Method

Eulerian Salt Budget and the Decompositions

The Lagrangian budget is based on the Eulerian volume integrated budget defined in the ECCOv4r4 product (I want to cite this note https://ecco-group.org/docs/evaluating_budgets_in_eccov4r3_updated_20220118.pdf but don't know how to, may be just Forget 2015?). The volume integrated budget equations are defined by discrete partial differential equations. In this study, we utilize the salt budget

$$V\frac{\Delta_E(z^*S)}{\Delta_E t} + \nabla_E \cdot (\mathscr{F}_{adv} + \mathscr{F}_{dif}) = VF_{salt} \tag{1}$$

and the volume budget

$$V\frac{\Delta_E z^*}{\Delta_E t} + \nabla_E \cdot (U_m) = P_{fr} \tag{2}$$

Where $\frac{\Delta_E *}{\Delta_E t} = \frac{*_{j+1} - *_j}{t_{j+1} - t_j}$ is the discrete time derivative, and $\nabla_E \cdot * = \Sigma_{x,y,z}(*_{i+0.5} - *_{i-0.5})$ is the discrete divergence of a vector. $z^* = \frac{z_D + \eta}{z_D}$ is the non-dimensional thickness explaining how much of the mean water column depth z_D is changed by sea surface height η change. S is the volume mean salinity, V is the average volume of the grid cell, $\mathscr{F}_{adv}, \mathscr{F}_{dif}$ are model the diagnosed advective and diffusive salt flux, respectively.U_m and P_{fr} are mass flux from Eulerian advection and freshwater forcing respectively.

The ECCO simulation uses the Gent-McWilliams advective parameterization, which adds another divergent-free mass transport (see supplementary material (bolus divergence handling)). We can reconstruct the mass budget equation 2 to use the total mass transport instead of only the Eulerian component.

$$\frac{\Delta_E z^*}{\Delta_E t} + \nabla_E \cdot U = P_{fr} \tag{3}$$

Where $U = U_m + U_b$ is the total mass flux. The advective flux $\mathscr{F}_{adv}$ in the salt budget equaiton 1 can be expressed as a function of the total mass flux.

$$\mathscr{F}_{adv} = US_w \tag{4}$$

Here, S_w is the interpolated salinity defined on the walls of grid cells. Its values are determined by the salinity in neighboring grid cells and the specific advection schemes. For all MITgcm advection schemes without flux limiter, the wall tracer concentration is a linear function of the concentration of neighboring concentration, and can be seen as an interpolation. ECCOv4r4 uses the third order upwind bias advection scheme in the vertical and the third order direct space-time scheme in the horizontal. For example, the interpolation scheme associated with the latter gives

$$S_w = S_{i-1} + d_0(S_i - S_{i-1}) + d_1(S_{i-1} - S_{i-2}), \forall U > 0; \tag{5}$$

and

$$S_w = S_i - d_0(S_i - S_{i-1}) - d_1(S_{i+1} - S_i), \forall U \leq 0; \tag{6}$$

where

$$d_0 = \frac{1}{6}(2 - |c|)(1 - |c|); d_1 = \frac{1}{6}(1 - |c|)(1 + |c|) \tag{7}$$

and c is the Courant number defined by the velocity and model temporal and spatial resolution as

$$c = \frac{U\Delta t}{\Delta x}. \tag{8}$$

Using the seaduck.topology module, we can find a interpolated salinity $S_{w,i}$ defined at cell walls by plugging in daily mean ECCO velocity and salinity and. Due to the non-linear nature of the advection term and the fact that we are using daily-mean data, we cannot recover the full advective flux by U and $S_{w,i}$, despite being a very good approximation. The difference, or the contribution of unresolved sub-daily advection, can be diagnosed by

$$F_{ua} = \frac{1}{t_{i+1} - t_i}\left(S_{w,i}\int_{t_i}^{t_{i+1}} U dt - \int_{t_i}^{t_{i+1}} \mathscr{F}_{adv} dt\right). \tag{9}$$

With the advection expressed as the product of mass flux and salinity plus a sub-daily correction, we can rewrite the salt budget equation as

$$V\frac{\Delta_E(z^*S)}{\Delta_E t} + \nabla_E \cdot (US_{w,i}) = V(F_{salt} + F_{dif,v} + F_{dif,h} + F_{ua}) \tag{10}$$

Subtract the product of volume mean salinity and the mass budget equation from this new form of salt budget equation we get

$$V\left(\frac{\Delta_E(z^*S)}{\Delta_E t} - S\frac{\Delta_E z^*}{\Delta_E t}\right) + g(U, S) = V(F_{salt} + F_{dif,v} + F_{dif,h} + F_{ua} + F_{fr}) \tag{11}$$

Where $F_{fr} = \frac{P_{fr}S}{V}$ is the effect of freshwater forcing that involves river runoffs,cryosphere, and most importantly the atmosphere. $g(U,S) = [\nabla_E \cdot (US_{w,i}) - S\nabla_E \cdot U]$ is equivalent to the familiar form of divergent-free advection $u \cdot \nabla s$. It will later be shown in supplementary material that the salinity change felt by the Lagrangian particles is equivalent to this term. $\left(\frac{\Delta_E(z^*S)}{\Delta_E t} - S\frac{\Delta_E z^*}{\Delta_E t}\right) \approx \frac{\Delta_E S}{\Delta_E t}$ is the approximated temporal salinity derivative. This is because $\frac{z^*_{j+1}-z^*_{j+1}}{z^*_j}$ and $\frac{S_{j+1}-S_{j+1}}{S_j}$ are both small, and $z^* \approx 1$. This approximation is well justified because, as we will see in (subsection: lhs interpolation), the difference is much smaller than other errors and the terms of interest in this study.

Now, we are going to apply a temporal averaging over the duration of the ECCO period (1992-2017) to both sides of the equation.

$$V\overline{\left(\frac{\Delta_E(z^*S)}{\Delta_E t} - S\frac{\Delta_E z^*}{\Delta_E t}\right)} + g(\overline{U},\overline{S}) + \overline{g(U',S')} = V(\overline{F_{salt}} + \overline{F_{dif,v}} + \overline{F_{dif,h}} + \overline{F_{ua}} + \overline{F_{fr}}) \tag{12}$$

Where $*' = * - \overline{*}$ is the temporal anomaly, and $g(U',S') = \nabla_E \cdot (U'S'_{w,i}) - S'\nabla_E \cdot U'$.Subtract the mean equation from the original equation we get the anomaly equation

$$V\left(\frac{\Delta_E(z^*S)}{\Delta_E t} - S\frac{\Delta_E z^*}{\Delta_E t}\right)' + g(U,S') = -g(U',\overline{S}) + \overline{g(U',S')} + V(F'_{salt} + F'_{dif,v} + F'_{dif,h} + F'_{ua} + F'_{fr}) \tag{13}$$

We also use a slightly different form of the above equation.

$$V\left(\frac{\Delta_E(z^*S)}{\Delta_E t} - S\frac{\Delta_E z^*}{\Delta_E t}\right)' + g(\overline{U},S') = -g(U',\overline{S}) - \left(g(U',S') - \overline{g(U',S')}\right) + V(F'_{salt} + F'_{dif,v} + F'_{dif,h} + F'_{ua} + F'_{fr}) \tag{14}$$

In 13 and 14, the LHS represent the salinity change felt by particles advected by time-dependent velocity and mean velocity, respectively.

Interpolation schemes and the Lagrangian salt budget

We use the python package seaduck [@seaduck] developed by the authors to simulate the particles and conduct budget analysis on them. Lagrangian particle trajectories, in the absence of stochastic diffusion, can be described by the Lagrangian flow map

$$\frac{dx}{dt} = u(x(a,t),t) \tag{15}$$

$$a = x(a, 0) \tag{16}$$

Following [@tracmass_mass], we use the mass/volume conserving "piecewise-stationary" scheme to interpolate the velocity and simulate the particles.In each grid cell, each velocity/mass flux component is linear interpolated along the component using daily mean mass flux defined on the cell walls. During every output time step (one day), the velocity is assumed to be the unchanged.

$$u(x,t) = U_{i-1,j} + (x^* - x^*_{i-1})(U_{i,j} - U_{i-1,j}), t_{j-0.5} < t < t_{j+0.5} \tag{17}$$

where $x^* = x/\Delta x$ is the spatial coordinate in a grid cell non-dimensionalized by the size of the grid cell. Although the mass flux used in ECCO is not fully divergent free (see equation mass budget), the divergent is very small and it is localized at the surface. A Boussinesq style approximation is used here: we assume the velocity field conserves mass/volume (see), and meanwhile preserve the changes to salinity due to volume change (freshwater forcing). Since the advective salt flux and mass flux are both piecewise stationary functions in time, the salinity interpolation needs to be piecewise in time as well. Although the details are determined by the salinity interpolation scheme which will be introduced in subsection , the salinity changed can be described in a general form as

$$\Delta s = \Sigma_i \int_{t_{t-0.5}}^{t_{i+0.5}} (u \cdot \nabla s) dt + \Sigma_{i+0.5} \Delta s_{i+0.5} \tag{18}$$

The first term represent the salinity change felt by the particles during each piecewise stationary period, during which the Eulerian salinity is constant.This part of salinity change is purely from the particle movement, i.e. advection. The second term is the Eulerian tendency term. All salinity changes happen in infinitely short time, and therefore does not involve movement of particles.

For a closed Lagrangian description of the salinity budget in between temporal jumps, we have

$$u \cdot \nabla s = -\frac{\partial s}{\partial t} + \Sigma_{i \in RHS} f_i \tag{19}$$

we put the $-\frac{\partial s}{\partial t}$ to the right as the "suppressed Eulerian tendency term". The Lagrangian scheme we are going to introduce in the following subsections has this Lagrangian budget 19 equation exactly closed. As we will show in later subsection , the cumulative effect of the suppressed Eulerian tendency term also roughly cancels out that of the salinity jump.

$$0 \approx \Sigma_i \int_{t_{t-0.5}}^{t_{i+0.5}} (-\frac{\partial s}{\partial t})dt + \Sigma_{i+0.5}\Delta s_{i+0.5} \tag{20}$$

Note that we can replace u, s with their decomposed components. Two alternative forms we used are the one focusing on salinity anomaly advection

$$\Delta s' = \Sigma_i \int_{t_{t-0.5}}^{t_{i+0.5}} (u \cdot \nabla s')dt + \Sigma_{i+0.5}\Delta s'_{i+0.5} \tag{21}$$

which correspond to 13 and the salinity anomaly advection by mean velocity

$$\overline{\Delta} s' = \Sigma_i \int_{t_{t-0.5}}^{t_{i+0.5}} (\overline{u} \cdot \nabla s')dt + \Sigma_{i+0.5}\Delta s'_{i+0.5} \tag{22}$$

which correspond to 14. The interpolation schemes we used for every terms in the salinity budget equation are consistent with the salt equation we described earlier. That is to say

$$\frac{1}{V}\int f_i dv = F_i, i \in RHS; \tag{23}$$

$$\frac{1}{V}\int \frac{\partial s}{\partial t} dv = \left(\frac{\Delta_E(z^* S)}{\Delta_E t} - S\frac{\Delta_E z^*}{\Delta_E t}\right) \tag{24}$$

$$\frac{1}{V}\int (u \cdot \nabla s)dV = g(U, S) \tag{25}$$

As a result, the Lagrangian scheme we are going to lay out in detail can be seen as a equivalent description of the finite-volume model's salt budget, aside from the mass/volume budget.

Salinity, and advection interpolation

The salinity on each cell walls are prescribed using the interpolation scheme defined by the model advection scheme, as described in the Eulerian budget (e.g. 5). Because neighboring grid cells share the same cell wall, this also ensures that there is no spatial jumps in salinity in between the grid cells. In the interior, the interpolation is a linear interpolation in time along streamlines. As particles travel along the streamlines, the rate of salinity change is constant during a time step. Since the salinity and velocity field is piecewise stationary, the rate of salinity change felt by the particles is, by definition, the advection term. The interpolation of advection scheme is thus defined by the salinity and velocity interpolation schemes. Advection along a streamline is given by the difference of salinity between the cell walls at which the streamline enters and leaves the grid cell divided by the temporal duration. It can be shown that (see subsection), even with small mass flux divergences, the volume integration of the advection term is equal to the advection term we discussed in the previous section.

$$\frac{1}{V}\int w_{adv}(\vec{x})dV = 1 \tag{26}$$

where $w_{adv}(\vec{x}) = (u \cdot \nabla s)/g(U, S)$ is the weight function for the advection distribution.

LHS term interpolation

The LHS of the equation is made up of the advection term and the Eulerian tendency term. The previous section has already defined the interpolation scheme for the advection term. In the volume averaged budget calculation, the Eulerian tendency term is determined by daily snapshots of salinity and sea surface height.The advective flux and, as a result, the salinity interpolation are from daily mean data. The change of daily mean salinity from step to step cannot be derived from the Eulerian tendency term in the salty budget equation. As a result, it is impossible to close a budget on the particles without knowledge of every model time step.

In this study, we make the simplest assumption that the Eulerian tendency term in a grid cell is uniform.

$$w_{et}(\vec{x}) = 1 \tag{27}$$

where w_{et} is the weight of the Eulerian tendency term, defined in a similar fashion as w_{adv}.This interpolation scheme is not going to exactly match the salinity change of the particles associated with Eulerian tendency.However, it is a simple and accurate approximation since the Eulerian tendency term is spatially smooth and the time scale (years) we are interested in is much longer than the output interval (a day). In addition, it can be readily seen that it satisfies 24

The error related to the Eulerian tendency term (the RHS of 20) has another component that arises from the Eulerian budget formulation, i.e.

$$\left(\frac{\Delta_E(z^*S)}{\Delta_E t} - S\frac{\Delta_E z^*}{\Delta_E t}\right) \neq \frac{\Delta_E S}{\Delta_E t} \tag{28}$$

This component is very small compared to the components we discussed above and is therefore ignored. As we see in the result section, the combined error (the RHS of 20) is still small compare to the other terms, which combined with the interpolation scheme we are going to discuss in the next subsection gives us a nearly closed budget on every Lagrangian particle.

RHS term interpolation

In this subsection, we consider the salinity change happened within a time step when the Eulerian salinity and velocity are independent of time.In order for the interpolation to be consistent with the Eulerian salt budget, we need the volume integrated contribution to match that diagnosed from the Eulerian/finite-volume perspective. An convenient form of equation is the following normalization condition.

$$\frac{1}{V}\int_{vol} w_{term} dV = 1 \tag{29}$$

To have a closed Lagrangian budget, we have

$$\Sigma w_i(\vec{x})F_i = w_{LHS}(\vec{x})F_{LHS} \tag{30}$$

where w_{LHS}, F_{LHS} are the spatial distribution and volume mean of the LHS terms, respectively. Since both the advection term and the Eulerian tendency term (see and) both satisfy the normalization condition, their weighted average also satisfy

$$\frac{1}{V}\int_{vol} w_{LHS} dV = 1, \tag{31}$$

and since the volume-mean budget is closed, we have

$$F_{LHS} = \Sigma F_i. \tag{32}$$

There is going to be infinitely many functions that satisfy 31 and 32 at the same time. In this study we use a family of them,

$$w_i(\vec{x}) = w_{i,prior}(\vec{x}) + \frac{F_i F_{LHS}}{\Sigma F_i^2}(w_{LHS}(\vec{x}) - w_{i,prior}(\vec{x})), i \in RHS, \tag{33}$$

where $w_{i,prior}$ is a prior weight function that satisfies the normalization condition. It is very easy to see that 33 does satisfy both conditions. In this work, we use $w_{prior} = t_{res}/t_s(\vec{x})$, where t_{res} is the mean residence time of the cell and t_s is the temporal length of the streamline.

There is another interpretation of this interpolation scheme. If one integrate 19 between a particle's entry and exit in a grid cell,

$$\Delta S + (\frac{\partial s}{\partial t})\Delta t = \Sigma F_i t_i \tag{34}$$

where t_i is the effective time defined by,

$$t_i = \int_{traj} w_i(\vec{x}) \frac{dl}{|u|} \quad (35)$$

Now, if we take the integrated form, and assume that the t_i follows normal ditribution $N(t_{res}, \sigma t_{res})$, where σ is a dimensionless constant that got canceled out in the later optimization process. The maximum likelihood estimate of the contribution of each terms give the same result of the above mentioned interpolation scheme. A more detailed proof can be found in .

Particle effective volume

The velocity we use to simulate the particles is the interpolated velocity from finite-resolution model. This means the flow map 15 is not tracking the original molecules, which will be the case if, in an ideal world, an infinite resolution flow field is used. Rather, the flow map serves as an alternative description of the model advective scheme. The velocity interpolation is piecewise in time and piecewise linear in space. It can be proved (see) that the flow map $X(a)$ with any fixed time is an absolute continuous function of the label a (see prove at the end of this subsection), which means the first derivative exists and it is Riemann integrable. The flow map maps a continuous domain $\Omega(0)$ to another continuous domain $\Omega(t)$. Now, consider integration of some property (e.g. freshwater forcing) over a subset (or full set) of the advected domain.

$$I = \int_{\omega \subseteq \Omega(t)} F(x(a,t),t) d^3X \quad (36)$$

All quantitative results in this paper can be writen under this general form. A example of this would be integrating the freshwater forcing in the intersection of a grid cell and advected domain, i.e. the instantaneous contribution of freshwater forcing from the grid cell. Because the flow map is absolute continuous, this general formulation can be rewritten as

$$I(a,t) = \int_{\omega(0) \subseteq \Omega(0)} F(x(a,t),t) \left| \frac{dx_i}{da_j} \right| d^3a \quad (37)$$

where $\left| \frac{dx_i}{da_j} \right|$ is the determinant of the flow map's Jacobian matrix. Under our interpolation scheme, F is always finite, which makes the integrand Riemann integratable.

$$I(a,t) = \lim_{N \to \infty} \Sigma_N F(x(a_i,t),t) \left| \frac{dx_i}{da_j} \right| (a_i,t) \frac{V(\Omega(0))}{N} \quad (38)$$

where a_i is a uniformly placed array of position or a set of random points drawn from a uniform distribution. Numerically, we can use this formulation to estimate the quantity by increasing the number of points used and look at the convergence. For conveneince, we are going to assume the determinant of the Jacobian is 1. This approximation is well justified in , since the velocity divergent is very small.

Supplementary Material

Relation between volume-averaged and transport-averaged description

In this section, we are going to describe the advection process from both volume averaged and transport averaged sense, and show that, under our setup, the two descriptions are related by the particle mean residence time, even when weak compressibility is present. First, we explain, what we mean by weak compressibility

$$|D| \ll T_{out}, T_{in}, \tag{39}$$

where

$$T_{in} = -\int\int_{\hat{n}\cdot u<0} dA \cdot u T_{out} = \int\int_{\hat{n}\cdot u>0} dA \cdot u D = T_{out} - T_{in} \tag{40}$$

are the transport into and out of the control volume (grid cell) and the volume integrated divergence. The single largest source of "compressibility" in ECCO is evaporation and precipitation, which is much smaller than the volume transport across grid cell walls. Given this reasonable assumption, the transport averaged residence time of the system is then given in a similar way as equilibrated system

$$t_{res} = \frac{2V}{T_{in} + T_{out}} \approx \frac{V}{T_{in}} \approx \frac{V}{T_{out}} \tag{41}$$

The transport average salinity change felt by the particles are given by

$$E[\Delta S] = S_{out} - S_{in} \tag{42}$$

where

$$S_{in} = -\frac{\int\int_{\hat{n}\cdot u<0} dA \cdot uS}{T_{in}} S_{out} = \frac{\int\int_{\hat{n}\cdot u>0} dA \cdot uS}{T_{out}} \tag{43}$$

are the transport-weighted-mean salinity of the incoming and outgoing flow. Our requirement of the interpolated salinity field, can then be expressed as

$$T_{out}S_{out} - T_{in}S_{in} = \int_{\Omega} \nabla \cdot (uS) dV \tag{44}$$

Now, we are going to prove that for a particle enters and leave a grid cell the mean salinity change is related to the volume mean advection term of the cell by

$$E[\Delta S] = t_{res}\frac{1}{V}\int_{\Omega} u \cdot \nabla S dV. \tag{45}$$

Take all of the transport averaged terms to the left hand side we got

$$\frac{E[\Delta S]}{t_{res}} = \frac{1}{V}\int_{\Omega} u \cdot \nabla S dV. \tag{46}$$

Now, the left hand side can be transformed into volume averaged terms with some assumptions

$$\begin{aligned}
LHS =& \frac{E[\Delta S]}{t_{res}} \\
=& \frac{\frac{T_{out}S_{out}}{T_{out}} - \frac{T_{in}S_{in}}{T_{in}}}{\frac{2V}{T_{in}+T_{out}}} \\
=& \frac{\int_{\Omega} \nabla \cdot (uS) dV}{V} \\
& + \left(\frac{T_m}{T_{out}} - 1\right)\frac{T_{out}S_{out}}{V} \\
& - \left(\frac{T_m}{T_{in}} - 1\right)\frac{T_{in}S_{in}}{V} \\
\approx& \frac{\int_{\Omega} \nabla \cdot (uS) dV}{V} - \frac{D(T_{out}S_{out} + T_{in}S_{in})}{2T_m V} \\
\approx& \frac{1}{V}\int_{\Omega} \nabla \cdot (uS) - S(\nabla \cdot u) dV \\
=& RHS
\end{aligned} \tag{47}$$

The second last step used the fact that salinity has a base value (around 35 psu) much larger than its variability (around or less than 0.1 psu within a grid cell of diameter one degree).

$$\frac{T_{out}S_{out} + T_{in}S_{in}}{2T_m} \approx S_{any} \tag{48}$$

Remove upstream bolus velocity error

The ECCOv4r4 bolus velocity product has an unexpected behavior at several grid cells in the dataset. By definition, the vertical component of bolus velocity is constructed such that the vector field is divergent free. However, at those grid cells the velocity is not divergent free to machine position. Those grid cells also share a common property that there are precisely two unmasked cells in the vertical column and they are all next to the coast. The diagnosed salinity fluxes is not affected by this strange behavior. However, when we are trying to recreate the fluxes using daily mean salinity and velocity, the divergence of bolus velocity result in a large divergence (convergence) in the recreated flux. This strange characteristic also result in unrealistic trajectories, backward particles entering the grid cell (10, 52 ,89) is unable to escape because all the velocity components are outward or zero. The authors consider the above mentioned behavior an upstream bug and make the following adjustments to the budget analysis: 1. Identified the grid cells that have large divergence by looking at differences in the difference in mean advection divergence. Where the difference is larger than $10^{-7} psu/s$, the cell is considered erroneous. 2. Separate the unresolved advection. This term, the difference between recreated divergence and diagnosed divergence, is the combined contribution of sub-daily fluctuation and the erroneous convergence. We separate the two spatially. Namely, difference at those identified erroneous cells are considered error and the rest is considered sub-daily contribution.3. We do not apply any special treatment for particle trajectories, even for particles that are going to stuck in the cells forever. 4. Salinity in the erroneous cells are set uniformly as the salinity of the wall with the largest transport. 5. The jump in salinity when particles enter or leave the cell is recorded as the contribution from this error term.

Prove of equivalency between state estimate and interpolation

We are going to look at the contribution of the terms to salinity, which is just

$$\Delta S = \Sigma_{i=1}^{N_{term}} C_i \tag{49}$$

Now, we assume the contribution is normally distributed

$$C_i \sim N(F_i t, \sigma F_i t) \tag{50}$$

Again, σ is a dimensionless constant that will be canceled out. We assume the terms are not correlated. The joint probability density function of the contributions is thus just a simple normal distribution.

$$P(C_1, C_2, ..., C_{n_{term}}) = \frac{1}{(\Pi_{i=1}^{n_{term}} F_i)(\sigma t\sqrt{2\pi})^{n_{term}}} \exp[-\Sigma_{i=1}^{n_{term}} \frac{(C_i - F_i t_i)^2}{2(\sigma F_i t)^2}] \tag{51}$$

The likelihood can also be described by this function. To make the notations simpler, we let

$$x_i = C_i - F_i t \tag{52}$$

The constraint on x_i is then

$$\Sigma_{i=1}^{n_{term}} x_i = \Delta S - \Sigma_{i=1}^{n_{term}} F_i t = C_e \tag{53}$$

which happens to be the error in the independent approach. The Likelihood function is then

$$L(x_1, x_2, ..., x_{n_{term}}) = \frac{1}{(\Pi_{i=1}^{n_{term}} F_i)(\sigma t\sqrt{2\pi})^{n_{term}}} \exp[-\Sigma_{i=1}^{n_{term}} \frac{x_i^2}{2(\sigma F_i t)^2}] \tag{54}$$

We can take the log and ignore the constants. It is then easy to see that to maximize the likelihood is just to minimize the following optimization function

$$l(x_1, x_2, ..., x_{n_{term}}) = \Sigma_{i=1}^{n_{term}} \frac{x_i^2}{F_i^2} \tag{55}$$

Now we use Lagrange multiplier method. The new optimization goal is then

$$l' = \Sigma_{i=1}^{n_{term}} \frac{x_i^2}{F_i^2} - \lambda(\Sigma_{i=1}^{n_{term}} x_i - C_e) \tag{56}$$

The maximum/minimum appears when all the partial derivatives are zero.

$$\frac{\partial l'}{\partial x_1} = \frac{2x_1}{F_1^2} - \lambda = 0 ... \frac{\partial l'}{\partial x_j} = \frac{2x_j}{F_j^2} - \lambda = 0 ... \frac{\partial l'}{\partial \lambda} = \Sigma_{i=1}^{n_{term}} x_i - C_e = 0 \tag{57}$$

This is very easy to solve, and the results are

$$\lambda = \frac{2C_e}{\Sigma_{i=1}^{n_{term}} F_i^2} x_j = \frac{F_j^2 C_e}{\Sigma_{i=1}^{n_{term}} F_i^2}, j = 1, 2, ...n_{term} \tag{58}$$

Going back to the contribution

$$C_j = F_j t + \frac{F_j^2}{\Sigma_{i=1}^{n_{term}} F_i^2}(\Delta S - t\Sigma_{i=1}^{n_{term}} F_i) \tag{59}$$

Divide $F_j t$ on both side we get,

$$w_j = 1 + \frac{F_j}{\Sigma_{i=1}^{n_{term}} F_i^2}(\frac{\Delta S}{t} - \Sigma_{i=1}^{n_{term}} F_i) \tag{60}$$

Where w_j is the non-dimensional distribution function, which is a constant along this trajectory (it could vary from trajectory to trajectories). We can express ΔS as

$$\Delta S = \int_{traj} u \cdot \nabla s \frac{dl}{|u|} = \int_{traj} F_{ADV} w_{ADV} \frac{dl}{|u|} \tag{61}$$

If we assume w_{ADV} is a constant along the trajectory, we have

$$\Delta S = F_{ADV} W_{ADV} t \tag{62}$$

Recall

$$F_{ADV} = \Sigma_{i=1}^{n_{term}} F_i \tag{63}$$

The expression for the distribution function is therefore just

$$w_j = 1 + \frac{F_j F_{ADV}}{\Sigma_{i=1}^{n_{term}} F_i^2}(w_{ADV} - 1) \tag{64}$$

Which is just the $p = 1$ case with different notation (I will fix later).

Prove for absolute continuity

Integrate the velocity in time

$$X_i(a,t) = a_i + \int_0^t u_i(\vec{X}(a,t'),t')dt' \tag{65}$$

Take the partial derivative on a_j of both sides

$$\frac{\partial X_i}{\partial a_j} = I_{ij} + \int_0^t \frac{\partial u_i}{\partial X_k} \frac{\partial X_k}{\partial a_j} dt' \tag{66}$$

Now we take the full time derivative on both sides.

$$\frac{d}{dt}\frac{\partial X_i}{\partial a_j} = \frac{\partial u_i}{\partial X_k}\frac{\partial X_k}{\partial a_j} \tag{67}$$

This is an initial value problem, and $\frac{\partial u_i}{\partial x_k}$ is bounded every where in the model domain under our velocity interpolation scheme. The initial condition is simply

$$\frac{\partial X_i(a,0)}{\partial a_j} = I_{ij} \tag{68}$$

Then the derivative of flow map can be expressed as a matrix exponential

$$\frac{\partial X_i}{\partial a_j}(a,t) = \exp[\int_0^t \frac{\partial u_i}{\partial X_j}(X(a,t'),t')dt'] \tag{69}$$

We can actually take another derivative on a_q.

$$\frac{\partial^2 X_i}{\partial a_j \partial a_q}(a,t) = \int_0^t \frac{\partial^2 u_i}{\partial X_k \partial X_p}\frac{\partial X_p}{\partial a_q}dt' \exp[\int_0^t \frac{\partial u_k}{\partial X_j}dt'] \tag{70}$$

The second derivative of velocity in our interpolation scheme looks like delta functions that have infinite value at cell boundaries. Note that

$$\frac{d}{dt}(\frac{\partial u_i}{\partial x_k}) = \frac{\partial^2 u_i}{\partial x_k \partial t} + \frac{\partial^2 u_i}{\partial x_k \partial x_p}u_p \tag{71}$$

The first integration on the left hand side thus becomes

$$\int_0^t \frac{d}{dt}(\frac{\partial u_i}{\partial x_k})\frac{1}{u_p}\frac{\partial X_p}{\partial a_q}dt' - \int_0^t \frac{\partial^2 u_i}{\partial x_k \partial t}\frac{1}{u_p}\frac{\partial X_p}{\partial a_q}dt' \tag{72}$$

It is not hard to show that both of the integrals are finite after some integration by part on the first one and righting the second one as delta functions in time.(Hopefully this is somewhat obvious at this point, because I am too tired to type more fractions in latex...) Until now, we actually proved that the flow map is $C^2 $, this is a stronger condition than absolute continuity.

The determinant of flow map Jacobian

In the integration, the determinant of the Jacobian of the Lagrangian flow map can be expressed as

$$|J| = det\left(\frac{\partial x_i}{\partial a_j}\right) = \epsilon_{pqr}\frac{\partial x_1}{\partial a_p}\frac{\partial x_2}{\partial a_q}\frac{\partial x_3}{\partial a_r} \tag{73}$$

Taking the Lagrangian derivative of both sides, we have

$$\frac{d}{dt}|J| = \epsilon_{pqr}\frac{\partial u_1}{\partial a_p}\frac{\partial x_2}{\partial a_q}\frac{\partial x_3}{\partial a_r} + \epsilon_{pqr}\frac{\partial x_1}{\partial a_p}\frac{\partial u_2}{\partial a_q}\frac{\partial x_3}{\partial a_r} + \epsilon_{pqr}\frac{\partial x_1}{\partial a_p}\frac{\partial x_2}{\partial a_q}\frac{\partial u_3}{\partial a_r} \tag{74}$$

Expand the velocity derivative using the chain rule on the first term, we get

$$\epsilon_{pqr}\frac{\partial u_1}{\partial x_m}\frac{\partial x_m}{\partial a_p}\frac{\partial x_2}{\partial a_q}\frac{\partial x_3}{\partial a_r} = \epsilon_{pqr}\frac{\partial u_1}{\partial x_1}\frac{\partial x_1}{\partial a_p}\frac{\partial x_2}{\partial a_q}\frac{\partial x_3}{\partial a_r} = \frac{\partial u_1}{\partial x_1}|J| \tag{75}$$

This because if $m \neq 1$ there will be repeating terms under the Levi-Civita notation. This is also true for the other two terms in the derivative. As a result,

$$\frac{d}{dt}|J| = |J|\nabla \cdot u \tag{76}$$

For a divergent free velocity field, the determinant is a constant 1 as dictated by the initial condition. For convenience, we assume the determinant remains one through out the duration of the simulation, which is less than 10 years. This is a good approximation because the divergent is in general very small ($O(10^{-10})$) and are concentrated at the top grid. This approximation is also equivalent to assuming that there is a constant flux/volume assigned to a particle, which is a common assumption in Lagrangian particle literature (e.g. Exceptional freshening. . .)