PROCEEDINGS A

royalsocietypublishing.org/journal/rspa

Research

Cite this article: Constantin A, Johnson RS. 2021 On the propagation of waves in the atmosphere. *Proc. R. Soc. A* **477**: 20200424. https://doi.org/10.1098/rspa.2020.0424

Received: 29 May 2020
Accepted: 4 May 2021

Subject Areas:
applied mathematics, atmospheric science

Keywords:
atmospheric waves, asymptotic methods, differential equations

Author for correspondence:
Adrian Constantin
e-mail: adrian.constantin@univie.ac.at

THE ROYAL SOCIETY
PUBLISHING

On the propagation of waves in the atmosphere

Adrian Constantin[1] and Robin S. Johnson[2]

[1]Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria
[2]School of Mathematics, Statistics and Physics, Newcastle University, Newcastle NE1 7RU, UK

AC, 0000-0001-8868-9305

The leading-order equations governing the unsteady dynamics of large-scale atmospheric motions are derived, *via* a systematic asymptotic approach based on the thin-shell approximation applied to the ellipsoidal model of the Earth's geoid. We present some solutions of this single set of equations that capture properties of specific atmospheric flows, using field data to choose models for the heat sources that drive the motion. In particular, we describe standing-waves solutions, waves propagating towards the Equator, equatorially trapped waves and we discuss the African Easterly Jet/Waves. This work aims to show the benefits of a systematic analysis based on the governing equations of fluid dynamics.

1. Introduction

Waves play a fundamental role in the development and evolution of our atmosphere; it is generally accepted, for example, that Rossby waves are the most important large-scale waves (controlling the weather in the mid- and higher latitudes), although gravity (buoyancy) waves and Kelvin waves are also significant. Of course, there are many other ingredients that contribute to the motion of the atmosphere, both locally and globally; here, we focus on the more familiar and well-documented wave-like motions. The complexities of the unsteady flows, including wave motions, suggest—apparently—that these cannot be systematically extracted from the full set of governing equations. Indeed, the familiar way forward is to simplify, to extreme levels, both the geometry and the assumed background state (which is

© 2021 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/, which permits unrestricted use, provided the original author and source are credited.

then perturbed to produce wave motions, for example). Such an approach ignores the detailed structure of any ambient state of the atmosphere, which, in any event, must be regarded as representative of suitable averages over time and space, in order to produce fairly simple and accessible theoretical treatments.

It would seem that the conventional analyses have been developed because, it has been argued, neither the correct almost-spherical geometry nor a realistic background state could be incorporated within the system of relevant governing equations. Retaining the full nature of the Earth's curved-space geometry is essential for large-scale atmospheric flows. To go beyond the limitations of the flat geometry of the f-plane approximation, the typical approach consists in invoking a weak contribution from the curvature by using the β-plane approximation (see [1]). However, in contrast to the f-plane approximation, the β-plane approximation fails to represent a consistent approximation to the governing equations for geophysical flows at mid-latitudes and in polar regions; see [2]. A further typical simplification in research investigations is to ignore the fine structure of the density variation by introducing weighted averages or by relying on the Boussinesq approximation.

As demonstrated by [3] in the case of steady flow, a detailed and extensive description of the motion in an atmosphere that envelopes an ellipsoidal model of the geoid, superimposed on a stationary but variable background state, can be formulated and solved. We now show that the same approach can be adopted when time dependence is added to the system, which, in this initial phase of the development of a coherent mathematical description, is based on an idealized model of the atmosphere. It is therefore appropriate to set aside some of the more realistic but complicating properties of the atmosphere as we observe it, such as turbulence, instabilities and convection processes (see [4]). The plan for the future is to build on this robust theoretical basis, progressively adding more of the important elements that play a role in our atmosphere. Necessarily, our approach differs quite markedly from the more familiar developments that have been employed to describe the appearance and properties of waves in the atmosphere. At the heart of our method is the careful construction of an asymptotic solution driven primarily by the thin-shell approximation, retaining all the other physical attributes of the problem (such as a model for the accurate representation of the oblate-spherical geometry, variable background state, variable density and viscosity and general heat sources). Note that, because the oblateness is only slight for the Earth, it is convenient to incorporate this property as a small correction to the otherwise spherical coordinate system; importantly, we show that this contribution to the geometry uncouples from the thermodynamics and dynamics of the atmosphere, at leading order. The upshot is that the unsteady motion sits on a realistic background state, which results in a very different structure when, for example, waves are included. In particular, we see that any unsteady motion—especially waves—appear at the same order as the underlying dynamic-thermodynamic balance: they do not constitute small perturbations of some uniform (constant) state. This is an important and fundamental difference, which, when coupled with the overarching assumption of a thin-shell geometry—the only simplifying assumption we make for the geometry—for the atmosphere over the surface of the Earth, produces a novel approach to these problems. The result, we claim, provides a far more mathematically consistent presentation of the underlying structure of the atmosphere than anything attempted hitherto, but necessarily it provides a different-looking description for wave propagation even if, superficially, it is very familiar. Furthermore, the resulting system of equations is derived from, and is consistent with, a set of governing equations for the general fluid dynamics of the atmosphere.

The plan is to outline the derivation of the governing equations—the details can be found in [3]—and then we carefully describe how the unsteady motion, with waves, can be accommodated. The resulting system will be analysed and various special cases investigated. Note that throughout the last decades considerable insight into atmospheric flows has been gained by studying the plethora of empirical models that typically rely on observational data and heuristic simplifications of the governing equations. As for systematic asymptotic expansions (e.g. [5,6] and publications discussed therein), these do not start from the general set of fundamental, governing equations. However, all models that are consistent with physical reality

must, ultimately, be reflections of a common framework because, underlying them all, is just one set of governing equations. This fact is a strong motivation to derive a leading-order generic system of (reduced) equations, based on reliable and transparent approximation procedures, as pursued in the present paper. Unsurprisingly, we recover qualitative features that are similar to those encountered in classical models, but there are a few important corrections to be noted, especially with regards to the dispersion relation of zonally harmonic waves and concerning the meridional decay rate of equatorially trapped waves (see §5). This situation is somewhat similar to that encountered in the investigation of Rossby waves: the beta-plane approximation was the step forward from the f-plane approximation that revealed to Rossby [7] the presence of these waves, but Haurwitz [8] extended the accuracy (especially with regard to the meridional extent of the wave) by taking the spherical shape of the Earth into account (see also [9,10]).

2. Governing equations

Following [3], we model the atmosphere as a compressible, viscous fluid (using the familiar Navier–Stokes and mass conservation equations of fluid dynamics, allowing for variable density and describing a viscous fluid), coupled to an equation of state and a suitable version of the first law of thermodynamics. The air is treated as a single-component gas, but with appropriate heating (both external, e.g. solar radiation, and internal, e.g. latent heat). This choice of model is the simplest that we can envisage at this initial phase of the investigation. There is no doubt that further refinements are possible, which, for example, could be used to provide a more accurate description of the background state of the atmosphere, e.g. by including the effects of moisture or any horizontal variation in the eddy viscosity. Such extensions are something for the future, but they will need to be incorporated within the framework of a general set of governing equations that describe the fluid that represents the atmosphere.

The shape of the Earth's sea-level geopotential surface—essentially an oblate spheroid—is well-known from satellite data. It is usual, in the atmospheric sciences, to approximate this oblate spheroid by an ellipsoid obtained by rotating an ellipse, whose centre coincides with the centre of the Earth, about its semi-minor (polar) axis (of length $d'_P \approx 6357\,\mathrm{km}$), with a semi-major (equatorial) axis of length $d'_E \approx 6378\,\mathrm{km}$. (We use primes to denote physical [dimensional] variables; these will be removed when we introduce a suitable non-dimensionalization.) In particular, this accurately accounts for the Earth's equatorial bulge of approximately 21 km and, furthermore, this ensures that the largest departure from the true shape is about 100 m (as a depression to the south of India); see [11]. The Cartesian coordinates (X', Y', Z') of the ellipsoid, with the origin at the Earth's centre and the vertical axis through the North Pole, can be expressed in terms of the longitude and geodetic latitude φ and β, respectively, by

$$(X', Y', Z') = \frac{d'_E}{\sqrt{1 - \mathfrak{e}^2 \sin^2\beta}} \Big(\cos\beta\cos\varphi, \cos\beta\sin\varphi, (1-\mathfrak{e}^2)\sin\beta \Big),$$

where

$$\mathfrak{e} = \sqrt{1 - \left(\frac{d'_P}{d'_E}\right)^2} \approx 0.081$$

is the eccentricity. The coordinate system (φ, β, z'), z' being the vertical distance up from the surface of the ellipsoid, is associated with the ellipsoid, which is rotating about its polar axis with (constant) angular speed $\Omega' \approx 7.29 \times 10^{-5}\,\mathrm{rad\,s^{-1}}$. In this system, the unit tangent vectors at the surface of the ellipsoid are $(\mathbf{e}_\varphi, \mathbf{e}_\beta, \mathbf{e}_z)$; $\mathbf{e}_\varphi$ points from West to East along the geodetic parallel, $\mathbf{e}_\beta$ from South to North along the geodetic meridian and $\mathbf{e}_z$ points upwards (figure 1). The system is valid throughout the space, except along the direction of the polar axis, where this description fails because $\mathbf{e}_\varphi$ and $\mathbf{e}_\beta$ are not well-defined at the two Poles. Note that the direction of apparent gravity (seen by an observer in the rotating frame) is normal to the surface of the ellipsoid, which is the isosurface of the geopotential that defines zero elevation; the geopotential

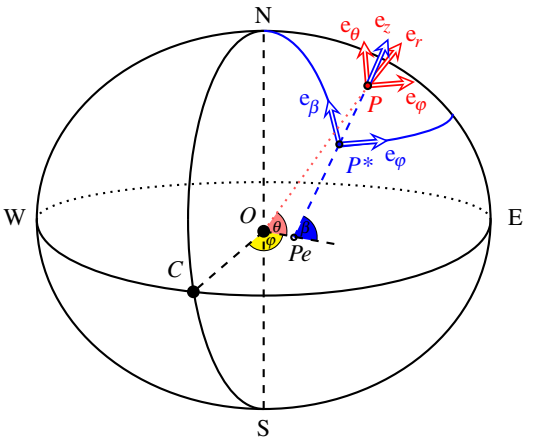

Figure 1. Away from the polar axis, we represent a point P in the atmosphere using the hybrid spherical-geopotential rotating coordinate system (φ, θ, z'), obtained from the spherical system $(e_\varphi, e_\theta, e_r)$ and the geopotential system $(e_\varphi, e_\beta, e_z)$. Here, φ and θ are the longitude and geocentric latitude of P, respectively, β is the geodetic latitude of the projection P^* of P on the ellipsoidal geoid and e_z points upwards along the normal P^*P to the geoid (which intersects the equatorial plane in the point P_e). The unit vectors (e_θ, e_r) are obtained by rotating the unit vectors (e_β, e_z) by the angle $(\beta - \theta)$, in the plane of fixed longitude φ. (Online version in colour.)

being the sum of the Newtonian (gravitational) potential and the centrifugal potential due to the Earth's diurnal rotation. The details of this geometric model are fully described in [3]. In summary, we transform from spherical coordinates, (φ, θ, r'), to the hybrid spherical-geopotential coordinates (φ, θ, z'), and couple this with a transformation of the velocity vector and the gravity term. Thus the components of these vectors, which are defined normal and tangential to the surface of the ellipsoid, are decomposed into components in the spherical system. As pointed out in [3], the advantage of using this spherical-geopotential hybrid rotating coordinate system rather than the conventional spherical potential approximation (described in [12]) is that the formulation retains the details of the curved-space geometry of the Earth and the leading-order (geometrical) correction terms apply to the background state of the atmosphere but do not interact (in the leading-order perturbation) with the dynamics of the atmosphere.

The relation between r' and z' is given by the identity

$$\frac{r'}{d'_E} = 1 + \varepsilon z - \frac{1}{2}\mathfrak{e}^2 \sin^2\theta - \mathfrak{e}^4\left(\frac{1}{2} - \frac{3}{8}\sin^2\theta\right)\sin^2\theta + \mathrm{O}(\mathfrak{e}^6, \varepsilon\mathfrak{e}^4, \varepsilon^2\mathfrak{e}^2, \varepsilon^3),$$

with $z' = \varepsilon z$ (see below). Furthermore, we must use a non-dimensionalization that enables us to put the equations in a form that is relevant for a discussion of atmospheric flows. To do this, we introduce a scale length, which we take to be the maximum height of the troposphere, H' (which is about 16 km at the Equator), and an associated speed scale $\Omega' H'$ (approx. $1.2\,\mathrm{m\,s^{-1}}$). Thus we define

$$z' = \varepsilon z \quad \text{where } \varepsilon = \frac{H'}{d'_E},\ t' = \frac{1}{\Omega'}\,t,\ \rho' = \bar{\rho}'\rho,\ \mu' = \bar{\mu}'\mu,$$

$$(u', v', w') = \Omega' H'\,(u, v\cos(\beta-\theta) + kw\sin(\beta-\theta), kw\cos(\beta-\theta) - v\sin(\beta-\theta)),$$

where the over-bar denotes average ($\bar{\rho}' \approx 0.8\,\mathrm{kg\,m^{-3}}$, $\bar{\mu}' \approx 2\times10^{-2}\,\mathrm{kg\,m^{-1}\,s^{-1}}$ being density and dynamic eddy viscosity, respectively) and the time scale $1/\Omega'$ measures time in units of about $3\frac{1}{2}$ h (this being appropriate for wave-like or other unsteady motions of the atmosphere). The velocity

in spherical coordinates is (u', v', w') and (u, v, w) is a non-dimensional velocity with w normal, and (u, v) tangential, to the ellipsoidal geoid, where

$$\beta - \theta = \mathfrak{e}^2 \sin\theta\cos\theta + \mathfrak{e}^4 \sin\theta\cos^3\theta - \varepsilon\mathfrak{e}^2 z\,\sin\theta\cos\theta + \mathrm{O}(\mathfrak{e}^6, \varepsilon\mathfrak{e}^4, \varepsilon^2\mathfrak{e}^2, \varepsilon^3).$$

The constant k measures the size of the velocity component normal to the ellipsoid, and its choice controls the type of problem to be examined. In the case of thin-shell theory (i.e. small ε, as we have here), the choice consistent with the equations (and kinematic boundary conditions, for example) is $k = \varepsilon$; we use this identification hereafter.

The pressure (p') and the temperature (T') are correspondingly non-dimensionalized:

$$p' = \bar{\rho}'(\Omega' d'_E)^2 p, \quad T' = \frac{(\Omega' d'_E)^2}{\mathfrak{R}'}\,T,$$

where $\mathfrak{R}' \approx 287\,\mathrm{m^2\,s^{-2}\,K^{-1}}$ is the universal gas constant and $(\Omega' d'_E)^2/\mathfrak{R}' \approx 800^\circ$ K. The dynamic eddy viscosity is assumed to vary only in the radial direction (this being the choice that was made in [3]); the transformation from r' to z' then shows that, for the viscous terms that appear in the Navier–Stokes equation, we have

$$\mu = m(z) + \mathrm{O}(\mathfrak{e}^6, \varepsilon\mathfrak{e}^4, \varepsilon^2\mathfrak{e}^2, \varepsilon^3).$$

We are now in a position to quote the governing equations, suitably approximated for small ε and small $\delta = \mathfrak{e}^2$, this latter parameter measuring the effects of small deviations of the ellipsoid from the spherical. The equations follow directly from those derived in [3]; here, though, we retain the time dependence; the errors in each equation are recorded and we write out each component of the Navier–Stokes equation. We obtain

$$\varepsilon\rho\frac{\partial u}{\partial t} - 2\varepsilon\rho v\sin\theta = -\frac{1-\varepsilon z+\delta\Delta}{\cos\theta}\frac{\partial p}{\partial\varphi} + \frac{\varepsilon}{R_e}\frac{\partial}{\partial z}\left(m\frac{\partial u}{\partial z}\right) + \mathrm{O}(\varepsilon^2, \varepsilon\delta, \delta^2), \tag{2.1}$$

$$\begin{aligned}\varepsilon^2\rho\frac{\partial v}{\partial t} + 2\varepsilon^2\rho u\sin\theta + \varepsilon\rho(1+\varepsilon z-\delta\Delta)\sin\theta\,\cos\theta \\ = -(1-\varepsilon z+\delta\Delta)\left[\varepsilon\frac{\partial p}{\partial\theta} + (2\delta\Delta+\delta^2 D)\frac{\partial p}{\partial z}\right] \\ - 2\delta\Delta\rho g[1+\delta(\sin\theta+\cos\theta)\cos\theta - 3\varepsilon z] + \frac{\varepsilon^2}{R_e}\frac{\partial}{\partial z}\left(m\frac{\partial v}{\partial z}\right) + \mathrm{O}(\varepsilon^3, \varepsilon^2\delta, \varepsilon\delta^2),\end{aligned} \tag{2.2}$$

$$-\varepsilon\rho\cos^2\theta = -\frac{\partial p}{\partial z} - \rho g(1-2\varepsilon z+2\delta\Delta) + \mathrm{O}(\varepsilon^2, \varepsilon\delta, \delta^2), \tag{2.3}$$

$$\frac{D\rho}{Dt} + \varepsilon\rho\left\{\frac{1}{\cos\theta}\left(\frac{\partial u}{\partial\varphi} + \frac{\partial}{\partial\theta}(v\cos\theta)\right) + \frac{\partial w}{\partial z}\right\} = \mathrm{O}(\varepsilon^2, \varepsilon\delta, \delta^2), \tag{2.4}$$

$$p = \rho T \tag{2.5}$$

and

$$c_p\frac{DT}{Dt} - \varepsilon\kappa\frac{\partial^2 T}{\partial z^2} - \frac{1}{\rho}\frac{\mathrm{D}p}{\mathrm{D}t} = \varepsilon\,Q(\varphi, \theta, z;\,\varepsilon, \delta) + \mathrm{O}(\varepsilon^2, \varepsilon\delta, \delta^2), \tag{2.6}$$

where

$$\frac{\mathrm{D}}{\mathrm{D}t} \equiv \frac{\partial}{\partial t} + \varepsilon\frac{u}{\cos\theta}\frac{\partial}{\partial\varphi} + \varepsilon v\frac{\partial}{\partial\theta} + \varepsilon w\frac{\partial}{\partial z},$$

$$\Delta(\theta) = \frac{1}{2}\sin\theta\cos\theta, \quad D(\theta) = \left(1 - \frac{3}{2}\sin^2\theta\right)\sin\theta\cos\theta.$$

The additional parameters that we have introduced here, based on the (constant) average acceleration of gravity at the surface of the Earth ($g' \approx 9.8\,\mathrm{m\,s^{-2}}$), the specific heat of air ($c_p' \approx 1500\,\mathrm{m^2\,s^{-2}\,K^{-1}}$) and the thermal diffusivity of predominantly dry air ($\nu' \approx 2\times10^{-5}\,\mathrm{m^2\,s^{-1}}$), are

$$R_e = \frac{\tilde{\rho}'\Omega' H'^2}{\tilde{\mu}'} \approx 7\times10^5, \quad g = \frac{g'H'}{(\Omega' d_E')^2} \approx 0.72,$$
$$c_p = \frac{c_p'}{\Re'} \approx 5.25, \qquad \kappa = \frac{\nu' c_p'}{\Re'\Omega' H'^2} \approx 6\times10^{-9},$$

all O(1), i.e. held fixed as the limiting process is performed. Note that the appearance of large Reynolds number, R_e, and small κ, suggests that additional approximations could be performed, but these are unnecessary (and would then require the introduction of viscous and thermal boundary layers, for example). We note that equation (2.6) is the first law of thermodynamics, with Q representing the (non-dimensional) heat sources (or sinks), expressing the change of total energy due to any heat exchanges. The second law of thermodynamics, setting limits for the transformations between heat energy and the sum of kinetic and potential energies (see [13,14]), will be not be of direct concern here: in the general discussion we do not, nor do we need to, specify the sources of the heat input since we do not address the issue of the climate being in a non-equilibrium thermodynamical state (see [15–17] for the challenges encountered in trying to exploit the second law to determine a lower bound to the entropy production). Furthermore, we choose to apply our results to regions of the Earth where the topography is relatively unchanging; the inclusion of significant orography, which will affect the bottom boundary condition and the local generation and structure of waves, is a refinement relegated to a future investigation.

3. Asymptotic structure of the solution

We seek a solution of (2.1)–(2.6) as an asymptotic expansion which begins

$$q(\varphi,\theta,z,t;\varepsilon,\delta) \sim q_0(\varphi,\theta,z) + \varepsilon\, q_1(\varphi,\theta,z,t) + \delta\,\widehat{q}_1(\varphi,\theta,z,t),$$

where q (and correspondingly q_n and $\widehat{q}_n$) represent each of the variables u, v, w, p, ρ, T. Furthermore, we assume that the boundary and initial conditions are consistent with these expansions, and that the heat-source term also follows this pattern. (The lack of any evident singularities in the velocity, temperature or pressure fields, providing that the neighbourhood of the Poles is avoided, and in the small-eccentricity corrections, indicates that this asymptotic structure is otherwise uniformly valid, i.e. higher-order terms remain small; more details are given in [3].) The leading-order problem (which is time-independent, being driven by $\partial\rho_0/\partial t = 0$ from (2.4)), then gives

$$\frac{\partial p_0}{\partial\varphi} = 0; \quad \rho_0 \sin\theta\cos\theta = -\frac{\partial p_0}{\partial\theta};\ 0 = -\frac{\partial p_0}{\partial z} - \rho_0 g;\ p_0 = \rho_0 T_0;$$
$$c_p L_0(T_0) - \kappa\,\frac{\partial^2 T_0}{\partial z^2} - \frac{1}{\rho_0} L_0(p_0) = Q_0(\varphi,\theta,z),$$

where

$$L_0 \equiv \left(\frac{u_0}{\cos\theta}\,\frac{\partial}{\partial\varphi} + v_0\,\frac{\partial}{\partial\theta} + w_0\,\frac{\partial}{\partial z}\right);$$

note that the third equation above appears twice in the set (2.1)–(2.6). There is a solution of this system (see [3]) in which the temperature decreases linearly with height, independently of the velocity field (which nevertheless does appear in the equations at this order): the classical adiabatic model of the lower atmosphere. This corresponds to an external heat source of zero ($Q_0 \equiv 0$): the heating is supplied by heat transfer from the surface of the Earth upwards into the atmosphere. (It should be noted, in our approach, that we determine the heat sources that are consistent with the motions that we describe; clearly such heat sources then drive the

well-behaved flow properties. There is therefore no need to add general, auxiliary conditions on the background heating that would ensure the existence of solutions.) The resulting solution is

$$T_0 = A - \frac{1}{c_p}\left(gz - \frac{1}{2}\cos^2\theta\right); \quad p_0 = B\left[A - \frac{1}{c_p}\left(gz - \frac{1}{2}\cos^2\theta\right)\right]^{c_p};$$
$$\rho_0 = B\left[A - \frac{1}{c_p}\left(gz - \frac{1}{2}\cos^2\theta\right)\right]^{c_p - 1}.$$

We impose the boundary condition

$$T_0 = T_{E0} \quad \text{at } [z = 0, \theta = \theta_0],$$

so that

$$A = T_{E0} - \frac{1}{2c_p}\cos^2\theta_0;$$

the constant B is fixed by knowing the pressure (or density) at ground level at some θ. We have chosen to work with a background state which is steady; other choices are possible, e.g. via the velocity field (u_0, v_0, w_0) or by allowing variations on a suitable, long time scale. As formulated here, we may choose to use the boundary conditions appropriate to the region of the Earth centred on $(z = 0, \theta = \theta_0)$ and to a particular season. This solution describes the background state, at leading order in ε and δ, for the troposphere, and it is only this section of the atmosphere that we shall discuss here; a related description for higher levels in the atmosphere can be found in [3].

The dominant effects of the ellipsoidal-geometry correction, as these distort the otherwise spherical solution, arise at O(δ); the equations at this order are

$$0 = -\left(\frac{\partial \widehat{p}_1}{\partial \varphi} - \Delta\frac{\partial p_0}{\partial \varphi}\right); \quad (\widehat{\rho}_1 + \Delta\rho_0)\sin\theta\cos\theta = -\left(\frac{\partial \widehat{p}_1}{\partial \theta} - \Delta\frac{\partial p_0}{\partial \theta}\right);$$
$$0 = -\left(\frac{\partial \widehat{p}_1}{\partial z} + \Delta\frac{\partial p_0}{\partial z}\right) - (\widehat{\rho}_1 - 2\Delta\rho_0)g; \quad \frac{\partial \widehat{\rho}_1}{\partial t} = 0;\ \widehat{p}_1 = \widehat{\rho}_1 T_0 + \rho_0 \widehat{T}_1;$$
$$c_p L_0(\widehat{T}_1) - \kappa\frac{\partial^2 \widehat{T}_1}{\partial z^2} - \frac{1}{\rho_0}L_0(\widehat{p}_1) + \frac{\widehat{\rho}_1}{\rho_0^2}\left(v_0\frac{\partial p_0}{\partial \theta} + w_0\frac{\partial p_0}{\partial z}\right) = \widehat{q}_1(\varphi, \theta, z, t);$$

these equations have been simplified by using the equations that describe the leading order. The solution of this system simply produces a (uniformly valid) small adjustment to the background state, as described in [3]. However, the solution for these temperature, pressure and density perturbations is time-independent, and so the first law of thermodynamics (at this order) shows that, in order to maintain this structure, a heat source that moves with the fluid is required, e.g. some latent-heat input. We do not pursue the details of this contribution here because it uncouples from the leading-order time-dependent element of the motion (e.g. waves) that we want to investigate.

The important time-dependence appears at O(ε), which is also where the dominant contribution to both the dynamics and thermodynamics is evident; this general structure, for steady flow, is discussed in [3]. The resulting equations are

$$\rho_0\frac{\partial u_0}{\partial t} - 2\rho_0 v_0\sin\theta = -\frac{1}{\cos\theta}\left(\frac{\partial p_1}{\partial \varphi} - z\frac{\partial p_0}{\partial \varphi}\right) + \frac{1}{R_e}\frac{\partial}{\partial z}\left(m\frac{\partial u_0}{\partial z}\right), \tag{3.1}$$

$$\rho_0\frac{\partial v_0}{\partial t} + 2\rho_0 u_0\sin\theta + (\rho_1 + z\rho_0)\sin\theta\cos\theta = -\left(\frac{\partial p_1}{\partial \theta} - z\frac{\partial p_0}{\partial \theta}\right) + \frac{1}{R_e}\frac{\partial}{\partial z}\left(m\frac{\partial v_0}{\partial z}\right), \tag{3.2}$$

$$-\rho_0\cos^2\theta = -\frac{\partial p_1}{\partial z} - g(\rho_1 - 2z\rho_0), \tag{3.3}$$

$$\cos\theta\frac{\partial \rho_1}{\partial t} + \frac{\partial}{\partial \varphi}(\rho_0 u_0) + \frac{\partial}{\partial \theta}(\rho_0 v_0\cos\theta) + \frac{\partial}{\partial z}(\rho_0 w_0\cos\theta) = 0, \tag{3.4}$$

$$p_1 = \rho_0 T_1 + \rho_1 T_0, \tag{3.5}$$

and

$$c_p \frac{\partial T_1}{\partial t} - \frac{1}{\rho_0} \frac{\partial p_1}{\partial t} = \mathfrak{Q}_0(\varphi, \theta, z, t), \tag{3.6}$$

where

$$Q \sim Q_0(\varphi, \theta, z) + \mathfrak{Q}_0(\varphi, \theta, z, t) + \delta \widehat{q}_1(\varphi, \theta, z, t)$$

with $Q_0 \equiv 0$ for the troposphere. For time-independent flows, equations (3.1)–(3.6) recover precisely those obtained in [3], but with (3.6) replaced by a more complete version of the first law of thermodynamics, which provides a mechanism for describing the heat sources associated with the motion. This difference arises because the leading terms in the time derivative vanish, so we must, perforce, go to the next order for the description of a steady dynamic-thermodynamic balance. The considerations in [3] describe the combined dynamic and thermodynamic properties of the steady troposphere; this further analysis has now demonstrated that, on time scales measured by $1/\Omega'$, time dependence appears at this same order. We can expect, therefore, to find unsteady solutions which may not be wave-like, but also (superimposed on any such motions) wave solutions, both being the same size (in the ε sense) and both appearing in the dominant dynamic-thermodynamic structure of the atmosphere and, significantly, not as a small perturbation to a constant state.

Our main aim—to put the theoretical study of the motions of our idealized model of the atmosphere on a solid foundation based on the general equations of fluid mechanics—is motivated, in part, by the reported observations and data from field studies (e.g. satellite observations, regional and global weather data, together with weather forecasts, climate modelling and reanalyses). Indeed, atmospheric waves—we are interested here in stable waves—are typically identified experimentally from satellite observations, after having subtracted the background flow field *via* appropriate filtering methods. Most studies use spectral analysis to isolate various wave modes, seeking agreement with structures revealed in theoretical studies; see the discussion in [18]. The system (3.1)–(3.6) shows that a careful asymptotic derivation can accommodate a dynamic-thermodynamic balance, which encompasses more behaviour (and is readily accessible) than simply perturbing a constant zonal flow (as described in [1,19,20]). The latter oversimplification of the dynamics was driven, presumably, by the need to use a very restrictive form of heat forcing, coupled to the limitations of a simplistic model of the background state.

4. Development of the unsteady problem at leading order

The main aim in this initial investigation is to examine the unsteady problem associated with the dominant dynamic-thermodynamic balance, which arises at $\mathrm{O}(\varepsilon)$ in our idealized model of the atmosphere. (This model excludes, therefore, the role of turbulence, instabilities, etc.; it is planned to extend our approach by adding more extensive attributes of the atmosphere in subsequent studies.) In particular, we seek solutions—hopefully in a manageable form—of two types: unsteady, but not wave-like, and wave solutions. However, before we proceed, it is expedient to make some small adjustments to the presentation of equations (3.1)–(3.6); this will ease the later analysis.

First, writing

$$\mathfrak{Q}_0 = \frac{\partial \mathfrak{q}_0}{\partial t} \quad \text{with } \mathfrak{q}_0(\varphi, \theta, z, t) = \int_0^t \mathfrak{Q}_0(\varphi, \theta, z, s)\, \mathrm{d}s,$$

equation (3.6) can be integrated to give

$$c_p T_1 - \frac{1}{\rho_0} p_1 = \mathfrak{q}_0(\varphi, \theta, z, t) + A_1(\varphi, \theta, z), \tag{4.1}$$

where $A_1(\varphi,\theta,z)$ is an arbitrary function, fixed by the initial data on the perturbation temperature and pressure. Now from equation (3.3), we see that

$$p_1 = \rho_0\{(gz^2 + z\cos^2\theta) + g\Pi_1 + B_1\}, \tag{4.2}$$

where

$$\Pi_1 = \int_0^z \left(\frac{\mathfrak{q}_0(\varphi,\theta,\xi,t) + A_1(\varphi,\theta,\xi)}{T_0(\theta,\xi)}\right) \mathrm{d}\xi, \tag{4.3}$$

and $B_1(\varphi,\theta,t)$ is an arbitrary function determined, for example, by the perturbation pressure on the ground. The expression for Π_1 measures the total heat input, weighted with respect to the background temperature, from the bottom of the atmosphere upwards, and over time. The complete description of the thermodynamic properties of the atmosphere, at this order, is then obtained from (3.5) and (3.6) in the form

$$T_1 = \frac{1}{c_p}\left(\frac{p_1}{\rho_0} + \mathfrak{q}_0 + A_1\right), \quad \rho_1 = \frac{1}{c_p T_0}[(c_p - 1)p_1 - (\mathfrak{q}_0 + A_1)\rho_0], \tag{4.4}$$

where p_1 is given by (4.2).

It is now convenient to introduce

$$F_1(\varphi,\theta,z,t) = g\Pi_1(\varphi,\theta,z,t) + B_1(\varphi,\theta,t), \tag{4.5}$$

and first we show how the heat forcing (4.5) can be related directly to a familiar descriptor used in atmospheric studies. To do this, we introduce the potential temperature, $\mathfrak{T} = Tp^{-1/c_p}$, and the square of the (non-dimensional) Brunt–Väisälä frequency, N, where

$$N^2 = \frac{g}{\varepsilon}\frac{\partial \ln(\mathfrak{T})}{\partial z}; \tag{4.6}$$

see [19]. (We have introduced $\Omega'\sqrt{d'_E/H'}$ for the non-dimensionalization leading to N.) This is used as a measure of the static stability of the atmosphere, the adiabatic perturbation of a fluid parcel about its equilibrium position being governed by

$$\frac{\mathrm{D}^2}{\mathrm{D}t^2}(\zeta) = -N^2(\zeta), \tag{4.7}$$

for small vertical displacements ζ. Considering only the unsteady perturbation, i.e. the $\mathrm{O}(\varepsilon)$ terms (because the $\mathrm{O}(\delta)$ terms uncouple from any wave-like motion and contribute only to the background state), we have

$$\mathfrak{T} = Tp^{-1/c_p} = (T_0 + \varepsilon T_1 + \cdots)(p_0 + \varepsilon p_1 + \cdots)^{-1/c_p},$$

which gives

$$\ln(\mathfrak{T}) = \varepsilon\left(\frac{T_1}{T_0} - \frac{1}{c_p}\frac{p_1}{p_0}\right) + o(\varepsilon),$$

and then invoking (4.3)–(4.4) we obtain, at this order

$$\frac{1}{\varepsilon}\frac{\partial \ln(\mathfrak{T})}{\partial z} = \frac{\partial}{\partial z}\left(\frac{T_1}{T_0} - \frac{1}{c_p}\frac{p_1}{p_0}\right) = \frac{\partial}{\partial z}\left(\frac{\mathfrak{q}_0 + A_1}{c_p T_0}\right) = \frac{1}{c_p}\frac{\partial^2 \Pi_1}{\partial z^2} = \frac{1}{gc_p}\frac{\partial^2 F_1}{\partial z^2}.$$

Therefore, we may identify

$$N^2 = \frac{1}{c_p}\frac{\partial^2 F_1}{\partial z^2}. \tag{4.8}$$

It is usual to assume that N^2 is a constant (see [19]), in which case the general explicit solution of equation (4.7) follows directly, the necessary and sufficient condition for stability being $N^2 > 0$. The assumption of constant N^2 is clearly an oversimplification. While an explicit, general formula for the solution of equation (4.7), for N^2 varying in the vertical direction, is not available, the comparison method shows that stability holds if $N^2 > 0$ (and $N = 0$ implies neutral stability, while instability corresponds to $N^2 < 0$); see [21].

Using (4.5), equations (3.1) and (3.2) become

$$\rho_0 \frac{\partial u_0}{\partial t} - 2\rho_0 v_0 \sin\theta = -\frac{\rho_0}{\cos\theta}\frac{\partial F_1}{\partial \varphi} + \frac{1}{R_e}\frac{\partial}{\partial z}\left(m\frac{\partial u_0}{\partial z}\right) \tag{4.9}$$

and

$$\rho_0 \frac{\partial v_0}{\partial t} + 2\rho_0 u_0 \sin\theta = -\rho_0 \frac{\partial F_1}{\partial \theta} + \frac{\rho_0 \sin\theta\cos\theta}{g}\frac{\partial F_1}{\partial z} + \frac{1}{R_e}\frac{\partial}{\partial z}\left(m\frac{\partial v_0}{\partial z}\right), \tag{4.10}$$

respectively, together with

$$\frac{\partial}{\partial \varphi}(\rho_0 u_0) + \frac{\partial}{\partial \theta}(\rho_0 v_0 \cos\theta) + \frac{\partial}{\partial z}(\rho_0 w_0 \cos\theta) = \frac{\rho_0 \cos\theta}{g}\left(\frac{\partial^2 F_1}{\partial z \partial t} - \frac{c_p - 1}{c_p}\frac{g}{T_0}\frac{\partial F_1}{\partial t}\right), \tag{4.11}$$

from equation (3.4). Moreover, because the background state is a function of

$$\zeta = gz - \frac{1}{2}\cos^2\theta$$

only, it is useful to transform from (φ, θ, z, t) coordinates to $(\varphi, \theta, \zeta, t)$; equations (4.9)–(4.11) then become

$$\rho_0 \frac{\partial u_0}{\partial t} - 2\rho_0 v_0 \sin\theta = -\frac{\rho_0}{\cos\theta}\frac{\partial F_1}{\partial \varphi} + \frac{g^2}{R_e}\frac{\partial}{\partial \zeta}\left(M\frac{\partial u_0}{\partial \zeta}\right), \tag{4.12}$$

$$\rho_0 \frac{\partial v_0}{\partial t} + 2\rho_0 u_0 \sin\theta = -\rho_0 \frac{\partial F_1}{\partial \theta} + \frac{g^2}{R_e}\frac{\partial}{\partial \zeta}\left(M\frac{\partial v_0}{\partial \zeta}\right) \tag{4.13}$$

and

$$\frac{\partial}{\partial \varphi}(\rho_0 u_0) + \frac{\partial}{\partial \theta}(\rho_0 v_0 \cos\theta) + \frac{\partial}{\partial \zeta}(\rho_0 v_0 \sin\theta\cos^2\theta + g\rho_0 w_0 \cos\theta) = \cos\theta\frac{\partial^2(\rho_0 F_1)}{\partial \zeta \partial t}, \tag{4.14}$$

respectively, where $m(z) \equiv M(\zeta + (1/2)\cos^2\theta)$. This completes the development of the system that we plan to discuss here, namely, equations (4.12)–(4.14), in conjunction with (4.2) and (4.4).

5. Solutions of the time-dependent problem

We now turn to the main thrust of this presentation: the construction of some explicit solutions, all derived from our overarching set of general equations that arise at $O(\varepsilon)$. In particular, we examine certain types of wave-like motion; among the plethora of unsteady motions, this is a reasonable testing ground for the applicability of our equations. Although we strongly advocate an investigation, using numerical techniques, of the complete system (4.2), (4.4) and (4.12)–(4.14), for realistic forcing functions and flow properties, e.g. suitable, variable eddy viscosity, our aim here is to produce analytical results. This will, we believe, both confirm and explain some familiar theories (but now placed in a wider context, both in terms of minimal simplifying assumptions and precise approximations and errors) and also generate new observations about the large-scale structures of the moving atmosphere. There are clearly two main avenues to explore: either assume viscous flow or take the inviscid limit; the former produces closed-form solutions only in special cases (e.g. $M =$ constant), but the latter offers many possibilities. Indeed, because of the very large Reynolds number (see §2), it might be argued that, in any event, the role of viscosity can be ignored. So we look, firstly, at some inviscid flows. A comment about viscous flows will be presented in §7.

The inviscid limit of equations (4.12)–(4.14) can be written as

$$\frac{\partial U}{\partial t} - 2V\sin\theta = -\frac{1}{\cos\theta}\frac{\partial F}{\partial \varphi}, \tag{5.1}$$

$$\frac{\partial V}{\partial t} + 2U\sin\theta = -\frac{\partial F}{\partial \theta} \tag{5.2}$$

and

$$\frac{\partial U}{\partial \varphi} + \frac{\partial}{\partial \theta}(V\cos\theta) + \frac{\partial}{\partial \zeta}(V\sin\theta\cos^2\theta + gW\cos\theta) = \cos\theta\frac{\partial^2 F}{\partial \zeta \partial t}, \tag{5.3}$$

where we have introduced

$$\rho_0(u_0, v_0, w_0, F_1) = (U, V, W, F), \tag{5.4}$$

and so the explicit dependence on the background state is suppressed. This system provides a complete description of the (inviscid) velocity field at this order, for a given forcing F; this F can, in turn, be used to produce the perturbation to the thermodynamic state (given by p_1, ρ_1, T_1), together with the identification of any required heat sources. It is of some significance that equations (5.1) and (5.2) take a very familiar form (see [19]), but the forcing here drives the leading-order velocity field, not a perturbation of it about some uniform state. On the other hand, if we assume that $W = 0$ (which is not a consistent choice within the thin-shell approximation), then equation (5.3), combined with the other two, produces a necessary constraint on F in order for such a solution to exist: so, we must expect that we can allow only very special heat sources to drive the atmospheric flow, in this case. We are now in a position to explore, in some detail, what equations (5.1)–(5.3) tell us about unsteady motions in the atmosphere: we outline the type of results that can be obtained by presenting a few examples, together with some general observations and a case study.

(a) Solution harmonic in time

The simplest solution to seek is one in which all the variables are proportional to $e^{-\mathrm{i}\omega t}$, for some real constant ω; so we write

$$(U, V, W, F) = (\widehat{U}, \widehat{V}, \widehat{W}, \widehat{F})\,\mathrm{e}^{-\mathrm{i}\omega t},$$

where each of the 'hatted' functions depends on (φ, θ, ζ). From equations (5.1) and (5.2), we find directly that

$$\widehat{U} = \frac{1}{\omega^2 - 4\sin^2\theta}\left\{2\sin\theta\,\frac{\partial \widehat{F}}{\partial \theta} - \mathrm{i}\,\frac{\omega}{\cos\theta}\,\frac{\partial \widehat{F}}{\partial \varphi}\right\} \tag{5.5}$$

and

$$\widehat{V} = \frac{1}{\omega^2 - 4\sin^2\theta}\left\{-\mathrm{i}\omega\,\frac{\partial \widehat{F}}{\partial \theta} - 2\tan\theta\,\frac{\partial \widehat{F}}{\partial \varphi}\right\}. \tag{5.6}$$

It should be noted that, if this horizontal velocity field is to be defined for $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$ (so avoiding the Poles; see below), we must have $|\omega| > 2$. We now use (5.5) and (5.6) in (5.3), and so obtain the equation relating $\widehat{W}$ and $\widehat{F}$:

$$\begin{aligned} g\cos\theta\,\frac{\partial \widehat{W}}{\partial \zeta} &+ \frac{1}{\omega^2 - 4\sin^2\theta}\,\frac{\partial}{\partial \varphi}\left(2\sin\theta\,\frac{\partial \widehat{F}}{\partial \theta} - \mathrm{i}\,\frac{\omega}{\cos\theta}\,\frac{\partial \widehat{F}}{\partial \varphi}\right) \\ &- \frac{\partial}{\partial \theta}\left\{\frac{1}{\omega^2 - 4\sin^2\theta}\left(\mathrm{i}\,\omega\cos\theta\,\frac{\partial \widehat{F}}{\partial \theta} + 2\sin\theta\,\frac{\partial \widehat{F}}{\partial \varphi}\right)\right\} \\ &- \frac{\sin\theta\cos\theta}{\omega^2 - 4\sin^2\theta}\,\frac{\partial}{\partial \zeta}\left(\mathrm{i}\,\omega\cos\theta\,\frac{\partial \widehat{F}}{\partial \theta} + 2\sin\theta\,\frac{\partial \widehat{F}}{\partial \varphi}\right) = -\mathrm{i}\,\omega\cos\theta\,\frac{\partial \widehat{F}}{\partial \zeta}. \end{aligned} \tag{5.7}$$

Since $\widehat{W} = 0$ at ground level, from equation (5.7) we may determine the vertical velocity component, $w_0 = \widehat{W}/\rho_0$, given the forcing function $\widehat{F}$ (which, in turn, produces the horizontal velocity components and the perturbation to the thermodynamic state, and can be used to identify the associated heat sources). Note that the neighbourhood of the Poles must be avoided, because this is where $\widehat{W}$ is undefined, according to equation (5.7). To be more precise, we cannot allow θ to be in an $\mathrm{O}(\varepsilon)$-neighbourhood of the Poles, for general $\widehat{F}$, because the asymptotic expansions will not then be valid. Of course, the choice of $\widehat{F}$ may include the requirement that $\widehat{W}$ remains bounded—a reasonable condition on physical grounds—close to the Poles.

We can investigate a little further: equation (5.7) admits solutions for $\widehat{F}$ harmonic in φ and ζ (but not in θ); we set

$$\widehat{F} = \bar{F}(\theta)\, \mathrm{e}^{\mathrm{i}(k\varphi + l\zeta)},$$

where k and l are constants (and we are interested in real k, so that we have a component of propagation in the azimuthal direction, but l may be complex-valued); we do not impose the same structure on $\widehat{W}$ since it would amount to $\widehat{W} \equiv 0$, given that $\widehat{W} = 0$ at ground level. The resulting equation relating $\bar{F}$ and $\widehat{W}$ is

$$\begin{aligned} &\frac{\mathrm{d}^2\bar{F}}{\mathrm{d}\theta^2} + \left\{\left(\mathrm{i}l + \frac{8}{\omega^2 - 4\sin^2\theta}\right)\sin\theta\cos\theta - \tan\theta\right\}\frac{\mathrm{d}\bar{F}}{\mathrm{d}\theta} \\ &\quad + \left\{-\frac{k^2}{\cos^2\theta} + 2\,\frac{k}{\omega} + 2\mathrm{i}\,\frac{lk}{\omega}\,\sin^2\theta - \mathrm{i}l(\omega^2 - 4\sin^2\theta) + 16\,\frac{k}{\omega}\,\frac{\sin^2\theta}{\omega^2 - 4\sin^2\theta}\right\}\bar{F} \\ &\quad + \frac{\mathrm{i}g}{\omega}\,(\omega^2 - 4\sin^2\theta)\,\frac{\partial\widehat{W}}{\partial\zeta} = 0. \end{aligned} \tag{5.8}$$

(i) Reappraisal of the classical dispersion relation

While the dependence on θ in equation (5.8) makes any general further development challenging, we can be confident that we have the correct representation of the behaviour in the meridional direction. It is therefore this equation that should be examined if we require a consistent representation in the neighbourhood of particular θ; this we now do by seeking a harmonic solution close to $\theta = \theta_0$. To accomplish this, we evaluate all the coefficients in equation (5.8) on $\theta = \theta_0$, and then seek a zonally harmonic solution:

$$\bar{F} = F_0\, \mathrm{e}^{\mathrm{i}m(\theta - \theta_0)},$$

where m is a constant, not necessarily real. The resulting 'dispersion' relation becomes

$$\begin{aligned} F_0\Bigg\{&2kl\sin^2\theta_0\cos\theta_0 + \frac{8\omega m\sin\theta_0\cos^2\theta_0}{\omega^2 - 4\sin^2\theta_0} - \omega m\sin\theta_0 - \omega l(\omega^2 - 4\sin^2\theta_0)\cos\theta_0 \\ &+ \mathrm{i}\left(\frac{\omega k^2}{\cos\theta_0} + \omega m^2\cos\theta_0 + \omega ml\sin\theta_0\cos^2\theta_0 - 2k\cos\theta_0 - \frac{16k\sin^2\theta_0\cos\theta_0}{\omega^2 - 4\sin^2\theta_0}\right)\Bigg\} \\ &+ g\cos\theta_0\,(\omega^2 - 4\sin^2\theta_0)\,\frac{\partial\widehat{W}}{\partial\zeta} = 0. \end{aligned} \tag{5.9}$$

We should note that, if we seek such a solution, harmonic in (φ, ζ, t) and locally harmonic in θ, which is then evaluated for zero heat input, i.e. $F_0 = 0$, then there is no motion at this order. Equation (5.9) is not a dispersion relation in any conventional sense, unless some very special assumptions are made about $\partial\widehat{W}/\partial\zeta/F_0$; rather, this equation determines $\widehat{W}$, given F_0 and all the wavenumbers and the frequency. However, our approach enables us to explain how a familiar result can be recovered. In equation (5.9), we ignore all the terms that are generated by the existence of a background state or that arise from the spherical geometry other than the term involving $\partial\widehat{W}/\partial\zeta$; in addition, we assume that there is no contribution from a heat source, which then leaves us with

$$F_0\left\{\frac{\omega k^2}{\cos\theta_0} + \omega m^2\cos\theta_0\right\} - \mathrm{i}g\cos\theta_0\,(\omega^2 - 4\sin^2\theta_0)\,\frac{\partial\widehat{W}}{\partial\zeta} = 0.$$

Even though θ_0 may take any value, we must insist that $\cos\theta_0 = 1$ in this context; finally, we also assume that there is an auxiliary condition (which is not available at this order), which gives

$$\frac{\frac{\partial\widehat{W}}{\partial\zeta}}{F_0} = -\frac{\omega l}{gN^2}$$

and so

$$\omega\{l^2\omega^2 - [(k^2+m^2)N^2] + 4l^2\sin^2\theta_0]\} = 0,$$

the standard result involving the Brunt–Väisälä frequency, N; see, for example, [19]. In other words, the familiar result has no bearing on, or relevance to, the movement of the atmosphere: it cannot be recovered in any systematic analysis of the governing equations. (The errors arise as follows: (i) it is inconsistent to replace $\cos\theta$ by 1 in the evaluation of a term such as $(\partial/\partial\theta)\{\cos\theta(\partial/\partial\theta)(v\cos\theta)\}$, (ii) there is no background state and (iii) there is no equation available, at this order, to allow the introduction of N^2.)

We now return to the general time-harmonic version of our equations, (5.5)–(5.7), and seek solutions that are exponential in φ and ζ:

$$\widehat{F}(\varphi,\theta,\zeta) = \mathrm{e}^{\alpha\varphi+\gamma\zeta} f(\theta), \tag{5.10}$$

where α and γ are constants (which may be chosen to be real or complex, as appropriate, in specific cases). Thus equations (5.5) and (5.6) become

$$\widehat{U} = \frac{1}{\omega^2 - 4\sin^2\theta}\left(2\sin\theta\,\frac{\mathrm{d}f}{\mathrm{d}\theta} - \mathrm{i}\,\frac{\alpha\omega}{\cos\theta}f\right)\mathrm{e}^{\alpha\varphi+\gamma\zeta} \tag{5.11}$$

and

$$\widehat{V} = -\frac{1}{\omega^2 - 4\sin^2\theta}\left(\mathrm{i}\omega\,\frac{\mathrm{d}f}{\mathrm{d}\theta} + 2\alpha\tan\theta f\right)\mathrm{e}^{\alpha\varphi+\gamma\zeta}, \tag{5.12}$$

respectively; correspondingly, equation (5.7), after an integration in ζ and imposing the boundary condition $\widehat{W} = 0$ at the bottom of the atmosphere (i.e. on $\zeta = -\frac{1}{2}\cos^2\theta$), becomes

$$\begin{aligned}\widehat{W} = \frac{1}{g\gamma\cos\theta}\Bigg\{&\frac{\sin(2\theta)}{2(\omega^2 - 4\sin^2\theta)}\left(\mathrm{i}\omega\gamma\cos\theta\,\frac{\mathrm{d}f}{\mathrm{d}\theta} + 2\alpha\gamma f\sin\theta\right)\\ &- \frac{1}{\omega^2 - 4\sin^2\theta}\left(2\alpha\sin\theta\,\frac{\mathrm{d}f}{\mathrm{d}\theta} - \frac{\mathrm{i}\omega\alpha^2}{\cos\theta}f\right) - \mathrm{i}\omega\gamma f\cos\theta\\ &+ \frac{\mathrm{d}}{\mathrm{d}\theta}\left[\frac{1}{\omega^2 - 4\sin^2\theta}\left(\mathrm{i}\omega\cos\theta\,\frac{\mathrm{d}f}{\mathrm{d}\theta} + 2\alpha f\sin\theta\right)\right]\Bigg\}\left(\mathrm{e}^{\gamma\zeta} - \mathrm{e}^{-\frac{1}{2}\gamma\cos^2\theta}\right)\mathrm{e}^{\alpha\varphi}.\end{aligned} \tag{5.13}$$

(ii) Standing waves

As a simple illustration of time-harmonic solutions, we consider flows oscillating in the region of the tropics (so we take $|\theta| \leq 23^\circ$). Let $\widehat{W} = 0$ (so there is no motion in the vertical direction and then the bottom boundary condition is automatically satisfied), with a heat forcing $\widehat{F}$ that depends on only θ (so we take $\alpha = \gamma = 0$ in (5.10)). With these conditions used in equation (5.13), we find that

$$\frac{\mathrm{d}f}{\mathrm{d}\theta} = \frac{\omega^2 - 4\sin^2\theta}{\cos\theta} f_0 \mathrm{e}^{\mathrm{i}\theta_0},$$

for some real constants $f_0 > 0$ and $\theta_0 \in [0, 2\pi)$, and then, using (5.11) and (5.12) in (5.5) and (5.6), respectively, we obtain

$$\widehat{U} = 2f_0\,\mathrm{e}^{\mathrm{i}(\theta_0-\omega t)}\tan\theta, \quad \widehat{V} = \frac{\omega f_0}{\cos\theta}\,\mathrm{e}^{\mathrm{i}(-(\pi/2)+\theta_0-\omega t)} \quad \text{and} \quad \widehat{W} = 0.$$

This solution represents a standing wave, i.e. oscillating in time but not propagating. We note that, even though the expressions derived above exhibit no variation in the vertical direction, (5.4) shows that there is a vertical structure, by virtue of ρ_0, in the horizontal velocity field (u_0, v_0).

(iii) Modulated travelling wave

Waves propagating towards the Equator (for example, as shown in figure 2) can be obtained from our system (5.11)–(5.13) by setting $\alpha = 0$ (so there is no component of propagation in the azimuthal

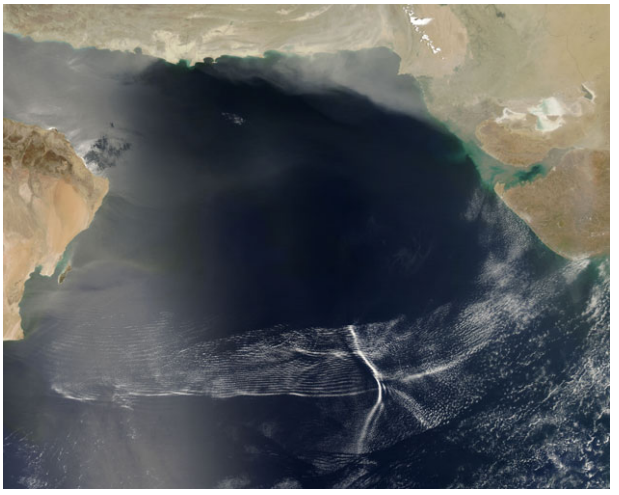

Figure 2. Natural-colour photograph of atmospheric waves spanning the Arabian Sea from Oman to India, taken on 8 May 2007, by NASA's Acqua satellite (Image Credit: NASA). A low-level, high-speed jet stream (the Findlater Jet; see figure 3) cuts across the centre of the wave pattern, creating a distinctive parabolic-shaped cloud; clouds of dust from Iran, Pakistan and Oman are also visible. The air flow roughens the surface of the ocean: bands of light and dark water that mimic the wave pattern are visibile near the shore of Oman, where the sunlight reflected off the water is picked up directly by the satellite's sensor (with a rough water surface dispersing light and thus creating the dark bands, while calm water is brighter). (Online version in colour.)

direction) and then by choosing a suitable $f(\theta)$. A simple choice, which produces no motion on the Equator, is

$$\frac{1}{\omega^2 - 4\sin^2\theta}\frac{\mathrm{d}f}{\mathrm{d}\theta} = A\mathrm{e}^{\mathrm{i}\lambda\theta}\sin\theta,$$

where A is an arbitrary (complex) constant and λ some real constant. Then, from (5.11) and (5.12), and reinstating the harmonic time-dependence, we obtain a modulated wave motion described by

$$U = 2A\mathrm{e}^{\mathrm{i}(\lambda\theta-\omega t)+\gamma\zeta}\sin^2\theta \quad \text{and} \quad V = -\mathrm{i}\omega A\mathrm{e}^{\mathrm{i}(\lambda\theta-\omega t)+\gamma\zeta}\sin\theta,$$

where γ may be chosen to be complex-valued (to produce an exponential behaviour with height, together with a tilt in the direction of propagation). Indeed, we could simplify further by imposing the condition $|\gamma| \to 0$, which produces the simpler solution

$$U = 2A\mathrm{e}^{\mathrm{i}(\lambda\theta-\omega t)}\sin^2\theta \quad \text{and} \quad V = -\mathrm{i}\omega A\mathrm{e}^{\mathrm{i}(\lambda\theta-\omega t)}\sin\theta$$

and

$$W = \frac{\mathrm{i}\omega A}{g}\frac{\zeta + \frac{1}{2}\cos^2\theta}{\cos\theta}\left(\cos(2\theta) + \frac{1}{2}\mathrm{i}\lambda\sin(2\theta)\right)\mathrm{e}^{\mathrm{i}(\lambda\theta-\omega t)}.$$

This solution, we observe, describes W varying linearly with height, but our development makes clear that many other choices are available, for which suitable heat sources can be identified.

The above analysis demonstrates that we can reasonably expect to match what is observed. Of course, the identification of flow structures and processes from the various available satellite data is rather subjective [22], but such exercises nevertheless provide major sources of insight. In the case of the wave pattern visible in figure 2, it is observed, during May, that the surface temperature of the north Indian Ocean becomes the highest among the world's ocean surface temperatures, often in excess of 28°C over the Arabian Sea, where evaporation largely exceeds precipitation; see [23]. The sea-surface temperature is comparable with the daily maximal land temperature in the coastal areas adjacent to the Arabian Sea, with a significant drop in temperature at night over the land. These changes in the temperature over space and time can give rise to wave patterns

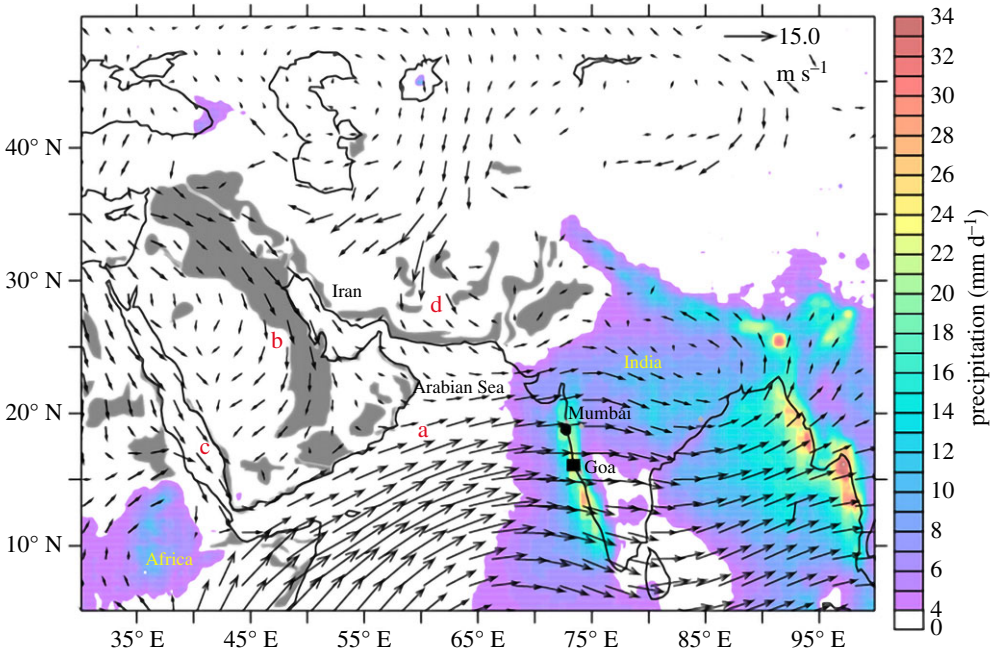

Figure 3. Climatology (2003–2014) of the major winds at the 850 hPa level (about 1.5 km above sea level) over the northern Indian Ocean during the summer monsoon period, with the major dust-source regions shown in grey: (*a*) the Findlater Jet (see [24]); (*b*) the Shamal winds; (*c*) the Red Sea Winds; (*d*) the Levar Winds. Reproduced from [25], CC by Springer Nature. (Online version in colour.)

of the type shown in figure 2. In addition, there is a strong eastward flow near the Equator, as shown in figure 3. In an attempt to recover this general flow pattern, we make a different ansatz (but still invoking $|\gamma| \to 0$ for simplicity), which produces an appropriate solution in equatorial regions: set $\alpha = ia$ where a is a real constant, and choose $f(\theta) = b$, where b is a complex constant. This results in the solution

$$U = \frac{\omega a b}{(\omega^2 - 4\sin^2\theta)\cos\theta}\, e^{i(a\varphi - \omega t)} \quad \text{and} \quad V = -\frac{2iab\tan\theta}{\omega^2 - 4\sin^2\theta}\, e^{i(a\varphi - \omega t)},$$

which describes a wave propagating in the azimuthal direction (with $V = 0$ and $U \neq 0$ along the Equator), and

$$W = \frac{iab\left(\zeta + \frac{1}{2}\cos^2\theta\right)}{g(\omega^2 - 4\sin^2\theta)^2\cos^2\theta}\left[2(\omega^2 + 4\sin^2\theta)\cos^2\theta - \omega a(\omega^2 - 4\sin^2\theta)\right] e^{i(a\varphi - \omega t)}.$$

These two examples demonstrate that our single formulation can capture different modes and types of wave propagation in the atmosphere, avoiding the need for *ad hoc* modelling.

(b) Equatorially trapped waves

Due to the nature of their energy sources, as well as the small Coriolis forcing, large-scale equatorial atmospheric flows present specific structural features. Synoptic-scale disturbances of the atmosphere outside the tropics are mainly driven by horizontal temperature gradients, the primary energy source being the potential energy associated with the latitudinal temperature gradient [19]. In the tropics, however, horizontal temperature gradients are very small, and the energy harnessed by atmospheric flows comes from diabatic heating due to latent heat release, mostly occurring in association with moist convection. The cloudiest regions on Earth (with the oceans significantly cloudier than land) are the tropics, and the higher temperatures

in these regions enhance water evaporation into the air, so that the atmosphere becomes very humid. The rapid drop of temperature and pressure with altitude causes some of the water vapour in the air to condense, forming clouds (on average 19°K colder than the ground) and, if enough water condenses, the cloud droplets can become large enough to fall as rain. There is a considerable release of heat during precipitations, which alters the thermal equilibrium and triggers an adjustment process by means of fast propagating waves (with speeds of the order of $20\,\mathrm{m\,s^{-1}}$; see [26]). This local reaction of the atmospheric flow may also induce remote responses through the excitation of equatorial waves. In particular, equatorially trapped waves are related to the heat forcing that occurs over the Indonesian 'maritime continent', an area straddling the Equator in the middle of the warmest body of water (between the Indian and western Pacific warm pools), whose land-sea heat-capacity differences (due to the mixture of ocean and islands with long coastlines) generate the world's largest rainfall on diurnal cycles, with enormous energy release by convective condensation. With all this in mind, we investigate how our system of equations might model some aspects of these phenomena.

We start by imposing a vanishing meridional velocity component (a reasonable assumption in equatorial regions), and then the system (5.1)–(5.3) reduces to

$$\frac{\partial U}{\partial t} = -\frac{1}{\cos\theta}\,\frac{\partial F}{\partial \varphi}, \tag{5.14}$$

$$2U\sin\theta = -\frac{\partial F}{\partial \theta} \tag{5.15}$$

and

$$\frac{\partial U}{\partial \varphi} + g\cos\theta\,\frac{\partial W}{\partial \zeta} = \cos\theta\,\frac{\partial^2 F}{\partial \zeta\,\partial t}. \tag{5.16}$$

We now seek a solution generated by a heat forcing $\tilde{F}(\eta,\theta,\zeta)$, where $\eta = \varphi - ct$, which is azimuthally localized (that is, it vanishes for $|\eta-\eta_0|\geq\alpha$, for suitable η_0 and some $\alpha>0$, which is taken to be fairly small); furthermore, this function is to be symmetric about the Equator and describes a heat source that moves in the zonal direction at constant speed, c. From (5.15), we obtain

$$U(\eta,\theta,\zeta) = -\frac{1}{2\sin\theta}\,\frac{\partial \tilde{F}}{\partial \theta}(\eta,\theta,\zeta), \tag{5.17}$$

and the meridional symmetry of $\tilde{F}$ ensures (by l'Hospital's rule) that U is not singular along $\theta=0$. Combining (5.14) with (5.17) yields the differential equation

$$\frac{\partial}{\partial\theta}\left(\frac{\partial \tilde{F}}{\partial \eta}\right) = -\left(\frac{2}{c}\tan\theta\right)\frac{\partial \tilde{F}}{\partial \eta}, \tag{5.18}$$

which can be integrated directly to give

$$\frac{\partial \tilde{F}}{\partial \eta}(\eta,\theta,\zeta) = A(\eta,\zeta)\,[\cos\theta]^{2/c},$$

for some function $A\geq 0$ (because the heat forcing decreases poleward) and such that $A(\eta,\zeta)=0$ for $|\eta-\eta_0|\geq\alpha$. Since $\tilde{F}(\eta,\theta,\zeta)$ must vanish for $|\eta-\eta_0|\geq\alpha$, integration gives

$$\tilde{F}(\eta,\theta,\zeta) = B(\eta,\zeta)\,[\cos\theta]^{2/c}, \tag{5.19}$$

where $B(\eta,\zeta)=\int_{\eta_0-\alpha}^{\eta} A(s,\zeta)\,\mathrm{d}s$ and then

$$\int_{\eta_0-\alpha}^{\eta_0+\alpha} A(s,\zeta)\,\mathrm{d}s = 0, \tag{5.20}$$

to ensure that $B(\eta,\zeta)=0$ for $|\eta-\eta_0|\geq\alpha$. From (5.17), we obtain

$$U(\eta,\theta,\zeta) = \frac{1}{c}\,B(\eta,\zeta)\,[\cos\theta]^{(2/c)-1}. \tag{5.21}$$

Since $A\geq 0$ and the heat forcing decreases away from the Equator, from (5.19), we infer that necessarily $c>0$. Furthermore, (5.21) demonstrates two essential properties: eastward

propagation (because $c > 0$) and equatorial trapping. To gain insight into the meridional decay (i.e. equatorial trapping) of the flow, note that the non-dimensional speed c corresponds to $c\Omega' d'_E$ (in $\mathrm{m\,s^{-1}}$), and then $c \approx 23$ produces a realistic value of $20\,\mathrm{m\,s^{-1}}$. For this value, (5.21) predicts a threefold reduction at 12 degrees of latitude, while at the boundary of the tropics, corresponding to a latitude of about 23 degrees, the flow attains merely 2% of its equatorial strength. On the other hand, combining (5.16) and (5.19)–(5.21) yields

$$\frac{\partial W}{\partial \zeta} = -\frac{1}{gc}\,[\cos\theta]^{(2/c)-2}\left\{A + c^2\cos^2\theta\,\frac{\partial A}{\partial \zeta}\right\}. \tag{5.22}$$

The fact that $W = 0$ at ground level $\zeta = -(1/2)\cos^2\theta$ rules out the possibility that $\partial W/\partial\zeta < 0$ somewhere at ground level. Since (5.20) ensures the existence of regions where $A > 0$, we conclude from (5.22) that A (and therefore the heat forcing) must feature variations in the vertical direction, with $\partial A/\partial\zeta < 0$ at ground level. Field data (see fig. 12 in [27]) confirms that the heat forcing typically decreases in the lower and in the upper tropical troposphere, and it increases with respect to height at mid-altitudes (roughly between 2 and 8 km).

The classical theoretical approach for describing equatorially trapped waves (see [19]) concentrates on their horizontal structure, using a shallow water model (for a fluid system of mean depth h_e in a motionless basic state) and the equatorial β-plane approximation. While the predictions made by relying on these approximations (non-dispersive, eastward propagating trapped waves) are somewhat similar to ours, there are significant differences. Firstly, approximating $\sin\theta$ and $\cos\theta$ by the first two terms in their Taylor expansions near $\theta = 0$ leads to an exponential meridional decay rate; ours is not so severe because we have not approximated the θ-dependence. Secondly, the wave speed is taken to be $c = \sqrt{gh_e}$, i.e. that for shallow water gravity waves, although an adequate choice of h_e is somewhat elusive. In contrast to this, since $F = \rho_0 F_1$, we deduce from (4.5) that the value of F at ground level is determined by that of the perturbation pressure, which therefore, due to (5.19), determines the speed c. The third aspect is more fundamental: the classical approach ignores variations of the heat forcing in the vertical direction, while the above discussion of the implications of (5.22) shows that this is an essential aspect of the dynamics. In this context, a simple example of type (5.19)–(5.20) that captures the main observed features is

$$F(\varphi,\theta,\zeta) = \mathfrak{a}(\varphi)\,H(\zeta)\,[\cos\theta]^{2/c}, \tag{5.23}$$

where, in analogy to the considerations discussed in [28],

$$\mathfrak{a}(\varphi) = \begin{cases} \cos\left[\dfrac{\pi}{2\alpha}\,(\varphi-\varphi_0)\right], & |\varphi-\varphi_0| < \alpha, \\ 0, & |\varphi-\varphi_0| \geq \alpha, \end{cases}$$

and where $H(\zeta)$ is the restriction of a cubic polynomial in ζ to the interval $[-(1/2), g]$, which decreases for small and large values of ζ and increases in the middle part of the interval (thus capturing the previously discussed typical monotonicity with height of the heat forcing throughout the tropical troposphere). Such a polynomial can be easily obtained by means of Lagrange interpolation, using field data to identify the relevant values at the top and bottom of the troposphere, and at the two critical points in-between, where the monotonicity of the heat-forcing-profile changes.

(c) No additional constraints: general unsteady motion

The horizontal, inviscid velocity field can be obtained directly from equations (5.1) and (5.2) without imposing any additional constraints. For fixed (φ,θ,ζ), we write (5.1)–(5.2) as the

inhomogeneous linear differential system

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} 0 & 2\sin\theta \\ -2\sin\theta & 0 \end{pmatrix}\begin{pmatrix} U \\ V \end{pmatrix} + \begin{pmatrix} -\dfrac{1}{\cos\theta}\dfrac{\partial F}{\partial \varphi} \\ -\dfrac{\partial F}{\partial \theta} \end{pmatrix},$$

whose general solution is given by the variation-of-constants formula (see [29])

$$\begin{pmatrix} U(\varphi,\theta,\zeta,t) \\ V(\varphi,\theta,\zeta,t) \end{pmatrix} = \mathrm{e}^{\mathcal{A}t}\begin{pmatrix} C(\varphi,\theta,\zeta) \\ D(\varphi,\theta,\zeta) \end{pmatrix} + \int_0^t \mathrm{e}^{\mathcal{A}(t-s)}\begin{pmatrix} -\dfrac{1}{\cos\theta}\dfrac{\partial F}{\partial \varphi}(\varphi,\theta,\zeta,s) \\ -\dfrac{\partial F}{\partial \theta}(\varphi,\theta,\zeta,s) \end{pmatrix}\mathrm{d}s, \tag{5.24}$$

where C and D are arbitrary functions, and

$$\mathrm{e}^{\mathcal{A}t} = \sum_{k\geq 0}^{\infty} \frac{t^k}{k!}\mathcal{A}^k \quad \text{with } \mathcal{A} = \begin{pmatrix} 0 & 2\sin\theta \\ -2\sin\theta & 0 \end{pmatrix}.$$

The fundamental solution of the associated homogeneous constant-coefficient system

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} 0 & 2\sin\theta \\ -2\sin\theta & 0 \end{pmatrix}\begin{pmatrix} U \\ V \end{pmatrix}$$

can be found directly as

$$\mathrm{e}^{\mathcal{A}t} = \begin{pmatrix} \cos(2t\sin\theta) & \sin(2t\sin\theta) \\ -\sin(2t\sin\theta) & \cos(2t\sin\theta) \end{pmatrix},$$

and consequently

$$\begin{aligned} U(\varphi,\theta,\zeta,t) &= C(\varphi,\theta,\zeta)\,\cos(2t\sin\theta) + D(\varphi,\theta,\zeta)\,\sin(2t\sin\theta) \\ &\quad - \int_0^t \left(\frac{\cos[2(t-s)\sin\theta]}{\cos\theta}\frac{\partial F}{\partial \varphi} + \sin[2(t-s)\sin\theta]\frac{\partial F}{\partial \theta}\right)\mathrm{d}s \end{aligned} \tag{5.25}$$

and

$$\begin{aligned} V(\varphi,\theta,\zeta,t) &= D(\varphi,\theta,\zeta)\,\cos(2t\sin\theta) - C(\varphi,\theta,\zeta)\,\sin(2t\sin\theta) \\ &\quad + \int_0^t \left(\frac{\sin[2(t-s)\sin\theta]}{\cos\theta}\frac{\partial F}{\partial \varphi} - \cos[2(t-s)\sin\theta]\frac{\partial F}{\partial \theta}\right)\mathrm{d}s. \end{aligned} \tag{5.26}$$

Corresponding solutions for W can be obtained from equation (5.3). Furthermore, we see that this describes the solution that arises when the external forcing is zero ($F \equiv 0$): it is unsteady but of a very specific form. This formulation, therefore, enables the solution to be found for any given forcing, the choice of which can be guided by the available data; this is clearly an area for extensive investigation.

6. Case study: the African Easterly Jet/Waves

During the Northern Hemisphere summer, strong heating of the Saharan region in North Africa, and the relatively cool and moist air to its south (in the Gulf of Guinea), creates a situation in which the usual north–south horizontal temperature gradient is reversed in the lower troposphere above the Sahel region, while in the upper troposphere the insolation-induced horizontal decrease of temperature with increasing latitude persists. Observations show the appearance of a strong westward jet, called the African Easterly Jet (AEJ), whose core on the western coast of Africa is near 15° N, at a height of about 4 km (where the reversal of the temperature gradient in the middle troposphere occurs); see [30,31]. Denoting by a bracket the temporal mean (recall that

$3\frac{1}{2}$ h corresponds to a unit interval for the non-dimensional variable t), in the time-periodic setting, we see that (5.1)–(5.3) yield the steady system

$$-2\langle V\rangle \sin\theta = -\frac{1}{\cos\theta}\frac{\partial\langle F\rangle}{\partial\varphi},$$

$$2\langle U\rangle \sin\theta = -\frac{\partial\langle F\rangle}{\partial\theta}$$

and

$$\frac{\partial\langle U\rangle}{\partial\varphi} + \frac{\partial\langle V\rangle}{\partial\theta}\cos\theta - \langle V\rangle \sin\theta + \frac{\partial\langle V\rangle}{\partial\zeta}\sin\theta\cos^2\theta + g\frac{\partial\langle W\rangle}{\partial\zeta}\cos\theta = 0,$$

so that

$$\langle U\rangle = -\frac{1}{2\sin\theta}\frac{\partial\langle F\rangle}{\partial\theta}, \tag{6.1}$$

$$\langle V\rangle = \frac{1}{2\sin\theta\cos\theta}\frac{\partial\langle F\rangle}{\partial\varphi} \tag{6.2}$$

and

$$\frac{\partial\langle W\rangle}{\partial\zeta} = \frac{1}{2g}\left(\frac{1}{\sin^2\theta}\frac{\partial\langle F\rangle}{\partial\varphi} - \frac{\partial^2\langle F\rangle}{\partial\varphi\partial\zeta}\right). \tag{6.3}$$

Since $W = 0$ at the ground level ($\zeta = -(1/2)\cos^2\theta$), integration of equation (6.3) gives

$$\langle W\rangle(\varphi,\theta,\zeta) = -\frac{1}{2g}\frac{\partial\langle F\rangle}{\partial\varphi}\Big|_{\zeta=-(1/2)\cos^2\theta}^{\zeta} + \frac{1}{2g\sin^2\theta}\int_{-(1/2)\cos^2\theta}^{\zeta}\frac{\partial\langle F\rangle}{\partial\varphi}(\varphi,\theta,s)\,\mathrm{d}s. \tag{6.4}$$

The relations (6.1)–(6.3) enable us to infer the basic dynamics of the AEJ from the observed behaviour of the heat forcing F. The heat-flux contribution from clouds is negligible, the main source being the ground-level temperature gradient, distributed in the lower troposphere by dry convection and diffusion (rather than by condensational heating); see the discussion in [30]. Going north from the Guinean coast towards the Sahara, in the region from 10° N to 20° N, the surface temperature increases gradually by about 10° K (see the data in [30]). Thus $\frac{\partial\langle F\rangle}{\partial\theta} > 0$ and (6.1) predicts a westward zonal flow component for the AEJ. On the other hand, the surface temperature in the lower-troposphere Sahel region between 15° W and 10° E increases (weakly) eastwards (see the data in [30]). The fact that $\partial\langle F\rangle/\partial\varphi > 0$ in the lower troposphere yields, from (6.2), that the meridional flow velocity of the AEJ is northward. Field data (see [30]) also show that near 15° N, between 15° W and 10° E, the positive longitudinal gradient of the temperature increases with height near the ground and then starts to decrease towards the mid-troposphere altitude, where it reverses sign: the classic Hadley circulation coexists with a second but shallower overturning circulation in the lower part of the troposphere (see [32]). Due to (6.2), at a fixed latitude θ, we can track the variation of $\partial\langle F\rangle/\partial\varphi$ with height by the corresponding variation of the meridional velocity component, V. Field data for the AEJ (see [30,33]) show that V increases linearly with height in the lower third of the troposphere, from about 1–2 m s^{-1} near the ground, to a maximum of about 6–7 m s^{-1} at an altitude of about 4 km. Given that one non-dimensional speed unit corresponds to $\Omega' H' \sim 1.16$ m s^{-1}, this means that the vertical slope of $\partial\langle F\rangle/\partial\varphi$ near the ground is less than 15. Since $\sin(15°) \sim 0.258$, we see that, for the AEJ, $1/\sin^2\theta$ exceeds the vertical slope of $\partial\langle F\rangle/\partial\varphi$ in the lower third of the troposphere. Consequently, (6.3) confirms that uplifting ($W > 0$) occurs near the ground. We have presented the main qualitative features of the AEJ, but detailed data about the temperature (which is readily available, but its use is outside the scope of this initial investigation) would provide the basis for a quantitative analysis, producing a more comprehensive description.

African easterly waves (AEWs) are westward moving oscillatory disturbances of the AEJ, initiated by mesoscale convective systems over Central Africa and propagating in the lower troposphere; see [34]. The wave periods are within the range of 2–9 days and the wavelengths are about 2000 km, with maximal speeds of about 11 m s^{-1}. A model for these AEWs can be obtained from the system (5.1)–(5.3) by constructing time-dependent perturbations $(\tilde{U},\tilde{V},\tilde{W})$ of the AEJ mean flow, $(\langle U\rangle,\langle V\rangle,\langle W\rangle)$, driven by a suitable time-dependent heat forcing $\tilde{F}$, a perturbation of

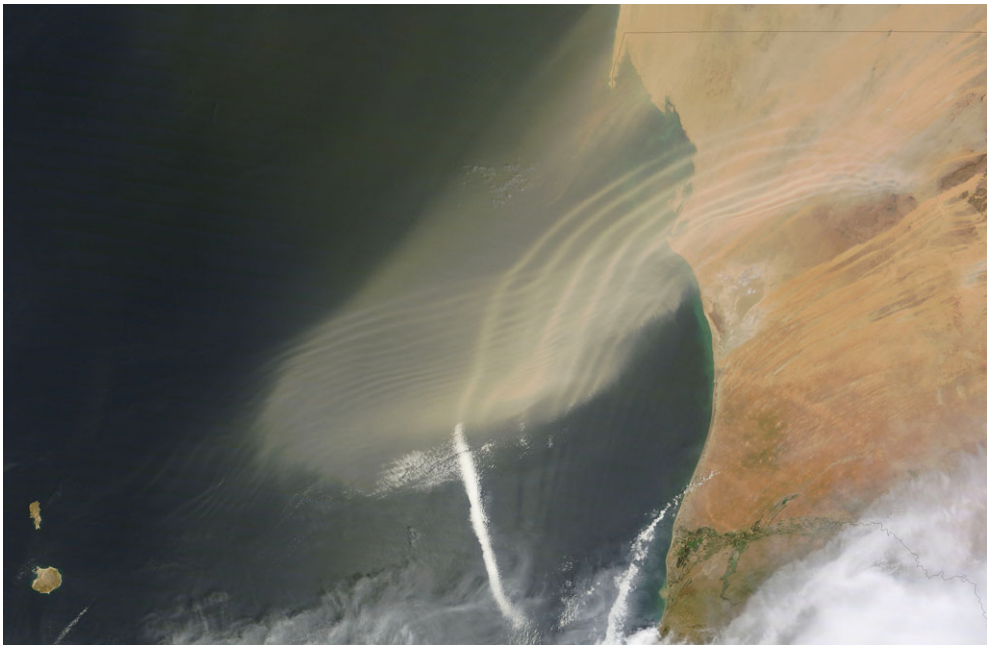

Figure 4. Natural-colour photograph of the north-western coastal region of Africa, taken on 23 September 2011, by NASA's Terra satellite (Image Credit: NASA). This photograph captures two atmospheric waves made visible by dust (sand) blowing from the Sahara Desert, these waves being along the tracks of the two predominant air flows in this region. The waves with longer wavelength (of about 20 km) propagate at about 10 m s^{-1} in the southwest-northeast direction of the African Easterly Jet (AEJ)—a permanent air flow in the lower troposphere, thermally induced by the contrast between the warm Sahara Desert area to the north and the cool Gulf of Guinea to the south—upward to 6 km above the ground, with the core at about 3 km above the ground (where the vertically sheared AEJ is most intense; see [35]). The shorter waves propagating northward at about 3 m s^{-1} arise as perturbations of an ageostrophic meridional circulation in the upper troposphere, mainly in the region of 8–12 km above ground; see [36]. Note that the westward travelling African Easterly Waves, arising as lower-tropospheric three-dimensional synoptic-scale disturbances of the AEJ (see [37]), with a maximal speed of about 11 m s^{-1} (attained around 4 km altitude) and wavelengths of about 2000 km, are not visible in the photograph. These more intricate wave patterns are identified by using theoretical predictions to interpret field data, for example, by performing a wavenumber-frequency spectrum analysis to isolate statistically significant spectral peaks that correspond to the available dispersion relations (see [38]). (Online version in colour.)

the mean heat forcing $\langle F \rangle$. With the mean flow determined by (6.1)–(6.3), the waves are therefore solutions of the system

$$\frac{\partial \tilde{U}}{\partial t} - 2\tilde{V}\sin\theta = -\frac{1}{\cos\theta}\frac{\partial \tilde{F}}{\partial \varphi},$$

$$\frac{\partial \tilde{V}}{\partial t} + 2\tilde{U}\sin\theta = -\frac{\partial \tilde{F}}{\partial \theta}$$

and

$$\frac{\partial \tilde{U}}{\partial \varphi} + \frac{\partial}{\partial \theta}(\tilde{V}\cos\theta) + \frac{\partial}{\partial \zeta}(\tilde{V}\sin\theta\cos^2\theta + g\tilde{W}\cos\theta) = \cos\theta\,\frac{\partial^2 \tilde{F}}{\partial \zeta\,\partial t},$$

whose general solution was described in §5*c*. Waves with shorther wavelengths than the AEWs are also observed (figure 4); the above approach also applies in this case.

7. Viscous flow

We return to the original set of equations that define the velocity field at this order, namely equations (4.12)–(4.14). An analysis of this system is essential if viscous effects are thought to play

a role in the atmospheric flows of interest, which is certainly the case within the atmospheric boundary layer (see the discussion in [39]). The general procedure is exactly as before: solve the first two for the horizontal components, and then the third gives the vertical component (or provides an auxiliary condition, if we choose to set $w_0 = 0$). The simplest way to proceed, as was done in [3], is to introduce the complex velocity:

$$u_0 + \mathrm{i}v_0 = Z_0(\varphi, \theta, \zeta\, t);$$

equations (4.12) and (4.13) then give

$$\frac{\partial Z_0}{\partial t} + 2\mathrm{i}Z_0 \sin\theta = \frac{g^2}{\rho_0 R_e}\frac{\partial}{\partial \zeta}\left(M\frac{\partial Z_0}{\partial \zeta}\right) - \frac{1}{\rho_0}\left(\frac{1}{\cos\theta}\frac{\partial}{\partial \varphi} + \mathrm{i}\frac{\partial}{\partial \theta}\right)(\rho_0 F_1). \tag{7.1}$$

This can be rewritten in a slightly more convenient form by setting

$$Z_0(\varphi, \theta, \zeta\, t) = \mathrm{e}^{-2\mathrm{i}t\sin\theta} Y_0(\varphi, \theta, \zeta\, t),$$

to obtain

$$\frac{\partial Y_0}{\partial t} - \frac{g^2}{\rho_0 R_e}\frac{\partial}{\partial \zeta}\left(M\frac{\partial Y_0}{\partial \zeta}\right) = -\frac{\mathrm{e}^{-2\mathrm{i}t\sin\theta}}{\rho_0}\left(\frac{1}{\cos\theta}\frac{\partial}{\partial \varphi} + \mathrm{i}\frac{\partial}{\partial \theta}\right)(\rho_0 F_1). \tag{7.2}$$

By means of a Liouville substitution, equation (7.2) can be transformed into a (complex-valued) Bessel equation (after the removal of a time-dependent factor), whose general solution can be expressed in terms of Bessel functions of the first kind (at least, for some reasonable choices of M); see [3]. It is clear that the development of these solutions is quite involved but readily accessible, and the results should provide important insights into these flows; this is another area that is left for future investigation.

8. Discussion

The ideas developed in this paper have shown that the theory presented in [3] can be extended to include time dependence (on a suitable time scale, of course). In particular, therefore, we invoke an asymptotic method that is driven by the thin-shell approximation, which represents a minimal assumption for the atmosphere enveloping the Earth. (It is expedient also to take the shape of the Earth as ellipsoidal, and use almost-spherical coordinates; the role of this approximation is simply to adjust the nature of the background state, without affecting the leading-order unsteady description of the atmosphere: the uncoupling at this order is an important property of the underlying system. Thus this aspect of the problem is included in the general formulation, but ignored in the main thesis that we propound here.) As we found in our earlier work [3], the real surprise is that the thin-shell approximation, without recourse to any other assumptions about the nature of the flow, leads to a complete description of the dominant dynamic and thermodynamic elements that are needed to describe the atmosphere. This earlier work showed how standard steady-state models that are used for specific problems in the atmosphere (e.g. Ekman and geostrophic flows, Hadley cell circulation) can be recovered from one set of simple governing equations, and with many accurate interpretations that underlie the flow configurations (such as the precise nature of the heating that is needed to drive the motion). In the current study, our set of governing equations, (4.12)–(4.14), with (4.2) and (4.4), provide the basis for a systematic study of unsteady motion and, most particularly, of various modes of wave propagation in the troposphere.

The development presented here follows the path originated in [3]: a steady background state of the atmosphere (described by equations in which the velocity field appears at this same order) is perturbed by using the thin-shell parameter, with time included in the perturbation system. The time scale—a few hours—is that associated with the rotation of the Earth, which is the appropriate choice for reasonably large-scale wave-like motions of the atmosphere. The upshot is that we have a perturbation state that is unsteady, contains the main contributors to the dynamics and thermodynamics, as well as viscosity, compressibility, the Earth's rotation and the underlying spherical geometry. In addition, the formulation allows for the identification of the

heat sources that are needed to initiate and maintain the motion, and this connection is altogether transparent. One main advantage of the careful and systematic construction of a set of equations with a robust pedigree is that we can hope to develop numerical solutions of this simpler system, driven by reliable data. Our systematic approach is based on theoretical first principles, augmented by phenomenological insights (regarding the nature of the heat forcing), and avoids *ad hoc* assumptions. This conceptually coherent modelling brings to light the significant dependency of the atmospheric flow on the heat forcing, avoiding adjustments of model parameterizations that often mask underlying deficiencies of models, which are not derived systematically from the governing equations, a procedure that may fail to capture the relevant physical processes even if, due to parameter calibration, they might exhibit agreement between simulation and data.

Our equations admit solutions that are harmonic in time, and so the majority of our examples are based on solutions that possess this property. A familiar model for simple, oscillatory flows involves the Brunt–Väisälä frequency (N), so we started by investigating how such flows appear in our system. On the one hand, we see that N^2 is directly and simply related to our forcing function (F_1) and so its form for various flows can be interpreted in terms of the associated heat sources. However, the standard analysis, which leads to the usual dispersion relation involving N^2, cannot be obtained in any systematic way (as explained in §5a(i)): we conclude that any such 'derivations', and conclusions based on the results, remain highly suspect. Fortunately, there are many other avenues that we can explore which lead to relevant and reliable results, with applications to the motion of the atmosphere.

The main thrust of the examples that we have presented here is for inviscid flows; viscous effects have been included in the formulation, but are weak (and relevant only in thin boundary layers) and are therefore set aside in much of what we present. The examples are intended to show what is possible; clearly, many other choices can be made, suitable for other flow configurations. (It should be noted that, in all cases, the existence of a background state, which varies in $\zeta = gz - (1/2)\cos^2\theta$, ensures that there is always a variation in the vertical direction, with an associated distortion in the meridional direction.) The system that describes the dynamic-thermodynamic balance, with heat sources, admits solutions that are harmonic in t, φ and ζ, but not in θ: the dependence on θ necessarily involves some modulation in the meridional direction. With these points in mind, we have obtained a standing-wave solution, which oscillates in time only, as well as solutions that describe waves propagating in the meridional direction, and waves in the azimuthal direction. The former provides a model for waves that propagate towards the Equator (over the Arabian Sea); the latter is a solution of the type observed to be moving in an equatorial direction over the Indian Ocean. These solutions, and all similar ones, possess a vertical structure, provided in part by the existence of the background state, and all satisfying the condition $W = 0$ at the bottom of the atmosphere. We expect that many other examples, relevant to other modes and directions of propagation, with varying complexities of modulation in the meridional direction, can be identified; this is an area that needs further investigation.

Another example, which shows in greater depth the advantages of a general, all-embracing system of equations, arises when we consider waves that are trapped in the neighbourhood of the Equator. In contrast to the familiar approach, we need make no assumptions about the behaviour in the meridional direction: f- or β-plane approximations are altogether unnecessary. Simply by seeking a solution that imposes no motion in the meridional direction, together with a suitable heat forcing, leads to a solution that is restricted to a neighbourhood of the Equator, the trapping being represented by a power-law behaviour (and not exponential as in the simple theories). This example, perhaps more so than the previous ones described above, shows one of the significant advantages of our formulation: we can identify and interpret the heat sources required to maintain the motions.

One final example, discussed in some detail, shows how a time-averaged version of the equations, for periodic solutions, can be used to provide some general observations without recourse to the construction of explicit solutions. In particular, this procedure was applied to a description of the AEJ, showing how the temperature profile in the troposphere necessarily produces the observed properties of this atmospheric flow. Indeed, this calculation can be

extended to include the African Easterly Wave as a time-dependent perturbation of the AEJ, although the details have not been pursued here.

There are many different types of unsteady motion in the atmosphere (including various instabilities) see [40]; we have chosen, as simple examples, a few modes of wave propagation (of which there are many) in this initial investigation. These have been chosen to demonstrate how details of the motion can be extracted and interpreted and, where appropriate, how general properties can be identified without the need for explicit solutions. In addition to these examples, we have also shown how the general time-dependent problem can be formulated, avoiding any assumptions about the nature of the time-like behaviour. Indeed, if the heat forcing is given, then the complete dynamical structure of the atmosphere can be calculated (at leading order). Furthermore, as with all the examples discussed here, the specification of the heat forcing, F_1, can be used to provide the thermodynamic elements associated with the motion and lead to the interpretation of the required heat sources. All this, we emphasize, is based on a single set of simplified equations that are robustly connected to the underlying, governing equations for a fluid. Finally, we have outlined how the corresponding viscous problem can be solved, although the details are less readily accessible; more information on the viscous problem can be found in [3].

What we have developed and described here should be, we submit, the basis for the study of unsteady atmospheric flows. On the one hand, we have a set of equations derived carefully (using precise asymptotic methods) incorporating minimal simplifying assumptions; this system provides a reliable starting point for investigation, analysis and interpretation. On the other hand, particularly with the ready availability of extensive and reliable data, these equations can be used to generate numerical solutions, which, in turn, can become the seed for numerical studies of the original, full set of equations. In either case, we have shown how familiar results, and new results, can be obtained directly and systematically from the underlying governing equations. Furthermore, any simpler model should be tested against this system: can such models be obtained by making additional assumptions consistent with these equations, working altogether systematically? We submit that the only reliable validation of model equations is their (asymptotic) derivation from a set of general governing equations, along the lines that we have presented here.

Data accessibility. All data are provided in full in the paper.

Authors' contributions. The authors contributed equally to this study. The authors contributed equally to this study, conceiving its scientific content, collecting the data and drafting the manuscript. Both authors gave final approval for publication and agree to be held accountable for the work performed therein.

Competing interests. We declare we have no competing interests.

Funding. No funding has been received for this article.

Acknowledgements. The authors are grateful for helpful comments from the referees.

References

1. Vallis GK. 2017 *Atmospheric and oceanic fluid dynamics*. Cambridge, UK: Cambridge University Press.
2. Dellar PJ. 2011 Variations on a beta-plane: derivation of non-traditional beta-plane equations from Hamilton's principle on a sphere. *J. Fluid Mech.* **674**, 174–195. (doi:10.1017/S0022112010006464)
3. Constantin A, Johnson RS. 2021 On the modelling of large-scale atmospheric flows. *J. Differ. Equ.* **285**, 751–798. (doi:10.1016/j.jde.2021.03.019)
4. Ghil M, Robertson AW. 2002 'Waves' vs. 'particles' in the atmosphere's phase space: a pathway to long-range forecasting? *Proc. Natl Acad. Sci. USA* **99**, 2493–2500. (doi:10.1073/pnas.012580899)
5. Klein R. 2010 Scale-dependent models for atmospheric flows. *Annu. Rev. Fluid Mech.* **42**, 249–274. (doi:10.1146/annurev-fluid-121108-145537)
6. Majda A. 2003 *Introduction to PDEs and waves for the atmosphere and ocean*. Courant Lecture Notes in Mathematics, vol. 9. Amer. Math. Soc. Providence, RI: New York University.

7. Rossby CG. 1939 Relations between variations in the intensity of the zonal circulation of the atmosphere and the displacements of the semipermanent centers of action. *J. Mar. Res.* **2**, 38–55. (doi:10.1357/002224039806649023)
8. Haurwitz B. 1940 The motion of atmospheric disturbances on the spherical Earth. *J. Mar. Res.* **3**, 254–267.
9. Craig RA. 1945 A solution of the nonlinear vorticity equation for atmospheric motion. *J. Meteor.* **2**, 175–178. (doi:10.1175/1520-0469(1945)002<0175:ASOTNV>2.0.CO;2)
10. Neamtan SM. 1946 The motion of harmonic waves in the atmosphere. *J. Atmos. Sci.* **3**, 53–56. (doi:10.1175/1520-0469(1946)003<0053:TMOHWI>2.0.CO;2)
11. Gill AE. 1982 *Atmosphere-ocean dynamics: an introductory text*. New York, NY: Academic Press.
12. White AA, Hoskins BJ, Roulstone I, Staniforth A. 2005 Consistent approximate models of the global atmosphere: shallow, deep, hydrostatic, quasi-hydrostatic and non-hydrostatic. *Quart. J. R. Met. Soc.* **131**, 2081–2107. (doi:10.1256/qj.04.49)
13. Atkins P. 2010 *The laws of thermodynamics: a very short introduction*. Oxford, UK: Oxford University Press.
14. Curry JA, Webster PJ. 1999 *Thermodynamics of atmospheres and oceans*. New York, NY: Academic Press.
15. Ghil M, Lucarini V. 2020 The physics of climate variability and climate change. *Rev. Mod. Phys.* **92**, 035002. (doi:10.1103/RevModPhys.92.035002)
16. Lucarini V. 2009 Thermodynamic efficiency and entropy production in the climate system. *Phys. Rev. E* **80**, 021118. (doi:10.1103/PhysRevE.80.021118)
17. Peixoto JP, Oort AH. 1992 *Physics of climate*. New York, NY: Springer.
18. Yang GY, Hoskins BJ, Slingo J. 2003 Convectively coupled equatorial waves: a new methodology for identifying wave structures in observational data. *J. Atmos. Sci.* **60**, 1637–1654. (doi:10.1175/1520-0469(2003)060<1637:CCEWAN>2.0.CO;2)
19. Holton JR, Hakim GJ. 2013 *An introduction to dynamic meteorology*. New York, NY: Academic Press.
20. Marshall J, Plumb RA. 2016 *Atmosphere, ocean and climate dynamics: an introductory text*. New York, NY: Academic Press.
21. Swanson CA. 1968 *Comparison and oscillation theory of linear differential equations*. New York, NY: Academic Press.
22. Birch CE, Reeder MJ, Berry GJ. 2014 Wave-cloud lines over the Arabian Sea. *J. Geophys. Res.: Atmos.* **119**, 4447–4457. (doi:10.1002/2013JD021347)
23. Shenoi SSC, Shankar D, Shetye SR. 2002 Differences in heat budgets of the near-surface Arabian Sea and Bay of Bengal: implications for the summer monsoon. *J. Geophys. Res.: Oceans* **107**, 1–14. (doi:10.1029/2000JC000679)
24. Findlater J. 1969 A major low-level air current near the Indian Ocean during the northern summer. *Quart. J. R. Met. Soc.* **95**, 362–380. (doi:10.1002/qj.49709540409)
25. Ramaswamy V, Muraleedharan PM, Prakash Babu C. 2017 Mid-troposphere transport of Middle-East dust over the Arabian Sea and its effect on rainwater composition and sensitive ecosystems over India. 13676: *Sci. Rep.* **131**, 13676.
26. Wang B. 2003 Kelvin waves. In *Encyclopedia of Meteorology* (ed. J Holton), pp. 1062–1067. New York, NY: Academic Press.
27. Schumacher C, Houze Jr. RA, Kraucunas I. 2004 The tropical dynamical response to latent heating estimates derived from the TRMM precipitation radar. *J. Atmos. Sci.* **61**, 1341–1358. (doi:10.1175/1520-0469(2004)061<1341:TTDRTL>2.0.CO;2)
28. Gill AE. 1980 Some simple solutions for heat-induced tropical circulation. *Quart. J. R. Met. Soc.* **106**, 447–462. (doi:10.1002/qj.49710644905)
29. Meiss JD. 2017 *Differential dynamical systems*. Philadelphia, PA: SIAM.
30. Cook KH. 1999 Generation of the African Easterly Jet and its role in determining West African precipitation. *J. Clim.* **12**, 1165–1184. (doi:10.1175/1520-0442(1999)012<1165:GOTAEJ>2.0.CO;2)
31. Schlueter A, Fink AH, Knippertz GJ. 2019 A systematic comparison of tropical waves over Northern Africa. Part II: dynamics and thermodynamics. *J. Clim.* **32**, 2605–2625. (doi:10.1175/JCLI-D-18-0651.1)
32. Zhang C, Woodworth P, Gu G. 2006 The seasonal cycle in the lower troposphere over West Africa from sounding observations. *Quart. J. R. Met. Soc.* **132**, 2559–2582. (doi:10.1256/qj.06.23)

33. Parker DJ, Thorncroft CD, Burton RR, Diongue-Niang A. 2005 Analysis of the African Easterly Jet using aircraft observations from the JET2000 experiment. *Quart. J. R. Met. Soc.* **131**, 1461–1482. (doi:10.1256/qj.03.189)
34. Bain CL, Williams KD, Milton SF, Heming JT. 2014 Objective tracking of African Easterly Waves in Met Office models. *Quart. J. R. Met. Soc.* **140**, 47–57. (doi:10.1002/qj.2110)
35. Braun SA. 2010 Reevaluating the role of the Saharan Air Layer in Atlantic tropical cyclogenesis and evolution. *Mon. Weather Rev.* **138**, 2007–2037. (doi:10.1175/2009MWR3135.1)
36. Besson L, Lemaitre Y. 2014 Mesoscale convective systems in relation to African and Tropical Easterly Jets. *Mon. Weather Rev.* **142**, 3224–3242. (doi:10.1175/MWR-D-13-00247.1)
37. Kiladis GN, Thorncroft CD, Hall NHJ. 2006 Three-dimensional structure and dynamics of African Easterly Waves. Part I: observations. *J. Atmos. Sci.* **63**, 2212–2230. (doi:10.1175/JAS3741.1)
38. Wheeler M, Kiladis GN. 1999 Convectively coupled equatorial waves: analysis of clouds and temperature in the wavenumber-frequency domain. *J. Atmos. Sci.* **56**, 374–399. (doi:10.1175/1520-0469(1999)056<0374:CCEWAO>2.0.CO;2)
39. Constantin A, Johnson RS. 2019 Atmospheric Ekman flows with variable eddy viscosity. *Bound. Layer Meteorol.* **170**, 395–414. (doi:10.1007/s10546-018-0404-0)
40. Palmer PI. 2017 *The atmosphere: a very short introduction*. Oxford, UK: Oxford University Press.