

**I would like to thank the referees for the fruitful comments.**

**Referee 2**

It seems the title for the vertical axis in Figs 2 (up) and 4 as "amplitude function" should be better, because of "amplitude" is the maximal point on the wave shape, and you use this term "amplitude" for  $A_0$  in text.

**Thanks, agree, changed to "Amplitude function" where needed (Figs 2 upper, 4, 6, and 7).**

**Referee 1**

I worked quite hard to try to make this happen with the parameters the author provides (and is to be commended for taking care to ensure this is the case), and was not successful. It is possible that the free surface which the author accounts for in his asymptotic procedure, but which is absent in the formulation of the DJL equation I use, is essential for this, but I cannot see a physical reason why.

**There is the physical reason which is discussed in any classical course on internal waves. The reason is that the boundary conditions at the free surface add additional nonlinearity. That is why the numerical models for waves propagating between rigid lids on which you are working do not repeat my results.**

The manuscript, in my opinion, starts to drift when the topic of trapped cores is brought up. Trapped cores either invalidate the assumptions used to derive the DJL equation, or are a dynamical feature that naturally evolves, for example when a wave shoals. They have been extensively discussed in the literature, and none of this discussion appears in either the text or the references. I understand the author wants to present his ideas about wave stability and hence is not obliged to fully review the literature, but some commentary would help the reader orient the present study.

**I believe that any additional discussion about transient effects and shoaling is beyond the scope of the present study. The study on the waves with vortex core has been cited in the paper.**

I thus hope the author will adopt a subset of the suggestions below, and after this I think the manuscript can be published.

**Thank you; please find my answers to the minor suggestions below.**

1) The title seems excessively general. At the very least "in a stratified fluid" should be added.

**The article refers not only to stratified flows. Similar result is valid for Rossby waves and inertial waves in swirling flows so the title reflects this point. Appropriate reference has been added.**

2) “then” refers to a comparison in time as in “I ate lunch and then I ate supper”. “than” is the correct word when the author states “the family of solutions is richer than two-humped structures”. A similar error occurs at other points in the article.

**Sorry for that typo. Corrected.**

3) The description of what we did in Dunphy et al isn't quite right. It might make sense to describe Lamb and Wan's work first since their result is what allowed us to construct the multi-scaled solitary wave solution. The part about nearly identical profiles isn't really relevant to that aspect of the study. I suppose the author feels it is important to mention since he argues that very small differences in density profiles can make the difference in whether multiscaled waves do or do not exist in his formulation. In the two pycnocline example from Lamb and Wan we were following, small differences in stratification made no difference to the calculation.

**I implemented the suggestion given in your initial comment given below  
“I mention above is the same code that we reported in Dunphy et al 2011, and the main point there was not that multi-scale solitary waves actually occur in nature, but that the spectral methods we implemented allow for even something as finely balanced as one of these waves to be computed in minutes”**

**The major point is that Dunphy et al 2011 presented a numerical code able to deal with fine details of stratification. That is why I referred this paper first. I dealt with situation without 2 pycnoclines but with the fine structure of stratification as discussed in the revised paper. I am hoping it explains my presentation.**

4) On page 3 the discussion of the Weierstrass approximation theorem has been expanded but I still think it's a bit unclear. At the very least the author should state that  $f(z)$  is now taken to be a polynomial (may be before equation (11)). It would be helpful to tell the reader whether going from the general expression (11) to the specific result (17) is algorithmic, or whether it just kind of worked out for this choice of  $f(z)$ .

**It is clearly explained in the paper that I presented all formulae for the wave characteristics based on the polynomial profile of stratification.**

I also think the section heading “Multiscaling” occurs in an odd place, since the first sentence ties in very nicely to the last paragraph of the previous section.

**The last paragraph before the “Multiscaling” section presents the general result. All material presented in “Multiscaling” related to the specific cases when multiscaling**

**was found. The word “paragraph” in the referee review contains a typo. We are all making such mistakes.**

Finally, it would be very useful to have a table with  $f(z)$ ,  $P_N(z)$  for the various special cases discussed, possible with a column for relevant figures.

**Tables are necessary if no formula can be presented. I presented all assumptions and formulae necessary to calculate the wave characteristics.**

5) Is phase velocity the correct term? I am not aware of a group velocity for solitary waves, so wouldn't propagation speed be easier to understand?

**Phase velocity is a classical definition discussed in any course on nonlinear waves so the term should be used as is.**

6) Both the figure captions, and the discussion of the figures in the text are very brief. Figure 1 is very useful. I would tighten the axes to show how special the region required really is, and I would add a vertical line at the value of  $\alpha$  used on page 4. Then I would describe this in the text.

**I believe that a brief discussion is in line with the length of this short paper. The referee agreed that the length of the paper is appropriate and ticked the relevant line in the assessment of the paper.**

7) On page 7 the author states that waves “evolve”. This is misleading, since there is no temporal evolution, merely a tracing of the form that the wave takes in parameter space.

**The word “evolve“ is replaced by “change”.**

English: “In a Russian journal”, “dissimilated” should be changed to “disseminated” or some similar word, “Assumption of small, albeit finite”, “a priori” not “a priory”, “Multisclaed” not “Multscaled”, “fourth order” not “forth order”

**Corrected. The word “Multisclaed” in the referee review contains a typo. As I mentioned before such small mistakes are done by everybody. The expression “small, albeit finite amplitude” was used, for example, in the book “IUTAM Symposium on Laminar-Turbulent Transition and Finite Amplitude Solutions edited by Tom Mullin, R. R. Kerswell.”**

**Finally, the list of corrected typos is**

- 1) “than” line 20, p.1 and line 11 p.2.
- 2) “change” line 6, p.7.
- 3) “disseminated” line 21, p.1.
- 4) “a priori” line 15, p.2.
- 5) “fourth order” line 11,p.9.

