin{aligned}
\boldsymbol{\omega} = \nabla\times\boldsymbol{v} &= \frac{1}{r^2\cos\vartheta}\begin{vmatrix} \mathbf{i}\,r\cos\vartheta & \mathbf{j}\,r & \mathbf{k} \\ \partial/\partial\lambda & \partial/\partial\vartheta & \partial/\partial r \\ ur\cos\vartheta & rv & w \end{vmatrix} \\
&= \mathbf{i}\frac{1}{r}\left(\frac{\partial w}{\partial\vartheta} - \frac{\partial(rv)}{\partial r}\right) - \mathbf{j}\frac{1}{r\cos\vartheta}\left(\frac{\partial w}{\partial\lambda} - \frac{\partial}{\partial r}(ur\cos\vartheta)\right) \\
&\quad + \mathbf{k}\frac{1}{r\cos\vartheta}\left(\frac{\partial v}{\partial\lambda} - \frac{\partial}{\partial\vartheta}(u\cos\vartheta)\right).
\end{aligned} \tag{2.59}$$

Making the traditional and shallow fluid approximations, the horizontal components of (2.58) may be written

$$\frac{\partial\boldsymbol{u}}{\partial t} + (\boldsymbol{f} + \mathbf{k}\zeta)\times\boldsymbol{u} + w\frac{\partial\boldsymbol{u}}{\partial z} = -\frac{1}{\rho}\nabla_z p - \frac{1}{2}\nabla\boldsymbol{u}^2, \tag{2.60}$$

where $\boldsymbol{u} = (u, v, 0)$, $\boldsymbol{f} = \mathbf{k}\,2\Omega\sin\vartheta$, ∇_z is the horizontal gradient operator (the gradient at a constant value of z), and using (2.59), ζ is given by

$$\zeta = \frac{1}{a\cos\vartheta}\frac{\partial v}{\partial\lambda} - \frac{1}{a\cos\vartheta}\frac{\partial}{\partial\vartheta}(u\cos\vartheta) = \frac{1}{a\cos\vartheta}\frac{\partial v}{\partial\lambda} - \frac{1}{a}\frac{\partial u}{\partial\vartheta} + \frac{u}{a}\tan\vartheta. \tag{2.61}$$

The separate components of the momentum equation are given by:

$$\frac{\partial u}{\partial t} - (f+\zeta)v + w\frac{\partial u}{\partial z} = -\frac{1}{a\rho\cos\vartheta}\left(\frac{1}{\rho}\frac{\partial p}{\partial \lambda} + \frac{1}{2}\frac{\partial \boldsymbol{u}^2}{\partial \lambda}\right), \tag{2.62}$$

and

$$\frac{\partial v}{\partial t} + (f+\zeta)u + w\frac{\partial v}{\partial z} = -\frac{1}{a}\left(\frac{1}{\rho}\frac{\partial p}{\partial \vartheta} + \frac{1}{2}\frac{\partial \boldsymbol{u}^2}{\partial \vartheta}\right). \tag{2.63}$$

Related expressions are given in problem 2.3, and we treat vorticity at greater length in chapter 4.

2.2.7 Angular momentum

The zonal momentum equation can be usefully expressed as a statement about axial angular momentum; that is, angular momentum about the rotation axis. The zonal angular momentum per unit mass is the component of angular momentum in the direction of the axis of rotation and it is given by, without making any shallow atmosphere approximation,

$$m = (u + \Omega r\cos\vartheta)r\cos\vartheta. \tag{2.64}$$

The evolution equation for this quantity follows from the zonal momentum equation and has the simple form

$$\frac{\mathrm{D}m}{\mathrm{D}t} = -\frac{1}{\rho}\frac{\partial p}{\partial \lambda}, \tag{2.65}$$

where the material derivative is

$$\frac{\mathrm{D}}{\mathrm{D}t} = \frac{\partial}{\partial t} + \frac{u}{r\cos\vartheta}\frac{\partial}{\partial \lambda} + \frac{v}{r}\frac{\partial}{\partial \vartheta} + w\frac{\partial}{\partial r}. \tag{2.66}$$

Using the mass continuity equation, (2.65) can be written as

$$\frac{\mathrm{D}\rho m}{\mathrm{D}t} + \rho m\nabla\cdot\boldsymbol{v} = -\frac{\partial p}{\partial \lambda} \tag{2.67}$$

or

$$\frac{\partial \rho m}{\partial t} + \frac{1}{r\cos\vartheta}\frac{\partial(\rho u m)}{\partial \lambda} + \frac{1}{r\cos\vartheta}\frac{\partial}{\partial \vartheta}(\rho v m\cos\vartheta) + \frac{\partial}{\partial z}(\rho m w) = -\frac{\partial p}{\partial \lambda}. \tag{2.68}$$

This is an angular momentum conservation equation.

If the fluid is confined to a shallow layer near the surface of a sphere, then we may replace r, the radial coordinate, by a, the radius of the sphere, in the definition of m, and we define $\widetilde{m} \equiv (u + \Omega a\cos\vartheta)a\cos\vartheta$. Then (2.65) is replaced by

$$\frac{\mathrm{D}\widetilde{m}}{\mathrm{D}t} = -\frac{1}{\rho}\frac{\partial p}{\partial \lambda}, \tag{2.69}$$

where now

$$\frac{\mathrm{D}}{\mathrm{D}t} = \frac{\partial}{\partial t} + \frac{u}{a\cos\vartheta}\frac{\partial}{\partial \lambda} + \frac{v}{a}\frac{\partial}{\partial \vartheta} + w\frac{\partial}{\partial z}. \tag{2.70}$$

Using mass continuity, (2.69) may be written as

$$\frac{\partial \rho\widetilde{m}}{\partial t} + \frac{u}{a\cos\vartheta}\frac{\partial \widetilde{m}}{\partial \lambda} + \frac{v}{a}\frac{\partial \widetilde{m}}{\partial \vartheta} + w\frac{\partial \widetilde{m}}{\partial z} = -\frac{1}{\rho}\frac{\partial p}{\partial \lambda}, \tag{2.71}$$

which is the appropriate angular momentum conservation equation for a shallow atmosphere.

** From angular momentum to the spherical component equations*

An alternative way of deriving the three components of the momentum equation in spherical polar coordinates is to *begin* with (2.65) and the principle of conservation of energy. That is, we take the equations for conservation of angular momentum and energy as true a priori and demand that the forms of the momentum equation be constructed to satisfy these. Expanding the material derivative in (2.65), noting that $\mathrm{D}r/\mathrm{D}t = w$ and $\mathrm{D}\cos\vartheta/\mathrm{D}t = -(v/r)\sin\vartheta$, immediately gives (2.47a). Multiplication by u then yields

$$u\frac{\mathrm{D}u}{\mathrm{D}t} - 2\Omega uv\sin\vartheta + 2\Omega uw\cos\vartheta - \frac{u^2 v\tan\vartheta}{r} + \frac{u^2 w}{r} = -\frac{u}{\rho r\cos\vartheta}\frac{\partial p}{\partial\lambda}. \tag{2.72}$$

Now suppose that the meridional and vertical momentum equations are of the form

$$\frac{\mathrm{D}v}{\mathrm{D}t} + \text{Coriolis and metric terms} = -\frac{1}{\rho r}\frac{\partial p}{\partial\vartheta} \tag{2.73a}$$

$$\frac{\mathrm{D}w}{\mathrm{D}t} + \text{Coriolis and metric terms} = -\frac{1}{\rho}\frac{\partial p}{\partial r}, \tag{2.73b}$$

but that we do not know what form the Coriolis and metric terms take. To determine that form, construct the kinetic energy equation by multiplying (2.73) by v and w, respectively. Now, the metric terms must vanish when we sum the resulting equations along with (2.72), so that (2.73a) must contain the Coriolis term $2\Omega u\sin\vartheta$ as well as the metric term $u^2\tan\vartheta/r$, and (2.73b) must contain the term $-2\Omega u\cos\phi$ as well as the metric term u^2/r. But if (2.73b) contains the term u^2/r it must also contain the term v^2/r by isotropy, and therefore (2.73a) must also contain the term vw/r. In this way, (2.47) is precisely reproduced, although the sceptic might argue that the uniqueness of the form has not been demonstrated.

A particular advantage of this approach arises in determining the appropriate momentum equations that conserve angular momentum and energy in the shallow-fluid approximation. We begin with (2.69) and expand to obtain (2.50a). Multiplying by u gives

$$u\frac{\mathrm{D}u}{\mathrm{D}t} - 2\Omega uv\sin\vartheta - \frac{u^2 v\tan\vartheta}{a} = -\frac{u}{\rho a\cos\vartheta}\frac{\partial p}{\partial\lambda}. \tag{2.74}$$

To ensure energy conservation, the meridional momentum equation must contain the Coriolis term $2\Omega u\sin\vartheta$ and the metric term $u^2\tan\vartheta/a$, but the vertical momentum equation must have neither of the metric terms appearing in (2.47c). Thus we deduce the following equations:

$$\frac{\mathrm{D}u}{\mathrm{D}t} - \left(2\Omega\sin\vartheta + \frac{u\tan\vartheta}{a}\right)v = -\frac{1}{\rho a\cos\vartheta}\frac{\partial p}{\partial\lambda}, \tag{2.75a}$$

$$\frac{\mathrm{D}v}{\mathrm{D}t} + \left(2\Omega\sin\vartheta + \frac{u\tan\vartheta}{a}\right)u = -\frac{1}{\rho a}\frac{\partial p}{\partial\vartheta}, \tag{2.75b}$$

$$\frac{\mathrm{D}w}{\mathrm{D}t} = -\frac{1}{\rho}\frac{\partial \rho}{\partial r} - g. \tag{2.75c}$$

This equation set, when used in conjunction with the thermodynamic and mass continuity equations, conserves appropriate forms of angular momentum and energy. In the hydrostatic approximation the material derivative of w in (2.75c) is *additionally* neglected.

Thus, the hydrostatic approximation is mathematically and physically consistent with the shallow-fluid approximation, but it is an additional approximation with slightly different requirements that one may choose, rather than being required, to make. From an asymptotic perspective, the difference lies in the small parameter necessary for either approximation to hold, namely:

$$\text{shallow fluid and traditional approximations:} \qquad \gamma \equiv \frac{H}{a} \ll 1, \tag{2.76a}$$

$$\text{small aspect ratio for hydrostatic approximation:} \qquad \alpha \equiv \frac{H}{L} \ll 1, \tag{2.76b}$$

where L is the horizontal scale of the motion and a is the radius of the Earth. For hemispheric or global scale phenomena $L \sim a$ and the two approximations coincide. (Requirement (2.76b) for the hydrostatic approximation is derived in section 2.7.)

2.3 CARTESIAN APPROXIMATIONS: THE TANGENT PLANE

2.3.1 The f-plane

Although the rotation of the Earth is central for many dynamical phenomena, the sphericity of the Earth is not always so. This is especially true for phenomena on a scale somewhat smaller than global where the use of spherical coordinates becomes awkward, and it is more convenient to use a locally Cartesian representation of the equations. Referring to Fig. 2.4 we will define a plane tangent to the surface of the Earth at a latitude ϑ_0, and then use a Cartesian coordinate system (x, y, z) to describe motion on that plane. For small excursions on the plane, $(x, y, z) \approx (a\lambda \cos\vartheta_0, a(\vartheta - \vartheta_0), z)$. Consistently, the velocity is $\boldsymbol{v} = (u, v, w)$, so that u,v and w are the components of the velocity *in the tangent plane,* in approximately in the east–west, north–south and vertical directions, respectively.

The momentum equations for flow in this plane are then

$$\frac{\partial u}{\partial t} + (\boldsymbol{v}\cdot\nabla)u + 2\Omega^y w - 2\Omega^z v = -\frac{1}{\rho}\frac{\partial p}{\partial x}, \tag{2.77a}$$

$$\frac{\partial v}{\partial t} + (\boldsymbol{v}\cdot\nabla)v + 2\Omega^z u = -\frac{1}{\rho}\frac{\partial p}{\partial y}, \tag{2.77b}$$

$$\frac{\partial w}{\partial t} + (\boldsymbol{v}\cdot\nabla)w + 2(\Omega^x v - \Omega^y u) = -\frac{1}{\rho}\frac{\partial p}{\partial z} - g, \tag{2.77c}$$

where the rotation vector $\boldsymbol{\Omega} = \Omega^x\mathbf{i} + \Omega^y\mathbf{j} + \Omega^z\mathbf{k}$ and $\Omega^x = 0$, $\Omega^y = \Omega\cos\vartheta_0$ and $\Omega^z = \Omega\sin\vartheta_0$. If we make the traditional approximation, and so ignore the components of $\boldsymbol{\Omega}$ not in the direction of the local vertical, then

$$\frac{\mathrm{D}u}{\mathrm{D}t} - f_0 v = -\frac{1}{\rho}\frac{\partial p}{\partial x}, \tag{2.78a}$$

$$\frac{\mathrm{D}v}{\mathrm{D}t} + f_0 u = -\frac{1}{\rho}\frac{\partial p}{\partial y}, \tag{2.78b}$$

$$\frac{\mathrm{D}w}{\mathrm{D}t} = -\frac{1}{\rho}\frac{\partial p}{\partial z} - \rho g. \tag{2.78c}$$

where $f_0 = 2\Omega^z = 2\Omega\sin\vartheta_0$. Defining the horizontal velocity vector $\boldsymbol{u} = (u, v, 0)$, the first two equations may also be written as

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f}_0 \times \boldsymbol{u} = -\frac{1}{\rho}\nabla_z p, \tag{2.79}$$

where $\mathrm{D}\boldsymbol{u}/\mathrm{D}t = \partial \boldsymbol{u}/\partial t + \boldsymbol{v} \cdot \nabla \boldsymbol{u}$, $\boldsymbol{f}_0 = 2\Omega \sin\vartheta_0 \mathbf{k} = f_0 \mathbf{k}$, and $\mathbf{k}$ is the direction perpendicular to the plane (it does not change its orientation with latitude). These equations are, evidently, exactly the same as the momentum equations in a system in which the rotation vector is aligned with the local vertical, as illustrated in the right-hand panel in Fig. 2.4 (on page 60). They will describe flow on the surface of a rotating sphere to a good approximation provided the flow is of limited latitudinal extent so that the effects of sphericity are unimportant; we have made what is known as the *f-plane* approximation since the Coriolis parameter is a constant. We may in addition make the hydrostatic approximation, in which case (2.78c) becomes the familiar $\partial p/\partial z = -\rho g$.

2.3.2 The beta-plane approximation

The magnitude of the vertical component of rotation varies with latitude, and this has important dynamical consequences. We can approximate this effect by allowing the effective rotation vector to vary. Thus, noting that, for small variations in latitude,

$$f = 2\Omega \sin\vartheta \approx 2\Omega \sin\vartheta_0 + 2\Omega(\vartheta - \vartheta_0)\cos\vartheta_0, \tag{2.80}$$

then on the tangent plane we may mimic this by allowing the Coriolis parameter to vary as

$$\boxed{f = f_0 + \beta y}, \tag{2.81}$$

where $f_0 = 2\Omega \sin\vartheta_0$ and $\beta = \partial f/\partial y = (2\Omega\cos\vartheta_0)/a$. This important approximation is known as the *beta-plane*, or *β-plane,* approximation; it captures the the most important *dynamical* effects of sphericity, without the complicating *geometric* effects, which are not essential to describe many phenomena. The momentum equations (2.78) are unaltered except that f_0 is replaced by $f_0 + \beta y$ to represent a varying Coriolis parameter. Thus, sphericity combined with rotation is dynamically equivalent to a *differentially rotating* system. For future reference, we write down the β-plane horizontal momentum equations:

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f} \times \boldsymbol{u} = -\frac{1}{\rho}\nabla_z p, \tag{2.82}$$

where $\boldsymbol{f} = (f_0 + \beta y)\widehat{\mathbf{k}}$. In component form this equation becomes

$$\frac{\mathrm{D}u}{\mathrm{D}t} - fv = -\frac{1}{\rho}\frac{\partial p}{\partial x}, \qquad \frac{\mathrm{D}v}{\mathrm{D}t} + fu = -\frac{1}{\rho}\frac{\partial p}{\partial y}. \tag{2.83a,b}$$

The mass conservation, thermodynamic and hydrostatic equations in the β-plane approximation are the same as the usual Cartesian, f-plane, forms of those equations.

2.4 EQUATIONS FOR A STRATIFIED OCEAN: THE BOUSSINESQ APPROXIMATION

The density variations in the ocean are quite small compared to the mean density, and we may exploit this to derive somewhat simpler but still quite accurate equations of motion. Let us first examine how much density does vary in the ocean.

2.4.1 Variation of density in the ocean

The variations of density in the ocean are due to three effects: the compression of water by pressure (which we denote as $\Delta_p\rho$), the thermal expansion of water if its temperature changes ($\Delta_T\rho$), and the haline contraction if its salinity changes ($\Delta_S\rho$). How big are these? An appropriate equation of state to approximately evaluate these effects is the linear one

$$\rho = \rho_0\left[1 - \beta_T(T - T_0) + \beta_S(S - S_0) + \frac{p}{\rho_0 c_s^2}\right], \tag{2.84}$$

where $\beta_T \approx 2\times 10^{-4}\,\mathrm{K}^{-1}$, $\beta_S \approx 10^{-3}\,\mathrm{psu}^{-1}$ and $c_s \approx 1500\,\mathrm{m\,s^{-1}}$ (see the table on page 35). The three effects may then be evaluated as follows.

Pressure compressibility. We have $\Delta_p\rho \approx \Delta p/c_s^2 \approx \rho_0 gH/c_s^2$. where H is the depth and the pressure change is quite accurately evaluated using the hydrostatic approximation. Thus,

$$\frac{|\Delta_p\rho|}{\rho_0} \ll 1 \qquad \text{if} \qquad \frac{gH}{c_s^2} \ll 1, \tag{2.85}$$

or if $H \ll c_s^2/g$. The quantity $c_s^2/g \approx 200\,\mathrm{km}$ is the density scale height of the ocean. Thus, the pressure at the bottom of the ocean (say $H = 10\,\mathrm{km}$ in the deep trenches), enormous as it is, is insufficient to compress the water enough to make a significant change in its density. Changes in density due to dynamical variations of pressure are small if the Mach number is small, and this is also the case.

Thermal expansion. We have $\Delta_T\rho \approx -\beta_T\rho_0\Delta T$ and therefore

$$\frac{|\Delta_T\rho|}{\rho_0} \ll 1 \qquad \text{if} \qquad \beta_T\Delta T \ll 1. \tag{2.86}$$

For $\Delta T = 20\,\mathrm{K}$, $\beta_T\Delta T \approx 4\times 10^{-3}$, and evidently we would require temperature differences of order β_T^{-1}, or 5000 K to obtain order one variations in density.

Saline contraction. We have $\Delta_S\rho \approx \beta_S\rho_0\Delta S$ and therefore

$$\frac{|\Delta_S\rho|}{\rho_0} \ll 1 \qquad \text{if} \qquad \beta_S\Delta S \ll 1. \tag{2.87}$$

As changes in salinity in the ocean rarely exceed 5 psu, for which $\beta_S\Delta S = 5\times 10^{-3}$, the fractional change in the density of seawater is correspondingly very small.

Evidently, fractional density changes in the ocean are very small.

2.4.2 The Boussinesq equations

The *Boussinesq equations* are a set of equations that exploit the smallness of density variations in many liquids.[4] To set notation we write

$$\rho = \rho_0 + \delta\rho(x,y,z,t) \tag{2.88a}$$

$$= \rho_0 + \hat{\rho}(z) + \rho'(x,y,z,t) \tag{2.88b}$$

$$= \tilde{\rho}(z) + \rho'(x,y,z,t), \tag{2.88c}$$

where ρ_0 is a constant and we assume that

$$|\hat{\rho}|, |\rho'|, |\delta\rho| \ll \rho_0. \tag{2.89}$$

We need not assume that $|\rho'| \ll |\hat{\rho}|$, but this is often the case in the ocean. To obtain the Boussinesq equations we will just use (2.88a), but (2.88c) will be useful for the anelastic equations considered later.

Associated with the reference density is a reference pressure that is defined to be in hydrostatic balance with it. That is,

$$p = p_0(z) + \delta p(x, y, z, t) \tag{2.90a}$$

$$= \tilde{p}(z) + p'(x, y, z, t), \tag{2.90b}$$

where $|\delta p| \ll p_0$, $|p'| \ll \tilde{p}$ and

$$\frac{\mathrm{d}p_0}{\mathrm{d}z} \equiv -g\rho_0, \qquad \frac{\mathrm{d}\tilde{p}}{\mathrm{d}z} \equiv -g\tilde{\rho}. \tag{2.91a,b}$$

Note that $\nabla_z p = \nabla_z p' = \nabla_z \delta p$ and that $p_0 \approx \tilde{p}$ if $|\hat{\rho}| \ll \rho_0$.

Momentum equations

To obtain the Boussinesq equations we use $\rho = \rho_0 + \delta\rho$, and assume $\delta\rho/\rho_0$ is small. Without approximation, the momentum equation can be written as

$$(\rho_0 + \delta\rho)\left(\frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} + 2\boldsymbol{\Omega} \times \boldsymbol{v}\right) = -\nabla\delta p - \frac{\partial p_0}{\partial z}\mathbf{k} - g(\rho_0 + \delta\rho)\mathbf{k}, \tag{2.92}$$

and using (2.91a) this becomes, again without approximation,

$$(\rho_0 + \delta\rho)\left(\frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} + 2\boldsymbol{\Omega} \times \boldsymbol{v}\right) = -\nabla\delta p - g\delta\rho\mathbf{k}. \tag{2.93}$$

If density variations are small this becomes

$$\boxed{\frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} + 2\boldsymbol{\Omega} \times \boldsymbol{v} = -\nabla\phi + b\mathbf{k}}, \tag{2.94}$$

where $\phi = \delta p/\rho_0$ and $b = -g\,\delta\rho/\rho_0$ is the *buoyancy*. Note that we should not and do not neglect the term $g\,\delta\rho$, for there is no reason to believe it to be small ($\delta\rho$ may be small, but g is big). Equation (2.94) is the momentum equation in the Boussinesq approximation, and it is common to say that the Boussinesq approximation ignores all variations of density of a fluid in the momentum equation, except when associated with the gravitational term.

For most large-scale motions in the ocean the *deviation* pressure and density fields are also approximately in hydrostatic balance, and in that case the vertical component of (2.94) becomes

$$\frac{\partial\phi}{\partial z} = b. \tag{2.95}$$

A condition for (2.95) to hold is that vertical accelerations are small *compared to $g\,\delta\rho/\rho_0$, and not compared to the acceleration due to gravity itself.* For more discussion of this point, see section 2.7.

Mass Conservation

The unapproximated mass conservation equation is

$$\frac{\mathrm{D}\delta\rho}{\mathrm{D}t} + (\rho_0 + \delta\rho)\nabla\cdot\boldsymbol{v} = 0. \tag{2.96}$$

Provided that time scales advectively — that is to say that $\mathrm{D}/\mathrm{D}t$ scales in the same way as $\boldsymbol{v}\cdot\nabla$ — then we may approximate this equation by

$$\boxed{\nabla\cdot\boldsymbol{v} = 0}\,, \tag{2.97}$$

which is the same as that for a constant density fluid. This *absolutely does not* allow one to go back and use (2.96) to say that $D\delta\rho/Dt = 0$; the evolution of density is given by the thermodynamic equation in conjunction with an equation of state, and this should not be confused with the mass conservation equation. Note also that in eliminating the time-derivative of density we eliminate the possibility of sound waves.

Thermodynamic equation and equation of state

The Boussinesq equations are closed by the addition of an equation of state, a thermodynamic equation and, as appropriate, a salinity equation. Neglecting salinity for the moment, a useful starting point is to write the thermodynamic equation, (1.116), as

$$\frac{\mathrm{D}\rho}{\mathrm{D}t} - \frac{1}{c_s^2}\frac{\mathrm{D}p}{\mathrm{D}t} = \frac{\dot{Q}}{(\partial\eta/\partial\rho)_p T} \approx -\dot{Q}\left(\frac{\rho_0\beta_T}{c_p}\right) \tag{2.98}$$

using $(\partial\eta/\partial\rho)_p = (\partial\eta/\partial T)_p(\partial T/\partial\rho)_p \approx c_p/(T\rho_0\beta_T)$. Given the expansions (2.88a) and (2.90a), (2.98) can be written to a good approximation as

$$\frac{\mathrm{D}\delta\rho}{\mathrm{D}t} - \frac{1}{c_s^2}\frac{\mathrm{D}p_0}{\mathrm{D}t} = -\dot{Q}\left(\frac{\rho_0\beta_T}{c_p}\right), \tag{2.99}$$

or, using (2.91a),

$$\frac{\mathrm{D}}{\mathrm{D}t}\left(\delta\rho + \frac{\rho_0 g}{c_s^2}z\right) = -\dot{Q}\left(\frac{\rho_0\beta_T}{c_p}\right), \tag{2.100}$$

as in (1.119). The severest approximation to this is to neglect the second term in brackets on the left-hand side, and noting that $b = -g\delta\rho/\rho_0$ we obtain

$$\boxed{\frac{\mathrm{D}b}{\mathrm{D}t} = \dot{b}}\,, \tag{2.101}$$

where $\dot{b} = g\beta_T\dot{Q}/c_p$. The momentum equation (2.94), mass continuity equation (2.97) and thermodynamic equation (2.101) then form a closed set, called the *simple Boussinesq equations.*

A somewhat more accurate approach is to include the compressibility of the fluid that is due to the hydrostatic pressure. From (2.100), the potential density is given by $\delta\rho_{\mathrm{pot}} = \delta\rho + \rho_0 gz/c_s^2$, and so the *potential buoyancy,* that is the buoyancy based on potential density, is given by

$$b_\sigma \equiv -g\frac{\delta\rho_{\mathrm{pot}}}{\rho_0} = -\frac{g}{\rho_0}\left(\delta\rho + \frac{\rho_0 gz}{c_s^2}\right) = b - g\frac{z}{H_\rho}, \tag{2.102}$$

where $H_\rho = c_s^2/g$. The thermodynamic equation, (2.100), may then be written

$$\frac{\mathrm{D}b_\sigma}{\mathrm{D}t} = \dot{b}_\sigma, \tag{2.103}$$

where $\dot{b}_\sigma = \dot{b}$. Buoyancy itself is obtained from b_σ by the 'equation of state', $b = b_\sigma + gz/H_\rho$.

In many applications we may need to use a still more accurate equation of state. In that case (and see section 1.6.2) we replace (2.101) by the thermodynamic equations

$$\boxed{\frac{\mathrm{D}\theta}{\mathrm{D}t} = \dot{\theta}, \qquad \frac{\mathrm{D}S}{\mathrm{D}t} = \dot{S}}, \tag{2.104a,b}$$

where θ is the potential temperature and S is salinity, along with an equation of state. The equation of state has the general form $b = b(\theta, S, p)$, but to be consistent with the level of approximation in the other Boussinesq equations we replace p by the hydrostatic pressure calculated with the reference density, that is by $-\rho_0 gz$, and the equation of state then takes the general form

$$\boxed{b = b(\theta, S, z)}. \tag{2.105}$$

An example of (2.105) is (1.156), taken with the definition of buoyancy $b = -g\delta\rho/\rho_0$. The closed set of equations (2.94), (2.97), (2.104) and (2.105) are the *general Boussinesq equations*. Using an accurate equation of state and the Boussinesq approximation is the procedure used in many comprehensive ocean general circulation models. The Boussinesq equations, which with the hydrostatic and traditional approximations are often considered to be the oceanic primitive equations, are summarized in the shaded box on the next page.

* *Mean stratification and the buoyancy frequency*

The processes that cause density to vary in the vertical often differ from those that cause it to vary in the horizontal. For this reason it is sometimes useful to write $\rho = \rho_0 + \widehat{\rho}(z) + \rho'(x, y, z, t)$ and define $\widetilde{b}(z) \equiv -g\widehat{\rho}/\rho_0$ and $b' \equiv -g\rho'/\rho_0$. Using the hydrostatic equation to evaluate pressure, the thermodynamic equation (2.98) becomes, to a good approximation,

$$\frac{\mathrm{D}b'}{\mathrm{D}t} + N^2 w = 0, \tag{2.106}$$

where $\mathrm{D}/\mathrm{D}t$ remains a three-dimensional operator and

$$N^2(z) = \left(\frac{\mathrm{d}\widetilde{b}}{\mathrm{d}z} - \frac{g^2}{c_s^2}\right) = \frac{\mathrm{d}\widetilde{b}_\sigma}{\mathrm{d}z}, \tag{2.107}$$

where $\widetilde{b}_\sigma = \widetilde{b} - gz/H_\rho$. The quantity N^2 is a measure of the mean stratification of the fluid, and is equal to the vertical gradient of the mean potential buoyancy. N is known as the buoyancy frequency, something we return to in section 2.9. Equations (2.106) and (2.107) also hold in the simple Boussinesq equations, but with $c_s^2 = \infty$.

Summary of Boussinesq Equations

The simple Boussinesq equations are, for an inviscid fluid:

momentum equations: $$\frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} + \boldsymbol{f} \times \boldsymbol{v} = -\nabla\phi + b\mathbf{k}, \tag{B.1}$$

mass conservation: $$\nabla \cdot \boldsymbol{v} = 0, \tag{B.2}$$

buoyancy equation: $$\frac{\mathrm{D}b}{\mathrm{D}t} = \dot{b}. \tag{B.3}$$

A more general form replaces the buoyancy equation by:

thermodynamic equation: $$\frac{\mathrm{D}\theta}{\mathrm{D}t} = \dot{\theta}, \tag{B.4}$$

salinity equation: $$\frac{\mathrm{D}S}{\mathrm{D}t} = \dot{S}, \tag{B.5}$$

equation of state: $$b = b(\theta, S, \phi). \tag{B.6}$$

Energy conservation is only assured if $b = b(\theta, S, z)$.

2.4.3 Energetics of the Boussinesq system

In a uniform gravitational field but with no other forcing or dissipation, we write the simple Boussinesq equations as

$$\frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} + 2\boldsymbol{\Omega} \times \boldsymbol{v} = b\mathbf{k} - \nabla\phi, \qquad \nabla \cdot \boldsymbol{v} = 0, \qquad \frac{\mathrm{D}b}{\mathrm{D}t} = 0. \tag{2.108a,b,c}$$

From (2.108a) and (2.108b) the kinetic energy density evolution (cf. section 1.10) is given by

$$\frac{1}{2}\frac{\mathrm{D}v^2}{\mathrm{D}t} = bw - \nabla \cdot (\phi\boldsymbol{v}), \tag{2.109}$$

where the constant reference density ρ_0 is omitted. Let us now define the potential $\Phi \equiv -z$, so that $\nabla\Phi = -\mathbf{k}$ and

$$\frac{\mathrm{D}\Phi}{\mathrm{D}t} = \nabla \cdot (\boldsymbol{v}\Phi) = -w, \tag{2.110}$$

and using this and (2.108c) gives

$$\frac{\mathrm{D}}{\mathrm{D}t}(b\Phi) = -wb. \tag{2.111}$$

Adding (2.111) to (2.109) and expanding the material derivative gives

$$\frac{\partial}{\partial t}\left(\frac{1}{2}\boldsymbol{v}^2 + b\Phi\right) + \nabla \cdot \left[\boldsymbol{v}\left(\frac{1}{2}\boldsymbol{v}^2 + b\Phi + \phi\right)\right] = 0. \tag{2.112}$$

This constitutes an energy equation for the Boussinesq system, and may be compared to (1.186). (Also see problem 2.14.) The energy density (divided by ρ_0) is just $\boldsymbol{v}^2/2 + b\Phi$. What does the term $b\Phi$ represent? Its integral, multiplied by ρ_0, is the potential energy of the flow minus that of the basic state, or $\int g(\rho - \rho_0)z\,\mathrm{d}z$. If there were a heating term on the right-hand side of (2.108c) this would directly provide a source of potential energy, rather than internal energy as in the compressible system. Because the fluid is incompressible, there is no conversion from kinetic and potential energy into internal energy.

** Energetics with a general equation of state*

Now consider the energetics of the general Boussinesq equations. Suppose first that we allow the equation of state to be a function of pressure; the equations are motion are then (2.108) except that (2.108c) is replaced by

$$\frac{\mathrm{D}\theta}{\mathrm{D}t} = 0, \qquad \frac{\mathrm{D}S}{\mathrm{D}t} = 0, \qquad b = b(\theta, S, \phi). \tag{2.113a,b,c}$$

A little algebraic experimentation will reveal that no energy conservation law of the form (2.112) generally exists for this system! The problem arises because, by requiring the fluid to be incompressible, we eliminate the proper conversion of internal energy to kinetic energy. However, if we use the approximation $b = b(\theta, S, z)$, the system does conserve an energy, as we now show.[5]

Define the potential, Π, as the integral of b at constant potential temperature and salinity; that is

$$\Pi(\theta, S, z) \equiv -\int_a^z b\,\mathrm{d}z', \tag{2.114}$$

where a is any constant, so that $\partial\Pi/\partial z = -b$. Taking the material derivative of the left-hand side gives

$$\frac{\mathrm{D}\Pi}{\mathrm{D}t} = \left(\frac{\partial \Pi}{\partial \theta}\right)_{S,z} \frac{\mathrm{D}\theta}{\mathrm{D}t} + \left(\frac{\partial \Pi}{\partial S}\right)_{\theta,z} \frac{\mathrm{D}S}{\mathrm{D}t} + \left(\frac{\partial \Pi}{\partial z}\right)_{\theta,S} \frac{\mathrm{D}z}{\mathrm{D}t} = -bw, \tag{2.115}$$

using (2.113a,b). Combining (2.115) and (2.109) gives

$$\frac{\partial}{\partial t}\left(\frac{1}{2}\boldsymbol{v}^2 + \Pi\right) + \nabla \cdot \left[\boldsymbol{v}\left(\frac{1}{2}\boldsymbol{v}^2 + \Pi + \phi\right)\right] = 0. \tag{2.116}$$

Thus, energetic consistency is maintained with an arbitrary equation of state, provided the pressure is replaced by a function of z. If b is not an explicit function of z in the equation of state, the conservation law is identical to (2.112).

2.5 EQUATIONS FOR A STRATIFIED ATMOSPHERE: THE ANELASTIC APPROXIMATION

2.5.1 Preliminaries

In the atmosphere the density varies significantly, especially in the vertical. However deviations of both ρ and p from a statically balanced state are often quite small, and the relative vertical variation of potential temperature is also small. We can usefully exploit these observations to give a somewhat simplified set of equations, useful both for theoretical and numerical analyses because sound waves are eliminated by way of an ‘anelastic’ approximation.[6] To begin we set

$$\rho = \widetilde{\rho}(z) + \delta\rho(x, y, z, t), \qquad p = \widetilde{p}(z) + \delta p(x, y, z, t), \tag{2.117a,b}$$

where we assume that $|\delta\rho| \ll |\widetilde{\rho}|$ and we define $\widetilde{p}$ such that

$$\frac{\partial \widetilde{p}}{\partial z} \equiv -g\widetilde{\rho}(z). \tag{2.118}$$

The notation is similar to that for the Boussinesq case except that, importantly, the density basic state is now a (given) function of vertical coordinate. As with the Boussinesq case, the idea is to ignore dynamic variations of density (i.e., of $\delta\rho$) except where associated with gravity. First recall a couple of ideal gas relationships involving potential temperature, θ, and entropy s (divided by c_p, so $s \equiv \log\theta$), namely

$$s \equiv \log\theta = \log T - \frac{R}{c_p}\log p = \frac{1}{\gamma}\log p - \log\rho, \tag{2.119}$$

where $\gamma = c_p/c_v$, implying

$$\delta s = \frac{1}{\theta}\delta\theta = \frac{1}{\gamma}\frac{\delta p}{p} - \frac{\delta\rho}{\rho} \approx \frac{1}{\gamma}\frac{\delta p}{\widetilde{p}} - \frac{\delta\rho}{\widetilde{\rho}}. \tag{2.120}$$

Further, if $\widetilde{s} \equiv \gamma^{-1}\log\widetilde{p} - \log\widetilde{\rho}$ then

$$\frac{\mathrm{d}\widetilde{s}}{\mathrm{d}z} = \frac{1}{\gamma\widetilde{p}}\frac{\mathrm{d}\widetilde{p}}{\mathrm{d}z} - \frac{1}{\widetilde{\rho}}\frac{\mathrm{d}\widetilde{\rho}}{\mathrm{d}z} = -\frac{g\widetilde{\rho}}{\gamma\widetilde{p}} - \frac{1}{\widetilde{\rho}}\frac{\mathrm{d}\widetilde{\rho}}{\mathrm{d}z}. \tag{2.121}$$

In the atmosphere, the left-hand side is, typically, much smaller than either of the two terms on the right-hand side.

2.5.2 The momentum equation

The exact inviscid horizontal momentum equation is

$$(\widetilde{\rho} + \rho')\left(\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f}\times\boldsymbol{u}\right) = -\nabla_z\delta p. \tag{2.122}$$

Neglecting ρ' where it appears with $\widetilde{\rho}$ leads to

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f}\times\boldsymbol{u} = -\nabla_z\phi, \tag{2.123}$$

where $\phi = \delta p/\widetilde{\rho}$, and this is similar to the corresponding equation in the Boussinesq approximation.

The vertical component of the inviscid momentum equation is, without approximation,

$$(\widetilde{\rho} + \delta\rho)\frac{\mathrm{D}w}{\mathrm{D}t} = -\frac{\partial\widetilde{p}}{\partial z} - \frac{\partial\delta p}{\partial z} - g\widetilde{\rho} - g\delta\rho = -\frac{\partial\delta p}{\partial z} - g\delta\rho. \tag{2.124}$$

using (2.118). Neglecting $\delta\rho$ on the left-hand side we obtain

$$\frac{\mathrm{D}w}{\mathrm{D}t} = -\frac{1}{\widetilde{\rho}}\frac{\partial\delta p}{\partial z} - g\frac{\delta\rho}{\widetilde{\rho}} = -\frac{\partial}{\partial z}\left(\frac{\delta p}{\widetilde{\rho}}\right) - \frac{\delta p}{\widetilde{\rho}^2}\frac{\partial\widetilde{\rho}}{\partial z} - g\frac{\delta\rho}{\widetilde{\rho}}. \tag{2.125}$$

This is not a useful form for a gaseous atmosphere, since the variation of the mean density cannot be ignored. However, we may eliminate $\delta\rho$ in favour of δs using (2.120) to give

$$\frac{\mathrm{D}w}{\mathrm{D}t} = g\delta s - \frac{\partial}{\partial z}\left(\frac{\delta p}{\widetilde{\rho}}\right) - \frac{g}{\gamma}\frac{\delta p}{\widetilde{p}} - \frac{\delta p}{\widetilde{\rho}^2}\frac{\partial\widetilde{\rho}}{\partial z}, \tag{2.126}$$

and using (2.121) gives

$$\frac{\mathrm{D}w}{\mathrm{D}t} = g\delta s - \frac{\partial}{\partial z}\left(\frac{\delta p}{\widetilde{\rho}}\right) + \frac{\mathrm{d}\widetilde{s}}{\mathrm{d}z}\frac{\delta p}{\widetilde{\rho}}. \tag{2.127}$$

What have these manipulations gained us? Two things:

(i) The gravitational term now involves δs rather than $\delta \rho$ which enables a more direct connection with the thermodynamic equation.

(ii) The potential temperature scale height ($\sim 100\,\mathrm{km}$) in the atmosphere is much larger than the density scale height ($\sim 10\,\mathrm{km}$), and so the last term in (2.127) is small.

The second item thus suggests that we choose our reference state to be one of constant potential temperature (see also problem 2.19). The term $\mathrm{d}\widetilde{s}/\mathrm{d}z$ then vanishes and the vertical momentum equation becomes

$$\boxed{\frac{\mathrm{D}w}{\mathrm{D}t} = g\delta s - \frac{\partial \phi}{\partial z}}, \tag{2.128}$$

where $\phi = \delta p/\widetilde{\rho}$ and $\delta s = \delta\theta/\theta_0$, where θ_0 is a constant. If we define a buoyancy by $b_a \equiv g\delta s = g\delta\theta/\theta_0$, then (2.123) and (2.128) have the same form as the Boussinesq momentum equations, but with a slightly different definition of buoyancy.

2.5.3 Mass conservation

Using (2.117a) the mass conservation equation may be written, without approximation, as

$$\frac{\partial \delta\rho}{\partial t} + \nabla \cdot [(\widetilde{\rho} + \delta\rho)\boldsymbol{v}] = 0. \tag{2.129}$$

We neglect $\delta\rho$ where it appears with $\widetilde{\rho}$ in the divergence term. Further, the local time derivative will be small if time itself is scaled advectively (i.e., $T \sim L/U$ and sound waves do not dominate), giving

$$\nabla \cdot \boldsymbol{u} + \frac{1}{\widetilde{\rho}}\frac{\partial}{\partial z}(\widetilde{\rho} w) = 0. \tag{2.130}$$

It is here that the eponymous 'anelastic approximation' arises: the elastic compressibility of the fluid is neglected, and this serves to eliminate sound waves. For reference, in spherical coordinates the equation is

$$\frac{1}{a\cos\vartheta}\frac{\partial u}{\partial \lambda} + \frac{1}{a\cos\vartheta}\frac{\partial}{\partial\vartheta}(v\cos\vartheta) + \frac{1}{\widetilde{\rho}}\frac{\partial(w\widetilde{\rho})}{\partial z} = 0. \tag{2.131}$$

In an ideal gas, the choice of constant potential temperature determines how the reference density $\widetilde{\rho}$ varies with height. In some circumstances it is convenient to let $\widetilde{\rho}$ be a constant, ρ_0 (effectively choosing a different equation of state), in which case the anelastic equations become identical to the Boussinesq equations, albeit with the buoyancy interpreted in terms of potential temperature in the former and density in the latter. Because of their similarity, the Boussinesq and anelastic approximations are sometimes referred to as the strong and weak Boussinesq approximations, respectively.

2.5.4 Thermodynamic equation

The thermodynamic equation for an ideal gas may be written

$$\frac{\mathrm{D}\ln\theta}{\mathrm{D}t} = \frac{\dot{Q}}{Tc_p}. \tag{2.132}$$

In the anelastic equations, $\theta = \widetilde{\theta} + \delta\theta$, where $\widetilde{\theta}$ is constant, and the thermodynamic equation is

$$\frac{\mathrm{D}\delta s}{\mathrm{D}t} = \frac{\widetilde{\theta}}{Tc_p}\dot{Q}. \tag{2.133}$$

Summarizing, the complete set of anelastic equations, with rotation but with no dissipation or diabatic terms, is

$$\boxed{\begin{gathered}\frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} + 2\boldsymbol{\Omega}\times\boldsymbol{v} = \mathbf{k}b_a - \nabla\phi \\ \frac{\mathrm{D}b_a}{\mathrm{D}t} = 0 \\ \nabla\cdot(\widetilde{\rho}\boldsymbol{v}) = 0\end{gathered}}\,, \tag{2.134a,b,c}$$

where $b_a = g\delta s = g\delta\theta/\widetilde{\theta}$. The main difference between the anelastic and Boussinesq sets of equations is in the mass continuity equation, and when $\widetilde{\rho} = \rho_0 =$ constant the two equation sets are formally identical. However, whereas the Boussinesq approximation is a very good one for ocean dynamics, the anelastic approximation is much less so for large-scale atmosphere flow: the constancy of the reference potential temperature state is not a particularly good approximation, and the deviations in density from its reference profile are not especially small, leading to inaccuracies in the momentum equation. Nevertheless, the anelastic equations have been used very productively in limited area 'large-eddy simulations' where one does not wish to make the hydrostatic approximation but where sound waves are unimportant.[7] The equations also provide a good jumping-off point for theoretical studies and for the still simpler models of chapter 5.

2.5.5 * Energetics of the anelastic equations

Conservation of energy follows in much the same way as for the Boussinesq equations, except that $\widetilde{\rho}$ enters. Take the dot product of (2.134a) with $\widetilde{\rho}\boldsymbol{v}$ to obtain

$$\widetilde{\rho}\frac{\mathrm{D}}{\mathrm{D}t}\left(\frac{1}{2}\boldsymbol{v}^2\right) = -\nabla\cdot(\phi\widetilde{\rho}\boldsymbol{v}) + b_a\widetilde{\rho}w. \tag{2.135}$$

Now, define a potential $\Phi(z)$ such that $\nabla\Phi = -\mathbf{k}$, and so

$$\widetilde{\rho}\frac{\mathrm{D}\Phi}{\mathrm{D}t} = -w\widetilde{\rho}. \tag{2.136}$$

Combining this with the thermodynamic equation (2.134b) gives

$$\widetilde{\rho}\frac{\mathrm{D}(b_a\Phi)}{\mathrm{D}t} = -wb_a\widetilde{\rho}. \tag{2.137}$$

Adding this to (2.135) gives

$$\widetilde{\rho}\frac{\mathrm{D}}{\mathrm{D}t}\left(\frac{1}{2}\boldsymbol{v}^2 + b_a\Phi\right) = -\nabla\cdot(\phi\widetilde{\rho}\boldsymbol{v}), \tag{2.138}$$

or, expanding the material derivative,

$$\frac{\partial}{\partial t}\left[\widetilde{\rho}\left(\frac{1}{2}\boldsymbol{v}^2 + b_a\Phi\right)\right] + \nabla\cdot\left[\widetilde{\rho}\boldsymbol{v}\left(\frac{1}{2}\boldsymbol{v}^2 + b_a\Phi + \phi\right)\right] = 0. \tag{2.139}$$

This equation has the form

$$\frac{\partial E}{\partial t} + \nabla\cdot[\boldsymbol{v}(E + \widetilde{\rho}\phi)] = 0, \tag{2.140}$$

where $E = \widetilde{\rho}(\boldsymbol{v}^2/2 + b_a\Phi)$ is the energy density of the flow. This is a consistent energetic equation for the system, and when integrated over a closed domain the total energy is evidently conserved. The total energy density comprises the kinetic energy and a term $\widetilde{\rho}b_a\Phi$, which is analogous to the potential energy of a simple Boussinesq system. However, it is not exactly equal to potential energy because b_a is the buoyancy based on potential temperature, not density; rather, the term combines contributions from both the internal energy and the potential energy into an enthalpy-like quantity.

2.6 CHANGING VERTICAL COORDINATE

Although using z as a vertical coordinate is a natural choice given our Cartesian worldview, it is not the only option, nor is it always the most useful one. Any variable that has a one-to-one correspondence with z in the vertical, so any variable that varies monotonically with z, could be used; pressure and, perhaps surprisingly, entropy, are common choices. In the atmosphere pressure almost always falls monotonically with height, and using it instead of z provides a useful simplification of the mass conservation and geostrophic relations, as well as a more direct connection with observations, which are often taken at fixed values of pressure. (In the ocean pressure coordinates are essentially almost the same as height coordinates, because density is almost constant.) Entropy seems an exotic vertical coordinate, but it is very useful in adiabatic flow and we consider it in chapter 3.

2.6.1 General relations

First consider a general vertical coordinate, ξ. Any variable Ψ that is a function of the coordinates (x, y, z, t) may be expressed instead in terms of (x, y, ξ, t) by considering z to be function of the independent variables (x, y, ξ, t); that is, we let $\Psi(x, y, \xi, t) = \Psi(x, y, z(x, y, \xi, t), t)$. Derivatives with respect to z and ξ are related by

$$\frac{\partial\Psi}{\partial\xi} = \frac{\partial\Psi}{\partial z}\frac{\partial z}{\partial\xi} \quad\text{and}\quad \frac{\partial\Psi}{\partial z} = \frac{\partial\Psi}{\partial\xi}\frac{\partial\xi}{\partial z}. \tag{2.141a,b}$$

Horizontal derivatives in the two coordinate systems are related by the chain rule,

$$\left(\frac{\partial\Psi}{\partial x}\right)_\xi = \left(\frac{\partial\Psi}{\partial x}\right)_z + \left(\frac{\partial z}{\partial x}\right)_\xi\frac{\partial\Psi}{\partial z}, \tag{2.142}$$

and similarly for time.

The material derivative in ξ coordinates may be derived by transforming the expression in z coordinates using the above expressions (problem 2.22). However, because (x, y, t, ξ) are independent coordinates, and noting that the 'vertical velocity' in ξ coordinates is just $\dot{\xi}$ (i.e., $\mathrm{D}\xi/\mathrm{D}t$, just as the vertical velocity in z coordinates is $w = \mathrm{D}z/\mathrm{D}t$), we can write down

$$\frac{\mathrm{D}\Psi}{\mathrm{D}t} = \frac{\partial \Psi}{\partial t} + \boldsymbol{u} \cdot \nabla_\xi \Psi + \dot{\xi}\frac{\partial \Psi}{\partial \xi}, \tag{2.143}$$

where ∇_ξ is the gradient operator at constant ξ. The operator $\mathrm{D}/\mathrm{D}t$ is physically the same in z or ξ coordinates because it is the total derivative of some property of a fluid parcel, and this is independent of the coordinate system. However, the individual terms within it will differ between coordinate systems.

2.6.2 Pressure coordinates

Let us now transform the ideal gas primitive equations from height coordinates to pressure coordinates, (x, y, p, t). In z coordinates the equations are

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f} \times \boldsymbol{u} = -\frac{1}{\rho}\nabla p, \qquad \frac{\partial p}{\partial z} = -\rho g, \tag{2.144a}$$

$$\frac{\mathrm{D}\theta}{\mathrm{D}t} = 0, \qquad \frac{\mathrm{D}\rho}{\mathrm{D}t} + \rho \nabla \cdot \boldsymbol{v} = 0, \tag{2.144b}$$

where $p = \rho R T$ and $\theta = T\,(p_R/p)^{R/c_p}$, and p_R is the reference pressure. These are respectively the horizontal momentum, hydrostatic, thermodynamic and mass continuity equations. The analogue of the vertical velocity is $\omega \equiv Dp/Dt$, and the advective derivative itself is given by

$$\frac{\mathrm{D}}{\mathrm{D}t} = \frac{\partial}{\partial t} + \boldsymbol{u} \cdot \nabla_p + \omega \frac{\partial}{\partial p}. \tag{2.145}$$

To obtain an expression for the pressure force, now let $\xi = p$ in (2.142) and apply the relationship to p itself to give

$$0 = \left(\frac{\partial p}{\partial x}\right)_z + \left(\frac{\partial z}{\partial x}\right)_p \frac{\partial p}{\partial z}, \tag{2.146}$$

which, using the hydrostatic relationship, gives

$$\left(\frac{\partial p}{\partial x}\right)_z = \rho \left(\frac{\partial \Phi}{\partial x}\right)_p, \tag{2.147}$$

where $\Phi = gz$ is the *geopotential*. Thus, the horizontal pressure force in the momentum equations is

$$\frac{1}{\rho}\nabla_z p = \nabla_p \Phi, \tag{2.148}$$

where the subscripts on the gradient operator indicate that the horizontal derivatives are taken at constant z or constant p. Also, from (2.144a), the hydrostatic equation is just

$$\frac{\partial \Phi}{\partial p} = -\alpha. \tag{2.149}$$

The mass conservation equation simplifies attractively in pressure coordinates, if the hydrostatic approximation is used. Recall that the mass conservation equation can be derived from the material form

$$\frac{\mathrm{D}}{\mathrm{D}t}(\rho\,\delta V) = 0, \tag{2.150}$$

where $\delta V = \delta x\,\delta y\,\delta z$ is a volume element. But by the hydrostatic relationship $\rho\delta z = (1/g)\delta p$ and thus

$$\frac{\mathrm{D}}{\mathrm{D}t}(\delta x\,\delta y\,\delta p) = 0. \tag{2.151}$$

This is completely analogous to the expression for the material conservation of volume in an incompressible fluid, (1.15). Thus, without further ado, we write the mass conservation in pressure coordinates as

$$\nabla_p \cdot \boldsymbol{u} + \frac{\partial \omega}{\partial p} = 0, \tag{2.152}$$

where the horizontal derivative is taken at constant pressure. The primitive equations in pressure coordinates, equivalent to (2.144) in height coordinates, are thus:

$$\boxed{\begin{array}{cc} \dfrac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f}\times\boldsymbol{u} = -\nabla_p\Phi, & \dfrac{\partial\Phi}{\partial p} = -\alpha \\ \dfrac{\mathrm{D}\theta}{\mathrm{D}t} = 0, & \nabla_p\cdot\boldsymbol{u} + \dfrac{\partial\omega}{\partial p} = 0 \end{array}}\,, \tag{2.153}$$

where $\mathrm{D}/\mathrm{D}t$ is given by (2.145). The equation set is completed with the addition of the ideal gas equation and the definition of potential temperature. These equations are isomorphic to the hydrostatic general Boussinesq equations (see the shaded box on page 72) with $z \leftrightarrow -p$, $w \leftrightarrow -\omega$, $\phi \leftrightarrow \Phi$, $b \leftrightarrow \alpha$, and an equation of state $b = b(\theta, z) \leftrightarrow \alpha = \alpha(\theta, p)$. In an ideal gas, for example, $\alpha = -(\theta R/p_R)(p_R/p)^{1/\gamma}$.

The main practical difficulty with the pressure-coordinate equations is the lower boundary condition. Using

$$w \equiv \frac{\mathrm{D}z}{\mathrm{D}t} = \frac{\partial z}{\partial t} + \boldsymbol{u}\cdot\nabla_p z + \omega\frac{\partial z}{\partial p}, \tag{2.154}$$

and (2.149), the boundary condition of $w = 0$ at $z = z_s$ becomes

$$\frac{\partial \Phi}{\partial t} + \boldsymbol{u}\cdot\nabla_p\Phi - \alpha\omega = 0 \tag{2.155}$$

at $p(x, y, z_s, t)$. In theoretical studies, it is common to assume that the lower boundary is in fact a constant pressure surface and simply assume that $\omega = 0$, or sometimes the condition $\omega = -\alpha^{-1}\partial\Phi/\partial t$ is used. For realistic studies (with general circulation models, say) the fact that the level $z = 0$ is not a coordinate surface must be properly accounted for. For this reason, and especially if the lower boundary is uneven because of the presence of topography, so-called *sigma coordinates* are sometimes used, in which the vertical coordinate is chosen so that the lower boundary is a coordinate surface. Sigma coordinates may use height itself as a vertical measure (typical in oceanic applications) or use pressure (typical in atmospheric applications). In the latter case the vertical coordinate is $\sigma = p/p_s$ where $p_s(x, y, t)$ is the surface pressure. The difficulty of applying (2.155) is replaced by a prognostic equation for the surface pressure, derived from the mass conservation equation (problem 2.24).

2.6.3 Log-pressure coordinates

A variant of pressure coordinates is *log-pressure* coordinates, in which the vertical coordinate is $Z = -H \ln(p/p_R)$ where p_R is a reference pressure (say 1000 mb) and H is a constant (for example the scale height RT_s/g) so that Z has units of length. (Uppercase letters are conventionally used for some variables in log-pressure coordinates, and these are not to be confused with scaling parameters.) The 'vertical velocity' for the system is now

$$W \equiv \frac{\mathrm{D}Z}{\mathrm{D}t}, \tag{2.156}$$

and the advective derivative is

$$\frac{\mathrm{D}}{\mathrm{D}t} \equiv \frac{\partial}{\partial t} + \boldsymbol{u} \cdot \nabla_p + W \frac{\partial}{\partial Z}. \tag{2.157}$$

It is straightforward to show (problem 2.25) that the primitive equations of motion in these coordinates are:

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f} \times \boldsymbol{u} = -\nabla_Z \Phi, \qquad \frac{\partial \Phi}{\partial Z} = \frac{RT}{H}, \tag{2.158a}$$

$$\frac{\mathrm{D}\theta}{\mathrm{D}t} = 0, \qquad \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial W}{\partial Z} - \frac{W}{H} = 0. \tag{2.158b}$$

The last equation may be written $\nabla_Z \cdot \boldsymbol{u} + \rho_R^{-1} \partial(\rho_R W)/\partial z = 0$, where $\rho_R = \exp(-z/H)$, so giving a form similar to the mass conservation equation in the anelastic equations. Note that integrating the hydrostatic equation between two pressure levels gives, with $\Phi = gz$,

$$z(p_2) - z(p_1) = \frac{R}{g} \int_{p_1}^{p_2} T \, \mathrm{d} \ln p. \tag{2.159}$$

Thus, the thickness of the layer is proportional to the average temperature of the layer.

2.7 SCALING FOR HYDROSTATIC BALANCE

In this section we consider one of the most fundamental balances in geophysical fluid dynamics, hydrostatic balance, and in the next section we consider another fundamental balance, geostrophic balance. The corresponding states, hydrostasy and geostrophy, are not exactly realized, but their approximate satisfaction has profound consequences on the behaviour of the atmosphere and ocean. We first encountered hydrostatic balance in section 1.3.4; we now look in more detail at the conditions required for it to hold.

2.7.1 Preliminaries

Consider the relative sizes of terms in (2.77c):

$$\frac{W}{T} + \frac{UW}{L} + \frac{W^2}{H} + \Omega U \sim \left| \frac{1}{\rho} \frac{\partial p}{\partial z} \right| + g. \tag{2.160}$$

For most large-scale motion in the atmosphere and ocean the terms on the right-hand side are orders of magnitude larger than those on the left, and therefore must be approximately equal. Explicitly, suppose $W \sim 1\,\mathrm{cm\,s^{-1}}$, $L \sim 10^5\,\mathrm{m}$, $H \sim 10^3\,\mathrm{m}$, $U \sim 10\,\mathrm{m\,s^{-1}}$, $T = L/U$. Then by substituting into (2.160) it seems that the pressure term is the only one which could balance the gravitational term, and we are led to approximate (2.77c) by,

$$\frac{\partial p}{\partial z} = -\rho g. \tag{2.161}$$

This equation, which is a vertical momentum equation, is known as *hydrostatic balance.*

However, (2.161) is not always a useful equation! Let us suppose that the density is a constant, ρ_0 . We can then write the pressure as

$$p(x, y, z, t) = p_0(z) + p'(x, y, z, t), \tag{2.162}$$

where

$$\frac{\partial p_0}{\partial z} \equiv -\rho_0 g. \tag{2.163}$$

That is, p_0 and ρ_0 are in hydrostatic balance. The inviscid vertical momentum equation becomes, without approximation,

$$\frac{\mathrm{D}w}{\mathrm{D}t} = -\frac{1}{\rho_0}\frac{\partial p'}{\partial z}. \tag{2.164}$$

Thus, for constant density fluids, the gravitational term has no dynamical effect: there is no buoyancy force, and the pressure term in the horizontal momentum equations can be replaced by p'. Hydrostatic balance, and in particular (2.163), is certainly not an appropriate vertical momentum equation in this case. If the fluid is stratified, we should therefore subtract off the hydrostatic pressure associated with the mean density before we can determine whether hydrostasy is a useful *dynamical* approximation, accurate enough to determine the horizontal pressure gradients. This is automatic in the Boussinesq equations, where the vertical momentum equation is

$$\frac{\mathrm{D}w}{\mathrm{D}t} = -\frac{\partial \phi}{\partial z} + b, \tag{2.165}$$

and the hydrostatic balance of the basic state is already subtracted out. In the more general equation,

$$\frac{\mathrm{D}w}{\mathrm{D}t} = -\frac{1}{\rho}\frac{\partial p}{\partial z} - g, \tag{2.166}$$

we need to compare the advective term on the left-hand side with the pressure variations arising from horizontal flow in order to determine whether hydrostasy is an appropriate vertical momentum equation. Nevertheless, if we only need to determine the pressure for use in an equation of state then we simply need to compare the sizes of the dynamical terms in (2.77c) with g itself, in order to determine whether a hydrostatic approximation will suffice.

2.7.2 Scaling and the aspect ratio

In a Boussinesq fluid we write the horizontal and vertical momentum equations as

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + \boldsymbol{f} \times \boldsymbol{u} = -\nabla\phi, \qquad \frac{\mathrm{D}w}{\mathrm{D}t} = -\frac{\partial \phi}{\partial z} - b. \tag{2.167a,b}$$

With $\boldsymbol{f} = 0$, (2.167a) implies the scaling

$$\phi \sim U^2. \tag{2.168}$$

If we use mass conservation, $\nabla_z \cdot \boldsymbol{u} + \partial w/\partial z = 0$, to scale vertical velocity then

$$w \sim W = \frac{H}{L}U = \alpha U, \tag{2.169}$$

where $\alpha \equiv H/L$ is the aspect ratio. The advective terms in the vertical momentum equation all scale as

$$\frac{\mathrm{D}w}{\mathrm{D}t} \sim \frac{UW}{L} = \frac{U^2 H}{L^2}. \tag{2.170}$$

Using (2.168) and (2.170) the ratio of the advective term to the pressure gradient term in the vertical momentum equations then scales as

$$\frac{|\mathrm{D}w/\mathrm{D}t|}{|\partial\phi/\partial z|} \sim \frac{U^2 H/L^2}{U^2/H} \sim \left(\frac{H}{L}\right)^2. \tag{2.171}$$

Thus, the condition for hydrostasy, that $|\mathrm{D}w/\mathrm{D}t|/|\partial\phi/\partial z| \ll 1$, is:

$$\boxed{\alpha^2 \equiv \left(\frac{H}{L}\right)^2 \ll 1}. \tag{2.172}$$

The advective term in the vertical momentum may then be neglected. Thus, hydrostatic balance is a *small aspect ratio approximation.*

We can obtain the same result more formally by non-dimensionalizing the momentum equations. Using uppercase symbols to denote scaling values we write

$$\begin{gathered}(x, y) = L(\hat{x}, \hat{y}), \qquad z = H\hat{z}, \qquad \boldsymbol{u} = U\hat{\boldsymbol{u}}, \qquad w = W\widehat{w} = \frac{HU}{L}\widehat{w}, \\ t = T\hat{t} = \frac{L}{U}\hat{t}, \qquad \phi = \Phi\hat{\phi} = U^2\hat{\phi}, \qquad b = B\hat{b} = \frac{U^2}{H}\hat{b},\end{gathered} \tag{2.173}$$

where the hatted variables are non-dimensional and the scaling for w is suggested by the mass conservation equation, $\nabla_z \cdot \boldsymbol{u} + \partial w/\partial z = 0$. Substituting (2.173) into (2.167) (with $\boldsymbol{f} = 0$) gives us the non-dimensional equations

$$\frac{\mathrm{D}\hat{\boldsymbol{u}}}{\mathrm{D}\hat{t}} = -\nabla\hat{\phi}, \qquad \alpha^2\frac{\mathrm{D}\widehat{w}}{\mathrm{D}\hat{t}} = -\frac{\partial\hat{\phi}}{\partial\hat{z}} - \hat{b}, \tag{2.174a,b}$$

where $\mathrm{D}/\mathrm{D}\hat{t} = \partial/\partial\hat{t} + \hat{u}\partial/\partial\hat{x} + \hat{v}\partial/\partial\hat{y} + \widehat{w}\partial/\partial\hat{z}$ and we use the convention that when ∇ operates on non-dimensional quantities the operator itself is non-dimensional. From (2.174b) it is clear that hydrostatic balance pertains when $\alpha^2 \ll 1$.

2.7.3 * Effects of stratification on hydrostatic balance

To include the effects of stratification we need to involve the thermodynamic equation, so let us first write down the complete set of non-rotating dimensional equations:

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} = -\nabla_z\phi, \qquad \frac{\mathrm{D}w}{\mathrm{D}t} = -\frac{\partial\phi}{\partial z} + b', \tag{2.175a,b}$$

$$\frac{\mathrm{D}b'}{\mathrm{D}t} + wN^2 = 0, \qquad \nabla\cdot\boldsymbol{v} = 0. \tag{2.176a,b}$$

We have written, without approximation, $b = b'(x,y,z,t) + \tilde{b}(z)$, with $N^2 = \mathrm{d}\tilde{b}/\mathrm{d}z$; this separation is useful because the horizontal and vertical buoyancy variations may scale in different ways, and often N^2 may be regarded as given. (We have also redefined ϕ by subtracting off a static component in hydrostatic balance with $\tilde{b}$.) We non-dimensionalize (2.176) by first writing

$$\begin{aligned} &(x,y) = L(\hat{x},\hat{y}), \qquad z = H\hat{z}, \qquad \boldsymbol{u} = U\hat{\boldsymbol{u}}, \quad w = W\widehat{w} = \epsilon\frac{HU}{L}\widehat{w}, \\ &t = T\hat{t} = \frac{L}{U}\hat{t}, \qquad \phi = U^2\hat{\phi}, \qquad b' = \Delta b\hat{b} = \frac{U^2}{H}\hat{b}', \qquad N^2 = \overline{N}^2\hat{N}^2, \end{aligned} \tag{2.177}$$

where ϵ is, for the moment, undetermined, $\overline{N}$ is a representative, constant, value of the buoyancy frequency and Δb scales only the horizontal buoyancy variations. Substituting (2.177) into (2.175) and (2.176) gives

$$\frac{\mathrm{D}\hat{\boldsymbol{u}}}{\mathrm{D}\hat{t}} = -\nabla_z\hat{\phi}, \qquad \epsilon\alpha^2\frac{\mathrm{D}\widehat{w}}{\mathrm{D}t} = -\frac{\partial\hat{\phi}}{\partial\hat{z}} + \hat{b}' \tag{2.178a,b}$$

$$\frac{U^2}{\overline{N}^2H^2}\frac{\mathrm{D}\hat{b}'}{\mathrm{D}\hat{t}} + \epsilon\widehat{w}\hat{N}^2 = 0, \qquad \nabla\cdot\hat{\boldsymbol{u}} + \epsilon\frac{\partial\widehat{w}}{\partial\hat{z}} = 0. \tag{2.179a,b}$$

where now $\mathrm{D}/\mathrm{D}\hat{t} = \partial/\partial\hat{t} + \hat{\boldsymbol{u}}\cdot\nabla_z + \epsilon\partial/\partial\hat{z}$. To obtain a non-trivial balance in (2.179a) we choose $\epsilon = U^2/(\overline{N}^2H^2) \equiv Fr^2$, where Fr is the *Froude number,* a measure of the stratification of the flow. The vertical velocity then scales as

$$W = \frac{Fr\,UH}{L} \tag{2.180}$$

and if the flow is highly stratified the vertical velocity will be even smaller than a pure aspect ratio scaling might suggest. (There must, therefore, be some cancellation in horizontal divergence in the mass continuity equation; that is, $|\nabla_z\cdot\boldsymbol{u}| \ll U/L$.) With this choice of ϵ the non-dimensional Boussinesq equations may be written:

$$\frac{\mathrm{D}\hat{\boldsymbol{u}}}{\mathrm{D}\hat{t}} = -\nabla_z\hat{\phi}, \qquad Fr^2\alpha^2\frac{\mathrm{D}\widehat{w}}{\mathrm{D}\hat{t}} = -\frac{\partial\hat{\phi}}{\partial\hat{z}} + \hat{b}' \tag{2.181a,b}$$

$$\frac{\mathrm{D}\hat{b}'}{\mathrm{D}\hat{t}} + \widehat{w}\hat{N}^2 = 0, \qquad \nabla\cdot\hat{\boldsymbol{u}} + Fr^2\frac{\partial\widehat{w}}{\partial\hat{z}} = 0. \tag{2.182a,b}$$

The non-dimensional parameters in the system are the aspect ratio and the Froude number (in addition to $\hat{N}$, but by construction this is just an order one function of z). From (2.181b) condition for hydrostatic balance to hold is evidently that

$$\boxed{Fr^2\alpha^2 \ll 1}\,, \tag{2.183}$$

so generalizing the aspect ratio condition (2.172) to a stratified fluid. Because Fr is a measure of stratification, (2.183) formalizes our intuitive expectation that the more stratified a fluid the more vertical motion is suppressed and therefore the more likely hydrostatic balance is to hold. Also note that (2.183) is equivalent to $U^2/(L^2\overline{N}^2) \ll 1$.

Suppose we solve the hydrostatic equations; that is, we omit the advective derivative in the vertical momentum equation, and by numerical integration we obtain $\boldsymbol{u}$, w and b. This flow is the solution of the non-hydrostatic equations in the small aspect ratio limit. The solution never violates the scaling assumptions, even if w seems large, because we can always rescale the variables in order that condition (2.183) is satisfied.

Why bother with any of this scaling? Why not just say that hydrostatic balance holds when $|Dw/Dt| \ll |\partial\phi/\partial z|$? One reason is that we do not have a good idea of the value of w from direct measurements, and it may change significantly in different oceanic and atmospheric parameter regimes. On the other hand the Froude number and the aspect ratio are familiar non-dimensional parameters with a wide applicability in other contexts, and which we can control in a laboratory setting or estimate in the ocean or atmosphere. Still, in scaling theory it is common that ascertaining which parameters are to be regarded as given and which should be derived is a choice, rather than being set a priori.

2.7.4 Hydrostasy in the ocean and atmosphere

Is the hydrostatic approximation in fact a good one in the ocean and atmosphere?

In the ocean

For the large-scale ocean circulation, let $N \sim 10^{-2}\,\mathrm{s^{-1}}$, $U \sim 0.1\,\mathrm{m\,s^{-1}}$ and $H \sim 1\,\mathrm{km}$. Then $Fr = U/(NH) \sim 10^{-2} \ll 1$. Thus, $Fr^2\alpha^2 \ll 1$ even for unit aspect-ratio motion. In fact, for larger scale flow the aspect ratio is also small; for basin-scale flow $L \sim 10^6\,\mathrm{m}$ and $Fr^2\alpha^2 \sim 0.01^2 \times 0.001^2 = 10^{-10}$ and hydrostatic balance is an extremely good approximation.

For intense convection, for example in the Labrador Sea, the hydrostatic approximation may be less appropriate, because the intense descending plumes may have an aspect ratio (H/L) of one or greater and the stratification is very weak. The hydrostatic condition then often becomes the requirement that the Froude number is small. Representative orders of magnitude are $U \sim W \sim 0.1\,\mathrm{m\,s^{-1}}$, $H \sim 1\,\mathrm{km}$ and $N \sim 10^{-3}\,\mathrm{s^{-1}}$ to $10^{-4}\,\mathrm{s^{-1}}$. For these values Fr ranges between 0.1 and 1, and at the upper end of this range hydrostatic balance is violated.

In the atmosphere

Over much of the troposphere $N \sim 10^{-2}\,\mathrm{s^{-1}}$ so that with $U = 10\,\mathrm{m\,s^{-1}}$ and $H = 1\,\mathrm{km}$ we find $Fr \sim 1$. Hydrostasy is then maintained because the aspect ratio H/L is much less than unity. For larger scale synoptic activity a larger vertical scale is appropriate, and with $H = 10\,\mathrm{km}$ both the Froude number and the aspect ratio are much smaller than one; indeed with $L = 1000\,\mathrm{km}$ we find $Fr^2\alpha^2 \sim 0.1^2 \times 0.1^2 = 10^{-4}$ and the flow is hydrostatic to a very good approximation indeed. However, for smaller scale atmospheric motions associated with fronts and, especially, convection, there can be little expectation that hydrostatic balance will be a good approximation.

For large-scale flows in both atmosphere and ocean, the conceptual simplifications afforded by the hydrostatic approximation can hardly be overemphasized.

Variable	Scaling symbol	Meaning	Atmos. value	Ocean value
(x, y)	L	Horizontal length scale	10^6 m	10^5 m
t	T	Time scale	1 day (10^5 s)	10 days (10^6 s)
(u, v)	U	Horizontal velocity	$10\,\mathrm{m\,s^{-1}}$	$0.1\,\mathrm{m\,s^{-1}}$
	Ro	Rossby number, U/fL	0.1	0.01

Table 2.1 Scales of large-scale flow in atmosphere and ocean. The choices given are representative of large-scale eddying motion in both systems.

2.8 GEOSTROPHIC AND THERMAL WIND BALANCE

We now consider the dominant dynamical balance in the horizontal components of the momentum equation. In the horizontal plane (meaning along geopotential surfaces) we find that the Coriolis term is much larger than the advective terms and the dominant balance is between it and the horizontal pressure force. This balance is called *geostrophic balance,* and it occurs when the Rossby number is small, as we now investigate.

2.8.1 The Rossby number

The *Rossby number* characterizes the importance of rotation in a fluid.[8] It is, essentially, the ratio of the magnitude of the relative acceleration to the Coriolis acceleration, and it is of fundamental importance in geophysical fluid dynamics. It arises from a simple scaling of the horizontal momentum equation, namely

$$\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{u} + \boldsymbol{f} \times \boldsymbol{u} = -\frac{1}{\rho}\nabla_z p, \tag{2.184a}$$

$$U^2/L \qquad fU \tag{2.184b}$$

where U is the approximate magnitude of the horizontal velocity and L is a typical length scale over which that velocity varies. (We assume that $W/H \lesssim U/L$, so that vertical advection does not dominate the advection.) The ratio of the sizes of the advective and Coriolis terms is defined to be the Rossby number,

$$\boxed{Ro \equiv \frac{U}{fL}}. \tag{2.185}$$

If the Rossby number is small then rotation effects are important, and as the values in Table 2.1 indicate this is the case for large-scale flow in both ocean and atmosphere.

Another intuitive way to think about the Rossby number is in terms of time scales. The Rossby number based on a time scale is

$$Ro_T \equiv \frac{1}{fT}, \tag{2.186}$$

where T is a time scale associated with the dynamics at hand. If the time scale is an advective one, meaning that $T \sim L/U$, then this definition is equivalent to (2.185). Now, $f = 2\Omega \sin\vartheta$, where Ω is the angular velocity of the rotating frame and equal to $2\pi \sin\vartheta / T_p$ where T_p is the period of rotation (24 hours). Thus,

$$Ro_T = \frac{T_p}{4\pi T \sin\vartheta} = \frac{T_i}{T}, \tag{2.187}$$

where $T_i = 1/f$ is the 'inertial time scale', about three hours in mid-latitudes. Thus, for phenomena with time scales much longer than this, such as the motion of the Gulf Stream or a mid-latitude atmospheric weather system, the effects of the Earth's rotation can be expected to be important, whereas a short-lived phenomena, such as a cumulus cloud or tornado, may be oblivious to such rotation. The expressions (2.185) and (2.186) of course, just approximate measures of the importance of rotation.

2.8.2 Geostrophic balance

If the Rossby number is sufficiently small in (2.184a) then the rotation term will dominate the nonlinear advection term, and if the time period of the motion scales advectively then the rotation term also dominates the local time derivative. The only term that can then balance the rotation term is the pressure term, and therefore we must have

$$\boldsymbol{f} \times \boldsymbol{u} \approx -\frac{1}{\rho}\nabla_z p, \tag{2.188}$$

or, in Cartesian component form

$$fu \approx -\frac{1}{\rho}\frac{\partial p}{\partial y}, \qquad fv \approx \frac{1}{\rho}\frac{\partial p}{\partial x}. \tag{2.189}$$

This balance is known as *geostrophic balance,* and its consequences are profound, giving geophysical fluid dynamics a special place in the broader field of fluid dynamics. We *define* the geostrophic velocity by

$$\boxed{fu_g \equiv -\frac{1}{\rho}\frac{\partial p}{\partial y}, \qquad fv_g \equiv \frac{1}{\rho}\frac{\partial p}{\partial x}}, \tag{2.190}$$

and for low Rossby number flow $u \approx u_g$ and $v \approx v_g$. In spherical coordinates the geostrophic velocity is

$$fu_g = -\frac{1}{\rho a}\frac{\partial p}{\partial \vartheta}, \qquad fv_g = \frac{1}{a\rho\cos\vartheta}\frac{\partial p}{\partial \lambda}, \tag{2.191}$$

where $f = 2\Omega \sin\vartheta$. Geostrophic balance has a number of immediate ramifications:

- ⋆ Geostrophic flow is parallel to lines of constant pressure (isobars). If $f > 0$ the flow is anticlockwise round a region of low pressure and clockwise around a region of high pressure (see Fig. 2.5).

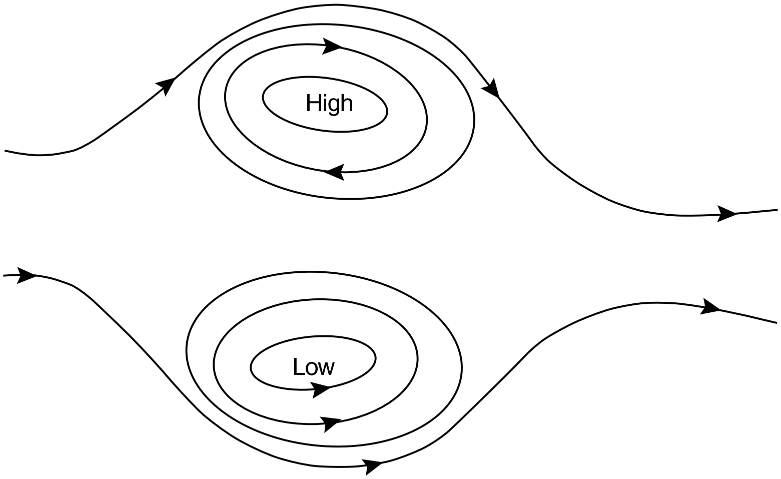


Fig. 2.5 Schematic of geostrophic flow with a positive value of the Coriolis parameter f. Flow is parallel to the lines of constant pressure (isobars). Cyclonic flow is anti-clockwise around a low pressure region and anticyclonic flow is clockwise around a high. If f were negative, as in the Southern Hemisphere, (anti)cyclonic flow would be (anti)clockwise.

- ⋆ If the Coriolis force is constant and if the density does not vary in the horizontal the geostrophic flow is horizontally non-divergent and

$$\boxed{\nabla_z \cdot \boldsymbol{u}_g = \frac{\partial u_g}{\partial x} + \frac{\partial v_g}{\partial y} = 0}. \tag{2.192}$$

We may define the *geostrophic streamfunction,* ψ, by

$$\psi \equiv \frac{p}{f_0 \rho_0}, \tag{2.193}$$

whence

$$u_g = -\frac{\partial \psi}{\partial y}, \qquad v_g = \frac{\partial \psi}{\partial x}. \tag{2.194}$$

The vertical component of vorticity, ζ, is then given by

$$\zeta = \mathbf{k} \cdot \nabla \times \boldsymbol{v} = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = \nabla_z^2 \psi. \tag{2.195}$$

- ⋆ If the Coriolis parameter is not constant, then cross-differentiating (2.190) gives, for constant density geostrophic flow,

$$v_g \frac{\partial f}{\partial y} + f \nabla_z \cdot \boldsymbol{u}_g = 0, \tag{2.196}$$

which implies, using mass continuity,

$$\beta v_g = f \frac{\partial w}{\partial z}. \tag{2.197}$$

where $\beta \equiv \partial f/\partial y = 2\Omega\cos\vartheta/a$. This geostrophic vorticity balance is sometimes known as 'Sverdrup balance', although the latter expression is better restricted to the case when the vertical velocity results from external agents, and specifically a wind stress, as considered in chapter 14.

2.8.3 Taylor–Proudman effect

If $\beta = 0$, then (2.197) implies that the vertical velocity is not a function of height. In fact, in that case none of the components of velocity vary with height if density is also constant. To show this, in the limit of zero Rossby number we first write the three-dimensional momentum equation as

$$\boldsymbol{f}_0 \times \boldsymbol{v} = -\nabla\phi - \nabla\chi, \tag{2.198}$$

where $\boldsymbol{f}_0 = 2\boldsymbol{\Omega} = 2\Omega\mathbf{k}$, $\phi = p/\rho_0$, and $\nabla\chi$ represents other potential forces. If $\chi = gz$ then the vertical component of this equation represents hydrostatic balance, and the horizontal components represent geostrophic balance. On taking the curl of this equation, the terms on the right-hand side vanish and the left-hand side becomes

$$(\boldsymbol{f}_0 \cdot \nabla)\boldsymbol{v} - \boldsymbol{f}_0\nabla \cdot \boldsymbol{v} - (\boldsymbol{v} \cdot \nabla)\boldsymbol{f}_0 + \boldsymbol{v}\nabla \cdot \boldsymbol{f}_0 = 0. \tag{2.199}$$

But $\nabla \cdot \boldsymbol{v} = 0$ by mass conservation, and because $\boldsymbol{f}_0$ is constant both $\nabla \cdot \boldsymbol{f}_0$ and $(\boldsymbol{v} \cdot \nabla)\boldsymbol{f}_0$ vanish. Thus

$$(\boldsymbol{f}_0 \cdot \nabla)\boldsymbol{v} = 0, \tag{2.200}$$

which, since $\boldsymbol{f}_0 = f_0\mathbf{k}$, implies $f_0\partial\boldsymbol{v}/\partial z = 0$, and in particular we have

$$\frac{\partial u}{\partial z} = 0, \quad \frac{\partial v}{\partial z} = 0, \quad \frac{\partial w}{\partial z} = 0. \tag{2.201}$$

A different presentation of this argument proceeds as follows. If the flow is exactly in geostrophic and hydrostatic balance then

$$v = \frac{1}{f_0}\frac{\partial\phi}{\partial x}, \qquad u = -\frac{1}{f_0}\frac{\partial\phi}{\partial y}, \qquad \frac{\partial\phi}{\partial z} = -g. \tag{2.202a,b,c}$$

Differentiating (2.202a,b) with respect to z, and using (2.202c) yields

$$\frac{\partial v}{\partial z} = \frac{-1}{f_0}\frac{\partial g}{\partial x} = 0, \quad \frac{\partial u}{\partial z} = \frac{1}{f_0}\frac{\partial g}{\partial y} = 0. \tag{2.203}$$

Noting that the geostrophic velocities are horizontally non-divergent ($\nabla_z \cdot \boldsymbol{u} = 0$), and using mass continuity then gives $\partial w/\partial z = 0$, as before.

If there is a solid horizontal boundary anywhere in the fluid, for example at the surface, then $w = 0$ at that surface and thus $w = 0$ everywhere. Hence the motion occurs in planes that lie perpendicular to the axis of rotation, and the flow is effectively two dimensional. This result is known as the *Taylor–Proudman effect,* namely that for constant density flow in geostrophic and hydrostatic balance the vertical derivatives of the horizontal and the vertical velocities are zero.[9] At zero Rossby number, if the vertical velocity is zero somewhere in the flow, it is zero everywhere in that vertical column; furthermore, the horizontal flow has no vertical shear, and the fluid moves like a slab. The effects of rotation have provided a *stiffening* of the fluid in the vertical.

In neither the atmosphere nor the ocean do we observe precisely such vertically coherent flow, mainly because of the effects of stratification. However, it is typical of geophysical fluid dynamics that the assumptions underlying a derivation are not fully satisfied, yet there are manifestations of it in real flow. Thus, one might have naïvely expected, because $\partial w/\partial z = -\nabla_z \cdot \boldsymbol{u}$, that the scales of the various variables would be related by $W/H \sim U/L$. However, if the flow is rapidly rotating we expect that the horizontal flow will be in near geostrophic balance and therefore nearly divergence free; thus $\nabla_z \cdot \boldsymbol{u} \ll U/L$, and $W \ll HU/L$.

2.8.4 Thermal wind balance

Thermal wind balance arises by combining the geostrophic and hydrostatic approximations, and this is most easily done in the context of the anelastic (or Boussinesq) equations, or in pressure coordinates. For the anelastic equations, geostrophic balance may be written

$$-fv_g = -\frac{\partial \phi}{\partial x} = -\frac{1}{a\cos\vartheta}\frac{\partial \phi}{\partial \lambda}, \qquad fu_g = -\frac{\partial \phi}{\partial y} = -\frac{1}{a}\frac{\partial \phi}{\partial \vartheta}. \tag{2.204a,b}$$

Combining these relations with hydrostatic balance, $\partial\phi/\partial z = b$, gives

$$\boxed{\begin{aligned} -f\frac{\partial v_g}{\partial z} &= -\frac{\partial b}{\partial x} = -\frac{1}{a\cos\lambda}\frac{\partial b}{\partial \lambda} \\ f\frac{\partial u_g}{\partial z} &= -\frac{\partial b}{\partial y} = -\frac{1}{a}\frac{\partial b}{\partial \vartheta} \end{aligned}}\,. \tag{2.205a,b}$$

These equations represent *thermal wind balance,* and the vertical derivative of the geostrophic wind is the 'thermal wind'. Eq. (2.205b) may be written in terms of the zonal angular momentum as

$$\frac{\partial m_g}{\partial z} = -\frac{a}{2\Omega\tan\vartheta}\frac{\partial b}{\partial y}, \tag{2.206}$$

where $m_g = (u_g + \Omega a\cos\vartheta)a\cos\vartheta$. Potentially more accurate than geostrophic balance is the so-called cyclostrophic or gradient-wind balance, which retains a centrifugal term in the momentum equation. Thus, we omit only the material derivative in the meridional momentum equation (2.50b) and obtain

$$2u\Omega\sin\vartheta + \frac{u^2}{a}\tan\vartheta \approx -\frac{\partial \phi}{\partial y} = -\frac{1}{a}\frac{\partial \phi}{\partial \vartheta}. \tag{2.207}$$

For large-scale flow this only differs significantly from geostrophic balance very close to the equator. Combining cyclostrophic and hydrostatic balance gives a modified thermal wind relation, and this takes a simple form when expressed in terms of angular momentum, namely

$$\frac{\partial m^2}{\partial z} \approx -\frac{a^3\cos^3\vartheta}{\sin\vartheta}\frac{\partial b}{\partial y}. \tag{2.208}$$

If the density or buoyancy is constant then there is no shear and (2.205) or (2.208) give the Taylor–Proudman result. But suppose that the temperature falls in the poleward direction. Then thermal wind balance implies that the (eastward) wind will increase with

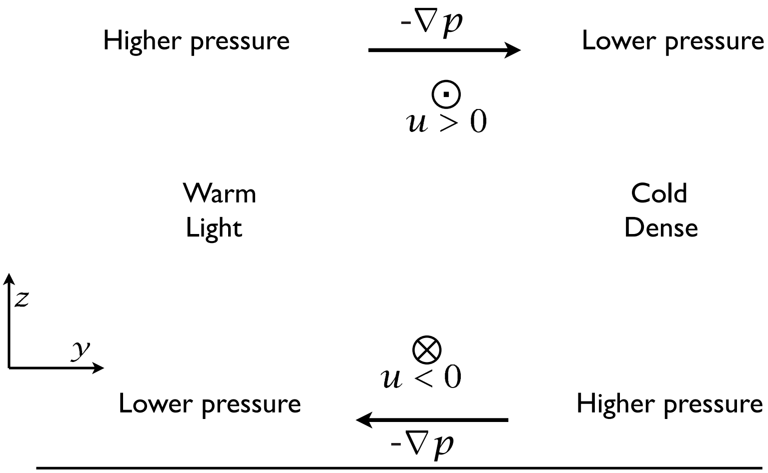


Fig. 2.6 The mechanism of thermal wind. A cold fluid is denser than a warm fluid, so by hydrostasy the vertical pressure gradient is greater where the fluid is cold. Thus, the pressure gradients form as shown, where 'higher' and 'lower' mean relative to the average at that height. The horizontal pressure gradients are balanced by the Coriolis force, producing (for $f > 0$) the horizontal winds shown (⊗ into the paper, and ⊙ out of the paper). Only the wind *shear* is given by the thermal wind.

height — just as is observed in the atmosphere! In general, a vertical shear of the horizontal wind is associated with a horizontal temperature gradient, and this is one of the most simple and far-reaching effects in geophysical fluid dynamics. The underlying physical mechanism is illustrated in Fig. 2.6.

Pressure coordinates

In pressure coordinates geostrophic balance is just

$$\boldsymbol{f} \times \boldsymbol{u}_g = -\nabla_p \Phi, \tag{2.209}$$

where Φ is the geopotential and ∇_p is the gradient operator taken at constant pressure. If f if constant, it follows from (2.209) that the geostrophic wind is non-divergent on pressure surfaces. Taking the vertical derivative of (2.209) (that is, its derivative with respect to p) and using the hydrostatic equation, $\partial\Phi/\partial p = -\alpha$, gives the thermal wind equation

$$\boldsymbol{f} \times \frac{\partial \boldsymbol{u}_g}{\partial p} = \nabla_p \alpha = \frac{R}{p}\nabla_p T, \tag{2.210}$$

where the last equality follows using the ideal gas equation and because the horizontal derivative is at constant pressure. In component form this is

$$-f\frac{\partial v_g}{\partial p} = \frac{R}{p}\frac{\partial T}{\partial x}, \qquad f\frac{\partial u_g}{\partial p} = \frac{R}{p}\frac{\partial T}{\partial y}. \tag{2.211}$$

In log-pressure coordinates, with $Z = -H\ln(p/p_R)$, thermal wind is

$$\boldsymbol{f} \times \frac{\partial \boldsymbol{u}_g}{\partial Z} = -\frac{R}{H}\nabla_Z T. \tag{2.212}$$

The physical meaning in all these cases is the same: a horizontal temperature gradient, or a temperature gradient along an isobaric surface, is accompanied by a vertical shear of the horizontal wind.

2.8.5 * Effects of rotation on hydrostatic balance

Because rotation inhibits vertical motion, we might expect it to affect the requirements for hydrostasy. The simplest setting in which to see this is the rotating Boussinesq equations, (2.167). Let us non-dimensionalize these by writing

$$
\begin{gathered}
(x,y) = L(\hat{x},\hat{y}), \qquad z = H\hat{z}, \qquad \boldsymbol{u} = U\hat{\boldsymbol{u}}, \qquad t = T\hat{t} = \frac{U}{L}\hat{t}, \qquad \boldsymbol{f} = f_0\hat{\boldsymbol{f}}, \\
w = \frac{\beta H U}{f_0}\widehat{w} = \hat{\beta}\frac{HU}{L}\widehat{w}, \qquad \phi = \Phi\hat{\phi} = f_0 U L\hat{\phi}, \qquad b = B\hat{b} = \frac{f_0 u L}{H}\hat{b},
\end{gathered}
\tag{2.213}
$$

where $\hat{\beta} \equiv \beta L/f_0$. (If $\boldsymbol{f}$ is constant, then $\hat{\boldsymbol{f}}$ is a unit vector in the vertical direction.) These relations are the same as (2.173), except for the scaling for w, which is suggested by (2.197), and the scaling for ϕ and b', which are suggested by geostrophic and thermal wind balance.

Substituting into (2.167) we obtain the following scaled momentum equations:

$$
\boxed{Ro\frac{\mathrm{D}\hat{\boldsymbol{u}}}{\mathrm{D}\hat{t}} + \hat{\boldsymbol{f}}\times\hat{\boldsymbol{u}} = -\nabla\hat{\phi}, \qquad Ro\,\hat{\beta}\alpha^2\frac{\mathrm{D}\widehat{w}}{\mathrm{D}\hat{t}} = -\frac{\partial\phi}{\partial z} - \hat{b}}\,. \tag{2.214a,b}
$$

Here, $\mathrm{D}/\mathrm{D}\hat{t} = \partial/\partial\hat{t} + \hat{\boldsymbol{u}}\cdot\nabla_z + \hat{\beta}\partial/\partial\hat{z}$ and $Ro = U/(f_0 L)$. There are two notable aspects to these equations. First and most obviously, when $Ro \ll 1$, (2.214a) reduces to geostrophic balance, $\boldsymbol{f}\times\boldsymbol{u} = -\nabla\hat{\phi}$. Second, the material derivative in (2.214b) is multiplied by three non-dimensional parameters, and we can understand the appearance of each as follows.

(i) The aspect ratio dependence (α^2) arises in the same way as for non-rotating flows — that is, because of the presence of w and z in the vertical momentum equation as opposed to (u,v) and (x,y) in the horizontal equations.

(ii) The Rossby number dependence (*Ro*) arises because in rotating flow the pressure gradient is balanced by the Coriolis force, which is Rossby number larger than the advective terms.

(iii) The factor $\hat{\beta}$ arises because in rotating flow w is smaller than u by the $\hat{\beta}$ times the aspect ratio.

The factor $Ro\,\hat{\beta}\alpha^2$ is very small for large-scale flow; the reader is invited to calculate representative values. Evidently, a rapidly rotating fluid is more likely to be in hydrostatic balance than a non-rotating fluid, other conditions being equal. The combined effects of rotation and stratification are, not surprisingly, quite subtle and we leave that topic for chapter 5.

2.9 STATIC INSTABILITY AND THE PARCEL METHOD

In this and the next couple of sections we consider how a fluid might oscillate if it were perturbed away from a resting state. Our focus is on vertical displacements, and the restoring force is gravity, and we will neglect the effects of rotation, and indeed initially we will neglect horizontal motion entirely. Given that, the simplest and most direct way to approach the problem is to consider from first principles the pressure and gravitational forces on a displaced parcel. To this end, consider a fluid at rest in a constant gravitational field, and therefore in hydrostatic balance. Suppose that a small parcel of the fluid is adiabatically displaced upwards by the small distance δz, without altering the overall pressure field; that is, the fluid parcel instantly assumes the pressure of its environment. If after the displacement the parcel is lighter than its environment, it will accelerate upwards, because the upward pressure gradient force is now greater than the downward gravity force on the parcel; that is, the parcel is *buoyant* (a manifestation of Archimedes' principle) and the fluid is *statically unstable.* If on the other hand the fluid parcel finds itself heavier than its surroundings, the downward gravitational force will be greater than the upward pressure force and the fluid will sink back towards its original position and an oscillatory motion will develop. Such an equilibrium is *statically stable.* Using such simple 'parcel' arguments we will now develop criteria for the stability of the environmental profile.

2.9.1 A simple special case: a density-conserving fluid

Consider first the simple case of an incompressible fluid in which the density of the displaced parcel is conserved, that is $\mathrm{D}\rho/\mathrm{D}t = 0$ (and refer to Fig. 2.7 setting $\rho_\theta = \rho$). If the environmental profile is $\widetilde{\rho}(z)$ and the density of the parcel is ρ then a parcel displaced from a level z [where its density is $\widetilde{\rho}(z)$] to a level $z + \delta z$ [where the density of the parcel is still $\widetilde{\rho}(z)$] will find that its density then differs from its surroundings by the amount

$$\delta\rho = \rho(z+\delta z) - \widetilde{\rho}(z+\delta z) = \widetilde{\rho}(z) - \widetilde{\rho}(z+\delta z) = -\frac{\partial\widetilde{\rho}}{\partial z}\delta z. \tag{2.215}$$

The parcel will be heavier than its surroundings, and therefore the parcel displacement will be stable, if $\partial\widetilde{\rho}/\partial z < 0$. Similarly, it will be unstable if $\partial\widetilde{\rho}/\partial z > 0$. The upward force (per unit volume) on the displaced parcel is given by

$$F = -g\delta\rho = g\frac{\partial\widetilde{\rho}}{\partial z}\delta z, \tag{2.216}$$

and thus Newton's second law implies that the motion of the parcel is determined by

$$\rho(z)\frac{\partial^2\delta z}{\partial t^2} = g\frac{\partial\widetilde{\rho}}{\partial z}\delta z, \tag{2.217}$$

or

$$\frac{\partial^2\delta z}{\partial t^2} = \frac{g}{\widetilde{\rho}}\frac{\partial\widetilde{\rho}}{\partial z}\delta z = -N^2\delta z, \tag{2.218}$$

where

$$N^2 = -\frac{g}{\widetilde{\rho}}\frac{\partial\widetilde{\rho}}{\partial z} \tag{2.219}$$

is the *buoyancy frequency*, or the *Brunt–Väisälä frequency,* for this problem. If $N^2 > 0$ then a parcel displaced upward is heavier than its surroundings, and thus experiences a restoring force; the density profile is said to be stable and N is the frequency at which the fluid parcel oscillates. If $N^2 < 0$, the density profile is unstable and the parcel continues to ascend and convection ensues. In liquids it is often a good approximation to replace $\widetilde{\rho}$ by ρ_0 in the denominator of (2.219).

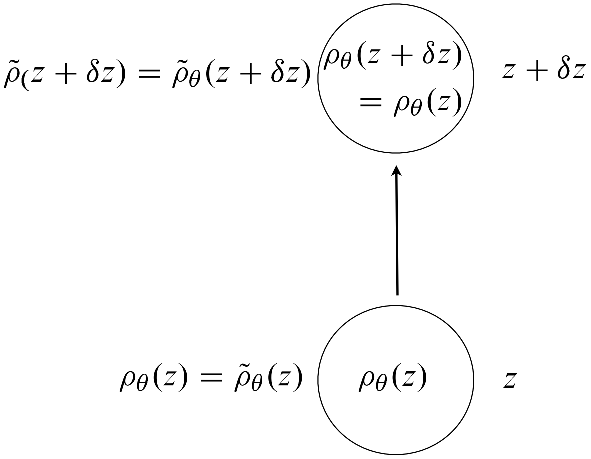

$\tilde{\rho}(z+\delta z) = \tilde{\rho}_\theta(z+\delta z)$
$\rho_\theta(z+\delta z)$
$= \rho_\theta(z)$
$z+\delta z$
$\rho_\theta(z) = \tilde{\rho}_\theta(z)$
$\rho_\theta(z)$
z

Fig. 2.7 A parcel is adiabatically displaced upward from level z to $z + \delta z$, preserving its potential density, which it takes from the environment at level z. If $z + \delta z$ is the reference level, the potential density there is equal to the actual density. The parcel's stability is determined by the difference between its density and the environmental density [see (2.220)]; if the difference is positive the displacement is stable, and conversely.

2.9.2 The general case: using potential density

More generally, in an adiabatic displacement it is *potential density,* ρ_θ, and not density itself that is materially conserved. Consider a parcel that is displaced adiabatically a vertical distance from z to $z + \delta z$; the parcel preserves its potential density, and let us use the pressure at level $z + \delta z$ as the reference level. The *in situ* density of the parcel at $z + \delta z$, namely $\rho(z + \delta z)$, is then equal to its potential density $\rho_\theta(z + \delta z)$ and, because ρ_θ is conserved, this is equal to the potential density of the environment at z, $\tilde{\rho}_\theta(z)$. The difference in *in situ* density between the parcel and the environment at $z + \delta z$, $\delta\rho$, is thus equal to the difference between the potential density of the environment at z and at $z + \delta z$. Putting this together (and see Fig. 2.7) we have

$$\begin{aligned}\delta\rho &= \rho(z+\delta z) - \tilde{\rho}(z+\delta z) = \rho_\theta(z+\delta z) - \tilde{\rho}_\theta(z+\delta z)\\ &= \rho_\theta(z) - \tilde{\rho}_\theta(z+\delta z) = \tilde{\rho}_\theta(z) - \tilde{\rho}_\theta(z+\delta z),\end{aligned} \tag{2.220}$$

and therefore

$$\delta\rho = -\frac{\partial\tilde{\rho}_\theta}{\partial z}\delta z, \tag{2.221}$$

where the derivative on the right-hand side is the environmental gradient of potential density. If the right-hand side is positive, the parcel is heavier than its surroundings and the displacement is stable. Thus, the conditions for stability are:

$$\boxed{\begin{array}{ll}\text{stability}: & \dfrac{\partial\tilde{\rho}_\theta}{\partial z} < 0\\ \text{instability}: & \dfrac{\partial\tilde{\rho}_\theta}{\partial z} > 0\end{array}}\ . \tag{2.222a,b}$$

That is, *the stability of a parcel of fluid is determined by the gradient of the locally-referenced potential density.* The equation of motion of the fluid parcel is

$$\frac{\partial^2\delta z}{\partial t^2} = \frac{g}{\rho}\left(\frac{\partial\tilde{\rho}_\theta}{\partial z}\right)\delta z = -N^2\delta z, \tag{2.223}$$

where, noting that $\rho(z) = \widetilde{\rho}_\theta(z)$ to within $O(\delta z)$,

$$\boxed{N^2 = -\frac{g}{\widetilde{\rho}_\theta}\left(\frac{\partial \widetilde{\rho}_\theta}{\partial z}\right)}. \tag{2.224}$$

This is a general expression for the buoyancy frequency, true in both liquids and gases. It is important to realize that the quantity $\widetilde{\rho}_\theta$ is the *locally-referenced* potential density of the environment, as will become more clear below.

An ideal gas

In the atmosphere potential density is related to potential temperature by $\rho_\theta = p_R/(\theta R)$. Using this in (2.224) gives

$$\boxed{N^2 = \frac{g}{\widetilde{\theta}}\left(\frac{\partial \widetilde{\theta}}{\partial z}\right)}, \tag{2.225}$$

where $\widetilde{\theta}$ refers to the environmental profile of potential temperature. The reference value p_R does not appear, and we are free to choose this value arbitrarily — the surface pressure is a common choice. The conditions for stability, (2.222), then correspond to $N^2 > 0$ for stability and $N^2 < 0$ for instability. In the troposphere (the lowest several kilometres of the atmosphere) the average N is about $0.01\,\mathrm{s}^{-1}$, with a corresponding period, $(2\pi/N)$, of about 10 minutes. In the stratosphere (which lies above the troposphere) N^2 is a few times higher than this.

A liquid ocean

No simple, accurate, analytic expression is available for computing static stability in the ocean. If the ocean had no salt, then the potential density referenced to the surface would generally be a measure of the sign of stability of a fluid column, if not of the buoyancy frequency. However, in the presence of salinity, the surface-referenced potential density is not necessarily even a measure of the sign of stability, because the coefficients of compressibility β_T and β_S vary in different ways with pressure. To see this, suppose two neighbouring fluid elements at the surface have the same potential density, but different salinities and temperatures, and displace them both adiabatically to the deep ocean. Although their potential densities (referenced to the surface) are still equal, we can say little about their actual densities, and hence their stability relative to each other, without doing a detailed calculation because they will each have been compressed by different amounts. It is the profile of the *locally-referenced* potential density that determines the stability.

An approximate expression for stability that is sometimes useful arises by noting that in an adiabatic displacement

$$\delta\rho_\theta = \delta\rho - \frac{1}{c_s^2}\delta p = 0. \tag{2.226}$$

If the fluid is hydrostatic $\delta p = -\rho g \delta z$ so that if a parcel is displaced adiabatically its density changes according to

$$\left(\frac{\partial \rho}{\partial z}\right)_{\rho_\theta} = -\frac{\rho g}{c_s^2}. \tag{2.227}$$

and so, if T and L_c are assumed to be constant,

$$\theta_{eq} = \theta \exp\left(\frac{L_c w}{c_p T}\right). \tag{2.243}$$

The equivalent potential temperature so defined is approximately conserved during condensation, the approximation arising going from (2.242) to (2.243). It is a useful expression for diagnostic purposes, and in constructing theories of convection, but it is not accurate enough to use as a prognostic variable in a putatively realistic numerical model. The 'equivalent temperature' may be defined in terms of the equivalent potential temperature by

$$T_{eq} \equiv \theta_{eq}\left(\frac{p}{p_R}\right)^{\kappa}. \tag{2.244}$$

2.10 GRAVITY WAVES

The parcel approach to oscillations and stability, while simple and direct, is divorced from the fluid-dynamical equations of motion, making it hard to include other effects such as rotation, or to explore the effects of possible differences between the hydrostatic and non-hydrostatic cases. To remedy this, we now use the equations of motion to analyse the motion resulting from a small disturbance.

2.10.1 Gravity waves and convection in a Boussinesq fluid

Let us consider a Boussinesq fluid, at rest, in which the buoyancy varies linearly with height and the buoyancy frequency, N, is a constant. Linearizing the equations of motion about this basic state gives the linear momentum equations,

$$\frac{\partial u'}{\partial t} = -\frac{\partial \phi'}{\partial x}, \qquad \frac{\partial w'}{\partial t} = -\frac{\partial \phi'}{\partial z} + b', \tag{2.245a,b}$$

the mass continuity and thermodynamic equations,

$$\frac{\partial u'}{\partial x} + \frac{\partial w'}{\partial z} = 0, \qquad \frac{\partial b'}{\partial t} + w'N^2 = 0, \tag{2.246a,b}$$

where for simplicity we assume that the flow is a function only of x and z. A little algebra gives a single equation for w',

$$\left[\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2}\right)\frac{\partial^2}{\partial t^2} + N^2\frac{\partial^2}{\partial x^2}\right] w' = 0. \tag{2.247}$$

Seeking solutions of the form $w' = \mathrm{Re}\, W \exp[\mathrm{i}(kx + mz - \omega t)]$ (where Re denotes the real part) yields the dispersion relationship for gravity waves:

$$\boxed{\omega^2 = \frac{k^2 N^2}{k^2 + m^2}}. \tag{2.248}$$

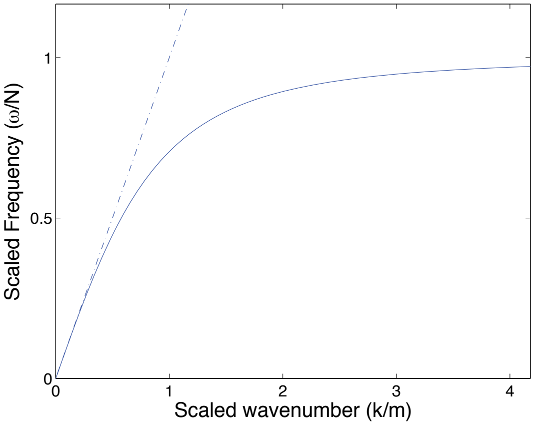


Fig. 2.8 Scaled frequency, ω/N, plotted as a function of scaled horizontal wavenumber, k/m, using the full dispersion relation of (2.248) (solid line, asymptoting to unit value for large k/m) and with the hydrostatic dispersion relation (2.252) (dashed line, tending to ∞ for large k/m).

The frequency (see Fig. 2.8) is thus always less than N, approaching N for small horizontal scales, $k \gg m$. If we neglect pressure perturbations, as in the parcel argument, then the two equations,

$$\frac{\partial w'}{\partial t} = b', \qquad \frac{\partial b'}{\partial t} + w'N^2 = 0, \tag{2.249}$$

form a closed set, and give $\omega^2 = N^2$.

If the basic state density increases with height then $N^2 < 0$ and we expect this state to be unstable. Indeed, the disturbance grows exponentially according to $\exp(\sigma t)$ where

$$\sigma = \mathrm{i}\omega = \frac{\pm k\widetilde{N}}{(k^2 + m^2)^{1/2}}, \tag{2.250}$$

where $\widetilde{N}^2 = -N^2$. Most convective activity in the ocean and atmosphere is, ultimately, related to an instability of this form, although of course there are many complicating issues — water vapour in the atmosphere, salt in the ocean, the effects of rotation and so forth.

Hydrostatic gravity waves and convection

Let us now suppose that the fluid satisfies the hydrostatic Boussinesq equations. The linearized two-dimensional equations of motion become

$$\frac{\partial u'}{\partial t} = -\frac{\partial \phi'}{\partial x}, \qquad 0 = -\frac{\partial \phi'}{\partial z} + b', \tag{2.251a}$$

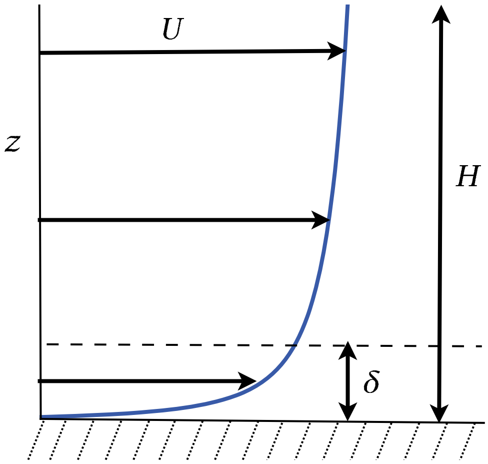
$$\frac{\partial u'}{\partial x} + \frac{\partial w'}{\partial z} = 0, \qquad \frac{\partial b'}{\partial t} + w'N^2 = 0, \tag{2.251b}$$

where these are the horizontal and vertical momentum equations, the mass continuity equation and the thermodynamic equation respectively. A little algebra gives the dispersion relation,

$$\omega^2 = \frac{k^2 N^2}{m^2}. \tag{2.252}$$

The frequency and, if N^2 is negative the growth rate, is unbounded for as $k/m \to \infty$, and the hydrostatic approximation thus has quite unphysical behaviour for small horizontal scales (see also problem 2.11).[11]

Fig. 2.10 An idealized boundary layer. The values of a field, such as velocity, U, may vary rapidly in a boundary in order to satisfy the boundary conditions at a rigid surface. The parameter δ is a measure of the boundary layer thickness, and H is a typical scale of variation away from the boundary.



2.12 THE EKMAN LAYER

In the final topic of this chapter, we return to geostrophic flow and consider the effects of friction. The fluid fields in the interior of a domain are often set by different physical processes than those occurring at a boundary, and consequently often change rapidly in a thin *boundary layer,* as in Fig. 2.10. Such boundary layers nearly always involve one or both of viscosity and diffusion, because these appear in the terms of highest differential order in the equations of motion, and so are responsible for the number and type of boundary conditions that the equations must satisfy — for example, the presence of molecular viscosity leads to the condition that the tangential flow (as well as the normal flow) must vanish at a rigid surface.

In many boundary layers in non-rotating flow the dominant balance in the momentum equation is between the advective and viscous terms. In some contrast, in large-scale atmospheric and oceanic flow the effects of rotation are large, and this results in a boundary layer, known as the *Ekman layer,* in which the dominant balance is between Coriolis and frictional or stress terms.[12] Now, the direct effects of molecular viscosity and diffusion are nearly always negligible at distances more than a few millimetres away from a solid boundary, but it is inconceivable that the entire boundary layer between the free atmosphere (or free ocean) and the surface is only a few millimetres thick. Rather, in practice a balance occurs between the Coriolis terms and the forces due to the stress generated by small-scale turbulent motion, and this gives rise to a boundary layer that has a typical depth of a few tens to several hundreds of metres. Because the stress arises from the turbulence we cannot with confidence determine its precise form; thus, we should try to determine what general properties Ekman layers may have that are *independent* of the precise form of the friction.

The atmospheric Ekman layer occurs near the ground, and the stress at the ground itself is due to the surface wind (and its vertical variation). In the ocean the main Ekman layer is near the surface, and the stress at ocean surface is largely due to the presence of the overlying wind. There is also a weak Ekman layer at the bottom of the ocean, analogous to the atmospheric Ekman layer. To analyse all these layers, let us assume the following.

- ⋆ The Ekman layer is Boussinesq. This is a very good assumption for the ocean, and a reasonable one for the atmosphere if the boundary layer is not too deep.

Part II

INSTABILITIES, WAVE–MEAN FLOW INTERACTION AND TURBULENCE

Oh brave new world, That has such people in't!
William Shakespeare, *The Tempest,* c. 1611.

CHAPTER
NINE

Geostrophic Turbulence and Baroclinic Eddies

GEOSTROPHIC TURBULENCE is turbulence in flows that are stably stratified and in near-geostrophic balance. Like any problem in turbulence it is difficult, and a real 'solution' — meaning an accurate, informative statement about average states, without computation of the detailed evolution — may be out of our reach, and may not exist. Nevertheless, and ironically, it is sometimes easier to say something interesting about geostrophic turbulence than about incompressible isotropic two- or three-dimensional turbulence. In the latter class of problems there is nothing else to understand other than the problem of turbulence itself; on the other hand, rotation and stratification give one something else to grasp, a nettle though it may be, and it becomes possible to address geophysically interesting phenomena without having to solve the whole turbulence problem. Furthermore, in inhomogeneous geostrophic turbulence, asking questions about the *mean fields* is meaningful and useful, whereas this is trivial in isotropic turbulence.

Geostrophic turbulence is not restricted to quasi-geostrophic flow; indeed, the large scale turbulence of the Earth's ocean and atmosphere is sometimes simply called 'macroturbulence'. Nevertheless, the quasi-geostrophic equations retain advective nonlinearity in the vorticity equation, and they capture the constraining effects of rotation and stratification that are so important in geophysical flows in a simple and direct way; for these reasons the quasi-geostrophic equations will be our main tool. Let us consider the effects of rotation first, then stratification.

9.1 EFFECTS OF DIFFERENTIAL ROTATION IN TWO-DIMENSIONAL TURBULENCE

In the limit of motion of a scale much shorter than the deformation radius, and with no topography, the quasi-geostrophic potential vorticity equation, (5.118), reduces to the two-dimensional equation,

$$\frac{\mathrm{D}q}{\mathrm{D}t} = 0, \tag{9.1}$$

where $q = \zeta + f$. This is the perhaps the simplest equation with which to study the effects of rotation on turbulence. Suppose first that the Coriolis parameter is constant, so that $f = f_0$. Then (9.1) becomes simply the two-dimensional vorticity equation

$$\frac{\mathrm{D}\zeta}{\mathrm{D}t} = 0. \tag{9.2}$$

Thus constant rotation has *no* effect on purely two-dimensional motion. Flow that is already two-dimensional — flow on a soap film, for example — is unaffected by rotation. (In the ocean and atmosphere, or in a rotating tank, it is of course the effects of rotation that lead to the flow being quasi-two dimensional in the first instance.)

Suppose, though, that the Coriolis parameter is variable, as in $f = f_0 + \beta y$. Then we have

$$\frac{\mathrm{D}}{\mathrm{D}t}(\zeta + \beta y) = 0 \qquad \text{or} \qquad \frac{\mathrm{D}\zeta}{\mathrm{D}t} + \beta v = 0. \tag{9.3a,b}$$

If the asymptotically dominant term in these equations is the one involving β, then we obtain $v = 0$. Put another way, if β is very large, then the meridional flow v must be correspondingly small to ensure that βv is bounded. Any flow must then be predominantly *zonal*. This constraint may be interpreted as a consequence of angular momentum and energy conservation as follows. A ring of fluid encircling the Earth at a velocity u has an angular momentum per unit mass $a\cos\theta(u + \Omega a\cos\theta)$, where θ is the latitude and a is the radius of the Earth. Moving this ring of air polewards (i.e., giving it a meridional velocity) while conserving its angular momentum requires that its zonal velocity and hence energy must increase. Unless there is a source for that energy the flow is constrained to remain zonal.

9.1.1 The wave–turbulence cross-over

Scaling

Let us now consider how turbulent flow might interact with Rossby waves. We write (9.1) in full as

$$\frac{\partial\zeta}{\partial t} + \boldsymbol{u}\cdot\nabla\zeta + \beta v = 0. \tag{9.4}$$

If $\zeta \sim U/L$ and if $t \sim T$ then the respective terms in this equation scale as

$$\frac{U}{LT} \qquad \frac{U^2}{L^2} \qquad \beta U. \tag{9.5}$$

How time scales (i.e., advectively or with a Rossby wave frequency scaling) is determined by which of the other two terms dominates, and this in turn is scale dependent. For large scales the β-term will be dominant, and at smaller scales the advective term is dominant. The cross-over scale, denoted L_R, is called the *Rhines scale* and is given by[1]

$$\boxed{L_R \sim \sqrt{\frac{U}{\beta}}}\,. \tag{9.6}$$

The U in (9.6) should be interpreted as the root-mean-square velocity at the energy containing scales, not a mean or translational velocity. (We will refer to the specific scale $\sqrt{U/\beta}$ as the Rhines scale, and more general scales involving a balance between nonlinearity and β, discussed below, as the β-scale, and denote them L_β.)

This is not a unique way to arrive at a β-scale, since we have chosen the length scale that connects vorticity to velocity to also be the β-scale, and it is by no means clear a priori that this should be so. If the two scales are different, the three terms in (9.4) scale as

$$\frac{\mathcal{Z}}{T} \;:\; \frac{U\mathcal{Z}}{L} \;:\; \beta U, \tag{9.7}$$

respectively, where $\mathcal{Z}$ is the scaling for vorticity [i.e., $\zeta = \mathcal{O}(\mathcal{Z})$]. Equating the second and third terms gives the scale

$$L_{\beta Z} = \frac{\mathcal{Z}}{\beta}. \tag{9.8}$$

Nevertheless, (9.6) and (9.8) both indicate that at some *large* scale Rossby waves are likely to dominate whereas at small scales advection, and turbulence, dominates.

Another heuristic way to derive (9.6) is by a direct consideration of time scales. Ignoring anisotropy, Rossby wave frequency is β/k and an inverse advective time scale is Uk, where k is the wavenumber. Equating these two gives an equation for the Rhines wavenumber

$$k_R \sim \sqrt{\frac{\beta}{U}}. \tag{9.9}$$

This equation is the inverse of (9.6), but note that factors of order unity cannot be revealed by simple scaling arguments such as these. The cross-over between waves and turbulence is reasonably sharp, as indicated in Fig. 9.1.

Turbulent phenomenology

Let us now try to go beyond elementary scaling arguments and examine wave–turbulence cross-over using the phenomenology of two-dimensional turbulence. We will suppose that the fluid is stirred at some well-defined scale k_f, producing an energy input ε. Then (assuming no energy is lost to smaller scales) energy cascades to large scales at that same rate. At some scale, the β-term in the vorticity equation will start to make its presence felt. By analogy with the procedure for finding the viscous dissipation scale in turbulence, we can find the scale at which linear Rossby waves dominate by equating the inverse of the turbulent eddy turnover time to the Rossby wave frequency. The eddy-turnover time is

$$\tau_k = \varepsilon^{-1/3} k^{-2/3}, \tag{9.10}$$

and equating this to the inverse Rossby wave frequency k/β gives the β-wavenumber and its inverse, the β-scale:

$$\boxed{k_\beta \sim \left(\frac{\beta^3}{\varepsilon}\right)^{1/5}, \qquad L_\beta \sim \left(\frac{\varepsilon}{\beta^3}\right)^{1/5}}. \tag{9.11a,b}$$

In a real fluid these expressions are harder to evaluate than (9.9), since it is generally much easier to measure velocities than energy transfer rates, or even vorticity. On the other hand,

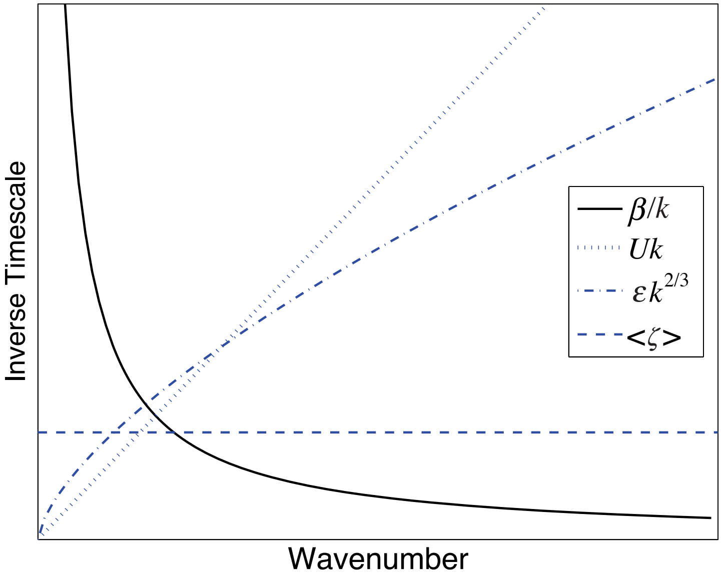


Fig. 9.1 Three estimates of the wave–turbulence cross-over, in wavenumber space. The solid curve is the frequency of Rossby waves, proportional to β/k. The other three curves are various estimates of the inverse turbulence time scale, or 'turbulence frequency'. These are the turbulent eddy transfer rate, proportional to $\varepsilon k^{2/3}$ in a $k^{-5/3}$ spectrum; the simple estimate Uk where U is an root-mean-squared velocity; and the mean vorticity, which is constant. Where the Rossby wave frequency is larger (smaller) than the turbulent frequency, i.e., at large (small) scales, Rossby waves (turbulence) dominate the dynamics.

(9.11) is perhaps more satisfactory from the point of view of turbulence, and ε may be determined by processes largely independent of β, whereas the magnitude of the eddies at the energy containing scales is likely to be a function of β. We also note that the scale given by (9.11b) is not necessarily the energy-containing scale, and may in principle differ considerably from the scale given by (9.9). This is because the inverse cascade is not necessarily *halted* at the scale (9.11b) — this is just scale at which Rossby waves become important. Energy may continue to cascade to larger scales, albeit anisotropically as discussed below, and so the energy containing scale may be larger.

9.1.2 Generation of zonal flows and jets

None of the effects discussed so far takes into account the anisotropy inherent in Rossby waves, and such anisotropy can give rise to predominantly zonal flows and jets. To understand this, let us first note that energy transfer will be relatively inefficient at those scales where linear Rossby waves dominate the dynamics. But the wave–turbulence boundary is not isotropic; the Rossby wave frequency is quite anisotropic, being given by

$$\omega = -\frac{\beta k^x}{{k^x}^2 + {k^y}^2}. \tag{9.12}$$

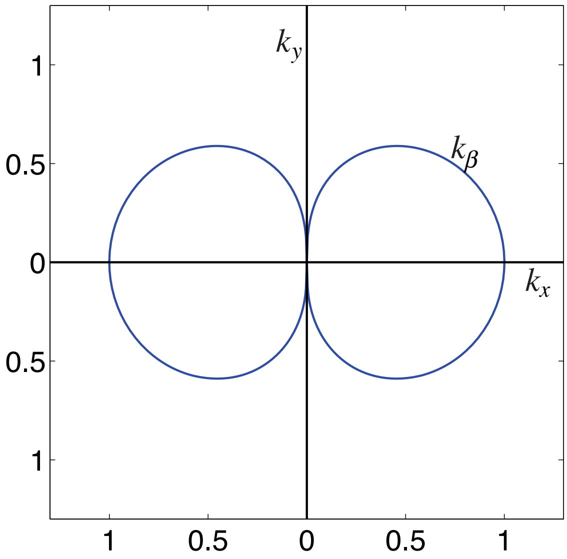


Fig. 9.2 The anisotropic wave–turbulence boundary k_β, in wave-vector space calculated by equating the turbulent eddy transfer rate, proportional to $k^{2/3}$ in a $k^{-5/3}$ spectrum, to the Rossby wave frequency $\beta k^x/k^2$, as in (9.14). Within the dumbbell Rossby waves dominate and energy transfer is inhibited. The inverse cascade plus Rossby waves thus leads to a generation of zonal flow.

If, albeit a little crudely, we suppose that the turbulent part of the flow remains isotropic, the wave–turbulence boundary is then given by equating the inverse of (9.10) with (9.12); that is solution of

$$\varepsilon^{1/3}k^{2/3} = \frac{\beta k^x}{k^2}, \tag{9.13}$$

where k is the isotropic wavenumber. Solving this gives expressions for the x- and y-wavenumber components of the wave–turbulence boundary, namely

$$k^x_\beta = \left(\frac{\beta^3}{\varepsilon}\right)^{1/5}\cos^{8/5}\theta, \qquad k^y_\beta = \left(\frac{\beta^3}{\varepsilon}\right)^{1/5}\sin\theta\cos^{3/5}\theta, \tag{9.14}$$

where the polar coordinate is parameterized by the angle $\theta = \tan^{-1}(k^y/k^x)$. This rather uninformative-looking formula is illustrated in Fig. 9.2. (Slight variations on this theme are produced by using different expressions for the turbulence time scale; see problem 9.1.)

What occurs physically? The region inside the dumbbell shapes in Fig. 9.2 is dominated by Rossby waves, where the natural frequency of the oscillation is *higher* than the turbulent frequency. If the flow is stirred at a wavenumber higher than this the energy will cascade to larger scales, but because of the frequency mismatch the turbulent flow will be unable to efficiently excite modes within the dumbbell. Nevertheless, there is still a natural tendency of the energy to seek the gravest mode, and it will do this by cascading toward the $k^x = 0$ axis; that is, toward zonal flow. Thus, the combination of Rossby waves and turbulence will lead to the formation of zonal flow and, potentially, zonal jets.[3]

Figure 9.3 illustrates this mechanism; it shows the freely evolving (unforced, inviscid) energy spectrum in a simulation on a β-plane, with an initially isotropic spectrum. The energy implodes, cascading to larger scales but avoiding the region inside the dumbbell and piling up at $k^x = 0$. In physical space this mechanism manifests itself by the formation of zonally elongated structures and jets, in both freely-decaying and forced-dissipative simulations (Fig. 9.4 and Fig. 9.5).

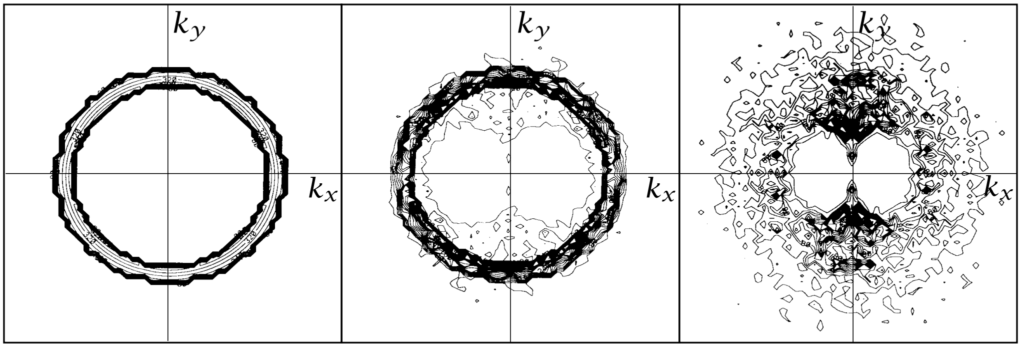


Fig. 9.3 Evolution of the energy spectrum in a freely evolving two-dimensional simulation on the β-plane. The panels show contours of energy in wavenumber (k_x, k_y) space at successive times. The initial spectrum is isotropic. The energy 'implodes', but its passage to large scales is impeded by the β-effect, and second and third panels show the spectrum at later times, illustrating the dumbbell predicted by (9.14) and Fig. 9.2.[2]

9.1.3 † Joint effect of β and friction

The β term does not remove energy from a fluid. Thus, if energy is being added to a fluid at some small scales, and the energy is cascading to larger scales, then the β-effect does not of itself halt the inverse cascade, it merely deflects the cascade such that the flow becomes more zonal. Suppose that the fluid obeys the barotropic vorticity equation,

$$\frac{\partial \zeta}{\partial t} + J(\psi, \zeta) + \beta \frac{\partial \psi}{\partial x} = F - r\zeta + \nu \nabla^2 \zeta, \tag{9.15}$$

where the viscosity, ν, is small and acts only to remove enstrophy, and not energy, at very small scales. The forcing, F, supplies energy at a rate ε and this is cascaded upscale and removed by the linear drag term $-r\zeta$, where the drag coefficient r is a constant. If the friction is sufficiently large, then the energy is removed before it feels the effect of β. A scaling of (9.15) suggests that the relative importance of the β-effect relative to friction is parameterized by the non-dimensional number $\beta L/r$, where L is the length scale of the energy containing modes. This length scale is not known a priori, and a more predictive measure of the importance of friction and β is obtained by comparing the frictional wavenumber (8.74), and the β-wavenumber, (9.11). The ratio of these scales, γ, is given by

$$\gamma = \left(\frac{\beta \varepsilon^{1/2}}{r^{5/2}} \right)^{3/5}. \tag{9.16}$$

When γ is large the β-effect will be felt, but when γ is small frictional effects will dominate at large scales.

Even if γ is large, then frictional effects must still be important somewhere, in order that energy may be removed. Forming an energy equation from (9.15) by multiplying by $-\psi$ and spatially integrating (and neglecting viscosity) we find the energy balance

$$\varepsilon = -\frac{1}{A} \int_A \psi F \, \mathrm{d}A = \frac{r}{A} \int_A (\nabla \psi)^2 \, \mathrm{d}A = 2r\overline{E}, \tag{9.17}$$

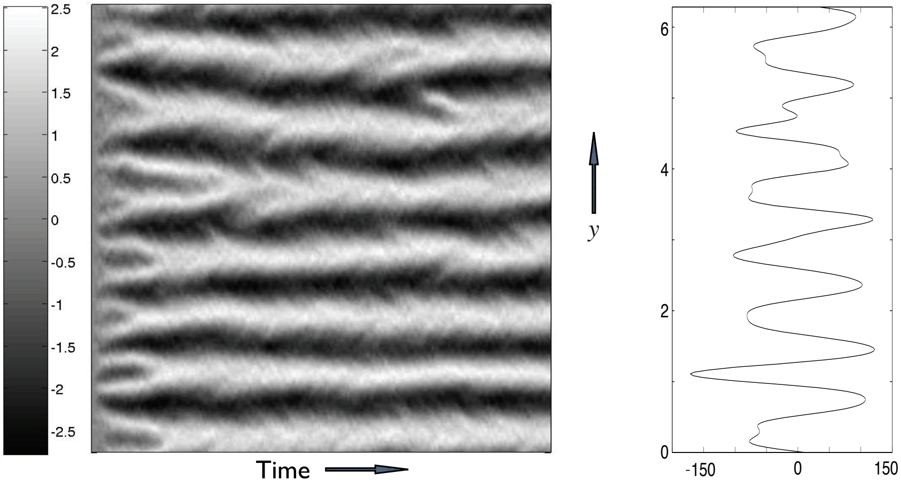


Fig. 9.5 Left: Grey-scale image of zonally averaged zonal velocity ($\overline{u}$) as a function of time and latitude (Y), produced in a simulation forced around wavenumber 80 and with $k_\beta = \sqrt{\beta/U} \approx 10$ (in a square domain of side 2π). Right: Values of $\partial^2\overline{u}/\partial y^2$ as a function of latitude, late in the integration. Jets form very quickly from the random initial conditions, and are subsequently quite steady.[4]

where $\overline{E}$ is the average energy of the fluid per unit mass. Using (9.6) with $U = \sqrt{2\overline{E}}$, where $\overline{E}$ is obtained from (9.17), we obtain the Rhines scale

$$L_R = \left(\frac{\varepsilon}{r\beta^2}\right)^{1/4}. \tag{9.18}$$

What do the two scales, (9.11b) and (9.18), represent? The first is the scale at which the β-effect is first felt by the inverse cascade, and this parameterizes the overall size of the dumbbell of Fig. 9.2. It is this scale that is most relevant for large-scale meridional mixing, and so for the meridional heat transport in the atmosphere. But if the inverse cascade continues past this scale, especially in the zonal direction, then (9.18) may characterize the largest scale reached by the inverse cascade, and so the meridional scale of the jets.[5]

9.2 STRATIFIED GEOSTROPHIC TURBULENCE

9.2.1 An analogue to two-dimensional flow

Now let us consider stratified effects in a simple setting, using the quasi-geostrophic equations with constant Coriolis parameter and constant stratification.[6] The (dimensional) unforced and inviscid governing equation may then be written as

$$\frac{\mathrm{D}q}{\mathrm{D}t} = 0, \qquad q = \nabla^2\psi + Pr^2\frac{\partial^2\psi}{\partial z^2}, \tag{9.19a}$$

where $Pr = f_0/N$ is the *Prandtl ratio* (and Pr/H is the inverse of the deformation radius) and $\mathrm{D}/\mathrm{D}t = \partial/\partial t + \boldsymbol{u}\cdot\nabla$ is the two-dimensional material derivative. The vertical boundary

Part III

LARGE-SCALE ATMOSPHERIC CIRCULATION

I think the causes of the general trade-winds have not been fully explained by any of those who have wrote on that subject... That the action of the Sun is the original cause of these Winds, I think all are agreed.
George Hadley, *Concerning the Cause of the General Trade Winds,* 1735.

CHAPTER
ELEVEN

The Overturning Circulation: Hadley and Ferrel Cells

IN THIS CHAPTER AND THE TWO FOLLOWING we discuss the large-scale circulation, and in particular the *general circulation,* of the atmosphere, this being the mean flow on scales from the synoptic eddy scale — about 1000 km — to the global scale. In this chapter we focus on the dynamics of the Hadley Cell and then, rather descriptively, on the mid-latitude overturning cell or the Ferrel Cell. The latter provides a starting point for chapter 12 which discusses the dynamics of the extratropical zonally averaged circulation. Finally, in chapter 13, we consider the deviations from zonal symmetry, or more specifically the stationary wave pattern, and the stratosphere. In these three chapters we will use many of the tools developed in the previous chapters, but those readers who already have some acquaintance with geophysical fluid dynamics may simply wish to jump in here.

The atmosphere is a terribly complex system, and we cannot hope to fully explain its motion as the analytic solution to a small set of equations. Rather, a full understanding of the atmosphere requires describing it in a consistent way on many levels simultaneously. One of these levels involves simulating the flow by numerically solving the governing equations of motion as completely as possible, for example by using a comprehensive General Circulation Model (GCM). However, such a simulation brings problems of its own, including the problem of understanding the simulation, and discerning whether it is a good representation of reality. Thus, in this chapter and the two following we concentrate on simpler, more conceptual models. We begin this chapter with a brief observational overview of some of the pre-eminent large-scale features of the atmosphere, concentrating on the zonally averaged fields.[1]

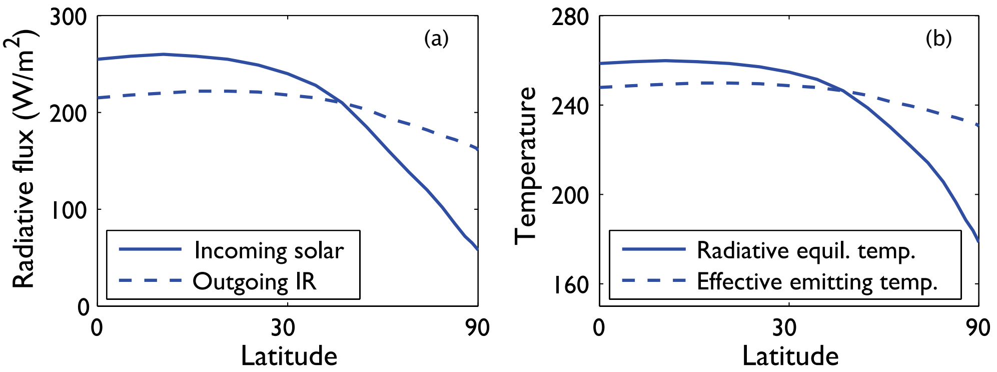


Fig. 11.1 (a) The (approximate) observed net average incoming solar radiation and outgoing infrared radiation at the top of the atmosphere, as a function of latitude (plotted on a sine scale). (b) The temperatures associated with these fluxes, calculated using $T = (R/\sigma)^{1/4}$, where R is the solar flux for the radiative equilibrium temperature and R is the infrared flux for the effective emitting temperature. Thus, the solid line is an approximate radiative equilibrium temperature

11.1 BASIC FEATURES OF THE ATMOSPHERE

11.1.1 The radiative equilibrium distribution

A gross but informative measure characterizing the atmosphere, and the effects that dynamics have on it, is the pole-to-equator temperature distribution. The *radiative equilibrium* temperature is the hypothetical, three-dimensional, temperature field that would obtain if there were no atmospheric or oceanic motion, given the composition and radiative properties of the atmosphere and surface. The field is a function only of the incoming solar radiation at the top of the atmosphere, although to evaluate it entails a complicated calculation, especially as the radiative properties of the atmosphere depend on the amount of water vapour and cloudiness in the atmosphere. (The distribution of absorbers is usually taken to be that which obtains in the observed, moving, atmosphere, in order that the differences between the calculated radiative equilibrium temperature and the observed temperature are due to fluid motion.)

A much simpler calculation that illustrates the essence of the situation is to first note that at the top of the atmosphere the globally averaged incoming solar radiation is balanced by the outgoing infrared radiation. If there is no lateral transport of energy in the atmosphere or ocean then *at each latitude* the incoming solar radiation will be balanced by the outgoing infrared radiation, and if we parameterize the latter using a single latitudinally-dependent temperature we will obtain a crude radiative-equilibrium temperature for the atmospheric column at each latitude. Specifically, a black body subject to a net incoming radiation of S (watts per square metre) has a radiative-equilibrium temperature T_{rad} given by $\sigma T_{rad}^4 = S$, this being Stefan's law with Stefan-Boltzmann constant $\sigma = 5.67 \times 10^{-8}\,\mathrm{W\,m^{-2}\,K^{-4}}$. Thus, for the Earth, we have, at each latitude,

$$\sigma T_{rad}^4 = S(\vartheta)(1 - \alpha), \tag{11.1}$$

where α is the albedo of the Earth and $S(\vartheta)$ is the incoming solar radiation at the top of the atmosphere, and its solution is shown in Fig. 11.1. The solid lines in the two panels show the net solar radiation and the solution to (11.1), T_{rad}; the dashed lines show the observed outgoing infrared radiative flux, I, and the effective emitting temperature associated with it, $(I/\sigma)^{1/4}$. The emitting temperature does not quantitatively characterize that temperature at the Earth's surface, nor at any single level in the atmosphere, because the atmosphere is not a black body and the outgoing radiation originates from multiple levels. Nevertheless, the qualitative point is evident: the radiative equilibrium temperature has a much stronger pole-to-equator gradient than does the effective emitting temperature, indicating that there is a poleward transport of heat in the atmosphere–ocean system. More detailed calculations indicate that the atmosphere is further from its radiative equilibrium in winter than summer, indicating a larger heat transport. The transport occurs because polewards moving air tends to have a higher static energy ($c_pT + gz$ for dry air; in addition there is some energy transport associated with water vapour evaporation and condensation) than the equatorwards moving air, most of this movement being associated with the large-scale circulation. The radiative forcing thus seeks to maintain a pole-to-equator temperature gradient, and the ensuing circulation seeks to reduce this gradient.

11.1.2 Observed wind and temperature fields

The observed zonally averaged temperature and zonal wind fields are illustrated in Fig. 11.2. The vertical coordinate is log pressure, multiplied by a constant factor $H = RT_0/g = 7.5\,\mathrm{km}$, so that the ordinate is similar to height in kilometres. [In an isothermal hydrostatic atmosphere $(RT_0/g)\,\mathrm{d}\ln p = -\,\mathrm{d}z$, and the value of H chosen corresponds to $T_0 = 256\,\mathrm{K}$.] To a good approximation temperature and zonal wind are related by thermal wind balance, which in pressure coordinates is

$$f\frac{\partial u}{\partial p} = \frac{R}{p}\frac{\partial T}{\partial y}. \tag{11.2}$$

In the lowest several kilometres of the atmosphere temperature falls almost monotonically with latitude and height, and this region is called the *troposphere*. The temperature in the lower troposphere in fact varies more rapidly with latitude than does the effective emitting temperature, T_E, the latter being more characteristic of the temperature in the mid-to-upper troposphere. The meridional temperature gradient is much larger in winter than summer, because in winter high latitudes receive virtually no direct heating from the Sun. It is also strongest at the edge of the subtropics, and here it is associated with a zonal jet, particularly strong in winter. There is no need to 'drive' this wind with any kind of convergent momentum fluxes: given the temperature, the flow is a consequence of thermal wind balance, and to the extent that the upper troposphere is relatively frictionless there is no need to maintain it against dissipation. Of course just as the radiative-equilibrium temperature gradient is much larger than that observed, so the zonal wind shear associated with it is much larger than that observed. Thus, the overall effect of the atmospheric and oceanic circulation, and in particular of the turbulent circulation of the mid-latitude atmosphere, is to *reduce* the amplitude of the vertical shear of the eastward flow by way of a poleward heat transport. Observations indicate that about two-thirds of this transport is effected by the atmosphere, and about a third by the ocean, more in low latitudes.[2]

Above the troposphere is the *stratosphere,* and here temperature typically increases with height. The boundary between the two regions is called the *tropopause,* and this varies

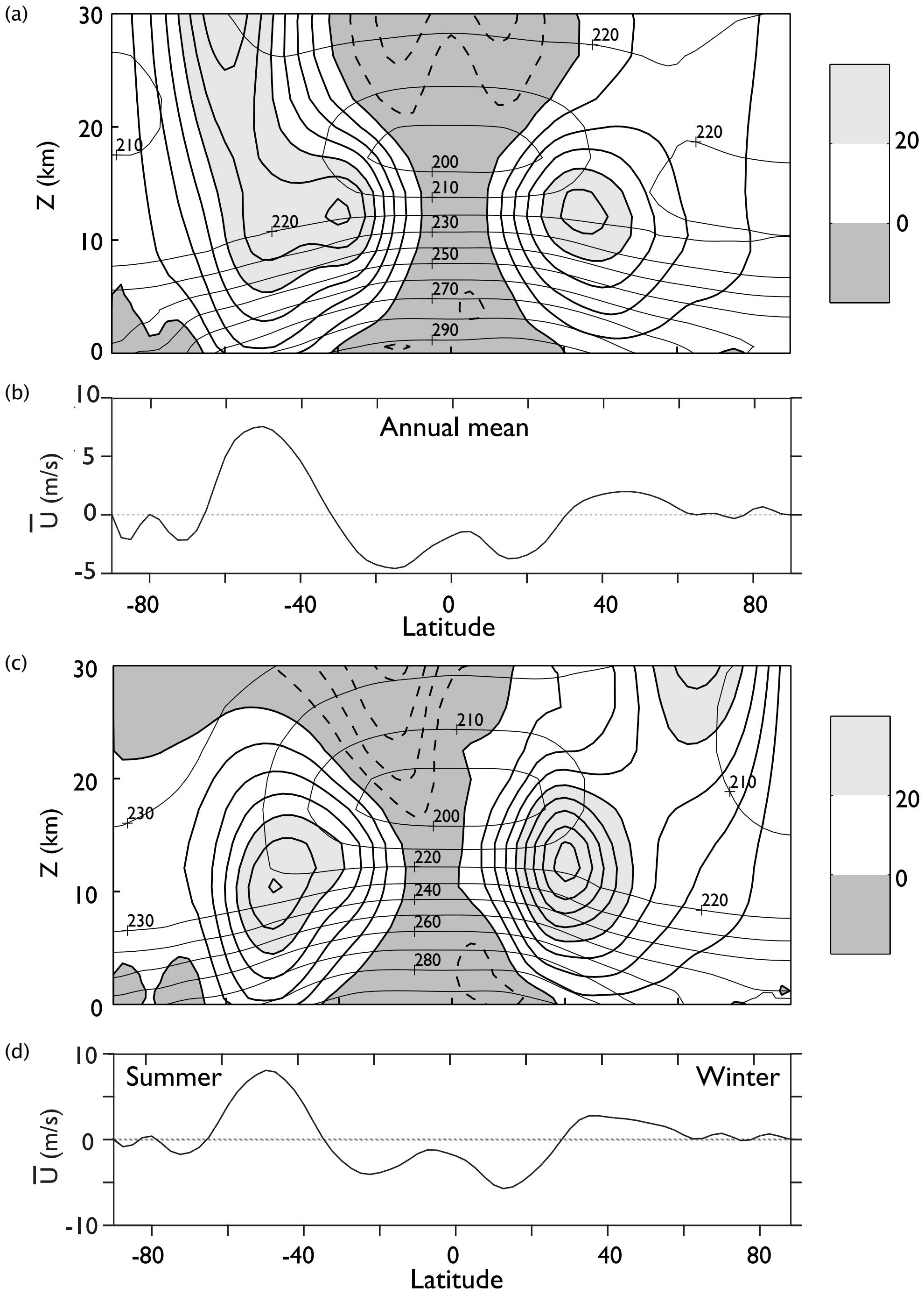


Fig. 11.2 (a) Annual mean, zonally averaged zonal wind (heavy contours and shading) and the zonally averaged temperature (lighter contours). (b) Annual mean, zonally averaged zonal winds at the surface. (c) and (d) Same as (a) and (b), except for northern hemisphere winter (DJF). The wind contours are at intervals of $5\,\mathrm{m\,s^{-1}}$ with shading for eastward winds above $20\,\mathrm{m\,s^{-1}}$ and for all westward winds, and the temperature contours are labelled. The ordinate of (a) and (c) is $Z = -H \log(p/p_R)$, where p_R is a constant, with scale height $H = 7.5\,\mathrm{km}$.

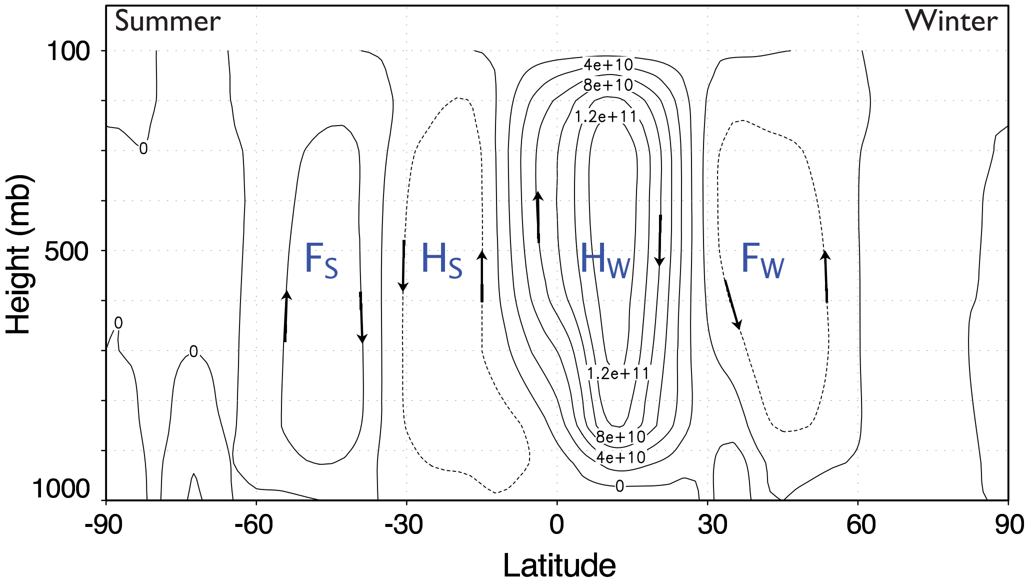


Fig. 11.3 The observed, zonally averaged, meridional overturning circulation of the atmosphere, in units of $\mathrm{kg\,s^{-1}}$, averaged over December–January–February (DJF). In each hemisphere note the presence of a direct *Hadley Cell* (H_W and H_S in winter and summer) with rising motion near the equator, descending motion in the subtropics, and an indirect *Ferrel Cell* (F_W and F_S) at mid-latitudes. There are also hints of a weak direct cell at high latitudes. The winter Hadley Cell is far stronger than the summer one.

in height from about 16 km in the tropics to about 8 km in polar regions. We consider the maintenance of this stratification in section 12.5.

The surface winds typically have, going from the equator to the pole, an E–W–E (easterly–westerly–easterly) pattern, although the polar easterlies are weak and barely present in the Northern Hemisphere. (Meteorologists use ‘westerly’ to denote winds from the west, that is eastward winds; similarly ‘easterlies’ are westward winds. We will use both ‘westerly’ and ‘eastward’, and both ‘easterly’ and ‘westward’, and the reader should be comfortable with all these terms.) In a given hemisphere, the surface winds are stronger in winter than summer, and they are also consistently stronger in the Southern Hemisphere than in the Northern Hemisphere, because in the former the surface drag is weaker because of the relative lack of continental land masses and topography. The surface winds are *not* explained by thermal wind balance. Indeed, unlike the upper level winds, they must be maintained against the dissipating effects of friction, and this implies a momentum convergence into regions of surface westerlies and a divergence into regions of surface easterlies. Typically, the maxima in the eastward surface winds are in mid-latitudes and somewhat polewards of the subtropical maxima in the upper-level westerlies and at latitudes where the zonal flow is a little more constant with height. The mechanisms of the momentum transport in the mid-latitudes and the maintenance of the surface westerly winds are the topics of section 12.1.

Some Features of the Large-scale Atmospheric Circulation

From Figures 11.1–11.3 we see or infer the following.

- A pole–equator temperature gradient that is much smaller than the radiative equilibrium gradient.
- A troposphere, in which temperature generally falls with height, above which lies the stratosphere, in which temperature increases with height. The two regions are separated by a tropopause, which varies in height from about 16 km at the equator to about 6 km at the pole.
- A monotonically decreasing temperature from equator to pole in the troposphere, but a weakening and sometimes reversal of this above the tropopause.
- A westerly (i.e., eastward) tropospheric jet. The time and zonally averaged jet is a maximum at the edge or just polewards of the subtropics, where it is associated with a strong meridional temperature gradient. In mid-latitudes the jet has a stronger barotropic component.
- An E–W–E (easterlies–westerlies–easterlies) surface wind distribution. The latitude of the maximum in the surface westerlies is in mid-latitudes, where the zonally averaged flow is more barotropic.

11.1.3 Meridional overturning circulation

The observed (Eulerian) zonally averaged meridional overturning circulation is illustrated in Fig. 11.3. The figure shows a streamfunction, Ψ for the vertical and meridional velocities such that, in the pressure coordinates used in the figure,

$$\frac{\partial \Psi}{\partial y} = \overline{\omega}, \qquad \frac{\partial \Psi}{\partial p} = -\overline{v}. \tag{11.3}$$

where the overbar indicates a zonal average. In each hemisphere there is rising motion near the equator and sinking in the subtropics, and this circulation is known as the *Hadley Cell.*[3] The Hadley Cell is a thermally direct cell (i.e., the warmer fluid rises, the colder fluid sinks), is much stronger in the winter hemisphere, and extends to about 30°. In mid-latitudes the sense of the overturning circulation is apparently reversed, with rising motion in the high-mid-latitudes, at around 60° and sinking in the subtropics, and this is known as the *Ferrel Cell.* However, as with most pictures of averaged streamlines in unsteady flow, this gives a misleading impression as to the actual material flow of parcels of air because of the presence of eddying motion, and we discuss this in the next chapter. At low latitudes the circulation is more nearly zonally symmetric and the picture does give a qualitatively correct representation of the actual flow. At high latitudes there is again a thermally direct cell (although it is weak and not always present), and thus the atmosphere is often referred to as having a three-celled structure.

11.1.4 Summary

Some of the main features of the zonally averaged circulation are summarized in the shaded box on the preceding page. We emphasize that the zonally averaged circulation is not synonymous with a zonally symmetric circulation, and the mid-latitude circulation is highly asymmetric. Any model of the mid-latitudes that did not take into account the zonal asymmetries in the circulation — of which the weather is the main manifestation — would be seriously in error. This was first explicitly realized in the 1920s, and taking into account such asymmetries is the main task of the dynamical meteorology of the mid-latitudes, and is the subject of the next chapter. On the other hand, the large-scale tropical circulation of the atmosphere is to a large degree zonally symmetric or nearly so, and although monsoonal circulations and the Walker circulation (a cell with rising air in the Eastern Pacific and descending motion in the Western Pacific) are zonally asymmetric, they are also relatively weaker than typical mid-latitude weather systems. Indeed the boundary between the tropics and mid-latitude may be usefully defined by the latitude at which such zonal asymmetries become dynamically important on the large scale and this boundary, at about 30° on average, roughly coincides with the latitude at which the mean meridional overturning circulation vanishes. We begin our dynamical description with a study of the low-latitude zonally symmetric atmospheric circulation.

11.2 A STEADY MODEL OF THE HADLEY CELL

> *Ceci n'est pas une pipe.*
> René Magritte (1898–1967), title of painting.

11.2.1 Assumptions

Let us try to construct a zonally symmetric model of the Hadley Cell,[4] recognizing that such a model is likely applicable mainly to the tropical atmosphere, this being more zonally symmetric than the mid-latitudes. We will suppose that heating is maximum at the equator, and our intuitive picture, drawing on the observed flow of Fig. 11.3, is of air rising at the equator and moving polewards at some height H, descending at some latitude ϑ_H, and returning equatorwards near the surface. We will make three major assumptions:

- *(i)* that the circulation is steady;
- *(ii)* that the polewards moving air conserves its axial angular momentum, whereas the zonal flow associated with the near-surface, equatorwards moving flow is frictionally retarded and is weak;
- *(iii)* that the circulation is in thermal wind balance.

We also assume the model is symmetric about the equator (an assumption we relax in section 11.4). These are all reasonable assumptions, but they cannot be rigorously justified; in other words, we are constructing a *model* of the Hadley Cell, schematically illustrated in Fig. 11.4. The model defines a limiting case — steady, inviscid, zonally-symmetric flow — that cannot be expected to describe the atmosphere quantitatively, but that can be analysed fairly completely. Another limiting case, in which eddies play a significant role, is described in section 11.5. The real atmosphere may defy such simple characterizations, but the two limits provide invaluable benchmarks of understanding.

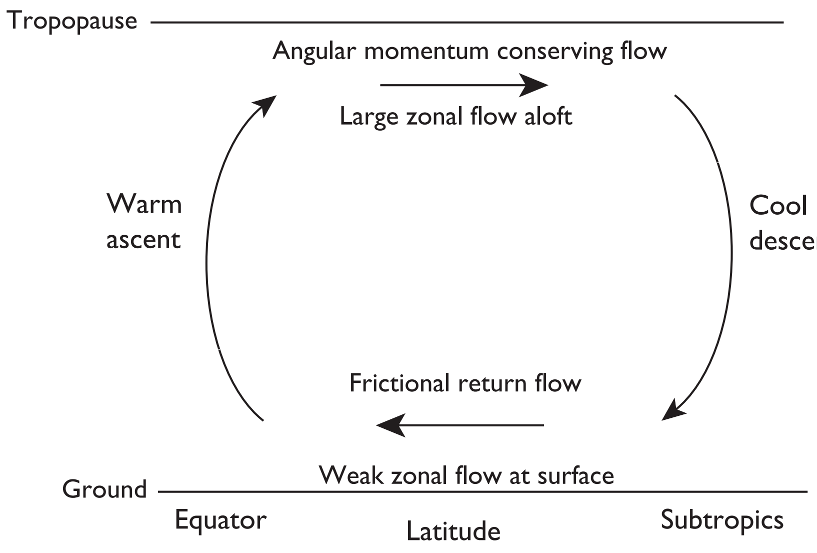


Fig. 11.4 A simple model of the Hadley Cell. Rising air near the equator moves polewards near the tropopause, descending in the subtropics and returning near the surface. The polewards moving air conserves its axial angular momentum, leading to a zonal flow that increases away from the equator. By the thermal wind relation the temperature of the air falls as it moves polewards, and to satisfy the thermodynamic budget it sinks in the subtropics. The return flow at the surface is frictionally retarded and small.

11.2.2 Dynamics

We now try to determine the strength and poleward extent of the Hadley circulation in our steady model. For simplicity we will work with a Boussinesq atmosphere, but this is not an essential aspect. We will first derive the conditions under which conservation on angular momentum will hold, and then determine the consequences of that.

The zonally averaged zonal momentum equation may be easily derived from (2.50a) and/or (2.62) and in the absence of friction it is

$$\frac{\partial \overline{u}}{\partial t} - (f + \overline{\zeta})\overline{v} + \overline{w}\frac{\partial \overline{u}}{\partial z} = -\frac{1}{\cos^2 \vartheta}\frac{\partial}{\partial \vartheta}(\cos^2 \vartheta\, \overline{u'v'}) - \frac{\partial \overline{u'w'}}{\partial z}, \tag{11.4}$$

where $\overline{\zeta} = -(a\cos\vartheta)^{-1}\partial_y(\overline{u}\cos\vartheta)$ and the overbars represent zonal averages. If we neglect the vertical advection and the eddy terms on the right-hand side, then a steady solution, if it exists, obeys

$$(f + \overline{\zeta})\overline{v} = 0. \tag{11.5}$$

Presuming that the meridional flow $\overline{v}$ is non-zero (an issue we address in section 11.2.8) then $f + \overline{\zeta} = 0$, or equivalently

$$2\Omega \sin\vartheta = \frac{1}{a}\frac{\partial \overline{u}}{\partial \vartheta} - \frac{\overline{u}\tan\vartheta}{a}. \tag{11.6}$$

At the equator we shall assume that $\overline{u} = 0$, because here parcels have risen from the surface

Part IV

LARGE-SCALE OCEANIC CIRCULATION

No fairer destiny [has] any physical theory, than that it should of itself point out the way to the introduction of a more comprehensive theory, in which it lives on as a limiting case.

Albert Einstein, *Relativity, the Special and the General Theory,* 1916.

CHAPTER SIXTEEN

The Wind- and Buoyancy-Driven Ocean Circulation

IN THIS CHAPTER we try to understand the combined effect of wind and buoyancy forcing in setting the three-dimensional structure of the ocean. There are three main topics we will consider.

(i) The *main thermocline*, the region in the upper 1 km or so of the ocean where density and temperature change most rapidly.

(ii) The 'wind-driven' overturning circulation. That is, we look into whether and how the ocean might maintain a deep overturning circulation that owes its existence to the effects of wind at the surface, as well as buoyancy forcing at the surface, and that persists even as the diapycnal diffusivity in the ocean interior goes to zero.

(iii) The circulation of the flow in a channel, as a model of the Antarctic Circumpolar Current (ACC).

16.1 THE MAIN THERMOCLINE: AN INTRODUCTION

In the previous chapter we saw that a fluid that is differentially heated from above will develop both an overturning circulation and a region near the surface where the temperature changes rapidly. To examine this in more detail, we consider the circulation in a closed, single hemispheric basin, and again suppose that there is a net surface heating at low latitudes and a net cooling at high latitudes that maintains a meridional temperature gradient at the surface. We presume, ab initio, that there is a single overturning cell, with water rising at low latitudes before returning to polar regions, illustrated schematically in Fig. 16.1.

At lower latitudes the surface water is warmer than the cold water in the abyss. Thus

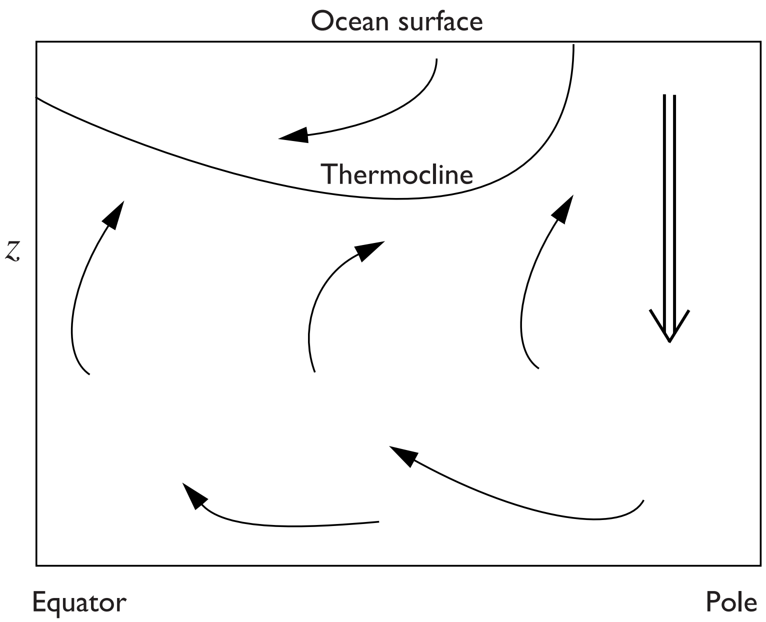


Fig. 16.1 Cartoon of a single-celled meridional overturning circulation, with a wall at the equator. Sinking is concentrated at high latitudes and upwelling spread out over lower latitudes. The thermocline is the boundary between the cold abyssal waters, with polar origins, and the warmer near-surface subtropical water. Wind forcing in the subtropical gyre mechanically pushes the warm water down, increasing the depth of the thermocline.

there must be a vertical temperature gradient everywhere except possibly at the highest latitudes where the cold dense water sinks. This temperature gradient is called the *thermocline,* and is illustrated in Fig. 16.2. In purely buoyancy-driven flows the thickness of the thermocline is determined by way of an advective–diffusive balance, and proportional to some power of the thermal diffusivity as we considered in section 15.7. Let us revisit this issue, first by way of a simple kinematic model.

16.1.1 A simple kinematic model

The fact that cold water with polar origins upwells into a region of warmer water suggests that we consider the simple one-dimensional advective–diffusive balance,

$$w\frac{\partial T}{\partial z} = \kappa\frac{\partial^2 T}{\partial z^2}, \tag{16.1}$$

where w is the vertical velocity, κ is a diffusivity and T is temperature. In mid-latitudes, where this might hold, w is positive and the equation represents a balance between the upwelling of cold water and the downward diffusion of heat. If w and κ are given constants, and if T is specified at the top ($T = T_T$ at $z = 0$) and if $\partial T/\partial z = 0$ at great depth ($z = -\infty$)

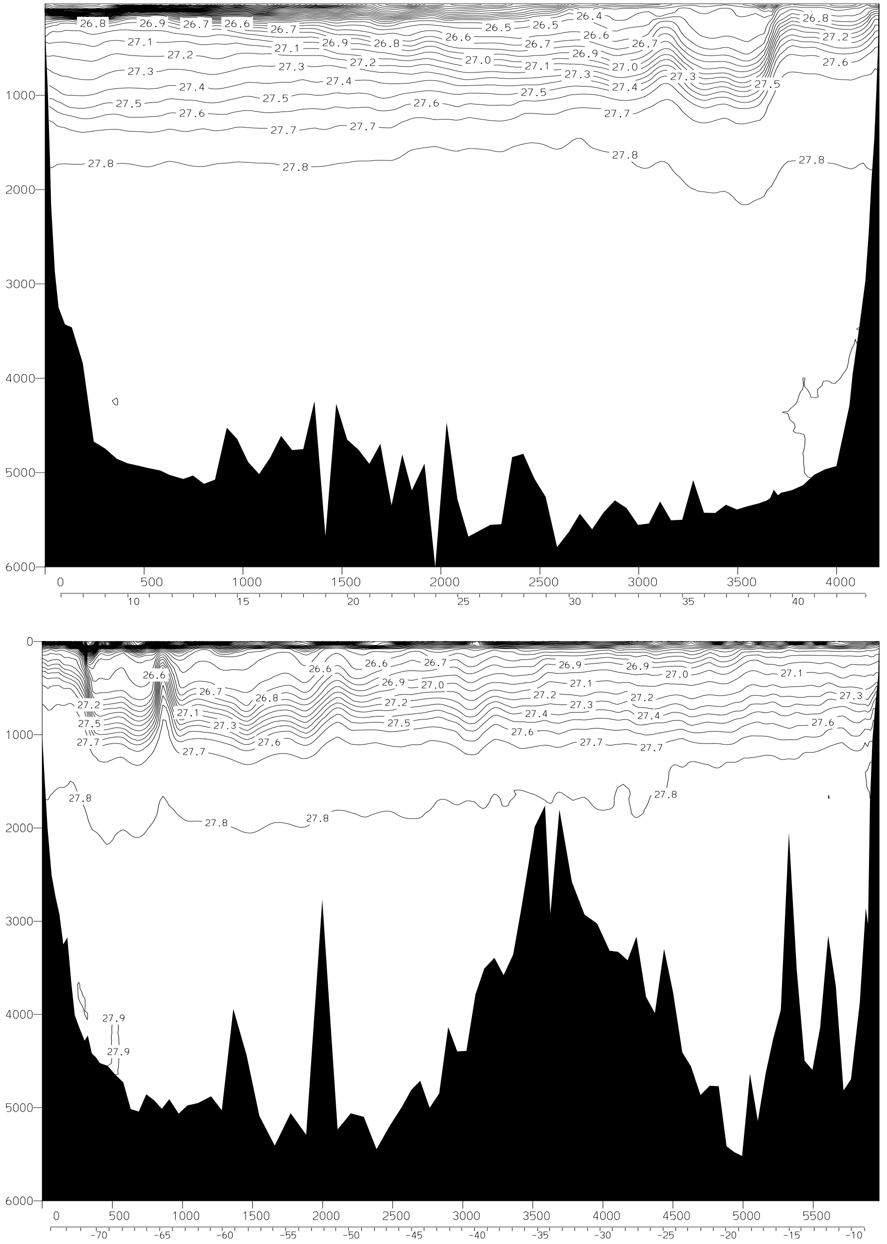


Fig. 16.2 Sections of potential density (σ_θ) in the North Atlantic. Upper panel: meridional section at $53°$W, from $5°$N to $45°$N, across the subtropical gyre. Lower panel: zonal section at $36°$N, from about $75°$W to $10°$W. A front is associated with the western boundary current and its departure from the coast near $40°$N. In the upper northwestern region of the subtropical thermocline there is a region of low stratification known as MODE water: isopycnals above this outcrop in the subtropical gyre and are 'ventilated'; isopycnals below the MODE water outcrop in the subpolar gyre or ACC.[1]

then the temperature falls exponentially away the surface according to

$$T = (T_T - T_B)\,\mathrm{e}^{wz/\kappa} + T_B, \tag{16.2}$$

where T_B is a constant. This expression cannot be used to estimate how the thermocline depth scales with either w or κ, because the magnitude of the overturning circulation depends on κ (section 15.7). However, it is reasonable to see if the observed ocean is broadly consistent with this expression. The diffusivity κ can be measured; it is an eddy diffusivity, maintained by small-scale turbulence, and measurements produce values that range between $10^{-5}\,\mathrm{m^2\,s^{-1}}$ in the main thermocline and $10^{-4}\,\mathrm{m^2\,s^{-1}}$ in abyssal regions over rough topography and in and near continental margins, with still higher values locally.[2] The vertical velocity is too small to be measured directly, but various estimates based on deep water production suggest a value of about $10^{-7}\,\mathrm{m\,s^{-1}}$. Using this and the smaller value of κ in (16.2) gives an e-folding vertical scale, κ/w, of just 100 m, beneath which the stratification is predicted to be very small (i.e., nearly uniform potential density). Using the larger value of κ increases the vertical scale to 1000 m, which is probably closer to the observed value for the total thickness of the thermocline (look at Fig. 16.2), but using such a large value of κ in the main thermocline is not supported by the observations. Similarly, the deep stratification of the ocean is rather larger than that given by (16.1), except with values of diffusivity on the large side of those observed.[3] Thus, there are two conclusions to be drawn.

(i) The observed thickness of the thermocline is somewhat larger than what one might infer from observed values of the diffusivity and overturning circulation.

(ii) The observed deep stratification is somewhat larger than what one might infer from the advective–diffusive balance (16.1) with observed values of diffusivity and overturning circulation.

Of course the model itself, (16.1), is overly simple but these conclusions suggest that additional physical factors may play a role in thermocline dynamics. Mechanical forcing, and in particular the wind, is one such: the wind-stress curl forces water to converge in the subtropical Ekman layer, thereby forcing relatively warm water to downwell and meet the upwelling colder abyssal water at some finite depth, thus deepening the thermocline from its purely diffusive value. Indeed, in so far as we can separate the two effects of wind and diffusion, we can say that the strength of the wind influences the *depth* at which the thermocline occurs, whereas the strength of the diffusivity influences the *thickness* of the thermocline. The influence of the wind on the abyssal circulation is not quite as direct, but we will find in section 16.5 that it will enable both a circulation and deep stratification to persist even in the absence of diffusion.

16.2 SCALING AND SIMPLE DYNAMICS OF THE MAIN THERMOCLINE

We now begin to consider the dynamics that produce an overturning circulation and a thermocline. The Rossby number of the large-scale circulation is small and the scale of the motion large, and the flow obeys the planetary-geostrophic equations:

$$\boldsymbol{f} \times \boldsymbol{u} = -\nabla\phi, \qquad \frac{\partial\phi}{\partial z} = b, \tag{16.3a,b}$$

$$\nabla \cdot \boldsymbol{v} = 0, \qquad \frac{\mathrm{D}b}{\mathrm{D}t} = \kappa \frac{\partial^2 b}{\partial z^2}. \tag{16.4a,b}$$

We suppose that these equations hold below an Ekman layer, so that the effects of a wind stress may be included by specifying a vertical velocity at the top of the domain. The diffusivity, κ, is, as we noted above, an eddy diffusivity, but since its precise form and magnitude are uncertain we must proceed with due caution, and a useful practical philosophy is to try to ignore dissipation and viscosity where possible, and to invoke them only if there is no other way out. Let us therefore scale the equations in two ways, with and without diffusion; these scalings will be central to our theory.

16.2.1 An advective scale

As usual we denote (with one or two exceptions) scaling values with capital letters and non-dimensional values with a hat, so that, for example, $u = U\hat{u}$ and $u = \mathcal{O}(U)$. Let us ignore the diffusive term in (16.4b) and try to construct a scaling estimate for the depth of the wind's influence.

If there is upwelling ($w > 0$) from the abyss, and Ekman downwelling ($w < 0$) at the surface, there is some depth D_a at which $w = 0$. By cross-differentiating (16.3a) we obtain $\beta v = -f\nabla_z \cdot \boldsymbol{u}$, and combining this with (16.4a) gives the familiar geostrophic vorticity equation and corresponding scaling

$$\beta v = f\frac{\partial w}{\partial z} \qquad \rightarrow \qquad \beta V = f\frac{W}{D_a}. \tag{16.5}$$

Here, D_a is the unknown depth scale of the motion, L is the horizontal scale of the motion, which we take as the gyre or basin scale, and V is a horizontal velocity scale. (It is reasonable to suppose that $V \sim U$, where U is the zonal velocity scale, and henceforth we will denote both by U.) The appropriate vertical velocity to use is that due to Ekman pumping, W_E; we will assume (and demonstrate later) that this is much larger than the abyssal upwelling velocity, which in any case is zero by assumption at $z = -D_a$. With this, (16.5) yields the Sverdrup-balance estimate

$$U = \frac{f}{\beta}\frac{W_E}{D_a}. \tag{16.6}$$

We may determine an appropriate value of U using the thermal wind relation, which from (16.3) is

$$\boldsymbol{f} \times \frac{\partial \boldsymbol{u}}{\partial z} = -\nabla b \qquad \rightarrow \qquad \frac{U}{D_a} = \frac{1}{f}\frac{\Delta b}{L}, \tag{16.7}$$

where Δb is the scaling value of variations of buoyancy in the horizontal. Assuming the vertical scales are the same in (16.6) and (16.7) then eliminating U gives

$$\boxed{D_a = W_E^{1/2}\left(\frac{f^2 L}{\beta \Delta b}\right)^{1/2}}. \tag{16.8}$$

[This is essentially the same as the estimate (14.130).] If we relate U and W_E using mass conservation, $U/L = W_E/D_a$, instead of using Sverdrup balance, then we write L in place of f/β and (16.8) becomes $D_a = (W_E f L^2/\Delta b)^{1/2}$, which is not qualitatively different for large scales. The important aspect of these equations is that the depth of the wind-influenced region increases with the magnitude of the wind stress (because $W_E \propto \mathrm{curl}_z \boldsymbol{\tau}$) and decreases with the meridional temperature gradient. The former dependence is reasonably intuitive, and the latter arises because as the temperature gradient increases the associated thermal wind-shear U/D_a correspondingly increases. But the horizontal transport (the product UD_a) is fixed by mass conservation; the only way that these two can remain consistent is for the vertical scale to decrease. Taking $W_E = 10^{-6}\,\mathrm{m\,s^{-1}}$, $\Delta b = g\Delta\rho/\rho_0 = g\beta_T\Delta T \sim 10^{-2}\,\mathrm{m\,s^{-2}}$, $L = 5000\,\mathrm{km}$ and $f = 10^{-4}\,\mathrm{s^{-1}}$ gives $D_a = 500\,\mathrm{m}$. Such a scaling argument cannot be expected to give more than an estimate of the depth of the wind-influenced region; nevertheless, because D_a is much less than the ocean depth, the estimate does suggest that the wind-driven circulation is predominantly an upper-ocean phenomenon.

16.2.2 A diffusive scale

The estimate (16.8) cares nothing about the thermodynamic equation; if we do take into account thermodynamics, with non-zero diffusivity, we recover the model of section 15.7. Thus, briefly, the scaling follows from advective–diffusive balance in the thermodynamic equation, the linear geostrophic vorticity equation, and thermal wind balance:

$$w\frac{\partial b}{\partial z} = \kappa\frac{\partial^2 b}{\partial z^2}, \qquad \beta v = f\frac{\partial w}{\partial z}, \qquad f\frac{\partial \boldsymbol{u}}{\partial z} = \mathbf{k}\times\nabla b, \tag{16.9a,b,c}$$

with corresponding scales

$$\frac{W}{\delta} = \frac{\kappa}{\delta^2}, \qquad \beta U = \frac{fW}{\delta}, \qquad \frac{U}{\delta} = \frac{\Delta b}{fL}, \tag{16.10a,b,c}$$

where δ is the vertical scale. Because there is now one more equation than in the advective scaling theory we cannot take the vertical velocity as a given, otherwise the equations would be overdetermined. We therefore take it to be the abyssal upwelling velocity, which then becomes part of the *solution,* rather than being imposed. From (16.10) we obtain the diffusive vertical scale,

$$\boxed{\delta = \left(\frac{\kappa f^2 L}{\beta\Delta b}\right)^{1/3}}. \tag{16.11}$$

With $\kappa = 10^{-5}\,\mathrm{m^2\,s^{-2}}$ and with the other parameters taking the values given following (16.8), (16.11) gives $\delta \approx 150\,\mathrm{m}$ and, using (16.10a), $W \approx 10^{-7}\,\mathrm{m\,s^{-1}}$, which is an order of magnitude smaller than the Ekman pumping velocity W_E.

A wind-influenced diffusive scaling

The scaling above assumes that the length scale over which thermal wind balance holds is the gyre scale itself. In fact, there is another length scale that is more appropriate, and this leads to a slightly different scaling for the thickness of the thermocline. To obtain this scaling, we first note that the depth of the subtropical thermocline is not constant: it shoals up to the east because of Sverdrup balance, and it may shoal up polewards as the curl of the wind stress falls (and is zero at the poleward edge of the gyre). Thus, referring to Fig. 16.3, the appropriate horizontal length scale $\widetilde{L}$ is given by

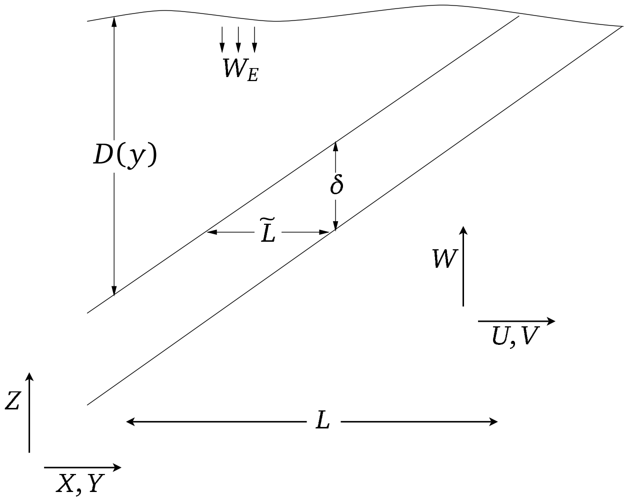


Fig. 16.3 Scaling the thermocline. The diagonal lines mark the diffusive thermocline of thickness δ and depth $D(y)$. The advective scaling for $D(y)$, i.e., D_a, is given by (16.8), and the diffusive scaling for δ is given by (16.13).

$$\widetilde{L} = \delta \frac{L}{D_a}. \tag{16.12}$$

This is no longer an externally imposed parameter, but must be determined as part of the solution. Using $\widetilde{L}$ instead of L as the length scale in the thermal wind equation (16.10c) gives, using (16.8), the modified diffusive scale

$$\boxed{\delta = \kappa^{1/2} \left(\frac{f^2 L}{\Delta b \, \beta D_a} \right)^{1/2} = \kappa^{1/2} \left(\frac{f^2 L}{\Delta b \, \beta W_E} \right)^{1/4}}. \tag{16.13}$$

Substituting values of the various parameters results in a thickness of about 100–200 m. The thermocline thickness now scales as $\kappa^{1/2}$. The interpretation of this scale and that of (16.11) is that the thickness of the thermocline scales as $\kappa^{1/3}$ in the absence of a wind stress, but scales as $\kappa^{1/2}$ if a wind stress is present that can provide a finite slope to the base of the thermocline that is independent of κ, and this is confirmed by numerical simulations.[4] From (16.9a) the vertical velocity, and hence the meridional overturning circulation, no longer scales as $\kappa^{2/3}$ but as

$$W = \frac{\kappa}{\delta} \propto \kappa^{1/2}. \tag{16.14}$$

16.2.3 Summary of the physical picture

What do the vertical scales derived above represent? The wind-influenced scaling, D_a, is the depth to which the directly wind-driven circulation can be expected to penetrate. Thus, over this depth we can expect to see wind-driven gyres and associated phenomena. At greater depths lies the abyssal circulation, and this is not wind-driven in the same sense. Now, in general, the water at the base of the wind-driven layer will not have the same thermodynamic properties as the upwelling abyssal water — this being cold and dense, whereas the water in the wind-driven layer is warm and subtropical (look again at Fig. 16.1). The thickness δ characterizes the diffusive transition region between these two water masses and in