Revisiting barotropic instability from the perspective of wave evolution theory

Yaokun Li[1]

[1]College of Global Change and Earth System Science, Faculty of Geographical Science, Beijing Normal University, Beijing
100875, China

Correspondence to: Yaokun Li (liyaokun@bnu.edu.cn)

Abstract. The instability of Rossby waves has been a long-standing topic in dynamical meteorology. The classic theoretical analysis had provided in–depth physical understanding of the problem. However, developing a systematic and quantitative comprehension of wave energy and amplitude evolution remains challenging. With an eye to such issues, this investigation provides a novel and practicable algorithm to solve the wave action conservation equation. Theoretical analysis establishes that wave packet energy evolves through two competing factors: direct proportionality to intrinsic frequency and inverse proportionality to group velocity magnitude. Energy density attains extremal values at turning points where group velocity magnitudes become extremized. To ensure a ray can be reflected by a turning point, zonal phase speed must be smaller than an upper limit determined by dispersion relation at the turning point. Crucially, a specific zonal phase speed range emerges below this maximum threshold where concurrent transient growth of both wave energy and amplitude occurs when a ray is moving toward the turning point, with the upper limit corresponding to optimal wave development conditions. Numerical experiments on a prototype westerly jet reveal distinct instability mechanisms: substantial transient growth–capable of triggering nonmodal instability–arises when rays moves toward turning points, while exponential amplification characteristic of modal instability develops at inflection points with positive energy growth rate. The derived zonal phase speed thresholds and transient growth metrics form a diagnostic framework applicable to observed atmospheric flows, enabling quantitative evaluation of both modal and nonmodal instability potentials. By unifying wave evolution dynamics with classical instability criteria, this work provides an operational bridge between theoretical predictions and real–world flow diagnostics.

1 Introduction

Rossby waves are among the most fundamental dynamical features in the upper atmosphere. A key characteristic of Rossby waves, which has profound implications for their modulation of weather systems, is their propagation (Segalini et al., 2024). Accompanied by transfers of heat, moisture, and momentum in propagation and amplification, they may play critical roles in shaping the weather and climate of polar regions (Woollings et al., 2023); the occurrence of extreme weather events in mid–latitudes (e.g., Ali et al., 2022; Jiménez-Esteve et al., 2022; Kornhuber et al., 2019); and tropical cyclone activity (e.g., Aiyyer and Wade, 2021; Riboldi et al., 2019). In some of these examples, amplified Rossby waves were thought to be

30 connected to extreme events through teleconnection. As Rossby waves intensify beyond a threshold, they can undergo wave breaking, triggering instability, a fundamental mechanism driving the formation of energy-rich eddies in the atmosphere (Read et al., 2020). Therefore, it is of great significance to study the development and instability of Rossby waves in the propagation process.

For barotropic instability specifically, classical normal mode theory has established well-known necessary conditions,
35 including the Rayleigh–Kuo criterion which mandates critical points where the absolute vorticity gradient of the basic flow reaches an extremum (Kuo, 1949; Rayleigh, 1880), Fjørtoft's complementary condition (Fjørtoft, 1950), and Howard's semicircle theorem (Howard, 1961). These foundational criteria are now enshrined in standard geophysical fluid dynamics textbooks (e.g., Pedlosky, 1987; Vallis, 2017). These classic necessary conditions, however, cannot be directly applied to identify instability since they only describe the characteristics of instability when it has already occurred. By applying the
40 concept of wave overreflection (Jones, 1968) where reflected waves exhibit amplified energy relative to incident waves via mean flow energy extraction mechanisms, Lindzen (1974) postulated that sustained overreflection, when confined by reflecting surfaces and quantized through boundary condition satisfaction, could generate unstable eigenmodes. Notably, the overreflection framework generalizes conventional modal instability by encompassing the broader class of phase–coherent ("quantization") amplified waves (e.g., Lindzen, 1974; Lindzen and Barker, 1985; Lindzen and Rosenthal, 1976; Lindzen
45 and Tung, 1978; Takehiro and Hayashi, 1992). Tung (1981) rigorously re–examined the quantization condition, demonstrating its sufficiency for barotropic instability when formulated as a spectral problem involving square–integrable real functions satisfying rigid boundary conditions with negative–definite Rayleigh quotients. Due to its mathematical complexity, this approach positions wave overreflection (or "wave geometry") as a complementary instability paradigm (Lindzen, 1988).

50 In addition to persistent amplification, Rossby waves may also undergo significant transient amplification before asymptotic decay. This transient growth mechanism (Pierrehumbert, 1983), first identified in the pioneering work of Orr (1907), constitutes a distinct instability paradigm—nonmodal instability—when exceeding critical thresholds. Its application to Rossby waves has been extensively discussed (Boyd, 1983; Farrell, 1984, 1985, 1982, 1988; Tung, 1983; Yamagata, 1976). Crucially, nonmodal instability can explain empirically observed turbulence onset at Reynolds numbers deemed
55 stable by normal mode theory (e.g., Boyd, 2018). A generalized stability theory constructed by Farrell and Ioannou (1996), emphasizing the central role of nonnormality in linearized dynamical systems, then applies to both the transient growth dynamics and normal mode stability of time–independent flows.

Compared to the fast phase propagation, the amplitude of a Rossby wave generally varies slowly, giving rise to the so-called Rossby wave packet, w to extreme weather and presumably has an influence on the predictability of mid-latitude
60 weather systems (Wirth et al., 2018). Therefore, the evolution and instability of Rossby waves can also be investigated by applying wave packet theory. For instance, Rossby waves in the westerly jet must exhibit northwest–southeast oriented phase tilts (south of the jet stream where the horizontal shear of the basic flow is positive) or northeast–southwest tilts (north

of the jet) to enable positive energy extraction from the basic flow's meridional shear (Chen and Chao, 1983; Pedlosky, 1987; e.g., Zeng, 1983). On the other hand, propagation characteristics emerge naturally through ray tracing theory, which defines

65   wave packet trajectories via local group velocity tangents. Early work by Yamagata (1976) and Hoskins and Karoly (1981) applied this formalism to barotropic Rossby waves. Ray tracing theory has been widely applied (e.g., Lu and Boyd, 2008) due to its successful explanation of atmospheric teleconnection patterns (e.g., Wallace and Gutzler, 1981).

Previous successes have revealed an inverse proportionality between wave amplitude and meridional wavenumber along rays, though they notably left unresolved the explicit energy evolution along propagation paths. Recently, Li et

70   al.(2021a) developed a diagnostic method correlating ray divergence with group velocity magnitude variation. Their subsequent studies implemented this approach to systematically characterize wave energy and amplitude evolution across diverse basic flow configurations (Li et al., 2021b, 2022; Li and Kang, 2022). These investigations revealed robust transient growth when rays approach turning latitudes, yet three fundamental questions remained unresolved: (1) criteria for concurrent energy–amplitude enhancement, (2) the role of critical layers in modulating growth thresholds, and (3)

75   quantitative evaluation of transient growth magnitudes for synoptic–scale Rossby wave packets. These unresolved aspects represent pivotal barriers to developing a unified instability theory through wave evolution diagnostics, motivating our systematic integration of wave evolution theory with barotropic instability diagnostics.

2 Wave evolution theory

The linearized vorticity equation for an incompressible barotropic atmosphere is typically expressed as

80

$$\left(\frac{\partial}{\partial t}+\bar{u}\frac{\partial}{\partial x}\right)\nabla^2\psi+\beta^*\frac{\partial\psi}{\partial x}=0,\tag{1}$$

where ψ is the perturbation stream function and $\bar{u}=\bar{u}(y)$ represents the zonal basic flow. $\nabla^2=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}$ is the

Laplace operator. $\beta^*=\beta-\frac{d^2\bar{u}}{dy^2}$ is the meridional gradient of absolute vorticity. $\beta=\frac{df}{dy}$ is the Rossby parameter and

assumed constant here and f is the Coriolis parameter.

Following the multi–scale asymptotic method (e.g., Pedlosky, 1987), the slowly varying wave train solution to Eq. (1)

85   is expressed as

$$\psi=\Psi(X,Y,T)e^{i(kx+ly-\omega t)},\tag{2}$$

where Ψ denotes the slowly varying amplitude, k and l are the zonal and meridional wavenumbers and ω is the frequency. Here $(X,Y,T) = (\varepsilon x, \varepsilon y, \varepsilon t)$ represent slow spatial/temporal variables, while (x,y,t) are fast

$\dfrac{\mathrm{D}_g}{\mathrm{D}T} = \dfrac{\partial}{\partial T} + \mathbf{c_g} \cdot \nabla$ spatial/temporal variables. The small dimensionless parameter $\varepsilon \ll 1$ quantifies the slowness of the

90 field's spatial and temporal variations. The slowly varying amplitude Ψ can be expanded asymptotically as

$$\Psi(X,Y,T) = \Psi_0(X,Y,T) + \varepsilon\Psi_1(X,Y,T) + \varepsilon^2\Psi_2(X,Y,T) + \cdots. \tag{3}$$

Substituting the wave train solution (Eq. (2)) into the linearized vorticity equation (Eq. (1)) and retaining zeroth–order terms in the asymptotic expansion yields the local dispersion relation:

$$\omega = k\bar{u} - \frac{\beta^* k}{K^2} \equiv \Omega(y,k,l), \tag{4}$$

95 where $K = \sqrt{k^2 + l^2}$ denotes the total wavenumber. Notably, this local dispersion relation mirrors the form of the standard plane wave solution, where the amplitude is assumed constant. This equivalence implies that the amplitude may be treated as locally constant at the leading order, such that the wave train approximates a plane wave in the lowest–order approximation.

2.1 Ray tracing theory

The local group velocity is defined as

$$
\begin{aligned}
c_{g,x} &\equiv \frac{\partial \Omega}{\partial k} = c + \frac{2\beta^* k^2}{K^4} \\
c_{g,y} &\equiv \frac{\partial \Omega}{\partial l} = \frac{2\beta^* kl}{K^4}
\end{aligned}, \tag{5}
$$

where $c = \omega/k$ denotes the zonal phase speed. Correspondingly, a ray is defined as the trajectory traced by a fluid parcel advected with the group velocity, satisfying

$$\frac{\mathrm{D}_g x}{\mathrm{D}T} \equiv \frac{\partial \Omega}{\partial k} = c_{g,x}, \frac{\mathrm{D}_g y}{\mathrm{D}T} \equiv \frac{\partial \Omega}{\partial l} = c_{g,y}, \tag{6}$$

where represents material derivative following the group velocity vector $\mathbf{c_g}$. Additionally, the following kinematic relations

105 hold

$$\frac{\mathrm{D}_g k}{\mathrm{D}T} \equiv -\frac{\partial \Omega}{\partial X} = 0, \frac{\mathrm{D}_g l}{\mathrm{D}T} \equiv -\frac{\partial \Omega}{\partial Y}, \frac{\mathrm{D}_g \omega}{\mathrm{D}T} \equiv \frac{\partial \Omega}{\partial T} = 0. \tag{7}$$

As Lighthill (2001) suggested, a wave packet evolving in position and wavenumber following Eqs. (6) and (7) may be interpreted as a "quasi–particle" whose energy–momentum relationship at every spatial point precisely mirrors the local frequency–wavenumber relationship (i.e., the local dispersion relation). This equivalence implies that wave packets exhibit
110 particle–like behavior, propagating along ray trajectories defined by the group velocity. Consequently, one can trace a wave packet's evolution from one location to another using Eqs. (6) and (7) without solving the full wave equations (Vallis, 2017). Ray tracing proves invaluable for two complementary reasons: First, it provides critical information on the spatial distribution of the wavenumber vector; second, it demonstrates that wave energy transport occurs along these ray trajectories (Lighthill, 2001).

115 Equation (7) demonstrates that zonal wavenumber (k) and frequency (ω) are conserved along a ray trajectory, as the dispersion relation does not explicitly depend on X and T. In contrast, meridional wavenumber (l) evolves dynamically along the ray path, leading to two critical meridional locations for ray propagation: a turning point ($y = y_t$) where $l = 0$ and meridional group velocity vanishes ($c_{g,y} = 0$), halting southward/northward propagation; a critical point ($y = y_c$), characterized by divergent meridional wavenumber ($l^2 \to \infty$) where group velocity aligns with zonal phase speed

120 ($\mathbf{c_g} = c\mathbf{i}$). Notably, the WKB solutions (Eqs. (2) and (3)) break down near turning points, where amplitude exhibit Airy function behavior in the meridional direction (Hoskins and Karoly, 1981). To reconcile this, asymptotic matching techniques employ a "patching" wave function in the turning point neighborhood, connecting the WKB solution in the "classical" region ($l^2 > \varepsilon$ where $\varepsilon \ll 1$) with the Airy function solution (e.g., Griffiths and Schroeter, 2018). Post–reflection from the turning point, the WKB approximation regains validity, enabling continued application of ray tracing beyond the critical

125 locus.

Previous studies (Li et al., 2021a; e.g., Yang and Hoskins, 1996) have categorized three distinct propagation regimes for Rossby wave packets. The evolution–dispersion (ED) regime, bounded between a single turning point and a critical point, typically exhibits initial propagation toward the turning point, followed by reflection and asymptotic approach to the critical point. Within this regime, Li et al. (2021a) observed potential transient amplification of both wave energy and amplitude

130 during the turning point approach. The wave guide (WG) regime, confined between two turning points, supports quasi–periodic ray reflections that enable long–distance propagation with minimal meridional wavenumber variation. Note that WG regime predicted by ray tracing theory may be a poor predictor for fairly narrow and strong jets where WKB approximation becomes invalid (Wirth, 2020). In contrast, the bidirectional dissipation (BD) regime, spanning dual critical points, drives unidirectional ray propagation toward any a critical point without reflection.

135 When the meridional gradient of absolute vorticity vanishes ($\beta^* = 0$, a condition corresponding to the Rayleigh–Kuo necessary instability criterion) at an inflection point ($y = y_i$), the dispersion relation Eq. (4) simplifies to $\omega = \overline{u}k$, rendering meridional wavenumber mathematically indeterminate. However, by applying the L'Hospital's rule to resolve this

indeterminate form, we derive a finite critical total wavenumber K_i (meridional shear of the basic flow is generally not set

to zero at the inflection point, that is, $\mathrm{d}\bar{u}/\mathrm{d}y\big|_{y=y_i} \neq 0$) satisfying

$$K_i^2 \equiv \lim_{y \to y_i} \frac{\beta^*}{\bar{u} - c} = \lim_{y \to y_i} \frac{\mathrm{d}\beta^*/\mathrm{d}y}{\mathrm{d}\bar{u}/\mathrm{d}y} = -\frac{\mathrm{d}^3\bar{u}/\mathrm{d}y^3\big|_{y=y_i}}{\mathrm{d}\bar{u}/\mathrm{d}y\big|_{y=y_i}} < \infty \ , \tag{8}$$

where $K_i > k$ ensures real meridional wavenumber solutions ($l_i = \pm\sqrt{K_i^2 - k^2}$) and physically admissible wave–like

structures. At $y = y_i$, zonal group velocity converges to zonal phase speed ($c_{g,x} = c$) while meridional group velocity

vanishes identically ($c_{g,y} = 0$), confining ray trajectories to purely zonal propagation along the inflection point.

Note that a fundamental discrepancy exists between the propagation regimes of ray tracing theory and the wave
geometry framework of overreflection theory, despite their shared dependence on critical and inflection points. In
overreflection dynamics (Lindzen, 1988; e.g., Lindzen and Tung, 1978), incident waves cannot reach critical points within
finite temporal scales unless inflection points coexist—a configuration enabling energy exchange between waves and the
mean flow.This contrasts fundamentally with ray theory, where critical point accessibility requires strict coincidence with
inflection points (namely, $y_i = y_c$) to permit finite–time arrival. Pedlosky (1987) elucidates a key constraint: marginally
stable waves, adjacent to unstable waves, must exhibit either vanishing Reynolds stress jumps or their self–canceling
summation. This condition holds unless the meridional gradient of absolute vorticity vanishes at y_c, that is, $\beta^*(y_c) = 0$.
Besides, the overreflection paradigm does not emphasize transient growth during ray propagation toward turning points.

2.2 Wave action conservation

Taking the first–order asymptotic approximation, the wave action conservation equation is expressed as:

$$\frac{\partial A}{\partial T} + \nabla \cdot (\mathbf{c_g} A) = 0 \,, \tag{9}$$

where $A \equiv E/\omega'$ denotes wave action density, $\omega' = \omega - \bar{u}k$ is intrinsic frequency, and the wave energy density E is

defined as

$$E = \frac{1}{4} K^2 \Psi_0^2. \tag{10}$$

Along ray trajectories, Eq. (9) simplifies to

$$\frac{1}{A}\frac{\mathrm{D}_g A}{\mathrm{D}T} = -\nabla \cdot \mathbf{c_g}. \tag{11}$$

This implies that the divergence of the group velocity ($\nabla \cdot \mathbf{c_g}$) serves as the sole source term governing variations in wave action density. However, direct evaluation of $\nabla \cdot \mathbf{c_g}$ using group velocity expression (Eq. (5)) remains impracticable as determining the partial derivatives in $\nabla \cdot \mathbf{c_g}$ requires knowledge of neighboring ray solutions (Lighthill, 2001).

Precise diagnostics of wave packet energy require solving the wave action conservation equation (Eq. (9) or Eq. (11)). In this framework, a wave packet propagating along a ray is conceptualized as a "quasi−particle" (e.g., Lighthill, 2001), where the spatial extent of this "quasi−particle" exceeds the wavelength scale while remaining negligible compared to the slowly varying background flow. Consequently, when viewed at the basic flow scale, the wave packet manifests as a dimensionless point (e.g., Bretherton and Garrett, 1968). Therefore, while ray tracing theory accurately describes the wave packet's path, it cannot resolve the intra−packet wave structure (e.g., Vallis, 2017). This scale separation implies that the divergence of the group velocity vector depends exclusively on variations in its magnitude—owing to suppressed lateral energy fluxes (Li et al., 2021a)—and can be expressed as

$$\nabla \cdot \mathbf{c_g} = \lim_{\delta T \to 0} \frac{\left|\mathbf{c_g}\right|_{T+\delta T} - \left|\mathbf{c_g}\right|_T}{\left|\mathbf{c_g}\right| \delta T} = \frac{1}{\left|\mathbf{c_g}\right|}\frac{\mathrm{D}_g \left|\mathbf{c_g}\right|}{\mathrm{D}T} = \frac{\mathrm{D}_g \ln\left|\mathbf{c_g}\right|}{\mathrm{D}T}, \tag{12}$$

where T denotes time, δT is a small time interval, and $\left|\mathbf{c_g}\right| = \sqrt{c_{g,x}^2 + c_{g,y}^2}$ represents the magnitude of the group velocity. It is critical to emphasize that this approximation (Eq. (12)) is derived at the basic flow scale, explicitly neglecting the wave packet's finite spatial extent.

Substituting Eq. (12) into Eq. (11) yields

$$\frac{A(t)}{A(0)} = \frac{\left|\mathbf{c_g}(0)\right|}{\left|\mathbf{c_g}(t)\right|}, \tag{13}$$

where $A(0)$ and $\left|\mathbf{c_g}(0)\right|$ denote the initial wave action density and initial group velocity magnitude, respectively. This relation reveals that the rate of change in wave action density is inversely proportional to the rate of group velocity magnitude variation along the ray trajectory. By invoking the definition of wave action density, we immediately derive

$$\frac{E(t)}{E(0)} = \frac{\omega'(t)}{\omega'(0)}\frac{\left|\mathbf{c_g}(0)\right|}{\left|\mathbf{c_g}(t)\right|}, \tag{14}$$

where $E(0)$ and $\omega'(0)$ represent the initial wave energy density and initial intrinsic frequency, respectively. This equation further demonstrates that the rate of change in wave energy density scales directly with the rate of intrinsic frequency variation and inversely with the rate of group velocity magnitude variation along the ray path.

185   Notably, these relations can be alternatively derived using the plane wave approximation: for a slowly varying wave train locally approximating a plane wave (constant amplitude), wave action density (and hence wave energy density) should exhibit no local variation, reducing Eq. (9) to $\nabla \cdot (A\mathbf{c_g}) = 0$. This implies that $E/\omega' \mathbf{c_g}$ is a solenoidal vector (Lighthill, 2001), meaning the flux of wave action through a thin ray tube remains conserved. Consequently, $E/\omega' \mathbf{c_g}$ is constant along the ray, yielding $E \sim \omega'/|\mathbf{c_g}|$—a result fully consistent with the earlier derivation.

3 Wave energy and amplitude variations

While substituting intrinsic frequency and group velocity expressions into Eq. (14) permits quantitative analysis of wave energy evolution along ray trajectories, this approach proves analytically intractable for better understanding the evolution dynamics. To circumvent this complexity, we will instead transform Eq. (11) to

$$\frac{1}{E}\frac{\mathrm{D}_g E}{\mathrm{D}T} = \left(\frac{2kl}{K^2}\frac{\mathrm{d}\overline{u}}{\mathrm{d}Y} - \frac{\mathrm{D}_g \ln|\mathbf{c_g}|}{\mathrm{D}T} \right) \equiv \xi(T), \tag{15}$$

195   whose solution is

$$E = E(0)\exp\left(\int_0^T \xi(\tau)\mathrm{d}\tau \right). \tag{16}$$

Equation (15) demonstrates that the temporal evolution of wave energy is governed by the balance between two competing mechanisms. The first term on the RHS quantifies energy extraction from the basic flow, where Rossby waves absorb kinetic energy through positive meridional shear ($\mathrm{d}\overline{u}/\mathrm{d}Y > 0$) when exhibiting northwest–southeast phase tilts ($kl > 0$ termed

200   leading structures). Conversely, northeast–southwest phase tilts ($kl < 0$ for trailing structures) facilitate energy transfer to basic flow, a cornerstone of classical Rossby wave evolution theory (Chen and Chao, 1983; Pedlosky, 1987; Zeng, 1983). The second term represents energy divergence/convergence driven by group velocity magnitude modulation. Accelerating rays induce energy dispersion, while decelerating rays promote energy accumulation.

     According to Eq. (5), group velocity magnitude square can be written as

$$|\mathbf{c_g}|^2 = \left(c + \frac{2\beta^* k^2}{K^4} \right)^2 + \left(\frac{2\beta^* kl}{K^4} \right)^2 = c^2 + 4k^2\frac{(\overline{u}-c)\overline{u}}{K^2}. \tag{17}$$

205

Then its variation along a ray can be expanded as

$$
\begin{aligned}
\frac{\mathrm{D}_g \left| \mathbf{c_g} \right|^2}{\mathrm{D}T} &= \frac{\mathrm{D}_g c^2}{\mathrm{D}T} + 4k^2 \frac{\mathrm{D}_g}{\mathrm{D}T} \left(\frac{\bar{u}}{K} \right)^2 - 4k^2 \frac{\mathrm{D}_g}{\mathrm{D}T} \left(\frac{\bar{u}c}{K^2} \right) \\
&= \frac{4k^2}{K^2} (2\bar{u} - c) \frac{\mathrm{D}_g \bar{u}}{\mathrm{D}T} + \frac{4k^2}{K^4} (c - \bar{u}) \bar{u} \frac{\mathrm{D}_g K^2}{\mathrm{D}T} \\
&= \frac{4k^2}{K^2} (2\bar{u} - c)(\bar{u} - c) \frac{2kl}{K^2} \frac{\mathrm{d}\bar{u}}{\mathrm{d}Y} - \frac{4k^2}{K^4} (\bar{u} - c) \bar{u} \left(-2kl \frac{\mathrm{d}\bar{u}}{\mathrm{d}Y} + \frac{\mathrm{d}\beta^*}{\mathrm{d}Y} \frac{2kl}{K^2} \right) \\
&= \frac{4k^2}{K^2} (\bar{u} - c)(3\bar{u} - c) \frac{2kl}{K^2} \frac{\mathrm{d}\bar{u}}{\mathrm{d}Y} + \frac{4k^2}{K^4} (\bar{u} - c) \bar{u} \frac{2kl}{K^2} \frac{\mathrm{d}^3 \bar{u}}{\mathrm{d}Y^3}
\end{aligned}
\tag{18}
$$

The second term on the RHS of Eq. (18), representing the third–order derivative of basic flow, becomes negligible under slowly varying basic flow approximation. Dividing Eq. (18) by $2 \left| \mathbf{c_g} \right|^2$, we can immediately derive an approximate expression for the ray divergence term in Eq. (15), that is

$$
\frac{\mathrm{D}_g \ln \left| \mathbf{c_g} \right|}{\mathrm{D}T} = \frac{\dfrac{2k^2}{K^2} (\bar{u} - c)(3\bar{u} - c)}{c^2 + \dfrac{4k^2}{K^2} (\bar{u} - c)\bar{u}} \frac{2kl}{K^2} \frac{\mathrm{d}\bar{u}}{\mathrm{d}Y} .
\tag{19}
$$

Substituting Eq. (19) into Eq. (15), we can derive

$$
\xi = \frac{c^2 - \dfrac{2k^2}{K^2} (\bar{u} - c)^2}{c^2 + \dfrac{4k^2}{K^2} (\bar{u} - c)\bar{u}} \frac{2kl}{K^2} \frac{\mathrm{d}\bar{u}}{\mathrm{d}Y} \equiv \lambda \left(c, \bar{u} \right) \frac{2kl}{K^2} \frac{\mathrm{d}\bar{u}}{\mathrm{d}Y} .
\tag{20}
$$

Eq. (20) further reveals that the two competing mechanisms governing the temporal evolution of wave energy can be synthesized into a single proportionality relation to the energy absorption mechanism with an effective coefficient λ that is dependent on zonal phase speed and basic flow. When $\lambda > 0$, energy absorption from the basic flow (mediated by leading structures in positive wind shear regions or trailing structures in negative wind shear regions) outpaces divergence of energy, resulting in net energy amplification. Conversely, $\lambda < 0$ indicates that energy divergence dominates over absorption, driving wave energy decay. This unified framework explicitly links local flow conditions (zonal phase speed, shear magnitude) to the net energetic outcome, providing a critical diagnostic for instability potential in barotropic wave systems.

3.1 Toward a turning point

Let's consider a case where a ray with a fixed zonal wavenumber propagates from an initial location y_0 toward a northern turning point at y_t ($c_{g,y} > 0$ and hence $l > 0$) that is situated south of a westerly jet ($\mathrm{d}\bar{u}/\mathrm{d}Y > 0$) with $\beta^* > 0$.

Besides, the meridional wavenumber monotonically decreases in $[y_0, y_t]$. When zonal phase speed is zero ($c = 0$ for

225 stationary waves), it is easy to derive $\lambda(0, \bar{u}) = -\dfrac{1}{2} < 0$. Then Eq. (20) becomes

$$\xi = -\frac{kl}{K^2}\frac{\mathrm{d}\bar{u}}{\mathrm{d}Y} < 0 , \tag{21}$$

since both meridional wavenumber and wind shear are positive when the ray propagates toward the northern turning point. It

reveals that stationary wave energy decays monotonically to a minimum at the turning point due to vanishing meridional

wavenumber ($l = 0$) at y_t. While the WKB approximation breaks down within a neighborhood to the turning point (e.g.,

230 $[y_t - \delta, y_t]$ where $\delta \ll 1$), it remains asymptotically valid at the boundary of the neighborhood (at $y = y_t - \delta$). This

dichotomy implies that stationary wave energy cannot physically attain the theoretical minimum value predicted at the

turning point; however, this minimum serves as a critical indicator of energy thresholds prior to reflection.

When zonal phase speed

$$c < \left(1 - \frac{K_0}{K_0 + \sqrt{2}k}\right)\bar{u}(y_0) \equiv c_a , \tag{22}$$

235 where $K_0 = \sqrt{k^2 + l_0^2}$ is the initial total wavenumber and l_0 is the initial meridional wavenumber at y_0, it is easy to derive

$$c < c_a \le \left(1 - \frac{K}{K + \sqrt{2}k}\right)\bar{u}(y) , \tag{23}$$

holds when the ray propagates toward the northern turning point. It demonstrates that $\lambda < 0$ and hence $\xi < 0$ on the ray

trajectory, meaning the wave energy will decay monotonically to a minimum at y_t.

Similarly, when

240 $$c > \left(2 - \sqrt{2}\right)\bar{u}(y_t) \equiv c_b , \tag{24}$$

it is easy to derive

$$c > c_b \ge \left(1 - \frac{K}{K + \sqrt{2}k}\right)\bar{u}(y) , \tag{25}$$

holds when the ray is on the way to y_t. It declares that $\lambda > 0$ and hence $\xi > 0$, indicating a monotonically amplifying

wave energy. Furthermore, the faster c is, the larger λ (and hence ξ) becomes, and the more powerful the wave energy

245 can enhance.

When zonal phase speed lies within the range (c_a, c_b), there will exist an intermediate point y_m within the range (y_0, y_t), allowing wave energy to first intensify to a maximum at y_m before decaying to a minimum at the northern turning point y_t. Finally, to make sure the ray can arrive at the northern turning point, there exists an upper limit for zonal group velocity, which can be expressed as

250

$$c_{\max} = \bar{u}(y_t) - \frac{\beta^*(y_t)}{k^2} < \bar{u}(y_t),$$ (26)

according to the dispersion relation Eq. (4). Note that we have preset $\beta^* > 0$ in above analysis.

To analyze amplitude variation, we consider the case where zonal phase speed falls within the range (c_a, c_b). From Eq. (10), the wave energy variation at y_m is

$$\left.\frac{D_g E}{D T}\right|_{y=y_m} = \frac{1}{4}\left.\frac{D_g K^2}{D T} \Psi_0^2\right|_{y=y_m} + \frac{1}{4}\left.\frac{D_g \Psi_0^2}{D T} K^2\right|_{y=y_m} = 0,$$ (27)

255   where the first term on the RHS represents meridional wavenumber reduction, generating negative contributions to energy growth. This necessitates positive contributions from the second term — the amplitude growth — to enable net zero amplification. Crucially, amplitude does not approach extremum at y_m. On the other hand, the first term vanishes at y_t so that the second must vanish simultaneously, meaning amplitude to attain its maximum at the turning point. This demonstrates a monotonic amplitude increase throughout the ray's northward trajectory.

260   To summarize, the evolutionary characteristics of wave energy and amplitude within a region (y_0, y_t) exhibit dependent behaviour on zonal phase speed. For slow zonal phase speed range ($0 < c < c_a$), wave energy undergoes monotonic decay to a minimum at the turning point (y_t) while amplitude conversely amplifies steadily to its maximum at this turning point. For intermediate range ($c_a < c < c_b$), energy demonstrates nonmonotonic evolution—first intensifying to a peak at a certain point (y_m) in the region ($y_0 < y_m < y_t$) before decaying to a minimum at y_t, while amplitude monotonically amplifying

265   to its maximum at y_t. For fast range ($c_b < c < c_{\max}$), concurrent monotonic amplification of both energy and amplitude achieves their respective maxima at y_t, where maximum intensification occurs as c approaches to its upper bound $c_{\max}$ when accessible. Notably, a wave whose zonal phase speed lies in the fast range can develop significantly. This is strongly consistent with the bounds on zonal phase speed in the normal mode instability theory (e.g., Howard, 1961; Pedlosky, 1987). From a mathematical standpoint, rays initiating propagation from critical points — where the meridional wavenumber

270   diverges ($l^2 \to \infty$)—exhibit unbounded amplification of both energy and amplitude. This divergence, however, violates the

scale separation prerequisite of the WKB approximation. Physical realizability therefore imposes constraints on meridional wavenumber as well as zonal wavenumber to ensure large–scale wave dynamics.

3.2 Toward an inflection point

Let's consider a case where a ray, with an initial meridional wavenumber l_0, is just moving from an initial point y_0 to an

275   inflection point y_i where the critical total wavenumber is K_i and corresponding meridional wavenumber is $\pm l_i$ according

to Eq. (8). When the ray arrives at y_i, meridional group velocity equals zero and zonal group velocity equals zonal phase

speed due to $\beta^* = 0$ there. It means the ray will horizontally move along y_i with a constant speed c. According to Eq.

(15), wave energy at y_i can be expressed as

$$E \sim \exp\left(\sigma_i T\right),\tag{28}$$

280   where $\sigma_i = \dfrac{2kl_i}{K_i^2}\dfrac{d\bar{u}}{dY}\bigg|_{y=y_i}$ is a constant growth rate, the sign of which depends on the critical meridional wavenumber and

meridional shear of the basic flow. Notably, a positive (negative) growth rate demonstrates exponential increase (decrease) of wave energy, and hence amplitude. The unlimited amplification of both wave energy and amplitude undoubtedly means modal instability, consistent with the critical layer problem in the classic instability theory (e.g., Pedlosky, 1987).

4 Quantitative analysis for a westerly jet prototype

285   To deepen our quantitative analysis of wave evolution, we utilize a prototypical westerly jet configuration previously examined by Kuo (1973), which writes as

$$\bar{u} = u_0 \operatorname{sech}^2\left(\frac{y}{d_0}\right),\tag{29}$$

where u_0 is the magnitude and d_0 is a measure for width of the westerly jet (Fig. 1a). Correspondingly, the meridional gradient of absolute vorticity β^* becomes

290    $$\beta^* = \beta_0 - u_0 \operatorname{sech}^2\left(\frac{y}{d_0}\right)\frac{1}{d_0^2}\left[4 - 6\operatorname{sech}^2\left(\frac{y}{d_0}\right)\right].\tag{30}$$

It is easy to identify that β^* has a maximum $M = \beta_0 + \dfrac{2u_0}{d_0^2}$ at $y = 0$ and tends to equal to β_0 when $y \to \pm\infty$ (Fig.

1b). Besides, β^* also has two equal minima $m = \beta_0 - \dfrac{2}{3}\dfrac{u_0}{d_0^2}$ at $y = d_0 \ln\sqrt{5 \pm 2\sqrt{6}}$.
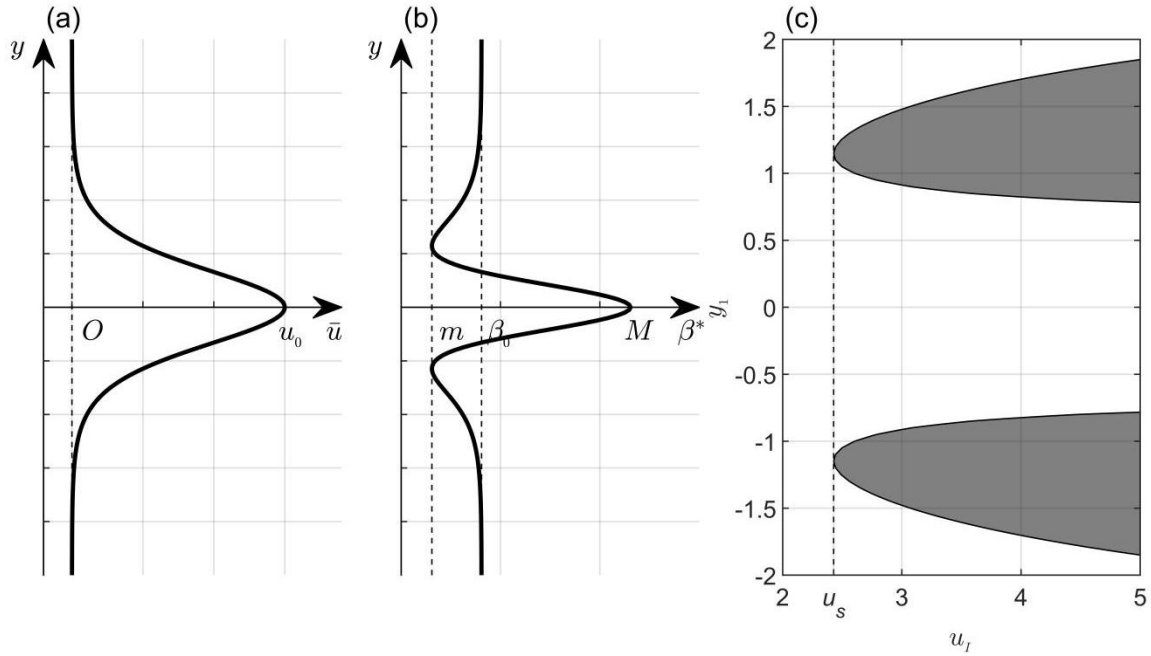
Figure 1: (a) Schematic spatial distribution of the westerly jet $\overline{u}$ defined by Eq. (29), where u_0 denotes the jet magnitude. (b) Schematic profile of the meridional gradient of absolute vorticity β^* as derived from Eq. (30), where m, M represent the minimum and maximum values of β^*, respectively, and β_0 denotes the Rossby parameter. (c) Shaded distribution of negative β_1^* plotted against non‒dimensional meridional location y_1 and westerly jet magnitude u_I. Negative β_1^* occurs when $u_I > \dfrac{3}{2}\beta_1 \equiv u_s$.

By setting $L = 10^6\,\mathrm{m}$ is the characteristic horizontal length; $U = 10\,\mathrm{m\,s^{-1}}$ is the characteristic value of the basic flow, and by setting

$$y_1 = \frac{y}{L}, \overline{u}_1 = \frac{\overline{u}}{U}, \omega_1 = \frac{L}{U}\omega, \beta_1 = \frac{\beta_0 L^2}{U}, k_1 = kL, l_1 = lL \ , \qquad (31)$$

where subscript "1" denote non–dimensional variables, the dispersion relation Eq. (4) is non–dimensionalized to

$$\omega_1 = \bar{u}_1 k_1 - \frac{\beta_1^* k_1}{K_1^2}, \tag{32}$$

where $\beta_1^* = \beta_1 - \dfrac{d^2 \bar{u}_1}{dy_1^2}$ is the non–dimensional gradient of absolute vorticity. The location where $\beta_1^* = 0$ (inflection

305   points) is crucial for instability problem. According to Eq. (30), $\beta_1^* < 0$ begins to emerge when non–dimensional

magnitude of the westerly jet $u_I \equiv u_0 / U > \dfrac{3}{2}\beta_1 \equiv u_s \approx 2.428$ (Fig. 1c), satisfying the Rayleigh–Kuo condition for

normal mode instability. In other words, there would be no normal model instability when $u_I < u_s$. However, nonmodal

instability caused by transient growth may still be possible. The wavelength of Rossby waves can vary from synoptic scale

of 103km order to planetary scale of 10^4km order. Correspondingly, the non–dimensional wavenumbers k_1 and l_1 can vary

310   from 2π to around 0.2π. Note that we have set matched sizes for zonal and meridional wavenumbers to make sure

analysis to be physically meaningful. The period for Rossby waves can vary from orders of 10^1 days to 10^2 days or longer

(infinity for stationary waves). The corresponding non–dimensional frequency ω_1 varies from around 0.7272 to zero (for

stationary waves).

4.1 Critical locations

315   When $\beta_1^* > 0$, the dispersion relation reduces to

$$c_1 = \bar{u}_1(y_t) - \frac{\beta_1^*(y_t)}{k_1^2} \equiv c_t < \bar{u}_1, \tag{33}$$

at turning point y_t. As analyzed in the previous section, c_t is the maximum zonal phase speed for a ray that can approach a

turning point. Substituting expressions of $\bar{u}_1$ and β_1^*, it is easy to that if $k_1^2 < 8$ (or $k_1 < \sqrt{8} \equiv k_c \approx 2.83$.

Corresponding wavelength is around 2200km), c_t has three extrema, two of which, labeled as c_M, are the same maxima

320   and the left of which, labeled as c_m, is the local minimum, that is

$$\begin{cases} c_M = \dfrac{3}{16}\dfrac{A}{B}\dfrac{u_I}{k_1^2} - \dfrac{\beta_1}{k_1^2}, y_1 = \ln\sqrt{\dfrac{20 - k_1^2 \pm 4\sqrt{3}\left(8 - k_1^2\right)}{k_1^2 + 4}} \equiv y_M \\ c_m = u_I - \dfrac{\beta_1 + 2u_I}{k_1^2}, y_1 = 0 \end{cases}, \tag{34}$$

where $A = C^2\left(k_1^2 + 4\right)^2, B = \left(\sqrt{3}\sqrt{8 - k_1^2} - 6\right)^4, C = 20 - k_1^2 - 4\sqrt{3}\sqrt{8 - k_1^2}$ for simplicity. If $k_1^2 > 8$, only c_m

survives and it becomes a maximum. Besides, to make sure a ray can arrive at a turning point, c_M should be larger than zero

if there are three extrema and c_m should be larger than zero if there are only one extremum. For the former case, $c_M > 0$

means $\dfrac{3}{16}\dfrac{A}{B}u_I > \beta_1$. For the latter case, $c_m > 0$ requires $u_I\left(k_1^2 - 2\right) > \beta_1$. At critical point y_c, the dispersion relation

reduces to

$$c_1 = \bar{u}_1\left(y_c\right) \equiv c_p \le u_I. \tag{35}$$

Zonal phase speed at critical point equals basic flow there. It is clear that c_p has a maximum, which equals the magnitude of

the westerly jet (u_I).

The propagation of a ray within distinct evolutionary regimes–ED, BD, and WG–is determined by its zonal phase speed.

Specifically, when $\beta_1^* > 0$, if $k < k_c$ (Fig. 2a), rays with zonal phase speed in the range $\left(0, c_M\right)$ propagate within the ED

region; those in $\left(c_M, c_p\right)$ occupy the BD region; and those in $\left(c_m, c_M\right)$ enter in the WG region, where their meridional

propagation is confined to specific meridional bounds. By contra $\left(c_m, c_p\right)$ st, if $k > k_c$ (Fig. 2b), rays with zonal phase

speed in $\left(0, c_m\right)$ propagate within the ED region, while those in fall into the BD region. When $\beta_1^* < 0$, intersection points

between turning and critical points emerge at $\beta_1^* = 0$. Owing to the symmetry properties of the westerly jet, we restrict our

analysis to regions ($y_1 < 0$) with positive meridional shear. Let the two intersection points be located at $-y_i$ and $-y_j$

($-y_i < -y_j < 0$), with corresponding zonal phase speeds labeled c_i and c_j ($c_i < c_j$), as illustrated in Fig. 2c and 2d.

When $-y_j < -y_M$, the propagation scenario (Fig. 2c) mirrors the $\beta_1^* > 0$ case in Fig. 2a, that is, rays with zonal phase

speed in $(0, c_M)$ propagates within the ED region. Conversely, when $-y_j > -y_M$, the zonal phase speed range defining

340   the ED regions shrinks to $(0, c_j)$ while the remaining (c_j, c_M) range transfers to a WG region (Fig. 2d).
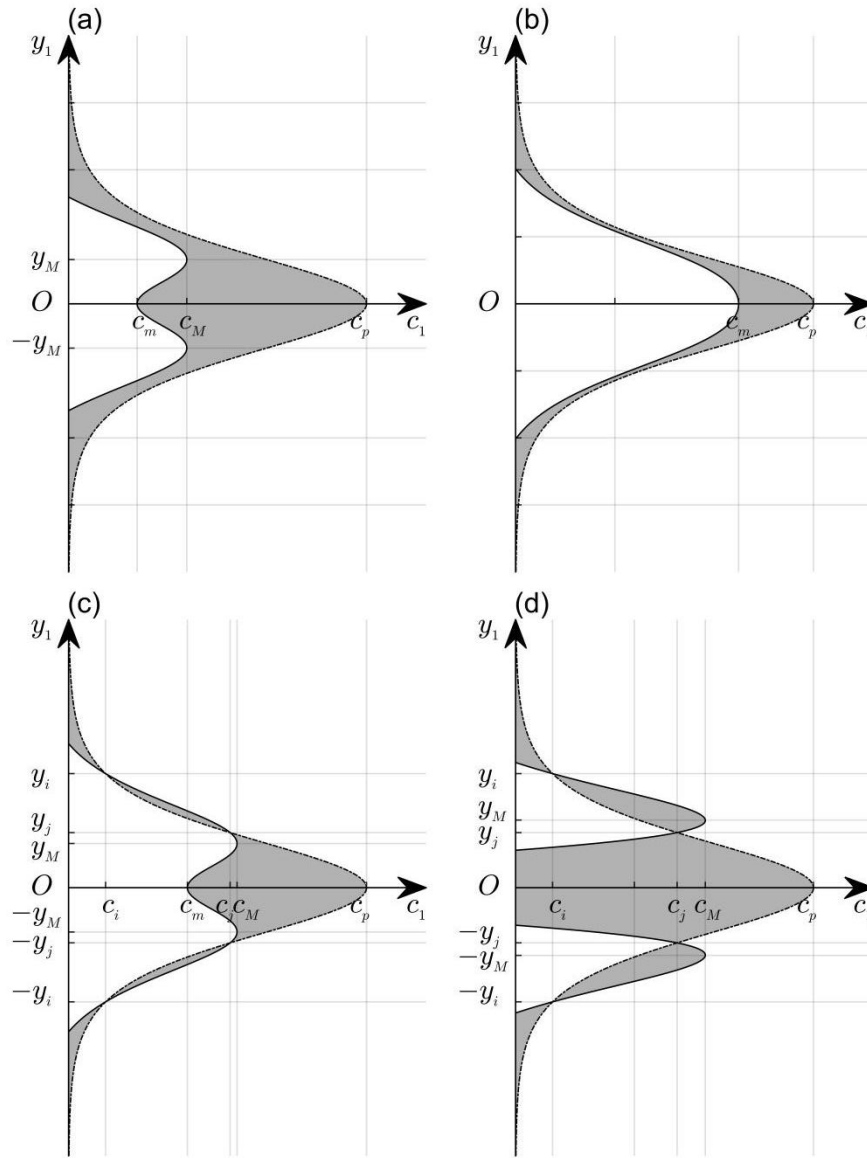
Figure 2: Schematic illustration of shaded propagation regions bounded by turning points (solid curve) and critical points (dash−dotted curves) under varying β_1^* and zonal wavenumber conditions: (a) $\beta_1^* > 0$ and $k_1 < \sqrt{8}$; (b) $\beta_1^* > 0$ and $k_1 > \sqrt{8}$; (c) $\beta_1^* < 0$, $k_1 < \sqrt{8}$, and $-y_j < -y_M < 0$; (d) $\beta_1^* < 0$, $k_1 < \sqrt{8}$, and $-y_M < -y_j < 0$. Intersection points of inflection

345   **and critical points occur at y_i and y_j ($y_i > y_j > 0$). Labels y_M and c_M denote the meridional location of turning point**

curve maximum and its corresponding zonal phase speed under condition $k_1 < \sqrt{8}$. c_m represents maximum zonal phase speed

associated with turning point curve under condition $k_1 > \sqrt{8}$. c_i and c_j correspond to zonal phase speeds at intersection points

y_i and y_j, respectively. c_p denotes the zonal phase speed at the critical point curve maximum.

4.2 Wave evolution in the ED region

350   Building on the preceding analysis, the zonal phase speed range defining the ED region is determined by critical thresholds

c_m, c_M, or c_j, which we collectively denote as $c_{\max}$ for simplicity (where $c_{\max}$ corresponds to c_m, c_M, or c_j

depending on the basic flow configuration). Consistent with prior theoretical analysis, this interval $\left(0, c_{\max}\right)$ can be

partitioned into three contiguous subranges – $\left(0, c_a\right)$, $\left(c_a, c_b\right)$, and $\left(c_b, c_{\max}\right)$ – as illustrated in Fig. 3. Within these

subranges, wave energy evolution follows distinct patterns: in $\left(0, c_a\right)$, energy decays monotonically to a minimum at the

355   turning point; in $\left(c_a, c_b\right)$, it first amplifies to a maximum before decaying to a local minimum at the turning point; and in

$\left(c_b, c_{\max}\right)$, it amplifies monotonically, peaking at the turning point. Amplitude, by contrast, demonstrates uniform

monotonic growth across the entire $\left(0, c_{\max}\right)$ range, reaching its maximum precisely at the turning point. The amplification

of wave energy toward the turning points aligns with transient growth mechanisms in barotropic Rossby waves within

uniformly sheared basic flows (e.g., Tung, 1983; Yamagata, 1976). The present study extends these findings to more

360   generalized basic flows, where analytical solutions for initial – value problems remain intractable, thereby enhancing the

applicability of transient growth theory to non–uniform shear environments.

Figure 3 further reveals that both wave energy and amplitude at the turning point are monotonically increasing

functions of zonal phase speed, which is consistent with prior theoretical prediction. For a wave with fixed zonal

wavenumber, this implies that maximum transient growth of both energy and amplitude occurs when the zonal phase speed

365   reaches $c_{\max}$, the upper bound of the ED region. In other words, $c_{\max}$ corresponds to the zonal phase speed at which waves

with fixed wavenumbers exhibit the most significant development (Fig. 3a) or even transition to potential instability (Fig.

3b). To focus our analysis, we examine the evolution of wave energy and amplitude specifically at the upper bound $c_{\max}$,

which is readily identifiable (Fig. 4) owing to the analytical westerly jet model. As illustrated in Fig. 4, the

$c_{\max}\left(k_1, u_I\right) = 0$ curve exhibits monotonic decay with increasing zonal wavenumber. Below this curve, the dashed contour

370   indicates that rays do not encounter any turning point along their trajectories. Above the curve, the farther a point lies from

the curve, the more rapidly the energy amplification rate becomes. Additionally, a second critical curve ($c_M = c_j$) emerges

(red solid curve in Fig. 4), above which the amplification rate become independent of zonal wavenumber and instead

increases monotonically with the intensity of the westerly jet. This configuration corresponds to scenarios where $c_{\max}$ assumes the value of c_j rather than c_m or c_M as depicted in Fig. 2d.

375   Figure 5 further illustrated the maxima of wave energy (Fig. 5a) and amplitude at turning points for ray with zonal phase speed $c_{\max}$. Given the amplitude's evolutionary pattern closely mirrors wave energy but with reduced magnitude, we focus our analysis on amplitude dynamics. For weak jet intensities—particularly when $u_I < 2.428$ (precluding normal mode instability)—transient growth upon reaching the turning point remains moderate. Consequently, Rossby waves exhibit moderate growth without transitioning to instability. By contrast, for jet intensities exceeding a critical threshold (e.g.,

380    $u_I > 3.0$), an optimal zonal wavenumber (or corresponding wavelength) emerges, coinciding with the strongest transient amplitude growth. For instance, when $u_I = 4.0$, a ray with $k_1 = 1$ (wavelength is around 6000km) amplifies to around 7.4 times its initial amplitude upon reaching the turning point (wave energy has a larger amplification). Although amplitude decays post–reflection from the turning point, the transient amplification is sufficient large to potentially trigger nonmodal instability. As jet intensity increases, the optimal zonal wavenumber shifts upward. When $u_I = 5.0$, this optimal

385   wavenumber corresponds to $k_1 \simeq 1.1$ (wavelength is around 5500km). Notably, the optimal zonal wavenumber lies above the $c_j = c_M$ defined boundary from Fig. 4 (reproduced in Fig. 5 for comparison). This indicates that the optimal zonal wavenumber resides within the ray propagation regime characterized in Fig. 2d.
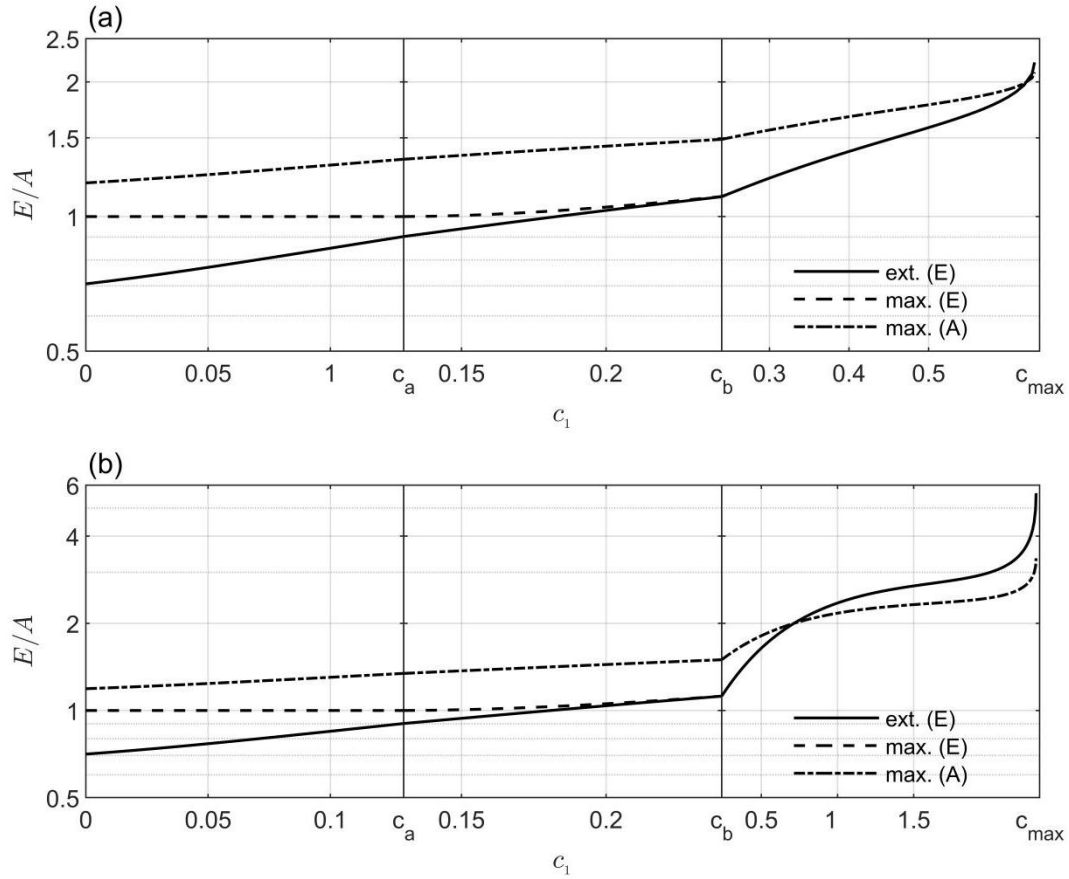
Figure 3: Variation of wave energy at turning points (solid curve), maximum wave energy (dashed curve), and maximum amplitude (dash−dotted curve) with zonal phase speed c_1. (a) For a case where $u_I = 1.5$; (b) for a case where $u_I = 4.0$. The wavelength is fixed as 3000km (corresponding to a zonal wavenumber $k \simeq 2.1$). Symbols $c_a \simeq 0.13$, $c_b \simeq 0.24$ denote divisions of zonal phase range $(0, c_{max})$ where $c_{max} \simeq 0.63$ specific to (a) and $c_{max} \simeq 2.3$ specific to (b). In $(0, c_a)$, energy decays monotonically to minima at turning points; in (c_a, c_b) it first amplifies to maxima before decaying to local minima at turning points; and in (c_b, c_{max}), it amplifies monotonically, peaking at turning points. By contrast, amplitude demonstrates uniform monotonic growth across in $(0, c_{max})$, reaching its maxima at turning points. Initial values of both wave energy and amplitude are normalized to unity.
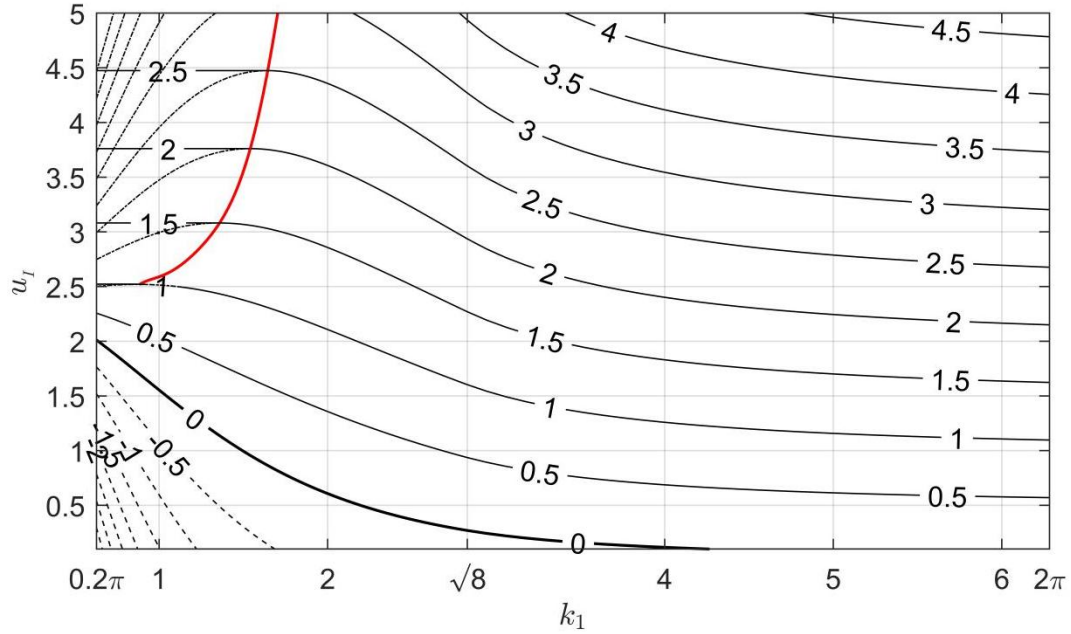
Figure 4: Contour plot of c_{max} (solid contour lines) as a function of zonal wavenumber k_1 and westerly jet magnitude u_I. The dashed contour denotes regions where the ED propagation regime does not exist due to $c_{max} < 0$. The red curve marks the critical threshold $c_M = c_j$. Above this curve, the dash–dotted contours represent c_M is superseded by c_j as the effective upper bound of the ED region due to the propagation dynamics depicted in Fig. 2d.
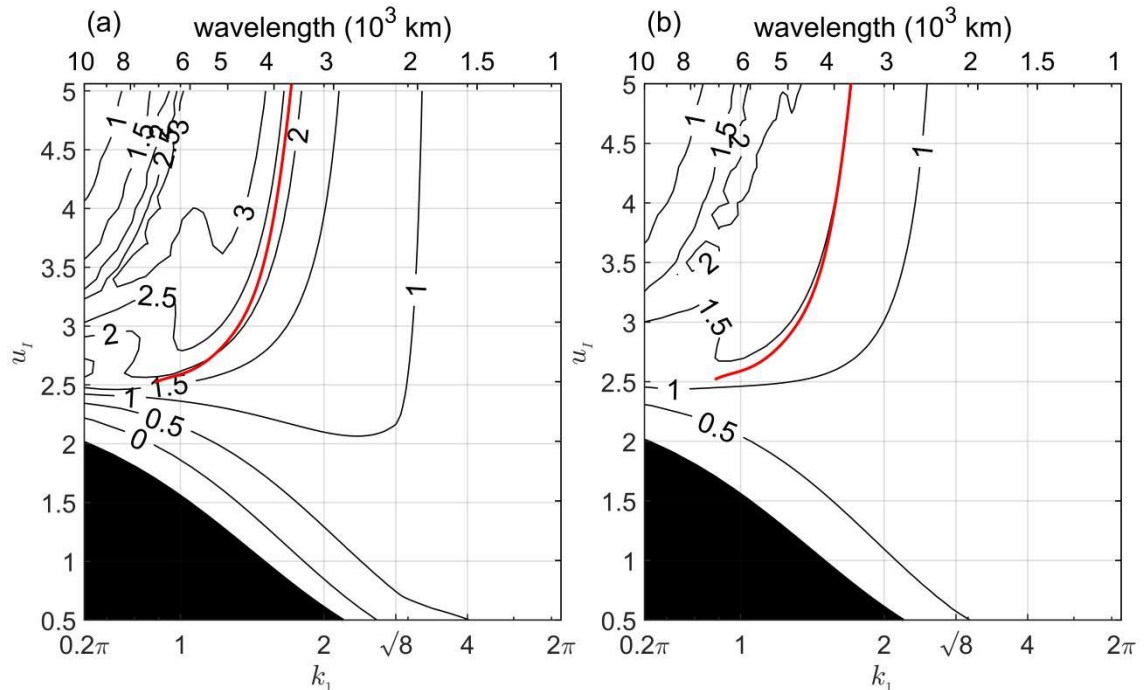
Figure 5: Logarithm of wave energy (a) and amplitude (b) at turning points for rays with zonal phase speed $c_{\max}$ **(as defined in Fig. 4). The regions where the ED regime is absent is masked. Initial wave energy and amplitude are normalized to unity.**

405 **4.3 The critical layer problem**

When a ray traverses a critical layer—a phenomenon extensively studied in geophysical fluid dynamics (e.g., Pedlosky, 1987)—we also focus on the region south of the jet center where meridional shear is positive. Consider a ray (solid curve in Fig. 6a) with zonal phase seed $c_{\max}$ (equivalent to c_j) propagating southward ($c_{g,y} < 0$ hence initial meridional wavenumber $l_0 < 0$) from an initial location $-y_0$ toward the inflection point $-y_i$ (also the critical point). Along this

410 trajectory, the negative meridional wavenumber (solid curve in Fig. 6b) increases monotonically to the negative critical value $-l_i$, then stabilizes as dictated by Eq. (8). This negative wavenumber corresponds to a negative growth rate, driving decay of both wave energy and amplitude (solid curve in Fig. 6c). Similarly, a northward–propagating ray (dashed curve in Fig. 6a) originating from $-y_0$ must also exhibit a negative meridional wavenumber (dashed curve in Fig. 6b) to satisfy the positive meridional group velocity constraint. Upon reaching the inflection point $-y_i$, this negative critical meridional wavenumber

415 $-l_i$ similarly induces a negative growth rate, resulting in declining energy and amplitude (dashed curve in Fig. 6c). In contrast, rays propagating southward from an initial location $-y_0$ ($-y_0 < -y_i$) encounter a southern turning point, where they undergo significant transient amplification. This pre–turning–point growth—sufficient to trigger nonmodal instability— renders reflection dynamics at the turning point less critical to the overall instability process.

Consider a ray originating exactly at the inflection point ($-y_0 = -y_j$). Owing to $c_{g,y} = 0$, the ray propagates zonally

420 at phase speed $c_{\max}$ (Fig. 7a). Although the initial meridional wavenumber l_0 at the inflection point is unconstrained, its evolution remains governed by the second equation in Eq. (7), that is, $\dfrac{\mathrm{D}_g l}{\mathrm{D}T} \equiv -\dfrac{\partial \Omega}{\partial Y}$. Within the prototype westerly jet,

$|l_0| < l_i$ drives $\dfrac{\mathrm{D}_g l}{\mathrm{D}T} > 0$, while $|l_0| > l_i$ induces $\dfrac{\mathrm{D}_g l}{\mathrm{D}T} < 0$. Consequently, as Fig. 7b illustrated, if $|l_0| < l_i$ (dashed curve) or $l_0 > l_i$ (solid curve), the meridional wavenumber approaches l_i ; if $l_0 < -l_i$ (dash–dotted curve), it monotonically decrease toward negative infinity. Correspondingly, wave energy exhibits exponential growth when meridional wavenumber

425 approaches l_i (solid and dashed curves in Fig. 7c) but decays exponentially when the meridional wavenumber tends toward negative infinity (dash–dotted curve in Fig. 7c).
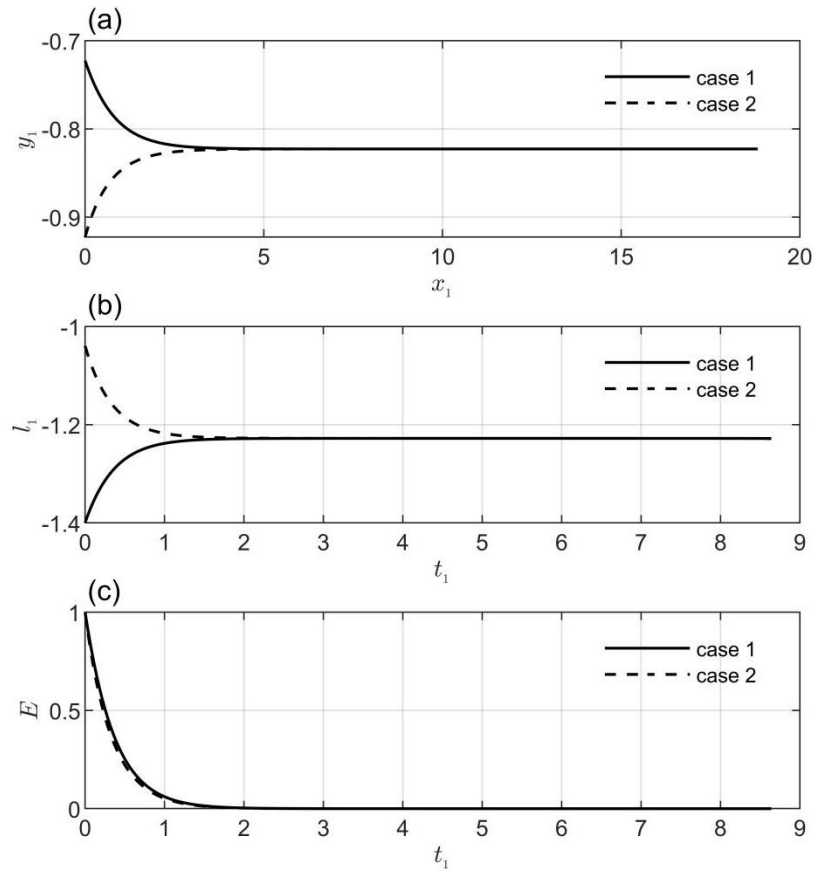
Figure 6: (a) Ray trajectories; (b) evolution of meridional wavenumber along rays; (c) wave energy evolution along rays. The solid (dashed) curve corresponds to a ray propagating southward (northward) the inflection point $y_j = -0.8227$ from an initial location $y_0 = y_i \pm 0.1$. Zonal wavenumber $k_1 = 1$; zonal phase speed $c_{max} = 2.1692$; and westerly jet magnitude $u_I = 4$.
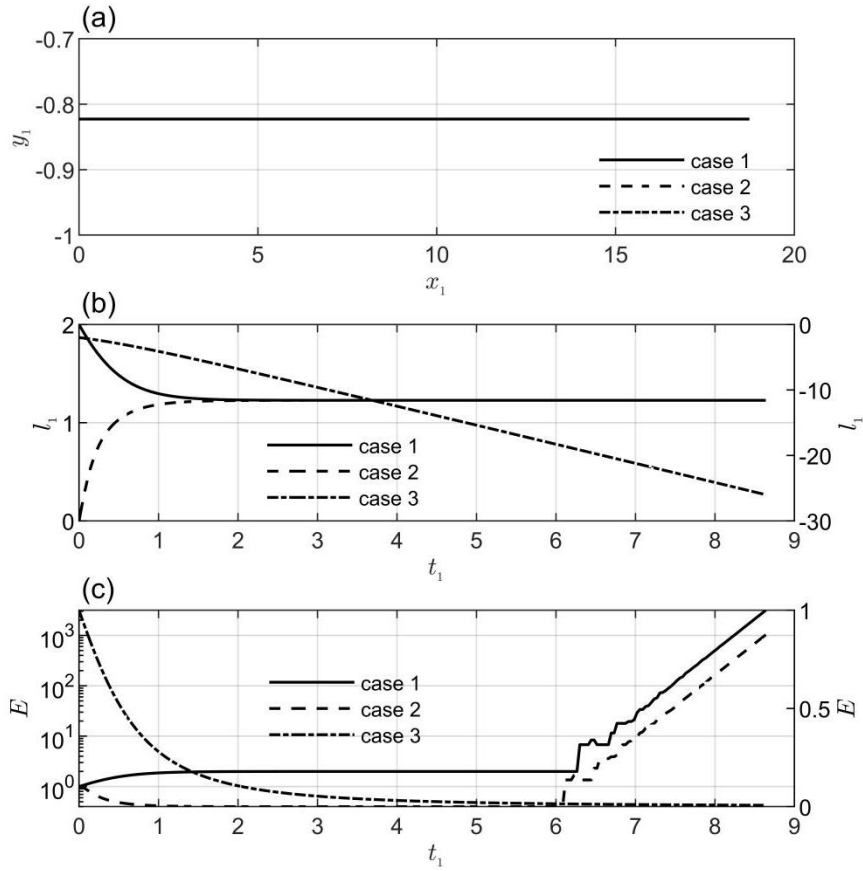
Figure 7: (a) Ray trajectories; (b) evolutionary of meridional wavenumber along ray trajectories; (c) wave energy evolution along ray trajectories. Initial meridional wavenumbers are assigned values of 2, 0, and −2, denoted by solid, dashed, and dash−dotted curves (labeled as Case 1, Case 2, and Case 3, respectively). Right−hand y−axes in panels (b) and (c) correspond to Case 3 for scale clarity. Zonal wavenumber $k_1 = 1$; zonal phase speed $c_{\max} = 2.1692$; initial meridional location $y_0 = -y_j = 0.8227$; and westerly jet magnitude $u_I = 4$, respectively.

4 Conclusions and discussion

The explicit physical understanding behind Rossby wave instability have been known for several decades. However, a general comprehension of Rossby wave instability in generalized basic flow setups is still an open question, which may help to interpret synoptic scale variability such as extreme weather events. This study re−examines classical barotropic instability through the lens of Rossby wave ray−tracing theory and wave action conservation. Analytical solutions reveal that wave energy evolution along ray trajectories exhibits two key dependencies: directly proportional to intrinsic frequency and inversely proportional to group velocity magnitude. Notably, wave energy attains extrema at turning points where group velocity magnitude reaches extremal values. A critical constraint emerges for rays to encounter a turning point during

445   propagation, that is, zonal phase speed must remain below an upper limit determined by dispersion relation at the turning point.

Three distinct energy evolution patterns are theoretically identified: (1) energy minima at turning points, (2) pre–turning–point maxima followed by decline to local minima, and (3) energy maxima at turning points under continuously accelerating zonal phase speed. On the contrast, wave amplitude maximizes at turning points. These findings suggest that

450   there exists a zonal phase speed range where simultaneous energy and amplitude growth means significant development of Rossby waves. Particularly, waves propagate with zonal phase speed equal to the upper limit is the most favorably developing. This provides dynamical support for Howard's (1961) semicircle theorem governing unstable waves.

The inherent symmetry of ray trajectories about a turning point creates complementary energy–amplitude evolution: growth towards a turning point will be counterbalanced by subsequent decay. This mechanism explains the transient growth

455   phenomenon first identified by Orr (1907) and later recognized as crucial for initiating nonmodal instability (Boyd, 1983; Farrell, 1984; e.g., Trefethen et al., 1993). At inflection points where meridional gradient of absolute vorticity vanishes, rays exhibit horizontal propagation with zonal phase speed that equals to zonal group velocity. Here, wave energy growth rate is determined by the product between critical meridional wavenumber and wind shear, with positive growth rate directly triggering exponential amplification characteristic of modal instability.

460   Quantitative calculation of a prototype westerly jet reveal critical insights into transient Rossby wave dynamics. When the Rayleigh–Kuo instability criterion remains unsatisfied, transient growth for Rossby waves with synoptic wavelengths proves to be moderate–a finding that aligns with their typical developmental (rather than destabilizing) behavior in such configuration. When the Rayleigh–Kuo criterion becomes satisfied, transient growth intensifies substantially to induce potential nonmodal instability when ray trajectories approach turning points, particularly when inflection points are

465   accessible for rays. At inflection points, where absolute vorticity gradients vanish, the emergence of constant positive growth rate will drive exponential energy accumulation, a hallmark of modal instability. Crucially, our analysis demonstrates that inflection points, if accessible for rays, serve as a primary standard for instability manifestation. This mechanism operates through sustained exponential growth characteristic of modal instability when rays approach inflection points with positive growth rate on the one hand, and via enhanced transient growth to trigger possible nonmodal instability when rays

470   propagating toward turning points on the other hand.

I envisage some future applications of the proposed analysis approach. For example, it provides a systematic assessment for Rossby wave evolution and instability in different basic flow profiles, especially in real–world atmospheric flows. By specifying known basic flow, practitioners can systematically identify turning points governing wave reflection, critical points regulating phase speed matching, and inflection points controlling modal instability onset. This diagnostic capability

475   enables two practical applications: (1) optimization of disturbance parameters through precise determination of wavelength/wavenumber combinations that maximize transient amplification potential, and (2) quantitative tracking of energy–amplitude co–evolution via closed–form analytical expressions. Besides, the methodology permits natural extension to baroclinic systems. Such extensions could bridge current barotropic insights with more realistic atmospheric scenarios

where vertical shear and stratification dominate Rossby wave evolution. This theoretical foundation ultimately provides a pathway for developing unified understanding of instability, either modal or nonmodal, in both barotropic and baroclinic flows.

Data availability

No datasets were generated or analyzed during the current study.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

This study is supported by the National Natural Science Foundation of China (Grants 42275051).

References

Aiyyer, A. and Wade, T.: Acceleration of tropical cyclones as a proxy for extratropical interactions: synoptic-scale patterns and long-term trends, Weather and Climate Dynamics, 2, 1051–1072, https://doi.org/10.5194/wcd-2-1051-2021, 2021.

Ali, S. M., Röthlisberger, M., Parker, T., Kornhuber, K., and Martius, O.: Recurrent Rossby waves and south-eastern Australian heatwaves, Weather and Climate Dynamics, 3, 1139–1156, https://doi.org/10.5194/wcd-3-1139-2022, 2022.

Boyd, J.: The continuous spectrum of linear Couette flow with the beta effect, Journal of the Atmospheric Sciences, 40, 2304–2308, 1983.

Boyd, J. P.: Nonmodal Instability, in: Dynamics of the Equatorial Ocean, edited by: Boyd, J. P., Springer Berlin Heidelberg, Berlin, Heidelberg, 311–327, https://doi.org/10.1007/978-3-662-55476-0_15, 2018.

Bretherton, F. P. and Garrett, C. J. R.: Wavetrains in inhomogeneous moving media, Proceedings of the Royal Society of London. Series A, 302, 529–554, https://doi.org/10.1098/rspa.1968.0034, 1968.

Chen, Y. Y. and Chao, J. P.: Conservation of wave action and development of spiral Rossby waves, SSPC-B0, 13, 663–672, https://doi.org/10.1360/zb1983-13-7-663, 1983.

Farrell, B.: Modal and Non-Modal Baroclinic Waves, Journal of the Atmospheric Sciences, 41, 668–673, https://doi.org/10.1175/1520-0469(1984)041%253C0668:MANMBW%253E2.0.CO;2, 1984.

Farrell, B.: Transient Growth of Damped Baroclinic Waves, Journal of the Atmospheric Sciences, 42, 2718–2727, https://doi.org/10.1175/1520-0469(1985)042%253C2718:TGODBW%253E2.0.CO;2, 1985.

505 Farrell, B. F.: The Initial Growth of Disturbances in a Baroclinic Flow, Journal of the Atmospheric Sciences, 39, 1663–1686, https://doi.org/10.1175/1520-0469(1982)039%253C1663:TIGODI%253E2.0.CO;2, 1982.

Farrell, B. F.: Optimal excitation of perturbations in viscous shear flow, Phys. Fluids, 31, 2093–2102, https://doi.org/10.1063/1.866609, 1988.

Farrell, B. F. and Ioannou, P. J.: Generalized Stability Theory. Part I: Autonomous Operators, Journal of the Atmospheric
510 Sciences, 53, 2025–2040, https://doi.org/10.1175/1520-0469(1996)053%253C2025:GSTPIA%253E2.0.CO;2, 1996.

Fjørtoft, R.: Application of integral theorems in deriving criteria of stability for laminar flows and for the baroclinic circular vortex, Geofis. Publ. 17, No. 6., 1950.

Griffiths, D. J. and Schroeter, D. F.: Introduction to Quantum Mechanics, 3rd edition., Cambridge University Press, Cambridge, UK, https://doi.org/10.1017/9781316995433, 2018.

515 Hoskins, B. J. and Karoly, D. J.: The Steady Linear Response of a Spherical Atmosphere to Thermal and Orographic Forcing, Journal of the Atmospheric Sciences, 38, 1179–1196, https://doi.org/10.1175/1520-0469(1981)038%253C1179:TSLROA%253E2.0.CO;2, 1981.

Howard, L. N.: Note on a paper of John W. Miles, Journal of Fluid Mechanics, 10, 509–512, https://doi.org/10.1017/S0022112061000317, 1961.

520 Jiménez-Esteve, B., Kornhuber, K., and Domeisen, D. I. V.: Heat Extremes Driven by Amplification of Phase-Locked Circumglobal Waves Forced by Topography in an Idealized Atmospheric Model, Geophysical Research Letters, 49, e2021GL096337, https://doi.org/10.1029/2021GL096337, 2022.

Jones, W. L.: Reflexion and stability of waves in stably stratified fluids with shear flow: a numerical study, Journal of Fluid Mechanics, 34, 609–624, https://doi.org/10.1017/S0022112068002119, 1968.

525 Kornhuber, K., Osprey, S., Coumou, D., Petri, S., Petoukhov, V., Rahmstorf, S., and Gray, L.: Extreme weather events in early summer 2018 connected by a recurrent hemispheric wave-7 pattern, Environ. Res. Lett., 14, 054002, https://doi.org/10.1088/1748-9326/ab13bf, 2019.

Kuo, H.: DYNAMIC INSTABILITY OF TWO-DIMENSIONAL NONDIVERGENT FLOW IN A BAROTROPIC ATMOSPHERE, Journal of the Atmospheric Sciences, 6, 105–122, https://doi.org/10.1175/1520-
530 0469(1949)006%253C0105:DIOTDN%253E2.0.CO;2, 1949.

Kuo, H. L.: On a simplified radiative-conductive heat transfer equation, pure and applied geophysics, 109, 1870–1876, https://doi.org/10.1007/BF00876111, 1973.

Li, Y. and Kang, Y.: Energy Dispersion of Westward Propagating Rossby Waves in Tropical Easterlies, SOLA, 18, 76–80, https://doi.org/10.2151/sola.2022-013, 2022.

535 Li, Y., Chao, J., and Kang, Y.: Variations in Wave Energy and Amplitudes along the Energy Dispersion Paths of Nonstationary Barotropic Rossby Waves, Adv. Atmos. Sci., 38, 49–64, https://doi.org/10.1007/s00376-020-0084-9, 2021a.

Li, Y., Chao, J., and Kang, Y.: Variations in Wave Energy and Amplitudes along the Ray Paths of Barotropic Rossby Waves in Horizontally Non-Uniform Basic Flows, Atmosphere, 12, 458, https://doi.org/10.3390/atmos12040458, 2021b.

Li, Y., Chao, J., and Kang, Y.: Variations in Amplitudes and Wave Energy along the Energy Dispersion Paths for Rossby
Waves in the Quasigeostrophic Barotropic Model, Adv. Atmos. Sci., 39, 876–888, https://doi.org/10.1007/s00376-021-1244-2, 2022.

Lighthill, J.: Waves in Fluids, 2nd ed., Cambridge University Press, Cambridge, UK, 524 pp., 2001.

Lindzen, R. S.: Stability of a Helmholtz Velocity Profile in a Continuously Stratified, Infinite Boussinesq Fluid--Applications to Clear Air Turbulence., Journal of the Atmospheric Sciences, 31, 1507–1514, https://doi.org/10.1175/1520-0469(1974)031%253C1507:SOAHVP%253E2.0.CO;2, 1974.

Lindzen, R. S.: Instability of plane parallel shear flow (toward a mechanistic picture of how it works), pure and applied geophysics, 126, 103–121, https://doi.org/10.1007/BF00876917, 1988.

Lindzen, R. S. and Barker, J. W.: Instability and wave over-reflection in stably stratified shear flow, Journal of Fluid Mechanics, 151, 189–217, https://doi.org/10.1017/S0022112085000921, 1985.

Lindzen, R. S. and Rosenthal, A. J.: On the instability of Helmholtz velocity profiles in stably stratified fluids when a lower boundary is present, Journal of Geophysical Research (1896-1977), 81, 1561–1571, https://doi.org/10.1029/JC081i009p01561, 1976.

Lindzen, R. S. and Tung, K. K.: Wave Overreflection and Shear Instability., Journal of the Atmospheric Sciences, 35, 1626–1632, https://doi.org/10.1175/1520-0469(1978)035%253C1626:WOASI%253E2.0.CO;2, 1978.

Lu, C. and Boyd, J. P.: Rossby Wave Ray Tracing in a Barotropic Divergent Atmosphere, Journal of the Atmospheric Sciences, 65, 1679–1691, https://doi.org/10.1175/2007JAS2537.1, 2008.

Orr, W. M.: The Stability or Instability of the Steady Motions of a Perfect Liquid and of a Viscous Liquid. Part I: A Perfect Liquid, Proceedings of the Royal Irish Academy. Section A: Mathematical and Physical Sciences, 27, 9–68, 1907.

Pedlosky, J.: Geophysical Fluid Dynamics, Springer New York, NY, 710 pp., https://doi.org/10.1007/978-1-4612-4650-3, 1987.

Pierrehumbert, R. T.: Bounds on the Growth of Perturbations to Non-Parallel Steady Flow on the Barotropic Beta Plane, Journal of the Atmospheric Sciences, 40, 1207–1217, https://doi.org/10.1175/1520-0469(1983)040%253C1207:BOTGOP%253E2.0.CO;2, 1983.

Rayleigh, Lord: On the Stability, or Instability, of certain Fluid Motions, Proceedings of the London Mathematical Society, s1-11, 57–72, https://doi.org/10.1112/plms/s1-11.1.57, 1880.

Read, P., Kennedy, D., Lewis, N., Scolan, H., Tabataba-Vakili, F., Wang, Y., Wright, S., and Young, R.: Baroclinic and barotropic instabilities in planetary atmospheres: energetics, equilibration and adjustment, Nonlinear Processes in Geophysics, 27, 147–173, https://doi.org/10.5194/npg-27-147-2020, 2020.

Riboldi, J., Grams, C. M., Riemer, M., and Archambault, H. M.: A Phase Locking Perspective on Rossby Wave Amplification and Atmospheric Blocking Downstream of Recurving Western North Pacific Tropical Cyclones, Monthly Weather Review, 147, 567–589, https://doi.org/10.1175/MWR-D-18-0271.1, 2019.

Segalini, A., Riboldi, J., Wirth, V., and Messori, G.: A linear assessment of barotropic Rossby wave propagation in different background flow configurations, Weather and Climate Dynamics, 5, 997–1012, https://doi.org/10.5194/wcd-5-997-2024, 2024.

575   Takehiro, S.-I. and Hayashi, Y.-Y.: Over-reflection and shear instability in a shallow-water model, Journal of Fluid Mechanics, 236, 259–279, https://doi.org/10.1017/S0022112092001411, 1992.

Trefethen, L. N., Trefethen, A. E., Reddy, S. C., and Driscoll, T. A.: Hydrodynamic Stability Without Eigenvalues, Science, 261, 578–584, 1993.

Tung, K. K.: Barotropic Instability of Zonal Flows, Journal of Atmospheric Sciences, 38, 308–321,
580   https://doi.org/10.1175/1520-0469(1981)038%253C0308:BIOZF%253E2.0.CO;2, 1981.

Tung, K. K.: Initial-value problems for Rossby waves in a shear flow with critical level, Journal of Fluid Mechanics, 133, 443–469, https://doi.org/10.1017/S0022112083002001, 1983.

Vallis, G. K.: Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-Scale Circulation, 2nd ed., Cambridge University Press, Cambridge, https://doi.org/10.1017/9781107588417, 2017.

585   Wallace, J. M. and Gutzler, D. S.: Teleconnections in the Geopotential Height Field during the Northern Hemisphere Winter, MONTHLY WEATHER REVIEW, 109, 784–812, https://doi.org/10.1175/1520-0493(1981)109%253C0784:TITGHF%253E2.0.CO;2, 1981.

Wirth, V.: Waveguidability of idealized midlatitude jets and the limitations of ray tracing theory, Weather and Climate Dynamics, 1, 111–125, https://doi.org/10.5194/wcd-1-111-2020, 2020.

590   Wirth, V., Riemer, M., Chang, E. K. M., and Martius, O.: Rossby Wave Packets on the Midlatitude Waveguide—A Review, Monthly Weather Review, 146, 1965–2001, https://doi.org/10.1175/MWR-D-16-0483.1, 2018.

Woollings, T., Li, C., Drouard, M., Dunn-Sigouin, E., Elmestekawy, K. A., Hell, M., Hoskins, B., Mbengue, C., Patterson, M., and Spengler, T.: The role of Rossby waves in polar weather and climate, Weather and Climate Dynamics, 4, 61–80, https://doi.org/10.5194/wcd-4-61-2023, 2023.

595   Yamagata, T.: On trajectories of Rossby wave-packets released in a lateral shear flow, Journal of the Oceanographical Society of Japan, 32, 162–168, https://doi.org/10.1007/BF02107270, 1976.

Yang, G.-Y. and Hoskins, B. J.: Propagation of Rossby Waves of Nonzero Frequency, Journal of the Atmospheric Sciences, 53, 2365–2378, https://doi.org/10.1175/1520-0469(1996)053%253C2365:PORWON%253E2.0.CO;2, 1996.

Zeng, Q.-C.: The Evolution of a Rossby-Wave Packet in a Three-Dimensional Baroclinic Atmosphere, Journal of the
600   Atmospheric Sciences, 40, 73–84, https://doi.org/10.1175/1520-0469(1983)040%253C0073:TEOARW%253E2.0.CO;2, 1983.