On the Lagrangian Dynamics of Atmospheric Zonal Jets and the Permeability of the Stratospheric Polar Vortex

I. I. Rypina, M. G. Brown, and F. J. Beron-Vera
Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, Florida

H. Koçak
Departments of Computer Science and Mathematics, University of Miami, Coral Gables, Florida

M. J. Olascoaga and I. A. Udoydchenkov
Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, Florida

(Manuscript received 17 May 2006, in final form 17 January 2007)

ABSTRACT

The Lagrangian dynamics of zonal jets in the atmosphere are considered, with particular attention paid to explaining why, under commonly encountered conditions, zonal jets serve as barriers to meridional transport. The velocity field is assumed to be two-dimensional and incompressible, and composed of a steady zonal flow with an isolated maximum (a zonal jet) on which two or more traveling Rossby waves are superimposed. The associated Lagrangian motion is studied with the aid of the Kolmogorov–Arnold–Moser (KAM) theory, including nontrivial extensions of well-known results. These extensions include applicability of the theory when the usual statements of nondegeneracy are violated, and applicability of the theory to multiply periodic systems, including the absence of Arnold diffusion in such systems. These results, together with numerical simulations based on a model system, provide an explanation of the mechanism by which zonal jets serve as barriers to the meridional transport of passive tracers under commonly encountered conditions. Causes for the breakdown of such a barrier are discussed. It is argued that a barrier of this type accounts for the sharp boundary of the Antarctic ozone hole at the perimeter of the stratospheric polar vortex in the austral spring.

1. Introduction

It is now generally accepted that during the austral winter and spring the Southern Hemisphere stratospheric polar vortex provides an effective barrier to meridional transport of passive tracers. During this period winds at high latitudes in the Southern Hemisphere throughout most of the stratosphere are characterized by a nearly zonal jet; the polar vortex can be defined as the region poleward of the jet core, and available evidence suggests that the transport barrier is nearly coincident with the jet core. The Southern Hemisphere polar vortex often persists throughout the fall, winter, and spring, being strongest in the winter and early spring. The polar vortex in the Northern Hemisphere is generally weaker than its Southern Hemisphere counterpart, and is usually present only in the winter and early spring. The Northern Hemisphere stratospheric polar vortex is less effective as a meridional transport barrier than its Southern Hemisphere counterpart, both because it is generally weaker and shorter lived, and because Rossby wave perturbations to the polar vortex are generally stronger in the Northern Hemisphere winter than in the Southern Hemisphere winter. A more complete discussion of these topics can be found in Andrews et al. (1987), Bowman (1993), Dahlberg and Bowman (1994), McIntyre (1989), and Holton et al. (1995).

Much recent interest in the stratospheric polar vortices derives from observations of the Antarctic ozone hole and, to a lesser extent, its generally less well-
defined Northern Hemisphere counterpart. The annual formation of the Antarctic ozone hole is controlled by chemical processes in the stratosphere (Lefevre et al. 1994; Solomon 1999; Webster et al. 1993). These processes will not be discussed here except to note that both cold temperatures and sunlight are required to trigger the chemical reactions that lead to ozone depletion. These conditions lead to ozone depletion over the pole in the late winter and early spring when the polar vortex is strong. The focus of the work reported here is explaining the mechanism by which the stratospheric polar vortex provides a barrier to the meridional transport of a passive tracer, such as ozone concentration. Such a barrier provides an explanation of why, under typical austral early spring conditions in the middle and upper stratosphere, the ozone-depleted air does not spread via turbulent mixing to midlatitudes.

Our explanation of the mechanism by which the polar vortex acts as a barrier to meridional transport relies heavily on results relating to Hamiltonian dynamical systems. In particular, the Kolmogorov–Arnold–Moser (KAM) theory (see, e.g., Arnold et al. 1986) plays a central role in our work. Details will be presented below, but a synopsis of our argument can be given now. According to the each of many variants of the KAM theorem, if a steady streamfunction is subjected to certain classes of time-dependent perturbations, some nonchaotic trajectories (which lie on tori in a higher-dimensional phase space) survive in the perturbed system. We argue that, under most conditions, the invariant tori that are most likely to survive in the perturbed system are those in close proximity to the core of the zonal jet, and that these provide a barrier to meridional transport. The relationship between our explanation of the mechanism leading to the transport barrier and the “potential vorticity barrier” explanation (Juckes and McIntyre 1987; McIntyre 1989) will be discussed in the final section of the paper.

Many authors (Bowman 1993; Chen 1994; Delshams and de la Llave 2000; Haynes and Shuckburgh 2000; Joseph and Legras 2002; Koh and Legras 2002; Koh and Plumb 2000; Mizuta and Yoden 2001, 2002; Ngan and Shepherd 1997, 1999a,b; Pierce and Fairlie 1993; Pierrerehumbert 1991; Yang 1998) have discussed the relevance of and/or have explicitly applied dynamical system tools to the study of the dynamics of the meridional transport barrier at the perimeter of the polar vortex. Indeed, some of these authors have suggested that the transport barrier at the perimeter of the polar vortex may be composed of KAM invariant tori. In this paper we consider several aspects of the connection between KAM invariant tori and the polar vortex transport barrier that have not previously been discussed: 1) we point out that at least one recent variant of the KAM theorem applies to those tori in the vicinity of the jet core at the perimeter of the polar vortex; 2) we explain why the tori most resistant to breaking are those in close proximity to the jet core; 3) we discuss the applicability of KAM theory to multiperiodic systems; and 4) we explain why, in multiperiodic systems, KAM invariant tori constitute true transport barriers. Thus, while the idea that we are pursuing is not entirely new, the results presented serve to significantly narrow the gap between theory and observation.

There is also a close connection between our work and that of Bowman (1996), in spite of the fact that the arguments presented in that paper are unrelated to dynamical systems. In both Bowman’s (1996) observationally based study and our more idealized study, the streamfunction is decomposed into a steady background on which a sum of traveling Rossby waves is superimposed. Our model is loosely motivated using dynamical arguments and chosen, in part, because rigorous mathematical results are available for streamfunctions of this general form. Using entirely different arguments than those given by Bowman (1996), we provide an explanation for his observation that, for a moderate strength perturbation, the transport barrier in the proximity of the jet core is expected to break down when one of the Rossby waves included in the perturbation has a phase speed close to that of the wind speed at the core of the jet.

The remainder of this paper is organized as follows. In the next section, a simple analytic form of the streamfunction is derived. This consists of a steady background flow—a zonal jet—on which two traveling Rossby waves are superimposed. In a reference frame moving at the phase speed of one of the Rossby waves, the flow consists of a steady background flow on which a time-periodic perturbation is superimposed. In section 3, the Lagrangian motion in such a model is discussed with the aid of two variants of the KAM theorem. We explain why, under typical conditions, particle trajectories near the core of the zonal jet in the perturbed system lie on KAM invariant tori, which provide a barrier to meridional transport. In section 4, we consider a more general model of the streamfunction, consisting of a zonal jet on which three or more traveling Rossby waves are superimposed. The Lagrangian motion is discussed with the aid of yet another variant of the KAM theorem. It is argued that the conclusions of section 3 are unchanged for a more general multiperiodic perturbation. In section 5 we summarize and discuss our results. Our KAM theorem–based explanation of the meridional transport barrier is contrasted to the
potential vorticity barrier explanation, and suggestions for future work are presented.

2. A simple, dynamically motivated model of the streamfunction

Our study focuses on elucidating the mechanism by which the zonal jet at the edge of the stratospheric polar vortex serves as a barrier to the meridional transport. Because of our focus on the zonal jet, it is natural to make use of a $\beta$-plane approximation with $\beta = (2\Omega / r_0) \cos \varphi_o$ defined at the latitude $\varphi_o$ of the core of the zonal jet. Here $\Omega = 2\pi/(1\text{day})$ is the angular frequency of the earth and $r_0 = 6371 \text{ km}$ is the earth’s radius. We shall assume that $\varphi_o = 60^\circ$ so $\beta = 1.14 \times 10^{-11} \text{s}^{-1} \text{m}^{-1}$. Also, our interest is in Lagrangian motion over time scales of a few months or less. This is sufficiently short that the diabatic processes can be neglected. The assumption of flow on an isentropic surface, together with incompressibility, allows the introduction of the streamfunction, $\psi(x, y, t)$, $u = -\partial \psi / \partial y$, $v = \partial \psi / \partial x$, with $x$ increasing to the east from an arbitrarily chosen longitude and $y$ increasing to the north from $\varphi_o$. The Lagrangian equations of motion are then

$$\frac{dx}{dt} = -\frac{\partial \psi}{\partial y}, \quad \frac{dy}{dt} = \frac{\partial \psi}{\partial x}.$$  

(1)

It is well known that these equations have Hamiltonian form, $H(p, q, t) \leftrightarrow \psi(x, y, t)$. This connection is exploited extensively in sections 3 and 4.

Consistent with our focus on zonal jets we shall assume that

$$\psi(x, y, t) = \psi_0(y) + \psi_1(x, y, t),$$  

(2)

where $\psi_0(y) = -d\psi_0/dy$ has a single extremum—a maximum—at $y = 0$. We now outline the steps of a derivation of a particular choice of $\psi_0(y)$ and $\psi_1(x, y, t)$. The same streamfunction has been previously used by del-Castillo-Negrete and Morrison (1993), and Kovályov (2000). Our presentation follows that of del-Castillo-Negrete and Morrison; more details can be found in that work. The simple analytical expressions for $\psi_0(y)$ and $\psi_1(x, y, t)$ that are presented below [see Eq. (12)] are far too simple to mimic the complexity of realistic stratospheric flows. But our model of the streamfunction is dynamically motivated and has approximately the correct length and time scales. This model streamfunction is used to produce numerical simulations to illustrate some important qualitative features of more realistic flows. In spite of its simplicity, our analytic model of the streamfunction includes all of the essential qualitative features of the stratospheric polar vortex that are needed to understand why it acts as a meridional transport barrier.

Consistent with our assumption of 2D incompressible flow on a $\beta$ plane, conservation of potential vorticity $Q = \nabla^2 \psi + \beta y$ dictates that

$$\frac{\partial Q}{\partial t} - \frac{\partial \psi}{\partial y} \frac{\partial Q}{\partial y} + \frac{\partial \psi}{\partial x} \frac{\partial Q}{\partial y} = 0.$$  

(3)

Substitution of (2) into (3) yields, after linearization (treating $\psi_1$ as a small perturbation to $\psi_0$),

$$\frac{\partial}{\partial t} \nabla^2 \psi_1 + u_0(y) \frac{\partial}{\partial x} \nabla^2 \psi_1 + [\beta - u_0''(y)] \frac{\partial \psi_1}{\partial x} = 0.$$  

(4)

Throughout this paper primes are used to denote differentiation with respect to a single independent variable: $u_0''(y) = du_0/dy$, etc. The assumption that $\psi_1$ has the form of a zonally propagating wave $\psi_1 = \phi(y) \exp[i(k(x - c\tau))]$ (or a superposition of such waves) yields the Rayleigh–Kuo equation:

$$[u_0(y) - c[\phi''(y) - k^2 \phi(y)] + [\beta - u_0''(y)] \phi(y) = 0.$$  

(5)

Problems associated with critical layers, where $u_0(y) = c$, and stability considerations lead to difficulties finding physically relevant solutions to this equation. We consider here the Bickley jet velocity profile:

$$u_0(y) = U_0 \text{sech}^2 \left( \frac{y}{L} \right).$$  

(6)

where $U_0$ and $L$ are constants. It was first shown by Lipp (1962) that for this velocity profile the Rayleigh–Kuo equation (5) admits two symmetric neutrally stable (I$m = 0$) solutions,

$$\phi_i(y) = A_i U_0 L \text{sech}^2 \left( \frac{y}{L} \right),$$  

(7)

where the $A_i (i = 1, 2)$ are dimensionless amplitudes. It is straightforward to verify that (6) and (7) constitute a solution to (5) provided

$$U_0 L^2 k^2 = 6c$$  

(8)

and

$$6c^2 - 4U_0 c + \beta U_0 L^2 = 0.$$  

(9)

The condition for the existence of two neutrally stable waves is

$$\beta L^2 / U_0 < 2/3.$$  

(10)

When this inequality is satisfied, (9) has two real roots; the corresponding wavenumbers are given by (8). It should be noted that for these solutions to (5), $[\beta - u_0''(y)]/|u_0(y) - c|$ is bounded at the critical layers; this
is a necessary condition for the existence of neutrally stable solutions (Kuo 1949).

The environment is defined by the parameters \( \beta, U_0, \)
and \( L, \) which, via (8) and (9), fix \( c_1, k_1, c_2, \) and \( k_2 \)
provided (10) is satisfied. But, because of the periodic boundary conditions in \( x, \) only a discrete set of \( k \) are allowed. At \( \varphi_0 = 60^\circ \) (this choice fixes \( \beta \) as noted above) these are

\[
k_n = \frac{2n}{r_e}, \quad n = 1, 2, \ldots
\]

For most choices of \( U_0 \) and \( L, \) (11) conflicts with (8) and (9). This issue was discussed by del-Castillo-Negrete and Morrison (1993) who argued that initial disturbances for which (8) and (9) are inconsistent with (11) should relax, via a barotropic-instability-induced decrease in \( U_0 \) and increase in \( L, \) to a state for which (8), (9), and (11) are self consistent. We avoid this issue by choosing \( U_0 \) and \( L, \) which correspond to such a self-consistent state. Specifically, we have chosen \( U_0 = 62.66 \text{ m s}^{-1}, \) \( L = 1770 \) km, corresponding to zonal wavenumbers \( n = 2 \) and \( n = 3. \) These waves have eastward-propagating phase speeds, \( c_2/U_0 = 0.205 \) and \( c_3/U_0 = 0.461. \) The streamfunction is then

\[
\psi(x, y, t) = -U_0L \tanh\left(\frac{y}{L}\right)
+ A_3U_0L \sech^2\left(\frac{y}{L}\right) \cos[k_3(x - c_3t)]
+ A_2U_0L \sech^2\left(\frac{y}{L}\right) \cos[k_2(x - c_2t)].
\]

An important observation is that the time dependence associated with one of the two Rossby waves in (12) can be eliminated by viewing the flow in a reference frame moving at the phase speed of that wave. The choice of which wave to absorb into the background flow is arbitrary. In the reference frame moving at speed \( c_3 \) the streamfunction is

\[
\psi(x, y, t) = c_3y - U_0L \tanh\left(\frac{y}{L}\right)
+ A_3U_0L \sech^2\left(\frac{y}{L}\right) \cos(k_3x)
+ A_2U_0L \sech^2\left(\frac{y}{L}\right) \cos(k_2x - \sigma_2t),
\]

where \( \sigma_2 = c_2k_2 - c_3k_2 = k_2(c_2 - c_3). \) Note that \( \sigma_2 \) is negative because in the reference frame moving at the faster \( n = 3 \) wave, the \( n = 2 \) wave has westward-propagating phases.

### 3. A steady background flow subject to a periodic perturbation

In this section we consider streamfunctions of the general form:

\[
\psi(x, y, t) = \psi_0(x, y) + \psi_1(x, y, \sigma t), \tag{14}
\]

where \( \psi_1 \) is a periodic function of \( t \) with period \( 2\pi/\sigma. \) Equation (13) is a special case of Eq. (14). All of the concepts discussed in this section apply to the more general class of (14). The particular form in (13) is used for numerical simulations to illustrate the relevant important concepts. In the section that follows we consider a slightly larger and more geophysically relevant class of streamfunctions corresponding to a multiperiodic perturbation \( \psi_1. \) It will be seen that almost all of the results presented in this section generalize in a straightforward fashion to multiperiodically perturbed systems.

We begin with a discussion of the importance of the background steady contribution to the streamfunction \( \psi_0(x, y) \) in (14). If, as we have assumed, in the rest frame the streamfunction has the form of a zonal flow on which a sum of zonally propagating Rossby waves are superimposed [as in (12)], the problem of lack of uniqueness of \( \psi_0(x, y) \) arises immediately; the choice of which traveling wave to absorb into the background is arbitrary. Because the purely zonal rest frame contribution to \( \psi_0(x, y) [-U_0L \tanh(y/L) \text{ in (13)}] \) is always present and is generally larger than whichever traveling wave contribution is absorbed into \( \psi_0, \) we first discuss the special case \( \psi = \psi_0(y). \)

The Hamiltonian form of the Lagrangian equations of motion in (1) was noted earlier. The special case \( \psi = \psi_0(y) \) corresponds, trivially, to the so-called action-angle representation of the motion in which \((p, q) \to (I, \theta), H(p, q) \to \mathcal{H}(I). \) The equations of motion in terms of action-angle variables \((I, \theta) \) are \( dI/dt = -\partial \mathcal{H}/\partial \theta = 0, \) \( d\theta/dt = \partial \mathcal{H}/\partial I = \omega(I); \) these equations can be trivially integrated. Note that \( I \) and \( \theta \) are defined in such a way that the motion is \( 2\pi \) periodic in \( \theta \) with angular frequency \( \omega(I). \) When \( \psi = \psi_0(y), \) we may take \( I = -yR, \) \( \theta = x/R \) and \( \mathcal{H} = \psi_0, \) where \( R = r_e \cos \varphi. \) With these simple substitutions, the original Lagrangian equations of motion in (1) have the action-angle form. For systems of this type, \( \omega(I) \) is simply a relabelling of \( u(y) \) and \( T(I) = 2\pi/\omega(I) \) is the time required for a trajectory to circle the earth. Action-angle variables and, in particular, the quantity \( d\omega(I)/dI \) play a crucial role in much of the discussion that follows.

The choice \( \psi_0(y) = -U_0L \tanh(y/L) \) corresponds to \( \mathcal{H}(I) = U_0L \tanh[I/(RL)], \) \( \omega(I) = \partial \mathcal{H}/\partial I = (U_0/R) \sech^2[I/(RL)]. \) Figure 1a shows the corresponding
streamlines in the \((x, y)\) plane (the trivial \(x\) dependence is included for comparison to Fig. 3), and plots of \(u_0(y) = -\partial \psi_0/\partial y = U_0 \tanh(y/L), \omega(I), \) and \(\omega'(I)\). Note that at the jet core, \(T(I)\) has a local minimum, \(\omega(I)\) has a local maximum, and \(\omega'(I) = 0\).

Now consider a superposition of the background zonal jet \(\psi_0(y)\) and one of the two Rossby wave perturbations included in (12). In the reference frame moving at the phase speed of the Rossby wave, the flow is steady. The corresponding streamfunction is

\[
\psi_0(x, y) = cy - U_0 L \tanh \left( \frac{y}{L} \right) + AU_0 L \sech \left( \frac{y}{L} \right) \cos(kx),
\]

where the phase speed \(c\), wavenumber \(k\), and the dimensionless wave amplitude \(A\) are written without subscripts. Nondimensionalization \([\psi \rightarrow \psi/(U_0 L), x \rightarrow kx, y \rightarrow y/L]\) reveals that there are two irreducible dimensionless parameters: \(A\) and \(c/U_0\). A bifurcation diagram in \((A, c/U_0)\) for this system is shown in Fig. 2. This figure shows that there are three regions, corresponding to topologically distinct streamfunction structures, and two critical curves that separate these regions. Level surfaces of \(\psi\) in each of the three regions and on the two critical lines are shown in the figure. Holding \(A \neq 0\) constant while \(c/U_0\) is increased reveals all possible streamfunction topologies. For small \(c/U_0\) the streamfunction is characterized by hyperbolic heteroclinic chains both above and below a spatially periodic eastward jet near \(y = 0\). As \(c/U_0\) is increased, a critical value is encountered at which the two hyperbolic heteroclinic chains merge and the eastward jet disappears. A further increase of \(c/U_0\) leads to the formation of homoclinic hyperbolic chains above and below a westward jet. As a second critical value of \(c/U_0\) is passed, the hyperbolic homoclinic chains are destroyed via saddle-node annihilation. For large \(c/U_0\) the flow is everywhere westward without stagnation points. Similar behavior was noted previously by del-Castillo-Negrete and Morrison (1993) using essentially the same model.

The importance of Fig. 2 is that it shows that depending on the choice of \(A\) and \(c/U_0\), the zonal jet may be strong, weak, or absent entirely. If \(A\) and \(c/U_0\) correspond to a pair that lies on the critical line at which the two hyperbolic heteroclinic chains merge, the eastward jet disappears and the chain of unstable and stable manifolds near \(y = 0\) is unstable to an arbitrarily small time-dependent perturbation. For the stratospheric polar vortex the relevant (usually) domain of the \((A, c/U_0)\) parameter space is small values of both parameters (see, e.g., Bowman 1996). It should be emphasized, however, that when more than one Rossby wave is superimposed on the background zonal jet, as in (13), the choice of which Rossby wave to absorb into the background is arbitrary. Under such conditions the claim that the background flow topology corresponds to the small \(c/U_0\) region in Fig. 2 is justified only if this is true for \((A_i, c_i/U_0)\) for all of the waves present. Although the preceding discussion was motivated by a
particular model streamfunction, Eq. (15), the qualitative features that we have described are expected to be broadly applicable to Rossby wave perturbations to eastward zonal jets.

For any steady streamfunction $\psi = \psi(x, y)$ the Lagrangian equations of motion (1) can be transformed to action-angle form. The equations of motion in action-angle form are identical to the equations described above, but the transformation from $\psi(x, y)$ to $\overline{H}(I)$ is more complicated than the trivial relabeling of coordinates described above. More generally, $I(\overline{H}) = 2\pi^{-1} \oint x(y, \overline{H}) \, dy = -2\pi^{-1} \oint y(x, \overline{H}) \, dx$ where $\overline{H} = \psi$ and the integral is around a closed loop in $(x, y)$, and $\theta = \partial G / \partial I$ with $G(y, I) = \int_x x(y', \overline{H}) \, dy'$. The equivalence of the two forms of $I(\overline{H})$ given above follows from integration by parts. Note also that on a given level surface of $\psi$, $x(y)$ or $y(x)$ may be multivalued, dictating that some care be exercised when using these equations, and that there is flexibility in choosing the lower limit in the integral defining $G$. It is often necessary to define action-angle variables in different regions of $(x, y)$ in a piecewise fashion. We emphasize, however, that once this is done the form of the equations of motion in action-angle variables is that given above. It is important to keep in mind that $I$ is simply a label for a particular trajectory or, equivalently, for a particular level surface of $\psi(x, y)$.

For $A = 0.3$, $c/U_0 = 0.461$, plots of $|u| = (u^2 + v^2)^{1/2}$, $\omega(I)$, $T(I)$, and $\omega'(I)$ are shown in Fig. 3 for trajectories in the vicinity of the jet core only. Note that, in qualitative agreement with Fig. 1, Fig. 3 shows that in the vicinity of the jet core, $T(I)$ has a local minimum, $\omega(I)$ has a local maximum, and $\omega'(I) = 0$. These features play an important role in the considerations that follow.

We now turn our attention to periodically perturbed systems of the form (14), of which (13) is a special case. With $x$ and $y$ bounded, trajectories lie in a three-dimensional bounded phase space $(x, y, t \mod 2\pi/\sigma)$. The usual way to view trajectories in such a system is to construct a Poincaré section, which is a slice of the 3D space corresponding to $t \mod 2\pi/\sigma = \text{const}$. Three examples, corresponding to the system described by (13) with three choices of the perturbation strength $A_2$, are shown in Fig. 4. On these plots regular (nonchaotic) trajectories appear as discretely sampled smooth curves, while chaotic trajectories appear as sets of discrete samples that fill areas. In the $A_2 = 0$ limit all trajectories are nonchaotic; each curve seen in Fig. 4a can be thought of as a 2D slice of a torus in $(x, y, t \mod 2\pi/\sigma)$. For small perturbation strength $A_2$ some of the unperturbed tori are seen to survive, while other tori break up forming chains of island-like structures that are surrounded by chaotic seas. In general, as the perturbation strength increases more tori are destroyed and the motion becomes increasingly chaotic.

An additional Poincaré section is shown in Fig. 5. The parameters used to construct that figure are identical to those used to construct the middle panel in Fig. 4 except that in Fig. 5 $U_0$ is decreased to 41.31 m s$^{-1}$ so that $c_3/U_0$ is increased to 0.7. (Note that with this change the flow is no longer a dynamically self-consistent Bickley jet.) The transition from the middle panel in Fig. 4 to Fig. 5 shows, in a very idealized flow, qualitatively what happens during the spring warming of the polar vortex, which is associated with a decrease in $U_0$, and illustrates the importance of the background flow. The transition from Fig. 4 to Fig. 5 corresponds to a shift to the right in the bifurcation diagram shown in Fig. 2. Figure 4 corresponds to conditions in the region on the left in the bifurcation diagram, characterized by the presence of an eastward central jet. Figure 5 corresponds to conditions in the central region of the bifurcation diagram, but very close to the bifurcation curve on the left. The loss of stability of the central region in Fig. 5 is caused by the change in the background flow topology seen in Fig. 2, and in particular, the lack of stability of motion in the vicinity of hyperbolic chains under perturbation. The transition from Fig. 4 (middle panel) to Fig. 5 is consistent with the work of Bowman.
(1996), discussed above, focusing on the breakdown of the transport barrier at the perimeter of the polar vortex when \( c/U_0 \) exceeds a threshold that is close to unity.

As a preliminary to our discussion of KAM theory it is instructive to make some seemingly trivial comments about the geometry of systems of the form in (14) and the Poincaré sections shown in Figs. 4 and 5. The tori of the unperturbed system can be thought of as either 1D surfaces in \((x, y)\) or 2D surfaces in the 3D space \((x, y, t \mod 2\pi/H)\). Because the unperturbed streamfunction does not depend on \(t\) the latter view seems like an unnecessary complication, but it turns out to be very useful. The regular trajectories in the perturbed systems shown in Fig. 4 lie on the tori (of the second type) of the unperturbed system that survive under perturbation. Because each surviving torus is a 2D surface in the 3D space \((x, y, t \mod 2\pi/H)\), each such torus divides the 3D space into disjoint “inside” and “outside” regions.

A consequence of this is that in the \((x,y)\) plane each surviving torus represents an impenetrable barrier to transport.

The survival of some of the tori of the unperturbed system under small perturbation, as illustrated in Fig. 4, is predicted by the KAM theorem (see, e.g., Arnold et al. 1986; Poschel 2001). Before giving a more precise statement of the theorem, it is instructive to note that the mechanism that leads to the destruction of the tori of the unperturbed system is the excitation of resonances by the time-periodic perturbation. Resonances are excited when the ratio of the frequency of the perturbation \(\omega\) to the frequency of the motion on the unperturbed torus \(\omega(I)\) is the ratio of integers. Generically, a continuum of \(\omega(I)\) values are present. Under such conditions a fixed \(\sigma\) excites infinitely many resonances. In practice, however, the low-order resonances (e.g., 2:1) are the most important.

With this as a heuristic background, a form of the KAM theorem suitable for systems of the form (14) can now be stated. According to the theorem, tori in the vicinity of those tori in the unperturbed system for which \(\omega(I)/\sigma\) is sufficiently irrational, survive in the perturbed system provided that the strength of the perturbation is sufficiently weak and a nondegeneracy condition, \(\det(\partial\omega/\partial I) \neq 0\), is satisfied. The condition that \(\omega(I)/\sigma\) is sufficiently irrational is expressed quantitatively by a Diophantine condition; this will not be discussed further as this condition is not central to our arguments.

On the other hand, the nondegeneracy condition plays a critical role in our arguments. In the simplest form of the theorem this condition is \(\omega'(I) \neq 0\) [or in higher dimensions \(\det(\partial\omega/\partial I) \neq 0\)]. This condition guarantees the invertibility of \(\omega(I)\), whose importance stems in part from the fact that the theorem guarantees that the torus corresponding to that value of \(I\) for which

Fig. 4. Poincaré sections corresponding to the system described by Eqs. (1) and (13) with \(c_3/U_0 = 0.461\) and \(A_3 = 0.3\) for three values of \(A_2\): (top) 0, (middle) 0.1, and (bottom) 0.7. Note the robustness of the tori in the vicinity of the jet core.

Fig. 5. Poincaré section corresponding to the system described by Eqs. (1) and (13) with \(c_3/U_0 = 0.70, A_3 = 0.3\) and \(A_2 = 0.1\). Note that the central barrier seen in the middle plot in Fig. 4 has been lost as a result of an increase in \(c_3/U_0\).
\( \omega(I)/\sigma \) is sufficiently irrational survives in the perturbed system. This was the form of the nondegeneracy condition in Kolmogorov’s (1954) original statement of the theorem. Already in his original proof of the theorem, Arnold (1963) noted (in a footnote) that an alternate form of the nondegeneracy condition, the isoenergetic condition, could be used instead. Subsequently, Bruno (1992) and Russmann (1989) announced forms of the theorem that employ less restrictive nondegeneracy conditions.

The Russmann nondegeneracy condition is of particular interest in the present study. The condition is most naturally stated in words: for an autonomous system with \( N + 1 \) degrees of freedom the image of the frequency map \( I \rightarrow \omega(I) \) may not lie on any hyperplane of dimension \( N \) that passes through the origin. To our knowledge, all published formulations of the KAM theorem to date that make use of the Russmann nondegeneracy condition apply to autonomous systems. To apply such a result to (14) this system must first be written as an equivalent autonomous two degree-of-freedom system with a bounded phase space. The required transformation is a special case of the transformation described at the end of the following section. After performing this transformation, the Russmann condition reduces to a statement that in the 2D \((\omega, \sigma)\) space, the locus of points \((\omega(I), \sigma)\) must not fall on a line that passes through the origin. This condition is violated only when \( \omega(I) = \text{const} \). For our purposes, the significance of the Russmann nondegeneracy condition is that, for systems of the form in (14), it is satisfied in a domain that includes an isolated zero of \( \omega'(I) \). The price paid for making use of the relatively weak Russmann nondegeneracy condition is that a slightly less strong form of the theorem is proved. The Kolmogorov condition \( \omega'(I) \neq 0 \) can be used to prove that the torus corresponding to a particular value of \( I \) in the unperturbed system survives with unchanged frequency in the perturbed system. In contrast, when the Russmann form of the theorem is applied in a domain that includes \( I = I_0 \), where \( \omega'(I_0) = 0 \) (an isolated zero), the theorem only guarantees the existence of tori in the perturbed system whose frequencies are close to those in the unperturbed system. Thus, when the Russmann form of the theorem is applied, it is not appropriate to refer to the survival of tori corresponding to a particular value of \( I \) (see, e.g., Sevryuk 1995, 2006, for a more complete discussion of this issue). For our purposes this distinction is unimportant in that it makes no difference whether the \( I = I_0 \) torus survives under perturbation; it is important to know only that some nearby tori are present in the perturbed system as any such torus provides a barrier to transport.

Because of its generality, we have chosen to emphasize the importance of the Russmann nondegeneracy condition in our discussion of the stability of tori in the vicinity of that for which \( \omega'(I) = 0 \). It is worth noting, however, that other arguments have been used to establish essentially the same result for area-preserving mappings (Delsmans and de la Llave 2000; Simó 1998).

Consider again the Poincaré sections shown in Fig. 4. This figure shows that not only do the tori corresponding to trajectories near the jet core [where \( \omega'(I) = 0 \)] persist in the perturbed system, but these tori appear to be the most resistant to breaking. Numerical simulations based on model systems reveal that, in general, tori near that for which \( \omega'(I) = 0 \) tend to be the most resistant to breaking. Interestingly, for the model parameters used to construct Fig. 4, \( \sigma/\omega \) at the jet core is approximately 0.95. The significance of this ratio is its closeness to unity. On two nearby tori, one on each side of the jet core, the strongest possible \((1:1)\) resonance is excited. In spite of this, tori in the vicinity of the jet core are seen to be preserved for moderate perturbation strengths. The reason for the surprising stability of tori near that for which \( \omega'(I) = 0 \) will be described below. We emphasize, however, that the stability of tori near that for which \( \omega'(I) = 0 \) is not absolute. If, for instance, \( \sigma \) happens to be identical to \( \omega(I) \) at the jet core, where \( \omega'(I) = 0 \), thereby exciting a \((1:1)\) resonance on the jet core, tori near the jet core are not among the last to break up as the perturbation strength increases.

The fact that tori for which \( \omega'(I) = 0 \) are strongly resistant to breaking has been noted in the mathematical literature; Gaidashev and Koch (2004) refer to the “remarkable stability” of such tori. Systems which satisfy this condition are generally described as “shearless” or “nontwist” in the mathematical literature, and have been extensively studied in recent years (see, e.g., del-Castillo-Negrete and Morrison 1993; Dullin and Meiss 2003; Morozov 2002).

We turn our attention now to resonance widths as a means to explain the remarkable stability of the tori satisfying the nontwist condition. Resonance widths are important because when neighboring resonances overlap, the intervening tori generally break up; the widely used Chirikov definition of chaos is based on overlapping resonances (see, e.g., Chirikov 1979; Chirikov and Zaslavsky 1972; Lichtenberg and Lieberman 1983). Recall that resonances are excited on the tori for which \( \omega(I)/\sigma \) is rational. Resonance widths are controlled by the degree of rationality of \( \omega/\sigma \), the perturbation strength and a simple geometric factor, which we now consider. A simple analysis (see, e.g., Abdullaev 1993) reveals that resonance widths scale like \( \Delta I \sim (\omega(I))^{-1/2} \) or \( \Delta \omega \sim (\omega'(I))^{1/2} \). Because resonances are excited at
discrete values of $\omega$, it is the width $\Delta \omega$, rather than $\Delta t$, that is important in determining whether neighboring resonances overlap. Because $\Delta \omega \sim |\omega'(I)|^{1/2}$ small values of $\omega'(I)$ are generally associated with small resonance widths, and generally more surviving KAM tori. [The resonance width estimates just quoted follow from a simple perturbation analysis. When $\omega'(I) = 0$ at the resonance, the width of the resonance $\Delta \omega$ depends on $\omega'(I)$ at the resonance. The exact form of this expression is not essential to our argument. What is important is the observation that $\Delta \omega$ is small when $\omega'$ is small.]

In the vicinity of the jet core a narrow band of $\omega$ values will be present. Resonances will be excited in this band, but only for very special values of $\sigma$ will these be low-order resonances. The associated widths of these resonances are small owing to the smallness of $|\omega'(I)|^{1/2}$ in this region. As a result, mostly nonchaotic motion is preserved near the jet core, not because resonances are not excited, but because the corresponding resonance widths are usually so small that neighboring resonances do not overlap. Excitation of a low-order resonance very close to the jet core can overcome the smallness of $|\omega'(I)|^{1/2}$ and change this picture, so the stability of the tori near the jet core is not absolute.

In this section we have considered a steady zonal jet on which two traveling Rossby waves are superimposed. Either of the two Rossby waves can be absorbed into a modified steady background flow. We have shown, using well-known results relating to KAM theory, that, provided certain conditions are met, a typically narrow band of nonchaotic trajectories in the vicinity of the jet core, each lying on a KAM invariant torus, persists in the two-wave system and provides a barrier to meridional transport. The barrier is linked to the remarkable stability of KAM tori for which $\omega'(I)$ has an isolated zero. The conditions that need to be met for such a barrier to be present are 1) the rest frame phase speeds of both Rossby waves should not be comparable to the wind speed at the jet core; 2) the Rossby wave amplitudes must not be too large; and 3) low-order resonances in the immediate vicinity of the jet core in the moving frame must not be excited.

4. A steady background flow subject to a multiperiodic perturbation

In this section we consider streamfunctions of the form

$$\psi(x, y, t) = \psi_0(x, y) + \psi_1(x, y, \sigma_1 t, \ldots, \sigma_N t),$$

(16)

where $\psi_1$ is a multiperiodic function with constituent periods $2\pi/\sigma_i$, $i = 1, 2, \ldots, N$. It should be noted that a steady zonal flow on which a sum of $N + 1$ zonally propagating Rossby waves is superimposed has the form (16) when viewed in the reference frame moving at the phase speed of one of the Rossby waves. The $N = 1$ problem treated in the previous section is seen to be a special case of the problem treated here. In this section we show that most of the results discussed in the previous section carry over to the larger and more realistic class of problems considered here with only minor modification.

An important observation relating to systems of the form (16) is that one need only consider frequencies that are incommensurable (i.e., have the property that the ratio of all pairs of frequencies is irrational). Consider, for example, a multiperiodic function with periods of 4 and 6 days. This function is a simple periodic function with a period of 12 days. In general, a reduction of the number of frequencies can be achieved whenever two or more of the frequencies are commensurable. Thus, without loss of generality, it may be assumed that $\sigma_1, \sigma_2, \ldots, \sigma_N$ are incommensurable (i.e., $\psi_1$ is a quasiperiodic with $N$ incommensurable frequencies).

Systems of the form (16) have been intensively studied in recent years. A proof of the KAM theorem for such systems has been provided by Jorba and Simó (1996). Several points relating to the Jorba–Simo work are noteworthy. First, the theorem is formulated as a nonautonomous perturbation to an autonomous one degree of freedom system, so the unperturbed Hamiltonian, which must satisfy a nondegeneracy condition, is the system defined by $\psi_0(x, y)$ (after transforming to the action-angle representation). Second, the nondegeneracy condition that the unperturbed Hamiltonian is assumed to satisfy is the Kolmogorov condition $\omega'(I) \neq 0$. Third, Diophantine conditions must be satisfied by both $\sigma_i/\omega$ and $\sigma_i/\sigma_j$ ($i \neq j$). The second point is of particular importance in the present study. Loosely speaking, the Jorba–Simo work shows that the principal difference between the periodic perturbation problem and the quasiperiodic perturbation problem is that in the former problem the surviving KAM tori undergo periodic oscillations in $(x, y)$, while in the latter problem the surviving KAM tori undergo quasiperiodic oscillations in $(x, y)$. Jorba and Simo refer to the latter motion as a “quasiperiodic dance.” For our purposes, this distinction is unimportant; in both cases the surviving KAM tori provide a barrier in $(x, y)$ to transport, as we shall describe in more detail later in this section.

All of the mathematical difficulties associated with a quasiperiodic perturbation $\psi_1$ are present even for $N = 2$. Because, among all $N \geq 2$, the $N = 2$ case is the most convenient choice for numerical purposes, it is natural to focus on that choice. With this in mind we have
chosen, for numerical purposes, to use the streamfunction

$$\psi(x, y, t) = c_3 y - U_0 L \tanh\left(\frac{y}{L}\right)$$

$$+ A_3 U_0 L \text{sech}^2\left(\frac{y}{2L}\right) \cos(k_3 x)$$

$$+ A_2 U_0 L \text{sech}^2\left(\frac{y}{2L}\right) \cos(k_2 x - \sigma_3 t)$$

$$+ A_1 U_0 L \text{sech}^2\left(\frac{y}{2L}\right) \cos(k_1 x - \sigma_1 t). \quad (17)$$

In the $A_1 = 0$ limit this is identical to the streamfunction described by Eq. (13). Note that physically Eq. (17) represents a zonal flow corresponding to $\psi_0(y) = -U_0 L \tanh(y/L)$ on which three traveling Rossby-like waves are superimposed in a reference frame moving with speed $c_3$, the phase speed of the zonal wavenumber-3 wave. For convenience, we have assumed that the new perturbation term corresponds to zonal wavenumber 1, $k_1 = 2\pi/[2\pi r_c \cos(60^\circ)] = 2/r_c$, and has the same $\text{sech}^2(y/L)$ meridional structure as the $k_2$ and $k_3$ modes. However, unlike the $k_2$ and $k_3$ modes, which had some dynamical justification, the $k_1$ mode is simply an ad hoc additive perturbation that is included to illustrate some properties of quasiperiodic systems. With this in mind we have chosen $\sigma_1/\sigma_2$ to be the golden mean $(\sqrt{5} - 1)/2$ (whose continued fraction representation identifies it as the most irrational real number).

Numerical simulations based on the system defined by (17) are shown in Figs. 6 and 7. Figure 6 shows the time evolution of two sets of air parcels at times ranging from $t = 0$ to $t = 81$ days. The initial conditions are chosen to fall on two zonal lines $y = \text{constant}$ on opposite sides of the zonal jet. It is seen that after 81 days each side of the jet is well stirred, as indicated by what appears to be random distributions of dots on each side of the jet, but there is no transport across an undulating barrier near the core of the jet. The cause of this behavior is a thin band of KAM invariant tori near the jet core that survive in the perturbed system and form a meridional transport barrier. This thin band of KAM invariant tori that separate the polar from the midlatitude region in our idealized system undulates in a quasiperiodic fashion in time; this is the quasiperiodic
dance referred to by Jorba and Simo. Further support for this interpretation of Fig. 6 is provided by the results shown in Fig. 7. In that figure, for the same model system, finite-time Lyapunov exponents are shown as a function of the initial condition for a set of air parcel trajectories that spans the zonal jet. This figure shows that the region in the immediate vicinity of the jet core is characterized by small Lyapunov exponent estimates. This behavior is consistent with the interpretation that there is a narrow band of surviving tori (on which motion is nonchaotic) in this region.

The qualitative features of Figs. 6 and 7 are consistent with available observational evidence. Consistent with Fig. 6 are satellite-based measurements of ozone distributions in the austral spring; see, for example, Bowman and Mangus (1993) or the National Aeronautics and Space Administration (NASA) Total Ozone Mapping Spectrometer (TOMS) Web site (online at http://jwocky.gsfc.nasa.gov/eptoms/ep_v8.html). These observations consistently reveal a sharp (subject to the horizontal resolution limitations of these measurements and the fact that most measurements of this type are integral measurements through the depth of the atmosphere) boundary between low ozone concentration air inside the polar vortex and high ozone concentration air outside the polar vortex. Aircraft-based measurements (see, e.g., Tuck 1989; Waugh et al. 1994), whose resolution is higher than the satellite-based measurements, reveal sharper boundaries between these regions in addition to finer-scale structure, often including additional boundaries that are linked to chemical processes. Our Fig. 7, which indicates that the perimeter of the polar vortex is a narrow nonchaotic barrier that separates two predominantly chaotic regions, is consistent with Fig. 8 in Koh and Legras (2002), Fig. 2 in Pierce and Fairlie (1993), the analysis of material line growth by Bowman (1993), and the observation by Chen (1994) that imbedded in the narrow barrier between the inside and outside of the vortex is a potential vorticity contour that grows at a locally minimal rate. Also, Paparella et al. (1997) show, using analyzed winds, that the vortex edge appears to be reliably identified as a maximum of kinetic energy (i.e., the core of the meandering jet, which is readily identified in Figs. 6 and 7). This behavior is in good qualitative agreement with our arguments (recall our Fig. 3) in that in the background environment the trajectory for which $\omega'(I) = 0$ is close to that for which the kinetic energy is maximum.

Figures 6 and 7 suggest that the most robust of the tori of the original system are those in the vicinity of the core of the jet where $\omega'(I) = 0$. This observation is not surprising as it is consistent with the discussion in the previous section relating to resonance widths. But the observation does serve to identify a weakness in our argument, however, inasmuch as the Jorba–Simo proof of the KAM theorem for quasiperiodic systems makes use of the simplest (Kolmogorov) nondegeneracy condition $\omega'(I) \neq 0$. Thus the Jorba–Simo form of the KAM theorem does not address the stability of tori near the jet core (i.e., those that are apparently the most stable). [One might argue that the theorem holds for tori that are arbitrarily close to that for which $\omega'(I) = 0$, but this is not entirely satisfactory in our view given our focus on the jet core.] What is needed to rigorously complete our argument is a proof of the KAM theorem for quasiperiodic systems (16) that makes use of a Russmann-like nondegeneracy condition rather than the Kolmogorov condition. So far as we are aware, such a proof has not been published to date.
Our numerical simulations, including but not limited to Figs. 6 and 7, strongly suggest that a form of the KAM theorem holds for quasiperiodic systems (16) for which the background \( \omega'(I) \) has an isolated zero.

We turn our attention now to justifying the claim, made above without proof, that for quasiperiodic systems (16) KAM tori provide a barrier to transport. Recall that for periodic systems (14) this property was established by noting that each KAM torus is a two-dimensional surface in the three-dimensional space \((x, y, t \bmod 2\pi/\sigma)\) that divides the 3D space into nonintersecting inside and outside regions. An extension of the same argument applies to the quasiperiodic problem. To see this, note first that the nonautonomous 1 degree-of-freedom system described by Eqs. (1) and (16) can be written as an equivalent autonomous \( N + 1 \) degree-of-freedom system:

\[
\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}, \quad i = 1, 2, \ldots, N + 1, \tag{18}
\]

where \( q_i = \sigma_i t, p_i = -\psi/\sigma_i, i = 1, 2, \ldots, N \), and \( q_{N+1} = y, p_{N+1} = x \), with

\[
H(p, q) = \psi(p_{N+1}, q_{N+1}; q_1, q_2, \ldots q_N) + \sum_{i=1}^{N} \sigma_i p_i. \tag{19}
\]

It is straightforward to verify that (18) and (19) reduce to \( dt/d\tau = 1 \), Eqs. (1), and \( d\psi/dt = \partial \psi/\partial t \). An important property of the transformed system (18) and (19) is that each trajectory is constrained by the presence of \( N \) integrals [sometimes called constants of the motion, i.e., \( N \) functions \( f_i(p, q), i = 1, 2, \ldots, N \) for which \( df_{i}/d\tau = 0 \)]. These integrals are \( H \) and \( f_j = q_j/\sigma_j - q_j/\sigma_N, i = 1, 2, \ldots, N - 1 \). If one additional independent integral can be found, the system (18) and (19) can be solved by quadratures and is said to be integrable. [This should come as no surprise because the original system (1) and (16) also lacks only one integral to render it integrable.]

For our purposes, the principal significance of the \( N \) integrals is that, because of their presence, each trajectory in the \( 2(N + 1) \)-dimensional phase space lies on a surface of dimension \( 2(N + 1) - N = N + 2 \). In a near-integrable system of this type in which both KAM tori and chaotic trajectories are present, the tori have dimension equal to the number of degrees of freedom, \( N + 1 \). In the \( (N + 2) \)-dimensional space that is filled by chaotic trajectories, the \( (N + 1) \)-dimensional KAM tori serve as impenetrable transport barriers. [The significance of these numbers is that the dimension of the KAM tori is one less than the dimension of the space that the chaotic trajectories fill. Note, for example, that in \( (x, y, z) \) the 1D circle \( x^2 + y^2 = 1, z = 0 \) divides the 2D \( z = 0 \) plane into nonintersecting inside and outside regions, but the same 1D circle does not divide the 3D \( (x, y, z) \) volume into nonintersecting inside and outside regions.]

The argument just given shows that in the system described by (18) and (19) Arnold diffusion does not occur. Loosely speaking, this is the process that allows chaotic trajectories to bypass KAM invariant tori. This process occurs in near-integrable autonomous systems with \( N \gtrsim 3 \) degrees of freedom which, under perturbation, are constrained by only one integral \( H \). For such systems phase space has dimension \( 2N \), chaotic trajectories lie on surfaces of dimension \( 2N - 1 \), while KAM invariant tori have dimension \( N \); for \( N \gtrsim 3 \) these tori do not serve as impenetrable barriers to the chaotic trajectories. The cause of the absence of Arnold diffusion in the system described by (18) and (19) is the integrals in addition to \( H \) that constrain the motion of all trajectories.

In this section we have argued that, with some minor modifications, the conclusions of the previous section carry over to a multiperiodic perturbation. Unlike the results of the previous section, however, the multiperiodic argument lacks mathematical rigor in that, to date, no proof of a KAM theorem for quasiperiodically perturbed Russmann-nondegenerate Hamiltonians has been published. Numerical simulations provide strong evidence that such a result holds. With this in mind, we state with some confidence that, provided the conditions stated at the end of section 3 are met, the qualitative features that were described in the previous section—the robust nature of nonchaotic trajectories near the jet core that serve to isolate chaotic trajectories on opposite sides of the jet—are expected to be seen whether there are 2 or 20 Rossby waves superimposed on the background zonal jet.

5. Summary and discussion

In this paper we have argued, using several nontrivial extensions of the basic KAM theorem, that, under commonly encountered conditions, the zonal jet at the periphery of the stratospheric polar vortex provides a robust barrier to the meridional transport of passive tracers. In the model employed, the perturbation to the background steady zonal jet was assumed to consist of a sum of traveling Rossby waves. The transport barrier comprises a typically narrow band of nonchaotic trajectories, each lying on a KAM invariant torus that is labeled by \( I \), which survive in the perturbed system. These tori tend to be the most resistant to break up under perturbation because they are close to the unperturbed streamline near the jet core for which \( \omega'(I) = 0 \) and because resonance widths \( \Delta \omega \) are approximately proportional to \( |\omega'(I)|^{1/2} \). The required extensions to the basic KAM theorem that we have made
use of to arrive at this conclusion are as follows: 1) applicability of the theorem to multiperiodic systems, including the absence of Arnold diffusion in such systems; and 2) applicability of the theorem when the usual (Kolmogorov) nondegeneracy condition \( \omega'(I) = 0 \) is violated. Our argument falls short of complete mathematical rigor because, to our knowledge, published proofs of the KAM theorem for quasiperiodically perturbed systems make use of the Kolmogorov nondegeneracy condition. (Note, however, that numerical simulations strongly suggest that the theorem is satisfied for Kolmogorov-degenerate streamfunctions subject to quasiperiodic perturbations, indicating that our argument is firmly, if not rigorously, grounded.) Also, it should be emphasized that even rigorous applicability of a form of the theorem to trajectories in the vicinity of the jet core will not guarantee that for all perturbations the corresponding tori will survive and provide a barrier to meridional transport. KAM invariant tori may not survive in the perturbed system for some combination of the following reasons: 1) the phase speed of one of the more energetic Rossby waves is close to the zonal velocity at the jet core; 2) the perturbation excites a low-order resonance on one of the tori in close proximity to that for which \( \omega'(I) = 0 \); or 3) the amplitude of the perturbation is too large. In spite of these caveats, our simulations suggest that, under conditions similar to those found in the austral winter and spring, the transport barrier near the core of the zonal jet at the perimeter of the polar vortex is very robust.

A consequence of a robust transport barrier of this type in the presence of mostly chaotic motion on each side of the barrier is that coarse-grained distributions of tracers, including the potential vorticity \( Q \), should be nearly homogeneous on each side of the barrier. Thus, one expects to see a near steplike \( Q \) structure across the transport barrier. Such a distribution is consistent with McIntyre’s (1989; see also Juckes and McIntyre 1987; Dritschel et al. 2007, and references contained therein) “potential vorticity barrier” explanation of the transport barrier. In the paragraphs that follow we give an expanded explanation of the potential vorticity barrier idea, and then argue that it is most appropriate to think of the zonal jet as the cause of the barrier and the \( Q \) jump as a consequence of the barrier. The cause versus effect arguments should not overshadow the important observation that the arguments that we have presented are consistent with the notion of a \( Q \) barrier.

Associated with the jump in \( Q \) across the transport barrier is a strong \( Q \) gradient that, in turn, is associated with a strong Rossby wave restoring force and therefore a tendency to resist large-scale deformation. This observation suggests that the \( Q \) barrier across the perimeter of the polar vortex is the cause of (or at least provides a positive feedback mechanism to help maintain) the associated meridional transport barrier. This argument was first given by Juckes and McIntyre (1987) who also noted that this mechanism must be augmented by some other mechanism, which they suggested is shear, to explain the impermeability of the barrier to small-scale intrusions. The most recent discussion of these ideas is that of Dritschel et al. (2007).

Coarse-grained homogenization of \( Q \) on each side of the transport barrier is dynamically consistent with a nearly zonal flow with \( u(y) = u(0) - \xi(0^+) y + \beta y^2/2 \) for \( y > 0 \) and \( u(y) = u(0) - \xi(0^-) y + \beta y^2/2 \) for \( y < 0 \) where the constants \( \xi(0^+) \) and \( \xi(0^-) \) are the values of the relative vorticity on opposite sides of the barrier. For this reason we have chosen to use a notation that focuses on the jump in relative vorticity across the barrier.) Assuming the width of the barrier is fixed, the \( Q \)-barrier/Rossby wave restoring force argument depends only on the magnitude of the \( Q \) jump \( |\xi(0^+) - \xi(0^-)| \) across the barrier. Thus, that argument suggests that meridional transport barriers should be found with all possible sign combinations of \( \xi(0^+) \), \( \xi(0^-) \), and \( u(0) \). But, so far as we are aware, full-physics (taken here to mean based on (3) or some extension thereof—e.g., to account for a finite deformation radius) numerical simulations (Danilov and Gryanik 2004; Danilov and Gurarie 2004; Dritschel et al. 2007; Haynes et al. 2007; Huang and Robinson 1998; Manfroi and Young 1999) reveal robust meridional \( Q \) barriers only when \( u(0) > 0 \), \( u'(0^-) = -\xi(0^-) > 0 \) and \( u'(0^+) = -\xi(0^+) < 0 \), and only when \( u(0) \) is sufficiently large, that is, only in the presence of a strong eastward jet. Stated somewhat differently, the potential vorticity inversion of a meridional \( Q \) step does not, in general, lead to a zonal jet, while numerical simulations suggest that, among the class of meridional \( Q \)-step flows, only the special case corresponding to an eastward zonal jet appears to be associated with a meridional transport barrier. It is also noteworthy that robust meridional transport barriers need not be associated with a jump in \( Q \), as illustrated by the background flow \( u(y) = u(0) + \beta y^2/2 \), \( u(0) < 0 \), corresponding to a westward zonal jet. Under perturbation this flow satisfies the kinematic conditions that we have identified as being associated with a robust meridional transport barrier near \( y = 0 \). Additionally, under perturbation this flow is dynamically consistent in a coarse-grained
sense with $Q = \text{constant}$ (the same value for positive and negative $y$). Note that to observe a transport barrier in this flow under perturbation, a tracer other than $Q$ would have to be used. The foregoing considerations suggest that it is most correct to think of the eastward zonal jet at the perimeter of the stratospheric polar vortex as being the cause of the barrier to meridional transport and the associated coarse-grained step in $Q$ as a consequence of the transport barrier. We say “most correct” because of the nonrigorous nature of the arguments presented and because, in the presence of a background eastward jet with an associated $Q$ step, Rossby wave radiation stresses (Dritschel et al. 2007) provide a positive feedback to help maintain the jet while the Rossby wave–restoring force mechanism acts to resist large-scale deformation. These feedbacks suggest that the distinction between cause and effect is unavoidably somewhat fuzzy.

The role of shear is less ambiguous. The quantity identified above as $\omega'(I)$ is a measure of shear. Unperturbed tori that satisfy the condition $\omega'(I) = 0$ are referred to as shearless or twistless tori. As noted above, nondegenerate resonance widths scale like $|\omega'(I)|^{1/2}$. Thus, as the magnitude of the shear $|\omega'(I)|$ increases, more resonances overlap and more tori—which would otherwise act as transport barriers—are broken. Shear is thus a destabilizing influence. This simple observation is consistent with the arguments given above and our assertion that the cause of the transport barrier is the eastward zonal jet at whose core is a shearless torus. Numerical simulations (Danilov and Gryanik 2004; Danilov and Gurarie 2004; Dritschel et al. 2007; Haynes et al. 2007; Huang and Robinson 1998; Manfroi and Young 1999) consistently show that transport barriers associated with eastward zonal jets are in close proximity to the shearless jet core. Indeed, Haynes et al. (2007) emphasizes the connection between the transport barrier and the shearless jet core. A caveat is that the kinematic arguments that we have presented do not address the dynamical stability of the flow. The extent to which our arguments can be reconciled with the shear sheltering mechanism of Hunt and Durbin (1999) is worthy of investigation.

Dynamical consistency between Eqs. (1) and (3) is important. Discussions of this topic can be found in Piershubert (1991), del-Castillo-Negrete and Morrison (1993), Brown and Samelson (1994), Ngan and Shepherd (1997), Balasuriya (2001), and Haynes et al. (2007). The foregoing arguments address dynamical consistency using ad hoc arguments and citing results from numerical simulations. It is very unsatisfying to have to resort to evidence based on full-physics numerical models to address a question as simple as the cause-and-effect question posed above. The use of full-physics numerical models is probably necessary, however, because simpler models, especially kinematic models, may be missing physics that is crucial to addressing the question posed. The underlying problem is the difficulty of finding nontrivial time-dependent solutions to (3). To further complicate this issue, analytically specified time-dependent flows that satisfy (3), if such solutions can be found, have properties (Brown and Samelson 1994) that are probably not representative of typical solutions to (3). Thus, resorting to evidence from full-physics numerical models and the use of ad hoc dynamical arguments may be unavoidable. The ad hoc dynamical arguments invoked above can be adapted to crudely account for non-$Q$-conserving situations such as a sudden significant warming (e.g., Robinson 1988; Shepherd et al. 2005) or cooling of the stratospheric polar vortex. These comments notwithstanding, it should be emphasized that the arguments presented in this paper are kinematic, with dynamical consistency addressed in a nonrigorous fashion.

Some comments on the limits of validity of the analysis presented are appropriate. Our use of the $\beta$-plane approximation is not essential. This approximation was introduced purely for convenience. The main price paid for this convenience is that our simulations do not capture the fundamental asymmetry of the polar vortex, with more vigorous stirring on the equatorward side (the surf zone) of the transport barrier. On the other hand, our assumption of a steady background flow is necessary, implicitly assuming that the polar vortex itself, which might be modeled as a patch of constant vorticity, is stable over a time scale of a month or more. The class of multiperiodic perturbations considered is surprisingly large. Recall that the number of terms $N$ included in the multiperiodic perturbation need only be finite and that we focused on the quasiperiodic case because this is the only mathematically nontrivial case to consider. Thus, with an appropriate choice of amplitudes and phases, the class of describable perturbations ranges from simple periodic to a deterministic wave packet (small $N$, narrow spectrum), to an approximation to white noise forcing [large $N$, flat and broad spectrum, random phases; see, e.g., Falkovich et al. (2001) for a discussion of the latter class of problems]. Finally, we note that, although KAM theory assumes that the perturbation strength $\varepsilon$ is small, numerical simulations show that invariant tori—and those in the vicinity of Kolmogorov-degenerate tori, in particular—persist for surprisingly large perturbation strengths. In the bottom panel in Fig. 4, for example, $\varepsilon = 0.7$. Thus, the small $\varepsilon$ assumption is less restrictive than one might expect.

The transport barrier at the perimeter of the strato-
spheric polar vortex that we have identified as being due to a thin band of KAM invariant tori can be described as a Lagrangian coherent structure. The subject of Lagrangian coherent structures has been extensively studied in recent years (see, e.g., Haller 2000, 2001; Haller and Yuan 2000, 2002; Malhotra and Wiggins 1998; Shadden et al. 2005). In most applications the Lagrangian coherent structures of interest are the stable and/or unstable manifolds of perturbed hyperbolic points. Unlike the KAM tori in our study, which constitute global barriers for transport, such invariant manifolds are barriers for transport only in a local sense and for a sufficiently short time. Also, while KAM tori are associated with regular motion, the stable and unstable manifolds are generically associated with chaotic motion in the vicinity of their points of intersections (homoclinic points).

A natural and important extension of the mostly theoretical work reported here is to use analyzed winds (following, e.g., Bowman 1993; Simmons et al. 2005; Koh and Legras 2002) to more thoroughly test the predictions made here. An empirical study of this type must employ spherical coordinates \([i.e., \phi = \psi(\lambda, \phi, t)]\) on a selected isentropic surface, where \(\lambda\) and \(\phi\) are longitude and latitude, respectively. Questions that could be addressed with such a model include the following. Is our hypothesized rest frame decomposition of \(\psi, \psi(\lambda, \phi, t) = \psi_0(\phi) + \psi_1(\lambda, \phi, t)\), where \(\psi_1\) is a superposition of zonally propagating Rossby waves, a good approximation? Are coarse-grained vorticity distributions consistent with those described above? Is Lagrangian motion predominantly chaotic, except in the vicinity of a thin wobbly transport barrier? Does the transport barrier coincide locally with a jet in \(Q\)? Does the transport barrier coincide locally with the jet core? Is the breakup of the transport barrier on a given isentropic surface caused by one of the mechanisms that we have described?\(^1\)

Acknowledgments. We thank the reviewers, all of whom made insightful comments that led to improvements to this paper. This work was supported by the U.S. National Science Foundation Grants CMG0417425 and OCE0648284.

---

\(^1\)Note added in proof: Two very recent publications address mathematical issues that are central to the results presented here. First, we (Rypina et al. 2007) have quantified the resonance width argument on which our robust transport barrier argument is based. Second, Sevryuk (2007) has provided a version of the KAM theorem for quasiperiodically perturbed one-degree-of-freedom Hamiltonian systems that satisfy the Russmann nondegeneracy condition. The results presented in both of these publications serve to strengthen the arguments made in this paper.

REFERENCES


Unauthenticated | Downloaded 10/21/23 03:25 PM UTC