2. THE TRANSPORT OPERATOR

2.1 Mean and turbulent components of transport

The transport component of the continuity equation,

$$\frac{\partial n}{\partial t} = -\nabla \bullet (n\mathbf{U}) \tag{2.1}$$

is called the *advection equation*. Its solution defines the transport operator **T**. To solve the advection equation we need information on the wind velocity **U**. Winds in the atmosphere are driven by pressure gradients, which are in turn driven by thermal gradients. In atmospheric dynamics models the wind velocity is calculated with the *Navier-Stokes equation*, which is the equivalent of the continuity equation applied to momentum. In the Navier-Stokes equation for momentum the transport term is of the form $\mathbf{U} \bullet \nabla \mathbf{U}$ and the associated nonlinearity produces chaos in the solution for **U**(**x**, *t*). This means that **U** is not steady but has large erratic fluctuations on all scales down to ~1 mm (in space) and ~10 Hz (in time). We refer to the erratic nature of the flow as *turbulence.*

Fluid dynamics models attempting to resolve turbulence in a deterministic sense must use grid resolution of the order of a mm. This is impractical on the scales that interest us, thus we need to treat turbulence in a statistical-ensemble manner. For a given model gridpoint and time step we want the time-averaged value <**U**> and a statistical measure of variability for the wind. We thus decompose **U**(**x**, t) as

$$\mathbf{U} = \langle \mathbf{U} \rangle + \mathbf{U}' \tag{2.2}$$

where **U**' is the *fluctuating* component of the wind such that <**U'**> = 0 over the time step. The number density in the advection equation can be decomposed similarly:

$$n = \langle n \rangle + n' \tag{2.3}$$

In the advection equation, we can only hope to solve for the time-averaged value <*n*>,

$$\frac{\partial \langle n \rangle}{\partial t} = -\nabla \bullet \langle n\mathbf{U} \rangle \tag{2.4}$$

and therefore we need to express the time-averaged advective flux $\mathbf{F}_{\mathbf{adv}} = -\nabla \langle n\mathbf{U} \rangle$ on the right-hand side in terms of <*n*>. We decompose this flux following (2.2) and (2.3):

$$\mathbf{F}_{\mathbf{adv}} = \langle n\mathbf{U} \rangle = \langle (\langle n \rangle + n')(\langle \mathbf{U} \rangle + \mathbf{U}') \rangle = \langle n \rangle \langle \mathbf{U} \rangle + \langle n'\mathbf{U}' \rangle \tag{2.5}$$

where $\mathbf{F_M} = \langle n \rangle \langle \mathbf{U} \rangle$ is the *mean advective flux* over the time step and $\mathbf{F_T} = \langle n' \mathbf{U}' \rangle$ is the *turbulent flux*. Thus the transport flux of a species in the atmosphere is the sum of two terms: $\mathbf{F_M}$ determined by the mean wind over the time step (for which information is available) and $\mathbf{F_T}$ determined by the covariance of n and $\mathbf{U}$ over the time step. The partitioning of the flux between mean advective and turbulent components obviously depends on the length of the time-averaging period. In models this period is generally at least one hour. In the vertical direction where mean flows are weak and where buoyancy produces large local accelerations, turbulence generally accounts for most of the flux. In the horizontal direction, turbulence is generally less important but may still dominate the flux if the mean wind is weak.

In solving the advection equation, rewritten now as

$$\frac{\partial \langle n \rangle}{\partial t} = -\nabla \bullet \langle n \rangle \langle \mathbf{U} \rangle - \nabla \bullet \langle n' \mathbf{U}' \rangle \tag{2.6}$$

a challenge is to describe the turbulent flux divergence term $-\nabla \bullet \langle n' \mathbf{U}' \rangle$ as a function of $\langle n \rangle$. Because we don't generally have observational information for this term, nor do we have the computational resources to calculate it deterministically, it must be *parameterized* based on some empirical evidence. I describe below two simple parameterizations of turbulence commonly used in atmospheric models.

2.2 Parameterizations of turbulence

2.2.1 Eddy diffusion

In the *eddy diffusion* (alternatively called *turbulent diffusion*) parameterization of turbulence, we approximate turbulence as a random process analogous to molecular diffusion. By analogy with Fick's law, we assume that the turbulent flux is proportional to the time-averaged mean gradient in the mixing ratio:

$$\mathbf{F_T} = -\mathbf{K} n_a \nabla \langle C \rangle \tag{2.7}$$

where $\mathbf{K}$ is an *eddy diffusion matrix* and $\nabla \langle C \rangle$ is the gradient in the time-averaged mixing ratio. The elements of the $\mathbf{K}$ matrix are the *eddy diffusion coefficients* ($\mathrm{cm^2\ s^{-1}}$). The diagonal elements describe transport along each spatial coordinate driven by the gradient in mixing ratio along that coordinate. The off-diagonal elements describe covariance between turbulent transport in different directions. In general this covariance is assumed to be zero, so that $\mathbf{K}$ is a diagonal matrix:

$$\mathbf{K} = \begin{pmatrix} K_x & 0 & 0 \\ 0 & K_y & 0 \\ 0 & 0 & K_z \end{pmatrix} \tag{2.8}$$

where K_x, K_y, K_z are the eddy diffusion coefficients in each direction to be determined empirically. Note that in equation (2.7) the concentration gradient is expressed in terms of mixing ratio rather than number density, because of the compressibility of air; the turbulent flux responds to a gradient in mixing ratio, not necessarily to a gradient in number density.

With the eddy diffusion parameterization of turbulent transport, the continuity equation becomes

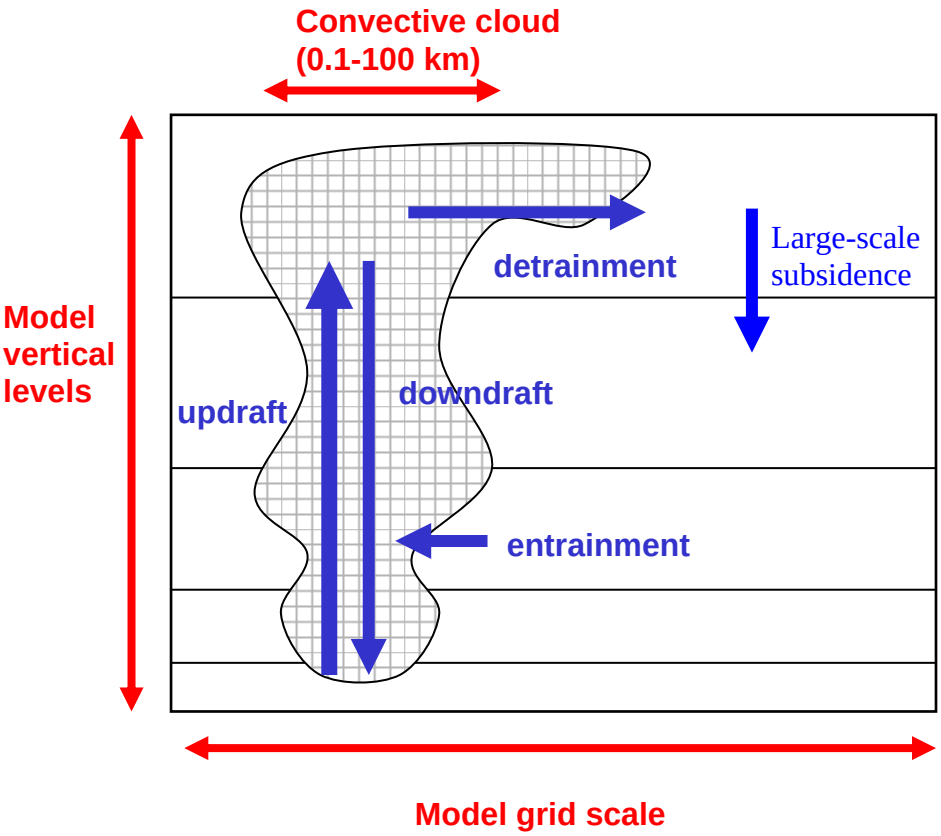
$$\frac{\partial \langle n \rangle}{\partial t} = -\nabla \bullet \langle n \rangle \langle \mathbf{U} \rangle + \nabla \bullet (\mathbf{K} n_a \nabla \langle C \rangle) + \langle P \rangle - \langle L \rangle \qquad (2.9)$$

where the concentration of the species is now solely described by time-averaged terms. In the rest of this chapter we will take n and C to represent the time-averaged terms and drop the bracket notation.

The eddy diffusion parameterization of turbulence has been quite successful, because observations show that **K** generally does not vary with the nature or the gradient of the species being transported. Textbook formulas are available for K_x, K_y, K_z as function of the surface fluxes of momentum and sensible heat, which measure respectively the mechanical and buoyant components of turbulence [Seinfeld and Pandis, chapter 18.7]. The vertical component K_z is of most interest because of the dominance of turbulent motions in that direction. In an unstable (buoyant) atmosphere K_z is of the order of 10^5-10^7 cm^2 s^{-1}, while in a stable (stratified) atmosphere it is of the order of 10^2-10^5 cm^2 s^{-1}. Even in a stratified atmosphere, turbulent transport is orders of magnitude faster than molecular diffusion.

2.2.2 Wet convective transport

The eddy diffusion parameterization assumes that turbulence occurs on sufficiently small scales that it can be treated as a diffusive process on the scale of the transport model. It is called a *local* description of turbulence because it involves only concentration gradients between adjacent gridboxes. That is generally found to be an acceptable assumption for dry boundary layer turbulence, and it is also a popular assumption in conceptual models because of its simplicity in implementation. However, it fails to describe motions that take place on subgrid scales (and are therefore "turbulent" from the standpoint of the model) but still exhibit structure on the grid scale. Such is the case for wet convective motions, where release of latent heat in a rising saturated air parcel leads to rapid cloud updrafts moving across several gridboxes in the vertical (see figure). The updraft is typically 1-10 km in horizontal extent, i.e., subgrid on the scale of global or mesoscale models. From air mass balance considerations it must be balanced by downdrafts within the cloud and by grid-scale subsidence outside the cloud. Mid-level entrainment and detrainment in the convective column further complicate the picture.



Wet convective motions cannot be represented by an eddy diffusion parameterization. Updrafts and downdrafts connect directly the lower and upper levels of the convective column, bypassing the mid-levels, whereas a diffusive parameterization would require gradual transport through the mid-levels. One therefore needs a *convective parameterization* in the model. The convective parameterization consists in the specification of updraft, downdraft, and entrainment/detrainment fluxes for the grid column as a function of the grid-resolved factors that drive convection (large-scale convergence and vertical instability). Note that the need for a parameterization of convection is not specific to atmospheric chemistry models; meteorological models also need it to simulate transport of momentum, heat, and water. In general, CTMs will attempt to replicate the parameterization of convective transport used in the parent meteorological model. Ideally this can be done with archived updraft, downdraft, and entrainment/detrainment fluxes. If such archives are not available, the convective fluxes must be estimated by applying the convective parameterization to the archived wind velocity and temperature fields. This post-diagnosis is subject to bias because the archived wind and temperature fields have generally been relaxed for convection. Convective transport in the atmospheric chemistry model is then underestimated.

2.3 Numerical solution of the advection equation

We now address the numerical solution of the advection equation in one dimension,

$$\frac{\partial C}{\partial t} = -u\frac{\partial C}{\partial x} \tag{2.10}$$

corresponding to the transport operator $\mathbf{T_x}$. Equation (2.10) is the *advective* form of the continuity equation; it has the advantage of only requiring information on u at the gridpoint of the model. Solution to the flux form will be discussed in section 2.3.2; it requires information on u midway between gridpoints, i.e., at gridbox boundaries.

A large number of schemes have been described in the literature to solve the advection equation. We present here four classic schemes that illustrate the general character of the solution (section 2.3.1), and also describe a "volume" scheme that has much application in current research models (section 2.3.2). A successful advection scheme must have several attributes. *Accuracy* is of course highly desired. *Stability* (the ability of the solution to remain bounded even after long integration times) is essential to ensure that the model does not diverge. *Mass conservation* is also essential; if the algorithm does not conserve mass, the ability to derive budgets will be compromised. *Positivity* is required to avoid non-physical solutions, and also in the case of reactive chemicals to avoid unstability in the chemical operator.

Stability generally requires that the dimensionless *Courant number* $\varepsilon = u\Delta x/\Delta t$, where Δx is the separation between gridpoints and Δt.is the time step, be less than unity. This can be a serious limitation on time step size, and we will see in section 2.3.3 how a semi-Lagrangian approach can help to circumvent this limitation. Postitivity often requires some adjustment to the solution to avoid negative ripples; this can cause non-linearity in the solution, so that transport of the sum of concentrations of two chemicals ($C_1 + C_2$) may yield slightly different results from transporting separately the two chemicals.

2.3.1 Classic schemes

We wish to solve the advection equation at a discrete set of gridpoints $i = 1,\ldots m$ separated by Δx. Let $C_{i,t}$ and $u_{i,t}$ represent the mixing ratio and the wind velocity at gridpoint i and time t. Starting from the Taylor expansion in space

$$C_{i\pm1,t} = C_{i,t} \pm \Delta x\frac{\partial C_{i,t}}{\partial x} + \frac{\Delta x^2}{2}\frac{\partial^2 C_{i,t}}{\partial x^2} \pm \frac{\Delta x^3}{6}\frac{\partial^3 C_{i,t}}{\partial x^3} + \ldots \tag{2.11}$$

we can write finite-difference forms for the spatial derivative $\partial C/\partial x$:

$$\frac{\partial C}{\partial x} = \frac{C_{i+1,t} - C_{i,t}}{\Delta x} + \mathrm{O}(\Delta x) \quad \text{(forward)}$$
$$\frac{\partial C}{\partial x} = \frac{C_{i,t} - C_{i-1}}{\Delta x} + \mathrm{O}(\Delta x) \quad \text{(backward)} \qquad (2.12)$$
$$\frac{\partial C}{\partial x} = \frac{C_{i+1,t} - C_{i-1}}{2\Delta x} + \mathrm{O}(\Delta x^2) \quad \text{(centered)}$$

where the centered form is obtained by subtracting the (-) from the (+) forms of equation (2.11). The centered form is second-order accurate and is therefore in general a better choice than the forward or backward forms. Similar finite difference forms can be written for the temporal derivative.

Differences between advection schemes largely reflect differences in the finite-difference forms for the derivatives. The *forward Euler* scheme uses the forward form of the temporal derivative and the centered form of the spatial derivative:

$$C_{i,t+\Delta t} = C_{i,t} + \frac{u\Delta t}{2\Delta x}(C_{i-1,t} - C_{i+1,t}) = C_{i,t} + \frac{\varepsilon}{2}(C_{i-1,t} - C_{i+1,t}) \qquad (2.13)$$

and has accuracy $\mathrm{O}(\Delta t, \Delta x^2)$. It would appear to be a straightforward method for solving the advection equation. However, it is unstable for all values of ε and so it should never be used. By using instead the centered form of the temporal derivative, one obtains the *leapfrog* scheme:

$$C_{i,t+\Delta t} = C_{i,t-\Delta t} + \varepsilon(C_{i-1,t} - C_{i+1,t}) \qquad (2.14)$$

which has accuracy $\mathrm{O}(\Delta t^2, \Delta x^2)$. This scheme is stable for $\varepsilon \le 1$. However it is *dispersive*, which means that it produces small waves or ripples downstream of an advected feature. The ripples are of particular concern when they produce negative concentrations, as happens frequently in regions of strong gradients. This problem is illustrated in Figure 1 with the advection of a triangular wave in a uniform and constant wind field.

A scheme that has the same order of accuracy as the leapfrog scheme but is somewhat less dispersive is the *Lax-Wendroff* scheme. Here we start from the Taylor expansion of the temporal derivative:

$$C_{i,t+\Delta t} = C_{i,t} + \Delta t \frac{\partial C_{i,t}}{\partial t} + \frac{\Delta t^2}{2}\frac{\partial^2 C_{i,t}}{\partial t^2} + \mathrm{O}(\Delta t^2) \qquad (2.15)$$

and we convert the partial derivatives in time on the right-hand side into partial derivatives in space, using equation (2.10). Note that (2.10) can be applied to any scalar conserved during motion (i.e., whose total derivative $d/dt = 0$). We apply it here to $\partial C / \partial t$ and further assume u to be uniform:

$$\frac{\partial}{\partial t}\frac{\partial C}{\partial t} = -u\frac{\partial}{\partial x}(\frac{\partial C}{\partial t}) = -u\frac{\partial}{\partial x}(-u\frac{\partial C}{\partial x}) = u^2\frac{\partial^2 C}{\partial x^2} \tag{2.16}$$

Substitution into (2.15) then yields:

$$C_{i,t+\Delta t} = C_{i,t} - u\Delta t\frac{\partial C_{i,t}}{\partial x} + \frac{u^2\Delta t^2}{2}\frac{\partial^2 C_{i,t}}{\partial x^2} + \mathrm{O}(\Delta x^3) \tag{2.17}$$

We then substitute into (2.17) the centered finite-difference form of $\partial C_{i,t}/\partial x$, and the finite-difference form for $\partial^2 C_{i,t}/\partial x^2$ which is

$$\frac{\partial^2 C_{i,t}}{\partial x^2} = \frac{\frac{C_{i+1,t}-C_{i,t}}{\Delta x} - \frac{C_{i,t}-C_{i-1,t}}{\Delta x}}{\Delta x} = \frac{C_{i+1,t}-2C_{i,t}+C_{i-1,t}}{\Delta x^2} \tag{2.18}$$

We thus obtain the equation for the Lax-Wendroff scheme:

$$C_{i,t+\Delta t} = C_{i,t} + \frac{\varepsilon}{2}(C_{i-1,t}-C_{i+1,t}) + \frac{\varepsilon^2}{2}(C_{i+1,t}-2C_{i,t}+C_{i-1,t}) \tag{2.19}$$

which has accuracy O(Δt^2, Δx^2) and is stable for $\varepsilon \le 1$. Figure 1 illustrates the advection of a triangular wave by this scheme. Although less dispersive than the leapfrog scheme it still generates spurious negative concentrations. It is also *diffusive*, that is, perturbations to the concentration field tend to become smoothed out over time. In fact, comparison of equations (2.19) and (2.13) indicates that the Lax-Wendroff scheme is in fact the forward Euler scheme with a diffusion term $\partial^2 C/\partial x^2$ added. One can show that the Lax-Wendroff scheme adds the minimum amount of diffusion to the forward Euler scheme in order to make it stable. In contrast, the leapfrog scheme is not diffusive at all.

Our last classic scheme is the *upstream* scheme, in which the spatial derivative is expressed as either the forward or backward finite-difference form depending on the direction of the flow:

$$\begin{aligned} C_{i,t+\Delta t} &= C_{i,t} - \varepsilon(C_{i+1,t}-C_i) \qquad \text{if } u \le 0 \\ C_{i,t+\Delta t} &= C_{i,t} - \varepsilon(C_{i,t}-C_{i-1,t}) \qquad \text{if } u \ge 0 \end{aligned} \tag{2.20}$$

This scheme has accuracy O(Δt, Δx) and is stable for $\varepsilon \le 1$. In contrast to the leapfrog and Lax-Wendorff schemes it is not dispersive, as shown in Figure 1, and therefore has the great advantage of not generate negative concentrations. However it is highly diffusive, far more so than the previous two schemes. This diffusivity results from the low accuracy in the approximation of the derivatives. More generally, however, schemes that are odd-order accurate tend to be diffusive, while schemes that are even-order accurate tend to be dispersive.

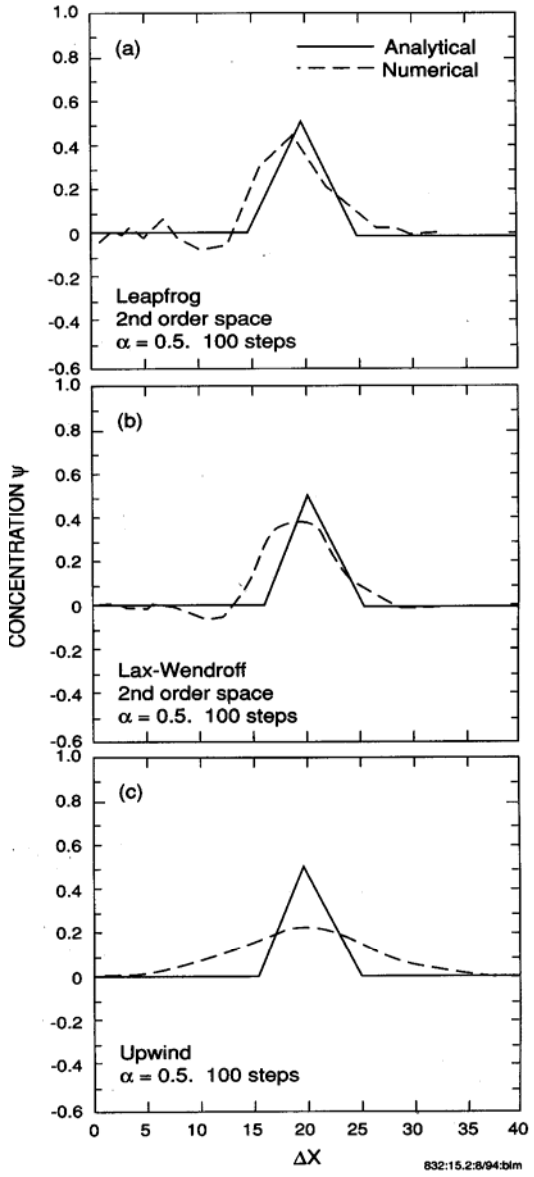


Figure 12.2. *Performance of three numerical schemes for the one-dimensional advection of a triangular function: (a) leap frog, (b) Lax-Wendroff, and (c) upwind methods (100 time steps with Courant number α = 0.5). The exact (analytical) solution is also shown (adapted from Rood, 1987).*

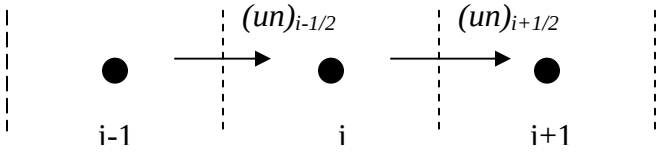
Exercise. Consider a square wave in a 1-D domain of gridpoints i separated by Δx, with initial conditions $C_{i,0} = 1$ for $i \in [5,10]$ and $C_{i,0} = 0$ for $i < 5$ or $i > 10$. Transport this wave for 100 time steps with $\varepsilon = 0.5$ and $\varepsilon = 0.1$, using each of the four classic schemes. Plot and discuss your results.

2.3.2 Volume schemes

"Volume schemes" describe a class of algorithms in which advection is treated as a mass flux between adjacent gridboxes. This treatment makes it particularly

convenient to conserve mass, as one can ensure that the mass flux out of a donor box matches the mass flux into the adjacent receptor box.

We consider here a volume scheme widely used in research and based on the classic upstream scheme (section 2.3.1). The mass fluxes are calculated at the boundaries *i-1/2* and *i+1/2* between adjacent gridboxes, as depicted in the diagram below:



The discretized flux form of the advection equation can then be written

$$\frac{n_{i,t+\Delta t} - n_{i,t}}{\Delta t} = \frac{(un)_{i-1/2,t} - (un)_{i+1/2,t}}{\Delta x} \tag{2.21}$$

To calculate the fluxes at the boundaries between gridboxes we need to have winds at the boundaries $i\pm1/2$ rather than at the gridpoints i. Such winds are sometimes available from the meteorological archives. If they are not available they should be requested! If we assume concentrations to be well-mixed within a gridbox, we can write $(un)_{i-1/2}$ as

$$(un)_{i-1/2} = \frac{u_{i-1/2} + |u_{i-1/2}|}{2} n_{i-1,t} + \frac{u_{i-1/2} - |u_{i-1/2}|}{2} n_{i,t} \tag{2.22}$$

with a similar form for $(un)_{i+1/2}$. Replacement in (2.20) defines an upstream scheme, in which fluxes are calculated solely on the basis of the upstream concentrations. The scheme is mass-conserving and stable as long as one does not remove more mass within a gridbox over the time step than there is initially, that is, as long as the Courant number is less than unity. However it is highly diffusive, reflecting the assumption of a uniform concentration within each gridbox. One can reduce the diffusivity at the cost of computational complexity by allowing for concentration gradients within a gridbox. In the *slopes scheme,* we assume a uniform gradient of concentrations within the gridbox:

$$n_{i,t}(x) = a_o + a_1(x - \frac{\Delta x}{2}) \qquad x \in [-\frac{\Delta x}{2}, \frac{\Delta x}{2}] \tag{2.23}$$

and advect the integral of the concentration over the appropriate gridbox fraction. At the end of every time step we recalculate the coefficients a_o and a_1 by adjusting for the masses transported in and out of the gridbox. This is done in practice by transporting the 0^{th} and 1^{st}-order *moments* of the distribution, where the j-th order moment $M_{j,i,t}$ has the form

$$M_{j,i,t} = \frac{1}{\alpha} \int_{-\Delta x/2}^{\Delta x/2} x^j n_{i,t}(x) dx \tag{2.24}$$

and α is a normalizing factor such that the moment has units of concentration. The suite of moments up to order j can thus fully describe a subgrid distribution $n_{i,t}(x)$ as a polynomial of order j. The *second-order-moments (SOM) scheme* is particularly popular as it resolves the subgrid distribution in terms of a parabola, thus allowing the preservation of extrema within a gridbox. In that case the advection scheme must transport three moments. In 3-D, the SOM scheme requires the transport of 10 moments (one 0^{th}-order, three 1^{st}-order, and six 2^{nd}-order); the three extra 2^{nd}-order moments are needed to handle cross-terms in *xy*, *xz*, and *yz* in the 2^{nd}-order polynomial. Thus it is computationally expensive, but it may allow a coarser model grid than would otherwise be considered acceptable.

2.3.3 Semi-Lagrangian algorithm

A general stability requirement for Eulerian approaches to solving the advection equation is a Courant number less than unity. This can be a problem when winds are strong or gridboxes are small in some part of the modeling domain (as is the case at polar latitudes in a latitude-longitude grid). One way to get around this requirement is with a *semi-Lagrangian* algorithm, in which the gridded concentration field at time $t+\Delta t$ is updated from that at time t by conducting Lagrangian back-trajectory calculations over $[t, t+\Delta t]$ for the air at each gridpoint $\mathbf{x}$ at time $t+\Delta t$. At time t this air was at location $\mathbf{x_o}$, given by

$$\mathbf{x_o} = \mathbf{x} + \int_{t+\Delta t}^{t} \mathbf{U}(\mathbf{x}, t)dt \qquad (2.25)$$

so that the mixing ratio field $C(\mathbf{x_o}, t)$ maps onto $C(\mathbf{x}, t+\Delta t)$. The locations $\mathbf{x_o}$ do not in general correspond to gridpoints, so that the mixing ratio field $C(\mathbf{x_o}, t)$ must be interpolated from the field $C(\mathbf{x}, t)$ derived at the previous time step.

The semi-Lagrangian algorithm has no Courant number constraints, which is a big advantage, and it is less diffusive than most schemes. It has two important disadvantages. The first is that it does not conserve mass. One generally addresses this problem by multiplying the concentration field obtained at time $t+\Delta t$ by a uniform scaling factor to enforce mass conservation. The second is that determination of $\mathbf{x_o}$ from a single back-trajectory, as described by (2.25), induces large errors in regions of shear or convergent flow where $\mathbf{U}$ is highly variable in space. An approach to address this problem would be to calculate a *cluster* of back-trajectories for an ensemble of points in the domain $[\mathbf{x} \pm \Delta\mathbf{x}/2]$ and thus derive a *probability density function* for $\mathbf{x_o}$. However, the computational cost is large.