Extension of Ekman (1905) wind-driven transport theory to the β-plane

Nathan Paldor[1] and Lazar Friedland[2]

[1]Fredy and Nadine Herrmann Institute of Earth Sciences, Hebrew University of Jerusalem, Jerusalem, Israel
[2]Racah Institute of Physics, Hebrew University of Jerusalem, Jerusalem, Israel

Correspondence: Nathan Paldor (nathan.paldor@mail.huji.ac.il)

Abstract. The seminal, Ekman (1905)'s, f-plane theory of wind driven transport at the ocean surface is extended to the β-plane by substituting the pseudo angular momentum for the zonal velocity in the Lagrangian equation. When the β term is added, the equations become nonlinear, which greatly complicates the analysis. Though rotation relates the momentum equations in the zonal and the meridional directions, the transformation to pseudo angular momentum greatly simplifies the longitudinal dynamics, which yields a clear description of the meridional dynamics in terms of a slow drift compounded by fast oscillations, which can then be applied to describe the motion in the zonal direction. Both analytical expressions and numerical calculations highlight the critical role of the equator in determining the trajectories of water columns forced by eastward directed (in the northern hemisphere) wind stress even when the water columns are initiated far from the equator. Our results demonstrate that the averaged motion in the zonal direction depends on the amplitude of the meridional oscillations and ~~it~~ is independent of the direction of the wind stress. The zonal drift is determined by a balance between the initial conditions and the magnitude of the wind stress so ~~for some initial conditions it is~~ it can be as large as the mean meridional motion i.e., the averaged flow direction is not necessarily perpendicular to the wind direction.

1 Introduction

The seminal theory of wind driven transport at the ocean surface was developed about 120 years ago by the Swedish oceanographer Vagn Walfrid Ekman for the highly idealized case of constant Coriolis frequency – the f-plane. The Ekman (1905) theory addresses the downward spiraling horizontal velocity in the ocean's surface and its vertical integral – the transport. Ekman's elegant solution of the problem has become a textbook material in physical oceanography, dynamical meteorology and geophysical fluid dynamics (see e.g. Gill, 1982; Pedlosky, 1987; Vallis, 2017). For uniform wind stress the dynamics on the f-plane consists of two parts: A steady flow to the right/left of the wind direction in the northern/southern hemisphere and inertial oscillations (of frequency f_0 – the constant Coriolis frequency). However, though it is one of the cornerstones of atmosphere and ocean dynamics, the theory was never extended to include the latitudinal increase in the Coriolis frequency,

known as the β effect, which is the focus of the present study. In contrast to the β-plane, in spherical coordinates the theory of wind-driven transport was studied numerically in Constantin and Johnson (2019) and Paldor (2002) but due to the complexity of the governing equations in these coordinates, the numerical solutions have not yielded analytic understanding. With the wind-driven dynamics on the f-plane fully understood and quantified, the β-plane offers an in-between set-up where analytical insight can complement the numerical solutions.

For given wind-stress forcing the known general differences between the dynamics on the f-plane and β-plane suggest heuristically that the extension of Ekman's transport theory to the β-plane should include the following qualitative elements:

1. An increase/decrease in mean meridional velocity for an eastward/westward directed stress due to the decrease/increase in Coriolis frequency when the water column moves southward/northward.

2. The frequency of oscillation about the mean velocity should decrease/increase (so oscillation period should increase/decrease) due to the decrease/increase in Coriolis frequency along the trajectory (for an eastward directed stress in the northern hemisphere while the opposite changes occur for westward directed tress and in the southern hemisphere).

3. Since the oscillation's frequency and amplitude are inversely correlated (energy flux is unchanged) a decrease in frequency should lead to an increase in amplitude and vise versa.

4. Since inertial oscillations, that form a perfectly circular motion on the f-plane, drift westward on the β-plane the averaged zonal motion should drift to the west. A heuristic reasoning of the westward drift in terms of the change in the radius of the inertia circle was proposed by Von Arx (1964) and complete quantitative theories of the drift were developed in Ripa (1997) and Paldor (2007).

The numerical solutions of the governing Lagrangian equations (see section 2 below) shown in figure 1 fully confirm the first 3 expectations listed above but contradict the fourth one – for both westward (right panel) and eastward (left panel) stresses, the trajectories drift to the east. From the particular example shown in figure 1 it is unclear whether the eastward transition is a general feature of the wind driven dynamics on the β-plane or a specific occurrence related to the particular choice of initial conditions and/or parameter values.

In addition to resolving the issue of the zonal drift and quantifying the various rates of changes, the present study also addresses the equatorial problem that exists only on the β-plane. This equatorial issue can be described as follows: An eastward directed stress in the northern hemisphere forces a net southward directed mean flow which, on the β-plane, is accompanied by a decrease in the Coriolis frequency. Thus, at some time the wind forced water column must find itself in a latitude where the Coriolis frequency vanishes – the equator. From that point onward the water column is subject to non-rotating dynamics and must move eastward at an accelerated velocity. In the rest of this work we will estimate the time it takes the water column to change its dynamics from rotating to non-rotating and analyze how the two dynamical regimes connect with one another.

The work is organized as follows: In section 2 we nondimensionalize the governing Lagrangian equations and simplify them by substituting the pseudo angular momentum for the zonal velocity. The simplified system is analyzed in section 3 and the work concludes with a discussion and summary in section 4.

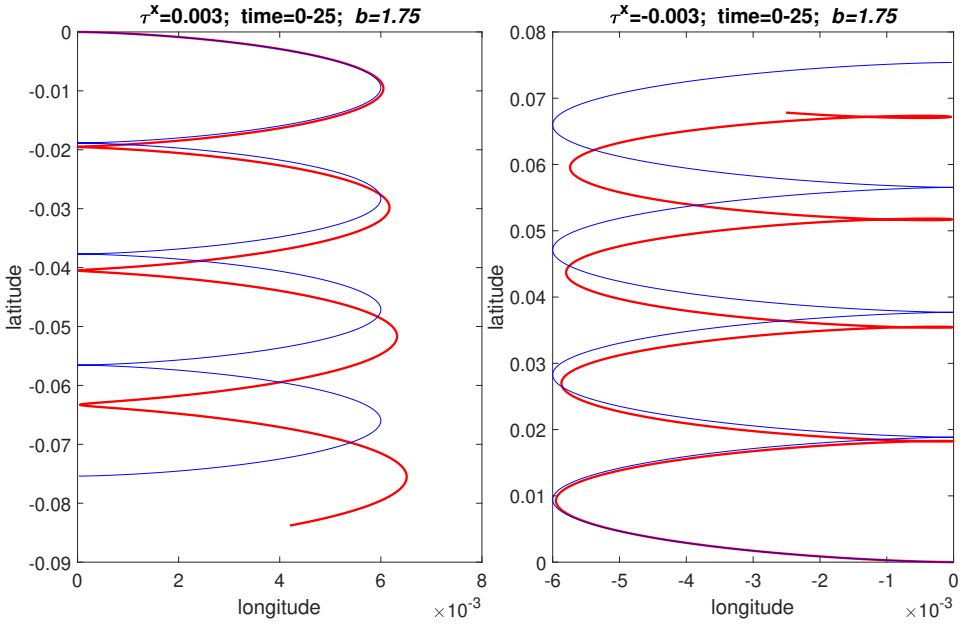


Figure 1. The (longitude, latitude) trajectories of water columns at the ocean surface subject to westward directed (right panel) and eastward directed (left panel) wind stress on the f-plane (blue curves) and on the β-plane (red curves). The time unit is the inverse of the mean Coriolis frequency and the longitude and latitude distances are scaled on Earth's radius. The value of b (scaled β) corresponds to 30° latitude. The scaling of the wind stress (τ^x) is detailed in section 2. Both trajectories start from $(x, y) = (0, 0)$ located at the bottom-right point in the right panel and at the upper-left point in the left panel

2 The Nondimensional Model

The time-dependent trajectory of a column of water in the surface Ekman layer forced by the overlying uniform wind stress on the f-plane is a fundamental problem of Physical Oceanography that is fully described in most textbooks (Gill, 1982; Pedlosky, 1987; Vallis, 2017). The governing Lagrangian equations that describe the dynamics of vertically ~~integrated~~ averaged horizontal velocity components consist of the momentum equations in the zonal and meridional directions and the (trivial) relations between these velocity components and the changes in the coordinate of the moving column i.e.:

$$\frac{dx}{dt} = U, \quad \frac{dy}{dt} = V, \quad \frac{dU}{dt} = fV + \frac{\tau^x}{\rho}\frac{\tau^x}{H\rho}, \quad \frac{dV}{dt} = -fU. \tag{1}$$

Here τ^x is the uniform zonally directed wind stress (which is positive/negative for eastward/westward directed wind, respectively), ρ is the water density, H is the depth (thickness) of the layer, $f = f_0 + \beta y$ is the Coriolis parameter (where $f_0 = 2\Omega\sin(\phi_0)$, $\beta = 2\Omega\cos(\phi_0)/$~~$a$~~ with ~~$a$~~ R_e with R_e and Ω – Earth's radius and rotation frequency, respectively and ϕ_0 – the latitude where the plane is tangential to Earth), U and V are the vertically ~~integrated~~ averaged horizontal velocity components in the eastward and northward directions, respectively, and x and y are the respective coordinates in these directions. The only

added complication of this system relative to that studied in details in e.g. chapter 9 of Gill (1982) is that here the Coriolis frequency, f, in the momentum equations is y-dependent.

The 4-dimensional system (1) can be easily integrated numerically but the general properties of its solutions can be best deciphered by reducing the number of its free parameters. This is done by scaling time, t, on $\frac{1}{f_0}$, x and y on ~~a~~ R_e so the velocity scale is ~~$f_0 a$ and the scale for the vertically integrated transport U and V is $f_0 a H$ where H is the depth (thickness) of the Ekman layer~~ $f_0 R_e$. With this scaling the nondimensional Coriolis frequency is $1 + by$ where ~~$b = \frac{\beta a}{f_0} = \cot(\phi_0)$~~ $b = \frac{\beta R_e}{f_0} = \cot(\phi_0)$ is the nondimensional β. The system is further simplified by replacing U by the pseudo angular momentum, defined as $D = U - y(1 + \frac{b}{2}y)$ in nondimensional units. As was shown by Paldor (2007) when $\tau^x = 0$ i.e., in the Inertial case, D is conserved. We note that in spherical coordinates the conservation of angular momentum, which is the spherical counterpart of D, yields a simple relation between the zonal velocity and the latitude (Paldor, 2001). Formally, a similar quantity relating the zonal velocity, U, and the meridional coordinate, y, can also be derived in Cartesian coordinates but, unlike spherical coordinates, this conserved quantity is not the angular momentum. With these changes system (1) transforms to:

$$\frac{dx}{dt} = D + y(1 + \frac{b}{2}y), \tag{2}$$

$$\frac{dy}{dt} = V, \tag{3}$$

$$\frac{dD}{dt} = \Gamma, \tag{4}$$

$$\frac{dV}{dt} = -(1 + by)(D + y(1 + \frac{b}{2}y)). \tag{5}$$

Here t, x, y and V denote the nondimensional counterparts of the dimensional variables denoted by the same symbols in system (1) and, as explained above, $D = U - y(1 + \frac{b}{2}y)$ is the nondimensional pseudo angular momentum. Equation (4) confirms that D is indeed conserved when $\Gamma = 0$. The solutions of this system are determined by the 4 required initial conditions and the 2 parameters: ~~$b = \frac{\beta a}{f_0} =$~~ $b = \frac{\beta R_e}{f_0} = \cot(\phi_0)$ – the nondimensional β and ~~$\Gamma = \frac{\tau^x}{\rho f_0^2 a H}$~~ $\Gamma = \frac{\tau^x}{H \rho f_0^2 R_e}$ – the constant, nondimensional, surface wind stress. The value of b at $\phi_0 = 30°$ is 1.75 and for realistic values of ~~$\tau^x/\rho \approx 2 \times 10^{-4} m^2 s^{-2}$~~ $\tau^x/\rho \approx 2 \times 10^{-4} m^2 s^{-2}$, $f_0 = 10^{-4} s^{-1}$ ~~and~~ , $H = 30m$, $\Gamma = 10^{-3}$ so the theory should be applicable to b of $O(1)$ and $\Gamma \ll 1$. The sign of Γ is that of τ^x – positive for eastward directed stress and negative for westward directed stress.

We solve this system by starting at the origin of the β-plane, i.e. $x(0) = y(0) = 0$ and assume that the initial $V(0)$ and $D(0) = U(0)$ are sufficiently small. The numerical solutions presented below are initiated with $D(0) = 0$ and $V(0) \neq 0$. However, the definition of D implies that trajectories emanating from $D(0) \neq 0$ and $y(0) = 0$ can also be calculated starting from $D(0) = 0$ and a suitable $y(0) \neq 0$. Note that the choice $D(0) = 0$ does not restrict the generality of our solutions since the shift of time from t to $t' = t + D(0)/\Gamma$ yields $D(t' = 0) = 0$ so $D(0) = 0$ can be assumed. The analysis of the solutions of system (2) – (5), including numerical examples, are presented in the next section.

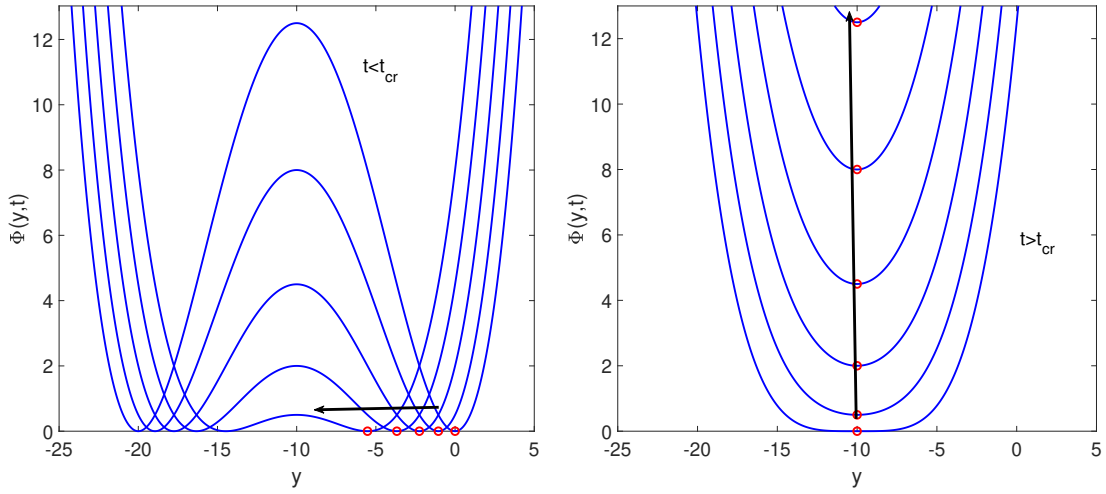


Figure 2. ~~The~~ A schematic demonstration of the change in the potential $\Phi(y,t)$ for $b = 0.1$ and $\Gamma = 0.1$ at $t = 0, 10, 20, ..., 100$. The direction of increase in time is indicated by the ~~green~~ black arrows for $t < t_{cr}$ (left panel) and $t > t_{cr}$ (right panel). The minima, y_m^+, of the potentials are indicted by red circles.

3 Analysis

The analysis of system (2)-(5) begins with the (V, y) subsystem, i.e. equations (3) and (5) along with the (trivial) solution $D = \Gamma t$ of (4). The derived solution of $y(t)$ will then be substituted in Eq. (2) to yield the zonal propagation speed. First, we combine equations (3) and (5) to the single second-order equation

$$\frac{d^2y}{dt^2} = -(1+by)\left[D + y(1+\frac{b}{2}y)\right]. \tag{6}$$

We will discuss solutions of this equation for initial conditions in the vicinity of $y = dy/dt = V = 0$ and assume that Γ is sufficiently small [for the smallness condition see equation (A3) in the Appendix]. We proceed by rewriting Eq. (6) as

$$\frac{d^2y}{dt^2} = -\frac{\partial\Phi(y,t)}{\partial y} \tag{7}$$

where

$$\Phi(y,t) = \frac{1}{2}\left[\Gamma t + y\left(1+\frac{1}{2}by\right)\right]^2. \tag{8}$$

Equation (7) describes the dynamics of a quasi-particle in a slowly (for small Γ) time varying quasi-potential well $\Phi(y,t)$. In figure 2 we illustrate this potential for $\Gamma = b = 0.1$ at times t = 0, 10, 20, ..., 100. The minima of these potentials, denoted collectively by y_m, are given by the 3 roots of:

$$\left.\frac{\partial\Phi}{\partial y}\right|_{y=y_m} = \left[\Gamma t + y_m\left(1+\frac{1}{2}by_m\right)\right](1+by_m) = 0. \tag{9}$$

Two cases should be considered depending on time being below or above the critical time

$$t_{cr} = \frac{1}{2b\Gamma}. \tag{10}$$

For $t < t_{cr}$, there exist two minima defined by $\Gamma t + y_m\left(1 + \frac{1}{2}by_m\right) = 0$, i.e.,

$$y_m^{\pm} = \frac{1}{b}(-1 \pm \sqrt{1 - 2b\Gamma t}) \tag{11}$$

while for $t > t_{cr}$, there exists a single minimum located at

$$y_m^0 = -\frac{1}{b}. \tag{12}$$

Figure 3 shows the numerical solutions of $y(t)$ when the column originates near y_m^+ i.e. $x = y = D = 0$ and $V = 0.002$. As predicted, the exact solution of $y(t)$ oscillates with small amplitude (that increases with the value of $V(t=0)$) about the evolution curves of y_m^+ and y_m^0 shown by the black curves. The direction of evolution of ~~y_m^+ in time~~ $y_m^+(t)$ and its transformation to ~~y_m^0 (~~the constant $y_m^0(t)$ at $t = t_{cr}$ correspond to the evolution of the red circles in figure 2 ~~) are indicated by the green arrows~~ for $t < t_{cr}$ (left panel) and $t > t_{cr}$ (right panel)~~in figure 2.~~. The averaged numerical solution is expected to deviate appreciably from the simple scenario shown here only for high oscillation amplitude near y_m^+.

The main idea of the following analysis is that since the system starts *near* $y = dy/dt = V = 0$, i.e. near the minimum of the potential, y_m^+, by the adiabatic theory (see pages 531-535 in Goldstein, 1980) it will *stay near* this minimum for $t < t_{cr}$. At $t = t_{cr}$, y_m^+ transforms into y_m^0 and therefore at all $t > t_{cr}$, the system remains near y_m^0. Thus, the column remains near the minimum of Φ at all times, while this minimum is slowly decreasing for $t < t_{cr}$ and stays constant for $t > t_{cr}$. Since the trajectory originates near the minimum y_m^+ and since for small Γ the variation of the potential is slow (see Eq. (9)), we expect the solution for y to be of the form

$$y = y_m(t) + \delta y \tag{13}$$

where $y_m(t)$ starts at y_m^+ and later (i.e. at $t = t_{cr}$) transforms into y_m^0 and δy is a small perturbation. We substitute this form of solution into Eq. (6) and rewrite the resulting equation as

$$\frac{d^2\delta y}{dt^2} = F - \omega_0^2(t)\delta y - A\delta y^2 - B\delta y^3 \tag{14}$$

where $F = -d^2y_m/dt^2$ is an inhomogeneous forcing term and the coefficients of the other 3 terms on the RHS of this equation are:

$$\omega_0^2(t) = \begin{cases} (1 + by_m^+)^2 = 1 - 2b\Gamma t, & t < t_{cr} \\ b\Gamma t - 1/2, & t > t_{cr} \end{cases} \tag{15}$$

$$A = \begin{cases} (3/2)b\omega_0, & t < t_{cr} \\ 0, & t > t_{cr} \end{cases} \tag{16}$$

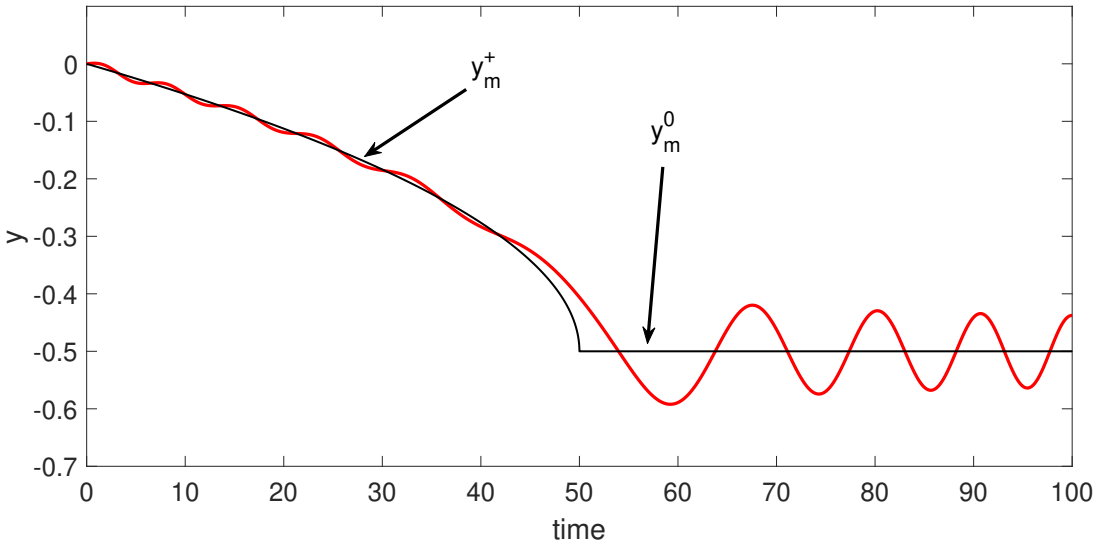


Figure 3. Numerical solutions of $y(t)$ in system (2) - (5) starting from ~~$x = y = V = D = 0$ (blue curves)~~ $x = 0 = y = D$ and ~~from $x = y = D = 0$, but $V = 0.05$ (red curves) for $b = 2$ and $\Gamma = 0.005$~~$V = 0.002$. The ~~green~~ black curve shows the evolution of ~~$y_m(t)$~~$y_m^+(t)$ given by equation (11) for $t < t_{cr}$ and of$y_m^0(t)$ given by equation (12) for $t > t_{cr}$. The values of $b = 2$ and $\Gamma = 0.001$ used here imply that the change between the two approximate solutions occurs at $t_{cr} = \frac{1}{2b\Gamma} = 50$.

and

$B = \frac{1}{2}b^2$.

In the present model, equation (11) ~~implies~~ imply $d^2y_m^+/dt^2 = -b\Gamma^2(1-2b\Gamma t)^{-3/2}$ so ~~$F = -d^2y_m^+/dt^2 > 0$~~ according to equation (15) $F = -d^2y_m^+/dt^2 = b\Gamma^2/\omega_0^3 > 0$ for $t < t_{cr}$. The second term on the RHS of Eq. (14) describes linear oscillations having slowly varying frequency $\omega_0(t)$, while the third and fourth terms represent the effect of small anharmonicity of the potential well near the minimum. Note that for $y_m = y_m^+$ the term $y_m^+(t)$ in Eq. (13) describes slow monotonic variation of the latitude shown by the ~~green curves in our example in~~ black arrows in figure 2 at $t < t_{cr}$. No such variation exists at $t > t_{cr}$ since then $y_m = y_m^0 = const$. As will be shown below, the nonlinear terms in (14) mostly affect the zonal drift in x.

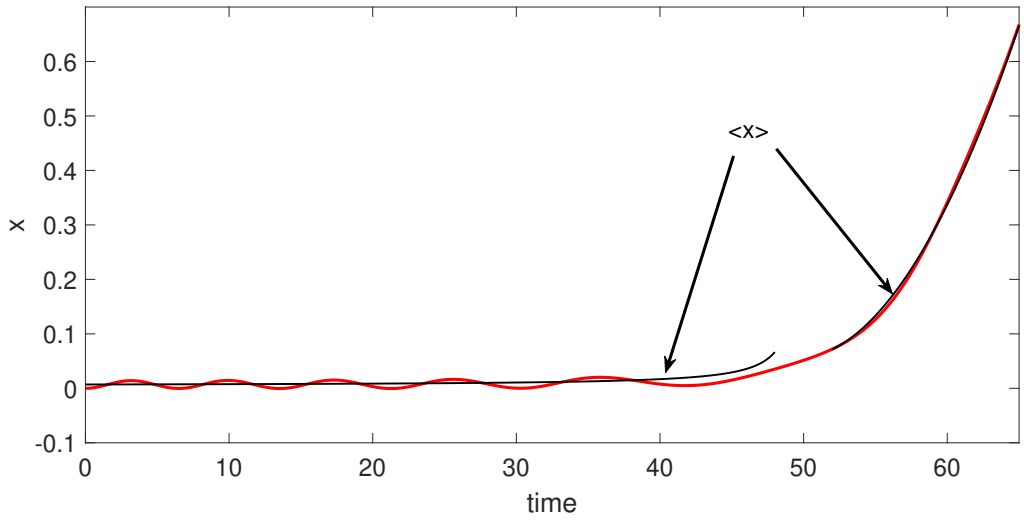
Importantly, for constant parameters ω_0, A and B the solution of Eq. (14) can be found in textbooks (see e.g. pages 86-87 in Landau and Lifshitz, 1982) and it has the form

$$\delta y = \frac{F}{\omega_0^2} + a\cos\psi - \frac{Aa^2}{2\omega_0^2} + \frac{Aa^2}{6\omega_0^2}\cos(2\psi) + O(a^3), \tag{17}$$

where $\psi = \omega t + \phi_0$, (ϕ_0 takes into account initial conditions~~,~~), a is the amplitude of the linear part of the δy and

$$\omega = \omega_0 + \left(\frac{3B}{8\omega_0} - \frac{5A^2}{12\omega_0^3}\right)a^2. \tag{18}$$

Therefore, δy includes harmonic oscillations of amplitude a and $O(a^2)$ corrections and oscillation frequency ω (that includes an $O(a^2)$ correction to ω_0). As is shown in the Appendix, when ω_0 is a slow function of time as in our case [$d\omega_0/dt \sim O(\Gamma)$],



~~Figure 3 shows the numerical solutions of $y(t)$ for two slightly different initial conditions. As predicted, both the blue curve (column initiated with $x = y = V = D = 0$) and red curve (column initiated with $x = y = D = 0$ and $V = 0.05$) oscillate about the evolution curve of y_m (green curve). The results show that away from the equator (i.e. for~~ (20) for $t < t_{cr}$ ~~) the amplitude of oscillations is larger when $V \neq 0$~~and by equation (21) for $t > t_{cr}$ where $t_{cr} = 50$ as in figure 3. The black curves terminate near the equator where the adiabaticity breaks down since the frequency, ω_0, tends to 0 there according to equation (15)

~~Figure 3 shows the numerical solutions of $y(t)$ for two slightly different initial conditions. As predicted, both the blue curve (column initiated with $x = y = V = D = 0$) and red curve (column initiated with $x = y = D = 0$ and $V = 0.05$) oscillate about the evolution curve of y_m (green curve). The results show that away from the equator (i.e. for~~ (20) for $t < t_{cr}$ ~~) the amplitude of oscillations is larger when $V \neq 0$~~and by equation (21) for $t > t_{cr}$ where $t_{cr} = 50$ as in figure 3. The black curves terminate near the equator where the adiabaticity breaks down since the frequency, ω_0, tends to 0 there according to equation (15)

Figure 4. Numerical solutions of $x(t)$ in system (2) - (5) starting from the same initial conditions as in figure 3 ~~for the blue and red curves~~. The ~~green~~ two black curves show the monotonic evolution ~~(averaged over oscillations)~~ of x described by ~~the theory developed here for two different values of $a(t = 0)$~~Eq.

~~Figure 3 shows the numerical solutions of $y(t)$ for two slightly different initial conditions. As predicted, both the blue curve (column initiated with $x = y = V = D = 0$) and red curve (column initiated with $x = y = D = 0$ and $V = 0.05$) oscillate about the evolution curve of y_m (green curve). The results show that away from the equator (i.e. for~~ (20) for $t < t_{cr}$ ~~) the amplitude of oscillations is larger when $V \neq 0$~~and by equation (21) for $t > t_{cr}$ where $t_{cr} = 50$ as in figure 3. The black curves terminate near the equator where the adiabaticity breaks down since the frequency, ω_0, tends to 0 there according to equation (15)

∴

the solution (17) remains the same, but ψ is replaced by ~~$\psi = \int \omega dt + \phi_0$~~ $\psi = \int \omega dt + \phi_0$ and the oscillation's amplitude a becomes a slow function of time such that $\omega a^2 = const$.

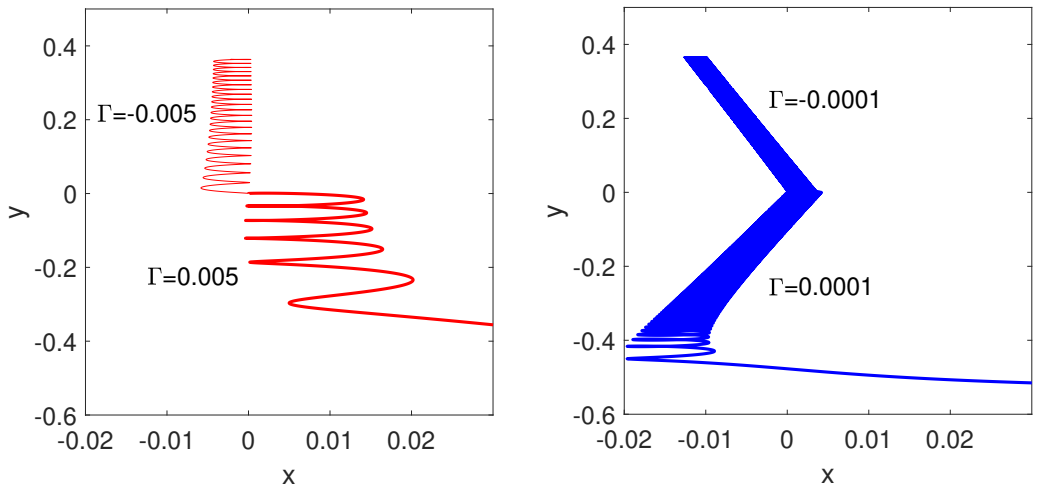


Figure 5. The different water column trajectories for $b = 2$, initial conditions $x = y = D = 0$ and ~~different initial velocities, V~~$V = 0.002$. Left panel (~~blue~~ red curves): ~~$V(t = 0) = 0.00$~~large Γ^2; Right panel (~~red~~ blue curves): ~~$V(t = 0) = 0.05$~~small Γ^2. The values of Γ are noted near each of the curves: thin curves denote negative (westward directed) stresses and thick curves denote positive (eastward directed) stresses. The integration time is $2t_{cr}$ in all cases but the curves terminate just prior to reaching the equator. Note that for $\Gamma = 0.00001$ the integration time, $\frac{1}{b\Gamma}$, is 5000 i.e. the columns in the right panel complete several thousand oscillations in the course of the integration

This completes our solution for the latitude, y, and we proceed to the longitudinal dynamics.

The dynamics in the zonal direction, x, is governed by Eq. (2) which after substitution of (13) becomes

$$\frac{dx}{dt} = \left[D + y_m(1 + \frac{b}{2}y_m)\right] + (1 + by_m)\delta y + \frac{b}{2}\delta y^2. \tag{19}$$

Here again we consider two cases. For $t < t_{cr}$, $D + y_m^+(1 + \frac{b}{2}y_m^+) = 0$ and, therefore, by averaging ~~in time (i.e. neglecting oscillatory components due to δy)~~ locally in time over a single oscillation and using Eqs. (17) and (15) we get

$$\frac{d\langle x\rangle}{dt} = -\frac{ba^2}{2} + \frac{F}{\omega_0}. \tag{20}$$

This equation shows that the average zonal drift is a nonlinear phenomenon in terms of the amplitude of oscillation. Since, as was shown above, $F > 0$ for $t < t_{cr}$, the drift is determined by the balance between $ba^2/2$ ~~(determined by the initial displacement from the fixed point $V = 0, D = 0, y = y_m^+$) and F ($\propto \Gamma^2/\omega_0$)~~and $F/\omega_0 = b\Gamma^2/\omega_0^4$. Thus, the sign (direction) of the zonal drift is independent of the sign of Γ. ~~In contrast, for~~

For $t > t_{cr}$, $y_m^0 = -1/b$, so $D + y_m^0(1 + \frac{b}{2}y_m^0) = \Gamma t - 1/(2b)$ and therefore,

$$\frac{d\langle x\rangle}{dt} = \Gamma t - \frac{1}{2b} + \frac{1}{4}ba^2. \tag{21}$$

Figure 3 displays numerical solutions of $x(t)$ in system (2) - (5) starting from the same initial conditions as in figure 3~~for the blue and red curves. The green~~ . The black curves show the monotonic evolution (averaged over oscillations) of x

described by the theory developed here~~for two different values of $a(t=0)$. As predicted, the~~. Since the trajectories originate in mid-latitudes, a westward directed wind-stress will always stir the trajectories away from the equator so Γ in equation (21) must be positive i.e. the long-term zonal drift on the equator ~~is positive and monotonic~~has to be directed eastward.

Figure 5 compares the $(x(t), y(t))$ trajectories emanating from ~~different initial velocities V for~~ $x = 0 = D = y$ and $V = 0.002$ for two pairs of eastward directed (~~solid~~ thick curves) and westward directed (thin curves) wind stresses~~of identical magnitude . The various curves clearly demonstrate~~. In the right panel the magnitude of the wind stress is small (=0.0001) and in the left panel the magnitude of the wind tress is large (=0.005). The two curves in each panel demonstrate that, as concluded above, the zonal drift is independent of the ~~effect of both the wind stress directionand the initial velocity. Trajectories emanating with $V = 0$ (left panel) are very similar to those on the f-plane until one reaches the equator (which is completely missing from the f-plane dynamics) where a trajectory consists of a steady translation and oscillations~~sign (direction) of the wind stress (since according to equation (20) it is proportional to $F \propto \Gamma^2$). A comparison between the trajectories in the two panels shows that for tiny wind stress (right panel) the trajectories drift westward as in the force-free, inertial, oscillations while with the increase in the magnitude of the wind-stress (left panel) the zonal drift is directed eastward. In accordance with the intuits presented in the Introduction on the f-plane solution the oscillation's (inertial) frequency changes with latitude i.e. increasing/decreasing in northward/westward directed trajectories while the oscillation's amplitude follow the opposite pattern. ~~As expected, the zonal drift of the oscillations in the right panel of figure 5 is independent of the sign of Γ as it is determined by the balance between the initial deviation from $V = 0$ and Γ^2.~~

4 Discussion and Summary

The two simple limits of $b = 0$ (Ekman transport on the f-plane) and $\Gamma = 0$ (inertial trajectories on the β-plane) should be discussed as special cases of the ~~involved theorypresented here~~present theory. These limits are well known in physical oceanography but they were never presented as limits of a single dynamical system.

In the $b = 0$ limit (wind forced transport on the f-plane) the potential in (8) becomes $\Phi(y) = \frac{1}{2}(D + y)^2$ (recall: $D = \Gamma t$). This potential has a single minimum at $y^f = -D$ and the frequency of oscillation near this point is $\omega^f = 1$. Near y^f the potential, Φ, is identical to that of Harmonic Oscillator: $\frac{1}{2}(y - y^f)^2$. ~~The substitution $b = 0$ leaves equation unchanged so $D = \Gamma t$ i.e.~~ Thus, the minimum $y^f = -D$ must decrease (or increase, depending on the sign of Γ) indefinitely at ~~the same rate as~~ a rate that equals Γ ~~. Thus~~ i.e. the potential Φ simply translates in the $+y$ or $-y$ directions without changing its shape.

The $\Gamma = 0$ limit (inertial trajectories on the β-plane) implies, according to equation (4), that D is conserved ~~i.e. it is time-independent. Thus,~~ so system (2) - (5) has two conserved quantities – D and the energy – E. With the increase in the initial energy (say by increasing $V(t = 0)$) the inertial trajectory will oscillate in (V, y) while drifting westward (see Ripa, 1997; Paldor, 2007) as on the sphere (Paldor, 2001). Equation (20) and the trajectories shown in figure 5 show that the long-term westward drift on the β-plane when $\Gamma \neq 0$ is slower than in the inertial, $\Gamma = 0$, case and it is independent of the sign of Γ.

The solutions of the nonlinear system (2) - (5) are determined by the 2 initial conditions ~~($x = 0 = D$~~ $V(t=0)$ and $y(t=0)$ (recall: $x(t=0) = 0 = D(t=0)$ can be assumed without loss of generality since x does not affect the dynamics and D can be translated in time) and the values of the two parameters, b and Γ (that represent the dimensional parameters β and τ^x, respectively) for a total of 4 parameters! Thus, these solutions display a ~~wide~~ range of temporal evolution and this work describes and analyzes the general properties of these solutions and illustrates them in numerical examples. In particular, the westward drift of the trajectories can be eastward (as in figure 1 and the left panel of figure 5) or westward (as in the right panel of figure 5. The sensitive dependence of the drift on parameter values (including initial conditions) is a defining property of nonlinear systems such as that studied here.

The intent of the analysis in this work is to provide an overview of the complex phenomena that result from the extension of Ekman's theory to the β-plane. In particular, this work shows that the zonal drift is independent of the sign of τ^x but depends on a (previously unknown) balance between $V(0)^2$ (or the initial displacement from $(y_m^+)^2$) and $(\tau^x)^2$. The values of the parameters used in the numerical results presented here were chosen so as to highlight the phenomena being discussed while still being realistic. Thus, with the velocity scale of $f_0 R_e = 640 m/s$ the value of $V(0) = 0.002$ used in figures 3 – 5 corresponds to a dimensional velocity of about $1 ms^{-1}$. Trajectories of much higher oscillation amplitudes will be encountered with higher $V(0)$ values.

The symmetry between y_m^+ and y_m^- in the present theory suggests that for the same wind stress, Γ, the southern hemisphere's fixed point will also move towards the equator, i.e. northward. However, in all other respects the evolution near y_m^- is identical to that described above for y_m^+.

The importance of latitudes where the curl of the wind-stress vanishes, that play a fundamental role in (Stommel, 1948) vorticity based theory of wind driven ocean gyres, can not be captured in extensions of the present Lagrangian theory. However, extensions of the present new Lagrangian theory on the β-plane can include variable zonal wind stress, $\tau^x(y)$, which can highlight the role played by latitudes where the wind-stress itself vanishes. ~~Furthermore, the extension of the present study~~ The application of the concepts developed here to spherical geometry ~~using the concepts developed here~~ and to wind-driven circulation over the continental shelf (where the sloping bottom yields the topographic β-effect) is an interesting ~~and valuable focus of a future study~~goal for future studies.

Appendix A: Adiabatic evolution of meridional oscillations and initial conditions

In this appendix we discuss adiabatic (slow) evolution of linear longitudinal oscillations described by [see Eq. (14)]

$$\frac{d^2\delta y}{dt^2} = -\omega_0^2(t)\delta y + \frac{F}{\omega_0^2} \tag{A1}$$

and seek solution of this equation of form

$$\delta y = a(t)\cos\Big(\int_0^t \omega_0(t)dt + \phi_0\Big) + \frac{F}{\omega_0^2}. \tag{A2}$$

Here ϕ_0 is added to take into account initial conditions and we assume that the change of ω_0 during one period $2\pi/\omega_0$ of oscillations is small, i.e.

$$\frac{d\omega_0}{dt}\frac{2\pi}{\omega_0} \ll \omega_0 \tag{A3}$$

which is guaranteed if Γ is sufficiently small. This is our adiabaticity criterion. A similar condition, $\frac{da}{dt}\frac{2\pi}{\omega_0} \ll a$, is also assumed for the amplitude of oscillations. Next, we substitute (A2) into (A1) and neglect d^2a/dt^2 to get

$$2\frac{da}{dt}\omega_0 + a\frac{d\omega_0}{dt} = 0, \tag{A4}$$

yielding

$$\omega_0 a^2 = I = Const. \tag{A5}$$

The constant I (the action) is given by initial conditions. When the nonlinear terms in (14) are included in the analysis, all the derivation of weakly nonlinear solution as described in pages 86-87 of Landau and Lifshitz (1982) is not affected by the replacement of the linear component $a\cos(\omega t + \phi_0)$ by $a(t)\cos(\int_0^t \omega(t)dt + \phi_0)$ in the adiabatic problem which is the basis of solution (17) in section 3.

Finally, the action I, which remains constant all all times, can be calculated from the initial conditions, $\delta y(0)$ and $V(0) = d(\delta y)/dt|_{t=0}$. Using (A2) we have $\delta y(0) = a(0)\cos\phi_0 + F(0)/\omega_0^2(0)$ and $V(0) = -a(0)\omega_0(0)\sin\phi_0$. Then

$$a(0)^2 = \left(\delta y(0) - \frac{F(0)}{\omega_0^2(0)}\right)^2 + \left(\frac{V(0)}{\omega_0(0)}\right)^2 \tag{A6}$$

and

$$I = \omega_0(0)\left(\delta y(0) - \frac{F(0)}{\omega_0^2(0)}\right)^2 + \frac{V^2(0)}{\omega_0(0)}. \tag{A7}$$

The case depicted in figures 3 and 3 has $\delta y(0) = 0$, so one gets $a^2(0) = \frac{F^2(0)+V^2(0)}{\omega_0^2(0)}$ and $I = \omega_0(0)a^2(0)$.

Author contributions. The research on the problem was initiated by NP, who also proposed the transformation to the pseudo angular momentum while LF proposed the application of the adiabaticity theory. Both authors contributed equally to the numerical calculations and manuscript preparation.

Competing interests. The authors declare that they have no conflict of interests

Acknowledgements. The authors are happy to declare that no funding was used in this research.

References

Constantin, A. and Johnson, R.: Ekman-type solutions for shallow-water flows on a rotating sphere: a new perspective on a classical problem, Physics of Fluids, 31, 021 401, 2019.

Ekman, V. W.: On the influence of the earth's rotation on ocean-currents., Ark. Mat. Astr. Fys., 2, 1905.

Gill, A. E.: Atmosphere-ocean dynamics, vol. 30, Academic Press, 1982.

Goldstein, H.: Classical mechanics, Eddison-Wesley Publishing Company, 1980.

Landau, L. D. and Lifshitz, E. M.: Mechanics, Butterworth-Heinemann, Oxford, England, 3 edn., 1982.

Paldor, N.: The zonal drift associated with time-dependent particle motion on the earth, Quarterly Journal of the Royal Meteorological Society, 127, 2435–2450, 2001.

Paldor, N.: The transport in the Ekman surface layer on the spherical Earth, Journal of marine research, 60, 47–72, 2002.

Paldor, N.: Inertial particle dynamics on the rotating earth, Lagrangian Analysis and Prediction of Coastal and Ocean Dynamics, 119, 135, 2007.

Pedlosky, J.: Geophysical fluid dynamics, vol. 710, Springer, 1987.

Ripa, P.: "Inertial" Oscillations and the β-Plane Approximation (s), Journal of physical oceanography, 27, 633–647, 1997.

Stommel, H.: The westward intensification of wind-driven ocean currents, Eos, Transactions American Geophysical Union, 29, 202–206, 1948.

Vallis, G. K.: Atmospheric and oceanic fluid dynamics, Cambridge University Press, 2017.

Von Arx, W. S.: An introduction to physical oceanography, USA, 1964.