Methods and Applications of Analysis.
Vol. 28, No. 3, pp. 387–400, September 2021

CONVERGENCE RATES OF THE SINGULAR LIMIT FOR EQUATORIAL SHALLOW-WATER DYNAMICS*

SONG JIANG†, QIANGCHANG JU‡, AND XIN XU§

Dedicated to Professor Ling Hsiao on the occasion of her 80th birthday

Abstract. We study the singular limit of the equatorial shallow-water system which describes the motion of the atmosphere/ocean in the equatorial zone. Based on the convergence result of Dutrifoy, Majda and Schochet [Comm. Pure Appl. Math., 2009, 62(3): 322-333], we further obtain the convergence rate estimates of the solutions in the case of well-prepared initial data.

Key words. Equatorial shallow-water system, Singular limit, Convergence rates.

Mathematics Subject Classification. 35B25, 35L45, 35Q35.

1. Introduction. The study of singular limit problems for atmosphere and ocean models is of great importance to climate science. For flows in mid-latitudes, the variations of the Coriolis force due to the curvature of the Earth are usually neglected. Mathematically, this leads to a singular limit problem with constant coefficients [9, 10, 14] and it has been justified with great generality [1, 5, 6, 11, 12, 15]. However, for flows in the equatorial zone, one has to take into account the variations of the Coriolis force since the Coriolis force totally degenerates at the equator. This makes the problem more intricate as one is faced with a singular limit problem with variable coefficient.

In this article, we consider the following non-viscous equatorial shallow-water system

$$\begin{cases} \partial_t \vec{v} + \vec{v}\cdot\nabla\vec{v} + \dfrac{1}{\varepsilon}(y\vec{v}^{\perp} + \nabla h) = 0, \\ \partial_t h + \vec{v}\cdot\nabla h + h\nabla\cdot\vec{v} + \dfrac{1}{\varepsilon}\nabla\cdot\vec{v} = 0. \end{cases} \tag{1.1}$$

Here $\vec{v} = (u, v)(t, x, y)$ is the horizontal velocity, $\vec{v}^{\perp} = (-v, u)$, and $h = h(t, x, y)$ is the height. The longitude x is periodic and the distance to equator y will be assumed to vary in $\mathbb{R}$. Namely, we will consider (1.1) in the spatial domain $\mathbb{T}\times\mathbb{R}$. The system (1.1) is written in non-dimensional variables under the assumption that both the Froude number and the height fluctuations are of order ε that is assumed to be a small parameter in this paper.

*Received February 17, 2020; accepted for publication August 11, 2020. This research was supported by the NSFC-ISF joint research program (Grant No. 11761141008). The research of Jiang was supported by NSFC (Grant No. 11631008), and National Key R&D Program (2020YFA0712200), National Key Project (GJXM92579) and the Sino-German Science Center (Grant No. GZ1465). Ju was supported by NSFC (Grant No. 12070144). Xu was supported by Project funded by China Postdoctoral Science Foundation (Grant No. 2021T140633 and 2021M693028), NSFC (Grant No. 12001506) and the Natural Science Foundation of Shandong Province (Grant No. ZR2020QA014). The authors thank Professor Steve Schochet from Tel-Aviv University for helpful discussions.

†Institute of Applied Physics and Computational Mathematics, Beijing 100088, P. R. China (jiang@iapcm.ac.cn).

‡Institute of Applied Physics and Computational Mathematics, Beijing 100088, P. R. China (ju_qiangchang@iapcm.ac.cn).

§Corresponding author. School of Mathematical Sciences, Ocean University of China, Qingdao 266041, P. R. China (xx@ouc.edu.cn).

To transform (1.1) into a multiscale system, we allow $\vec{v}$ and h to be functions of $X = \varepsilon x \in \mathbb{T}$ as well as of $x \in \mathbb{T}$. Thus, system (1.1) becomes

$$\begin{cases} \partial_t u + uu_x + \varepsilon u u_X + v u_y + h_X + \dfrac{1}{\varepsilon}(-yv + h_x) = 0, \\ \partial_t v + uv_x + \varepsilon u v_X + v v_y + \dfrac{1}{\varepsilon}(yu + h_y) = 0, \\ \partial_t h + uh_x + \varepsilon u h_X + v h_y + (1+\varepsilon h) u_X \\ \qquad\qquad\qquad + h(u_x + v_y) + \dfrac{1}{\varepsilon}(u_x + v_y) = 0. \end{cases} \tag{1.2}$$

The spatial domain is $\mathbb{T} \times \mathbb{T} \times \mathbb{R}$ and the system is equipped with the following initial conditions

$$u|_{t=0} = u_{0\varepsilon}, \qquad v|_{t=0} = v_{0\varepsilon}, \qquad h|_{t=0} = h_{0\varepsilon}. \tag{1.3}$$

The behaviour of solutions to (1.2) as the small parameter ε goes to zero has been studied by several authors. The first rigorous proof of this singular limit problem was given by Majda and Dutrifoy in [2]. They proved that solutions to (1.2) depending on X but not on x exist on a uniform time interval, and converge as $\varepsilon \to 0$ to the solution of the following long wave equations

$$\begin{cases} \partial_t u + h_X - yV = 0, \\ \partial_t h + u_X + V_y = 0, \\ yu + h_y = 0, \end{cases} \tag{1.4}$$

for some Lagrangian multiplier V. They also investigated the case in which solutions depend on x but not X in [3]. And the limit equations are identified as some zonal jets. For the viscous equatorial shallow-water equations, Gallagher and Saint-Raymond [7] obtained both weak and strong convergence results by giving a detailed study of the resonances in the limiting process.

As mentioned before, in the viewpoint of singular limit of hyperbolic PDEs, the system (1.2), as ϵ tends to zero, turns to be a singular limit with a large non-constant operator. This makes it more difficult to obtain uniform estimates of solutions in the usual Sobolev spaces since differentiating the equations with respect to y leads to the terms of order $O(\varepsilon^{-1})$. To overcome such difficulties, by a clever use of the special structure of the equations and in the spirit of the classical singular limit theory [9, 10], Dutrifoy, Majda and Schochet [4] succeed in obtaining the uniform estimates in some modified Sobolev space $\widetilde{W}_p$ (see Theorem 1.1 for the definition). Later, this result was generalized to the ill-prepared initial data case by Mullaert [13].

On the other hand, we should mention that convergence rates are of particular interest in the fields like weather prediction in which the slowly-varying part of solutions is of primary interest. More detailed knowledge of the rate of convergence on the slow part of solutions would also improve our understanding of the connection between the original system and the limit system.

In the present paper, we study convergence rates as $\epsilon \to 0$ of the solutions of the equatorial shallow-water system to the solutions of the long wave system, an issue that is not considered in [4, 7, 3, 2]. We focus on the case of well-prepared initial data in this paper (also see Remark 1.3 below).

To state our main result, we first review the theorem by Dutrifoy, Majda and Schochet for convenience.

THEOREM 1.1 ([4]). *For $n \geq 3$, suppose that the initial data $(u_{0\varepsilon}, v_{0\varepsilon}, h_{0\varepsilon}) \in \tilde{W}_{2n}$ with $\tilde{W}_p$ being the Banach space of functions having finite norm*

$$\|w\|^2_{\tilde{W}_p} \equiv \sum_{j+k+l+m\leq p} \|y^j \partial_x^k \partial_X^l \partial_y^m w\|^2_{L^2}.$$

Suppose further that the initial data are well-prepared in the sense that

$$(-yv_{0\varepsilon} + \partial_x h_{0\varepsilon}, yu_{0\varepsilon} + \partial_y h_{0\varepsilon}, \partial_x u_{0\varepsilon} + \partial_y v_{0\varepsilon}) = O(\varepsilon) \qquad \text{in } \tilde{W}_{2(n-1)}.$$

Then a solution

$$(u, v, h) \in C([0, T]; \tilde{W}_{2n}) \cap C^1([0, T]; \tilde{W}_{2(n-1)})$$

to the initial value problem (1.2)–(1.3) exists for a time $T > 0$ independent of ε and satisfies the following uniform estimate

$$\|(u, v, h)\|_{\tilde{W}_{2n}} + \|\partial_t(u, v, h)\|_{\tilde{W}_{2(n-1)}} \leq C.$$

Moreover, assume that

$$u_{0\varepsilon} \to u_0, \qquad h_{0\varepsilon} \to h_0 \qquad \text{in } \tilde{W}_{2(n-1)} \text{ as } \varepsilon \to 0,$$

then the solution (u, v, h) converges in $C([0, T]; \tilde{W}_{2(n-1)})$ to $(u^0, 0, h^0)$ as ε goes to zero, where u^0 and h^0 depend only on t, X and y, and satisfies the long wave equations (1.4) with initial conditions:

$$u^0|_{t=0} = u_0, \qquad h^0|_{t=0} = h_0.$$

The main result of this paper, Theorem 1.2 below, will show that the solution of (1.2), established in Theorem 1.1, will converge to a solution of the long wave system (1.4) at rate of ε under a slightly more regularity assumption on the initial data than in [4]. Before stating our main result, we list some notations that will be used throughout this paper. We use C to denote a general constant that may vary from line to line but is independent of the small parameter ε, and $\|\cdot\|$ the L^2 norm:

$$\|f\| = \Big(\int_{\mathbb{T}} \int_{\mathbb{R}} |f(t, X, y)|^2 dy dX \Big)^{\frac{1}{2}}.$$

THEOREM 1.2. *In addition to the assumptions of Theorem 1.1, suppose that $n \geq 4$ and that the initial data converges to the limit function at rate of ε, namely*

$$\|(u_{0\varepsilon} - u_0, h_{0\varepsilon} - h_0)\|_{\tilde{W}_2} = O(\varepsilon).$$

Then for all $t \in [0, T]$, there is a constant C_0 independent of ε, such that

$$\|\bar{u}(t, X, y) - u^0(t, X, y)\| + \|\bar{h}(t, X, y) - h^0(t, X, y)\| \leq C_0\varepsilon,$$

where $\bar{w}$ denotes the average with respect to x throughout this paper, i.e.

$$\bar{w} = \frac{1}{|\mathbb{T}|} \int_{\mathbb{T}} w(t, x, X, y) dx.$$

REMARK 1.3. By combining the techniques of the present paper with [13] and the time-averaging method [14], it is possible to obtain the convergence rates of the solution even for ill-prepared data. We shall leave this for future study.

We give a few words on the strategy used in the proof of Theorem 1.2. First, to prove the convergence rate of the solution, we need to obtain some useful equations for the difference between the solution to (1.2) and the solution to the limit system (1.4). However, since the large operator in (1.2) is not of constant coefficient, it seems that both averaging and projection methods [14, 8] are insufficient to get the desired equations. Here we shall exploit the special structures of the system to obtain the evolution equations (4.7) and (4.10), and it turns out that these two equations play a central role in our proof. Second, as shown by Dutrifoy, Majda and Schochet in [4], the uniform estimate in Theorem 1.1 is sufficient to get the convergence of the solution of (1.2) to a solution of (1.4). But this uniform estimate of the solution is inadequate to obtain a convergence rate. Therefore, some new uniform estimates on the solution have to be established in this paper. Finally, after reformulating (1.2) into a symmetric hyperbolic system in the next section, the proof of our result is reduced to establishing a convergence rate for l and r (see (2.1) for the definition of l and r). Based on the positivity of the operator $(2 - H)$ (see (2.4) for the definition of H) in (4.10), one can get a convergence rate of l by combining the new uniform estimates with the energy method. However, the method used to get the convergence rate of l fails to work for getting a convergence rate of r because the operator $H + 2$ in (4.7) is neither positive nor negative. To circumvent this difficulty, we study the spectral properties of the operator H in Lemma 4.2 where it is shown that $\|\cdot\| \le \|(H+2)\cdot\|$. Then we can obtain a convergence rate for r by showing that $(H + 2)r$ enjoys the same convergence rate.

The rest part of this paper is arranged as follows. In Section 2 we shall reformulate (1.2) into a symmetric hyperbolic system and give a equivalent theorem to Theorem 1.1 for the new system. Then, in Section 3, some new uniform estimates of the solution are derived and these estimates will be used to obtain the convergence rates of the solution. Finally, the proof of Theorem 1.2 is given in Section 4.

2. Reformulation. In this short section, following [4], we shall reformulate the system (1.2) into a symmetric one with an antisymmetric penalization. Setting

$$r = \frac{1}{\sqrt{2}}u + \frac{\sqrt{2}}{\varepsilon}(\sqrt{1+\varepsilon h} - 1), \qquad l = -\frac{1}{\sqrt{2}}u + \frac{\sqrt{2}}{\varepsilon}(\sqrt{1+\varepsilon h} - 1), \tag{2.1}$$

we get from (1.2) that

$$\begin{cases} r_t + \dfrac{3r-l}{2\sqrt{2}}(r_x + \varepsilon r_X) + r_X + v r_y + \dfrac{r+l}{4}v_y + \dfrac{1}{\varepsilon}(r_x + \dfrac{v_y - yv}{\sqrt{2}}) = 0, \\ l_t + \dfrac{r-3l}{2\sqrt{2}}(l_x + \varepsilon l_X) - l_X + v l_y + \dfrac{r+l}{4}v_y + \dfrac{1}{\varepsilon}(-l_x + \dfrac{v_y + yv}{\sqrt{2}}) = 0, \\ v_t + \dfrac{r-l}{\sqrt{2}}(v_x + \varepsilon v_X) + v v_y + \dfrac{r+l}{4}(l_y + r_y) + \dfrac{1}{\varepsilon}(\dfrac{r_y + yr}{\sqrt{2}} + \dfrac{l_y - yl}{\sqrt{2}}) = 0, \end{cases} \tag{2.2}$$

with initial data

$$r|_{t=0} = r_{0\varepsilon}, \qquad l|_{t=0} = l_{0\varepsilon}, \qquad v|_{t=0} = v_{0\varepsilon}. \tag{2.3}$$

Denoting $U = (r, l, v)^\top$ and

$$L_\pm \equiv \frac{1}{\sqrt{2}}(\partial_y \mp y),$$

we can write (2.2) into the following compact form

$$U_t + A_1(U)U_x + A_2(U)U_X + A_3(U)U_y + \frac{1}{\varepsilon}MU = 0,$$

where A_j are symmetric matrices and M is given by

$$M \equiv \begin{bmatrix} \partial_x & 0 & L_+ \\ 0 & -\partial_x & L_- \\ L_- & L_+ & 0 \end{bmatrix}.$$

One can easily check that M is antisymmetric.

Now, for later use, let us introduce the operator

$$H \equiv L_-L_+ + L_+L_- = \partial_y^2 - y^2. \tag{2.4}$$

By a straightforward calculation, we find that

$$[L_+, L_-] = I, \qquad [H, L_\pm] = \mp 2L_\pm, \tag{2.5}$$

where $[\cdot,\cdot]$ is the commutator defined by $[A, B] = AB - BA$. For the system (2.2)–(2.3), we have the following uniform existence and convergence theorem corresponding to Theorem 1.1.

THEOREM 2.1 ([4]). *Suppose that the initial data $(r_{0\varepsilon}, l_{0\varepsilon}, v_{0\varepsilon}) \in \tilde{W}_{2n}$ with $n \geq 3$ and are well-prepared in the sense that the first order time-derivative of (r, l, v) at $t = 0$ is of order one. Then, a solution (r, l, v) to system (2.2)–(2.3) exists in a uniform time interval $[0, T]$ and satisfies the uniform estimate*

$$\|(r, l, v)\|_{\tilde{W}_{2n}} + \|\partial_t(r, l, v)\|_{\tilde{W}_{2(n-1)}} \leq C. \tag{2.6}$$

Assume further that the initial data

$$r_{0\varepsilon} \to r_0, \qquad l_{0\varepsilon} \to l_0, \qquad \text{in } \tilde{W}_{2(n-1)},$$

then as $\varepsilon \to 0$, the solution (r, l, v) converges to $(r^0, l^0, 0)$ in $C^0([0, T]; \tilde{W}_{2(n-1)})$, and

$$\varepsilon^{-1} v \to v^1,$$

where (r^0, l^0, v^1) satisfies

$$\begin{cases} \partial_t r^0 + r^0_X + L_+ v^1 = 0, \\ \partial_t l^0 - l^0_X + L_- v^1 = 0, \\ L_- r^0 + L_+ l^0 = 0, \end{cases} \tag{2.7}$$

with initial data (r_0, l_0).

Note that, by virtue of (2.5), the limit system (2.7) can be rewritten into

$$\begin{cases} \partial_t (2 + H) r^0 + r^0_X = 0, \\ \partial_t (2 - H) l^0 - l^0_X = 0. \end{cases} \tag{2.8}$$

As one will see later, this form of the limit equations turns to be more useful in obtaining a convergence rate of solutions.

3. Uniform Estimates. In this section we derive some uniform estimates of the solution by using (2.6) and (2.2). These estimates are indispensable to obtain a convergence rate in the next section.

The main result in this section is stated in the following lemma.

LEMMA 3.1. *Under the assumptions of Theorem 2.1, we have the following uniform estimate for the solution* (r, l, v)*:*

$$\|L_-r + L_+l\|_{\tilde{W}_{2(n-1)}} + \|v\|_{\tilde{W}_{2(n-2)}} + \|(r_x, l_x)\|_{\tilde{W}_{2(n-3)}} \le C\varepsilon. \tag{3.1}$$

Proof. First, from $(2.2)_3$ and the uniform estimate (2.6) we get

$$\begin{aligned}
&\|L_-r + L_+l\|_{\tilde{W}_{2(n-1)}} \\
&\quad = \varepsilon\|v_t + \frac{r-l}{\sqrt{2}}(v_x + \varepsilon v_X) + vv_y + \frac{r+l}{4}(l_y + r_y)\|_{\tilde{W}_{2(n-1)}} \\
&\quad \le \varepsilon\|v_t\|_{\tilde{W}_{2(n-1)}} + \varepsilon\|\frac{r-l}{\sqrt{2}}(v_x + \varepsilon v_X) + vv_y + \frac{r+l}{4}(l_y + r_y)\|_{\tilde{W}_{2(n-1)}} \\
&\quad \le C\varepsilon,
\end{aligned} \tag{3.2}$$

where we have used the product estimates in $\tilde{W}_{2n}$. Next, taking $L_-(2.2)_1 - L_+(2.2)_2 - \partial_x(2.2)_3$ and using $[L_+, L_-] = I$ of (2.5), one deduces that

$$\begin{aligned}
\varepsilon^{-1}v =& L_-\Big[r_t + \frac{3r-l}{2\sqrt{2}}(r_x + \varepsilon r_X) + r_X + vr_y + \frac{r+l}{4}v_y\Big] \\
&- L_+\Big[l_t + \frac{r-3l}{2\sqrt{2}}(l_x + \varepsilon l_X) - l_X + vl_y + \frac{r+l}{4}v_y\Big] \\
&- \partial_x\Big[v_t + \frac{r-l}{\sqrt{2}}(v_x + \varepsilon v_X) + vv_y + \frac{r+l}{4}(l_y + r_y)\Big].
\end{aligned}$$

Therefore, using the uniform estimate (2.6), we conclude

$$\begin{aligned}
\|v\|_{\tilde{W}_{2(n-2)}} \le& \varepsilon\|L_-\Big[r_t + \frac{3r-l}{2\sqrt{2}}(r_x + \varepsilon r_X) + r_X + vr_y + \frac{r+l}{4}v_y\Big]\|_{\tilde{W}_{2(n-2)}} \\
&+ \varepsilon\|L_+\Big[l_t + \frac{r-3l}{2\sqrt{2}}(l_x + \varepsilon l_X) - l_X + vl_y + \frac{r+l}{4}v_y\Big]\|_{\tilde{W}_{2(n-2)}} \\
&+ \varepsilon\|\partial_x\Big[v_t + \frac{r-l}{\sqrt{2}}(v_x + \varepsilon v_X) + vv_y + \frac{r+l}{4}(l_y + r_y)\Big]\|_{\tilde{W}_{2(n-2)}} \\
&\le C\varepsilon.
\end{aligned} \tag{3.3}$$

Now, from the first two equations of (2.2) one gets

$$\begin{cases}
r_x = -\varepsilon\big[r_t + \frac{3r-l}{2\sqrt{2}}(r_x + \varepsilon r_X) + r_X + vr_y + \frac{r+l}{4}v_y\big] - L_+v, \\
l_x = \varepsilon\big[l_t + \frac{r-3l}{2\sqrt{2}}(l_x + \varepsilon l_X) - l_X + vl_y + \frac{r+l}{4}v_y\big] + L_-v.
\end{cases} \tag{3.4}$$

Using (3.3), we infer from (3.4) that

$$\begin{aligned}
&\|r_x\|_{\tilde{W}_{2(n-3)}} + \|l_x\|_{\tilde{W}_{2(n-3)}} \\
&\qquad \le \varepsilon\big\|\big[r_t + \frac{3r-l}{2\sqrt{2}}(r_x+\varepsilon r_X) + r_X + vr_y + \frac{r+l}{4}v_y\big]\big\|_{\tilde{W}_{2(n-3)}} \\
&\qquad + \varepsilon\big\|\big[l_t + \frac{r-3l}{2\sqrt{2}}(l_x+\varepsilon l_X) - l_X + vl_y + \frac{r+l}{4}v_y\big]\big\|_{\tilde{W}_{2(n-3)}} \\
&\qquad + \|L_+v\|_{\tilde{W}_{2(n-3)}} + \|L_-v\|_{\tilde{W}_{2(n-3)}} \\
&\qquad \le C\varepsilon.
\end{aligned} \tag{3.5}$$

Putting the estimates (3.2), (3.3) and (3.5) together, we complete the proof of Lemma 3.1. □

4. Convergence Rates. Based on the uniform estimates in the previous section, we can obtain a convergence rate in this section.

First, we need to get the equations for the difference of the solution to (2.2) and the solution to the limit system (2.8). Averaging the first two equations in (2.2) with respect to x gives

$$\begin{cases} \bar{r}_t + \bar{r}_X + \dfrac{1}{\varepsilon}L_+\bar{v} = R_1, \\ \bar{l}_t - \bar{l}_X + \dfrac{1}{\varepsilon}L_-\bar{v} = R_2, \end{cases} \tag{4.1}$$

where

$$\begin{aligned}
R_1 &= -\overline{\frac{3r-l}{2\sqrt{2}}(r_x+\varepsilon r_X)} - \overline{vr_y} - \overline{\frac{r+l}{4}v_y}, \\
R_2 &= -\overline{\frac{r-3l}{2\sqrt{2}}(l_x+\varepsilon l_X)} - \overline{vl_y} - \overline{\frac{r+l}{4}v_y}.
\end{aligned}$$

Applying L_- to $(4.1)_1$ and L_+ to $(4.1)_2$, one sees that

$$\begin{cases} L_-\bar{r}_t + L_-\bar{r}_X + \frac{1}{\varepsilon}L_-L_+\bar{v} = L_-R_1, \\ L_+\bar{l}_t - L_+\bar{l}_X + \frac{1}{\varepsilon}L_+L_-\bar{v} = L_+R_2. \end{cases} \tag{4.2}$$

Recalling $[L_+, L_-] = I$ and (4.2), we find that

$$\partial_t(L_-\bar{r} - L_+\bar{l}) + \partial_X(L_-\bar{r} + L_+\bar{l}) - \frac{1}{\varepsilon}\bar{v} = L_-R_1 - L_+R_2. \tag{4.3}$$

Then, taking $(4.1)_1 + L_+(4.3)$ gives

$$\begin{aligned}
&\partial_t\bar{r} + \partial_t L_+(L_-\bar{r} - L_+\bar{l}) + \bar{r}_X \\
&\quad = -\partial_X L_+(L_-\bar{r} + L_+\bar{l}) + L_+L_-R_1 - L_+L_+R_2 + R_1.
\end{aligned} \tag{4.4}$$

By a direct calculation, we have

$$\begin{aligned}
\partial_t\bar{r} + \partial_t L_+(L_-\bar{r} - L_+\bar{l}) &= \partial_t\big[L_+L_-\bar{r} - L_-L_+\bar{r} + L_+L_-\bar{r} - L_+L_+\bar{l}\big] \\
&= \partial_t\big[L_+L_-\bar{r} + L_-L_+\bar{r} - 2L_-L_+\bar{r} + L_+L_-\bar{r} - L_+L_+\bar{l}\big] \\
&= \partial_t\big[H\bar{r} - 2L_-L_+\bar{r} + 2L_+L_-\bar{r} - L_+L_-\bar{r} - L_+L_+\bar{l}\big] \\
&= \partial_t\big[(H+2)\bar{r} - L_+(L_-\bar{r} + L_+\bar{l})\big].
\end{aligned} \tag{4.5}$$

Thus, substitution of (4.5) into (4.4) yields

$$\partial_t\big[(H+2)\bar r - L_+(L_-\bar r + L_+\bar l)\big] + \bar r_X \\ = -\partial_X L_+(L_-\bar r + L_+\bar l) + L_+L_-R_1 - L_+L_+R_2 + R_1, \tag{4.6}$$

which, combined with $(2.8)_1$, gives

$$\partial_t\big[(H+2)(\bar r - r^0) - L_+(L_-\bar r + L_+\bar l)\big] + \partial_X(\bar r - r^0) \\ = -\partial_X L_+(L_-\bar r + L_+\bar l) + L_+L_-R_1 - L_+L_+R_2 + R_1. \tag{4.7}$$

Next, taking $(4.1)_2 + L_-(4.3)$ gives

$$\partial_t \bar l + \partial_t L_-(L_-\bar r - L_+\bar l) - \bar l_X \\ = -\partial_X L_-(L_-\bar r + L_+\bar l) + L_-L_-R_1 - L_-L_+R_2 + R_2. \tag{4.8}$$

In a similar fashion, we can rewrite (4.8) into

$$\partial_t\big[(2-H)\bar l + L_-(L_-\bar r + L_+\bar l)\big] - \bar l_X \\ = -\partial_X L_-(L_-\bar r + L_+\bar l) + L_-L_-R_1 - L_-L_+R_2 + R_2. \tag{4.9}$$

Combining (4.9) with $(2.8)_2$, we obtain

$$\partial_t\big[(2-H)(\bar l - l^0) + L_-(L_-\bar r + L_+\bar l)\big] - \partial_X(\bar l - l^0) \\ = -\partial_X L_-(L_-\bar r + L_+\bar l) + L_-L_-R_1 - L_-L_+R_2 + R_2. \tag{4.10}$$

By virtue of the uniform estimates established in previous section and the positivity of the operator $(2-H)$, it is easy to see that to get the desired convergence rate in this paper, it is sufficient for us to get the same convergence rate for $\bar l(t,X,y)$ from (4.10). In fact, we have the following result:

THEOREM 4.1. *In addition to the assumptions in Theorem 2.1, assume further that* $n \geq 4$ *and*

$$\|l_{0\varepsilon} - l_0\|_{\tilde W_1} = O(\varepsilon), \tag{4.11}$$

then we have

$$\sup_{t\in[0,T]} \Big\{\|\bar l(t,X,y) - l^0(t,X,y)\| + \|\partial_y\big(\bar l(t,X,y) - l^0(t,X,y)\big)\|\Big\} \leq C\varepsilon.$$

Proof. Setting

$$f = (2-H)^{-1}L_-(L_-\bar r + L_+\bar l),$$

we can rewrite (4.10) into

$$\partial_t(2-H)(\bar l - l^0 + f) - \partial_X(\bar l - l^0 + f) \\ = -\partial_X L_-(L_-\bar r + L_+\bar l) + L_-L_-R_1 - L_-L_+R_2 + R_2 - \partial_X f. \tag{4.12}$$

Taking the inner product of (4.12) with $(\bar l - l^0 + f)$ results in

$$\frac{d}{dt}\int_{\mathbb T}\int_{\mathbb R}(2-H)(\bar l - l^0 + f)(\bar l - l^0 + f)dydX \\ \leq \|\partial_X L_-(L_-\bar r + L_+\bar l)\|\|(\bar l - l^0 + f)\| + \|\partial_X f\|\|(\bar l - l^0 + f)\| \\ + \|L_-L_-R_1 - L_-L_+R_2 + R_2\|\|(\bar l - l^0 + f)\|. \tag{4.13}$$

As $H = \partial_y^2 - y^2$, one has

$$\begin{aligned}
&\int_{\mathbb{T}}\int_{\mathbb{R}}(2-H)(\bar{l}-l^0+f)(\bar{l}-l^0+f)dydX \\
&\qquad = \int_{\mathbb{T}}\int_{\mathbb{R}}(2+y^2-\partial_y^2)(\bar{l}-l^0+f)(\bar{l}-l^0+f)dydX \\
&\qquad = \int_{\mathbb{T}}\int_{\mathbb{R}}(2+y^2)(\bar{l}-l^0+f)^2+|\partial_y(\bar{l}-l^0+f)|^2dydX.
\end{aligned}$$

Denoting

$$E^2(t) = \int_{\mathbb{T}}\int_{\mathbb{R}}(2+y^2)(\bar{l}-l^0+f)^2+|\partial_y(\bar{l}-l^0+f)|^2dydX,$$

we get from (4.13) that

$$\frac{d}{dt}E(t) \le Q_1 + Q_2 + Q_3, \tag{4.14}$$

where

$$\begin{aligned}
Q_1 &= \|\partial_X L_-(L_-\bar{r}+L_+\bar{l})\|, \\
Q_2 &= \|\partial_X f\|, \\
Q_3 &= \|L_-L_-R_1 - L_-L_+R_2 + R_2\|.
\end{aligned}$$

Next, we have to control each Q_j by exploiting the uniform estimate established in Lemma 3.1. For Q_1, we can use Lemma 3.1 to see that

$$\begin{aligned}
Q_1 &= \|\overline{\partial_X L_-(L_-r+L_+l)}\| \\
&\le \Big(\int_{\mathbb{T}}\int_{\mathbb{R}}\Big|\int_{\mathbb{T}}\partial_X L_-(L_-r+L_+l)dx\Big|^2 dydX\Big)^{\frac{1}{2}} \\
&\le \Big[\int_{\mathbb{T}}\int_{\mathbb{R}}\Big(\int_{\mathbb{T}}\big|\partial_X L_-(L_-r+L_+l)\big|dx\Big)^2 dydX\Big]^{\frac{1}{2}} \\
&\le \Big[\int_{\mathbb{T}}\int_{\mathbb{R}}|\mathbb{T}|\int_{\mathbb{T}}\big|\partial_X L_-(L_-r+L_+l)\big|^2 dxdydX\Big]^{\frac{1}{2}} \\
&\le |\mathbb{T}|^{\frac{1}{2}}\|(L_-r+L_+l)\|_{\tilde{W}_{2(n-1)}} \\
&\le C\varepsilon.
\end{aligned}$$

Then, from the boundedness of $(2-H)^{-1}$ we get

$$\begin{aligned}
Q_2 &= \|\partial_X(2-H)^{-1}L_-(L_-\bar{r}+L_+\bar{l})\| \\
&\le C\|\partial_X L_-(L_-\bar{r}+L_+\bar{l})\| \\
&\le C\|(L_-\bar{r}+L_+\bar{l})\|_{\tilde{W}_{2(n-1)}} \\
&\le C\varepsilon.
\end{aligned}$$

Finally, for Q_3 one sees that

$$Q_3 \le \|L_-L_-R_1\| + \|L_-L_+R_2\| + \|R_2\|,$$

where three terms on the right hand side can be bounded in a similar manner, and here we just show how to control the first term for simplicity. In view of the expression of R_1, we find that

$$\begin{aligned}\|L_-L_-R_1\| &= \|L_-L_-\overline{\Big(\frac{3r-l}{2\sqrt{2}}(r_x+\varepsilon r_X)} + \overline{vr_y} + \overline{\frac{r+l}{4}v_y\Big)}\| \\ &= \|\overline{L_-L_-\Big(\frac{3r-l}{2\sqrt{2}}(r_x+\varepsilon r_X) + vr_y + \frac{r+l}{4}v_y\Big)}\| \\ &\le C\|\frac{3r-l}{2\sqrt{2}}(r_x+\varepsilon r_X) + vr_y + \frac{r+l}{4}v_y\|_{\tilde{W}_2} \\ &\le C\|(r,l)\|_{\tilde{W}_2}(\|r_x\|_{\tilde{W}_2} + \varepsilon\|r_X\|_{\tilde{W}_2}) + C\|v\|_{\tilde{W}_2}\|r_y\|_{\tilde{W}_2} \\ &\quad + C\|(r,l)\|_{\tilde{W}_2}\|v_y\|_{\tilde{W}_2}.\end{aligned}$$

Substituting the uniform estimates (2.6) and (3.1) into the above inequality, we immediately have

$$\|L_-L_-R_1\| \le C\varepsilon.$$

In a similar way, one can show that $\|L_-L_+R_2\|$ and $\|R_2\|$ enjoy the same estimates. Hence, putting the estimates of Q_j into (4.14) and integrating the resulting inequality over $[0,t]$ for any $t \in [0,T]$, we conclude

$$\sup_{t\in[0,T]} E(t) \le E(0) + C\varepsilon. \tag{4.15}$$

Next, we proceed to estimate $E(0)$. Since we are considering the well-prepared initial data case, i.e. $\|U_t\|_{\tilde{W}_{2(n-1)}} = O(1)$, we get from the last equation in (2.2) that

$$\begin{aligned}&\|(L_-r+L_+l)\big|_{t=0}\|_{\tilde{W}_2} \\ &\quad \le \varepsilon\|v_t\big|_{t=0}\|_{\tilde{W}_2} + \varepsilon\|\Big(\frac{r-l}{\sqrt{2}}(v_x+\varepsilon v_X) + vv_y + \frac{r+l}{4}(l_y+r_y)\Big)\big|_{t=0}\|_{\tilde{W}_2} \\ &\quad \le C\varepsilon.\end{aligned} \tag{4.16}$$

Using the initial condition (4.11) in Theorem 4.1, we have

$$E(0) \le C\|l_{0\varepsilon} - l_0\|_{\tilde{W}_1} + C\|(L_-r+L_+l)\big|_{t=0}\|_{\tilde{W}_2} \le C\varepsilon.$$

Combining the above estimate with (4.15), one obtains

$$\sup_{t\in[0,t]} E(t) \le C\varepsilon,$$

from which, the convergence rate of $\bar{l}$ easily follows. Actually, we can obtain

$$\begin{aligned}\|\bar{l}-l^0\| + \|\partial_y(\bar{l}-l^0)\| &\le \|\bar{l}-l^0+f\| + \|\partial_y(\bar{l}-l^0+f)\| + \|(f,\partial_y f)\| \\ &\le E(t) + C\|(L_-\bar{r}+L_+\bar{l})\|_{\tilde{W}_2} \le C\varepsilon,\end{aligned}$$

which completes the proof of Theorem 4.1. □

So far, we have obtained the convergence rate for $\bar{l}$. To complete the proof, it remains to get the same convergence rate for $\bar{r}$. However, one can check that it is difficult to use a similar idea to obtain a convergence rate for $\bar{r}$. The reason lies in that the operator $H+2$ in (4.7) is neither positive nor negative. To get a convergence rate of $\bar{r}$, we shall make full use of the spectral property of the operator H. We start with the following lemma.

LEMMA 4.2. *Let $n \geq 3$ and $H = \partial_y^2 - y^2$ be a operator on $\tilde{W}_{2n}$, then it holds that (A)*

$$Ker(H+2) = \{0\}.$$

(B) For any $f \in \tilde{W}_{2n}$,

$$\|f\| \leq \|(H+2)f\|. \tag{4.17}$$

Proof. The proof of this lemma is based on the following parabolic cylinder functions. Let

$$\phi_n(y) = (2^n n!\sqrt{\pi})^{-\frac{1}{2}} H_n(y) e^{-\frac{y^2}{2}}, \qquad n = 0, 1, 2, \cdots,$$

with $H_n(y)$ being the Hermite polynomial of degree n

$$H_n(y) = (-1)^n e^{y^2} \frac{d^n}{dy^n} e^{-\frac{y^2}{2}}.$$

Since H only deals with y, we may assume that functions in $\tilde{W}_{2n}$ depends only on y for the sake of simplicity. Then one can easily check that $\{\phi_n(y)\}$ forms an orthonormal basis of $\tilde{W}_{2n}$. Moreover,

$$\begin{aligned}
&L_+\phi_n(y) = -\sqrt{n+1}\phi_{n+1}(y), \qquad n = 0, 1, 2, \cdots, \\
&L_-\phi_n(y) = \sqrt{n}\phi_{n-1}(y), \qquad n = 1, 2, \cdots, \\
&L_-\phi_0(y) = 0.
\end{aligned}$$

Hence, we have

$$H\phi_0(y) = L_+L_-\phi_0(y) + L_-L_+\phi_0(y) = -L_-\phi_1(y) = -\phi_0(y),$$

and for $n = 1, 2, \cdots$,

$$\begin{aligned}
H\phi_n(y) &= L_+L_-\phi_n(y) + L_-L_+\phi_n(y) \\
&= \sqrt{n}L_+\phi_{n-1}(y) - \sqrt{n+1}L_-\phi_{n+1}(y) \\
&= -n\phi_n(y) - (n+1)\phi_n(y) = -(2n+1)\phi_n(y).
\end{aligned}$$

Thus, for $n = 0, 1, 2, \cdots$, one has

$$(H+2)\phi_n(y) = -(2n-1)\phi_n(y). \tag{4.18}$$

For any $f \in \tilde{W}_{2n}$, we may decompose it as

$$f(y) = \sum_{n=0}^{\infty} f_n \phi_n(y).$$

Using (4.18), we obtain

$$(H+2)f(y) = -\sum_{n=0}^{\infty}(2n-1)f_n\phi_n(y).$$

Therefore, if $(H+2)f(y)=0$, f must be zero. This gives the statement (A). In addition, since the absolute value of all the eigenvalues of $H+2$ admits a positive lower bound 1, (B) follows easily. □

With the help of Lemma 4.2, we are now ready to get a convergence rate for $\bar{r}$.

THEOREM 4.3. *In addition to the assumptions in Theorem 2.1, assume further that $n \geq 4$ and*

$$\|r_{0\varepsilon} - r_0\|_{\tilde{W}_2} = O(\varepsilon), \tag{4.19}$$

then it holds that

$$\sup_{t\in[0,T]} \|\bar{r}(t,X,y) - r^0(t,X,y)\| \leq C\varepsilon.$$

Proof. Denoting

$$g = -(H+2)^{-1}L_+(L_-\bar{r} + L_+\bar{l})$$

we can rewrite (4.7) into

$$\begin{aligned}\partial_t(H+2)(\bar{r} - r^0 + g) + \partial_X(\bar{r} - r^0 + g)\\ = -\partial_X L_+(L_-\bar{r} + L_+\bar{l}) + L_+L_-R_1 - L_+L_+R_2 + R_1 + \partial_X g.\end{aligned} \tag{4.20}$$

Multiplying (4.20) by $(H+2)(\bar{r} - r^0 + g)$ and integrating it over $\mathbb{T}\times\mathbb{R}$, one finds that

$$\begin{aligned}&\frac{d}{dt}\|(H+2)(\bar{r} - r^0 + g)\|^2 + \int_{\mathbb{T}}\int_{\mathbb{R}} \partial_X(\bar{r} - r^0 + g)(H+2)(\bar{r} - r^0 + g)dydX\\ &\quad\leq \|\partial_X L_+(L_-\bar{r} + L_+\bar{l})\|\|(H+2)(\bar{r} - r^0 + g)\|\\ &\qquad + \|L_+L_-R_1 - L_+L_+R_2 + R_1 + \partial_X g\|\|(H+2)(\bar{r} - r^0 + g)\|.\end{aligned} \tag{4.21}$$

Now, let us calculate the second term on the left-hand side of (4.21). Recalling the expression of H, we see that

$$\begin{aligned}&\int_{\mathbb{T}}\int_{\mathbb{R}} \partial_X(\bar{r} - r^0 + g)(H+2)(\bar{r} - r^0 + g)dydX\\ &\quad= \int_{\mathbb{T}}\int_{\mathbb{R}} \partial_X(\bar{r} - r^0 + g)(\partial_y^2 - y^2 + 2)(\bar{r} - r^0 + g)dydX\\ &\quad= -\int_{\mathbb{T}}\int_{\mathbb{R}} \partial_X\partial_y(\bar{r} - r^0 + g)\partial_y(\bar{r} - r^0 + g)dydX\\ &\qquad + \int_{\mathbb{R}}(2-y^2)\int_{\mathbb{T}} \partial_X(\bar{r} - r^0 + g)(\bar{r} - r^0 + g)dXdy\\ &\quad= 0.\end{aligned}$$

Thus, we get from (4.21) that

$$\frac{d}{dt}\|(H+2)(\bar{r}-r^0+g)\| \\ \leq \|\partial_X L_+(L_-\bar{r}+L_+\bar{l})\| + \|L_+L_-R_1 - L_+L_+R_2 + R_1 + \partial_X g\|. \qquad (4.22)$$

Similarly to the estimate of Q_j in the proof of Theorem 4.1, one can see that the right-hand side of (4.22) can be bounded from above by $C\varepsilon$, using (2.6), (3.1) and the product estimates. Therefore, an integration of (4.22) over $[0,t]$ yields

$$\begin{aligned}\|(H+2)(\bar{r}-r^0+g)\| &\leq \|(H+2)(\bar{r}-r^0+g)\big|_{t=0}\| + C\varepsilon \\ &\leq \|(H+2)(\bar{r}-r^0)\big|_{t=0}\| + \|L_+(L_-\bar{r}+L_+\bar{l})\big|_{t=0}\| + C\varepsilon \\ &\leq C\|(r_{0\varepsilon}-r_0)\|_{\tilde{W}_2} + \|(L_-\bar{r}+L_+\bar{l})\big|_{t=0}\|_{\tilde{W}_1} + C\varepsilon \\ &\leq C\varepsilon,\end{aligned}$$

where we have used (4.19) and (4.16).

Finally, from (4.17) in Lemma 4.2 we get

$$\begin{aligned}\|\bar{r}-r^0\| &\leq \|(\bar{r}-r^0+g)\| + \|g\| \\ &\leq \|(H+2)(\bar{r}-r^0+g)\| + \|g\| \leq C\varepsilon,\end{aligned}$$

which completes the proof of Theorem 4.3. □

Proof of Theorem 1.2. It is easy to see that Theorem 1.2 follows from Theorems 4.1 and 4.3 immediately. We should point out here that the requirements (4.11) and (4.19) are strictly weaker than the assumptions of Theorem 1.2, but we get more conclusions in Theorems 4.1 and 4.3.

REFERENCES

[1] J.-Y. Chemin, B. Desjardins, I. Gallagher, and E. Grenier, *Mathematical Geophysics*, Oxford Lecture Series in Mathematics and Its Applications, 32. Clarendon Press, Oxford University Press, Oxford, 2006.

[2] A. Dutrifoy and A. Majda, *The dynamics of equatorial long waves: a singular limit with fast variable coefficients*, Comm. Math. Sci., 4:2 (2006), pp. 375–397.

[3] A. Dutrifoy and A. Majda, *Fast wave averaging for the equatorial shallow water equations*, Comm. Partial Diff. Eqs., 32:10 (2007), pp. 1617–1642.

[4] A. Dutrifoy, S. Schochet and A. Majda, *A simple justification of the singular limit for equatorial shallow-water dynamics*, Comm. Pure Appl. Math., 62:3 (2009), pp. 322–333.

[5] P. F. Embid and A. Majda, *Averaging over fast gravity waves for geophysical flows with arbitrary potential vorticity*, Comm. Partial Diff. Eqs., 21:3-4 (1996), pp. 619–658.

[6] P. F. Embid and A. Majda, *Low Froude number limiting dynamics for stably stratified flow with small or finite Rossby numbers*, Geophys. Astrophys. Fluid Dynam., 87:1-2 (1998), pp. 1–50.

[7] I. Gallagher and L. Saint-Raymond, *Mathematical study of the betaplane model: equatorial waves and convergence results [M]*, Société mathématique de France, 2006.

[8] S. Jiang, Q. Ju, and X. Xu, *Small Alfvén number limit for incompressible magneto-hydrodynamics in a domain with boundaries*, Sci. China Math., 62:11 (2019), pp. 2229–2248.

[9] S. Klainerman and A. Majda, *Singular limits of quasilinear hyperbolic systems with large parameters and the incompressible limit of compressible fluids*, Comm. Pure Appl. Math., 34:4 (1981), pp. 481–524.

[10] S. Klainerman and A. Majda, *Compressible and incompressible fluids*, Comm. Pure Appl. Math., 35:5 (1982), pp. 629–651.

[11] A. Majda, *Introduction to PDEs and Waves for the Atmosphere and Ocean. Courant Lecture Notes in Mathematics*, 9. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, R.I., 2003
[12] A. Majda and P. Embid, *Averaging over fast gravity waves for geophysical flows with unbalanced initial data*, Theoret. Comput. Fluid Dynam., 11:3-4 (1998), pp. 155–169.
[13] C. Mullaert, *Study of the non-viscous equatorial shallow water system*, J. Math. Pures Appl., 99:3 (2013), pp. 251–279.
[14] S. Schochet, *Fast singular limits of hyperbolic PDEs*, J. Diff. Eqs., 114:2 (1994), pp. 476–512.
[15] R. Temam and M. Ziane, *Some mathematical problems in geophysical fluid dynamics*, Handbook of Mathematical Fluid Dynamics, Vol. 3, pp. 535–657, North-Holland, Amsterdam, 2004.