Introduction to the Finite Volume Method

Conservation Equations

Generic Conservation Equation

Before introducing the finite volume method, we will first introduce a generic conservation equation of the form:

$$\frac{\partial \phi}{\partial t} + \nabla \cdot (\mathbf{u}\phi) + \nabla \cdot \mathbf{J}_\phi = S_\phi$$

The variables in this equation are defined as follows:

Variable	Description
ϕ	Generic variable
t	Time
$\mathbf{u}$	Velocity vector
$\mathbf{J}_\phi$	Diffusive flux of ϕ
S_ϕ	Volumetric source (or sink) of ϕ

This generic form can be used to represent almost any conservation law, a few of which will be discussed in the next section. This will be the general form that we will work with when developing our finite volume discretization scheme.

> **Tip:** Review the advection-diffusion equation (https://en.wikipedia.org/wiki/Convection%E2%80%93diffusion_equation), sometimes called the convection-diffusion equation, if it is unfamiliar to you.

Mass Conservation Equation

The mass conservation equation can be formulated by setting $\phi = \rho$, where ρ is the density of the substance. For mass conservation of a continuous phase, there is no diffusive flux, so $J_\rho = 0$. The mass conservation equation is therefore:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\mathbf{u}\rho) = S_\rho$$

The source term, S_ρ, is zero in many cases where there are no mass sources or sinks in the domain. A common special case for the mass conservation equation is that for incompressible constant density flows. In this case, since ρ is constant, $\partial\rho/\partial t = 0$, and ρ can be taken outside of the divergence term. Taking ρ to the right side of the equation results in:

$$\nabla \cdot \mathbf{u} = \frac{S_\rho}{\rho}$$

If there are no mass sources or sinks, then the mass conservation equation (or continuity equation) is:

$$\nabla \cdot \mathbf{u} = 0$$

It should be noted that the diffusive flux may be included for dispersed phases existing within a continuous phase (e.g. soluble particles in a liquid).

Momentum Conservation Equation

The momentum conservation equation is formed by letting $\phi = \rho \mathbf{u}$. The diffusive flux term is defined in general as $J_\mathbf{u} = -\nabla \cdot \sigma$, where σ is the fluid stress tensor. The momentum conservation equation can then be expressed in a general form as:

$$\frac{\partial (\rho \mathbf{u})}{\partial t} + \nabla \cdot (\rho \mathbf{u} \mathbf{u}) = \nabla \cdot \sigma + S_\mathbf{u}$$

The stress tensor for a fluid can be expressed in terms of the pressure and the viscous stress tensor:

$$\sigma = -p\mathbf{I} + \tau$$

where $\mathbf{I}$ is the identity matrix. Substituting this into the momentum conservation equation results in:

$$\frac{\partial (\rho \mathbf{u})}{\partial t} + \nabla \cdot (\rho \mathbf{u} \mathbf{u}) = -\nabla p + \nabla \cdot \tau + S_\mathbf{u}$$

For an incompressible Newtonian fluid, the stress tensor is simplified to:

$$\tau = \mu \left(\nabla \mathbf{u} + \nabla \mathbf{u}^T \right)$$

where μ is the dynamic viscosity of the fluid and the superscript T denotes the transpose of a matrix. With this definition (and the mass conservation equation for an incompressible flow with no sources/sinks), the momentum conservation equation for flow of an incompressible, Newtonian fluid with constant viscosity is given as:

$$\frac{\partial (\rho \mathbf{u})}{\partial t} + \nabla \cdot (\rho \mathbf{u} \mathbf{u}) = -\nabla p + \mu \nabla^2 \mathbf{u} + S_\mathbf{u}$$

Energy Conservation Equation

To develop an energy conservation equation, we let $\phi = \rho h$, where h is the specific enthalpy of the substance at a given state. With this definition, ϕ represents energy per unit volume. The diffusive flux is given by Fourier's law as $J = -k \nabla T$, where k is the thermal conductivity of the substance. Putting this all togehter, we arrive at:

$$\frac{\partial \left(\rho h \right)}{\partial t} + \nabla \cdot (\rho \mathbf{u} h) = \nabla \cdot (k \nabla T) + S_h$$

For an incompressible substance (where $h = c_p T$, under the assumption of constant specific heat capacity) with constant thermophysical properties (thermal conductivity and density) and no source terms, the energy equation is:

$$\frac{\partial T}{\partial t} + \nabla \cdot (\mathbf{u}T) = \alpha \nabla^2 T$$

where α is the thermal diffusivity.

Discretization of the Generic Conservation Equation

One-Way and Two-Way Coordinates

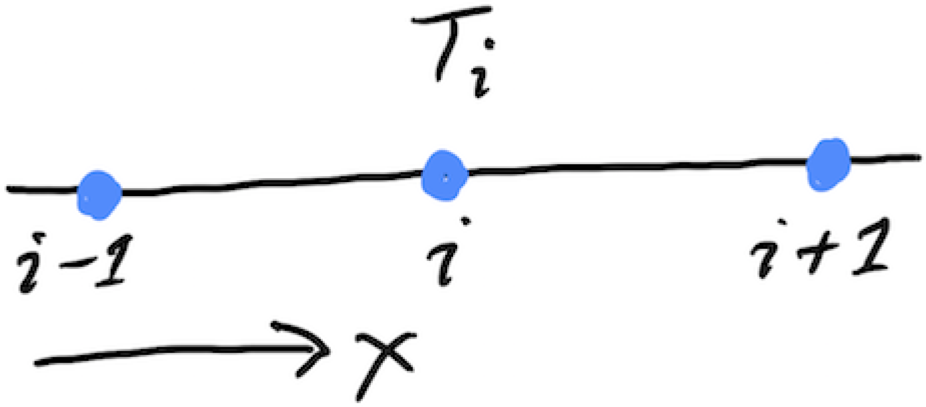
The dependent variable ϕ will, in general, be a function of three spatial coordinates, as well as time, i.e.:

$$\phi = \phi(\mathbf{x}, t)$$

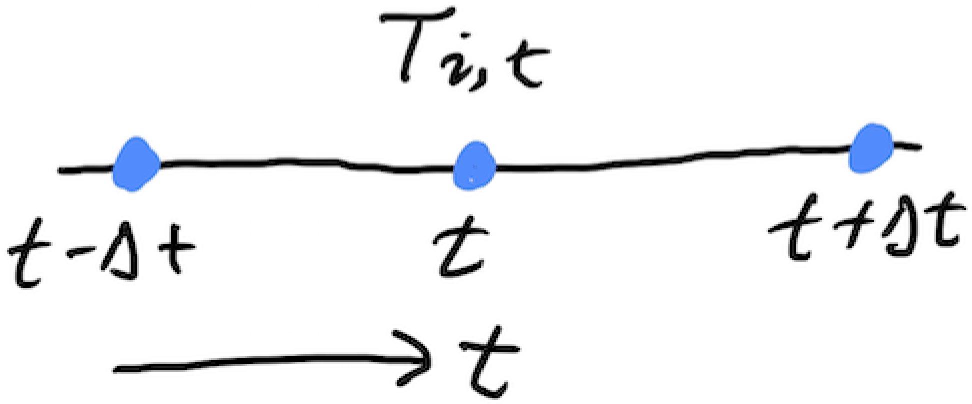
where $\mathbf{x} = (x, y, z)$ is the position vector. Considering the example of a variable that varies only in one dimension, coordinates can be considered either "one-way" or "two-way" depending on how the variable is influenced:

- One-way: changes in ϕ are influenced by changes on only one side of that location
- Two-way: changes in ϕ are influenced by changes on both sides of that location

For example, consider heat conduction, as shown below. The temperature T_i is influenced by changes in T on either side of the point i. Therefore, $\mathbf{x}$ is a two-way coordinate for heat conduction.



Taking another example of transient heat convection/conduction. The temperature at any given instant in time can only be influenced by changing the conditions that existed before that point in time. Therefore t is a one-way coordinate in this case.

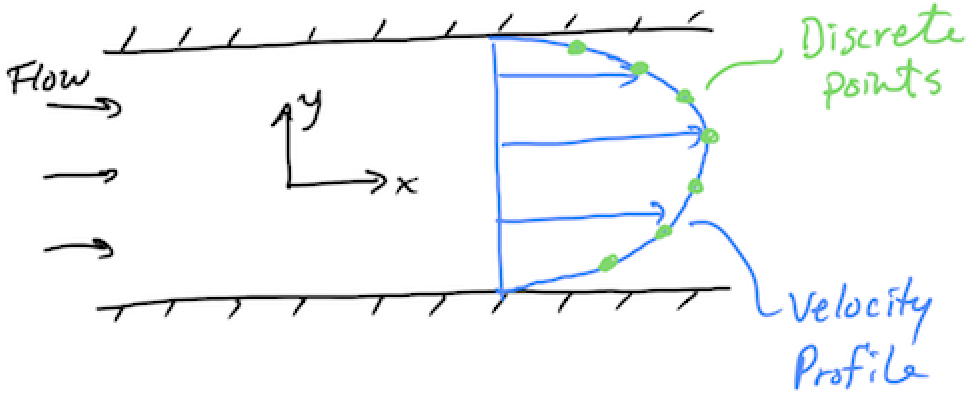


In terms of the generic conservation equation shown previously, we can consider the diffusion term to be two-way in space (as discussed above for heat conduction), while the convection term is one-way in space. This can be reasoned by considering the wind blowing over a source of smoke. The influence of the smoke can only be felt in the downwind direction, making this a one-way spatial process. Note that in the limit of a very slow wind, diffusion can cause the smoke to travel in the upwind direction, however the argument regarding the convection process still holds.

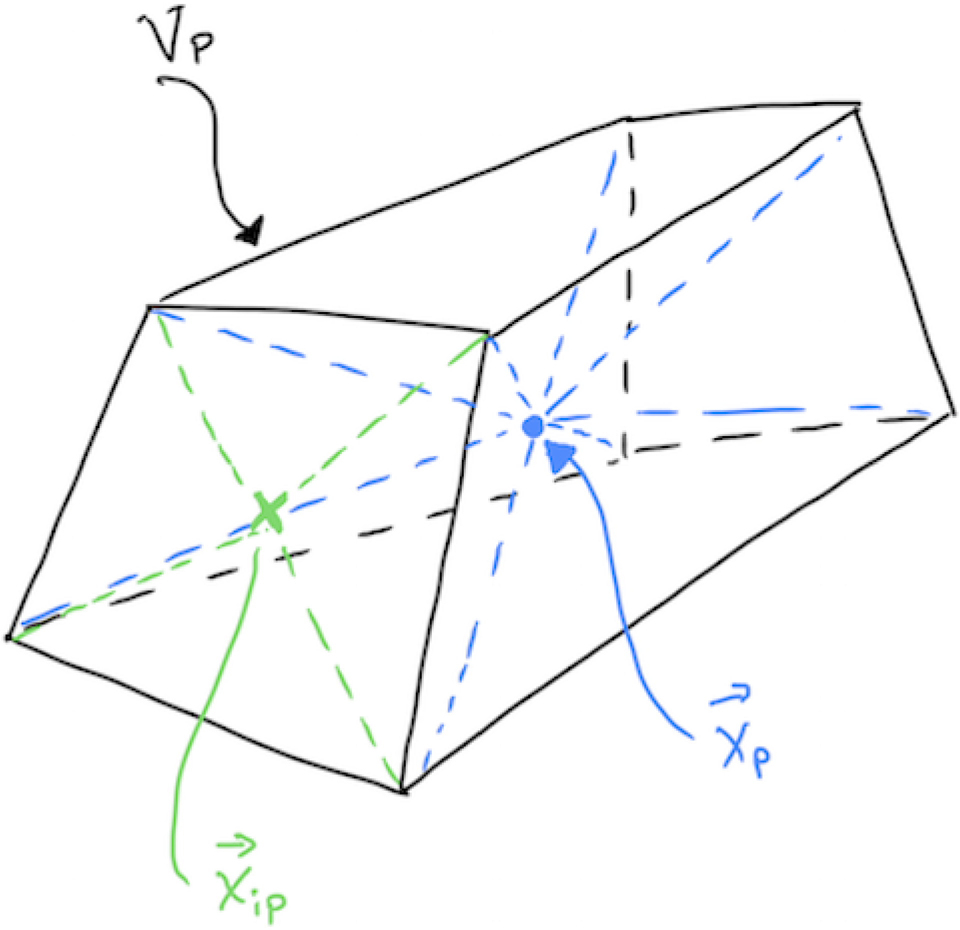
In mathematical terms, **parabolic** differential equations represent one-way behaviour, while **elliptic** differential equations are two-way. Without getting too deep into the mathematics, we will be considering two-way spatial behaviour and one-way time behaviour within the context of this course.

Discretization Concept

The concept behind the discretization of the governing PDEs is to replace their continuous solution with a discrete solution, at specific locations within the domain, that suitably approximates the continuous solution. This is similar to conducting an experiment, where, for example, the velocity can only be sampled at a finite number of locations to approximate the full profile. The figure below, shows the example of a velocity profile which is sampled at a specific set of points, which could either represent discrete experimental measurements, or discrete points in a numerical solution.



In the finite-volume method, the domain is discretized into a set of non-overlapping finite regions which fill the space of the domain under consideration. Discrete points are then located at the centroid of each control volume, the position of which is denoted $\mathbf{x}_P$, as shown in the figure below. The discrete solution is solved and stored at each of these points. Surrounding each one of these control volumes (or "cells"), there is a discrete set of faces. At the centre of each face is an "integration point" at which fluxes entering or leaving the cell are calculated. The position of each integration point on each face is denoted $\mathbf{x}_{ip}$. The governing equations are then integrated over a control volume, resulting in an integral conservation equation. Terms involving the divergence of a vector field are transformed into surface integrals using Gauss' (divergence) theorem. The final result is an equation expressing the balance of surface flux terms and volume source terms.



Tip: Review [Gauss' (divergence) theorem (https://en.wikipedia.org/wiki/Divergence_theorem)](https://en.wikipedia.org/wiki/Divergence_theorem), as this will be frequently used in our discretization processes.

Location of the Cell Centre and Face Integration Points

It was stated already that there are discrete points located at the centroid of each control volume, representing the locations of the solution variables. There are also points located on each face where fluxes are evaluated, which were also stated to be at the geometric centroid of the face. Here, we consider the importance of locating these points at their respective centroids, as a requirement for maintaining accuracy of the method.

As an illustrative example, let us consider the volume integral of a quantity ϕ. Suppose that we want to represent this integral in discrete form. Intuitively we might assume

$$\int_V \phi dV \approx \phi_P V_P$$

where ϕ_P is the value of ϕ at some internal location within V and V_P is the total volume of the cell, defined as

$$V_P = \int_V dV$$

To show this, we expand ϕ in a Taylor series about the point P. This results in

$$\int_V \phi dV \approx$$
$$\int_V \left[\phi_P + \nabla\phi_P \cdot (\mathbf{x} - \mathbf{x}_P) + \nabla\nabla\phi_P : (\mathbf{x} - \mathbf{x}_P)(\mathbf{x} - \mathbf{x}_P) \right.$$
$$\left. + O(\delta^3)\right] dV$$

where δ is used to represent the characteristic grid spacing. The quantity ϕ_P and its derivatives are constants, evaluated at the point P, and can therefore be taken outside of the integral. This results in

$$\int_V \phi dV \approx \phi_P V_P + \nabla\phi_P \cdot \int_V (\mathbf{x} - \mathbf{x}_P) dV + \nabla\nabla\phi_P :$$
$$\int_V (\mathbf{x} - \mathbf{x}_P)(\mathbf{x} - \mathbf{x}_P) dV + O(\delta^3) V_P$$

The first term on the right side of the previous equation matches with our intuitive assumption that the volume integral of ϕ equals $\phi_P V_P$, however there are additional terms appearing in the equation. The definition of a centroid requires that

$$\int_V (\mathbf{x} - \mathbf{x}_P) dV = 0$$

Therefore, provided that the point P is located at the centroid of the control-volume, the equation simplifies to

$$\int_V \phi dV \approx \phi_P V_P + \nabla\nabla\phi_P : \int_V (\mathbf{x} - \mathbf{x}_P)(\mathbf{x} - \mathbf{x}_P) dV$$
$$+ O(\delta^3) V_P$$

The order of magnitude of the integrand in the remaining integral has magnidude $O(\delta^2)$. Therefore, the final term can be neglected in comparison to this one, since as $\delta \to 0$, this term is negligible in comparison to the previous term. This results in

$$\int_V \phi dV \approx \left[\phi_P + O(\delta^2)\right] V_P$$

This result implies that there is a second-order error resulting from the approximation of the integral as $\phi_P V_P$, which is acceptable when the order of accuracy of the computation method is also second-order. That is to say, the error in this approximation is no greater than the approximations made elsewhere in the numerical method.

If the point P were not located at the centroid of the cell, then the second term in the Taylor expansion would not go to zero and the resulting approximation would have a first-order error, which would reduce the accuracy of the numerical method.

A similar analysis can be made to show that the integration points must also be located at the face centroids to ensure second-order accuracy, however, this will not be shown here.

> **Warning:** The conclusion of this analysis is that one must take care to properly calculate the cell and face centroids when developing the code for CFD analysis. Particularly for general polyhedral meshes with polygonal faces, one must not be tempted to make careless approximations, such as assuming the centroid to be at the geometric mean of the corner locations, as this will cause additional numerical error.

Transient Term

There are a number of different ways that the transient term can be discretized, depending on the order of accuracy and whether or not the discretization scheme is implicit or explicit. For this course we will focus only on implicit methods, although comparisons will be made with explicit methods later on. The general formulation for discretization of the transient term is obtained by integrating over the control volume V_P as well as the timestep $\Delta t = t_1 - t_0$ which results in

$$\int_{t_0}^{t_1} \int_V \frac{\partial \phi}{\partial t} dV dt \approx (\phi_P V_P)^{t_1} - (\phi_P V_P)^{t_0}$$

where the superscript indicates at which time value the quantities are evaluated. The type of transient discretization depends on how the variable is assumed to vary in time. This will be considered in a later lesson.

Advection Term

The general form of the advection term arises in a similar manner by integrating over the control volume V_P. In this case, Gauss' theorem is employed to convert the volume integral into a surface integral, according to

$$\int_V \nabla \cdot (\mathbf{u}\phi)\, dV = \int_S (\mathbf{u}\phi) \cdot \mathbf{n} dS$$

The surface integral is then approximated by a discrete summation over the faces surrounding the cell, each with area A_{ip}:

$$\int_S (\mathbf{u}\phi) \cdot \mathbf{n}_{ip} dS \approx \sum_{i=0}^{N_{ip}-1} \mathbf{u}_{ip} \cdot \mathbf{n}_{ip} \phi_{ip} A_{ip}$$

In order to complete the discretization of this term, methods for approximating $\mathbf{u}_{ip}$ and ϕ_{ip} must be specified. While the specification of $\mathbf{u}_{ip}$ can be done by any wide number of interpolation methods, interpolation of ϕ_{ip} must be considered carefully in order to obtain a stable numerical method. The topic of advection schemes will be covered separately in a later lesson.

Diffusion Term

The general form of the diffusion term for the quantity ϕ is $\nabla \cdot \mathbf{J}_\phi$. Integrating this over a control volume, using Gauss' theorem to convert the volume integral into a surface integral, results in:

$$\int_V \nabla \cdot \mathbf{J}_\phi dV = \int_S \mathbf{J}_\phi \cdot \mathbf{n} dS$$

The surface integral is then approximated by a discrete summation over the faces surrounding the cell:

$$\int_S \mathbf{J}_\phi \cdot \mathbf{n} dS \approx \sum_{i=0}^{N_{ip}-1} \mathbf{J}_{\phi,ip} \cdot \mathbf{n}_{ip} A_{ip}$$

The face flux, $\mathbf{J}_{\phi,ip}$, must be interpolated from the neighbouring cell values. Methods for conducting this interpolation will be discussed in later lessons.

Source Term

The source term is discretized by assuming that the source value is piecewise continuous, with one specific value being representative for each cell. This results in:

$$\int_V S_\phi dV \approx S_\phi V_P$$

In general, the source term may be dependent on ϕ itself, which requires linearization in most cases to obtain a stable numerical method.

Linearization

In general, the discretized terms above will depend non-linearly on the solution itself. Non-linearity can arise from a number of different factors including:

- Source terms that depend non-linearly on the primitive variable
- Non-linearities in the governing equations (i.e. advection term in momentum equation)
- Gradient correction terms that are required on non-orthogonal grids

To linearize, let us consider the governing PDE to be represented by the general differential operator

$$L(\phi^*) = 0$$

where ϕ^* represents the continuous solution to the PDE. To solve a PDE using the finite-volume method, the continuous solution, ϕ^*, is first approximated by a discrete solution vector $\phi \in \mathbb{R}^N$, where N is the number of control volumes in the discretized domain. The PDE is then integrated over each control volume and each term is approximated using the discrete solution ϕ. In general, the solution will not exactly satisfy the discretized equation; rather it will have a residual, $\mathbf{r} \in \mathbb{R}^N$. To solve the system of equations, Newton's method is used to linearize the problem. To do so, the residual vector is expanded in a Taylor series about the solution at a particular iteration, i, denoted ϕ_i. The goal is to find the solution where $\mathbf{r} = 0$. This procedure results in

$$\mathbf{r}(\phi_i) + \left.\frac{\partial \mathbf{r}}{\partial \phi}\right|_{\phi_i} (\phi - \phi_i) = 0$$

The Jacobian of the residual vector is denoted as

$$\mathbf{J}(\phi) = \frac{\partial \mathbf{r}}{\partial \phi}$$

which defines a linear system that is to be solved for the solution correction $\Delta\phi = (\phi - \phi_i)$, according to

$$\mathbf{J}\,(\phi_i)\,\Delta\phi = -\mathbf{r}(\phi_i)$$

The solution is then updated according to the fixed point iteration

$$\phi = \phi_i + \Delta\phi_i$$

What remains is to calculate both the residual vector and the Jacobian matrix. This will be considered term by term in the lessons that follow.

It can be useful to express the algebraic equation that defines the linear system for a particular control volume P as

$$a_P \delta\phi_P + \sum_{nb} a_{nb} \delta\phi_{nb} = -r_P$$

where the sum over nb represents the sum over all neighbouring cells (i.e. those sharing a face, where fluxes can be exchanged). The coefficients are defined as:

$$a_P = \frac{\partial r_P}{\partial \phi_P}$$

$$a_{nb} = \frac{\partial r_P}{\partial \phi_{nb}}$$

Four Basic Rules

There are four basic rules that have been outlined by Patankar (1980), which are useful when completing the discretization of the terms described previously. A summary of those rules is:

Rule 1: Consistency at control volume faces The flux through common faces that lie between cells must be the same when evaluated for each cell. The simplest way to ensure flux consistency is to use an identical expression for calculating the flux. The reason for this rule should be clear; if the flux between faces is not consistent, that means there is an artificial source of energy at the face.

Rule 2: Positive a_P coefficient and negative a_{nb} coefficients In situations involving convection and diffusion, it is reasonable to assume that an increase in ϕ and one control volume should lead to an increase in ϕ at neighbouring control volumes, where all other conditions are the unchanged. This requires that a_P has opposite sign from each of its a_{nb} coefficients, such that $\delta\phi_P$ and $\delta\phi_{nb}$ have the same sign, where r_P is assumed to be unchanged. Note that this rule is modified from that of Patankar, since we have kept all nb coefficients on the same side of the equation as the a_P coefficient. This is more natural in the context of the formulation given here, where the algebraic equations are formulated in terms of solution corrections rather than the solution values themselves. Although these approaches are closely related, there are some advantages to the correction-based approach (e.g. and initial guess of $\Delta\phi = 0$ can always be made since this is correct at convergece).

> **Tip:** Use Rule 2 to help debug your code, by ensuring that all coefficients have the correct sign.

Rule 3: Negative slope linearization of source terms Suppose we have a source term that depends on ϕ_P as $S_\phi = a + b\phi_P$. If this term is moved to the left side of the equation (as it will be, according to the sign convention we have adopted) the coefficient a_P could become negative if b is positive. Therefore we require that $b < 0$ (or we must neglect linearization of this term). This rule is not as arbitrary as it may appear, since most physical processes actually do have a negative slope relationship between the source term and the field variable. A positive slope relationship would be unstable since the source would cause the variable to increase, which further increases the source term. This would continue indefinitely and without bound. It would be possible to have a heat source that grows with temperature, but there would also need to be another source term for heat removal, to avoid an uncontrolled increase in temperature.

Rule 4: Sum of the neighbour coefficients Governing equations often contain only derivatives of the dependent variables, meaning both ϕ and $\phi + c$, where c is a constant, will satisfy the same governing equations. This implies that temperatures in both Celsius and Kelvin would both satisfy the same discretized equaitons. In order for this to be true, we may say that:

$$a_P = -\sum_{nb} a_{nb}$$

This is not always satisfied, since the linearization of a source term, as discussed above, only contributes to a_P. This is not a violation of the rule, but indicates that in such a case, the same equations cannot be used for both ϕ and $\phi + c$. This is quite reasonable, given the fact that one cannot change units from Celsius to Kelvin without modifying the source term coefficients appropriately.

> **Warning:** When solving equations with source terms, be aware of units.

Comparison to Other Methods

The finite volume discretization method is one of many possible methods that can be used to discretize a governing equation. Some other methods include finite difference and finite element methods. While we will not go into detail regarding these methods, we will highlight two important features of the finite volume method that make it particularly attractive for solving fluid mechanics problems:

- In the finite volume method, integral conservation of dependent variables is satisfied exactly over each control volume, any group of control volumes, and the entire domain, regardless of grid resolution. This cannot be guaranteed in general for other methods.

- The solution values are the values of the dependent variable at the control volume centroids; in the finite element method, finding the interpolation functions is a part of the solution.

Uncertainty and Error

There are a number of uncertainties and errors that arise from the discretization and solution of a system of equations using the finite volume method (or any numerical method for that matter).

Exercise: Read the article Uncertainty and Error in CFD Simulations (https://www.grc.nasa.gov/www/wind/valid/tutorial/errors.html) published on NASA's website.

Here we will mainly focus on acknowledged errors, since unacknowledged errors (programming or usage errors) should be checked through appropriate use of verification and validation cases. Descriptions of verification and validation are as follows:

- **Verification** involves checking that the computer code correctly solves the mathematical model; answers the question "am I solving the equations correctly?"
- **Validation** involves checking that the solution of the mathematical model agrees with reality; answers the question"am I solving the right equations?"

In terms of acknowledged errors, the following types of error are listed in the article linked above:

- Physical approximation error
    - Physical modelling error
    - Geometry modelling error
- Computer round-off error
- Iterative convergence error
- Discretization error
    - Spatial discretization error
    - Temporal discretization error

For now, we will not consider physical approximation error, since we will start with very simple model geometries without any complex physical models, such that physical approximation error should be negligible. Furthermore, we will not consider computer round-off error since Python implements 53-bits of precision for floating-point numbers (https://docs.python.org/3.6/tutorial/floatingpoint.html), meaning this error should be far less than other possible errors. Iterative convergence error is very important, however, discussion will be held for future lessons and will be illustrated by example. That leaves us to discuss discretization error.

Discretization error includes both spatial and temporal discretization errors which result from the fact that we can only evaluate our solution at a finite number of points in space and time. It is not possible to estimate discretization error *a priori*, so it is necessary to test the sensitivity of solutions to the selection of spatial and temporal discretization parameters. Spatial discretization error is most often evaluated using a **grid independence study** (also known as a grid convergence study). A typical grid independence study is outlined as follows:

1. Conduct a simulation on a coarse grid.
2. Evaluate a measurable quantity of interest (e.g. overall pressure drop or heat transfer rate).

3. Increase the number of control volumes in the domain by a significant amount (generally 1.5-2 times increase).
4. Compare the quantity of interest from the current simulation to that from the previous simulation.
5. If quantity changes by more than a specified tolerance, return to (3); otherwise the solution is independent of the grid to within the specified tolerance.

While the procedure above is an effective way to determine whether a solution is independent of the grid, more information can be gained on the discretization error using the "Grid Convergence Index" approach of Celic et al. (http://dx.doi.org/10.1115/1.2960953).

> **Exercise:** Read the article "Procedure for Estimation and Reporting of Uncertainty Due to Discretization in CFD Applications" (http://dx.doi.org/10.1115/1.2960953) published by the ASME Journal of Fluids Engineering.

Temporal discretization errors will be discussed later when we look into transient discretization schemes in more detail.

Next Steps

Now that you have read a little bit about the basics of the finite volume method, you are ready to move on to the next lesson on Steady One-Dimensional Heat Diffusion (2-SteadyDiffusion.ipynb) where we will implement code to solve this type of problem.

In [ ]:

```

```