Dear Reviewer:

Thank you very much for your insightful and constructive comments on my manuscript. I greatly appreciate the time and effort you have dedicated to reviewing my work, which has significantly helped improve the quality and clarity of the paper. I have carefully addressed all your concerns and revised the manuscript accordingly. Below is a summary of the key revisions made:

1. I have compared the case discussed in the manuscript with the cases for the observed westerly jet to stress that it can represent the basic feature of the real world atmosphere.

2. I have rewritten Section 3.1 to enhance logical flow and readability. A new schematic diagram (shown in Response Figure 5) has been added to illustrate key concepts.

3. I have added a new Appendix to introduce the detailed derivation for wave energy equation along rays.

4. I have added a paragraph to explicitly discuss the ray tracing limitations regarding wave superposition and tunneling, citing Harnik (2002) and Harnik & Heifetz (2007) to strengthen theoretical rigor.

5. I have carefully checked the manuscript to avoid mistakes and typos.

I believe these revisions have substantially improved the manuscript’s clarity, rigor, and accessibility. I hope the revised version meets your expectations and look forward to your further feedback.

Sincerely,

The author: Yaokun Li

1. A main point which bothers me is the relevance of the analysis in point 3.1 for the observed midlatitude flow (which I take it is what the paper is aimed at helping explain?) Specifically, the setup of a turning point south of the westerly jet, with the wave propagation to the south of the turning surface is opposite what is found in observations, for the case of waves propagating along a jet and being reflected equatorward from its poleward flank. A more realistic setup would be a turning point in a region of negative meridional shear, for example. It is also not explicitly stated if dU/dY and beta* are constant or not for this case. A schematic of the mean flow, where in the atmosphere this is relevant, and what phenomena does it represent are needed to convince the reader of the relevance of this analysis to the real atmosphere.

Response: Thank you for your insightful comment. To address your concern, I first present the distribution of the observed zonal mean zonal wind. As portrayed in Response Figure 1a, the major features of the annual, winter (December–January–February, DJF), and summer (June–July – August, JJA) mean zonal wind include two strong westerly jets dominating the subtropics in each hemisphere and a moderate easterly around the equator. Consequently, Eq. (29) in Section 4 serves as a valid theoretical prototype of the observed westerly jet. Notably, Eq. (29) is a widely applied westerly jet profile in theoretical analysis (e.g., Kuo 1973).

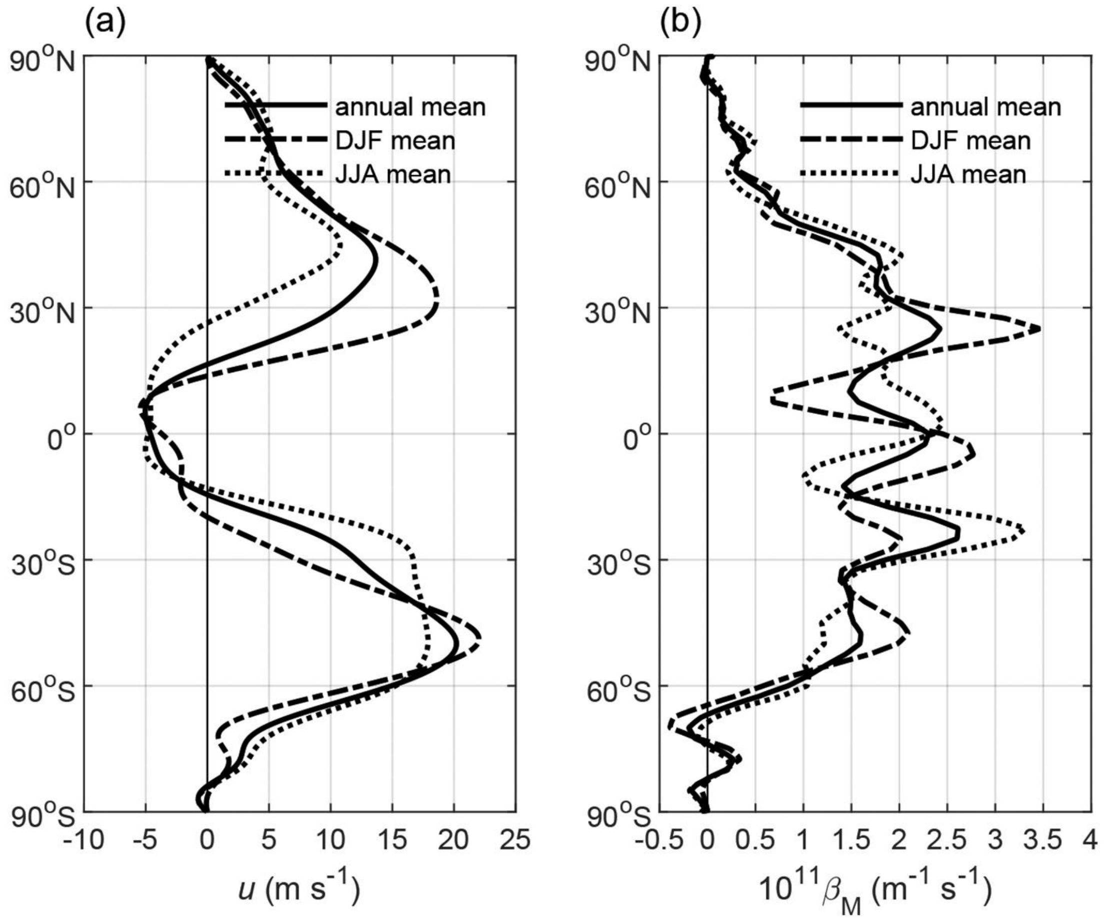
By applying the observed zonal mean zonal wind, we can readily identify the turning latitudes and critical latitudes that bound the propagative regions (see Response Figure 2). To illustrate, consider the DJF season: in the Northern Hemisphere winter, the westerly jet center is located around 30°N. The propagative regions (green-shaded in Response Figure 2) exhibit three distinct patterns: (1) for zonal wavenumbers 2 – 4, the propagative region is enclosed by two critical latitudes; (2) for wavenumbers 5 – 6, it is bounded by a southern critical latitude and a northern turning latitude (north of the jet center); and (3) for wavenumbers ⩾7, it is bounded by a southern critical latitude and a northern turning latitude (south of the jet center). Thus, the position of the turning latitude relative to the jet center depends on zonal wavenumber (Response Figure 2a), wave period (Response Figure 2b), and even the westerly jet profile (see Figure 2 in the submitted manuscript).

For the given profile of the westerly jet, we find that if the propagative regions for a ray are located south of the jet center, the turning latitude is also located south of the jet center (as shown in Figure 2 in the submitted manuscript). Therefore, we do not analyze the case in which a ray can cross the jet center. This case can, however, be theoretically analyzed as:

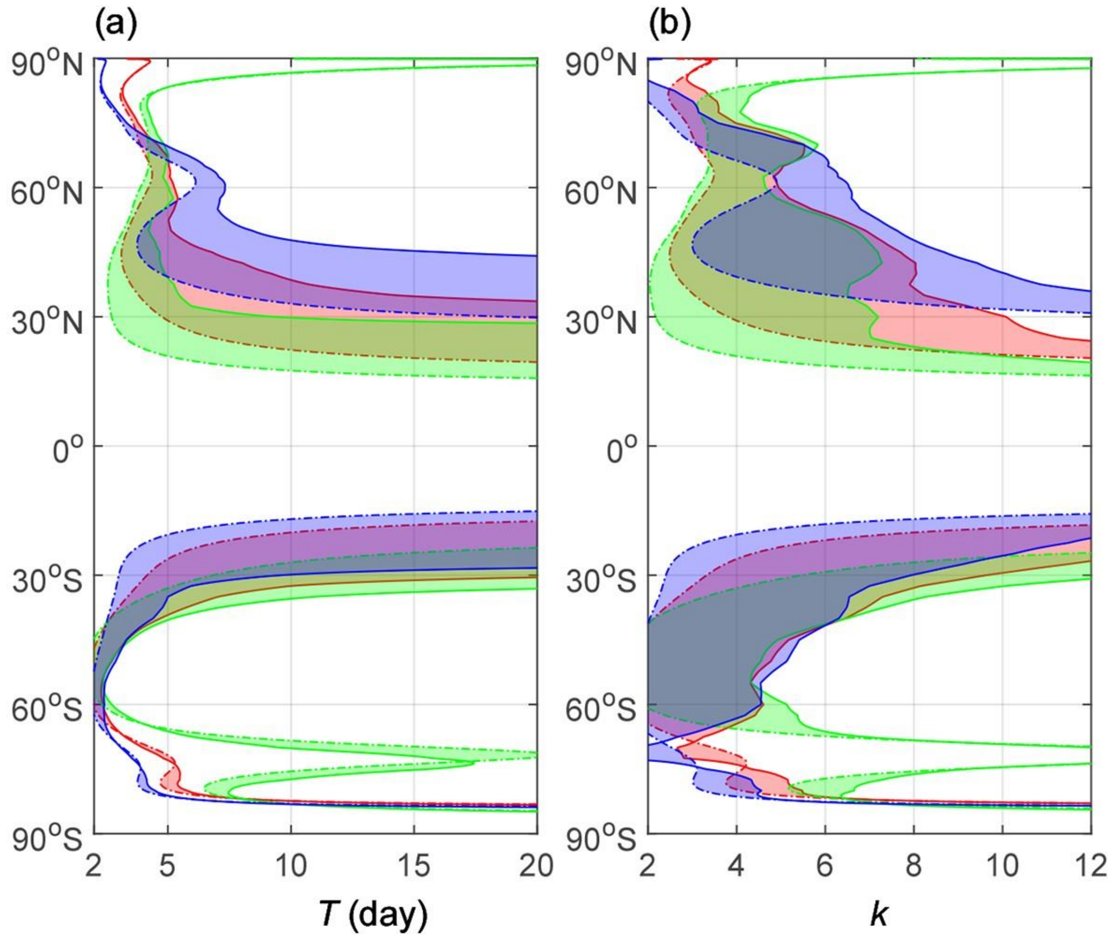
When the turning latitude lies north of the jet center, a ray (with positive initial zonal and meridional wavenumber) originating south of the jet center can cross the jet center before being reflected by the turning latitude. At the jet center, the meridional gradient of the westerly wind (dU/dY) vanishes. According to Eqs. (15) and (20) in the submitted manuscript, the wave energy density approaches an extreme value (since dU/dY > 0 south of the jet center and dU/dY < 0 north of the jet center). When continuing moving northward, the wave energy reaches another extreme value at the turning latitude (due to zero meridional wavenumber at the turning latitude). The above theoretical analysis becomes more complex when considering the modulation of λ, as indicated by Eq. (20), but the general behavior remains similar. I will explicitly analyze the case where λ varies.

Notably, in this theoretical framework, we only require the signs of dU/dY and β* rather than their specific values, confirming that these parameters are not assumed to be constant.

Following your comment and your below comments, I have carefully revised Section 3.1 to emphasize the theoretical setup is reasonable and to extend to cases where the turning latitude is located north of the jet center. I have also carefully revised the derivation to make sure the theoretical derivation is easier to follow. Thank you for your constructive comment.



Response Figure 1. (a) Meridional distribution of the annual (solid line), DJF (dash-dotted line), and JJA (dotted line) mean zonal wind, and (b) the corresponding meridional gradient of the potential vorticity. This figure is cited from Li et al. (2021).



Response Figure 2 Energy dispersion regions (shaded) bounded by a turning latitude (solid line) and by a wave trap line (dot-dashed line) for the annual (red), DJF (green) and JJA (blue) mean zonal wind: (a) for zonal wavenumber k = 8; (b) for a period of T = 10 days. This figure is cited from Li et al. (2021).

2. When the wave field consists of and incident and reflected wave, the group speed is not defined. Rather the wave consists of two waves with oppositely directed group velocities. Specifically, the sentence on lines 130-133: strictly speaking, the back and forth reflection allows the meridional wavenumber to keep constant, but only if you take into account the superposition of the two meridionally reflected waves. This means the solution is not a pure plane wave, and the relevance of the ray tracing formulation for understanding the actual amplitude is not clear. This point is discussed in Harnik (2002).

Response: Thank you for your insightful comment. I have read the paper by Harnik (2002) and regret not having consulted this work during the preparation of the manuscript. I fully agree with the conclusion that ray tracing theory fundamentally describes the trajectory of a wave packet for a quasi-plane wave with specified initial zonal/meridional wavenumbers and initial positions. In general, however, we do not observe isolated quasi-plane waves but rather superpositions of such waves. Consequently, I have added a paragraph in the ray tracing theory section to explicitly address this limitation. In the present study, following the classical theoretical framework, I focus on a single wave ray with a given wave vector propagating from a point source. Under this specific condition, the ray tracing method is equivalent to the wave packet path as Harnik (2002) suggested. Thank you again for your constructive feedback, which has strengthened the rigor of our analysis.

3. Paragraph on lines 144-152: another fundamental difference between ray tracing and the index-of refraction type of analysis that is at the heart of over-reflection theory is the inability of ray tracing to represent tunnelling - see e.g. Harnik (2002). Also, I am not sure I fully follow the sentence starting with "this contrasts" - over-reflection requires the waves to approach the critical surface via tunnelling through an evanescent region. This can happen also when the critical surface does not coincide with the turning surface (inflection point). See e.g. Harnik and Heifetz (2007).

Response: Thank you for insightful comment. I agree with Harnik's conclusion that rays cannot represent tunneling. Actually, this is consistent with my comparison here. Since rays cannot represent tunneling, it naturally cannot arrive at the critical level in a finite time. The expression here is confusing. I am sorry for that and I will made an explicit explanation below.

Firstly, the concepts such as the critical level where the zonal phase speed equals the basic flow ($\overline{u}-c=0$ at $y=y_{\text{turn}}$), the turning level where the refractive index vanishes ($l=0$ at $y=y_{\text{critical}}$), and the infection level where the meridional gradient of the absolute vorticity vanishes ($q_y=\beta-\partial^2\overline{u}/\partial y^2=0$ at $y=y_{\text{inflection}}$) are applicable for both overreflection and ray tracing.

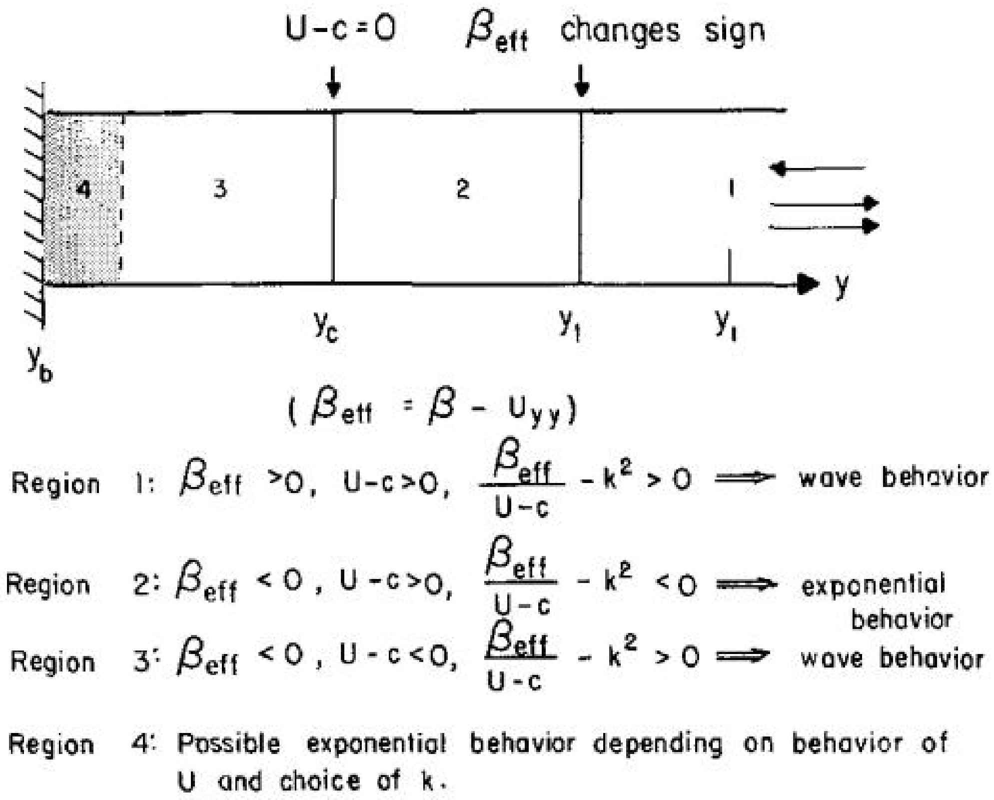
Secondly, Lindzen (1988) wrote: "there must at least be some $y=y_{\text{critical}}<y_1$ where $U=c$ in order to get overreflection. The simple existence of y_{critical} is, however, not enough. LINDZEN and TUNG (1978) showed that a necessary and sufficient condition for overreflection is that $\beta-U_{yy}<0$ at y_{critical} (or more generally that $\beta-U_{yy}$ have the opposite sign at y_{critical} that it has at y_1). Thus, to get overreflection, we need a $y=y_{\text{inflection}}$ ($y_{\text{critical}}<y_{\text{inflection}}<y_1$) where $\beta-U_{yy}=0$ -- i.e., an inflection point. The necessary condition for normal mode instability is a necessary and sufficient condition for wave overreflection". Correspondingly, "the geometry of wave propagation, implied by the above conditions, turns out to be important." The wave geometry can be clearly seen from Figure 2 in his paper. I also cite it here as Response Figure 3. As shown in the figure, we need a region 1 where the index of refraction is positive to enable wave propagation. Besides, we also need region 2 where the index of refraction is negative due to negative meridional gradient of the absolute vorticity. Furthermore, regions 3 where the index of refraction is positive is also important. **Note that in the wave geometry, the inflection level does not coincide with the critical level so that there exists a region 2, which is thought fundamental to overreflection.**

Thirdly, region 2, or called wave evanescence (EV) region in the investigation by Harnik and Heifetz (2007), is also necessary. This is consistent with Lindzen's conclusion. However, the turning level (TL) in the investigation by Harnik and Heifetz (2007) seems to coincide with the inflection level. To clearly show you, I also cite Figure 2 in their investigation as Response Figure 4 in this response file. Seen from the figure, the index of refraction is positive below the TL while negative above the TL. Therefore, we can infer that $n^2 = \frac{q_y}{U-c} - k^2 = 0$ at TL. On the other hand, q_y is negative below the TL while positive above the TL. Therefore, $q_y = \beta - U_{yy} = 0$ at TL. Since both two variables simultaneously equal zero at the TL, we may derive that $k = 0$ at the TL. This is a little confusing and maybe I missed something important.

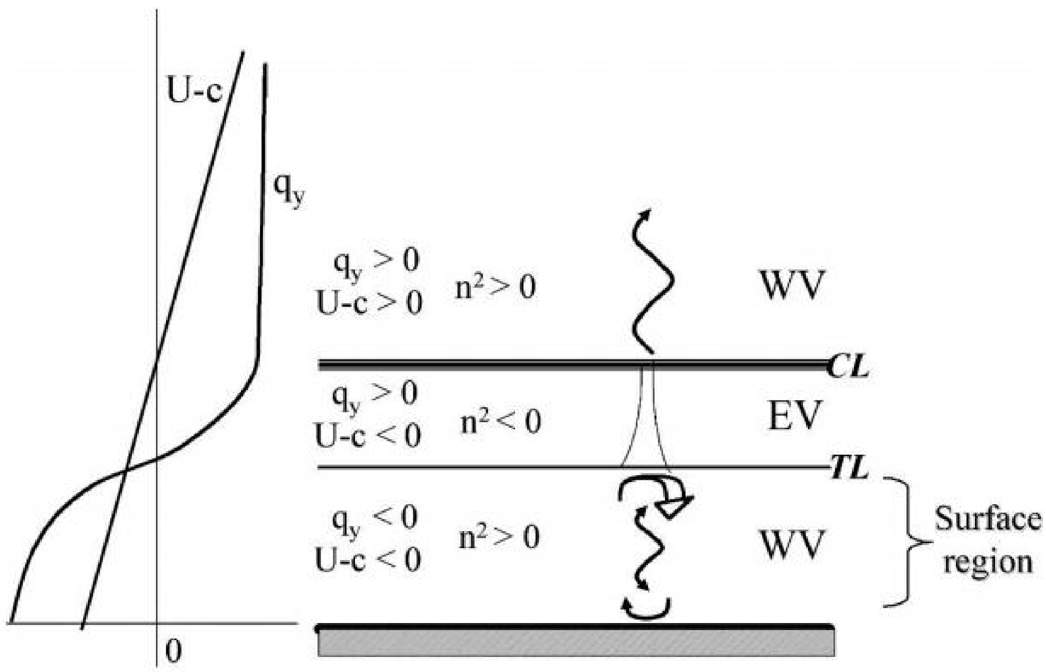
Fourthly, for the ray tracing theory, if there is no inflection level in the propagative region for a ray, the ray cannot arrive the critical level in a finite time span. In other words, the critical level is an asymptotic level for rays. In this investigation, I find that wave energy and amplitude can only have moderate increase so that instability is impossible. On the other hand, if an inflection level is in the propagative region, it must coincide with the critical level. If there is an inflection point at a location, say $y = y_{\text{inflection}}$, that is $q_y = 0$ at $y = y_{\text{inflection}}$, to make sure a positive index of refraction, that is $n^2 = \frac{q_y}{U-c} - k^2 > 0$, $U-c$ must be equal to zero at $y = y_{\text{inflection}}$ so that $\frac{q_y}{U-c}$ at $y = y_{\text{inflection}}$ can be defined. In such a case, a ray can arrive at the critical level (and the inflection level) in a finite time. Furthermore, I find that both wave energy and amplitude may have a substantial increase to induce possible instability.

To summarize, for overreflection, a region where the index of refraction is negative is necessary. However, for ray tracing, rays can never enter a region where the index of refraction is negative. I think this is the most dramatic difference between them. Actually, a ray cannot enter a region where the index of refraction is negative is equivalent to the ray cannot represent tunneling, as suggested by Harnik (2002), Harnik and Heifetz (2007).

To address your concern, I have revised the paragraph to compare with their differences through a clearer description. Now this paragraph becomes: “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 (e.g., Lindzen and Tung, 1978; Lindzen, 1988), incident waves can propagate across the evanescent region where the index of refraction is negative to realize overreflection amplification (e.g., Harnik, 2002; Harnik and Heifetz, 2007). However, ray tracing theory predicts rays can never enter such evanescent region no mater whether the inflection point exists or not. Furthermore, rays can arrive the critical point in finite time only if the critical point coincides with the inflection point. ”



Response Figure 3 Barotropic flow divided into regions according to wave propagation characteristics. This figure is cited from Lindzen (1988).



Response Figure 4 A schematic description of the wave geometry for an unstable normal mode. This figure is cited from Harnik and Heifetz (2007).

4. Sentence starting on line 168 through equation 12 - this is the central point of the underlying theory on which this paper is based. I confess that I have not fully internalized the argument of why div(Cg) is not well defined and why it can be replaced with equation 12, and it would require more time than I can spare for this. Given that this is published already, and that I can vaguely see that this could hold, I have continued the review assuming this is correct.

Response: Thank you for your insightful comment. Equation (12) and the theoretical treatment $\nabla \cdot$ Cg are fundamental for the manuscript. Here I will provide an explicit explanation.

Firstly, ray tracing method deems the wave packet as a "material particle" moving with the group velocity. As Lighthill (2001) wrote in his classic book "Waves in fluids": "waves behave like particles (moving along rays); indeed, a wave packet changes position and wavenumber which can be regarded as those for a 'particle' whose energy-momentum relationship at every position exactly parallels the frequency-wavenumber relationship (dispersion relationship) for the waves." It means we may use ray tracing method to track a wave packet from one location to another without solving the full wave equations of motion (Vallis 2017). Therefore, ray tracing is of value not only because it leads to information on the spatial distribution of the wavenumber vector but also because wave energy moves along rays (Lighthill 2001).

Secondly, let us consider the basic assumption of the slowly varying wave train which requires that a wave packet carries a wave with wavelength short enough so that ***locally*** the basic fields appears spatially uniform and ***locally*** the wave amplitude is very nearly constant over one wavelength (Pedlosky 1987). With above assumption, we can derive the ***local*** dispersion relation with has the same form with the plane wave with constant amplitude.

Since the group velocity over one wavelength is also ***locally*** spatial uniform, Pedlosky (1987) suggested that ***locally*** the divergence of group velocity vanishes so that the wave action density, E/ω' , does not change along rays, or, the wave energy density is proportional to the intrinsic frequency, that is, $E \sim \omega'$. It provides a simple method to evaluate the wave energy density. Furthermore, it also denotes that the wave energy density is ***locally*** determined by the intrinsic frequency which is only determined by the basic flow when the frequency and zonal wavenumber are fixed.

However, as Bretherton and Garrett (1968) commented, $E \sim \omega'$ is for the lowest (or zeroth) order approximation. It means the variation of amplitude, which is proportional to the small parameter ε (the corresponding first order terms), has been ignored, which is necessary to derive the zeroth order approximation equation (the local dispersion relation) and may not be enough for the first order approximation equation, which is just the wave action conservation.

Thirdly, to be accurate for the first order approximation, the ray divergence should be introduced to revise the above treatment. As pointed out previously, a wave packet along a moving ray has been regarded as a 'particle' . The spatial size of the 'particle' occupies is large compared to a wavelength but small compared to the slowly varying basic flow. **Therefore, viewed from the scale of the basic flow, a wave packet appears as a point with no spatial size** (e.g., Bretherton and Garrett 1968). **Actually, ray theory can only give the trajectory of a wave packet, not the detailed structure of the waves within the packet** (e.g., Vallis 2017).

Considering a particle (wave packet) with no spatial size moves with the group velocity, the divergence of group velocity is then only determined by the variation of the magnitude of the group velocity vector due to no lateral flux (Li et al. 2021), that is

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

where T is time and δT is a small time interval and $\left|\mathbf{c_g}\right|=\sqrt{c_{g,x}^2+c_{g,y}^2}$ denotes the magnitude of the group velocity vector. It should be stressed again that Eq. (1) is observed from the scale of the basic flow so that the size of the wave packet is ignored. With Eq. (1), the wave action conservation equation can be written as

$$\frac{\mathrm{D}_g}{\mathrm{D}T}\left(E/\omega'\right)+\left(E/\omega'\right)\frac{1}{\left|\mathbf{c_g}\right|}\frac{\mathrm{D}_g\left|\mathbf{c_g}\right|}{\mathrm{D}T}=0 \tag{2}$$

Then it is easy to derive its solution

$$E\sim\omega'/\left|\mathbf{c_g}\right| \tag{3}$$

Eq. (1) indicates that divergence of group velocity will be positive (negative) if the group velocity magnitude becomes faster and faster (slower and slower). Correspondingly, the wave action density will decrease (increase) according to Eq. (2). Physically, faster and faster (slower and slower) group velocity means the wave action density is divergent (convergent), which of course decreases (increases) the wave action density. Solution Eq. (3) means the wave energy density is proportional to the intrinsic frequency but inverse proportional to the magnitude of the group velocity.

Fourthly, if further considering the fact that a wave packet indeed occupies a certain spatial size, it seems that in principle one may directly calculate the divergence of group velocity from its expression (Bühler 2006). However, only the group velocity values on a ray are known. Neighbouring a ray, the group velocity values are unknown (Lighthill 2001). Physically, directly solving divergence of group velocity means the second order approximation should be introduced.

Finally, the solution Eq. (3) can also be derived from the multi-scale method (or WKB approximation), as emphasized by Lighthill (2001) in his seminal monograph. Under the zeroth-order approximation — valid at the slowly varying scale — the derived ***local*** dispersion relation mirrors the form of the plane wave solution. This equivalence implies that the wave is ***locally*** approximated as a plane wave with constant amplitude and wavenumbers at the slowly varying scale, leading to local conservation of wave action. Consequently, the partial derivative of wave action density with time vanishes, that is, wave action conservation equation

$$\frac{\partial}{\partial T}\left(\frac{E}{\omega'}\right)+\nabla\cdot\left(\mathbf{c_g}\frac{E}{\omega'}\right)=0 \tag{4}$$

can be simplified to

$$\nabla\cdot\left(\mathbf{c_g}\frac{E}{\omega'}\right)=0\,, \tag{5}$$

which means $\mathbf{c_g}\, E/\omega'$ is a solenoidal vector (Lighthill 2001). It further demonstrates that $E/\omega' \left|\mathbf{c_g}\right| = \text{constant}$ along a ray in terms of the cross-sectional area of a thin ray tube.

To sum up, the theoretical treatment ∇ • Cg has a solid mathematical and physical basis.

5. The derivation of equation 15 took some time, and the way it is phrased suggests only equation 11 is used, while actually also equation 12 and A=E/w' were used. I suggest at the very least stating this explicitly, but adding an appendix derivation will help.

Response: Thank you for your valuable feedback on the derivation of Equation (15). Here I will present the explicit derivation. The wave action conservation Eq. (4) in the above response can be expanded as

$$\frac{\partial}{\partial T}\left(\frac{E}{\omega'}\right)+\mathbf{c_g}\cdot\nabla\left(\frac{E}{\omega'}\right)+\left(\frac{E}{\omega'}\right)\nabla\cdot\mathbf{c_g}=0 \tag{6}$$

The first two terms are the individual derivative along rays. Therefore, Eq. (6) can be written as

$$\frac{\mathrm{D}_g}{\mathrm{D}T}\left(\frac{E}{\omega'}\right)+\left(\frac{E}{\omega'}\right)\nabla\cdot\mathbf{c_g}=0 \tag{7}$$

Note that

$$\begin{aligned}\frac{\mathrm{D}_g}{\mathrm{D}T}\left(\frac{E}{\omega'}\right)&=\frac{1}{\omega'}\frac{\mathrm{D}_g E}{\mathrm{D}T}+E\frac{\mathrm{D}_g}{\mathrm{D}T}\frac{1}{\omega'}\\&=\frac{1}{\omega'}\frac{\mathrm{D}_g E}{\mathrm{D}T}-\frac{E}{\omega'^2}\frac{\mathrm{D}_g\omega'}{\mathrm{D}T}\\&=\frac{1}{\omega'}\frac{\mathrm{D}_g E}{\mathrm{D}T}-\frac{E}{\omega'^2}\frac{\mathrm{D}_g}{\mathrm{D}T}\left(\omega-\bar{u}k\right)\end{aligned} \tag{8}$$

Furthermore, the frequency and zonal wavenumber are unchanged along rays. Therefore, Eq. (8) can be further expressed as

$$\begin{aligned}\frac{\mathrm{D}_g}{\mathrm{D}T}\left(\frac{E}{\omega'}\right)&=\frac{1}{\omega'}\frac{\mathrm{D}_g E}{\mathrm{D}T}+\frac{E}{\omega'^2}k\frac{\mathrm{D}_g\bar{u}}{\mathrm{D}T}\\&=\frac{1}{\omega'}\frac{\mathrm{D}_g E}{\mathrm{D}T}+\frac{E}{\omega'^2}k\left(\frac{\partial\bar{u}}{\partial T}+c_{g,x}\frac{\partial\bar{u}}{\partial X}+c_{g,y}\frac{\partial\bar{u}}{\partial Y}\right)\\&=\frac{1}{\omega'}\frac{\mathrm{D}_g E}{\mathrm{D}T}+\frac{E}{\omega'^2}kc_{g,y}\frac{\mathrm{d}\bar{u}}{\mathrm{d}Y}\\&=\frac{1}{\omega'}\frac{\mathrm{D}_g E}{\mathrm{D}T}-\frac{E}{\omega'}\frac{K^2}{\beta^* k}k\frac{2\beta^* kl}{K^4}\frac{\mathrm{d}\bar{u}}{\mathrm{d}Y}\\&=\frac{1}{\omega'}\frac{\mathrm{D}_g E}{\mathrm{D}T}-\frac{E}{\omega'}\frac{2kl}{K^2}\frac{\mathrm{d}\bar{u}}{\mathrm{d}Y}\end{aligned} \tag{9}$$

Note that we have applied the expression the meridional group velocity and intrinsic frequency, that is,

$$c_{g,y} = \frac{2\beta^* kl}{K^4}$$
$$\omega' = \omega - \bar{u}k = -\frac{\beta^* k}{K^2} \quad (10)$$

Substituting Eq. (9) into Eq. (7), we have

$$\frac{1}{\omega'}\frac{\mathrm{D}_g E}{\mathrm{D}T} - \frac{E}{\omega'}\frac{2kl}{K^2}\frac{\mathrm{d}\bar{u}}{\mathrm{d}Y} + \frac{E}{\omega'}\nabla \cdot \mathbf{c_g} = 0 \quad (11)$$

Simplifying it, we can derive Equation (15) in the submitted manuscript, that is

$$\frac{1}{E}\frac{\mathrm{D}_g E}{\mathrm{D}T} = \frac{2kl}{K^2}\frac{\mathrm{d}\bar{u}}{\mathrm{d}Y} - \nabla \cdot \mathbf{c_g} \quad (12)$$

In the revised manuscript, I have explicitly presented the above derivation in a new Appendix. Thank you for helping us improve the rigor and clarity of this section.

6. The explanation on lines 197-202 is a bit cumbersome. I think it is much simpler to state that the anomaly extract energy from the mean flow when the anomaly is tilted against the shear and it return energy to the flow when it is tilted with the shear. Similarly the sentence in parenthesis on line 216-217 can be changed to "mediated by structures tilting against the shear". Also the next sentence- "conversely, lambda<0 indicates a tilting with the shear... "

Response: Thank you for your constructive feedback. I have revised the relevant contents as you suggested.

7. Section 3.1: I don't understand the sentence on lines 231-232.

Response: I am sorry for the implicit expression here. In Section 3.1, I assume a simple case where an ED regime (enclosed by a critical latitude and a northern turning latitude in this case) is located south of the jet center. Therefore, the wind shear is always positive, that is $\frac{dU}{dy} > 0$.

Furthermore, I assume that the meridional gradient of the absolute vorticity is always larger than zero in the ED regime, that is, $\beta^* > 0$. I also set a ray starting from an initial location, say y_0, moves northward. According to the definition of the meridional group velocity, that is, $c_{g,y} = \frac{2\beta^* kl}{K^4}$, northward moving ray requires $kl > 0$, the leading structure or tilting against the shear.

Now, if the ray is for a stationary wave (the frequency is zero and hence zonal phase speed is zero, that is, $c = 0$), we can derive that

$$\xi = \frac{c^2 - \frac{2k^2}{K^2}(\bar{u}-c)^2}{c^2 + \frac{4k^2}{K^2}(\bar{u}-c)\bar{u}}\frac{2kl}{K^2}\frac{\mathrm{d}\bar{u}}{\mathrm{d}Y} = -\frac{kl}{K^2}\frac{\mathrm{d}\bar{u}}{\mathrm{d}Y} < 0 \quad (13)$$

It means the change rate of the wave energy will be always smaller than until the ray arrives at the turning latitude where zero meridional wavenumber causes zero change rate of the wave energy. Therefore, the wave energy of stationary waves will decrease to a minimum value when it moves northward to arrive at the turning latitude.

To avoid confusing and to following your below comment, I have rewritten Section 3.1 to make sure it is easier to follow. Thank you for helping me improving the quality of the manuscript.

8. Please add the derivation of conditions 23 and 25 to an appendix (and remove "It is easy to derive" - on lines 235 and 241). Please add a schematic illustrating the different cases discussed in page 10 - a schematic of the flow profile, the turning surface and what the different phase peed ranges imply about the value of U-c, and the implications for the wave geometry would be helpful. Line 248 and the discussion of an upper limit of group velocity is very confusing. Is c_max in equation 26 a phase speed or group speed? The only way to understand this is as follows: You state that it is there "to make sure the ray can arrive at the northern turning point". This as far as I understand requires a real meridional wavenumber l, or a positive l^2 (I can't think of any other way by which the phase speed affect the ability of the ray to reach the turning latitude except for the condition that l^2>0 south of this latitude.) Assuming l^2>0 south of the turning latitude, you then mean to say that the maximum zonal group speed c_gx value is obtained when l^2=0 but then the group speed would be U+beta*/k^2 and not U-beta*/k^2. The phase speed for this case will actually be as written in equation 26 but what insures that this is the maximal phase speed? without knowing how beta* varies in space (assuming dU/DY>0 as stated at the beginning of this subsection). Moreover, when I try to think about any conditions on the phase speed that will insure that the meridional wavenumber is indeed real south of the turning latitude I only come up with an opposite argument: If we compare the meridional wavenumber squared for two phase speeds, c1 > c2, then l^2 will be larger for c1, as long as c<u, which implies a lower limit for c in order to insure a positive l^2, not an upper limit. Please explain this argument more clearly, and again, supporting it by a schematic of a concrete flow configuration will help.

Response: Thank you for your valuable comments. I apologize for the lack of clarity in Section 3.1, which may have caused confusion. To address your comments, I have comprehensively revised Section 3.1 to enhance logical flow and readability. Additionally, a schematic diagram has been added to further facilitate understanding of the key concepts.

To analyze wave energy evolution, the wave energy equation along rays is quite fundamental. Therefore, I also express it here as (see Eq. (15) and Eq. (20) in the manuscript)

$$
\begin{aligned}
\frac{1}{E}\frac{\mathrm{D}_g E}{\mathrm{D}T} &= \left(\frac{2kl}{K^2}\frac{\mathrm{d}\bar{u}}{\mathrm{d}Y} - \frac{\mathrm{D}_g \ln\left|\mathbf{c}_{\mathbf{g}}\right|}{\mathrm{D}T}\right) \\
&\equiv \xi(T) \\
&= \frac{c^2 - \dfrac{2k^2}{K^2}(\bar{u}-c)^2}{c^2 + \dfrac{4k^2}{K^2}(\bar{u}-c)\bar{u}} \frac{2kl}{K^2}\frac{\mathrm{d}\bar{u}}{\mathrm{d}Y} \\
&\equiv \lambda(c,\bar{u})\frac{2kl}{K^2}\frac{\mathrm{d}\bar{u}}{\mathrm{d}Y}
\end{aligned}
\tag{14}
$$

As shown in the equation, the sign of the rate of change in wave energy depends on three terms, the coefficient λ, the wave structure kl, and the wind shear $\frac{\mathrm{d}\bar{u}}{\mathrm{d}Y}$, the latter two of which are relatively easy to identify.

In section 3.1, I present a case where a ray for a wave with a fixed zonal phase speed (labeled as c_0 or c as in the revised manuscript) propagates from its initial location to a northern turning latitude in an ED regime that is located south of the westerly jet (see Response Figure 5a, b). These conditions means that $kl > 0$ and $\frac{\mathrm{d}\bar{u}}{\mathrm{d}Y} > 0$. Therefore, the increase or decrease in wave energy is only determined by the sign of λ. Note that this case is the simplest. As I have responded in your first comment, this setting can also be found in observed westerly jets. Furthermore, based on this simple case, more complex cases (e.g., the turning latitude is located north of the jet center) can also be analyzed.

According to the expression of λ, its sign is determined by its numerator. Setting the numerator equals to zero, that is

$$c^2 - \frac{2k^2}{K^2}(\bar{u} - c)^2 = 0 \tag{15}$$

it is easy to solve its zero point (labeled as $c_{\lambda 0}$)

$$c_{\lambda 0}(y) = \frac{\sqrt{2}k}{K(y) + \sqrt{2}k}\bar{u}(y) \tag{16}$$

Then we know when $c_0 < c_{\lambda 0}$, $\lambda < 0$, wave energy will decrease and vice versa. On one hand, the zonal phase speed c_0 keeps unchanged along rays. On the other hand, the zero point $c_{\lambda 0}$ varies along rays since both westerly and total wavenumber varies along rays. Therefore, to compare their relative size, we should identify the variation feature of $c_{\lambda 0}$ at first.

According to our prescribed conditions, when the ray moves from its initial location (y_0) to the northern turning latitude (y_t), the wind speed is monotonically increasing while the total wavenumber is monotonically decreasing (due to decreasing meridional wavenumber). We can know that $c_{\lambda 0}$ monotonically increases when the ray moves from y_0 to y_t (see Response Figure 5c), that is

$$c_a \equiv \frac{\sqrt{2}k}{K(y_0) + \sqrt{2}k}\bar{u}(y_0) = c_{\lambda 0}(y_0) \le c_{\lambda 0}(y) \le c_{\lambda 0}(y_t) = \frac{\sqrt{2}k}{K(y_t) + \sqrt{2}k}\bar{u}(y_t) \equiv c_b$$

(17)

It means the values of $c_{\lambda 0}$ at y_0 and at y_t are its minimum and maximum. For simplicity, we have labeled the minimum and maximum as c_a and c_b. Then it would be easy to compare the relative size of c_0 and $c_{\lambda 0}$. (1) If $c_0 < c_a$, that is, the zonal phase speed is slower than the minimum of $c_{\lambda 0}$, the zonal phase speed will be smaller than all values of $c_{\lambda 0}$ in the range (y_0, y_t). Therefore, $\lambda < 0$ and hence $\frac{\mathrm{D}_g E}{\mathrm{D}T} < 0$ in (y_0, y_t). Correspondingly, wave energy will monotonically decrease from its initial value (say, equal 1) to a minimum at y_t where $\lambda = 0$ due to $l = 0$. (2) If $c_0 > c_b$, the zonal phase speed will be larger than $c_{\lambda 0}(y)$, leading $\lambda > 0$ and hence $\frac{\mathrm{D}_g E}{\mathrm{D}T} > 0$ in (y_0, y_t). Correspondingly, wave energy increases to a maximum at y_t where $\lambda = 0$ due to $l = 0$. (3) If $c_a < c_0 < c_b$, there will exist an intermediate location (say, $y_0 < y_m < y_t$) where $c_0 = c_{\lambda 0}(y_m)$. Then when the ray moves from y_0 to y_m, $\lambda > 0$; when the ray arrives at y_m, $\lambda = 0$; and when the ray continues to move from y_m to y_t, $\lambda < 0$. Correspondingly, the wave energy will increase to a maximum at y_m and then decrease to a minimum at y_t. Above features are illustrated in Response Figure 5d, e, f.

When the ray arrives at the turning latitude, the dispersion relation becomes

$$c_{yt} = \bar{u}(y_t) - \frac{\beta^*(y_t)}{k^2} \tag{18}$$

Note that c_{yt} prescribes the upper limit of the zonal phase speed of a wave whose ray can propagate in an ED regime (I used $c_{\max}$ to describe this upper limit in the submitted manuscript, which may be a little confusing. I have changed it to c_{yt} in the revised manuscript). If $c_0 > c_{yt}$, the ray will not be reflected by any turning point along its trajectory. Therefore the zonal phase speed range for a ray can propagate in an ED regime is $(0, c_{yt})$. Here qualitative analysis cannot provide the relative size of c_{yt} and c_a and c_b. Without loss of generality, we may assume that $c_{yt} > c_b$. Then the zonal phase speed range $\left(0, c_{yt}\right)$ can be divided into three parts: $\left(0, c_a\right)$, $\left(c_a, c_b\right)$, and $\left(c_b, c_{yt}\right)$. It is obvious that wave energy can have the most significant increase in the last part.

Up to now, we have identified the evolution feature of wave energy. When a westerly jet is given, and when the initial zonal and meridional wavenumber are given, we may change frequency (of course, we can not arbitrarily specify the frequency value. It must satisfy the dispersion relation) or zonal phase speed to calculate the energy evolution of rays with different zonal phase speed (e.g., see Figure 3 in the submitted manuscript). Note that here the zonal phase speed is not the propagation speed of the ray, but the phase speed of the wave.

Finally, let's consider a slightly complex case where the turning point lies north of the jet center, with other conditions remaining unchanged. Without loss of generality, we analyze the scenario where the zonal phase speed falls within the interval $\left(c_a, c_b\right)$. As the ray propagates from y_0 to y_m, wave energy increases to a maximum at y_m. When the ray travels from y_m to the jet center (say, y_z), wave energy declines from the maximum to a minimum at y_z where zero wind shear can also contribute to this minimum. As the ray proceeds from y_z to y_t, wave energy rebounds from the minimum to another maximum due to both negative wind shear and λ. Compared with the previous case, the most notable distinction is that wave energy can attain two distinct maxima as the ray propagates from the initial location to the turning point. Since analysis methodology remains consistent, we do not explicitly analyze this case.

Based on above content, I have totally rewritten Section 3.1. Since the derivation is clearer this time, I do not add them to an appendix. Thank you for helping me strengthen the presentation of this work.

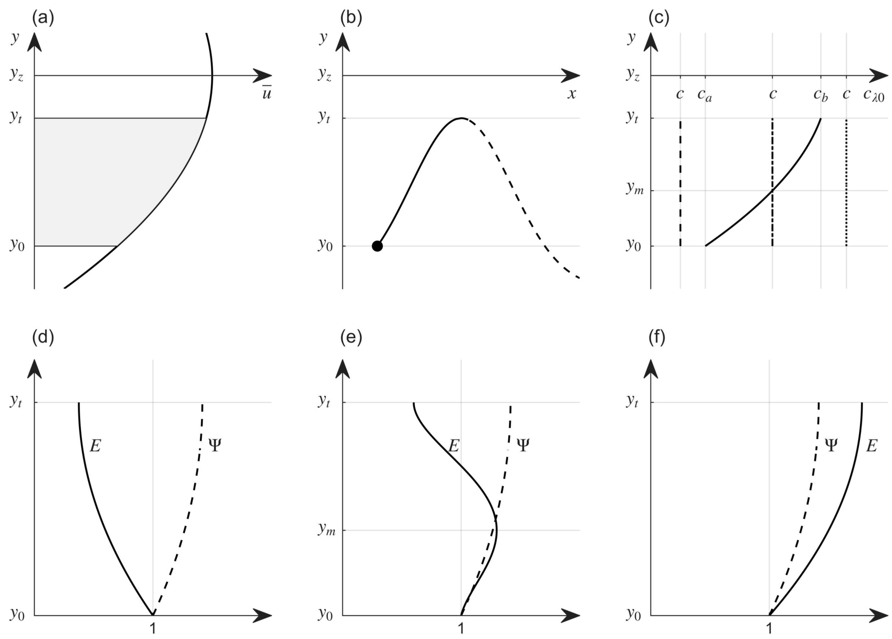
Belows are some concepts to help you understanding.

(1). When calculating ray trajectory, we should know the initial wavenumber (e.g., k0, l0), initial location (e.g., x0, y0), and initial frequency (e.g., w0). Since zonal wavenumber and frequency keep unchanged along a ray, the phase speed (c0=w0/k0) is also kept unchanged along a ray. However, the group speed varies along a ray. In above discussion, we have fixed k0, l0, x0, and westerly jet. Then we discuss the scenarios where zonal phase speed (or frequency) varies. When we give a specific value of zonal phase speed, we can calculate a specific meridional initial location (y0) according to the dispersion relation.

(2). A ray will be reflected to move toward another direction when it arrives at a location where its meridional wavenumber equals to zero. This location is called the turning latitude. A ray will move asymptotically toward the location where the meridional wavenumber square is infinity. This location is called a critical latitude. Then, there will be three types of propagative regions: (a) enclosed by two critical latitudes; (b) enclosed by two turning latitudes; and (c) enclosed by a critical latitude and a turning latitude. As we had introduced in section 2.1, these three types of propagative regions are named the bidirectional dissipation (BD) regime; the wave guide (WG) regime; and the evolution-dispersion (ED) regime, respectively. According to our previous investigations (e.g., Li et al. 2021), the wave energy can have a substantial increase in the ED regime. Therefore, We mainly focus on analyzing the theoretical wave energy evolution in the ED region in Section 3.

(3). The ED regime (see Response Figure 2) has an upper zonal phase speed limit. Saying that, we have meant the zonal wavenumber, the basic flow are both fixed. Then only the zonal phase speed (or frequency) can vary in the dispersion relation (note that meridional wavenumber equals zero at turning latitude and tends to be infinity at the critical latitude). That's the reason why I use $c_{\max}$ in the submitted manuscript (I have revised it to c_{yt} in the revised manuscript).

One of the most interesting finding in this manuscript is that we find c_{yt} corresponds with most substantial increase in both wave energy and amplitude. That's also the reason I use the notation $c_{\max}$.



Response Figure 5 Schematic illustration of an ED regime (a) that is located south of a westerly jet; a ray propagates northward from an initial location denoted by the black point (b); the distribution of the zero point of λ (c); and the cases of $c < c_a$ (d); $c_a < c < c_b$ (e); and $c > c_b$ (f), exhibited by the dashed, dash-dotted, and dotted vertical line in (c). The notation y_0, y_t and y_z denote the initial location, the turning location, and the jet center, respectively. y_m indicates a intermediate location where $c = c_{\lambda 0}(y_m)$. E and Ψ denote wave energy and amplitude, respectively.

9. Sentence on lines 256-7: I am not sure I follow. Why does dE/Dt have to be zero at the turning latitude? you said it vanishes at y_m..

Response: According to Equation (14) in this response, the variation of wave energy density along rays is (I copy it here)

$$\frac{1}{E}\frac{\mathrm{D}_g E}{\mathrm{D}T}=\left(\frac{2kl}{K^2}\frac{\mathrm{d}\bar{u}}{\mathrm{d}Y}-\frac{\mathrm{D}_g \ln|\mathbf{c}_g|}{\mathrm{D}T}\right)\equiv\xi(T)=\lambda(c,\bar{u})\frac{2kl}{K^2}\frac{\mathrm{d}\bar{u}}{\mathrm{d}Y} \quad (19)$$

At the turning latitude ($y=y_t$), the meridional wavenumber vanishes ($l=0$) so that the wave energy density has an extreme value at the turning latitude. In some cases, this is the only one extreme value for the wave energy. In some other cases ($c_a<c<c_b$ as analyzed above), the wave energy may have other extreme values. Wave energy reaches an extreme value at $y=y_m<y_t$ due to $\lambda=0$, then we have

$$\left.\frac{1}{E}\frac{\mathrm{D}_g E}{\mathrm{D}T}\right|_{y=y_m}=0 \quad (20)$$

Since $E=\frac{1}{4}\Psi^2K^2$ where Ψ and K are amplitude and total wavenumber, we can derive that

$$\left.\frac{\mathrm{D}_g E}{\mathrm{D}T}\right|_{y=y_m}=\left.\frac{1}{4}\frac{\mathrm{D}_g K^2}{\mathrm{D}T}\Psi_0^2\right|_{y=y_m}+\left.\frac{1}{4}\frac{\mathrm{D}_g \Psi_0^2}{\mathrm{D}T}K^2\right|_{y=y_m}=0 \quad (21)$$

Above Eq. (21) is just Eq. (27) in the submitted manuscript. At $y=y_m<y_t$, the derivative of total wavenumber (the first term at the right hand) is not equal to zero. Correspondingly, the derivative of amplitude (the second term at the right hand) is also not equal to zero. Therefore, amplitude does not approach extreme value at $y=y_m<y_t$ even if wave energy approaches. That's what I want to express there in the submitted manuscript.

I am sorry for the improper expressing that confuses you. Following your comments, I have totally rewritten Section 3.1. Thanks for your constructive comment.

10. Discussion from line 260 to end of section: It will help to have schematic of the wave geometry of the different phase speed cases, drawn on the mean flow (U and beta*). I was not able to follow the point about the fast phase speed range- I have a feeling I am missing something basic in the way you view the problem. Specifically the discussion of the critical latitude starting on line 269: unless beta*=0 this line separates wave propagation from evanescence. Also, while the EP fluxes suggests wave activity emanates from the critical line for unstable waves, why do the rays that emanate from it are most important for the instability to be able to exist? Instability should occur no matter where you place a wave source, no?

Response: I am sorry for not explicitly introducing my investigation. Your opinion involves two aspects, the former of which I have already provided a detailed response. For the latter, I firstly show you the case where β *>0. In such a case, there is no normal mode instability since the necessary condition is not satisfied. Now let's focus on the critical latitude where the meridional wavenumber is infinity. It is an asymptotic location for a ray. Therefore, a ray can never arrive such an asymptotic location in a finite time. Secondly, if β *=0 at y=yi, the dispersion relation

$$c = \overline{u} - \frac{\beta^*}{K^2} \tag{22}$$

reduces to

$$c = \overline{u} \tag{23}$$

Meanwhile, at the critical latitude (where meridional wavenumber is infinity), the dispersion relation also have the same form as Equation (23) shows (it is caused by infinity K2). Therefore, if an inflection latitude consists of a boundary of a ray, it must also be the critical latitude. At the inflection latitude, there is no definition for the meridional wavenumber. However, it does not mean meridional wavenumber can be taken arbitrary. To show this, we may firstly write the dispersion relation as

$$K^2 = \frac{\beta^*}{\overline{u} - c} \tag{24}$$

Then, it can be expressed as

$$K^2 = \frac{\beta^*}{\overline{u} - c} = \frac{\frac{\mathrm{d}\beta^*}{\mathrm{d}y}}{\frac{\mathrm{d}\overline{u}}{\mathrm{d}y}} = \frac{-\frac{\mathrm{d}^3\overline{u}}{\mathrm{d}y^3}}{\frac{\mathrm{d}\overline{u}}{\mathrm{d}y}} \equiv K_i^2 < \infty \tag{25}$$

at the inflection latitude. As shown in Equation (25), the total wavenumber can also be defined at the inflection latitude so does the meridional wavenumber (labeled as $l_i = \sqrt{K_i^2 - k^2}$). Therefore, a ray with an initial zonal and meridional wavenumber can arrive at the inflection latitude (also the critical latitude) where the meridional group velocity vanishes ($c_{g,y} = \frac{2\beta^* kl}{K^4} = 0$ due to $\beta^* = 0$). Correspondingly, the ray will move horizontally along the inflection latitude with zonal group velocity equal to the zonal phase speed ($c_{g,x} = c + \frac{2\beta^* k^2}{K^4} = c$ since $\beta^* = 0$). Above analysis means that when a ray arrives at the inflection latitude, it will move along the latitude so that it can never enter the region where $\beta^* < 0$. Similarly, if a ray propagates in the region where $\beta^* < 0$ (note that the zonal phase speed must be faster than the basic flow to make sure a positive index of refraction), when it arrives at the inflection latitude, it will also move along the latitude so that it can never enter the region where $\beta^* > 0$.

When the ray moves along the inflection latitude, energy equation (19) in this response becomes

$$\frac{1}{E}\frac{\mathrm{D}_g E}{\mathrm{D}T} = \left(\frac{2kl}{K^2}\frac{\mathrm{d}\overline{u}}{\mathrm{d}Y} - \frac{\mathrm{D}_g \ln\left|\mathbf{c}_\mathbf{g}\right|}{\mathrm{D}T}\right) = \frac{2kl}{K^2}\frac{\mathrm{d}\overline{u}}{\mathrm{d}Y} \tag{26}$$

Since the group velocity will keep constant. If the inflection latitude is south of the jet center, we have $\frac{\mathrm{d}\overline{u}}{\mathrm{d}Y} > 0$. Then if the meridional wavenumber is larger than zero, the right hand term will be lager than zero. If the meridional wavenumber equals l_i, the right hand term will become a constant positive value so that wave energy will increase exponentially without any limitation. This means normal mode instability.

According to Eq. (7) in the manuscript, the variation of meridional wavenumber is determined by

$$\frac{\mathrm{D}_g l}{\mathrm{D}T} \equiv -\frac{\partial \Omega}{\partial Y} = -\frac{\mathrm{d}\overline{u}}{\mathrm{d}Y}k + \frac{\mathrm{d}\beta^*}{\mathrm{d}Y}\frac{k}{K^2} = -\frac{\mathrm{d}\overline{u}}{\mathrm{d}Y}k - \frac{\mathrm{d}^3\overline{u}}{\mathrm{d}Y^3}\frac{k}{K^2} \tag{27}$$

For slowly varying basic flow, we may approximately write it as

$$\frac{\mathrm{D}_g l}{\mathrm{D}T} \approx -\frac{\mathrm{d}\overline{u}}{\mathrm{d}Y}k \tag{28}$$

It mean meridional wavenumber will monotonically decrease (we focus on discussion south of the jet center).

If a ray moves from an initial location just at the inflection latitude, the ray will always move along the inflection latitude (the trajectory is a horizontal line). Then if its initial meridional wavenumber is larger, $l_0 > l_i$, the meridional wavenumber will decrease until it decrease to equal to l_i. Then the meridional wavenumber will keep unchanged along the ray. Correspondingly, the wave energy will exponentially increase. If the initial meridional wavenumber is smaller, $l_0 < l_i$, then the meridional wavenumber will keep decreasing (since it cannot meet the critical value l_i). In such case, wave energy will eventually exponentially decrease. Therefore, whether a ray initially located at the inflection latitude can develop or not depends on the its initial wave structure. This is consistent with the critical layer problem in the classic instability theory (e.g., Pedlosky, 1987). Now I can address your concern. The ray starts to its journey from the inflection latitude is important because it is the only possibility that normal mode instability can occur in the ray tracing theory. Correspondingly, instability does not depends in the wave source where a ray starts to move. It depends on whether the wave source is on the inflection latitude or not (also the initial meridional wavenumber at the inflection latitude).

Minor:

Line 35 - the absolute vorticity (not its gradient) has an extremum, the gradient changes sign, no?

Response: Thanks for your careful reviewing. Here the gradient is redundant and should be removed. I have removed it in the revised manuscript and carefully checked the manuscript to avoid similar mistakes.

line 59 fix typo (w..)

Response: Thanks for your careful reviewing. I am sorry for the typo here. I have revised the sentence in the revised manuscript.

line 104- insert the definition of the material derivative

Response: Thanks for your careful reviewing. I have stress the definition of the material derivative along rays Line 124 in the revised manuscript.

Line 248 it should be "phase speed" not "group velocity" at the end of the line, no?

Response: Thanks for your careful reviewing. I am sorry for the typos here. I should be zonal phase speed rather than group velocity. Following your previous comments, I have rewritten Section 3.1.

line 318: it is easy to *show*

Response: Thanks for your careful reviewing. I have added the word “show” there. I have also carefully checked the manuscript to avoid similar typos.

References

Bretherton, F. P., and C. J. R. Garrett, 1968: 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.

Bühler, O., 2006: *A Brief Introduction to Classical, Statistical, and Quantum Mechanics*. American Mathematical Society, https://doi.org/10.1090/cln/013.

Harnik, N., 2002: The Evolution of a Stratospheric Wave Packet. *Journal of the Atmospheric Sciences*, **59**, 202–217, https://doi.org/10.1175/1520-0469(2002)059%3C0202:TEOASW%3E2.0.CO;2.

Harnik, N., and E. Heifetz, 2007: Relating Overreflection and Wave Geometry to the Counterpropagating Rossby Wave Perspective: Toward a Deeper Mechanistic Understanding of Shear Instability. *Journal of the Atmospheric Sciences*, **64**, 2238–2261, https://doi.org/10.1175/JAS3944.1.

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

Li, Y., J. Chao, and Y. Kang, 2021: 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.

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

Lindzen, R. S., 1988: 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.

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

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