Dr. Jeremy Roberts

NE 860

Numerical Neutron Transport

Lecture Notes and Problems

Kansas State University

Created Spring 2011. Last updated: April 15, 2015

Contents

Lecture | Introduction

This set of lectures derives from the class 22.106, Neutron Interactions and Applications, formerly taught at MIT. I wrote these while a TA under my advisor Dr. Benoit Forget, the instructor of the course. I continued to develop the notes as I taught 22.106 myself during my final semester of study as an Instructor-G (which happens also to be the last time the course was taught, for better or worse).

Although 22.106 covered a wide range of topics, ranging from the R-matrix theory to deterministic transport, these notes are limited to aspects of neutron transport that are important for reactor physics and shielding. The primary focus is on *numerical* methods, but some analytical techniques are developed where appropriate.

I've drawn from the original content of 22.106 along with a course I had previously at Wisconsin (with Dr. Douglass Henderson) and my own research. Some of the lectures are in better shape than others, and I intend to update them all as time allows and as I receive input from my own students.

J. Roberts
Manhattan, KS

Part I

Deterministic Transport

Lecture 1 | The Transport Equation

In this lecture, we introduce general *transport equations* and finish by developing the form used in neutron transport theory. In the lectures to follow, we shall describe other aspects of the neutron transport equation and methods by which it can be solved both analytically and numerically.

Transport Theory

Transport theory aims to describe mathematically the movement (i.e. "transport") of particles as they traverse a medium. For example, we might describe the transport of high energy gammas through a lead shield, or the movement of neutrons through uranium dioxide pellets. We might also describe the movement of particles in a dense gas as they navigate through a medium consisting of the gas itself.

In all cases, transport theory describes such processes in an *average* sense. For instance, we do not compute the individual trajectories of neutrons in a reactor via transport theory. That, instead, would require molecular dynamics, in which Newton's equation of force is solved for the many-bodied problem of all neutrons in the vicinity of interest (an essentially impossible problem), or perhaps the Monte Carlo methods described in previous lectures, where a sample of individual particles are tracked to approximate ensemble averages (a difficult, but as we've seen, tractable problem). Hence, the quantities we shall compute using the equations of transport theory should be recognized as expected and not exact values.

Fundamental Quantities

We begin by defining several fundamental quantities. It should be noted that the forms introduced at first are likely different from what you might have seen in a previous reactor physics course. The purpose here is two-fold. First, we wish to introduce the quantities and eventually the equations in a general way to make clear that transport theory is not restricted to the neutron transport equation. Second, for

those who might be familiar with, e.g., the Boltzmann equation of gas dynamics (and not neutron transport), the notation will be familiar and lead smoothly into our more familiar form.

We first define the *phase space density function*, the knowledge of which we can use to compute essentially all quantities of interest:

$$n(\mathbf{r},\mathbf{v},t)d^3rd^3v \equiv \quad \text{expected number of particles in } d^3r \text{ about } \mathbf{r} \text{ with velocity } dv \text{ about } \mathbf{v} \text{ at time } t.$$

It is often most convenient to break the velocity into its scalar (speed) and vector (direction) components. The scalar component is recast in the energy variable via $E = mv^2/2$, and the direction vector is defined $\mathbf{\Omega} = \mathbf{v}/|\mathbf{v}|$. The phase space density can then be rewritten as

$$n(\mathbf{r},\mathbf{\Omega},E,t)d^3rd\Omega dE \equiv \quad \text{expected number of particles in } d^3r \text{ about } \mathbf{r} \text{ going in the directions } d\Omega \text{ about } \mathbf{\Omega} \text{ with energy } dE \text{ about } E \text{ at time } t.$$

In this form, $n(\mathbf{r},\mathbf{\Omega},E,t)$ is often referred to as the angular density.

We can relate the phase space densities in terms of $\mathbf{v}$ and $(E,\mathbf{\Omega})$ via the relations

$$\begin{aligned} n(\mathbf{r},E,\mathbf{\Omega},t) &= (v/m)n(\mathbf{r},\mathbf{v},t) \\ n(\mathbf{r},E,\mathbf{\Omega},t) &= (1/mv)n(\mathbf{r},v,\mathbf{\Omega},t) \\ n(\mathbf{r},v,\mathbf{\Omega},t) &= v^2 n(\mathbf{r},\mathbf{v},t)\,, \end{aligned} \tag{1.1}$$

proofs of which are left as exercises.

Figure 1.1 depicts a schematic of the phase space used in terms of the position $\mathbf{r}$ and direction $\mathbf{\Omega}$. The position vector is further broken down into the polar angle θ and azimuthal angle ϕ. The differential solid angle element $d\Omega$ is also shown, and can be expressed in terms of θ and ϕ via

$$d\Omega = \sin\theta d\theta d\phi\,.$$

A quantity closely related to the phase space density (or current density) is the *angular current density*:

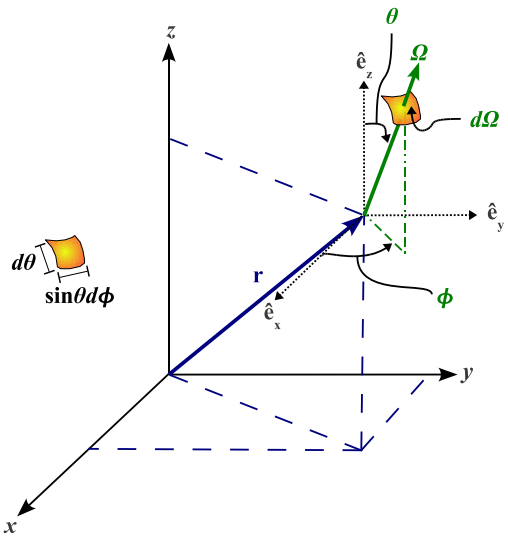

Figure 1.1: Schematic of Phase Space.

$\mathbf{j}(\mathbf{r},\mathbf{v},t)\cdot d\mathbf{S}d^3v = \mathbf{v}n(\mathbf{r},\mathbf{v},t)\cdot d\mathbf{S}d^3v \equiv$ expected number of particles that cross an area dS per second with velocity d^3v about $\mathbf{v}$ at time t.

We can also define *partial currents* with respect to a particular surface S defined by an outward normal vector $\hat{\mathbf{e}}_s$:

$J_\pm(\mathbf{r},t) = \pm\int_\pm d^3v\hat{\mathbf{e}}_s\cdot\mathbf{j}(\mathbf{r},\mathbf{v},t) \equiv$ the rate at which particles flow through S in the outward (+) or inward (-) direction.

The *current density* $\mathbf{J}(\mathbf{r},t)$ is defined by integrating the angular current density over all velocities. Then for our surface S, the net rate of particles passing outward through S is just $\mathbf{J}(\mathbf{r},t)\cdot\hat{\mathbf{e}}_s$. From our definition of partial currents, the net current passing outward must also be $J_+ - J_-$, yielding the useful identity

$$\mathbf{J}(\mathbf{r},t)\cdot\hat{\mathbf{e}}_s = J_+(\mathbf{r},t) - J_-(\mathbf{r},t)\,. \tag{1.2}$$

Of particular interest to us in the next section will be reaction rates, which can most easily be described using the *angular flux*

$$\psi(\mathbf{r},\mathbf{\Omega},E,t) = vn(\mathbf{r},\mathbf{\Omega},E,t) \equiv \quad \text{angular flux}$$

and *scalar flux*

$$\phi(\mathbf{r},E,t) = \int_{4\pi} d\Omega\psi(\mathbf{r},\mathbf{\Omega},E,t) \equiv \quad \text{scalar flux.}$$

A General Transport Equation

Consider an arbitrary volume V with a surface S. Our goal is to represent the time rate of change of the particle density $n(\mathbf{r},\mathbf{v},t)$ within the volume. Neglecting external forces, the only factors affecting the density are collisions within the volume that change a particle's velocity, the streaming of particles into and out of the surface, and any internal source of particles. This simple balance can be expressed mathematically as

$$\overbrace{\frac{\partial}{\partial t}\left(\int_V d^3r n(\mathbf{r},\mathbf{v},t)\right)}^{\text{total rate of change of } n \text{ in } V} = -\overbrace{\int_S dS\hat{\mathbf{e}}_S\cdot\mathbf{j}(\mathbf{r},\mathbf{v},t)}^{\text{streaming rate}} + \overbrace{\int_V d^3r\left(\frac{\partial n}{\partial t}\right)_{\text{coll}}}^{\text{collision rate}} + \underbrace{\int_V d^3r s(\mathbf{r},\mathbf{v},t)}_{\text{source emission rate}}\,, \tag{1.3}$$

where s represents a source inside the volume and $(\partial n/\partial t)_{\text{coll}}$ is the time rate of change due to collisions, specific forms of which are application-dependent and will be discussed below. Note the minus sign on the surface integral, i.e., the streaming term. Since the integral describes the net rate of neutrons going *out* of the surface, we negate it so that a positive net rate directed inward is a positive contribution to the total time rate of change of n in V.

Eq. (12.1) gives us a simple relation in terms of both volume and surface integrals. Our life is always easiest if we have the same integration on both sides. By the divergence (or Gauss') theorem, we can rewrite the streaming term

$$\int_S dS\hat{\mathbf{e}}_S \cdot \mathbf{j}(\mathbf{r},\mathbf{v},t) = \int_V d^3r \nabla \cdot \mathbf{j}(\mathbf{r},\mathbf{v},t)\,. \tag{1.4}$$

Because ∇ acts on $\mathbf{r}$ and not $\mathbf{v}$, we note $\nabla\cdot\mathbf{j} = \nabla\cdot(\mathbf{v}n) = \mathbf{v}\cdot\nabla n + \overbrace{n\nabla\cdot\mathbf{v}}^{\to 0} = \mathbf{v}\cdot\nabla n$. Hence, the streaming term becomes

$$\int_V d^3r \nabla \cdot \mathbf{j}(\mathbf{r},\mathbf{v},t) = \int_V d^3r \mathbf{v}\cdot\nabla n(\mathbf{r},\mathbf{v},t)\,. \tag{1.5}$$

For a constant volume, $(\partial/\partial t)\int_V d^3rn = \int_V d^3r(\partial n/\partial t)$, and so our balance equation can be rewritten as

$$\overbrace{\left(\int_V d^3r \frac{\partial}{\partial t} n(\mathbf{r},\mathbf{v},t)\right)}^{\text{total rate of change of } n \text{ in } V} = -\overbrace{\int_V d^3r \mathbf{v}\cdot\nabla n(\mathbf{r},\mathbf{v},t)}^{\text{streaming rate}} + \overbrace{\int_V d^3r \left(\frac{\partial n}{\partial t}\right)_{\text{coll}}}^{\text{collision rate}} + \underbrace{\int_V d^3r s(\mathbf{r},\mathbf{v},t)}_{\text{source emission rate}}\,. \tag{1.6}$$

For an arbitrary volume V, the integrands of Eq. (1.6) must vanish, yielding a general transport equation:

$$\frac{\partial}{\partial t} n(\mathbf{r},\mathbf{v},t) = -\mathbf{v}\cdot\nabla n(\mathbf{r},\mathbf{v},t) + \left(\frac{\partial n}{\partial t}\right)_{\text{coll}} + s(\mathbf{r},\mathbf{v},t)\,. \tag{1.7}$$

Even More Generality

We can skip the differential volume formulation by considering the material derivative of n (using Cartesian coordinates):

$$\begin{aligned}\frac{Dn}{Dt} &\equiv \frac{\partial n}{\partial t} + \frac{\partial x}{\partial t}\frac{\partial n}{\partial x} + \frac{\partial y}{\partial t}\frac{\partial n}{\partial y} + \frac{\partial z}{\partial t}\frac{\partial n}{\partial z} + \frac{\partial v_x}{\partial t}\frac{\partial n}{\partial v_x} + \frac{\partial v_y}{\partial t}\frac{\partial n}{\partial v_y} + \frac{\partial v_z}{\partial t}\frac{\partial n}{\partial v_z} \\ &= \frac{\partial n}{\partial t} + \mathbf{v}\cdot\nabla n + \mathbf{a}\cdot\nabla_{\mathbf{v}} n \\ &= \frac{\partial n}{\partial t} + \mathbf{v}\cdot\nabla n + \frac{\mathbf{F}}{m}\cdot\nabla_{\mathbf{v}} n\,,\end{aligned} \tag{1.8}$$

where $\nabla_{\mathbf{v}}$ is the gradient operator with respect to velocity components (rather than spatial coordinates). The material derivative is the total time rate of change, accounting for convective (streaming) effects as well as the influence of an external force $\mathbf{F}$, a factor we did not account for above.

This total rate of change must be balanced by sources and sinks, which are the collision and internal source terms. Hence, an even more general transport equation can be written

$$\frac{\partial n}{\partial t} + \mathbf{v}\cdot\nabla n + \frac{\mathbf{F}}{m}\cdot\nabla_{\mathbf{v}} n = \left(\frac{\partial n}{\partial t}\right)_{\text{coll}} + s\,. \tag{1.9}$$

Neutron Transport

To arrive at the neutron transport equation, we bring back the macroscopic cross-sections studied in Lecture **??**. Using our definition for the angular flux, the volumetric collision rate at a particular point in phase space and time is simply

$$R_{\text{coll}}(\mathbf{r},\boldsymbol{\Omega},E,t) = \psi(\mathbf{r},\boldsymbol{\Omega},E,t)\Sigma_t(\mathbf{r},E)\,. \tag{1.10}$$

However, we know that neutrons at one energy and angle can scatter into another energy and angle, and so in general, the rate at which neutrons at any angle and energy are scattered into a particular energy and angle is

$$R_{\text{in-scatter}}(\mathbf{r},\boldsymbol{\Omega},E,t) = \int_0^\infty dE' \int_{4\pi} d\Omega' \Sigma_s(\mathbf{r},\boldsymbol{\Omega}\cdot\boldsymbol{\Omega}',E'\to E)\psi(\mathbf{r},\boldsymbol{\Omega}',E',t)\,. \tag{1.11}$$

The time rate of change due to collisions is thus

$$\begin{aligned}\left(\frac{\partial n}{\partial t}\right)_{\text{coll}} &= -\psi(\mathbf{r},\boldsymbol{\Omega},E,t)\Sigma_t(\mathbf{r},E) \\ &\quad + \int_0^\infty dE' \int_{4\pi} d\Omega' \Sigma_s(\mathbf{r},\boldsymbol{\Omega}'\to\boldsymbol{\Omega},E'\to E)\psi(\mathbf{r},\boldsymbol{\Omega}',E',t)\,.\end{aligned} \tag{1.12}$$

Substituting this into the general transport equation, using $\psi = vn$, and neglecting any external forces, we find the neutron transport equation:

$$\begin{aligned}\frac{1}{v}\frac{\partial\psi}{\partial t} + \hat{\Omega}\cdot\nabla\psi + \Sigma_t\psi(\mathbf{r},\mathbf{\Omega},E,t) = \\ + \int_0^\infty dE' \int_{4\pi} d\Omega'\Sigma_s(\mathbf{r},\mathbf{\Omega}\cdot\mathbf{\Omega}',E'\to E)\psi(\mathbf{r},\mathbf{\Omega}',E',t) + s\,.\end{aligned} \tag{1.13}$$

Note, the source term s has been used to represent internal sources but could also account for other sources such as fission (though the specific form is, of course, hidden).

Assumptions for the Neutron Transport Equation

In writing down Eq. 1.13 as we have, a number of assumptions have been made explicitly or implicitly. These include:

1. The neutron density is large so that it makes sense to be computing for mean values (which is all transport equations can provide)
2. The neutrons are point particles, meaning that wave effects are insignificant
3. Collisions are well-defined, two-body interactions that occur instantaneously (delayed neutrons from fission, which are not covered here, are a notable exception and deserve special treatment)
4. Between collisions, neutrons stream with a constant velocity
5. Neutron-neutron interactions are negligible
6. The properties of the medium are assumed known and time-independent (burnup in a reactor is another exception)
7. The medium is taken to be isotropic (i.e. no directional-dependence)

Initial Conditions

This is an integro-differential equation in 7 variables: 3 in space, 2 in angle, 1 in energy, and time. Like all differential equations, the transport equation requires initial and boundary conditions.

Initial conditions for the transport equation are relatively straightforward. At some initial time t_0, an initial condition is expressed as

$$\psi(\mathbf{r},\mathbf{\Omega},E,t)|_{t=t_0} = f(\mathbf{r},\mathbf{\Omega},E)\,, \tag{1.14}$$

where f represents a known function of space, angle, and energy.

Time-dependent problems in neutron transport are often quite challenging due to the wide range of time scales involved. A good example comes from the study of reactor kinetics, where the time scales range from prompt neutron lifetimes (on the order of 10^{-5} seconds) to the delayed neutrons of longest lifetime (on the order of tens of seconds). Any numerical scheme is effectively limited by the smallest time scale, leading to a "stiff" problem.

Other neutron transport problems exhibit even more diverse time scales. The time scale for nuclear weapons is perhaps most easily quantified with "shakes", that is 10^{-8} seconds. Isotopic changes in nuclear reactors due to irradiation can have profound effects on time scales ranging from hours (xenon production) up to months or years (burnup).

Frequently, we are interested in steady state values, so that $\frac{\partial \psi}{\partial t} = 0$. If the source s in Eq. 1.13 includes an external source, then the problem is a *fixed source problem*. If s does not include an external source (and includes only e.g. a fission source), then the problem becomes an *eigenvalue problem*, which are studied in Lecture **??**.

Boundary Conditions

The most straightforward boundary condition to enforce is the *free surface* or *vacuum* condition. Physically, the condition represents the situation where no neutrons enter a volume from the outside. In other words, the volume of interest can be thought to exist in a void. Mathematically, the condition is expressed

$$\psi(\mathbf{r}, \mathbf{\Omega}, E, t) = 0\,, \quad \hat{\mathbf{n}} \cdot \mathbf{\Omega} < 0\,, \tag{1.15}$$

where $\hat{\mathbf{n}}$ is the unit *outward normal* vector to the surface of interest. Since $\hat{\mathbf{n}} \cdot \mathbf{\Omega}$ is just the cosine of the angle between the incident neutrons and the normal vector, we see the flux vanishes whenever that cosine is negative, or whenever the neutron direction is inward.

A point of warning: reentrant geometries must be avoided when using vacuum conditions. Unless treated with special care, reentrant geometries lead to inconsistency. Neutrons leaving one portion of the geometry could, in theory, reenter another portion, but since vacuum conditions disallow this, the true problem is not modeled correctly. A common example of this occurs when "squaring" an exterior cylindrical boundary, as exhibited in Figure 1.2.

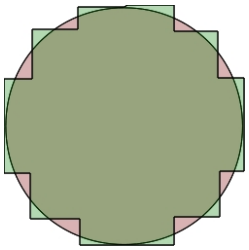

Figure 1.2: A reentrant square cyclinder.

Another useful boundary condition is simply to specify the incident flux when it is known:

$$\psi(\mathbf{r}_s, \mathbf{\Omega}, E, t) = f(\mathbf{r}_s, \mathbf{\Omega}, E, t)\,. \tag{1.16}$$

In this way, boundary sources can be defined.

A *reflective* or *specular* condition is such that

$$\psi(\mathbf{r}_s, \mathbf{\Omega}, E, t) = \psi(\mathbf{r}_s, \mathbf{\Omega}_R, E, t)\,, \quad \hat{\mathbf{n}} \cdot \mathbf{\Omega} < 0\,, \tag{1.17}$$

where $\mathbf{\Omega}_R$ is the (mirror) reflection of $\mathbf{\Omega}$. Reflective conditions are widely used in lattice physics, where pin cells or assemblies are modeled in an infinite array; the "infinite" is captured by the reflective conditions. See Figure 1.3.

A variation on reflective conditions is an *albedo* condition, where

$$\psi(\mathbf{r}_s, \mathbf{\Omega}, E, t) = \alpha\psi(\mathbf{r}_s, \mathbf{\Omega}_R, E, t)\,, \quad \hat{\mathbf{n}} \cdot \mathbf{\Omega} < 0\,. \tag{1.18}$$

Here, α is the "albedo" and quantifies the strength with which neutrons stream back into the system after streaming out. Historically, albedo conditions were highly useful since they can often capture the physics of reflectors with minimum computational cost. The albedos for many materials were precomputed (or found experimentally), effectively eliminating a significant portion of phase space in e.g. reactor analysis.

Another approach is to use *periodic* conditions, such that

$$\psi(\mathbf{r}_1, \mathbf{\Omega}, E, t) = \psi(\mathbf{r}_2, \mathbf{\Omega}, E, t)\,, \tag{1.19}$$

Figure 1.3 illustrates both reflective and periodic boundary conditions. Periodic conditions work well in infinite arrays that have assymetric unit cells (for which reflective conditions would represent an infinite but incorrect array).

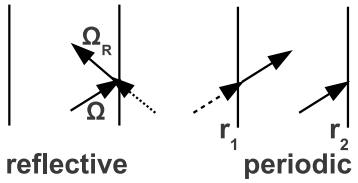

Figure 1.3: Reflective and periodic boundary conditions.

The final boundary condition we mention is the *white boundary condition*, a condition where all neutrons incident on a boundary reflect back isotropically in angle.

For this case,

$$\begin{aligned}\psi(\mathbf{r}_s, \mathbf{\Omega}, E, t) &= \frac{\int_{\hat{\mathbf{n}}\cdot\mathbf{\Omega}'>0} \hat{\mathbf{n}}\cdot\mathbf{\Omega}'\psi(\mathbf{r}, \mathbf{\Omega}', E, t)d\Omega'}{\int_{\hat{\mathbf{n}}\cdot\mathbf{\Omega}'>0} \hat{\mathbf{n}}\cdot\mathbf{\Omega}'d\Omega'} \\ &= \frac{J_+(\mathbf{r}_s, E, t)}{\int_{\hat{\mathbf{n}}\cdot\mathbf{\Omega}'>0} \hat{\mathbf{n}}\cdot\mathbf{\Omega}'d\Omega'}\,, \quad \hat{\mathbf{n}}\cdot\mathbf{\Omega}' < 0\,.\end{aligned} \tag{1.20}$$

Note the conditions on $\mathbf{\Omega}$ and $\mathbf{\Omega}'$. The first corresponds to the left hand side and is limited to $\hat{\mathbf{n}}\cdot\mathbf{\Omega}' < 0$, i.e. incoming directions. Contrarily, $\mathbf{\Omega}'$ is the dummy variable on the right hand side, and is always integrated over the domain where $\hat{\mathbf{n}}\cdot\mathbf{\Omega}' > 0$, i.e. outgoing directions. This is so because we integrate the entire outgoing neutron population (which is proportional to the outgoing partial current) and then redistribute that number uniformly over all incident directions, i.e. isotropically.

White boundary conditions have had use in lattice physics where an isotropic angular distribution is sometimes relatively accurate. In particular, the white boundary condition provides a useful fix for reflective conditions in Wigner-Seitz cells, which convert square pin cells into equivalent cylindrical cells, since cylindrical cells can be treated with 1-d methods. However, while in square cells the reflective conditions work fine, they do not work well in cylindrical geometries (see Figure 1.4), since neutrons entering at certain angles can spend too much time in the moderator before colliding. This consequently leads to overprediction of the moderator flux, an artifact known as the Newmarch effect. As a result, white conditions are used.

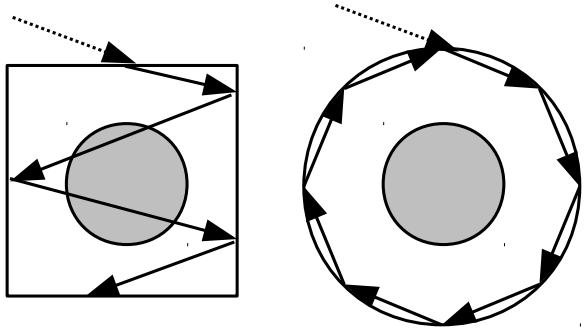

Figure 1.4: Square pin cell and equivalent Wigner-Seitz cell. Same incident direction and location.

Other Transport Equations

We finish this lecture by presenting in brief two other important transport equations.

Photon Transport

Photon transport is fundamental to radiation hydrodynamics (an integral aspect of "bomb" physics) and astrophysics. Photon transport can largely be divided into two classes of problems: *radiative transfer*, which consists of the propogation of soft (low energy) x-rays, and *high energy* photon transport, which can largely be treated as we do neutrons. We briefly describe the former.

The quantity of interest is the intensity, essentially an "energy angular flux", and is defined

$$I_\nu(\mathbf{r}, \mathbf{\Omega}, t) = (h\nu)cn(\mathbf{r}, \mathbf{\Omega}, E, t)\,, \tag{1.21}$$

where $h\nu$ is the photon energy. The "radiative transfer equation" is

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \Omega \cdot \nabla I_\nu = \rho(\varepsilon_\nu - \kappa_\nu I_\nu)\,, \tag{1.22}$$

where ε is a mass emission coefficient (a source term), and κ is a mass attenuation coefficient (a loss term). The radiative transfer equations are nonlinear due to the temperature-dependence of the underlying interaction coefficients (the "cross-sections), particularly the emission term (which is not explicitly represented in Eq. 1.22).

For the case of local thermodynamic equilibrium (LTE), Eq. 1.22 is simplified somewhat. Local thermodynamic equilibrium exists when the quantity $S_\nu = \varepsilon_\nu/\kappa_\nu = B_\nu$, where B_ν is the Planck distribution (i.e. the black body spectrum). In this case, Eq. 1.22 takes the form

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \Omega \cdot \nabla I_\nu = \rho\kappa_\nu(B_\nu - I_\nu)\,. \tag{1.23}$$

Radiative transfer is inherently a frequency-dependent process: the coefficients depend on the frequency, and a medium emits photons of a wide range of frequencies. A frequently used approximation is to neglect this dependence in what is called the *grey approximation*, similar to the one-speed studies in neutron transport we will study in the next several lectures. "Black bodies" are also often used; these are pure absorbers whose emission spectrum is the Planck distribution.

To determine the temperature (dependence on which is implicit in all the quantities of Eq. 1.22), an energy conservation equation is used. As an example, using the grey approximation and assuming temperature- and spatially-independent coefficients, Eq. 1.22 becomes

$$\frac{1}{c}\frac{\partial I}{\partial t} + \Omega \cdot \nabla I(\mathbf{r}, \mathbf{\Omega}, t) = \rho\kappa(acT^4(\mathbf{r}, t) - I)\,, \tag{1.24}$$

where a is the emissivity (as in the Stefan-Boltzmann law) and T is the temperature. The corresponding energy conservation equation is

$$\overbrace{c_v \frac{\partial T}{\partial t}}^{\text{energy rate of change}} = \overbrace{\rho\kappa}^{\text{abs. coef.}} \overbrace{\int_{4\pi} I(\mathbf{r}, \mathbf{\Omega}, t) d\Omega}^{\text{energy flux}} - \overbrace{\rho\kappa a c T^4(\mathbf{r}, t)}^{\text{loss due to emission}} + \overbrace{Q(\mathbf{r}, t)}^{\text{gains from outside}} , \tag{1.25}$$

where c_v is the specific heat and Q represents any external energy source.

Plasma Transport

Another area of interest for nuclear engineers is plasma physics. Let us apply Eq. 1.9 to electrons in a plasma, where we substitute in the Lorentz force for $\mathbf{F}$:

$$\frac{\partial n}{\partial t} + \mathbf{v} \cdot \nabla n + \frac{e}{m}\Big(\mathbf{E} + (\mathbf{v} \times \mathbf{B})\Big) \cdot \nabla_{\mathbf{v}} n = \Big(\frac{\partial n}{\partial t}\Big)_{\text{coll}} + s\,. \tag{1.26}$$

If we neglect sources and collisions, we arrive at the Vlasov equation:

$$\frac{\partial n}{\partial t} + \mathbf{v} \cdot \nabla n + \frac{e}{m}\Big(\mathbf{E} + (\mathbf{v} \times \mathbf{B})\Big) \cdot \nabla_{\mathbf{v}} n = 0\,. \tag{1.27}$$

Augmented with Maxwell's equations, the Vlasov equation gives a rather complete description of collisionless plasmas.

Further Reading

This lecture follows quite closely the treatment of transport theory in Chapter 1 of Duderstadt and Martin [18]. The reader is encouraged to read that chapter (uploaded to Stellar) and others (MIT libraries should have a copy for the eager beaver). Bell and Glasstone [9] give a more traditional derivation, as do Duderstadt and Hamilton [17].

The treatment of boundary conditions is rather straightforward, but the student may wish to consult e.g. Duderstadt and Martin [18] or Lewis and Miller [30]. The discussion of white boundary conditions and the Wigner-Seitz dilemma follows that of Hébert [21], and the Newmarch effect is identified by Stamm'ler and Abbate [43].

The discussion of photon and plasma transport largely follows that of Duderstadt and Martin [18]. The example grey approximation equations are given in a paper by Miller and Lewis [33], and there is a wealth of literature on the subject. For those interested in radiative transfer as it applies to atmospheres, see MIT course 12.815.

Exercises

1. **Variable Transformations**. Prove the relations given in Eq. 1.1.
2. **Duderstadt & Hamilton 4-3**. Suppose that the angular neutron density is given by
$$n(\mathbf{r}, \hat{\mathbf{\Omega}}) = \frac{n_0}{4\pi}(1 - \cos\theta)\,,$$
where θ is the angle between $\hat{\mathbf{\Omega}}$ and the z-axis. If A is the area perpendicular to the z-axis, then what is the number of neutrons passing through the area per second
   (a) per unit solid angle at an angle of 45^{o} with the z-axis
   (b) from the negative z to the positive z direction
   (c) net
   (d) total?
3. **Duderstadt & Hamilton 4-4**. In a spherical thermal reactor of radius R, it is found that the angular flux can be roughly described by
$$\psi(\mathbf{r}, E, \hat{\mathbf{\Omega}}) = \frac{\phi_0}{4\pi} E \exp\left(-\frac{E}{kT}\right) \frac{\sin(\pi r/R)}{r}\,.$$
Compute the total number of neutrons in the core.
4. **Duderstadt & Hamilton 4-5**. Demonstrate that in an isotropic flux, the partial current density in any direction is given by $J_+ = \phi/4$.

Lecture 2 | Neutron Scattering

In the last lecture, we introduced the transport equation and its specialization to neutrons. Over the course of the next several lectures, we'll study ways to treat each of the phase-space variables, the first of which will be the energy variable. However, before we dive into an explicit treatment of the energy variable, we need to review how neutrons change energy, i.e., how they slow down (or speed up). Chiefly, neutrons slow down by elastic scattering, and this lecture provides a brief overview of neutron scattering kinematics.

Scattering in the LAB Frame

Suppose an incoming neutron of energy E strikes a target of mass M at rest, as depicted in Fig. **??**, where the outgoing neutron and target have energies E' and E_A, respectively. Our goal is to define a relationship between E and E' given the scattering angle θ.

To relate E to E', we apply the laws of conservation of energy,

$$E = E' + E_A \,, \tag{2.1}$$

and (linear) momentum,

$$\begin{aligned} p &= p' \cos\theta + p_A \cos\phi \\ 0 &= p' \sin\theta - p_A \sin\phi \,. \end{aligned} \tag{2.2}$$

This is a set of three equations in four unknowns (E', E_A, θ, and ϕ). To simplify, recall the law of cosines, i.e.,

$$A^2 = B^2 + C^2 - 2BC\cos\alpha \,. \tag{2.3}$$

Here, the momentum equations lead to a similar triangle, and, hence, we have

$$p_A^2 = (p')^2 + p^2 - 2p'p\cos\theta \,. \tag{2.4}$$

Classically, the momentum and energy of a particle of mass m are related by

$$p^2/2m = E\,. \tag{2.5}$$

Hence,

$$2m_A E_A = 2mE' + 2mE - 4m\sqrt{EE'}\cos\theta\,, \tag{2.6}$$

or, by using Eq. (2.1), we have

$$m_A(E - E') = m(E' + E - 2\sqrt{EE'}\cos\theta)\,, \tag{2.7}$$

and rearranging the result gives

$$0 = E'(A+1) - \sqrt{E'}(2\sqrt{E}\cos\theta) - E(A-1)\,, \tag{2.8}$$

which is quadratic in $\sqrt{E'}$, and where $A = m_A/m$. Solving for E' and simplifying leads to

$$E' = \frac{E}{(A+1)^2}(\cos\theta + \sqrt{A^2 - \sin^2\theta})^2\,. \tag{2.9}$$

For a "grazing" collision, $\theta \approx 0$, and Eq. (2.9) gives

$$E' \approx \frac{E}{(A+1)^2}(1 + \sqrt{A^2})^2 \approx E\,, \tag{2.10}$$

as expected.

The minimum outgoing energy (for $A > 1$) corresponds to backward scattering, i.e., $\theta = \pi$, so that

$$E'_{\text{min}} = E\left(\frac{A-1}{A+1}\right)^2 = \alpha E\,. \tag{2.11}$$

Although this expression is valid for $A = 1$, i.e., when hydrogen is the target, the minimum occurs for $\theta = \pi/2$.

Scattering in the COM Frame

So we have E' in a one-to-one correspondence with θ, but these values are in the *l*aboratory frame of reference. Some analysis is a bit simpler in the center-of-mass frame, so we really want to relate θ to θ_c and E' to θ_c.

In Fig. **??**, v_c and V_c are the velocities of the neutron and target in the center-of-mass system.

The total system in each coordinate system must exhibit conserved momentum, i.e.,

$$(m + m_A)V_c m = mv + m_A V_a\,. \tag{2.12}$$

In the LAB frame, the target is stationary so that

$$V_c m = \frac{mv}{m + m_A} = \frac{v}{1 + A} \,. \tag{2.13}$$

To translate particle velocities from LAB to COM values requires we subtract V_{cm} from the LAB values, i.e.,

$$v_c = v - V_{cm} = \frac{A}{1 + A} v \tag{2.14}$$

and

$$V_c = -V_{cm} = -\frac{1}{1 + A} v \tag{2.15}$$

We now use these values in conservation equations to find the neutron energy before and after the collision, i.e.,

$$E_c = \frac{1}{2} m v_c^2 \frac{1 + A}{A} \tag{2.16}$$

and

$$E_c' = \frac{1}{2} m (v_c')^2 \frac{1 + A}{A} \,. \tag{2.17}$$

It follows that

$$v_c = v_c' = \frac{A}{1 + A} v \tag{2.18}$$

and

$$V_c = \frac{1}{1 + A} v \,. \tag{2.19}$$

Now, we look to develop relationships between (1) θ_c and θ and (2) E' and θ_c. Note that

$$v_c' \sin\theta_c = v' \sin\theta \tag{2.20}$$

and

$$v_c' \cos\theta_c + v_{cm} = v' \cos\theta \tag{2.21}$$

from which it follows that

$$\tan\theta = \frac{\sin\theta_c}{\cos\theta_c + \gamma} \,, \tag{2.22}$$

where

$$\gamma = 1/A \,. \tag{2.23}$$

Then, one can show that

$$\cos\theta = \frac{A \cos\theta_c + 1}{\sqrt{A^2 + 2A \cos\theta_c + 1}} \,, \tag{2.24}$$

proof of which is left as an exercise to the student.

Then, by squaring and adding Eqs. (2.20) and 2.21, one can show that

$$\left(\frac{v'}{v}\right)^2 = \frac{E'}{E} = \frac{A^2 + 2A\cos\theta_c + 1}{(1+A)^2} \tag{2.25}$$

or

$$\frac{E'}{E} = \frac{1}{2}[1 + \alpha + (1-\alpha)\cos\theta_c]\,. \tag{2.26}$$

Scattering Cross Sections

In the transport equation for neutrons, we encountered the macroscopic *double-differential scattering cross section* $\Sigma_s(E \to E', \Omega \to \Omega')$, where the spatial variable is suppressed. Let's consider the microscopic, double-differential scattering cross section

$$\sigma_s(E \to E', \Omega \to \Omega')\,. \tag{2.27}$$

By integrating over either Ω' or E', we have the (single) differential scattering cross sections

$$\sigma_s(E \to E') = \int_{4\pi} \sigma_s(E \to E', \Omega \to \Omega')d\Omega' \tag{2.28}$$

and

$$\sigma_s(E, \Omega \to \Omega') = \int_0^\infty \sigma_s(E \to E', \Omega \to \Omega')dE'\,, \tag{2.29}$$

or the (total) scattering cross section

$$\sigma_s(E) = \int_0^\infty \int_{4\pi} \sigma_s(E \to E', \Omega \to \Omega')d\Omega' dE'\,. \tag{2.30}$$

We can relate the differential scattering cross section $\sigma_s(E \to E')$ to the total scattering cross section $\sigma_s(E)$ with the form

$$\sigma_s(E \to E') = \sigma_s(E)P(E \to E')\,, \tag{2.31}$$

where $P(E \to E')$ is, loosely, the probability that a neutron of energy E scatters to energy E'. For elastic scattering, we can define $P(E \to E')$ explicitly, but for inelastic scattering, $P(E \to E')$ must be defined through experimental measurements.

Consider again the differential scattering cross section $\sigma_s(\Omega \to \Omega')$. If we assume the material in question is isotropic, i.e., that the scattering of neutrons does

not depend on the orientation of the neutron's direction of travel with respect to the material, then the scattering cross section depends only on the scattering angle, or

$$\begin{aligned}\sigma_s(\Omega \to \Omega') &= \sigma_s(E \to E', \Omega \cdot \Omega') \\ &= \sigma_s(E \to E', \cos\theta)/2\pi \\ &= \sigma_s(E \to E', \mu)/2\pi\,,\end{aligned} \tag{2.32}$$

where

$$\mu = \cos\theta\,, \tag{2.33}$$

and, hence,

$$d\mu = \sin\theta d\theta\,. \tag{2.34}$$

The factor of 2π comes from the fact that we must have

$$\int_{4\pi} \sigma_s(\Omega \to \Omega') d\Omega' = \int_0^{2\pi} \int_{-1}^{1} \frac{\sigma_s(E \to E', \mu)}{2\pi} d\mu d\phi\,. \tag{2.35}$$

We can use these expressions to define the probability $P(\Omega \cdot \Omega')$, i.e., the probability that a neutron scatters through the angle θ, or

$$P(\theta) 2\pi \sin\theta d\theta = \frac{\sigma_s(\theta)}{\sigma_s} 2\pi \sin\theta d\theta\,. \tag{2.36}$$

Because there is a one-to-one relationship between the outgoing energy E' and the scattering angle θ (either in the LAB or in the COM frames), we can relate the transfer probability $P(E \to E')$ to one for the scattering angle, i.e.,

$$P(E \to E') dE' = -\frac{\sigma_s(\theta)}{\sigma_s} 2\pi \sin\theta d\theta\,. \tag{2.37}$$

In the COM system, elastic scattering is isotropic for all but the highest energies, i.e.,

$$\sigma(\theta_c) \approx \text{constant} \equiv \frac{\sigma_s}{4\pi}\,. \tag{2.38}$$

Then, from Eq. (2.26), we have

$$\left|\frac{dE'}{d\theta_c}\right| = \frac{E(1-\alpha)\sin\theta_c}{2}\,, \tag{2.39}$$

and, hence,

$$P(E \to E') = \begin{cases} \frac{1}{E(1-\alpha)} & \alpha E \le E' \le E \\ 0 & \text{otherwise}\,. \end{cases} \tag{2.40}$$

Further Reading

Most of this development can be found in similar form in any good reactor physics text, e.g., Duderstadt and Hamilton[17].

Exercises

1. **Outgoing Energy in the LAB**. Prove Eq. (2.9)
2. **Relating LAB and COM**. Prove Eq. (2.24)
3. **Outgoing Energy in the COM**. Prove Eq. (2.26)
4. **The Average Outgoing Energy**. Prove that the average outgoing energy (in LAB) after a collision from $E \to E'$ is
$$\bar{E}' = \frac{1+\alpha}{2} E \,.$$
5. **Average Cosine of the Scattering Angle**. Prove that, for the case of isotropic scattering in the COM frame, the average cosine of the LAB scattering angle is given by
$$\bar{\mu} = \frac{2}{3A} \,.$$
Hence, in the LAB frame, neutron scattering tends to be anisotropic with a forward bias.
6. **Average Logarithmic Energy Loss**. A useful quantity is the so-called "lethargy," defined as
$$u = \ln E_0/E \tag{2.41}$$
for some reference energy E_0 (usually a large value like 10 MeV). Show that the average logarithmic energy loss is given by
$$\xi = \langle \ln(E/E') \rangle = 1 + \frac{\alpha}{1-\alpha} \ln \alpha \,. \tag{2.42}$$
7. **Average Logarithmic Energy Loss for Large** A. Show that $\xi \approx 2/(A + 2/3)$ for large A.
8. **Anisotropic Scattering**. Suppose that
$$\sigma(\theta_c) = \frac{\sigma_s}{4\pi}(1 + a\cos\theta_c) \,, \tag{2.43}$$
where θ_c is the scattering angle in the COM frame.
   (a) Determine $P(E \to E')$ and sketch it for $\alpha E \le E' \le E$ for the case of $a > 0$, $a = 0$, and $a < 0$.
   (b) Determine $P(\theta)$ (i.e., for the LAB), and plot over $0 \le \pi$.

Lecture 3 | Neutron Slowing Down

In the last lecture, we reviewed neutron scattering kinematics in order to define the energy transfer probability $P(E \rightarrow E')$ and the energy-dependence of differential scattering cross sections. In this lecture, we'll use these cross sections to determine the energy dependence of the neutron flux as neutrons "slow down" by consecutive collisions. We will limit ourselves to homogeneous media (for now) and use both a direct numerical approach and simpler, classical schemes.

Slowing Down in a Homogenous System

Suppose that the medium of interest is infinite in spatial extent, i.e., it has no boundaries, and it is homogeneous in makeup. Then, all spatial dependence vanishes, leading to

$$\Sigma_t(E)\psi(E,\hat{\mathbf{\Omega}}) = Q(E,\hat{\mathbf{\Omega}})\,, \tag{3.1}$$

and integration over angle yields

$$\begin{aligned}
&\Sigma_t(E)\phi(E) \\
&= \int_{4\pi} d\Omega \int_{4\pi} d\Omega' \int_0^\infty \Sigma_s(E',\hat{\mathbf{\Omega}}' \rightarrow E,\hat{\mathbf{\Omega}})\psi(E',\hat{\mathbf{\Omega}}')dE' \\
&+ \frac{\chi(E)}{k}\int_0^\infty \nu(E')\Sigma_f(E')\phi(E')dE' \\
&+ S_{\text{ext}}(E)\,.
\end{aligned} \tag{3.2}$$

If scattering is assumed to be isotropic in the LAB system, then $\Sigma_s(E',\hat{\mathbf{\Omega}}' \rightarrow E,\hat{\mathbf{\Omega}}) = \Sigma_s(E' \rightarrow E)/4\pi$ and

$$\begin{aligned}
\Sigma_t(E)\phi(E) = &\int_0^\infty \Sigma_s(E' \rightarrow E)\phi(E')dE' \\
&+ \frac{\chi(E)}{k}\int_0^\infty \nu(E')\Sigma_f(E')\phi(E')dE' \\
&+ S_{\text{ext}}(E)\,,
\end{aligned} \tag{3.3}$$

which is the infinite medium or spectrum equation. Here, $\phi(E)$ is referred to as the *spectrum*.

Between about 1 eV and 100 keV (the *resonance region*, which depends on nuclide), fission and inelastic scattering can be neglected, and the dominant interaction is elastic scattering. Hence, the spectrum equation (without an external source) simplifies to the the *slowing-down equation*

$$\Sigma_t(E)\phi(E) = \sum_k^{N_{\text{nucl}}} \frac{1}{1-\alpha_k} \int_E^{E/\alpha_k} N_k \sigma_{s,k}(E')\phi(E')\frac{dE'}{E'}\,, \tag{3.4}$$

where $\Sigma_t(E)$ is the total macroscopic cross section, N_k is the number density of nuclide k, $\sigma_{s,k}(E)$ is the microscopic elastic scattering cross section of nuclide k, $\alpha_k = (A-1)^2/(A+1)^2$ for a nuclide k of mass A (in amu), and N_{nucl} is the number of nuclides.

For the particular case of just one, purely-scattering nuclide, Eq. (3.4) simplifies to

$$\Sigma_s(E)\phi(E) = \frac{1}{1-\alpha} \int_E^{E/\alpha} \Sigma_s(E')\phi(E')\frac{dE'}{E'}\,, \tag{3.5}$$

and, by substitution, one can verify the corresponding solution is

$$\phi(E) = \frac{C}{E\Sigma_s(E)}\,, \tag{3.6}$$

where C is a constant defined by normalization. Because $\Sigma_s(E)$ is usually almost constant, the *slowing down spectrum* typically goes as $1/E$.

Direct Solution of the SDE

Let $E \to E_i$, where ΔE is small enough that $\Sigma(E)$ is well-approximated by interpolation between E_i and E_{i+1}. Then

$$\Sigma_t(E_i)\phi(E_i) = \sum_{j=i}^{N_E} \Sigma_s(E_j \to E_i)\phi(E_j)\Delta E_j + S(E_i)\,, \tag{3.7}$$

where an external source $S(E)$ has been included for generality, N_E is the number of discrete energy points used, and $\Delta E_j = E_{j+1} - E_j$.

For "continuous" treatment, the matrix implied by $\Sigma_s(E_j \to E_i)$ is far too large. Instead, we can evaluate its values on-the-fly by defining

$$\Sigma_{s,j\to i} \approx \begin{cases} \frac{\Sigma_s(E_j)}{(1-\alpha)E_j} & \alpha E_j \le E_i \le E_j \\ 0 & \text{otherwise}\,. \end{cases} \tag{3.8}$$

Remember, this approximation is valid only for *isotropic* (i.e., *s-wave*) scattering in the COM system. Typically, this is the case for all but the highest energies.

Substitution of Eq. (3.8) into Eq. (3.7) leads to

$$\Sigma_t(E_i)\phi(E_i) = \sum_{j=i}^{N_E}\sum_{k=1}^{N_{\text{nucl}}} \frac{\sigma_s(E_j)\phi(E_j)\Delta E_j}{(1-\alpha_k)E_j} + S(E_i)\,. \tag{3.9}$$

Note that the sum over energy includes *all* energies, but, in practice, the sum for some nuclides would be limited to a smaller number of energies for all but hydrogen.

This equation uses a one-sided *Riemann sum* for the integral. One could also use a *midpoint* or *trapezoid* rule. Knott and Yamamoto use the same form as Eq. 3.9 but substitute an average value $\bar{E}_j = \sqrt{E_j E_{j+1}}$ for E_j in the denominator.

To solve Eq. 3.9, rearrange to get

$$\phi_i = \frac{\sum\limits_{j=i+1}^{N_E}\sum\limits_{k=1}^{N_{\text{nucl}}} \frac{\sigma_s(E_j)\phi(E_j)\Delta E_j}{(1-\alpha_k)E_j} + S_j}{\Sigma_{t,i} - \Sigma_{s,i\to i}}\,, \tag{3.10}$$

where

$$\phi_N = \frac{S_N}{\Sigma_{t,N} - \Sigma_{s,N\to N}}\,. \tag{3.11}$$

Alternatively, neglect any external source and set ϕ_N directly. Note that the indexing used here is reversed from the standard use of E_1 indicating the largest energy. Rather, E_N is the largest energy. As basic implementation of the method described is provided in Listing 3.1. The implementation is limited to a single species with constant cross sections.

Listing 3.1: Direct Solution of the SDE

```
import numpy as np
def scatter_probability(E_i, E_j, alpha) :
    p = (1.0/E_j/(1.0-alpha)) * 1.0*((E_i >= alpha*E_j))
    return p
def compute_spectrum(E, Sigma_t, Sigma_s, alpha) :
    N_E = len(E)
    phi = np.zeros(N_E)
    phi[N_E-1] = 1.0
    for i in range(N_E-2, -1, -1) :
        Q_i = 0.0
        for j in range(N_E-1, i, -1) :
            dE = E[j] - E[j-1]
            E_bar = np.sqrt(E[j]*E[j-1])
            Q_i += phi[j] * Sigma_s * scatter_probability(E[i], E[j], alpha) * dE
        phi[i] = Q_i / Sigma_t
        print i, phi[i]
    return phi
E = np.logspace(-1, 2, 1000)
phi = compute_spectrum(E, 1.0, 1.0, 0.0)
```

Although the approach outlined (and implemented) is technically sound, it is not very efficient. The reader should consult Knott and Yamamoto (page 1031) for implementation ideas, especially for simplifying the construction of the in-scatter source term.

Classical Resonance Approximations

Although the numerical solution of the the slowing down equation is straightforward, in principle, it remains relatively expensive to solve for homogeneous media and, when applied to heterogeneous systems, analyses have historically been limited to small problems. As an alternative, several approximations can be made that lead to a direct, analytical solution.

Consider again the slowing-down equation,

$$\begin{aligned}\left(N_r\sigma_{t,r}(E)+\sum_{k\neq r}N_k\sigma_{s,k}\right)\phi(E) = \\ \frac{1}{1-\alpha_r}\int_E^{E/\alpha_r}N_k\sigma_{s,r}(E')\phi(E')\frac{dE'}{E'} \\ +\sum_{k\neq r}\frac{1}{1-\alpha_k}\int_E^{E/\alpha_k}N_k\sigma_{s,k}\phi(E')\frac{dE'}{E'}\,,\end{aligned} \tag{3.12}$$

modified such that terms related to a single resonator (identified by the r subscript) are separate from all other nuclides k, and where the non-resonator cross sections are assumed to be independent of energy and limited to elastic scattering, the first major approximations we shall make on our way to the narrow resonance (NR) and wide resonance (WR) approximations.

Next, we shall assume the "practical" width of a resonance is always much smaller than the energy lost by a neutron scattering with all non-resonant nuclides. Thus, for an integral of the form

$$\frac{1}{1-\alpha_k}\int_E^{E/\alpha_k}N_k\sigma_{s,k}\phi(E')\frac{dE'}{E'}\,, \tag{3.13}$$

most of the integration domain is far away from the resonance and, hence, we assume that the flux *within the integral* takes its asymptotic form, i.e., the form found for a pure scatterer (with a constant cross section), $\phi(E) \propto 1/E$. Therefore, the non-resonant scattering integral simplifies to

$$\frac{N_k\sigma_{s,k}}{1-\alpha_k}\int_E^{E/\alpha_k}\frac{1}{E'}\frac{dE'}{E'} = \frac{N_k\sigma_{s,k}}{E}\,. \tag{3.14}$$

Narrow Resonance Approximation

The NR approximation further assumes that energy loss of a neutron scattering with the resonator is also much smaller than the resonance width, which means the same approximation can be made for the resonance scattering integral as was made for the non-resonant scattering angle. With the additional assumption that $\sigma_{s,r}(E) \approx \sigma_{s,r}$, the slowing down equation simplies to

$$\left(N_r\sigma_{t,r}(E) + \sum_{k\neq r} N_k\sigma_{s,k}\right)\phi(E) = \frac{\Sigma_{s,r}}{E} + \sum_k \frac{\Sigma_{s,k}}{E}\,, \tag{3.15}$$

and, hence, the spectrum in the narrow resonance approximation is

$$\phi_{\text{NR}}(E) = \frac{\Sigma_{s,r} + \sum\limits_{k\neq r}\Sigma_{s,k}}{\Sigma_t(E)E}\,, \tag{3.16}$$

with arbitrary normalization.

Wide Resonance Approximation

Contrarily, the WR approximation assumes that the energy lost by a neutron scattering off the resonant nuclide is much *smaller* than the resonance width. The smallest such energy loss occurs in the limit $\alpha_r \to 1^*$ for which case the corresponding scattering integral simplifies to

$$\begin{aligned}\lim_{\alpha_r\to 1}&\left[\frac{1}{1-\alpha_r}\int_E^{E/\alpha_r} N_k\sigma_{s,r}(E')\phi(E')\frac{dE'}{E'}\right]\\ &\approx N_k\sigma_{s,r}(E)\phi(E)\lim_{\alpha_r\to 1}\int_E^{E/\alpha_r}\frac{dE'}{(1-\alpha_r)E'}\\ &= N_k\sigma_{s,r}(E)\phi(E)\,.\end{aligned} \tag{3.17}$$

It follows that

$$\phi_{\text{WR}} = \frac{\sum\limits_{k\neq r}\Sigma_{s,k}}{[\Sigma_t(E) - \Sigma_{s,r}(E)]E}\,. \tag{3.18}$$

Further Reading

To be continued.

*i.e., when the resonator has an infinite mass, which explains why WR sometimes is called the narrow resonance, infinite mass (NRIM) approximation)

Exercises

1. **Slowing Down in Purely Scattering Media**. Adapt the slowing-down code to treat a single medium with arbitrary mass number A. Using a point source at $E = 10$ keV, compute the spectrum and plot for $1 \text{ eV} \leq E \leq 10$ keV. For $A = 238$, zoom in and plot the rather strange ("Placzek") oscillations occuring near the source energy. Can you explain this behavior?
2. **Slowing Down in Arbitary Mixture**. Adapt the slowing-down code to treat an arbitary homogeneous mixture that uses real cross-section data from ENDF. Then
   (a) Find total and elastic scattering cross-section data for H-1 and U-238 (at room temperature) and develop a way to define that data on the same energy grid (sometimes called a "unionized" energy grid).
   (b) Use the slowing-down code to determine the spectrum for a 1000-to-1 mixture of H-to-U over the range $E \in [1, 100 \text{ eV}]$ with a point source at the upper limit. Use at least 2000 energy points evenly spaced on a log scale.
   (c) Plot the spectrum. Repeat for 100-to-1 and 10-to-1 ratios of H to U.
3. **The Narrow and Wide Resonance Approximations**. For the same three cases studied in Problem 2, determine the spectrum for both the narrow and wide resonance approximations. For the 10-to-1 case, plot the numerical, NR, and WR spectra normalized so that $\phi(1 \text{ eV})$ is 1.

Lecture 4 | Multigroup Method

In this lecture, the energy dependence of the particle flux is treated using the multigroup method. With knowledge of the energy spectrum (as computed, e.g., using the approach of Lecture 3), cross sections can be averaged over energy intervals in which various reaction rates are preserved. Not just any averaging will do, however; rather, we need to employ flux-weighted averaging.

To be continued.

Further Reading

To be continued.

Exercises

1. **Generation of Multigroup Constants**. For the same three cases studied in the last lecture, compute multigroup, microscopic, capture cross sections for U-238 over the ranges $E \in [1, 12]$, $E \in [12, 28]$, and $E \in [28, 50]$, all in eV. Note that these groups each contain one of the first three resonances of U-238. Use the numerical, NR, and WR spectra. Taking the numerical solution to be the reference, compute the relative error of the NR and WR results for each of the three groups and hydrogen concentrations.
2. **Self-Shielding Factors**. For the same three energy groups, compute the (Bondarenko) self-shielding factors using the three different spectra.

Lecture 5 | Analytical Solutions

In this lecture, we analyze the neutron transport equation analytically for several simple cases. In particular, we investigate neutron streaming in a vacuum and in a purely absorbing slab. The next lecture offers further analytical and semi-analytical treatments using the integral form of the transport equation. These two lectures ultimately show the difficulty with which realistic problems can be addressed by "pen and paper" and serve to motivate our later discussions of deterministic numerical methods.

One-Speed Transport

Before we consider solutions to the transport equation, we first eliminate the energy dependence. The reason for this is simple: *the energy is simply too hard to deal with directly*. The dependence of the various cross-sections on the energy is erratic, and, as we have seen in previous lectures, there are isotopes whose dependence on energy in certain energy ranges cannot even be resolved!

We can eliminate E in two ways. First, we can assume that ψ and the cross-sections are constant in energy within an energy range $E_g < E < E_{g-1}$; this is the multi-group method, which has been the workhorse of deterministic transport methods for decades*. A second, somewhat superficial approach is to multiply the energy-dependent transport equation by $\delta(E - E_0)$. Since $f(x)\delta(x - x_0) = f(x_0)$, we have for E_0 (or a groups g) the time- and energy-independent or *one-speed transport equation*:

$$\hat{\Omega} \cdot \nabla\psi(\mathbf{r}, \mathbf{\Omega}) + \Sigma_t\psi(\mathbf{r}, \mathbf{\Omega}) = \int_{4\pi} d\Omega' \Sigma_s(\mathbf{r}, \mathbf{\Omega} \cdot \mathbf{\Omega}')\psi(\mathbf{r}, \mathbf{\Omega}') + s(\mathbf{r}, \mathbf{\Omega}) \,. \qquad (5.1)$$

*A new methodology being developed here at MIT is a generalization of the multigroup method where instead of flat fluxes within groups (a "zeroth" order representation), the fluxes can have higher order dependences (linear, parabolic, etc.) using discrete Legendre polynomial expansions.

Streaming in Vacuum

Perhaps the easiest class of problems to consider are those whose medium is vacuum. In this case, there are no particle interactions, and all we need to do is follow particles along trajectories from sources. These trajectories are called *characteristics*, and in general, the *method of characteristics* is the mathematical technique we can use to find ψ. Note, a more general use of the *method of characteristics* applies to arbitrary media and is fast becoming the standard transport method in lattice physics.

The Streaming Term

Consider the streaming of neutrons in a sourceless vacuum:

$$\hat{\Omega} \cdot \nabla \psi(\mathbf{r}, \mathbf{\Omega}) = 0 \,. \tag{5.2}$$

What form does the streaming term $\hat{\Omega} \cdot \nabla \psi$ have? It depends crucially on the underlying coordinate systems.

It helps to note that $\hat{\mathbf{\Omega}} \cdot \nabla \psi$ is just the spatial rate of change of ψ along the direction of travel, i.e. along the characteristic. Suppose we have a particle originally at a location $\mathbf{r}_0$ going in a direction $\mathbf{\Omega}$. Then upon traveling a distance s, the location $\mathbf{r} = \mathbf{r}_0 + s\mathbf{\Omega}$. Accordingly, the spatial rate of change of ψ can be written

$$\frac{d}{ds} \psi(\mathbf{r}_0 + s\mathbf{\Omega}, \mathbf{\Omega}) = 0 \,. \tag{5.3}$$

We can also show more explicitly that $\frac{d\psi}{ds} = \hat{\Omega} \cdot \nabla \psi$. In general, the direction vector $\mathbf{\Omega}$ has the basic form

$$\mathbf{\Omega} = \mu \hat{\mathbf{e}}_\mu + \eta \hat{\mathbf{e}}_\eta + \xi \hat{\mathbf{e}}_\xi \,, \tag{5.4}$$

where μ, η, and ξ are directional cosines and the $\hat{\mathbf{e}}$'s are corresponding coordinate vectors. In general, $\mathbf{\Omega}$ depends on three spatial coordinates p_1, p_2, and p_3 and two angular coordinates, usually parameterized as the cosine of the polar angle $\chi = \cos(\theta)$[†] and the azimuthal angle ϕ. Hence,

$$\frac{d}{ds} = \frac{dp_1}{ds}\frac{\partial}{\partial p_1} + \frac{dp_2}{ds}\frac{\partial}{\partial p_2} + \frac{dp_3}{ds}\frac{\partial}{\partial p_3} + \frac{d\mu}{ds}\frac{\partial}{\partial \mu} + \frac{d\phi}{ds}\frac{\partial}{\partial \phi} \,. \tag{5.5}$$

The various derivatives in Eq. 5.5 may or may not vanish, depending on the geometry. Consider Cartesian geometry where $p_1 = x$ and so on. The Cartesian spatial and

[†]We use χ to represent the polar angle cosine in general rather than μ, since μ here will be the directional cosine with respect to the x axis.

angular system was given in Figure 1.1, where the polar angle was defined with respect to the z axis and the azimuth with respect to the x axis. Any incremental movement ds along the direction $\mathbf{\Omega}$ can be seen to change neither θ (nor its cosine ξ) nor ϕ, since the angular coordinate system is invariant as the particle moves. All this means is that $\mathbf{\Omega}$ at $\mathbf{r}_0$ is the same as the $\mathbf{\Omega}$ at $\mathbf{r}_0 + ds\mathbf{\Omega}$. Hence, $d\xi/ds = d\phi/ds = 0$ and

$$\frac{d}{ds} = \frac{dx}{ds}\frac{\partial}{\partial x} + \frac{dy}{ds}\frac{\partial}{\partial y} + \frac{dz}{ds}\frac{\partial}{\partial z}\,, \tag{5.6}$$

but dx/ds is just the directional cosine with respect to the x axis, μ, and likewise $dx/ds = \eta$ and $dz/ds = \xi$ so that

$$\frac{d}{ds} = \mu\frac{\partial}{\partial x} + \eta\frac{\partial}{\partial y} + \xi\frac{\partial}{\partial z} = \mathbf{\Omega}\cdot\nabla\,. \tag{5.7}$$

For other geometries, the streaming term is not as simple, since the angular coordinate system does depend on the position $\mathbf{r}$. The spherical spatial and angular system is given if Figure 5.1. The three spatial coordinates are r, θ_r and ϕ_r. The angular coordinate system is such that the polar angle is defined with respect to $\mathbf{r}$. The azimuth and secondary coordinates are defined somewhat arbitrarily.

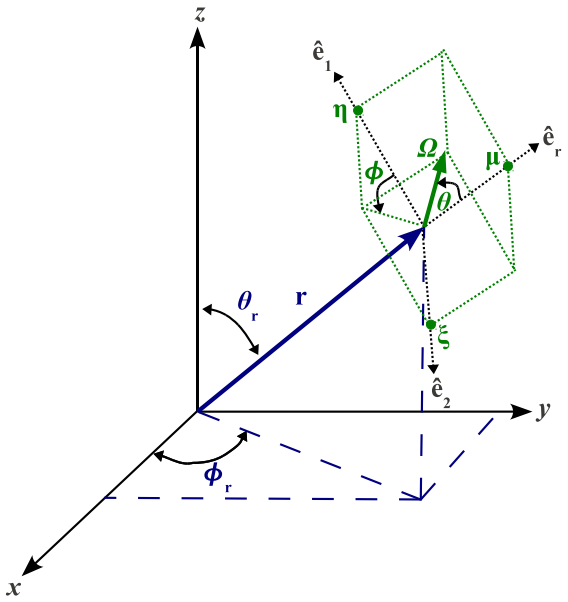

Figure 5.1: Spherical phase space.

The streaming operator in spherical coordinates is defined generally as

$$\frac{d}{ds} = \frac{dr}{ds}\frac{\partial}{\partial r} + \frac{d\theta_r}{ds}\frac{\partial}{\partial \theta_r} + \frac{d\phi_r}{ds}\frac{\partial}{\partial \phi_r} + \frac{d\mu}{ds}\frac{\partial}{\partial \mu} + \frac{d\phi}{ds}\frac{\partial}{\partial \phi}\,. \tag{5.8}$$

As a simple example, consider the case of 1-d transport in spherical coordinates, for which we assume the flux is independent of the spatial coordinates θ_r and ϕ_r, which eliminates the derivatives with respect to θ_r and ϕ_r. Moreover, if we look at the angular coordinates, we see that if r is the only spatial variable, then there should be dependence only on μ. A dependence on the azimuthal angle would require a non-uniform particle distribution in the other spatial directions. Hence, the derivative with respect to ϕ also vanishes, leaving

$$\frac{d}{ds} = \frac{dr}{ds}\frac{\partial}{\partial r} + \frac{d\mu}{ds}\frac{\partial}{\partial \mu}\,. \tag{5.9}$$

From Figure 5.1, we can immediately see that

$$\frac{dr}{ds} = \mu\,. \tag{5.10}$$

Less obvious is $d\mu/ds$. First note that $rd\theta = -ds\sin\theta$, which we can see directly from Figure 5.2. The negative sign arises because a positive ds leads to a decrease θ. Then, noting that $\mu = \cos\theta$ so that $d\mu = -\sin\theta d\theta$, we can show

$$\frac{d\mu}{ds} = \frac{1-\mu^2}{r}\,, \tag{5.11}$$

and

$$\frac{d\psi}{ds} = \mathbf{\Omega}\cdot\nabla_{1\mathrm{d}}\psi = \mu\frac{\partial\psi}{\partial r} + \frac{1-\mu^2}{r}\frac{\partial\psi}{\partial\mu}\,. \tag{5.12}$$

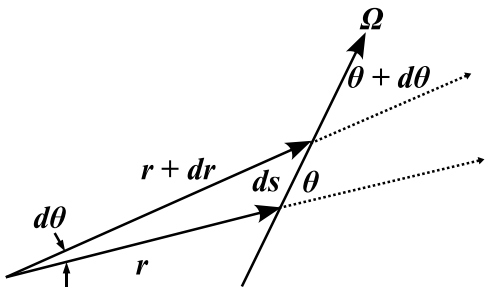

Figure 5.2: Change in r and θ as a particle streams.

Example 1: Meandering from a Plane Source in a Vacuum Slab

As a first example, consider the case of neutrons streaming in a 1-d slab for $x > 0$ where the source is an isotropic planar source at $x = 0$ which we will use as a boundary condition rather than a source term. The Cartesian streaming term above simplifies in 1-d to

$$\mu \frac{\partial \psi}{\partial x} = 0 \,, \tag{5.13}$$

subject to

$$\psi(0, \mu) = \frac{S_0}{2} \,. \tag{5.14}$$

In 1-d slab geometry, we essentially integrate out the 2π associated with the azimuthal angle. Hence, isotropic sources of strength S become $S/2$ instead of $S/4\pi$. The units of ψ are slightly different, and one must divide by 2π [rad] in order to obtain units appropriate in 3-d. Integrating the equations shows that

$$\psi(x, \mu) = c \,, \quad \mu > 0 \,, \tag{5.15}$$

for some constant c. At $x = 0$, we must have $\psi(0, \mu) = S/2$, and so for all $x > 0$, $\psi(x, \mu) = S/2$. The same holds for negative x and μ.

Example 2: Playing in Vacuum Outside a Spherical Shell Source

We now consider a neutrons streaming in a vacuum due to an isotropic spherical shell source of radius r_0. We focus only on $r > r_0$. The transport equation can be written

$$\mu \frac{\partial \psi}{\partial r} + \frac{1 - \mu^2}{r} \frac{\partial \psi}{\partial \mu} = \frac{S_0 \delta(r - r_0)}{4\pi r_0} \,. \tag{5.16}$$

From Eqs. 5.10 and 5.11, we have

$$ds = \frac{d\mu}{(1 - \mu^2)/r} = \frac{dr}{\mu} \,, \tag{5.17}$$

or

$$\frac{dr}{r} = \frac{\mu d\mu}{1 - \mu^2} \,. \tag{5.18}$$

We integrate from initial coordinates r_s and μ_s, as shown in Figure 5.3, and rearrange to obtain

$$\mu = \sqrt{1 - (1 - \mu_s^2) \left(\frac{r_s}{r} \right)^2} \,. \tag{5.19}$$

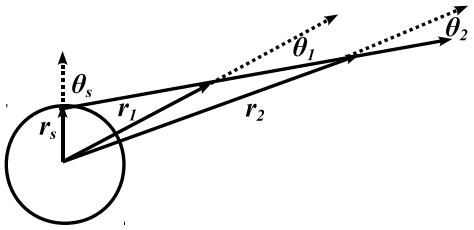

Figure 5.3: Streaming of particles from spherical shell source.

Eq. 5.19 shows explicitly how μ changes as a function of r. This phenomenon is often called angular redistribution and leads to collimation of the flux away from the source. Since $d\psi/ds$ must equal the source, we can write

$$ds = \frac{d\mu}{(1-\mu^2)/r} = \frac{d\psi}{\frac{S_0\delta(r-r_0)}{4\pi r_0}}\,, \tag{5.20}$$

or

$$\frac{d\psi}{dr} = \frac{S_0\delta(r-r_0)}{4\pi r_0}\,. \tag{5.21}$$

Then

$$\int_{r_s,\mu_s}^{r,\mu} \frac{d\psi}{dr} = \int_{r_s}^{r} dr \frac{S_0\delta(r-r_0)}{4\pi r_0}\,, \tag{5.22}$$

or

$$\psi(r,\mu) - \psi(r_s,\mu_s) = \frac{S_0}{8\pi r_s^2}\frac{1}{\sqrt{1-(1-\mu_s^2)}}\,. \tag{5.23}$$

We take the boundary flux to be $\psi(r_s,\mu_s) = 0$, which implies that μ_s is limited to $\mu_s \leq 0 \leq \pi/2$. If μ_s spanned through π, then particles could be born into the sphere and stream out of another location, a complication we care to avoid in this example. From Eq. 5.19, we have $(1-\mu_s^2) = (r/r_s)^2(1-\mu^2)$. We note that for $\mu_s = 0$, $\mu = \sqrt{1-(r_s/r)^2}$ and when $\mu_s = 1$, $\mu = 1$. Finally, we have

$$\psi(r,\mu) = \frac{S_0}{8\pi}\frac{1}{\sqrt{r_s^2 - r^2(1-\mu^2)}}\,, \quad \sqrt{1-\left(\frac{r_s}{r}\right)^2} \leq \mu \leq 1\,. \tag{5.24}$$

To illustrate the behavior of ψ, we have taken $r_s = 1$ and $S_0 = 8\pi$. Figure 5.4 shows ψ as a function of radius for several μ values. Of course, we see that as neutrons move farther from the source, the flux at larger θ values (smaller μ values) diminishes, as we expect due to angular redistribution. Figure 5.5 shows ψ as a function of μ for several values of r. We see effects of the same phenomenon, in that the angular distribution becomes more collimated about $\mu = 1$ for larger r.

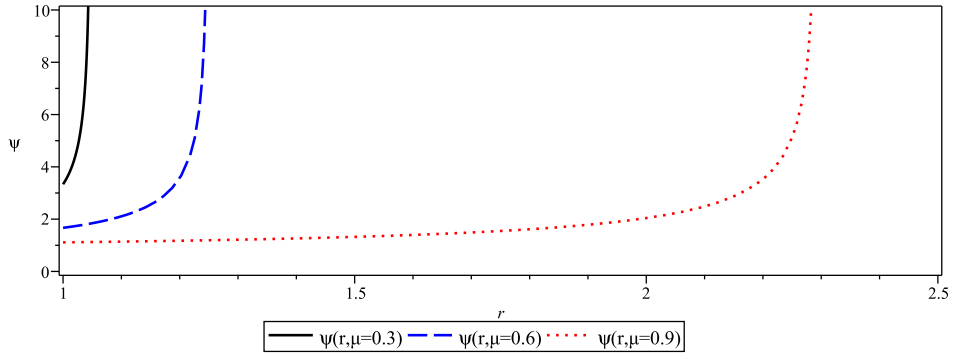

Figure 5.4: Angular flux as a function of radius for certain μ values.

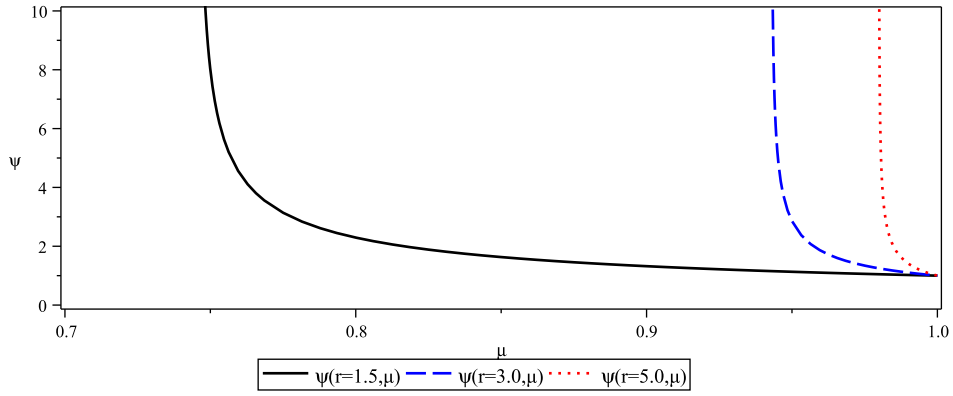

Figure 5.5: Angular flux as a function of μ at certain r values.

Example 3: A Purely Absorbing Slab

As a final example, we apply what we've learned in vacuum to transport in a purely absorbing slab. In this case, the 1-d transport equation is

$$\mu \frac{\partial \psi}{\partial x} + \Sigma_a(x)\psi(x,\mu) = S(x,\mu)\,. \tag{5.25}$$

As an example, we study a uniform slab of length L, subject to vacuum boundaries, and with a uniform, isotropic source of volumetric strength[‡] S_0. Our equation sim-

[‡]By volumetric strength, we mean the units are neutrons per unit volume per second.

plifies to

$$\mu \frac{\partial \psi}{\partial x} + \Sigma_a \psi(x, \mu) = \frac{S_0}{2} \,. \tag{5.26}$$

To solve the problem, we decompose ψ into ψ_+ for $\mu > 0$ and ψ_- for $\mu < 0$. In this way, we can start at one end of the slab and work our way across, essentially as would a neutron.

For $\mu > 0$, we divide Eq. 5.26 through by μ and compute the integrating factor

$$if = \exp \int_0^x dx \Sigma_a \mu = \exp \Sigma_a x/\mu \,. \tag{5.27}$$

Then we have

$$\frac{d}{dx}\left(\psi_+ e^{\Sigma_a x}\right) = \frac{S_0}{2\mu} e^{\Sigma_a x/\mu} \,, \tag{5.28}$$

and integrating from 0 to x yields

$$\psi_+ = \frac{S_0}{2\Sigma_a}\left(1 - e^{-\Sigma_a x/\mu}\right) , \tag{5.29}$$

where we note $\psi_+ 0, \mu = 0$ from the given boundary condition.

For $\mu < 0$, we do similarly. It helps in this case to use $-|\mu|$ in place of μ, as it can be easy to lose negative signs. Using the integrating factor $\exp{(L - x)/|\mu|}$, we find after integrating from L to x that

$$\psi_- = \frac{S_0}{2\Sigma_a}\left(1 - e^{-\Sigma_a (L-x)/|\mu|}\right) . \tag{5.30}$$

For the case of $L = 10$ and $\Sigma_a = 1$, Figure 5.6 shows ψ for several values of ψ. Note the symmetry, as should be expected.

Our next goal is to compute the scalar flux. By definition, the scalar flux in 1-d is

$$\phi(x) = \int_{-1}^{1} d\mu \psi(x, \mu) \,, \tag{5.31}$$

which can be broken into

$$\phi(x) = \int_{-1}^{0} d\mu \psi_-(x, \mu) + \int_0^1 d\mu \psi_+(x, \mu) \,. \tag{5.32}$$

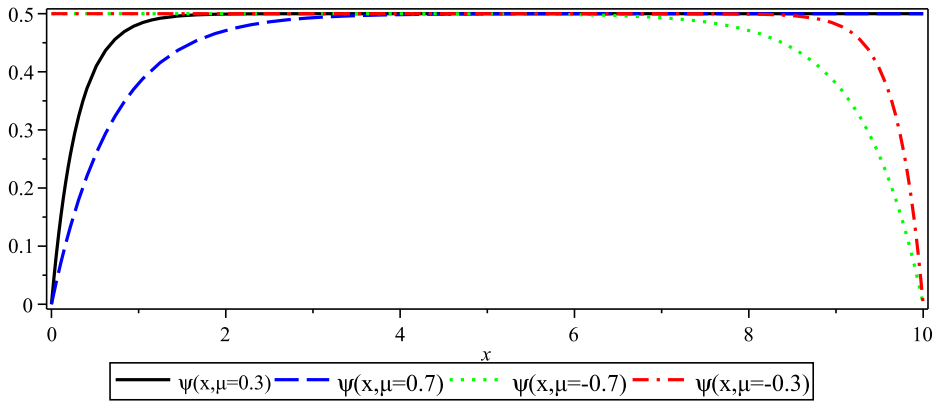

Figure 5.6: Angular flux at specific angles for the absorbing slab.

Inserting our expressions above, we have

$$\begin{aligned}\phi(x) &= \frac{S_0}{2\Sigma_a}\left(\int_{-1}^{1} d\mu - \int_{-1}^{0} d\mu e^{-\Sigma_a(L-x)/|\mu|} - \int_{0}^{1} d\mu e^{-\Sigma_a x/\mu}\right) \\ &= \frac{S_0}{2\Sigma_a}\left(2 - \int_{0}^{1} d\mu' e^{-\Sigma_a(L-x)/\mu} - \int_{0}^{1} d\mu e^{-\Sigma_a x/\mu}\right) \\ &= \frac{S_0}{2\Sigma_a}\left(2 - E_2(\Sigma_a(L-x)) - E_2(\Sigma_a x)\right).\end{aligned} \tag{5.33}$$

For the same numerical example, $\phi(x)$ is shown in Figure 5.7 along with the current density $\mathbf{J}(x)$, computation of which is left as an exercise.

Exponential Integrals

The functions $E_n(x)$ are called "exponential integrals" and are characteristic of slab problems. They are defined by

$$E_n(x) \equiv \int_0^1 d\mu \mu^{n-2} e^{-x/\mu}. \tag{5.34}$$

They also satisfy

$$E_n(x) = -\int dx E_{n-1}(x). \tag{5.35}$$

Several of the E_n functions are shown in Figure 5.8.

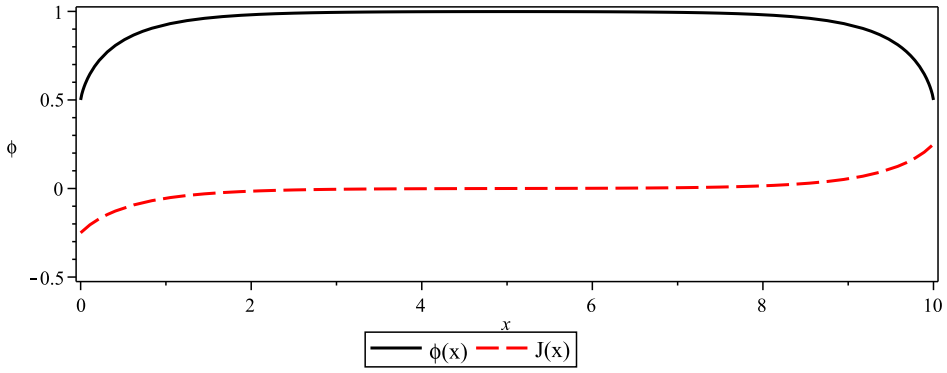

Figure 5.7: Scalar flux and current density.

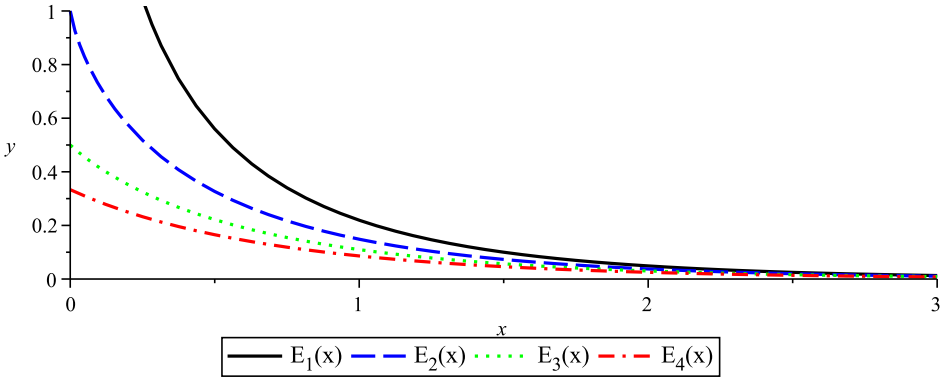

Figure 5.8: The first four exponential integral functions.

Further Reading

The problems discussed in this lecture have been rather elementary in nature. To explore more realistic problems—even just including scattering in slab geometry—would take us into a much more complicated mathematical domain involving integral transforms (with inverses using complex analysis), singular eigenfunction expansions, and more. Duderstadt and Martin [18] cover most of what has been discussed, and much more. The motivated student should look to that work (specifically Chapter 2), along with Bell and Glasstone (Chapter 2), Case and Zweifel [11], and Davison [15]. These are listed in reverse chronological order, and perhaps, in increasing order of difficulty. Davison in particular uses older notation, so the reader be aware!

Exercises

1. **Cylindrical Coordinates**. Derive the streaming term for the neutron transport equation in 1-d cylindrical geometry. Include diagrams to help explain your approach.
2. **Angular Current from Plane Source**. In Example 1, what is the angular current $\mathbf{j}$?
3. **Isotropic Current Conditions**. For Example 1, find ψ and $\mathbf{j}$ when the boundary condition is an isotropic angular current (rather than an isotropic angular flux).
4. **Flux Inside a Sphere**. Solve for the angular flux on the interior of the spherical shell from Example 2.
5. **Angular Redistribution in 1-d**. Explain in your own words why ψ in a 1-d spherical problem depends only on the radius r and cosine of the polar angle μ. Specifically, why is there no dependence on a second angular variable?
6. **Scalar Flux from Spherical Shell Source**. Solve for the scalar flux on the outside of the sphere in Example 2, i.e. integrate Eq. 5.24 over μ. What happens at $r = r_s$ and why? What is the limiting behavior of $\phi(r)$ as $r \to \infty$? In other words, what does the spherical source "look like" far from its surface?
7. **Current Density in a Slab**. Derive an expression for the current density $\mathbf{J}$ for a purely absorbing slab. Using the values from the example, generate the curves in Figure 5.7. Compute the total absorption rate in the slab and the leakage rate at the boundaries. Is it what you expect?
8. **Escape from Planet Slab**. Consider a uniform homogeneous slab of length L with Σ_t and Σ_s. Throughout the slab is a uniform isotropic source of strength S [n/s]. We are interested in the *escape probability*, $\mathrm{P_e}$, which is the probability that a neutron born uniformly and isotropically in the slab leaks without undergoing a collision. Derive an expression for $\mathrm{P_e}$ of a slab.

Lecture 6 | The Integral Form

In this lecture we consider the integral form of the transport equation in both general coordinates and the specific case of slab geometry. The integral form is useful in situations where only the scalar flux is required and is the foundation for the collision probability method, which we cover in the next lecture, as well as the method of characteristics, fast becoming the technique of choice in reactor analysis. We consider some analytical aspects of the integral equation and finish by discussing a simple numerical approach based on Neumann series expansions.

The Integral Transport Equation

You may have noticed from the previous two lectures that the angular dependence of the particle density is a relatively unique aspect of transport processes. Often, this angular dependence is the hardest aspect that we deal with directly, either analytically or numerically* It would be nice to eliminate the angular variable completely, and for certain problems, the *integral transport equation* allows us to do with a minimum of approximation.

Before we derive the integral form of the transport equation, it helps define a new quantity called the *emission density*,

$$Q(\mathbf{r}, \mathbf{\Omega}) = \int_{4\pi} d\Omega' \Sigma_s(\mathbf{r}, \mathbf{\Omega} \cdot \mathbf{\Omega}') \psi(\mathbf{r}, \mathbf{\Omega}') + S(\mathbf{r}, \mathbf{\Omega}) \,. \tag{6.1}$$

Essentially, Q is a generalized source containing any external source S and scattering source; of course, fission could also be included.

Recall the discussion of Eq. 5.3, where we consider a distance s from some reference point $\mathbf{r}_0$ along the characteristic $\mathbf{r}_0 + s\mathbf{\Omega}$. The full transport equation in

*Energy is by far more complex, as we've noted, and there are very few analytical problems where energy can be handled directly (slowing down in an infinite homogeneous medium problem is one example). Consequently, we typically use only the crudest representation possible in discretizing the energy variable, though much physics goes into generating the discrete data.

terms of s is

$$\frac{d}{ds}\psi(\mathbf{r}_0 + s\mathbf{\Omega}, \mathbf{\Omega}) + \Sigma_t(\mathbf{r}_0 + s\mathbf{\Omega})\psi = Q(\mathbf{r}_0 + s\mathbf{\Omega}, \mathbf{\Omega}) \,, \tag{6.2}$$

which can be called the *forward characteristic* equation, as we follow neutrons forward from the reference point. Similar, and which we use below, is the *backward characteristic* form, which represents following neutrons backward along the characteristic from a current point $\mathbf{r}$. Let $p = -s$. Then $d/dp = -d/ds$ and Eq. 6.2 becomes (after dropping the 0 subscript from $\mathbf{r}$)

$$-\frac{d}{dp}\psi(\mathbf{r} - p\mathbf{\Omega}, \mathbf{\Omega}) + \Sigma_t(\mathbf{r} - p\mathbf{\Omega})\psi = Q(\mathbf{r} - p\mathbf{\Omega}, \mathbf{\Omega}) \,. \tag{6.3}$$

We wish to integrate from $p = 0$ to some maximum distance (eventually to be either infinity or a global boundary). Introducing the integrating factor

$$if = e^{-\int_0^p \Sigma_t(\mathbf{r} - p'\mathbf{\Omega})dp'} \tag{6.4}$$

into Eq. 6.3, we have

$$-\frac{d}{dp}\Big(\psi(\mathbf{r}_0 - p\mathbf{\Omega}, \mathbf{\Omega})e^{-\int_0^p \Sigma_t(\mathbf{r} - p'\mathbf{\Omega})dp'}\Big) = Q(\mathbf{r} - p\mathbf{\Omega})e^{-\int_0^p \Sigma_t(\mathbf{r} - p'\mathbf{\Omega})dp'} \,, \tag{6.5}$$

and integrating

$$\begin{aligned} -\int_{\psi(p=0)}^{\psi(p'=p)} \left(d\Big(\psi(\mathbf{r} - p'\mathbf{\Omega}, \mathbf{\Omega})e^{-\int_0^{p'} \Sigma_t(\mathbf{r} - p''\mathbf{\Omega})dp''}\Big)\right) = \\ \int_0^p Q(\mathbf{r} - p'\mathbf{\Omega}, \mathbf{\Omega})e^{-\int_0^{p'} \Sigma_t(\mathbf{r} - p''\mathbf{\Omega})dp''}dp' \end{aligned} \tag{6.6}$$

yields

$$\begin{aligned} \psi(\mathbf{r}, \mathbf{\Omega}) - \psi(\mathbf{r} - p\mathbf{\Omega}, \mathbf{\Omega})e^{-\int_0^p \Sigma_t(\mathbf{r} - p''\mathbf{\Omega})dp''} = \\ \int_0^p Q(\mathbf{r} - p'\mathbf{\Omega}, \mathbf{\Omega})e^{-\int_0^{p'} \Sigma_t(\mathbf{r} - p''\mathbf{\Omega})dp''}dp' \,. \end{aligned} \tag{6.7}$$

We can simplify the notation somewhat by defining the *optical pathlength* τ, such that

$$\tau(\mathbf{r}, \mathbf{r} - p\mathbf{\Omega}) = \int_0^p \Sigma_t(\mathbf{r} - p'\mathbf{\Omega})dp' \,. \tag{6.8}$$

Then, the integral equation for the angular flux becomes

$$\psi(\mathbf{r}, \mathbf{\Omega}) = \int_0^p Q(\mathbf{r} - p'\mathbf{\Omega}, \mathbf{\Omega})e^{-\tau(\mathbf{r}, \mathbf{r} - p'\mathbf{\Omega})}dp' + \psi(\mathbf{r} - p\mathbf{\Omega}, \mathbf{\Omega})e^{-\tau(\mathbf{r}, \mathbf{r} - p\mathbf{\Omega})} \,. \tag{6.9}$$

In many cases, the emission density Q is assumed to be isotropic in the lab system. In this case, we can integrate out the angular dependence and arrive at an integral equation for the scalar flux often referred to as Peierl's equation. To do so, we let the integration bound p of Eq. 6.9 go to infinity. We assume the second term on the right hand side to vanish (i.e. either ψ vanishes at infinity, or equivalently, the optical path length $\tau(\mathbf{r}, \infty)$ goes to ∞, and so the exponential vanishes).

Letting $Q(\mathbf{r}, \mathbf{\Omega}) = Q(\mathbf{r} - p\mathbf{\Omega})/4\pi$, and substituting $\mathbf{r}' = \mathbf{r} - p\mathbf{\Omega}$, we integrate Eq. 6.9 over all angles to get

$$\phi(\mathbf{r}) = \int_{4\pi} d\Omega \int_0^\infty \frac{Q(\mathbf{r}')}{4\pi} e^{-\tau(\mathbf{r},\mathbf{r}')} dp\,. \tag{6.10}$$

Now note that $p = |\mathbf{\Omega} p| = |\mathbf{r} - \mathbf{r}'|$, and consequently $p^2 = |\mathbf{r} - \mathbf{r}'|^2$. Multiplying within the integrand by $1 = p^2/|\mathbf{r} - \mathbf{r}'|^2$ yields

$$\phi(\mathbf{r}) = \int_{4\pi} d\Omega \int_0^\infty \frac{Q(\mathbf{r}')}{4\pi} e^{-\tau(\mathbf{r},\mathbf{r}')} \frac{p^2 dp}{|\mathbf{r} - \mathbf{r}'|^2}\,, \tag{6.11}$$

and if the reader thinks this looks suspiciously like a volume integral in spherical coordinates, she would be right. Letting $dV' = 4\pi d\Omega dp p^2$, we have

$$\phi(\mathbf{r}) = \int_{V'} \frac{Q(\mathbf{r}') e^{-\tau(\mathbf{r},\mathbf{r}')} dV'}{4\pi |\mathbf{r} - \mathbf{r}'|^2}\,. \tag{6.12}$$

Integral Transport in Slab Geometry

We now look at the special case of Eq. 6.12 in slab geometry, for which the emission density is a function only of x, i.e. $Q(\mathbf{r}) = Q(x)$. We make use of cylindrical coordinates, with the axis taken to be x. Our goal will be to integrate out the radial (ρ) and azimuthal (ω) spatial components, leaving just the x dependence. The differential volume is then

$$dV' = \rho d\rho d\omega dx', \tag{6.13}$$

and Eq. 6.12 becomes

$$\begin{aligned} \phi(x) &= \int_{-\infty}^{\infty} dx' \int_0^\infty d\rho\rho \int_0^{2\pi} \frac{d\omega' Q(x') e^{-\tau(\mathbf{r},\mathbf{r}')}}{4\pi |\mathbf{r} - \mathbf{r}'|^2} \\ &= 2\pi \int_{-\infty}^{\infty} dx' \int_0^\infty d\rho\rho \frac{Q(x') e^{-\tau(\mathbf{r},\mathbf{r}')}}{4\pi |\mathbf{r} - \mathbf{r}'|^2}\,. \end{aligned} \tag{6.14}$$

We now need to express $\mathbf{r}$ and ρ in terms of x. Since the cross-sections (as quantified by τ) are really dependent only on x, we can relate the full distance $|\mathbf{r}' - \mathbf{r}|$

with its projection along the x axis via a directional cosine λ^{-1} such that

$$\lambda = \frac{|\mathbf{r}' - \mathbf{r}|}{|x' - x|} \tag{6.15}$$

and $\tau(\mathbf{r}', \mathbf{r}) = \lambda\tau(x', x)$. Moreover,

$$|\mathbf{r}' - \mathbf{r}|^2 = \rho^2 + |x' - x|^2 \tag{6.16}$$

which, using Eq. 6.15, can be rewritten as

$$\rho^2 = (\lambda^2 - 1)|x' - x|^2 \,, \tag{6.17}$$

and differentiating, we find

$$\rho d\rho = \lambda d\lambda |x' - x|^2 \,. \tag{6.18}$$

Noting that $\rho = 0$ corresponds to $\lambda = 1$, Eq. 6.14 can be written in terms of λ to give

$$\begin{aligned} \phi(x) &= 2\pi \int_{-\infty}^{\infty} dx' \int_1^{\infty} \lambda d\lambda \frac{|x' - x|^2}{4\pi|\mathbf{r}' - \mathbf{r}|^2} Q(x') e^{-\tau(x,x')} \\ &= 2\pi \int_{-\infty}^{\infty} dx' \int_1^{\infty} \lambda d\lambda \frac{1}{4\pi\lambda^2} Q(x') e^{-\tau(x,x')} \\ &= \int_{-\infty}^{\infty} dx' \frac{1}{2} \int_1^{\infty} \lambda d\lambda \frac{1}{\lambda} Q(x') e^{-\tau(x,x')} \end{aligned} \tag{6.19}$$

or

$$\phi(x) = \int_{-\infty}^{\infty} dx' \frac{1}{2} E_1(\tau(x, x')) Q(x') \,, \tag{6.20}$$

where we have used the E_1 function defined at the end of Lecture 5.

First-Flight Kernels

Eq. 6.20 gives us an example of the use of a *first-flight kernel* for the scalar flux, the general use of which takes the form

$$\phi(\mathbf{r}) = \int d^3\mathbf{r}' k(\mathbf{r}, \mathbf{r}') Q(\mathbf{r}') \tag{6.21}$$

for a kernel $k(\mathbf{r}, \mathbf{r}')$. For slab geometry, the first-flight kernel is seen to be

$$k_{\text{slab}}(x, x') = \frac{1}{2} E_1(\tau(x, x')) \,. \tag{6.22}$$

First-flight kernels have a particularly easy (and important!) physical interpretation. Consider Eq. 6.20 for the case of a purely absorbing medium. Then the emissivity Q consists only of external sources. To help visualize the problem, take Q to be a delta function at x_0, i.e. $Q(x) = Q_0\delta(x - x_0)$. Substituting this into Eq. 6.20 gives

$$\phi(x) = \frac{1}{2}E_1(\tau(x, x_0))Q_0\,. \tag{6.23}$$

Thus, the kernel $k(x, x')$ can be seen to give the contribution of the source particles born at x' to the flux at x. In other words, it gives to us the *uncollided flux*. For many systems, having the uncollided flux can be a good approximation for the total flux, and in some numerical schemes, it can be a good initial guess to help reduce computational time and numerical artifacts (e.g. the discrete ordinates method, discussed in Lecture 11).

If we look back at the Peierl's equation (Eq. 6.12), we find the fundamental first-flight kernel of the point source,

$$k_{\text{point}}(\mathbf{r}, \mathbf{r}') = \frac{e^{-\tau(\mathbf{r},\mathbf{r}')}}{4\pi|\mathbf{r} - \mathbf{r}'|^2}\,. \tag{6.24}$$

Two things are worth noting about Eq. 6.24. First, the first-flight kernels for all other geometrical configurations can be derived from this kernel. A second point, related to the first, is that the point kernel is closely related to the *Green's function* for the transport equation.

Green's Functions

A Green's function $G(x, x')$ for a linear differential operator[†] $L = L(x)$ is defined

$$LG(x, x') = \delta(x - x')\,. \tag{6.25}$$

A linear differential operator is any linear combination of basic differentiation operators. L could be d/dx or d^2/dx^2 or $d^2/dx^2 + d/dx$, and so on. The utility of G arises when we wish to solve the inhomogeneous differential equation

$$Lu(x) = f(x)\,. \tag{6.26}$$

If we multiply both sides of Eq. 6.25 by $f(x')$ and integrate over x', we find

$$\int LG(x, x')f(x')dx' = \int dx'\delta(x - x')f(x') = f(x)\,, \tag{6.27}$$

[†]We'll discuss the linearity of the transport equation in Lecture **??**, and we'll use operator notation extensively in Lecture 16.

but this suggests that

$$Lu(x) = \int LG(x,x')f(x')dx' = L\int G(x,x')f(x')dx'\,, \tag{6.28}$$

or

$$u(x) = \int G(x,x')f(x')dx'\,. \tag{6.29}$$

Hence, if we know $G(x,x')$, then we can solve the inhomogeneous equation for u.

What about the transport equation? Consider again Eq. 6.9, neglecting the second term, and letting $p \to \infty$, i.e.

$$\psi(\mathbf{r},\mathbf{\Omega}) = \int_0^\infty Q(\mathbf{r}-p\mathbf{\Omega},\mathbf{\Omega})e^{-\tau(\mathbf{r},\mathbf{r}-p\mathbf{\Omega})}dp\,. \tag{6.30}$$

Note that this is still integrating along the characteristic. It is more convenient to cast this as volume integral, similar to what we did above for Eq. 6.12. However, even in volume form, we still want the integration confined to the characteristic. By defining

$$\delta(\mathbf{\Omega}\cdot\mathbf{\Omega}') \equiv \delta(\mu-\mu')\delta(\phi-\phi')\,, \tag{6.31}$$

and

$$\mathbf{\Omega}_R \equiv \frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|}\,, \tag{6.32}$$

and recalling $p = |\mathbf{r}-\mathbf{r}'|$, we can rewrite Eq. 6.30 as

$$\psi(\mathbf{r},\mathbf{\Omega}) = \int_{4\pi} d\Omega_R \delta(\mathbf{\Omega}\cdot\mathbf{\Omega}_R)\int_0^\infty Q(\mathbf{r}',\mathbf{\Omega}_R)e^{-\tau(\mathbf{r},\mathbf{r}')}dp\,. \tag{6.33}$$

Using $dV' = p^2 dp d\Omega_R$, this becomes

$$\psi(\mathbf{r},\mathbf{\Omega}) = \int_{V'} dV' \frac{1}{|\mathbf{r}-\mathbf{r}'|^2}\delta(\mathbf{\Omega}\cdot\mathbf{\Omega}_R)Q(\mathbf{r}',\mathbf{\Omega}_R)e^{-\tau(\mathbf{r},\mathbf{r}')}\,. \tag{6.34}$$

Now let Q be a delta source at $\mathbf{r}_0$ emitting particles in direction $\mathbf{\Omega}_0$, or $Q = \delta(\mathbf{r}-\mathbf{r}_0)\delta(\Omega\cdot\Omega_0)$, similar to the right hand side of Eq. 6.25. Then

$$\begin{aligned}
\psi(\mathbf{r},\mathbf{\Omega}) &= \int_{V'} dV' \frac{e^{-\tau(\mathbf{r},\mathbf{r}')}}{|\mathbf{r}-\mathbf{r}'|^2}\delta(\mathbf{\Omega}\cdot\mathbf{\Omega}_R)\delta(\mathbf{r}-\mathbf{r}_0)\delta(\mathbf{\Omega}\cdot\mathbf{\Omega}_0) \\
&= \frac{e^{-\tau(\mathbf{r},\mathbf{r}_0)}}{|\mathbf{r}-\mathbf{r}_0|^2}\delta\left(\mathbf{\Omega}\cdot\frac{\mathbf{r}-\mathbf{r}_0}{|\mathbf{r}-\mathbf{r}_0|}\right)\delta(\mathbf{\Omega}\cdot\mathbf{\Omega}_0) \\
&= G_{\text{point}}(\mathbf{r},\mathbf{\Omega};\mathbf{r}_0,\mathbf{\Omega}_0)\,.
\end{aligned} \tag{6.35}$$

This is the Green's function for the angular flux, with which we can find ψ for any emission density $Q(\mathbf{r}, \mathbf{\Omega})$[‡]. We can use G_{point} to recover the scalar flux point kernel by letting Q be a unit isotropic point source, i.e. $Q = \delta(\mathbf{r} - \mathbf{r}_0)/4\pi$. Then

$$\begin{aligned}\psi(\mathbf{r}, \mathbf{\Omega}) &= \int_V dV' \int_{4\pi} d\Omega' G_{\text{point}}(\mathbf{r}, \mathbf{\Omega}; \mathbf{r}', \mathbf{\Omega}') \frac{\delta(\mathbf{r}' - \mathbf{r}_0)}{4\pi} \\ &= \frac{e^{-\tau(\mathbf{r}, \mathbf{r}_0)}}{4\pi|\mathbf{r} - \mathbf{r}_0|^2} \delta\left(\mathbf{\Omega} \cdot \frac{\mathbf{r} - \mathbf{r}_0}{|\mathbf{r} - \mathbf{r}_0|}\right).\end{aligned} \tag{6.36}$$

Integrating ψ over all angles yields

$$\begin{aligned}\phi(\mathbf{r}) &= \int_{4\pi} d\Omega \frac{e^{-\tau(\mathbf{r}, \mathbf{r}_0)}}{4\pi|\mathbf{r} - \mathbf{r}_0|^2} \delta\left(\mathbf{\Omega} \cdot \frac{\mathbf{r} - \mathbf{r}_0}{|\mathbf{r} - \mathbf{r}_0|}\right) \\ &= \frac{e^{-\tau(\mathbf{r}, \mathbf{r}_0)}}{4\pi|\mathbf{r} - \mathbf{r}_0|^2},\end{aligned} \tag{6.37}$$

which is indeed our point kernel for the scalar flux.

Neumann Series

Lecture 7 will cover the widely-used collision probability method that is based on the integral transport equation. Here, we investigate a less versatile yet quite enlightening method based on expansion of the scalar flux in a so-called Neumann[§] series for slab problems that include scattering.

Consider the integral equation in a slab of length L subject to vacuum boundary conditions:

$$\phi(x) = \int_0^L dx' \frac{1}{2} E_1(\tau(x, x'))(S(x') + \phi(x')\Sigma_s(x')). \tag{6.38}$$

The integration bounds are 0 and L, since all sources must vanish in the outside vacuum. Defining the operator K,

$$K\phi = \int_0^L dx' \frac{1}{2} E_1(\tau(x, x'))\phi(x)\Sigma_s(x), \tag{6.39}$$

we can rewrite the integral equation as

$$(I - K)\phi = K\frac{S}{\Sigma_s}. \tag{6.40}$$

[‡]Careful! If Q includes scattering, then it depends on ψ, and so a direct solution in this form is not possible

[§]Carl Neumann, not to be confused with John von Neumann.

Without going into formal detail (though it should seem "reasonable"), Eq. 6.40 can be solved by expanding

$$(I - K)^{-1} = I + K + K^2 + \dots , \tag{6.41}$$

where I is the identity operator, so that

$$\phi(x) = \sum_{n=0}^{\infty} K^n \frac{S(x)}{\Sigma_s(x)} . \tag{6.42}$$

Suppose we define

$$\phi_n(x) \equiv K^n \frac{S(x)}{\Sigma_s(x)} . \tag{6.43}$$

Then we see that $\phi_0 = K(S/\Sigma_s)$, $\phi_1 = K^2(S/\Sigma_s) = K(\phi_0)$ and so on. We recognize $\phi_0(x)$ as the uncollided flux, which is then used as input for ϕ_1. We recognize ϕ_1 as those neutrons already having undergone a single collision (and no more). In general, we call ϕ_n the nth collided flux. The Neumann series Eq. 6.42 just adds up all neutrons that have not collided, those that have collided once, those that have twice, and so on, thus capturing the entire population of neutrons in the system. Defining (and computing) ϕ_n in terms of the previous term ϕ_{n-1} sequence is called *Neumann iteration*.

A Numerical Example

We illustrate the use of Neumann iteration in a simple MATLAB code for a homogeneous slab problem with a uniform isotropic source. The integrals involved are approximated using the trapezoid rule, though the exercises explore other schemes. Please refer to the code comments in Listing 6.1 to understand the exact implementation of the algorithm.

One aspect of the algorithm that should be pointed out is use of the method of *subtraction of singularity*. In the numerical integration, we select some set of points x_i at which we compute the scalar flux. In a straight forward integration, this would require several evaluations of the term $E_1(\Sigma|x_i - x_i|)$, which blows up. As an alternative, one can write

$$\begin{aligned} \phi(x) &= \int_0^L k_{\text{slab}}(x, x')Q(x')dx' \\ &= \int_0^L k_{\text{slab}}(x, x')(Q(x') - Q(x))dx' + \int_0^L k_{\text{slab}}(x, x')Q(x)dx' . \end{aligned} \tag{6.44}$$

The second term of the final line can be integrated analytically, and you should do this to obtain the form used in the code. The first term vanishes identically for $x = x'$, which is why the code has the logical evaluation ***if****(i∼=0)* that if unsatisfied leaves $\phi_i = 0$.

Listing 6.1: Solution of Slab Problem via Neumann Series

```
function [phi,phiL] = neumann_slab(L,SigT,SigS,Q,N,M)
% function phi = neumann_slab(L,SigT,SigS,Q,N,M)
%   This function solves the one-speed transport equation in slab geometry
%   for the scalar flux using a Neumann series solution.  The integrals
%   are computed using the trapezoid rule.
%
%   The problem is a slab of width L, with uniform total cross-section
%   SigT, uniform (isotropic) scattering cross-section SigS, and
%   uniform isotropic source of volumetric strength Q.
%
%   Inputs:
%       L       -- slab width [cm]
%       SigT    -- total cross-section [1/cm]
%       SigS    -- scattering cross-section [1/cm]
%       Q       -- uniform isotropic source [1/cm^3-s]
%       N       -- number of points for evaluating flux
%       M       -- number of Neumann terms to compute
%   Output
%       phi     -- scalar flux [N,1]
%       phiL    -- Neumann terms [N,M+1] (M=0 yields uncollided only)
phi  = zeros(N,1);      % scalar flux
phiL = zeros(N,M+1);    % the Lth collided fluxes
QL   = Q*ones(N,1);     % source vector
% uncollided flux
phiL(:,1) = ithcollided(L,N,QL,SigT);
% collided fluxes
for i = 1:M
   QL          = SigS*phiL(:,i);             % compute new source
   phiL(:,i+1) = ithcollided(L,N,QL,SigT);   % get ith collided flux term
end
% scalar flux is just the sum
for i = 1:M+1
    phi = phi + phiL(:,i);
end
end % function neumann_slab

function phi = ithcollided(L,N,Q,SigT)
% function phi = ithcollided(L,N,Q,SigT)
%   This function computes a Neumann term for the given source.  The
%   exponential integral is computed using Matlab's symbolic function Ei,
%   which is pretty slow!  The method of subtraction of singularities has
%   been used.  The flux is computed as follows:
%     phi(x)   = int( k(x,x')*Q(x')dx', x=0,L )
%     phi(x_i) = h * ( 0.5*k(x_i,x_0)*Q(x_0) + 1.0*k(x_i,x_1) + ... )
%
%   Inputs:
%       L       -- slab width [cm]
%       N       -- number of points for evaluating flux
%       Q       -- source
%       SigT    -- total cross-section [1/cm]
%   Output
%       phi     -- the current Neumann term
phi = zeros(N,1);
h = L / (N-1);
% loop through all N
for i = 1:N
    xi = h*(i-1);
    phi(i) = 0;
```

```
    for j = 1:N % begin trapezoid integration
        if (i~=j) % if i=j, Q(i)=Q(j) and the term vanishes; avoids NaN!
            if (j==0 || j==N) % end points get coefficient of 0.5
                A = 0.5;
            else
                A = 1.0;
            end
            xj      = h*(j-1);
            phi(i)  = phi(i) + ...
                      A*h*mfun('Ei',1,(SigT*abs(xi-xj)))*(Q(j)-Q(i));
        end
    end
    % extra part from subtraction of the singularity
    phi(i) = phi(i) + 0.5*Q(i)/SigT * ...
             (2-mfun('Ei',2,SigT*(L-xi))-mfun('Ei',2,SigT*(xi)));
end
end % function ithcollided
```

Further Reading

The derivation of the integral equations generally follows that of Lewis and Miller [30]. First-flight kernels and Green's functions are covered in Duderstadt and Martin [18] and Case and Zweifel [11]. Solving for the scalar flux via Neumann iterations is discussed in Duderstadt and Martin [18], while its use as a general technique for integral equations is described e.g. by Arfken and Weber [8]. Implementation of the code and use of the subtraction of singularities method was inspired by notes of Prof. D. Henderson at Wisconsin. Any good numerical analysis textbook will provide more information on numerical integration, and the reader is encouraged to explore this fundamental topic.

Exercises

1. **Point-to-slab**. Show how k_{slab} can be generated using k_{point}.
2. **Line source**. Show that the first-flight kernel for an infinite isotropic line source is given by

$$k_{\text{line}}(r, r') = \frac{1}{2\pi r} Ki_1(\Sigma_t |r - r'|)\,, \tag{6.45}$$

where

$$Ki_n(x) = \int_0^{\pi/2} \cos^{n-1}(\theta) e^{-x/\cos(\theta)} d\theta = \int_0^{\infty} \frac{e^x \cosh(u)}{\cosh^n(u)} \tag{6.46}$$

are the Bickley-Naylor function of order n, and where $Ki_0(r) = K_0(r)$, the zeroth order modified Bessel function.

3. **Spherical shell**. Show that the first-flight kernel for a spherical shell is given by
$$k_{\text{sph}}(r, r') = \frac{r'}{2r} Ki_1(\Sigma_t|r - r'| - \Sigma_t|r + r'|)\,, \tag{6.47}$$
4. **Current density kernels**. In the lecture found the flux kernel for a plane source, and the kernels for line and spherical sources are given in the first two exercises. We can also derive kernels for the current density. For example, in a slab, there is a current kernel $\gamma_{\text{slab}}(x, x')$ such that
$$\mathbf{J}(x) = \int_0^L \gamma_{\text{slab}}(x, x')Q(x')dx'\,. \tag{6.48}$$
Derive this kernel $\gamma_{\text{slab}}(x, x')$.
5. **Using kernels**. Use k_{slab} to find the scalar flux for the example of Lecture 5, and use γ_{slab} to find the corresponding current density.
6. **Neumann Series**. Use the Neumann series code for the slab in Listing 6.1 to plot the scalar flux, uncollided flux, and the first five collided fluxes on the same graph using 20 spatial divisions for $L = 10$, $\Sigma_t = 1.0$, and (a) $\Sigma_s = 0.2$ and (b) $\Sigma_s = 0.8$. What changes for the case with higher scattering?
7. **Convergence**. For the same problem, modify the Neumann series code to compute as many collided fluxes are necessary so that
$$\frac{|\phi_n(x_i) - \phi_n(x_{i-1})|}{\phi_n(x_{i-1}} < \epsilon = 1 \times 10^{-6} \quad \forall\, x_i\,,$$
and comment on the results. (Hint, which norm on the x vector is appropriate?)
8. **Nonuniform medium**. Could the Neumann series code be easily modified to handle a nonuniform slab? Suggest an approach for this.
9. **Numerical integration**. Modify the Neumann series code to handle both Simpson's rule and Gaussian quadrature. Overviews of both can be found in numerical analysis texts. For the purely absorbing slab example of Lecture 5, compare the trapezoid, Simpson's, and Gaussian schemes to the analytical solution. Comment on which method appears to get it "right" with the least effort.
10. **Efficiency**. The Neumann code is slow due to the symbolic computation of the E_n functions. Modify the code to precomputing the various coefficients (indexed by i and j), and comment on any improvement. For even greater efficiency, find a way to evaluate the exponential functions numerically (via some form of approximation) that allows you compute $E_n(x)$ rapidly; see for example Hebert's *Applied Reactor Physics*. How does the numerical evaluation affect the accuracy of the result?

Lecture 7 | Collision Probability Method

In the last lecture, we discussed the integral transport equation, derived its form in slab geometry, and provided a simple numerical approach based on Neumann iteration and numerical quadrature for solving homogeneous slab problems. In this lecture, we discuss the rather versatile *collision probability method* in slab geometry, though the method is certainly applicable to higher dimensions. We finish by providing a simple code that serves as the basis for several exercises.

Collision Probabilities

Recall Eq. 6.20, the integral equation in slab geometry:

$$\phi(x) = \int_{-\infty}^{\infty} dx' \frac{1}{2} E_1(\tau(x,x'))Q(x')\,, \tag{6.20}$$

where the emission density Q contains any external and scattering sources, all assumed to be isotropic. Suppose we apply this equation to a finite slab of length L with vacuum boundaries, much as we did toward the end of Lecture 6, or

$$\phi(x) = \int_{0}^{L} dx' \frac{1}{2} E_1(\tau(x,x'))Q(x')\,. \tag{7.1}$$

Let us assume the slab can be divided into a number of regions in which all cross-sections and external sources uniform, as in Figure 7.1. Within a region i spanning $x_{i-1/2}$ to $x_{i+1/2}$, we define an average flux

$$\phi(x_i) = \phi_i \equiv \frac{1}{\Delta_i} \int_{x_{i-1/2}}^{x_{i+1/2}} \phi(x)dx\,, \tag{7.2}$$

where $\Delta_i \equiv x_{i+1/2} - x_{i-1/2}$. Substituting the expression for ϕ in Eq. 7.1 into Eq. 7.2 yields

$$\phi_i = \frac{1}{\Delta_i} \int_{x_{i-1/2}}^{x_{i+1/2}} dx \int_{0}^{L} dx' \frac{1}{2} E_1(\tau(x,x'))Q(x')\,. \tag{7.3}$$

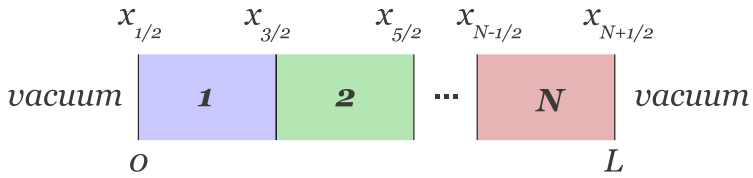

Figure 7.1: Slab discretization for CPM.

If we assume the flux within the cell is its average (a constant), then the emission density is also constant, and we find

$$\int_0^L dx' \frac{1}{2} E_1(\tau(x,x')) Q(x') dx \approx \sum_{i=1}^{N} Q_i \int_{x_{i-1/2}}^{x_{i+1/2}} dx' \frac{1}{2} E_1(\tau(x,x')) \,. \tag{7.4}$$

Substituting this into Eq. 7.3 (and being mindful of primes) yields

$$\begin{aligned} \phi_i &= \frac{1}{\Delta_i} \int_{x_{i-1/2}}^{x_{i+1/2}} dx \left(\sum_{i'=1}^{N} Q_{i'} \int_{x_{i'-1/2}}^{x_{i'+1/2}} dx' \frac{1}{2} E_1(\tau(x,x')) \right) \\ &= \frac{1}{\Delta_i} \sum_{i'=1}^{N} Q_{i'} \int_{x_{i-1/2}}^{x_{i+1/2}} dx \int_{x_{i'-1/2}}^{x_{i'+1/2}} dx' \frac{1}{2} E_1(\tau(x,x')) \end{aligned} \tag{7.5}$$

The collision probability method (as the name might imply) actually deals in terms of reactions rates rather than just the flux. If we multiply both sides of 7.5 by $\Sigma_i \Delta_i$, where Σ_i is the total cross-section in region i, we get an equation for $\phi_i \Sigma_i \Delta_i$, the total collision rate within region i. This equation is

$$\begin{aligned} \phi_i \Sigma_i \Delta_i &= \Sigma_i \sum_{i'=1}^{N} Q_{i'} \int_{x_{i-1/2}}^{x_{i+1/2}} dx \int_{x_{i'-1/2}}^{x_{i'+1/2}} dx' \frac{1}{2} E_1(\tau(x,x')) \\ &= \sum_{i'=1}^{N} Q_{i'} \Delta_{i'} \frac{\Sigma_i}{\Delta_{i'}} \int_{x_{i-1/2}}^{x_{i+1/2}} dx \int_{x_{i'-1/2}}^{x_{i'+1/2}} dx' \frac{1}{2} E_1(\tau(x,x')) \\ &= \sum_{i'=1}^{N} Q_{i'} \Delta_{i'} P_{ii'} \,, \end{aligned} \tag{7.6}$$

where

$$P_{ii'} \equiv \frac{\Sigma_i}{\Delta_{i'}} \int_{x_{i-1/2}}^{x_{i+1/2}} dx \int_{x_{i'-1/2}}^{x_{i'+1/2}} dx' \frac{1}{2} E_1(\tau(x,x')) \tag{7.7}$$

is the *first-flight collision probability*. Notice we rearranged terms such that $P_{ii'}$ is unitless. Moreover, $P_{ii'}$ can be interpreted as the probability that a neutron born uniformly and isotropically in a region i' makes its first collision in region i.

Solving the Equations

Eq. 7.6 provides a set of equation for the scalar flux (or rather, the collision rates) in terms of the region emission densities Q_i and collision probabilities $P_{ii'}$. In our case of slab geometry, we can compute the collision probabilities analytically. We simply list the results, leaving the derivations as exercises:

$$\begin{aligned} P_{ii} &= 1 - \frac{1}{2\Sigma_i\Delta_i}\Big(1 - 2E_3(\Sigma_i\Delta_i)\Big) \\ P_{ii'} &= \frac{1}{2\Sigma_i'\Delta_i'}\Big(E_3(\tau_{ii'}) - E_3(\tau_{ii'} + \Delta_i\Sigma_i) - E_3(\tau_{ii'} + \Delta_{i'}\Sigma_{i'}) \\ &\quad + E_3(\tau_{ii'} + \Delta_i\Sigma_i + \Delta_{i'}\Sigma_{i'})\Big)\,, \quad i \neq i'\,, \end{aligned} \tag{7.8}$$

where

$$\tau_{ii'} \equiv \begin{cases} \tau(x_{i+1/2}, x_{i'-1/2}) & i' > i \\ \tau(x_{i'+1/2}, x_{i-1/2}) & i' < i \end{cases}. \tag{7.9}$$

So how do we solve the equations once we have the $P_{ii'}$? We have a system of equations of the form

$$\Delta_i\Sigma_i\phi_i = \sum_{i'=1}^{N}(\Delta_{i'}\Sigma_{si'}\phi_{i'} + S_{i'}\Delta_{i'})P_{ii'}\,. \tag{7.10}$$

If we rearrange terms, bringing all ϕ's to the left hand side, we can write for $i = 1$

$$\Delta_1(\Sigma_1 - P_{11}\Sigma_{s1})\phi_1 - P_{12}\Delta_2\Sigma_{s2}\phi_2 + \ldots = S_1\Delta_1P_{11} + S_2\Delta_2P_{12} + \ldots \tag{7.11}$$

If we define $f_i = \Delta_i\Sigma_i\phi_i$ and $s_i = \sum_{i'} S_{i'}\Delta_{i'}P_{ii'}$, this becomes

$$\Big(1 - \frac{\Sigma_{s1}}{\Sigma_1}P_{11}\Big)f_1 + \Big(-\frac{\Sigma_{s2}}{\Sigma_2}P_{12}\Big)f_2 + \ldots = s_i\,. \tag{7.12}$$

Generalizing to matrix form, we have

$$\mathbf{H}\mathbf{f} = \mathbf{s}\,, \tag{7.13}$$

where

$$\mathbf{H} = \begin{bmatrix} 1 - \frac{\Sigma_{s1}}{\Sigma_1}P_{11} & -\frac{\Sigma_{s2}}{\Sigma_2}P_{12} & -\frac{\Sigma_{s3}}{\Sigma_2}P_{13} & \cdots \\ -\frac{\Sigma_{s1}}{\Sigma_1}P_{21} & 1 - \frac{\Sigma_{s2}}{\Sigma_2}P_{22} & -\frac{\Sigma_{s3}}{\Sigma_2}P_{23} & \cdots \\ & & & \ddots \end{bmatrix}. \tag{7.14}$$

For a purely absorbing case, Eq. 7.13 is quite simple, as $\mathbf{H}$ reduces to the identity matrix. Suppose we have a four region problem, with a source in just the third region. The resulting set of linear equations is simply

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta_1\Sigma_1\phi_1 \\ \Delta_2\Sigma_2\phi_2 \\ \Delta_3\Sigma_3\phi_3 \\ \Delta_4\Sigma_4\phi_4 \end{bmatrix} = \begin{bmatrix} S_3P_{13}\Delta_3 \\ S_3P_{23}\Delta_3 \\ S_3P_{33}\Delta_3 \\ S_3P_{43}\Delta_3 \end{bmatrix}, \tag{7.15}$$

the solution of which is simply $\phi_1 = (S_3P_{13}\Delta_3)/(\Delta_1\Sigma_1)$ and so forth.

Benefits and Limitations

Eq. 7.13 is powerful (as are integral methods in general) in that no angular approximation in the flux are made. However, the sources and scattering are assumed to be isotropic, which does limit to an extent the accuracy in angle.

Furthermore, the collision probability suffers from two rather significant deficiencies. First, the matrix $\mathbf{H}$ is in general a dense matrix. Hence, for large problems, the memory requirements are significant. Furthermore, the flat flux approximation within a region is only first-order accurate (meaning errors are proportional to Δ_i).

More Advanced Aspects

In the discussion above, we considered only vacuum conditions. The collision probability method has been most heavily used in lattice physics calculations, where the domain is often represented as an infinite array of identical unit cells via periodic boundary conditions.

To derive these conditions would be beyond our intended scope, but the reader might conclude the derivation would, theoretically, involve infinite sums over $P_{ii'}$, and this is indeed the case. The interested reader is encouraged to read Section 3.8.3 of Hebert [21] for discussion of a suitable numerical approach.

A Simple Code

Here, we provide a simple code for computing the scalar flux in a slab consisting of several regions of uniform composition, given in Listing 7.1. The fundamental components of the code are just those that compute τ and $P_{ii'}$ and should be readily understood. The reciprocity relations discussed in Lecture **??** are used to reduce some of the computation. As in the Neumann iteration code of the last lecture, MATLAB's built in symbolic function for the exponential integrals is used. This is very slow,

and so the student is encouraged to find efficient and accurate numerical techniques to evaluate these functions.

Listing 7.1: Solution of Slab Problem via CPM

```
function phi = cpm(Delta,SigT,SigS,Source)
% function phi = cpm(Delta,SigT,SigS,Source)
%   This function solves the one-speed transport equation in slab geometry
%   for the scalar flux using a the collision probability method.
%
%   Inputs: (all must be column vectors, i.e dimension [n,1]
%       Delta   -- region widths [cm]
%       SigT    -- region total cross-section [1/cm]
%       SigS    -- region scattering cross-section [1/cm]
%       Source  -- region uniform isotropic source [1/cm^3-s]
%   Output
%       phi     -- scalar flux in each region
tau = comptau(Delta,SigT);  % compute tau
P = ffcp(Delta,SigT,tau);   % compute collision probabilities
H = zeros(length(Delta));   % set up H and right hand side
s = zeros(length(Delta),1); % right hand side
for i = 1:length(Delta)
    for j = 1:length(Delta)
        H(i,j) = -SigS(j)/SigT(j)*P(i,j);        % add diagonal 1's below
        s(i) = s(i) + Source(j)*Delta(j)*P(i,j); % right hand side
    end
end
H = H + eye(length(Delta));     % add 1's to the diagonal
f = ( H \ s);                   % solve for the collision rates
phi = f ./ ( SigT .* Delta );   % solve for the flux
end

function tau = comptau(Delta,SigT)
% function tau = comptau(Delta,SigT)
%   This function computes the optical pathlengths between regions.
%
%   Inputs:
%       Delta   -- region widths [cm]
%       SigT    -- region total cross-section [1/cm]
%   Output:
%       tau     -- optical path length
tau = zeros(length(Delta));
DeltaSigT = Delta.*SigT;    % length of regions in mfp's
for i = 1:length(Delta)
    % tau(i,i) remains 0
    for ip = i+1:length(Delta) % only ip > i, since tau(x,x')=tau(x',x)
        % adding optical distance between x(i+1/2) and x(ip-1/2)
        tau(i,ip) = sum( DeltaSigT(i+1:ip-1) );
        tau(ip,i) = tau(i,ip);
    end
end
end

function P = ffcp(Delta,SigT,tau)
% function P = ffcp(Delta,SigT,tau)
%   This function computes the first-flight collision probabilities.  Note,
%   it uses MATLAB's symbolic function for E3, which is REALLY SLOW.  The
%   student is strongly encouraged to look up fast and accuracte numerical
%   approximations to the En functions.  See Hebert's textbook.
%
%   Inputs:
%       Delta   -- region widths [cm]
%       SigT    -- region total cross-section [1/cm]
%       tau     -- optical path length
%   Outputs:
%       P       -- first-flight collision probabilities
```

```
P = zeros(length(Delta));
for i = 1:length(Delta)
    P(i,i) = 1 - 0.5/SigT(i)/Delta(i) * ...
             ( 1 - 2*mfun('Ei', 3, SigT(i)*Delta(i)) );
    for j = 1:i-1
        P(i,j) = 0.5/SigT(j)/Delta(j) * ...
             ( mfun('Ei', 3, tau(i,j)) - ...
               mfun('Ei', 3, tau(i,j)+Delta(i)*SigT(i)) - ...
               mfun('Ei', 3, tau(i,j)+Delta(j)*SigT(j)) + ...
               mfun('Ei', 3, tau(i,j)+Delta(j)*SigT(j)+Delta(i)*SigT(i) ));
        % reciprocity in action
        P(j,i) = P(i,j)*(Delta(j)*SigT(j))/(Delta(i)*SigT(i));
    end
end
end
```

Further Reading

This lecture has largely followed Section 5-3 of Lewis and Miller [30]. There, the student will find some more details regarding periodic conditions for slab geometry as well as some information regarding a 2-d implementation of CPM in later sections. Hebert [21] also discusses CPM in several geometries and provides sample code relevant for both slab and a 2-d pin cell problems.

Exercises

1. **CPM and Scattering**. Use the CPM code to repeat Exercise 6 of Lecture 6. Does increase scattering have any effect on the computational time? Why or why not?
2. **Source Iteration**. In our (relatively standard) formulation of the collision probability method, we end up with a dense matrix $\mathbf{H}$. Reformulate the approach such that the scattering term is *not* brought to the left hand side. In this case, $\mathbf{H}$ is just the identity. However, the right hand side now contains an additional scattering source contribution, which depends on the (unknown) fluxes. Develop an approach such that the flux is initially guessed, the right hand side is formed, new fluxes are computed, and the right hand side is updated, yielding an iterative process called *source iteration*.
3. **CPM with Source Iteration**. Implement the *source iteration* you developed in the last exercise in the given CPM code. Repeat Exercise 6 from Lecture 6, using a stopping criterion of

$$\frac{|\phi_i^n - \phi_i^{n-1}|}{\phi_i^{n-1}} < \epsilon = 1 \times 10^{-6} \quad \forall\, i\,,$$

where n is the iteration index. Comparing to Exercise 1, is there any benefit to using source iteration in place of solving a dense matrix system directly? If not for this particular problem, underwhat conditions would you expect source iteration to be better?

4. **Understanding Periodic Conditions**. Read Section 3.8.3 of Hebert and provide a short synopsis of period conditions can be treated in slab geometry.
5. **Implementing Periodic Conditions**. Implement periodic conditions in the given CPM code, following the method described by Hebert.

Lecture 8 | CPM in Cylindrical Coordinates

In this lecture, the collision probability method is derived for 1-D cylindrical coordinates and specialized for the case of annular systems. For many years, this approach was fundamental for pin-cell spectrum calculations.

Sample Code

To be continued.

Further Reading

Lewis and Miller. Hebert. Carlvik 1964 paper.

Exercises

1. **White Boundary Conditions**.
   Collision probabilities found using the methods discussed in this lecture can be modified to account for white boundary conditions by setting

   $$R_i = \Sigma_i V_i - \sum_j \Sigma_j V_j P_{j\to i} \,,$$

   and defining updated collision probabilities

   $$\Sigma_i V_i P_{j\to i}^{\text{white}} = \Sigma_i V_i P_{j\to i} - \frac{R_i R_j}{\sum_k R_k} \,.$$

   Prove this expression by using reciprocity relations.

Lecture 9 | $\mathrm{P_N}$ Method and Diffusion

We continue in this lecture our treatment of the integro-differential form of the transport equation. Here, we turn to expansion of the angular variable in an orthogonal basis, the well-known Legendre polynomials in slab geometry and the spherical harmonics in multidimensional settings. This approach is typically referred to generically as the P_N method, which is the common symbol for the Legendre polynomials. We'll first start with a brief discussion of anisotropic scattering and how it can be handled using orthogonal expansions. We'll then derive the P_N equations for slab geometry, show how one formally arrives at the *neutron diffusion equation*, and provide a simple code that illustrates some concepts.

Anisotropic Scattering

In the past several lectures, we presented equations appropriate for scattering that is isotropic in the laboratory system. While it is often appropriate to assume isotropic scattering in the center of mass system (i.e. s-wave scattering), this corresponds to a forward peaked distribution in the laboratory system. This is especially true for light scatterers, which are important moderators. Since the neutron transport equation itself lives in the laboratory system, to account even for isotropic center of mass scattering requires we devise a way to include angular dependence in our scattering cross-section.

Recall that the macroscopic scattering cross-section can be expressed by $\Sigma_s(\hat{\Omega}' \rightarrow \hat{\Omega})$, which roughly quantifies the probability that a neutron going in direction $\hat{\Omega}'$ will collide and end up in the direction $\hat{\Omega}$. However, if the medium is isotropic, recall that this probability depends only on the cosine of the angle between the two direction vectors, i.e

$$\Sigma_s(\hat{\Omega}' \rightarrow \hat{\Omega}) = \Sigma_s(\hat{\Omega}' \cdot \hat{\Omega}) = \Sigma_s(\mu_0)\,, \tag{9.1}$$

where μ_0 is the cosine of the scattering angle. Note, this is not valid for a moving medium (perhaps a molten, flowing fuel) or single crystals.

The expression $\Sigma_s(\mu_0)$ is often expressed as an expansion in Legendre polynomials. This not only provides a satisfactory way to represent experimental data points in function form, but it also facilitates many analytic and numerical treatments of anisotropic scattering. Using the Legendre polynomials $P_l(\mu_0)$, we write

$$\Sigma_s(\mu_0) = \sum_{l=0}^{\infty} \frac{2l+1}{4\pi} \Sigma_{sl} P_l(\mu_0)\,, \tag{9.2}$$

where the Legendre coefficients (or *Legendre moments*) are defined

$$\Sigma_{sl} = 2\pi \int_{-1}^{1} \Sigma_s(\mu_0) P_l(\mu_0) d\mu_0\,. \tag{9.3}$$

Legendre Polynomials

So what are the Legendre polynomials? There are many ways to express them, including as the solution of the Legendre differential equation,

$$(1-x^2)\frac{d^2P_l}{dx^2} - 2x\frac{dP_l}{dx} + l(l+1)P_l(x) = 0\,. \tag{9.4}$$

This equation is related to Laplace's equation in spherical coordinates, for which full solutions take the form of the *spherical harmonics*. Additionally, the Legendre polynomials are orthogonal over $-1 < x < 1$, satisfying

$$\int_{-1}^{1} P_n(x) P_m(x) dx = \frac{2}{2n+1}\delta_{mn}\,, \tag{9.5}$$

where δ_{mn} is the Kronecker delta. In our derivation of the P_n equations, we'll use the Legendre recurrence relation,

$$(l+1)P_{l+1}(x) - (2l+1)xP_l(x) + lP_{l-1}(x) = 0\,. \tag{9.6}$$

For reference, the Legendre polynomials through $N = 5$ are given in Figure 9.1, and are defined

$$\begin{aligned}
P_0(x) &= 1 \\
P_1(x) &= x \\
P_2(x) &= \frac{1}{2}(3x^2 - 1) \\
P_3(x) &= \frac{1}{2}(5x^3 - 3x) \\
P_4(x) &= \frac{1}{8}(35x^4 - 30x + 3) \\
P_5(x) &= \frac{1}{8}(64x^5 - 70x^3 + 15x)\,.
\end{aligned} \tag{9.7}$$

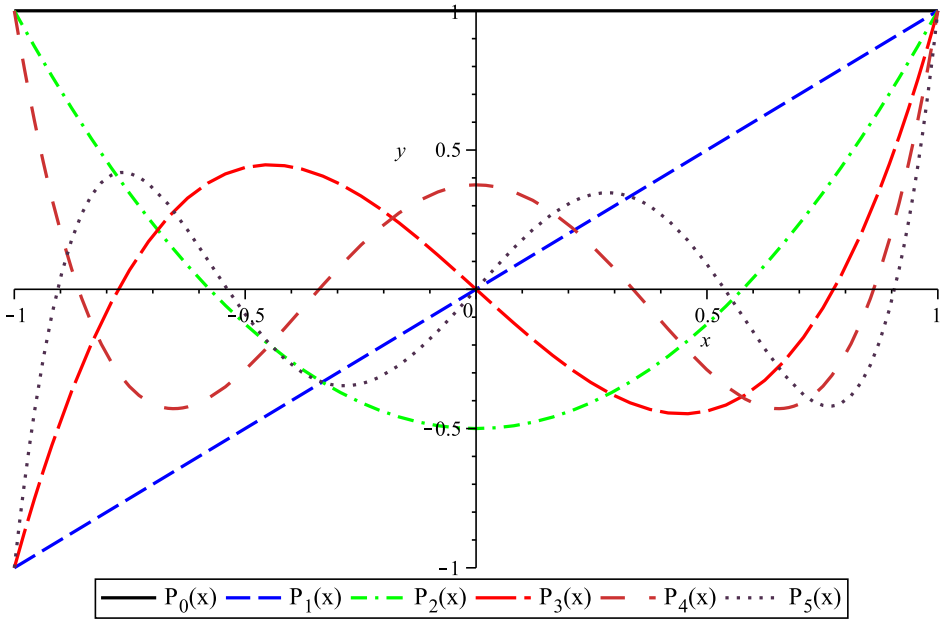

Figure 9.1: The first several Legendre polynomials, $P_N(x)$.

The Legendre polynomials are useful for expanding the scattering cross-section because they are orthogonal over the range $-1 \leq \mu \leq 1$. Moreover, expansions in the Legendre polynomials are unique in that the zeroth moment Σ_{s0} preserves the integral properties of the underlying distribution (i.e the mean), whereas the other moments merely provide shape. By looking at the P_N in Figure 9.1, this becomes more obvious when one notes only P_0 has a non-vanishing integral over the domain.

Expanding the Angular Dependence

To derive the P_N equations, we must be careful to choose a starting point consistent with our expansion of the scattering kernel defined by Eqs. 9.2 and 9.3. If we're not, it's easy to encounter a phantom 2π! Take as our starting point the monoenergetic transport equation in slab geometry with arbitary angular dependence of the scattering and source terms,

$$\mu \frac{\partial \psi}{\partial x} + \Sigma_t(x)\psi(x,\mu) = \int_0^{2\pi} d\phi' \int_{-1}^{1} d\mu' \Sigma_s(x,\mu_0)\psi(x,\mu') + S(x,\mu). \qquad (9.8)$$

We expand the source

$$S(x,\mu) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} S_n(x) P_n(\mu)\,, \tag{9.9}$$

where

$$S_n(x) = 2\pi \int_{-1}^{1} S(x,\mu) P_n(\mu) d\mu\,, \tag{9.10}$$

and the angular flux

$$\psi(x,\mu) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} \psi_n(x) P_n(\mu)\,, \tag{9.11}$$

where

$$\psi_n(x) = 2\pi \int_{-1}^{1} \psi(x,\mu) P_n(\mu) d\mu\,. \tag{9.12}$$

Before substituting these expressions, it helps to simplify the scattering expression. Using Eq. 9.2, we can write the scattering term of Eq. 9.8

$$\begin{aligned} &\int_0^{2\pi} d\phi' \int_{-1}^{1} d\mu' \Sigma_s(x,\mu_0)\psi(x,\mu') \\ &\qquad = \int_0^{2\pi} d\phi' \int_{-1}^{1} d\mu' \sum_{l=0}^{\infty} \frac{2l+1}{4\pi} \Sigma_{sl}(x) P_l(\mu_0)\psi(x,\mu')\,. \end{aligned} \tag{9.13}$$

Here, we make use of the *Legendre addition theorem*, which states

$$P_l(\mu_0) = P_l(\mu)P_l(\mu') + 2\sum_{m=1}^{l} \frac{(l-m)!}{(l+m)!} P_l^m(\mu) P_l^m(\mu') \cos(m(\phi-\phi'))\,, \tag{9.14}$$

where $P_l^m(\mu)$ are the *associated Legendre polynomials*, defined

$$P_l^m(\mu) = \sqrt{(1-\mu^2)^m} \frac{d^m P_l}{d\mu^m}\,. \tag{9.15}$$

Substituting this in, we find

$$\begin{aligned} &\int_0^{2\pi} d\phi' \int_{-1}^{1} d\mu' \sum_{l=0}^{\infty} \frac{2l+1}{4\pi} \Sigma_{sl}(x) P_l(\mu_0)\psi(x,\mu') \\ &\quad = \sum_{l=0}^{\infty} \frac{2l+1}{4\pi} \Sigma_{sl}(x) \int_{-1}^{1} d\mu' \psi(x,\mu') \int_0^{2\pi} d\phi' \Bigg(P_l(\mu)P_l(\mu') + \\ &\qquad\qquad 2\sum_{m=1}^{l} \frac{(l-m)!}{(l+m)!} P_l^m(\mu) P_l^m(\mu') \cos(m(\phi-\phi')) \Bigg)\,. \end{aligned} \tag{9.16}$$

Noting that $\int_0^{2\pi} \cos(m(\phi - \phi'))d\phi' = 0$, we find

$$\begin{aligned}
&\int_0^{2\pi} d\phi' \int_{-1}^{1} d\mu' \sum_{l=0}^{\infty} \frac{2l+1}{4\pi} \Sigma_{sl}(x) P_l(\mu_0) \psi(x, \mu') \\
&\quad = \sum_{l=0}^{\infty} \frac{2l+1}{4\pi} \Sigma_{sl}(x) \int_{-1}^{1} d\mu' \psi(x, \mu') \int_0^{2\pi} d\phi' \Big(P_l(\mu) P_l(\mu') \Big) \\
&\quad = \sum_{l=0}^{\infty} \frac{2l+1}{4\pi} \Sigma_{sl}(x) P_l(\mu) \left(2\pi \int_{-1}^{1} d\mu' \psi(x, \mu') P_l(\mu') \right) \\
&\quad = \sum_{l=0}^{\infty} \frac{2l+1}{4\pi} \Sigma_{sl}(x) P_l(\mu) \psi_l(x) \, .
\end{aligned} \tag{9.17}$$

Substituting the simplified scattering term of Eq. 9.17 and the expansions of Eqs. 9.9 and 9.11 into Eq. 9.8 yields

$$\begin{aligned}
&\sum_{n=0}^{\infty} \frac{2n+1}{4\pi} P_n(\mu) \Big(\mu \frac{\partial \psi_n(x)}{\partial x} + \Sigma_t(x) \psi_n(x) \Big) = \\
&\qquad \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} \Sigma_{sn}(x) P_n(\mu) \psi_n(x) + \sum_{n=0}^{\infty} \frac{2n+1}{4\pi} S_n(x) P_n(\mu) \, .
\end{aligned} \tag{9.18}$$

By rearranging the Legendre recurrence relation of Eq. 9.6, we can express the product $\mu P_n(\mu)$ as

$$\mu P_n(\mu) = \frac{1}{2n+1} \Big((n+1) P_{n+1}(\mu) + n P_{n-1}(\mu) \Big) \, . \tag{9.19}$$

Substituting this in for μP_n in the first term on the left of Eq. 9.18 and cancelling a few like terms yields

$$\begin{aligned}
&\sum_{n=0}^{\infty} \left(\Big((n+1) P_{n+1}(\mu) + n P_{n-1}(\mu) \Big) \frac{\partial \psi_n(x)}{\partial x} + (2n+1) P_n(\mu) \Sigma_t(x) \psi_n(x) \right) = \\
&\qquad \sum_{n=0}^{\infty} (2n+1) \Sigma_{sn}(x) P_n(\mu) \psi_n(x) + \sum_{n=0}^{\infty} (2n+1) S_n(x) P_n(\mu) \, .
\end{aligned} \tag{9.20}$$

To exploit the orthogonality property defined in Eq. 9.5, we multiply both sides of

Eq. 9.20 by $\frac{2m+1}{2}P_m$ and integrate the result over $-1 \leq \mu \leq 1$. Then we find

$$\int_{-1}^{1} d\mu \left\{ \sum_{n=0}^{\infty} \frac{2m+1}{2} P_m(\mu) \left(\overbrace{(n+1)P_{n+1}(\mu)\frac{\partial \psi_n}{\partial x}}^{\delta_{n+1,m} \rightarrow n=m-1} + \overbrace{nP_{n-1}(\mu)\frac{\partial \psi_n}{\partial x}}^{\delta_{n-1,m} \rightarrow n=m+1} \right) \right\} = m\frac{\partial \psi_{m-1}}{\partial x} + (m+1)\frac{\partial \psi_{m+1}}{\partial x}\,, \tag{9.21}$$

and

$$\int_{-1}^{1} d\mu \frac{2m+1}{2} P_m(\mu)(2n+1)P_n(\mu)\Sigma_t(x)\psi_n(x) = (2m+1)\Sigma_t(x)\psi_m(x)\,, \tag{9.22}$$

and similarly for the scattering and source terms. Together, we have

$$\frac{m+1}{2m+1}\frac{\partial \psi_{m+1}}{\partial x} + \frac{m}{2m+1}\frac{\partial \psi_{m-1}}{\partial x} + \Sigma_t(x)\psi_m(x) = \Sigma_{sm}(x)\psi_m(x) + S_m(x)\,,\ m = 0 \ldots \infty\,, \tag{9.23}$$

where we take $\psi_{-1} = 0$.

The P_N Equations

Eq. 9.23 is an infinite set of coupled differential equations that represents an exact treatment in angle. To yield a tractable problem, we use only the first $N+1$ equations and set

$$\psi_{N+1} = 0\,. \tag{9.24}$$

As an example, the P_1 approximation yields

$$\begin{aligned} n = 0: &\quad \frac{d\psi_1}{dx} + \Sigma_t(x)\psi_0(x) = \Sigma_{s0}\psi_0 + S_0 \\ n = 1: &\quad \frac{1}{3}\frac{d\psi_0}{dx} + \Sigma_t(x)\psi_1(x) = \Sigma_{s1}\psi_1 + S_1\,, \end{aligned} \tag{9.25}$$

which is a set of two equations in the two unknown Legendre moments ψ_0 and ψ_1. In general, the P_N approximation gives $N+1$ equations in an equal number of unknowns.

An interesting fact is that the P_N approximation is equalivalent to the S_{N+1} approximation in slab geometry if the Gauss-Legendre quadrature set is used, proof of which is left as an exercise for the simple case of the P_1 and S_2 equations. A similar equivalence can me made in certain cases for higher dimensions between the spherical harmonics equations and multidimensional S_N approximations. Doing so is one way to mitigate the numerical artifacts known as ray effects common to the S_N approximation. See the references in the Further Reading.

Diffusion Theory

The P_N equations give us one formal way to derive the diffusion equation. Suppose in Eq. 9.25 that we take the source to be isotropic so that $S_1 = 0$. Moreover, we define the first scattering momement to be $\Sigma_{s1} = \Sigma_{s0}\bar{\mu}$, where $\bar{\mu} = \int_{-1}^{1} d\mu\mu\Sigma_s(\mu) / \int_{-1}^{1} d\mu\Sigma_s(\mu)$.

Then, we rearrange Eq. 9.25 to find

$$\psi_1(x) = \frac{-1}{3(\Sigma_t - \Sigma_{s1}\mu)} \frac{d\psi_0}{dx} \equiv -D(x)\frac{\psi_0}{dx}\,, \tag{9.26}$$

so that

$$-\frac{d}{dx}D(x)\frac{\psi_0}{dx} + (\Sigma_t - \Sigma_{s0})\psi_0(x) = S_0(x)\,. \tag{9.27}$$

This is the diffusion equation, and we recognize $\psi_0(x)$ as the scalar flux $\phi(x)$, $\psi_1(x)$ as the current density $J(x)$, and D as the diffusion coefficient. Note that while diffusion theory explicitly assumes an isotropic source, it does treat linearly anisotropic scattering exactly.

Boundary Condition

In general, there are two approaches to defining boundary conditions for the P_N (or diffusion) equations:

1. conserve net current (Marshak)

2. obtain correct boundary flux for a finite number of μ values (Mark)

The P_N approximation consists of $N+1$ equations, which requires $N+1$ boundary conditions. Typically, it's desirable to spread these conditions evenly among surfaces. For a slab system, there are just two surfaces, and for odd N, there are an even number of conditions to be satisfied. Hence, we typically only use odd N approximations.

Marshak Conditions

The *Marshak* condition places a limit on the odd moments of the P_N expansion, as the odd moments drive net flows in angular space. Suppose we are given an incident boundary condition $B_L(\mu)$ on the left hand side, represented as

$$\psi(x_L, \mu) = B_L(\mu)\,, \;\; \mu > 0\,. \tag{9.28}$$

Then the Marshak condition states

$$\int_0^1 \psi(x_L, \mu) P_l(\mu) d\mu = \int_0^1 B_L(\mu) P_l(\mu) d\mu \,, \tag{9.29}$$

for $l = 1, 3, \ldots, N$. In general, the Marshak conditions give the best results for the P_N equations. Moreover, they capture the physically appealing notion of conservation of particles across the boundary.

Mark Condition

The Mark condition can be expressed as

$$\psi_l(x_L) = \int_{-1}^1 B_L(\mu) \delta(\mu - \mu_n) P_l(\mu) d\mu = B(\mu_n) P_l(\mu_n) \,, \tag{9.30}$$

for a desired set of μ_n. Typically, these μ_n are chosen to be symmetric for an equal number of μ_n per half-space. A common choice are the μ_n such that $P_{N+1}(\mu_n) = 0$, which should look strikingly similar to the condition for generating abscissa for the Gauss-Legendre quadrature set.

A Simple Code

To be continued.

Further Reading

To be continued.

Exercises

1. **$\mathbf{P_1}$ Boundary Conditions**. Derive the P_1 equations for isotropic scattering and an isotropic source. Derive the Marshak vacuum conditions for arbitrary left and right boundaries in a slab. Do the same for the Mark conditions.
2. **$\mathbf{P_2}$ Equations**. Show for the case of linearly anisotropic scattering and isotropic source that the P_2 equations can be written as a second order partial differential equation similar in form to the typical neutron diffusion equation.
3. **$\mathbf{P_3}$ Equations**. Starting from the 1-D transport equation with arbitrary anisotropic scattering and an isotropic source, derive the P_3 equations. Additionally, derive the Marshak vacuum conditions for arbitrary left and right boundaries in a slab.

4. **Numerical Solution of the $\mathbf{P_1}$ Equations**. Consider a slab of width 10 cm with $\Sigma_t = 1.0$ [cm^{-1}], and $c = \Sigma_s/\Sigma_t = 0.5$ (isotropic scattering in the lab system). A uniform, isotropic source of 1 [n/cm^2-s] is located in the first half of the slab, and both slab edges are subject to vacuum conditions. Write a code to solve the P_1 approximation to this problem using Marshak conditions. Use a finite-difference approximation in space. Plot $\psi(x, \mu)$ at $x = 0$, 2.5, 5.0, 7.5, and 10 [cm], and plot $\phi(x)$ over the whole slab.
5. **Numerical Solution of the $\mathbf{P_3}$ Equations**. For the same problem, write a code to solve the P_3 approximation using Marshak conditions.
6. **Diffusion via asymptotics**. Consider the 1-D transport equation in slab geometry. Suppose the angular flux varies significantly over ranges approximately equal to $l \gg \lambda = 1/\Sigma_t$. By letting $\epsilon = \lambda/l$, $z = x/l$, $\tilde{\Sigma}_t = \lambda\Sigma_t$, $\tilde{\Sigma}_s = \lambda\Sigma_s$, $\epsilon^2\tilde{\Sigma}_a = \tilde{\Sigma}_t - \tilde{\Sigma}_s$, and $\epsilon^2\tilde{S} = \lambda S$, we can define a scaled transport equation

$$\varepsilon\mu\frac{\partial\psi}{\partial z} + \tilde{\Sigma}_t\psi(z,\mu) = \frac{\tilde{\Sigma}_t - \tilde{\Sigma}_a\varepsilon^2}{2}\int_{-1}^{1}\psi(z,\mu')d\mu' + \frac{\varepsilon^2\tilde{S}}{2}\,, \tag{9.31}$$

   where, for simplicity, a homogeneous medium is assumed. Then let

$$\psi(x,\mu) = \sum_{n=0}^{\infty}\varepsilon^n\psi_n(x,\mu), \tag{9.32}$$

   which is a *perturbation series*. Then do the following:

   (a) Substitute this approximation into the scaled transport equation and find equations corresponding to ε^0, ε^1, and ε^2.

   (b) From the first two of these equations, show that $\psi_1 = \phi_0/2 - (\mu/2\Sigma_t)d\phi_0/dx$, where $\phi_n(x) = \int_{-1}^{1}\psi_n(x,\mu)$.

   (c) Show that the third equation leads directly to the diffusion equation once the variables and cross sections are converted back to their original forms.

   This problem provides a very quick taste of asymptotic analysis, a tool that Ed Larsen has used successfully for various problems of interest in neutron transport; see, e.g., Ref. [27].
7. **Legendre Addition Theorem**. Prove the Legendre polynomial addition theorem. You may use any resource you want, but make sure you understand all steps of the proof. A particularly straightforward approach begins as follows. Start with an expansion of an arbitrary function in the full spherical harmonics. Then, substitute the definition of the expansion coefficients back into the expansion. Noting that the integral and summation can be switched, what function must the summation be?

8. **Defining the Scattering Angle**. Prove $\mu_0 = \cos(\theta)\cos(\theta') + \sin(\theta)\sin(\theta')\cos(\phi - \phi')$, where μ_0 is the cosine of the scattering angle and (θ,ϕ) and (θ', ϕ') are the original and final angles, respectively. Include a diagram.

Lecture 10 | Numerical Diffusion

In Lecture 9, the diffusion approximation was formally derived. In this lecture, we present an approach for solving the diffusion equations numerically. Diffusion theory represents an important method by itself. Moreover, it finds widespread use in the acceleration of higher-order transport methods, including the nonlinear diffusion acceleration of Lecture 12 and the linear preconditioners of Lecture 14.

Mesh-Centered Diffusion Approximation

The diffusion approximation can also be coupled with the multigroup treatment in multiple spatial dimensions. Here, we limit the discussion to the multidimensional, one-group diffusion equation. A full multigroup treatment is left as an exercise.

Consider the one-group diffusion equation

$$-\nabla D(\vec{r})\nabla\phi(\vec{r}) + \Sigma_r(\vec{r})\phi(\vec{r}) = S(\vec{r})\,, \tag{10.1}$$

where Σ_r is the "removal" cross section. In the one-group approximation, $\Sigma_r = \Sigma_a$. To solve Eq. (10.1) numerically, we use a mesh-centered, finite-volume approach, in which cell materials and sources are taken to be constant [21]. By integrating Eq. (10.1) over the volume V_{ijk}, one obtains

$$\begin{aligned}\int_{x_{i-1/2}}^{x_{i+1/2}} dx \int_{y_{j-1/2}}^{y_{j+1/2}} dy \int_{z_{k-1/2}}^{z_{k+1/2}} dz \Big\{ -\nabla D(\vec{r})\nabla\phi(\vec{r}) + \Sigma_r(x,y,z)\phi(x,y,z) \Big\} \\ = \int_{x_{i-1/2}}^{x_{i+1/2}} dx \int_{y_{j-1/2}}^{y_{j+1/2}} dy \int_{z_{k-1/2}}^{z_{k+1/2}} dz\, S(x,y,z)\,,\end{aligned} \tag{10.2}$$

or

$$
\begin{aligned}
-D_{ijk}\Bigg[&\Delta_{y_j}\Delta_{z_k}\Big(\phi_x(x_{i+1/2},y_j,z_k)-\phi_x(x_{i-1/2},y_j,z_k)\Big)\\
&+\Delta_{x_i}\Delta_{z_k}\Big(\phi_y(x_i,y_{j+1/2},z_k)-\phi_y(x_i,y_{j-1/2},z_k)\Big)\\
&+\Delta_{x_i}\Delta_{y_j}\Big(\phi_z(x_i,y_j,z_{k+1/2})-\phi_z(x_i,y_j,z_{k-1/2})\Big)\Bigg]\\
&+\Delta_{x_i}\Delta_{y_j}\Delta_{z_k}\Sigma_{r,ijk}\phi_{ijk}=\Delta_{x_i}\Delta_{y_j}\Delta_{z_k}S_{ijz}\,,
\end{aligned}
\tag{10.3}
$$

where the cell-centered flux (equal to the average flux) is defined as

$$
\phi_{ijk}\equiv\frac{1}{\Delta_{x_i}}\frac{1}{\Delta_{y_j}}\frac{1}{\Delta_{z_k}}\int_{x_{i-1/2}}^{x_{i+1/2}}dx\int_{y_{j-1/2}}^{y_{j+1/2}}dy\int_{z_{k-1/2}}^{z_{k+1/2}}dz\,\phi(x,y,z)\,, \tag{10.4}
$$

the cell-average source is

$$
S_{ijk}\equiv\frac{1}{\Delta_{x_i}}\frac{1}{\Delta_{y_j}}\frac{1}{\Delta_{z_k}}\int_{x_{i-1/2}}^{x_{i+1/2}}dx\int_{y_{j-1/2}}^{y_{j+1/2}}dy\int_{z_{k-1/2}}^{z_{k+1/2}}dz\,S(x,y,z)\,, \tag{10.5}
$$

and, for example, $\phi_x(x_{i+1/2},y_j,z_k)$ is the derivative of ϕ with respect to x, averaged over y and z, and evaluated at $x=x_{i+1/2}$.

To evaluate the partial derivatives ϕ_x, ϕ_y, and ϕ_z in Eq. (10.3), Taylor-series expansions are employed. For the x-directed terms at the left (or west) boundary, let

$$
\phi(x_{i-1},y_j,z_k)\approx\phi(x_{i-1/2},y_j,z_k)-\frac{\Delta x_{i-1}}{2}\phi_x^+(x_{i-1/2},y_j,z_k) \tag{10.6}
$$

and

$$
\phi(x_i,y_j,z_k)\approx\phi(x_{i-1/2},y_j,z_k)+\frac{\Delta x_i}{2}\phi_x^-(x_{i-1/2},y_j,z_k)\,, \tag{10.7}
$$

with similar expressions for the right x boundary and the y- and z-directed terms. The $+$ and $-$ superscripts on the partial derivatives indicate forward and backward extrapolation from the cell midpoint, respectively. By continuity of net current, we must have

$$
D_{i-1,jk}\phi_x^+(x_{i-1/2},y_j,z_k)=D_{ijz}\phi_x^-(x_{i-1/2},y_j,z_k)\,. \tag{10.8}
$$

Multiplication of Eq. (10.6) by $D_{i-1,jk}/\Delta x_{i-1}$ and Eq. (10.7) by $D_{ijk}/\Delta x_i$, adding the results, and rearranging leads to

$$
\phi_{i-1/2,jk}=\frac{D_{i-1,jk}\phi_{i-1,jk}\Delta_{x_i}+D_{ijk}\phi_{ijk}\Delta x_{i-1}}{D_{i-1,jk}\Delta_{x_i}+D_{ijk}\Delta x_{i-1}}\,. \tag{10.9}
$$

Substituting this into the Eq. (10.7) gives

$$\begin{aligned}\phi_x(x_{i-1/2},y_j,z_k) &= \frac{2}{\Delta_{x_i}}\left(\phi_{ijk} - \frac{D_{i-1,jk}\phi_{i-1,jk}\Delta_{x_i} + D_{ijk}\phi_{ijk}\Delta x_{i-1}}{D_{i-1,jk}\Delta_{x_i} + D_{ijk}\Delta x_{i-1}}\right) \\ &= \frac{2}{\Delta_{x_i}}\left(\frac{D_{i-1,jk}\phi_{ijk}\Delta_{x_i} + D_{ijk}\phi_{i-1,j,k}\Delta x_{i-1}}{D_{i-1,jk}\Delta_{x_i} + D_{ijk}\Delta x_{i-1}}\right)\end{aligned} \tag{10.10}$$

or

$$\phi_x(x_{i-1/2},y_j,z_k) = 2D_{i-1,jk}\left(\frac{\phi_{ijk}-\phi_{i-1,jk}}{\Delta_{x_i}D_{i-1,jk} + \Delta x_{i-1}D_{ijk}}\right), \tag{10.11}$$

Similarly, one finds

$$\phi_x(x_{i+1/2},y_j,z_k) = 2D_{i+1,jk}\left(\frac{\phi_{i+1,jk}-\phi_{ijk}}{\Delta_{x_i}D_{i+1,jk} + \Delta x_{i+1}D_{ijk}}\right). \tag{10.12}$$

These equations and the equivalents for y and z are substituted into Eq. (10.3) to obtain a set of *internal equations*.

Internal Equation

Substitution of Eqs. (10.11) and (10.12) into Eq. (10.3) leads to

$$\begin{aligned}-D_{ijk}\Bigg\{\Delta_{y_j}\Delta z_j\Bigg[&2D_{i+1,jk}\left(\frac{\phi_{i+1,j}-\phi_{i,j}}{\Delta_{x_i}D_{i+1,jk}+\Delta x_{i+1}D_{ijk}}\right)+ \\ &2D_{i-1,jk}\left(\frac{\phi_{i-1,j}-\phi_{i,j}}{\Delta_{x_i}D_{i-1,jk}+\Delta x_{i-1}D_{ijk}}\right)\Bigg]\ldots+\Bigg\}+ \\ &\Delta_{x_i}\Delta_{y_j}\Delta_{z_k}\Sigma_{r,ijk}\phi_{ijk} = \Delta_{x_i}\Delta_{y_j}\Delta_{z_k}S_{ijk}\,.\end{aligned} \tag{10.13}$$

Equation (10.13) represents the balance of neutrons in the cell (i,j,k).

We can further simplify the notation. Let us define a coupling coefficient

$$\tilde{D}_{i+1/2,jk} \equiv \frac{2D_{i+1,jk}D_{ijk}}{\Delta_{x_i}D_{i+1,jk} + \Delta_{x_{i+1}}D_{ijk}}\,, \tag{10.14}$$

with similar coefficients for each direction. Then we can rewrite Eq. 10.13 as

$$\begin{aligned}
&\frac{\tilde{D}_{i+1/2,j,k}}{\Delta_{x_i}}\Big(\phi_{ijk}-\phi_{i+1,j,k}\Big)+\frac{\tilde{D}_{i-1/2,j,k}}{\Delta_{x_i}}\Big(\phi_{ijk}-\phi_{i-1,j,k}\Big)\\
+&\frac{\tilde{D}_{i,j+1/2,k}}{\Delta_{y_j}}\Big(\phi_{ijk}-\phi_{i,j+1,k}\Big)+\frac{\tilde{D}_{i,j-1/2,k}}{\Delta_{y_j}}\Big(\phi_{ijk}-\phi_{i,j-1,k}\Big)\\
+&\frac{\tilde{D}_{i,j,k+1/2}}{\Delta_{z_k}}\Big(\phi_{ijk}-\phi_{i,j,k+1}\Big)+\frac{\tilde{D}_{i,j,k-1/2}}{\Delta_{z_k}}\Big(\phi_{ijk}-\phi_{i,j,k-1}\Big)\\
&\qquad\qquad+\Sigma_{r,ijk}\phi_{ijk}=S_{ijk}\,.
\end{aligned}\tag{10.15}$$

Note that each term on the left with a coupling coefficient represents the net leakage from a surface divided by the area of that surface.

Boundary Equations

At the boundaries, we employ the albedo condition [21]

$$\frac{1}{2}D(x,y,z)\nabla\phi(x,y,z)\cdot\hat{n}(x,y,z)+\frac{1}{4}\frac{1-\alpha(x,y,z)}{1+\alpha(x,y,z)}\phi(x,y,z)=0\,,\tag{10.16}$$

where $\hat{n}$ is the outward normal, α describes the albedo condition ($\alpha=0$ for vacuum, and $\alpha=1$ for reflection).

In three dimensions, a mesh cell has six surfaces, some of which may be part of a global surface. As an example, we consider the west global boundary:

$$\text{west boundary}:\quad -\frac{1}{2}D_{1jk}\phi_x(x_{1/2},y_j,z_k)+\frac{1}{4}\frac{1-\alpha}{1+\alpha}\phi_{1/2,jk}=0\,.\tag{10.17}$$

Because this expression contains ϕ at the edge, we again need to employ Taylor expansions. For the west boundary, note that

$$\phi_{1jk}\approx\phi_{1/2,jk}+\frac{\Delta_{x_i}}{2}\phi_x(x_{1/2},y_j,z_k)\,,\tag{10.18}$$

and we rearrange to get

$$\phi_{1/2,jk}=\phi_{1jk}-\frac{\Delta_{x_i}}{2}\phi_x(x_{1/2},y_j,z_k)\,.\tag{10.19}$$

Placing this into the albedo condition yields

$$0=-\frac{1}{2}D_{1jk}\phi_x(x_{1/2},y_j,z_k)+\frac{1}{4}\frac{1-\alpha}{1+\alpha}\left(\phi_{1jk}-\frac{\Delta_{x_i}}{2}\phi_x(x_{1/2},y_j,z_k)\right),\tag{10.20}$$

and solving for $\phi_x(x_{1/2}, y_j, z_k)$ gives

$$\phi_x(x_{1/2}, y_j, z_k) = \frac{2(1-\alpha)\phi_{1jk}}{4(1+\alpha)D_{1jk} + \Delta_{x_1}(1-\alpha)}\,. \tag{10.21}$$

For the east, we similarly find

$$\phi_x(x_{I+1/2}, y_j, z_k) = -\frac{2(1-\alpha)\phi_{Ijk}}{4(1+\alpha)D_{Ijk} + \Delta_{x_I}(1-\alpha)}\,, \tag{10.22}$$

and likewise for the other surfaces. Each of these is placed into the proper partial derivative of Eq. (10.3). For example, the leakage contribution on the left hand side of Eq. (10.15) due to leakage from the west boundary is transformed as follows*:

$$\frac{\tilde{D}_{1/2,jk}}{\Delta_{x_1}}\Big(\phi_{0jk} - \phi_{1jk}\Big) \rightarrow \frac{2D_{1jk}(1-\alpha)\phi_{1jk}}{(4(1+\alpha)D_{1jk} + \Delta_{x_1}(1-\alpha))\Delta_{x_1}}\,. \tag{10.23}$$

Constructing the Diffusion Matrix

The equations for cell balance derived heretofore can be combined into a linear system for the cell fluxes. Generically, we can express this system as

$$\mathbf{L}\vec{\phi} = \vec{S}\,, \tag{10.24}$$

where $\mathbf{L}$ is the diffusion loss matrix. A straightforward implementation of the one-group diffusion loss matrix is presented in Listing 10.1. Note that several parameters and helper functions require definition. Although the code is set up for three dimensions, it can also be used for 1-D and 2-D problems by proper selection of boundary conditions.

Listing 10.1: Sample code for construction of one-group loss matrix.

```
L   = zeros((N, N))             # N x N loss matrix
remdims = [[1,2], [0,2], [0,1]] # For a given dim, provides remaining dims
# Loop over all cells
for n in range(0, N) :
  ijk = cell_to_ijk(n)          # Convert cardinal index to (i, j, k)
  w = [dx[ijk[0]], dy[ijk[1]], dz[ijk[2]]]
  # Loop over dimensions
  for dim in range(0, 3) :
    bound = [0, Nxyz[dim]]      # Minimum and maximum cell index in this dim
    # Loop over directions, + and -
    for dir in range(0, 2) :
      if ijk[dim] == bound[dir] :
        # Global boundary
        side = 3*dim + dir      # index of this side (0-5)
        B = beta[side]          # albedo for this side
        L[n][n] = L[n][n] + 2*D[n]*(1-B)/((4*(1+B)*D[n]+w[dir]*(1-B))*w[dir])
```

*Ignore the fact that ϕ_{0jk} doesn't exist!

```
    else :
      # Coupling to neighbor at cell m
      ijk_m = ijk
      ijk_m[dim] = ijk_m[dim] + (-1)**dir
      m = cell_from_ijk(ijk_m)
      w_m = [dx[ijk_m[0]], dy[ijk_m[1]], dz[ijk_m[2]]]
      D_tilde = 2*D[n]*D[m]/((w[dir]*D[m]+w_m[dir]*D[n])*w[dir])
      L[n][m] = -D_tilde / w[dir]
      L[n][n] = L[n][n] + D_tilde / w[dir]
# Add the removal cross-section component
L[n][n] = L[n][n] + Sigma_r[n]
```

Exercises

1. **Boundary Source**. Suppose a partial current $J_{jk}^{\text{west}-}$ were incident on the west boundary of cell $(1, j, k)$. Show that the corresponding current balance contribution (i.e., Eq. (10.23)) is unchanged but that the following source term arises:

$$Q_{1jk}^{\text{boundary}} = \frac{8D_{1jk}(1+\alpha)J_{jk}^{\text{west}-}}{(4(1+\alpha)D_{1jk} + \Delta_{x_1}(1-\alpha))\Delta_{x_1}} \tag{10.25}$$

2. **One-Group Diffusion**. For $\Sigma_t = 1.0$ and $\Sigma_s = 0.0$, compute the flux.
3. **Multigroup Diffusion**. Although a formal derivation of the multigroup diffusion approximation requires assumptions beyond the ones made in Lecture 9, the resulting set of multigroup diffusion equations is intuitive:

$$\begin{aligned} -\nabla D_g(\vec{r})\nabla\phi_g(\vec{r}) + \Sigma_{rg}(\vec{r})\psi_g(\vec{r}) &= \sum_{\substack{g'=1 \\ g'\neq g}}^{G} \Sigma_{sg\leftarrow g'}(\vec{r})\phi_{g'}(\vec{r}) \\ &+ \frac{\chi_g(\vec{r})}{k}\sum_{g'=1}^{G}\nu\Sigma_{fg'}(\vec{r})\phi_{g'}(\vec{r}) + S_g(\vec{r})\,. \end{aligned} \tag{10.26}$$

   Extend the routine proposed in Listing 10.1 to handle multigroup problems.

Lecture 11 | Discrete Ordinates Method

In the last two lectures, we dealt exclusively with the integral form of the transport equation and ways to solve it numerically. Now, we return to the integro-differential form of the equation and introduce the *discrete ordinates (S_N) method*, one of the most widely-used deterministic transport methods. We subsequently introduce the method of *source iteration* for treating the scattering source, for which two nonlinear acceleration schemes are presented in the next lecture.

A Discrete Angle Domain

The integral methods of the last two lectures are useful in many applications because they effectively eliminate the angular variable. If we are to treat the transport equation in its integro-differential form, we need a way to handle the angles. For the sake of brevity, we limit our discussion to transport in slab geometry with isotropic sources and scattering, for which the transport equation reduces to

$$\begin{aligned} \mu\frac{\partial\psi}{\partial x} + \Sigma_t(x)\psi(x,\mu) &= \frac{\Sigma_s(x)}{2}\int_{-1}^{1} d\mu'\psi(x,\mu') + \frac{S(x)}{2} \\ &= \frac{\Sigma_s(x)}{2}\phi(x) + \frac{S(x)}{2} \end{aligned} \tag{11.1}$$

The discrete ordinates method consists of requiring Eq. 11.1 to hold only for discrete values μ. In other words, we require

$$\mu_n\frac{\partial\psi}{\partial x} + \Sigma_t(x)\psi(x,\mu_n) = \frac{\Sigma_s(x)}{2}\phi(x) + \frac{S(x)}{2}\,. \tag{11.2}$$

What about ϕ? This is where the approximation is actually made. Because we have ψ only at discrete angles μ_n, we cannot perform the continuous integral. We do the next best thing and compute a weighted sum:

$$\phi(x) \approx \sum_{n=1}^{N} w_n\psi_n(x)\,, \tag{11.3}$$

where we have introduced the notation $\psi_n(x) \equiv \psi(x, \mu_n)$.

The only actual requirement on w_n is that they satisfy

$$\sum_{n=1}^{N} w_n = 2\,, \tag{11.4}$$

since this sum represents $\int_{-1}^{1} d\mu = 2$. Of course, we don't choose the weights arbitrarily, nor do we choose the cosines μ_n arbitrarily. The choice of both is the art of numerical quadrature.

Intermission: Numerical Quadrature (a.k.a. Numerical Integration)

You saw application of the *trapezoid rule* in the integral transport code of Lecture 6. You have probably seen the trapezoid rule in previous classes, and if not that, then certainly the even simpler *Riemann sum* approximation to integrals, here in the form of the *midpoint rule*:

$$\int_a^b f(x)dx \approx \Delta \sum_{n=0}^{N-1} f\Big(a + (n + 1/2)\Delta\Big)\,, \tag{11.5}$$

where

$$\Delta = \frac{b-a}{N}\,, \tag{11.6}$$

and N is the number of divisions in the center of which f is evaluated.

The general form of a quadrature rule is

$$\int_a^b f(x)dx \approx \sum_{n=0}^{N-1} w_n f(x_n)\,, \tag{11.7}$$

and we see that the midpoint rule chooses equal weights $w_n = w = \Delta$ and evaluates f at equally spaced points $x_n = a + (n + 1/2)\Delta$. The trapezoid rule also takes evenly spaced points, $x_n = a + n\Delta$, for $n = 0 \ldots N$, and with $w_0 = w_N = 0.5\Delta$, and $w_n = \Delta$, $0 < n < N$.

Other quadrature sets using evenly spaced points also exist. One general class consists of the *Newton-Cotes formulas*, which interpolate $f(x)$ using various order polynomials and thus allows for analytic integration. The midpoint rule and the trapezoid rule are the simplest examples, using zeroth order and first order polynomials. Another example is *Simpson's rule*, defined

$$\begin{aligned}\int_a^b f(x)dx \approx &\frac{\Delta}{3}\Big(f(a) + 4f(a + \Delta) + 2f(a + 2\Delta) + 4f(a + 3\Delta)\\ &+ 2f(a + 4\Delta) + \ldots + 4f(b - \Delta) + f(b)\Big)\,.\end{aligned} \tag{11.8}$$

Simpson's rule is based on a piece-wise quadratic interpolation of sets of three points and can often be a very good "quick and dirty" way to integrate numerically. The reader is encouraged to consult a numerical analysis text for more details.

In many cases, it is possible to use fewer, non-equally spaced points (with unique weights) to yield accurate integral approximations that would otherwise take many equally spaced points. One such scheme is Gauss-Legendre quadrature, which is the standard quadrature used to define ϕ in Eq. 11.3. The n-point Gauss-Legendre scheme is constructed so that polynomials of degree less than or equal to $2n-1$ are integrated exactly over the domain of interest.

Consider a 2-point Gauss-Legendre scheme, i.e. $\int_{-1}^{1} f(x)dx = w_1 f(x_1) + w_2 f(x_2)$. We have four unknowns, so we need four equations. The simplest case is to integrate the monomials 1, x, x^2, and x^3 over the desired range. Doing so, we get the set of equations

$$\begin{aligned}
\int_{-1}^{1} 1dx &= 2 = w_1 + w_2 \\
\int_{-1}^{1} x dx &= 0 = w_1 x_1 + w_2 x_2 \\
\int_{-1}^{1} x^2 dx &= \frac{2}{3} = w_1 x_1^2 + w_2 x_2^2 \\
\int_{-1}^{1} x^3 dx &= 0 = w_1 x_1^3 + w_2 x_2^3 \,.
\end{aligned} \tag{11.9}$$

This system is solved for $w_1 = w_2 = 1$, $x_1 = -1/\sqrt{3}$, and $x_2 = 1/\sqrt{3}$. Interestingly, x_1 and x_2 are the roots of the second-order Legendre polynomial $P_2(x) = x^2 - 1/3$. In fact, the x_i for any n-point Gauss-Legendre scheme are the n roots of the nth order Legendre polynomial—hence the name of the quadrature rule*. The weights are defined

$$w_i = \frac{2(1-x_i^2)}{(n+1)^2 (P_{n+1}(x_i))^2} \,. \tag{11.10}$$

As a reference, the Gauss-Legendre abscissa and weights are given in Table 11.1 for $N = 2$, 4, and 6. Note, odd orders are not used because the abscissa would then include $x = 0$. Since our goal is ultimately to integrate over μ, and since we shall encounter terms with $1/\mu$, we simply avoid the singularity by choosing even N.

*It should be noted that an entire class of Gaussian quadratures exists based on use of other sets of polynomials to approximate integrals where f is weighted by some weighting function associated with the polynomials. An important example is the Gauss-Chebyshev quadrature, where we approximate the integral $\int_{-1}^{1} \frac{f(x)}{\sqrt{1-x^2}} dx$. The x_i are found to be roots of the Chebyshev polynomials.

Table 11.1: Gauss-Legendre quadrature parameters.

N	μ_i	w_i
2	± 0.5773502691	1
4	± 0.8611363115	0.3478548451
	± 0.3399810435	0.6521451549
6	± 0.9342695142	0.1713244924
	± 0.6612093864	0.3607615730
	± 0.2386191860	0.4679139346

A Discrete Spatial Domain

We next deal with the spatial variable, using the finite volume method. Consider the discretization of Figure 11.1. The slab is discretized into I regions, each of which is assumed to have constant cross-sections and sources, denoted Σ_{ti} and so forth. The cell centers are given by x_i, and the cell edges by $x_{i\pm 1/2}$. For a slab defined over $0 \leq x \leq L$, $x_{1/2} = 0$ and $x_{I+1/2} = L$.

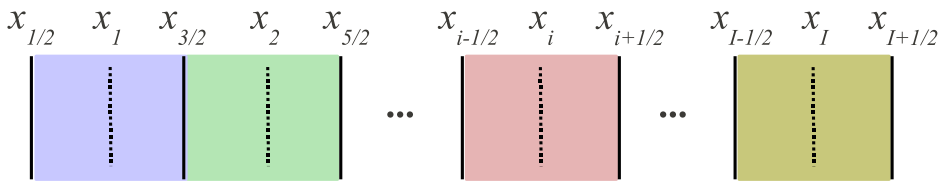

Figure 11.1: Discrete ordinates spatial discretization.

Now, we integrate Eq. 11.2 over cell i:

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} dx \Big(\mu_n \frac{\partial \psi_n}{\partial x} + \Sigma_t(x)\psi_n(x) \Big) = \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} dx \Big(\frac{\Sigma_s(x)}{2}\phi(x) + \frac{S(x)}{2} \Big) , \quad (11.11)$$

so that

$$\mu_n(\psi_{i+\frac{1}{2},n} - \psi_{i-\frac{1}{2},n}) + \Delta_i \Sigma_{ti}\psi_{i,n} = \frac{\Delta_i}{2}\Big(\Sigma_{is}\phi_i + S_i \Big) , \quad (11.12)$$

where $\Delta_i \equiv x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}$ and where the integrals have been approximated via the midpoint rule where needed. For the sake of brevity, we switch to the emission density, defining

$$Q_i = (\Sigma_{is}\phi_i + S_i)/2 , \quad (11.13)$$

so that

$$\mu_n(\psi_{i+\frac{1}{2},n} - \psi_{i-\frac{1}{2},n}) + \Delta_i \Sigma_{ti} \psi_{i,n} = \Delta_i Q_i \,. \tag{11.14}$$

Eq. 11.14 is a good first step, but we're not done. Imagine we know the angular flux entering the left side of slab i. We would like to be able to compute the angular flux leaving i to the right. Doing exactly this over the entire domain is referred to as "sweeping" and is a common technique for hyperbolic problems, as it allows us to propagate boundary conditions along the direction of neutron paths. The problem with Eq. 11.14 is that we have another unknown: $\psi_{i,n}$, the cell-centered angular flux. There are several approximations we can use, and we develop two: the diamond difference (DD) and step difference (SD) methods.

For both DD and SD, we define the cell-centered angular flux by

$$\psi_{i,n} = \frac{1+\alpha_n}{2}\psi_{i+\frac{1}{2},n} + \frac{1-\alpha_n}{2}\psi_{i-\frac{1}{2},n} \,, \tag{11.15}$$

where

$$\alpha_n = \begin{cases} 0 & \text{for DD} \\ |\mu_n|/\mu_n & \text{for SD} \end{cases} . \tag{11.16}$$

The DD method is probably the most ubiquitous of the simple schemes available. It essentially says the angular flux varies linearly within a cell, and it provides second-order accuracy. Unfortunately, the DD method can lead to negative fluxes for Δ_i too large, demonstration of which is left as an exercise. On the other hand, the SD method cannot yield negative fluxes, but is only first-order accurate. For the SD method, note that the cell-center angular flux is equal to the outgoing angular flux at the edge.

Sweeping Formulas

We hinted above that we can solve the discrete ordinates equations via use of sweeps across the domain. In slab geometry, this consists of sweeps to the right for $\mu_n > 0$ and sweeps left for $\mu_n < 0$. If we have vacuum or some other known boundary flux, we start with that condition, as reflective conditions require information from a sweep having come from the opposite direction. If all conditions are reflective, a guess must be made for one, which will likely increase the number of iterations required (we'll get to iterations below).

Substituting Eq. 11.15 into Eq. 11.14 and rearranging, we find for sweeping right

$$\psi_{i+\frac{1}{2},n} = A^+_{i,n}\psi_{i-\frac{1}{2},n} + B^+_{i,n}Q_i \,, \tag{11.17}$$

and sweeping left

$$\psi_{i-\frac{1}{2},n} = A^-_{i,n}\psi_{i+\frac{1}{2},n} + B^-_{i,n}Q_i \,, \tag{11.18}$$

where

$$A_{i,n}^{\pm} = \frac{2\mu_n \mp (1 \mp \alpha_m)\Sigma_{ti}\Delta_i}{2\mu_n \pm (1 \pm \alpha_m)\Sigma_{ti}\Delta_i}\,, \tag{11.19}$$

and

$$B_{i,n}^{\pm} = \frac{\pm 2\Delta_i}{2\mu_n \pm (1 \pm \alpha_m)\Sigma_{ti}\Delta_i}\,. \tag{11.20}$$

Source Iteration

When we introduce scattering into the emission density, the right hand side becomes dependent on the solution. A common way around this is the so-called source iteration method. The source iteration method is related to the Neumann iteration scheme discussed in Lecture 6, and a connection between the two is left as an exercise.

Source iteration is characterized by a right hand side that lags behind the left hand side. More explicitly, we solve the equation

$$\mu_n(\psi_{i+\frac{1}{2},n}^{m} - \psi_{i-\frac{1}{2},n}^{m}) + \Delta_i\Sigma_{ti}\psi_{i,n}^{m} = \Delta_i Q_i^{m-1}\,, \tag{11.21}$$

where m is the iteration index and

$$Q^{m-1} = \frac{1}{2}(\Sigma_{is}\phi_i^{m-1} + S_i)\,. \tag{11.22}$$

For $m = 1$, we need q^0. Often, ϕ^0 is simply set to zero, unless a better initial guess is easily made.

The iterations continue until the updated fluxes change insignificantly. This is usually quantified in a relative sense for the scalar flux by

$$\frac{\max|\phi^m - \phi^{m-1}|}{\phi^{m-1}} < \epsilon_\phi \tag{11.23}$$

for some small ϵ_ϕ.

Acceleration

A Simple Code

Here, we provide a simple discrete ordinates code for computing the angular and scalar flux in a slab, given in Listing 11.1. The input requires region (coarse mesh) edge locations, within which the number of fine meshes are defined. Each region is assigned a material index and volumetric source strength. The total and scattering cross-sections represent values for different materials; each material can be placed in

a region using the appropriate index. This quick input format allows one to investigate fairly easily a wide range of slab compositions.

Currently, the code is limited to the DD and S_4 approximations and vacuum boundaries. Convergence is tested for ϕ only. The code returns the mesh-centered scalar flux and mesh-edge angular flux. Within the solver, the notation follows the lecture notation relatively closely.

Listing 11.1: Solution of Slab Problem via Sn

```
function [phi,psi] = sn(EDGE,NFM,SigT,SigS,RegMat,Source)
% function [phi,psi] = sn(EDGE,NFM,Sigma,RegMat,Source)
%   This function solves the one-speed transport equation in slab geometry
%   using the discrete ordinates method.  The S4 and diamond difference
%   approximations are implemented, and only vacuum boundaries are treated.
%
%   Inputs:
%       EDGE    -- region edges
%       NFM     -- number of fine meshes per region
%       SigT    -- total cross-section for each material
%       SigS    -- scattering cross-section for each material
%       RegMat  -- which material goes in each region
%       Source  -- region uniform isotropic source [1/cm^3-s]
%   Output
%       phi     -- fine mesh-centered scalar flux
%       psi     -- fine mesh-edge angular flux

% assign angles and weights
mu = [ -0.8611363115 -0.3399810435  0.3399810435  0.8611363115 ];
wt = [  0.3478548451  0.6521451549  0.6521451549  0.3478548451 ];
% allocate
totNFM = sum(NFM);            % total number of fine meshes
psi    = zeros(totNFM+1,4);   % current angular flux
phi    = zeros(totNFM,1);     % current scalar flux
S      = zeros(totNFM,4);     % fine mesh source
Q      = S;                   % emission density
fmmid  = zeros(totNFM,1);     % fine mesh material id
% compute discretization
j = 0;
for i = 1:length(NFM)
    Delta( (j+1):(j+NFM(i)) )    = ( EDGE(i+1) - EDGE(i) ) / NFM(i);
    S( (j+1):(j+NFM(i)), : )     = Source(i)/2;
    fmmid( (j+1):(j+NFM(i))   )  = RegMat(i);
    j = sum(NFM(1:i));
end
% precompute coefficients, following Eqs. 20.16-20.20.
alpha = zeros(4,1);
A     = zeros(totNFM,4);
B     = A;
for i = 1:totNFM
    m = fmmid(i);
    for n = 1:4  % use the sign of mu to choose appropriate form
        smu    = sign(mu(n));
        denom  = 2*mu(n)+smu*(1+smu*alpha(n))*SigT(m)*Delta(i);
        A(i,n) = (2*mu(n)-smu*(1-smu*alpha(n))*SigT(m)*Delta(i)) / denom;
        B(i,n) = smu * 2 * Delta(i) / denom;
    end
end
% convergence parameters
eps_phi = 1e-5; max_it  = 200;
err_phi = 1;    it = 0;
% Begin source iterations
while (err_phi > eps_phi && it <= max_it )
    % Save old scalar flux
```

```
    phi0 = phi;
    % Update source
    for i = 1:totNFM
        Q(i,:) = S(i,:) + 0.5*SigS(fmmid(i))*phi(i);
    end
    % Perform sweeps
    for i = totNFM:-1:1  % right-to-left
        psi(i,1:2) = A(i,1:2).*psi(i+1,1:2) + B(i,1:2).*Q(i,1:2);
    end
    for i = 1:totNFM     % left-to-right
        psi(i+1,3:4) = A(i,3:4).*psi(i,3:4) + B(i,3:4).*Q(i,3:4);
    end
    % --- Insert Acceleration Here ---
    % Update phi (need cell-centered psi, so use Eq. 20.15)
    for i = 1:totNFM
         phi(i) = sum( wt(:) .* (0.5*( (1+alpha(:)).*psi(i+1,:)' + ...
                                       (1-alpha(:)).*psi(i,:)'    ) ));
    end
    % Update error and iteration counter
    err_phi =  max(  abs(phi-phi0)./phi0 );
    it = it + 1;
end
if (it <= max_it)
    disp(['Converged in ',num2str(it), ' iterations.'])
else
    disp('Failed to converge.')
end

end
```

Further Reading

A good introduction to the discrete ordinates method in one dimension can be found in Chapter 3 in Lewis and Miller [30]; higher dimensions are covered in Chapter 4 of the same work. The foundation of the method is usually credited to Chandrasekhar (in the context of stellar radiation) and can be found in his monograph [13].

Recent advances in the discrete ordinates method include employing advanced Krylov subspace methods, outlined by Warsa, Wareing, and Morel [45], variations of which are implemented in the state-of-the-art code Denovo [20]. Moreover, the discrete ordinates equations can be parallelized; the most popular approach is that of Koch, Baker, and Alcouffe [26] (also implemented in Denovo).

For more information on numerical quadrature, see any number of numerical analysis books. Also, MIT course 18.335 is also highly recommended, as it covers numerical quadrature and many other aspects of numerical methods (including a heavy focus on numerical linear algebra).

Exercises

1. **Negative fluxes**. Using the DD method, express the sweeping step relation in the form
$$\psi_{3/2,n} = A\psi_{1/2,n} + B\,,$$
where $\mu_n > 0$ and A and B are constants to be determined, and eliminate any α terms. Under what conditions can $\psi_{3/2,n}$ be negative? Under what conditions, if any, can $\phi_{1,n}$ be negative? While negative fluxes are not inherently bad numerically, they don't make much physical sense. Suggest a method for "fixing" negative fluxes.
2. **Step Difference**. Implement the step difference method.†
3. **Step Characteristic**. The step characteristic (SC) scheme is another differencing approach based on analytical integration of the transport equation over the fine mesh cell, yielding for example
$$\psi_{i+\frac{1}{2},n} = \psi_{i-\frac{1}{2},n} e^{-\frac{\Delta_i \Sigma_{ti}}{\mu}} + \frac{Q_i}{2\Sigma_{ti}}\left(1 - e^{-\frac{\Delta_i \Sigma_{ti}}{\mu}}\right).$$
   (a) Derive an expression for α_{mi} as in Eq. 12.5 for use with the SC method. Note that α in this case is also indexed by the fine mesh index i.
   (b) Implement the SC method.
4. **Accuracy**. Here, we want to investigate the accuracy of various difference methods. To do so, we will use a known analytical solution for as a reference. See e.g. Appendix A of LeVeque [29] for more on error analysis.
   (a) Using the S_6 approximation, solve *analytically* the discrete ordinates equations for the sample slab problem of Lecture 5. Implement this reference solution as a function in MATLAB (or your language of choice), so that $\phi(x)$ can be evaluated for any value of x.
   (b) Using the DD approximation, solve the S_6 equations using 10, 20, 40, 80, and 160, and 320 meshes. For each case, compute the absolute value of the maximum error between the DD and analytical solution for ϕ, i.e. find
$$e_I^{\text{DD}} = \max_{1 \le i \le I} \left| \phi_i^{\text{DD}} - \phi_i^{\text{ref}} \right|.$$
   (c) Do the same for the SD method.

†Note, in this and other exercises you are asked to use or modify the given S_N code. It is also acceptable to use your own code, either written from scratch or via modification of the one given.

(d) We say a method is pth order accurate if $e(\Delta) \propto \Delta^p$. Estimate p for both methods.

(e) Plot the errors for both methods as a function of Δ. Include also the functions $a\Delta^1$ and $b\Delta^2$, where a and b are constants chosen to yield a nice plot. Hint: use a log-log plot.

5. **Accuracy: Part Deux**. Here, we want to investigate the accuracy of the Gauss-Legendre quadrature and hence S_N order.

   (a) Compute $\phi(x)$ analytically for the sample slab problem of Lecture 5, and make this available as a function.

   (b) Compute ϕ analytically using the S_N method for orders 2, 4, 8, 16, 32, and 64. You will have to look up the quadrature parameters for $N > 4$.

   (c) Compute the errors in ϕ as in the last problem for $x = 5.0$ and plot as a function of N.

   (d) Estimate the "p" value as in the last problem and comment.

6. **Reflective Conditions**. Modify the given S_N code to handle reflective conditions, and solve the following problems...

7. **CMR**. Implement CMR in the given S_N code. Test it on the following problems, using a variety of coarse mesh sizes...

8. **CMFD**. Read the paper by Zhong et al. [49] and derive the CMFD equations in 1-d. Implement the method in the S_N code and compare it to CMR for the slab configuration given in the last problem.

9. **Discrete Ordinates Matrix**. The discrete ordinates equations are simply a set of coupled differential equations discretized in space and angle. Consequently, we can express the equations in matrix form. Consider a uniform slab (Σ_t and Σ_s) with a uniform isotropic source of volumetric strength S. Suppose we discretize this slab into three cells of equal width Δ. Using an S_2 approximation with the DD method, and assuming vacuum conditions:

   (a) Cast the sweep-based source iteration scheme in matrix form, expressing the right hand vector in terms of Q, the emission density.

   (b) Reformulate the equations so that the source term is brought to the left hand side, thus enabling a direct solution of the equations without source iteration.

   For $\Sigma_t = 1.0$, $\Sigma_s = 0.5$, $\Delta = 1.0$, and $S = 1.0$, verify your expressions from parts (a) and (b). For more on part (b), see the paper by Patton and Holloway [36].

Lecture 12 | Nonlinear Acceleration

The discrete ordinates method is in theory a very fast, memory-efficient technique. Unless the angular flux is needed explicitly, only one set of edge angular fluxes is needed at a time (e.g. the left edge flux for computing the right edge flux, which then can overwrite the left edge flux).

Despite the algorithmic simplicity of the method, the underlying source iteration scheme can be obnoxiously slow for problems with high scattering ratios. Several methods have been introduced over the years to accelerate the source convergence. Of these, diffusion synthetic acceleration (DSA) is probably the most powerful and widespread. The method essentially involves performing a diffusion solve to update the scalar flux, and has shown spectacular success for a wide variety of problems. DSA and related methods will be cast as "preconditioners" for linear solvers in the next lecture.

Here, we look at two relatively simple methods that have been quite successful: the *coarse mesh rebalance* (CMR) and *Nonlinear Diffusion Acceleration* (NDA) methods. Although the presentation is restricted to the one-group case, both methods apply readily to multigroup fixed-source and eigenvalue problems. In addition, both methods can be applied successfully to other transport discretization, including the method of characteristics.

Coarse Mesh Rebalance

CMR was probably the first wide-spread method for accelerating discrete ordinates calculations. Its foundation rests in the *neutron balance equation*, which we obtain in one dimension by integrating Eq. (11.1) over angle, leading to

$$\frac{\partial}{\partial x} J(x) + \Sigma_t(x)\phi(x) = \Sigma_s(x)\phi(x) + S(x)\,, \tag{12.1}$$

where $J = \mathbf{J} \cdot \hat{i}$ is the net current in the x direction. Now suppose we divide the domain into a number of coarse meshes, indexed by j, with cell edges at $x_{j-\frac{1}{2}}$ and

$x_{j+\frac{1}{2}}$. Subtracting the scattering source from both sides and integrating over the jth coarse mesh, we obtain

$$J_{j+\frac{1}{2}} - J_{j-\frac{1}{2}} + \int_{x_i} \Sigma_r(x)\phi(x)dx = \int_{x_i} S(x)dx\,, \tag{12.2}$$

where $\Sigma_t - \Sigma_s$ is the *removal* cross section. Expressing Eq. (1.2) as $J = J^+ - J^-$, where $+$ indicates to the right and $-$ to the left, we have

$$J^+_{j+\frac{1}{2}} - J^-_{j+\frac{1}{2}} - J^+_{j-\frac{1}{2}} + J^-_{j-\frac{1}{2}} + R_j = S_j\,, \tag{12.3}$$

where the partial currents are defined

$$J^{\pm}_{j+\frac{1}{2}} = \sum_{\mu_n \gtrless 0} \mu_n \psi_{j+\frac{1}{2},n}\,, \tag{12.4}$$

R_j is the coarse mesh-integrated removal rate, and S_j is the coarse mesh-integrated source.

Equation (12.3) must be satisfied by a fully converged fine mesh solution. For an unconverged iterate, which we denote $\psi^{m+\frac{1}{2}}$, Eq. (12.3) is generally not satisfied. To force the flux to satisfy neutron balance, we introduce multiplicative rebalance factors f_j such that

$$\psi^{m+1}_{i,n} = \begin{cases} f_j \psi^{m+\frac{1}{2}}_{i,n} & x_{j-\frac{1}{2}} < x_i < x_{j+\frac{1}{2}} \\ f_{j-1}\psi^{m+\frac{1}{2}}_{i,n} & x_i = x_{j+\frac{1}{2}} \text{ and } \mu_n > 0 \\ f_{j+1}\psi^{m+\frac{1}{2}}_{i,n} & x_i = x_{j+\frac{1}{2}} \text{ and } \mu_n < 0 \end{cases}. \tag{12.5}$$

Note the appropriate factor's index denotes from which coarse mesh the neutrons originate. For the case of vacuum boundaries, the corresponding incident partial current vanishes. For a reflective boundary at the left, $f_0 = f_1$, and similarly for the right boundary.

To illustrate the method, consider a slab with vacuum boundaries divided into three coarse meshes. Substituting the modified fluxes into the balance equation yields a set of three equations:

$$\begin{aligned} f_1 J^+_{3/2} - f_2 J^-_{3/2} + f_1 J^-_{1/2} + f_1 R_1 &= f_1 S_1 \\ f_2 J^+_{5/2} - f_3 J^-_{5/2} - f_1 J^+_{3/2} + f_2 J^-_{3/2} + f_2 R_2 &= f_2 S_2 \\ f_3 J^+_{7/2} - f_2 J^+_{5/2} + f_3 J^-_{5/2} + f_3 R_3 &= f_3 S_3\,, \end{aligned} \tag{12.6}$$

which yields a tridiagonal system written in condensed form as

$$\mathbf{A}\mathbf{f} = \mathbf{S}\,. \tag{12.7}$$

Solving this equation for $\mathbf{f}$ and updating to ψ^{m+1} allows us to recompute the scattering source, which—hopefully—converges faster than without the rebalancing scheme. Look to the Further Reading section for references that investigate stability of coarse mesh rebalance. Algorithm 1 shows how CMR (or any low order acceleration scheme) can be used within source iteration.

initialization;
while ψ^m *or* ϕ^m *not converged* **do**
  compute scattering source ;
  $\psi^{m+\frac{1}{2}} \leftarrow \text{sweep}(\psi^m)$;
  if *using CMR* **then**
    $\psi^{m+1} \leftarrow \text{cmr}(\psi^{m+\frac{1}{2}})$;
  else
    $\psi^{m+1} \leftarrow \psi^{m+\frac{1}{2}}$;
  end
end

Algorithm 1: Accelerated Source Iteration

Nonlinear Diffusion Acceleration

Nonlinear Diffusion (or Coarse-Mesh, Finite-Difference) Acceleration (NDA or CMFD) was originally developed by Kord Smith as a way to reduce the memory requirements of advanced, nodal-diffusion solvers [41]. The method also led to signficant speedups and has since been applied routinely in for lattice-physics computations based on the method of characteristics [**?**].

We begin along the same lines as we did for CMR by integrating the transport equation over a coarse mesh indexed by j, which leads to Eq. (12.2). Now, define the flux-weighted cross section

$$\Sigma_{t,j} = \frac{\int_{x_{j-1/2}}^{x_{j+1/2}} \Sigma_r(x)\phi^{n+1/2}dx}{\int_{x_{j-1/2}}^{x_{j+1/2}} \phi^{n+1/2}dx} , \tag{12.8}$$

with similar expressions for $\Sigma_{r,j}$ and $\Sigma_{s,j}$, along with the coarse-mesh-averaged flux

$$\phi_j = \frac{\int_{x_{j-1/2}}^{x_{j+1/2}} \phi(x)dx}{\int_{x_{j-1/2}}^{x_{j+1/2}} dx} \tag{12.9}$$

and source

$$S_j = \frac{\int_{x_{j-1/2}}^{x_{j+1/2}} S(x)dx}{\int_{x_{j-1/2}}^{x_{j+1/2}} dx} \,. \tag{12.10}$$

With these definitions, Eq. (12.2) can be written as

$$J_{j+1/2}^{n+1/2} - J_{j-1/2}^{n+1/2} + \Delta_j \Sigma_{r,j} \phi_j = \Delta_j S_j \tag{12.11}$$

Previously, we studied fine-mesh diffusion, which gave

$$\begin{aligned} J_{j+1/2}^{n+1/2} &= -\left(\frac{2D_j D_{j+1}}{D_j \Delta_{j+1} + D_{j+1}\Delta_j}\right)(\phi_{j+1} - \phi_j) \\ &= \tilde{D}_{j+1/2}(\phi_{j+1} - \phi_j) \,. \end{aligned} \tag{12.12}$$

where D_j is the diffusion coefficient for the jth coarse cell. Although the choice of D_j is not unique, a common approach is to set

$$D_j = 1/3\Sigma_{t,j} \tag{12.13}$$

or

$$D_j = \frac{\int_{x_{j-1/2}}^{x_{j+1/2}} D(x)\phi(x)dx}{\int_{x_{j-1/2}}^{x_{j+1/2}} \phi(x)dx} \,. \tag{12.14}$$

A better approach may be to weight $D(x)$ by $J(x)$ instead of $\phi(x)$, but in practice $J(x)$ is not usually computed.

Even with a flux- or current-weighted D_j, the current as defined by Eq. (12.12) is not generally equal to the current as computed directly from $\psi^{n+1/2}$, as in Eq. XXX, because the diffusion approximation is not guaranteed to yield the same solution as a transport approximation. Hence, even if the flux were perfectly converged, the diffusion approximation, coarse mesh or otherwise, will not reproduce the solution and, hence, does not preserve balance over the coarse cells.

We can enforce balance by letting

$$J_{j+1/2} = -\tilde{D}_{j+1/2}(\phi_{j+1} - \phi_j) - \hat{D}_{j+1/2}(\phi_{j+1} + \phi_j) \,, \tag{12.15}$$

where the nonlinear coupling coefficient $\hat{D}_{j+1/2}$ is defined as

$$\hat{D}_{j+1/2} = \frac{\tilde{D}_{j+1/2}(\phi_{j+1} - \phi_j) + J_{j+1/2}}{\phi_{j+1} + \phi_j)} \,. \tag{12.16}$$

For an albedo condition at the left, we can set

$$J_{1/2} = (\tilde{D}_{1/2} + \hat{D}_{1/2})\phi_1 \,, \tag{12.17}$$

where

$$\hat{D}_{1/2} = \frac{J_{1/2} - \tilde{D}_{1/2}\phi_1}{\phi_1} \,. \tag{12.18}$$

The resulting, one-dimensional diffusion equation is

$$\begin{aligned} -\tilde{D}_{j+1/2}(\phi_{j+1} - \phi_j) - \hat{D}_{j+1/2}(\phi_{j+1} + \phi_j) + \\ \tilde{D}_{j-1/2}(\phi_j - \phi_{j-1}) + \hat{D}_{j-1/2}(\phi_j + \phi_{j-1}) + \\ \Sigma_{r,j}\phi_j^{n+1}\Delta_j = S_j\Delta_j \,. \end{aligned} \tag{12.19}$$

Once solved, the coarse-mesh flux ϕ^n can be used to update the fine-mesh flux via

$$\phi^{n+1}(x) = \frac{\phi_j^n}{\phi_j^{n-1/2}}\phi(x) \quad x \in [x_{j-1/2}, x_{j+1/2}] \,. \tag{12.20}$$

With $\phi^n(x)$ found, the scattering source can be updated, another sweep can be performed to obtain $\psi^{n+3/2}$, and the process can be repeated.

Like CMR, if the NDA converges, it must converge to the right answer. However, that convergence is not guaranteed []. Sometimes, for a given $\tilde{D}$, the computed $\hat{D}$ leads to a diffusion matrix that is not diagonally dominant. While this does not prevent solution of the equations, the updated, coarse-mesh flux may contain unphysical features that lead to instability. Specifically, if $\hat{D}_{j+1/2} > \tilde{D}_{j+1/2}$, setting

$$\hat{D}_{j+1/2} = \tilde{D}_{j+1/2} = \frac{J_{j+1/2}}{2\phi_j} \tag{12.21}$$

ensures a diagonally-dominant matrix.

In addition, the use of D_j directly may be replaced by Larsen's "effective" diffusion coefficient, which may be particularly valuable for the step-characteristic scheme in use for both MOC and S_N calculations. Even though D_j is arbitary, its proper selection can help convergence.

Finally, it has been found that instabilities can be reduced significantly by using under-relaxation of the nonlinear coupling coefficient []. In other words, the coefficients are defined

$$\hat{D}_{j+1/2}^{n+1} = \alpha\hat{D}_{j+1/2}^{n+1} + (1-\alpha)\hat{D}_{j+1/2}^{n} \,, \tag{12.22}$$

where $\alpha \in [0, 1)$ is the relaxation parameter. Based on past experience, typical values for α range from 0.5 to 0.8, though the optimum value is highly problem dependent.

Further Reading

A review paper by Adams and Larsen [5] provides a survey of the many acceleration techniques available for the discrete ordinates equations, and with its several hundred references, is the place to look for further information. Lewis and Miller provides more information on coarse mesh rebalance, and Cefus and Larsen have assessed its stability [12]. Park and Cho [34] have suggested angular-dependent rebalance schemes, and their work is a good place to start to find references to other CMR variations. The discontinuity factors used in CMFD were first proposed by Smith in the context of homogenization [40], and he later proposed their use for acceleration [41]. A recent paper by Zhong et al. [49] provides a modern overview and advanced use of the approach along with a relatively detailed set of equations for implementation.

Lecture 13 | Iterative Solvers for Transport Equations

So far, three basic models for discretizing the transport equations have been discussed. The latter model, neutron diffusion, leads to a very straightforward system of linear equations for which standard techniques readily apply. The S_N and MOC transport schemes also lead to a set of linear equations. However, casting these equations in a standard form for solution by generic solvers is less obvious. Because our goal is ultimately to apply generic solvers to the equations, it is useful here first to provide a general representation of the equations.

Angular Flux Equations

Before proceeding, we adopt an operator notation following Larsen and Morel [28] to simplify somewhat the exposition to follow. Recall that the scalar flux is defined

$$\begin{aligned} \phi(\vec{r}) &= \int_{4\pi} d\Omega\, \psi(\vec{r}, \Omega) \\ &\approx \sum_{n=1}^{N} w_n \psi_n(\vec{r})\,. \end{aligned} \tag{13.1}$$

Introducing a *discrete-to-moment* operator $\mathcal{D}$, we enforce

$$\phi = \mathcal{D}\psi\,, \tag{13.2}$$

where space and angle indices are implicit. Moreover, we introduce a *moment-to-discrete* operator $\mathcal{M}$, defined

$$\psi = \mathcal{M}\phi\,, \tag{13.3}$$

where in general $\mathcal{D} \neq \mathcal{M}^{-1}$. Finally, defining the operator

$$\mathcal{L}(\,\cdot\,) \equiv \left(\hat{\Omega}\cdot\nabla + \Sigma_t(\vec{r})\right)(\,\cdot\,)\,, \tag{13.4}$$

the one group discretized transport equation becomes

$$\mathcal{L}\psi = \mathcal{M}\mathcal{S}\phi + q\,, \tag{13.5}$$

where $\mathcal{S}$ is the scattering operator (which can include fission). For the case of isotropic scattering in one group, we have $\mathcal{M} = \frac{1}{4\pi}$ while $\mathcal{S} = \Sigma_s(\vec{r})$.

For the multigroup problem, this notation generalizes to

$$\mathcal{L}_g\psi_g = \mathcal{M}\sum_{g'=1}^{G}\left(\mathcal{S}_{gg'} + \frac{1}{k}\mathcal{X}_g\mathcal{F}_{g'}\right)\phi_{g'} + q_g\,, \tag{13.6}$$

where the fission source has been explicitly represented. Here, $\mathcal{X}$ represents the fission spectrum χ in operator form.

Moments Equation

S_N in Operator Form

Equations 13.5 and 13.6 represent equations in which the unknown is the angular flux ψ. In general, and particularly for reactor physics, our interest is in computing reaction rates for which only the scalar flux is needed. Consequently, in practice the angular flux is rarely stored explicitly, but is rather computed on-the-fly during a sweep through the space-angle grid. This can be represented explicitly by manipulating the equations to be functions only of the scalar flux (and higher order flux moments if scattering is anisotropic).

To illustrate, consider Eq. 13.5. Multiplying through by the space-angle *transport sweep* operator $\mathcal{T} = \mathcal{D}\mathcal{L}^{-1}$, we have

$$\mathcal{D}\psi = \mathcal{D}\mathcal{L}^{-1}\mathcal{M}\mathcal{S}\phi + \mathcal{D}\mathcal{L}^{-1}q\,. \tag{13.7}$$

Recalling $\phi = \mathcal{D}\psi$ and rearranging yields

$$(\mathcal{I} - \mathcal{T}\mathcal{M}\mathcal{S})\phi = \mathcal{T}q\,, \tag{13.8}$$

or

$$\mathcal{A}_{\text{WG}}\phi = b\,, \tag{13.9}$$

where

$$\mathcal{A}_{\text{WG}} \equiv \mathcal{I} - \mathcal{T}\mathcal{M}\mathcal{S} \tag{13.10}$$

is the *within-group transport operator* and

$$b = \mathcal{T}q \tag{13.11}$$

represents the *uncollided* flux moments. Eq. 13.8 represents a linear equation in standard form for the scalar flux moments ϕ.

The operator $\mathcal{D}\mathcal{L}^{-1}$ has a very specific physical interpretation. The inverse $\mathcal{L}^{-1}$ represents the set of space-angle sweeps from one global boundary to another. The additional factor $\mathcal{D}$ implies that along the space-angle sweep, ϕ is updated, thus eliminating the need to store ψ. This is exactly how S_N and MOC solvers have traditionally been implemented.

The extension to multigroup yields a similar form

$$\left(\mathcal{I} - \mathcal{T}_{\text{MG}}\mathcal{M}_{\text{MG}}\left(\mathcal{S}_{\text{MG}} + \frac{1}{k}\mathcal{X}_{\text{MG}}\mathcal{F}_{\text{MG}}^{\mathsf{T}}\right)\right)\phi = \mathcal{T}_{\text{MG}}q_{\text{MG}}\,, \tag{13.12}$$

where $\mathcal{T}_{\text{MG}}$ and $\mathcal{M}_{\text{MG}}$ are block diagonal operators having blocks $\mathcal{T}_g$ and $\mathcal{M}$, respectively, $\mathcal{X}_{\text{MG}}$, $\mathcal{F}_{\text{MG}}$, and q_{MG} are the column vectors having elements $\mathcal{X}_g$, $\mathcal{F}_g$, and q_g respectively, and the multigroup scattering operator is defined

$$\mathcal{S}_{\text{MG}} = \begin{pmatrix} \mathcal{S}_{11} & \cdots & \mathcal{S}_{1G} \\ \vdots & \ddots & \vdots \\ \mathcal{S}_{G1} & \cdot & \mathcal{S}_{GG} \end{pmatrix}. \tag{13.13}$$

Similar to the within-group problem, we define the *multigroup transport operator* to be

$$\mathcal{A}_{\text{MG}} = \mathcal{I} - \mathcal{T}_{\text{MG}}\mathcal{M}_{\text{MG}}\left(\mathcal{S}_{\text{MG}} + \frac{1}{k}\mathcal{X}_{\text{MG}}\mathcal{F}_{\text{MG}}^{\mathsf{T}}\right). \tag{13.14}$$

MOC in Operator Form

While the operators and equations apply directly the S_N method, slight changes must be made to represent MOC. In MOC, we iterate on the quantities averaged over a *cell*, specifically the scalar flux moments ϕ_{cell}. However, the angular fluxes are computed for each track *segment*, and each cell can have several tracks crossing it. The average angular flux on each track segment in a given cell is used to generate the average scalar flux.

This can be cast in operator form for a one group problem as

$$\mathcal{L}\psi_{\text{seg}} = \mathcal{P}\mathcal{M}\mathcal{S}\phi_{\text{cell}} + \mathcal{P}q_{\text{cell}}\,, \tag{13.15}$$

where q_{cell} is the cell-wise external source and $\mathcal{P}$ is a cell-to-segment operator, "P" denoting prolongation, though not quite in the multigrid sense.

We can also define a segment-to-cell operator (restriction) $\mathcal{R}$ so that

$$\psi_{\text{cell}} = \mathcal{R}\psi_{\text{seg}}\,, \tag{13.16}$$

and

$$\phi_{\text{cell}} = \mathcal{D}\mathcal{R}\psi_{\text{seg}} \,. \tag{13.17}$$

Finally, the moments equation akin to Eq. 13.8 is

$$\left(\mathcal{I} - \mathcal{D}\mathcal{R}\mathcal{L}^{-1}\mathcal{P}\mathcal{M}\mathcal{S}\right)\phi_{\text{cell}} = \mathcal{D}\mathcal{R}\mathcal{L}^{-1}\mathcal{P}q_{\text{cell}} \,. \tag{13.18}$$

In the context of MOC, the operator

$$\mathcal{A}_{\text{WG-MOC}} = \mathcal{I} - \mathcal{D}\mathcal{R}\mathcal{L}^{-1}\mathcal{P}\mathcal{M}\mathcal{S} \tag{13.19}$$

is implied when discussing the within-group transport operator, and similarly for the multigroup variant.

Boundary Conditions

So far, we have neglected to treat boundary conditions. In particular, reflecting (or periodic, or white) boundary conditions create an additional set of unknowns, namely the incident (or exiting) boundary fluxes. In theory, treating these conditions is straightforward, but we refrain from further discussion for two reasons. First, adding boundary unknowns complicates presentation of the algorithms discussed below. Second, and more importantly, response function calculations use vacuum conditions only, and so treatment of reflecting boundaries is outside our scope. A more detailed discussion can be found in the work of Warsa *et al.* [45].

13.1 Transport Solvers

Consider the fixed source multigroup transport problem represented by

$$\mathcal{A}_{\text{MG}}\phi = b \,. \tag{13.20}$$

The traditional method for solving the multigroup equations is via a nested iteration in which a series of within-group equations is solved for each group, with the scattering (and possibly) fission sources updated between each group solve. More modern treatments view the equations as a complete set to be solved simultaneously.

Gauss-Seidel

The Gauss-Seidel method has long been used to solve the multigroup equations. For standard problems, the method first solves the fast (within-)group equation. The updated fast group flux can then be used to define the in-scatter source for the next group, and so on. Because this algorithm really implies inversion of each group

block (as defined in Eq. 13.12), the method is more accurately described as *block* Gauss-Seidel in energy.

Algorithm 2 describes a basic implementation of the method. Historically, convergence of the method is based on the difference of successive fluxes, either as an absolute norm as indicated or a norm of the relative point-wise difference. Assessing convergence in this way can be misleading. The difference between successive iterates can be much smaller than the difference between an iterate and the solution, sometimes resulting in premature convergence. Section **??** provides a numerical comparison of various convergence criteria.

Data: initial guess for group fluxes $\phi^{(0)}$, external source q, eigenvalue k, tolerance τ, maximum number of iterations N
Result: converged group fluxes ϕ
$n = 1$
while $\tau > \epsilon$ *and* $n < N$ **do**
  for g *from 1 to* G **do**
    // Compute in-scatter and in-fission sources
    $b_g = \mathcal{T}_g \left(\mathcal{M} \sum_{g'=1}^{G} \left(\mathcal{S}_{gg'} + \frac{1}{k} \mathcal{X}_g \mathcal{F}_{g'} \right) \phi_{g'}^{(n-1)} + q_g \right)$
    // Solve the within-group problem
    $\phi_g^{(n)} = \mathcal{A}_g^{-1} b_g$
  end
  $\tau = ||\phi^{(n)} - \phi^{(n-1)}||$
  $n = n + 1$
end

Algorithm 2: Gauss-Seidel Algorithm for the Multigroup Transport Equation

For problems in which there is no upscatter and no fission, Gauss-Seidel is an essentially exact scheme assuming the within-group equations are solved exactly. However, for cases with upscatter or fission, extra "upscatter" iterations are required, and in some cases, the convergence of Gauss-Seidel becomes prohibitively slow.

Source Iteration

The standard method for solving the within-group transport equation has been *source iteration*. The basic idea is a simple one: given an external source (including in-scatter and fission), we guess the flux, compute the within-group scattering source, solve for a new flux, and repeat until converged. Mathematically, source iteration is defined by the process

$$\phi^{(n)} = \mathcal{T}\mathcal{M}\mathcal{S}\phi^{(n-1)} + \mathcal{T}q\,. \tag{13.21}$$

However, recall that Richardson iteration for the system $\mathcal{A}x = b$ is just

$$x^{(n)} = (\mathcal{I} - \mathcal{A})x^{(n-1)} + b\,. \tag{13.22}$$

Since $\mathcal{I} - \mathcal{A}_{\text{WG}} = \mathcal{TMS}$, we see that source iteration is equivalent to Richardson iteration.

Source iteration has a particularly intuitive physical interpretation. Suppose our initial guess is zero. The resulting flux is just the right hand side, or the uncollided flux. Performing one iteration adds the contribution of neutrons having undergone a single collision. A second iteration includes the twice-collided neutrons, and so on. For systems in which neutrons undergo many scattering events, the physics suggests the process can be painfully slow to converge. The math verifies the physics [28]: the error $||\phi^{(n)} - \phi^{(n-1)}||$ of the infinite medium one group problem goes as the scattering ratio $c = \frac{\Sigma_s}{\Sigma_t}$, meaning that increased scattering leads to slower convergence.

Krylov Solvers

Because of the limitations of the Gauss-Seidel and source iteration schemes, much work has been done to apply modern linear solvers to transport problems. One of the most successful class of solvers studied consists of Krylov subspace methods.

Overview

One can categorize linear (and eigenvalue) solvers as being *stationary* or *nonstationary*. Stationary methods produce updated solutions using only a single previous solution, and the Gauss-Seidel and Richardson methods are both stationary. Other well-known examples include the Jacobi and successive over-relaxation (SOR) methods. On the other hand, nonstationary methods produce a solution based on two or more previous iterates (or the information used to create those iterates).

Krylov methods are nonstationary methods that rely on construction of a so-called *Krylov subspace* of dimension n, defined for an $m \times m$ operator $\mathcal{A}$ as

$$\mathcal{K}(n, x_0) \equiv \text{span}\{x_0,\, \mathcal{A}x_0,\, \mathcal{A}^2x_0,\, , \ldots ,\, , \mathcal{A}^{m-1}x_0\}\,, \tag{13.23}$$

for some initial, possibly random, vector x_0. The main idea of Krylov subspace methods is to find $x \in \mathcal{K}(m, x_0)$ that "best" solves the system of interest, be it an eigenproblem or linear system, where it is assumed that $m \ll n$.

Working with $\mathcal{K}(n, x_0)$ directly is difficult numerically, since repeated application of $\mathcal{A}$ sends the initial vector x_0 into the same direction, namely that of the dominant eigenvector of $\mathcal{A}$. Hence, the basis must be orthogonalized. The canonical

approach for nonsymmetric operators is Arnoldi's method, which by successive application of the modified Gram-Schmidt process yields the Arnoldi decomposition

$$\mathcal{A}\mathcal{V} = \mathcal{V}\mathcal{H} + fe_m^{\mathsf{T}}\,, \tag{13.24}$$

where $\mathbf{V} \in \mathbb{R}^{m\times n}$ consists of orthonormal columns, $\mathbf{H} \in \mathbb{R}^{n\times n}$ is an upper Hessenberg matrix, e_n is the n-vector of all zeros except a one in the nth location, and f is the residual, which is orthogonal to the columns of $\mathbf{V}$.

The most popular Krylov method for nonsymmetric linear systems (such as the transport equations above) is GMRES [**?**]. The basic idea of GMRES is straightforward: the nth step of GMRES produces a $n \times n$ Hessenberg matrix and the corresponding basis $\mathcal{V}$, and the approximate solution x_n is found by minimizing the residual norm $||r||_2 = ||\mathcal{A}x_n - b||_2$ for $x_i \in \mathcal{K}(n, x_0)$. In other words, x_n must be in the column space of $\mathcal{K}(n, x_0)$, mathematically expressed as $x_n = \mathcal{V}y$, where y satisfies

$$\begin{aligned} ||\mathcal{A}\mathcal{V}y - b||_2 &= ||\mathcal{V}^{\mathsf{T}}\mathcal{A}\mathcal{V}x_n - \mathcal{V}^{\mathsf{T}}b|| \\ &= ||\mathcal{H}y - \mathcal{V}^{\mathsf{T}}b||\,. \end{aligned} \tag{13.25}$$

The last equation shows that GMRES can be thought of as finding the best solution $x_n \in \mathcal{K}(n, x_0)$ in a least-squares sense.

Preconditioning

While Krylov methods are generally more robust than the classical stationary methods, their performance can be improved significantly via *preconditioning*. A preconditioner $\mathcal{M}$ is an operator whose inverse satisfies $\mathcal{M}^{-1} \approx \mathcal{A}^{-1}$ in some sense and is relatively inexpensive to apply.

A *left preconditioned* linear system is

$$\mathcal{M}^{-1}\mathcal{A}x = \mathcal{M}^{-1}b \tag{13.26}$$

while a *right preconditioned* system is

$$\mathcal{A}\mathcal{M}^{-1}y = b \tag{13.27}$$

with $x = \mathcal{M}^{-1}y$. The left preconditioned residual differs from the original residual but may be a better metric for convergence. The right preconditioned system preserves the original residual.

A preconditioner typically leads to a clustering of eigenvalues. As an extreme example, suppose that $\mathcal{M} = \mathcal{A}$. The preconditioned operator is then $\mathcal{A}\mathcal{M}^{-1} = \mathcal{I}$, for which all the eigenvalues are unity. Of course, to apply $\mathcal{M}^{-1}$ in this case

represents solving the original problem. While preconditioners cannot in general be expected to yield a set of eigenvalues equal to unity, any clustering typically improves convergence. Often, even pushing eigenvalues away from the origin tends to improve convergence [28]. Chapter **??** provides a relatively thorough development of several diffusion-based preconditioners for the transport equation.

Krylov Methods for the Transport Problems

Krylov solvers have been used to solve both the within-group [45] and multigroup [19] transport equations. For the multigroup equations in particular, the independent nature of the group-wise blocks makes parallelization in energy much more straightforward than for Gauss-Seidel.

For problems in which there is no fission and upscatter is limited to a subset of thermal groups, it is possible to solve the downscatter groups via Gauss-Seidel and to use a Krylov method on the thermal block of Eq. 13.12. Doing so can yield improved efficiency for some problems [19]. However, when fission is included, as is often true for response function generation, there is always thermal-to-fast coupling, and solving the full system via a Krylov method is to be preferred.

Lecture 14 | Linear Acceleration

While Krylov methods are generally more robust than the classical stationary methods, their performance can be improved significantly via *preconditioning*. A preconditioner $\mathcal{M}$ is an operator whose inverse satisfies $\mathcal{M}^{-1} \approx \mathcal{A}^{-1}$ in some sense and is relatively inexpensive to apply.

A *left preconditioned* linear system is

$$\mathcal{M}^{-1}\mathcal{A}x = \mathcal{M}^{-1}b \tag{14.1}$$

while a *right preconditioned* system is

$$\mathcal{A}\mathcal{M}^{-1}y = b \tag{14.2}$$

with $x = \mathcal{M}^{-1}y$. The left preconditioned residual differs from the original residual but may be a better metric for convergence. The right preconditioned system preserves the original residual.

A preconditioner typically leads to a clustering of eigenvalues. As an extreme example, suppose that $\mathcal{M} = \mathcal{A}$. The preconditioned operator is then $\mathcal{A}\mathcal{M}^{-1} = \mathcal{I}$, for which all the eigenvalues are unity. Of course, to apply $\mathcal{M}^{-1}$ in this case represents solving the original problem. While preconditioners cannot in general be expected to yield a set of eigenvalues equal to unity, any clustering typically improves convergence. Often, even pushing eigenvalues away from the origin tends to improve convergence [28]. Chapter **??** provides a relatively thorough development of several diffusion-based preconditioners for the transport equation.

Chapter **??** provided a short survey of trends in transport discretizations and solvers applicable to the fixed source multiplying problems relevant to response matrix methods. As noted, Krylov methods have quickly become important tools for a variety of transport problems, but to be successful, they require adequate *preconditioning*. This chapter details several preconditioners relevant to transport calculations for reactor physics, and specifically the comparatively small problems characteristics of response matrix methods.

14.1 Diffusion Synthetic Acceleration

Frequently, the most successful preconditioners are those that are based on *a priori* knowledge of the physics or structure of the problem. This has long been the case for accelerating transport problems, for which early techniques enforced balance over coarse regions in space, angle, and energy.

A related idea that has remained an important innovation is to apply a simple, low order approximation, invariably based on diffusion, to provide an efficient update or correction to an unconverged transport solution. Here, we provide a brief sketch of one diffusion-based scheme known as *diffusion synthetic acceleration* that has long been used but only recently as a preconditioner. For simplicity, the development is given for the one group case, following the excellent treatment of Larsen and Morel [28].

Suppose we perform one source iteration, so that

$$\phi^{n+\frac{1}{2}} = \mathcal{T}\mathcal{M}\mathcal{S}\phi^n + \mathcal{T}q\,. \tag{14.3}$$

where $\mathcal{S}$ is assumed to contain both the scatter and fission within-group terms. Note the half index. We subtract this equation from Eq. 13.8 to get

$$(\mathcal{I} - \mathcal{T}\mathcal{M}\mathcal{S})\phi - \phi^{n+\frac{1}{2}} = -\mathcal{T}\mathcal{M}\mathcal{S}\phi^n\,, \tag{14.4}$$

and adding $\mathcal{T}\mathcal{M}\mathcal{S}\phi^{n+\frac{1}{2}}$ to both sides and rearranging,

$$\epsilon = \overbrace{(\mathcal{I} - \mathcal{T}\mathcal{M}\mathcal{S})^{-1}\mathcal{T}\mathcal{M}}^{\text{what we approximate}}\overbrace{\mathcal{S}(\phi^{n+\frac{1}{2}} - \phi^n)}^{v} \tag{14.5}$$

where ϵ is the error. Note the error satisfies the transport equation

$$(\hat{\Omega}\cdot\nabla + \Sigma)\varepsilon - \frac{\Sigma_s}{4\pi}\epsilon = \frac{v}{4\pi}\,, \tag{14.6}$$

where

$$\epsilon = \int_{4\pi} d\Omega\,\varepsilon\,. \tag{14.7}$$

Solving the error equation is just as expensive as solving the original transport equation. As an alternative, we can use the diffusion approximation. In this case, we have

$$(-\nabla\cdot D\nabla + \Sigma - \Sigma_s)\epsilon = v\,, \tag{14.8}$$

or in operator form,

$$\epsilon = \mathcal{C}^{-1}v = \mathcal{C}^{-1}\mathcal{S}(\phi^{n+\frac{1}{2}} - \phi^n)\,. \tag{14.9}$$

This leads to the update

$$\begin{aligned}\phi \approx \phi^{n+1} &= \phi^{n+\frac{1}{2}} + \mathcal{C}^{-1}\mathcal{S}(\phi^{n+\frac{1}{2}} - \phi^n) \\ &= (\mathcal{I} + \mathcal{C}^{-1}\mathcal{S})\phi^{n+\frac{1}{2}} - \mathcal{C}^{-1}\mathcal{S}\phi^n \\ &= \Big(\mathcal{I} - (\mathcal{I} + \mathcal{C}^{-1}\mathcal{S})(\mathcal{I} - \mathcal{T}\mathcal{M}\mathcal{S})\Big)\phi^n + (\mathcal{I} + \mathcal{C}^{-1}\mathcal{S})\mathcal{T}q \\ &= (\mathcal{I} - \mathcal{P}^{-1}_{\text{WG-DSA}}\mathcal{A})\phi^n + \mathcal{P}^{-1}_{\text{WG-DSA}}\mathcal{T}q\,,\end{aligned} \tag{14.10}$$

which is the preconditioned source iteration (i.e. source iteration plus diffusion synthetic acceleration), and where

$$\mathcal{P}^{-1}_{\text{WG-DSA}} = (\mathcal{I} + \mathcal{C}^{-1}\mathcal{S}) \tag{14.11}$$

is the within-group diffusion preconditioning process. In standard left-preconditioner form, we have

$$\mathcal{P}^{-1}_{\text{WG-DSA}}\mathcal{A} = \mathcal{P}^{-1}_{\text{WG-DSA}}\mathcal{T}q\,. \tag{14.12}$$

Previous work indicates diffusion preconditioning coupled with a Krylov solver is less restrictive with respect to discretization and more effective for realistic problems [45].

14.2 Multigroup DSA

It is natural to extend the diffusion preconditioner to multigroup problems, yielding the process

$$\mathcal{P}^{-1}_{\text{MG-DSA}} \equiv \mathcal{I} + \mathcal{C}^{-1}\left(\mathcal{S} + \mathcal{X}\mathcal{F}^T\right)\,, \tag{14.13}$$

where $\mathcal{C}$ is the multigroup diffusion operator defined block-wise as

$$\mathcal{C}_{gg'} \equiv \delta_{gg'}\left(\nabla \cdot D_g(\vec{r})\nabla + \Sigma_{tg}(\vec{r})\right) - \Sigma_{sgg'}(\vec{r}) - \chi_g\nu\Sigma_{fg'}(\vec{r})\,. \tag{14.14}$$

An initial review of the literature yielded no application of multigroup diffusion as a preconditioner for multigroup transport problems. Related work addressed acceleration of outer Gauss-Seidel upscatter iterations [**?**] and fission iterations based on the rank one fission operator ($\mathcal{X}\mathcal{F}^{\mathsf{T}}$) [**?**], but all using an equivalent one group formulation. In general, the full multigroup diffusion problem is itself expensive for large problems (an issue addressed below), and this may explain why it has not been used extensively.

14.3 Coarse Mesh Diffusion Preconditioning

Preconditioning the multigroup equations with diffusion can lead to very large diffusion operators, and for many problems, the cost of constructing and, more importantly, applying the preconditioner becomes prohibitive. As an alternative, we investigate the use of coarse mesh diffusion preconditioners. Coarse mesh diffusion operators have long been central to acceleration techniques in reactor analysis, a chief example being the nonlinear diffusion acceleration scheme developed by Smith [**?**] for nodal diffusion methods and later extended to transport methods [**?**].

While no results could be found in the literature describing use of coarse mesh diffusion as a preconditioner, more general coarse mesh schemes in space, angle, and energy have been of substantial recent interest, particularly for multigrid preconditioning. Multigrid methods, like DSA (itself a two-grid method in angle), use a coarse grid solution to improve a fine grid solution by way of restriction (essentially averaging) and prolongation (essentially interpolation) operations. The idea is that slowly-varying, slowly-converging error modes become highly oscillatory, quickly-converging modes on the coarse mesh. While a complete description of multigrid methods is outside the present scope, the following sections describe implementation of a two-grid diffusion preconditioner. For a more complete overview of multigrid methods, the reader would be best served by one of the standard reviews available, *e.g.* Ref. [**?**].

A Spatial Two-Grid Multigroup Diffusion Preconditioner

In this work, we apply a two-grid spatial scheme to the diffusion preconditioner. Recent work suggests that multigrid methods in the energy variable can work very well [**?**, **?**]. While a slightly different context, nonlinear diffusion acceleration also typically uses a coarse energy mesh to great effect. However, our initial studies using a coarsened energy variable within a diffusion preconditioner suggests that the simultaneous restriction of space, angle, and energy may have inherent difficulties.

The coarse mesh diffusion preconditioner is a natural extension to Eq. 14.13 and is defined as

$$\mathcal{P}^{-1}_{\text{MG-DSA}} \equiv \mathcal{I} + \boldsymbol{P}\mathcal{C}_H^{-1}\boldsymbol{R}\left(\mathcal{S} + \mathcal{X}\mathcal{F}^T\right) , \tag{14.15}$$

where $\boldsymbol{P}$ and $\boldsymbol{R}$ are the *prolongation* and *restriction* operators, respectively, and $\mathcal{C}_H$ is the diffusion operator defined on the coarse spatial mesh.

Coarse Mesh Operator Homogenization

To define $\mathcal{C}_H$, cross sections must be *homogenized* over the fine mesh. The most physically-sound approach for producing averaged cross sections is to use flux-weighting

in a manner that preserves reaction rates. While a variety of such homogenization techniques exist, we use a rather conventional homogenization scheme that is simple to apply in preconditioning. For the particular case of group constant generation via an assembly-level lattice physics solver, the results of the scheme are referred to as *assembly homogenized cross sections* [**?**].

To illustrate, suppose we need the total cross section Σ_t averaged over coarse mesh cell j. Denoting the fine mesh flux in cell $i \in j$ to be ϕ_i, we have

$$\Sigma_{t,j} = \frac{\sum_{i\in j} V_i \phi_i \Sigma_{t,i}}{\sum_{i\in j} V_i \phi_i} \,, \tag{14.16}$$

where V_i is the fine mesh cell volume. If we define the coarse mesh flux to be the volume average

$$\phi_j = \frac{\sum_{i\in j} V_i \phi_i}{V_j} \,, \tag{14.17}$$

where

$$V_j = \sum_{i\in j} V_i \,, \tag{14.18}$$

then the total interaction rate in coarse cell j is

$$\sum_{i\in j} V_i \Sigma_{t,i} \phi_i = \Sigma_j \sum_{i\in j} V_i \phi_i = V_j \phi_j \Sigma_{t,j} \,. \tag{14.19}$$

Hence, the averaged quantities preserve the integrated reaction rate associated with the given fine mesh flux. All group constants, including the diffusion coefficient D, can be generated in this way.

The obvious problem with this approach is that the fine mesh flux ϕ_i is not known. The simplest approximation is to assume a constant flux, leading to volume-weighted cross sections. For preconditioning, this is a suitable approximation, since conserving true reaction rates is not a prerequisite. Unlike certain nonlinear acceleration techniques (*e.g.* CMFD) that rely on conservation to provide an integral form of the *solution* at each step, preconditioning—a linear process—only seeks to provide an additive improvement.

However, that certainly does *not* preclude use of more accurate flux shapes to achieve better results. In typical lattice physics applications, group constants are found by solving the transport equation in a pincell or assembly subject to reflecting conditions and using the resulting spectrum for weighting. For preconditioning, a simple pincell homogenization scheme could be used, in which 2-D pincell problems approximating parts of the full problem are solved, and the resulting region-wise volume-averaged spectra are used to produce homogenized materials in the appropriate cells. Analysis of this approach would be a natural extension to the present work.

One might note that the current flux iterate is freely available for use in homogenization; however, because the flux and homogenization process would change from step to step, the entire process would become nonlinear, and its application within the context of standard Krylov linear solvers would be suspect. Its use in solvers allowing variable preconditioners (*e.g.* FGMRES [**?**]) would be valuable future research.

Restriction

To restrict the fine mesh input vector for application of the inverse coarse mesh diffusion operator, a simple spatial average is used. As an example, consider a 1-D problem discretized into 6 fine meshes. The coarse meshing process is based on a level parameter l that defines the number of fine meshes per coarse mesh. Suppose $l = 2$, leading to the fine-to-coarse mapping $[0, 0, 1, 1, 2, 2]$. The restriction operator is defined

$$\boldsymbol{R} = \begin{pmatrix} v_0 & v_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & v_2 & v_3 & 0 & 0 \\ 0 & 0 & 0 & 0 & v_4 & v_5 \end{pmatrix} , \tag{14.20}$$

where

$$v_i = \frac{V_i}{V_j} , \qquad \text{for fine mesh } i \text{ in coarse mesh } j , \tag{14.21}$$

and hence $\sum_{i \in j} v_i = 1$.

Prolongation

To prolong the coarse mesh result back to the fine mesh, the coarse mesh value is distributed on the fine grid based on the (approximate) flux $\tilde{\phi}$ used to produce the coarse mesh diffusion operator. Given a coarse mesh value ϕ_j, the prolonged flux is defined as

$$\phi_{i \in j} = \phi_j \frac{l\tilde{\phi}_i}{\sum_{i \in j} \tilde{\phi}_i} . \tag{14.22}$$

For the example above, suppose each coarse region is assumed to have a fine mesh flux shape of $\tilde{\phi} = [a, b]$, where $a + b = 1$. The prolongation operator is defined

$$\boldsymbol{P} = \begin{pmatrix} 2a & 0 & 0 \\ 2b & 0 & 0 \\ 0 & 2a & 0 \\ 0 & 2b & 0 \\ 0 & 0 & 2a \\ 0 & 0 & 2b \end{pmatrix} . \tag{14.23}$$

In the constant flux approximation ($a = b = 0.5$), multiplication of $\boldsymbol{R} \in \mathbb{R}^{n\times m}$ by $\boldsymbol{P} \in \mathbb{R}^{m\times n}$ yields the identity matrix $\mathcal{I} \in \mathbb{R}^{n\times n}$.

Smoothing

While the coarse mesh scheme described so far represents a complete preconditioner, it can be significantly improved by a *smoothing* operation. The motivation for smoothing is that the coarse mesh solve damps low frequency error modes of the fine mesh problem but not high frequency modes. In more physical terms, a coarse mesh solve can be expected to get the gross shape right, but not the finer details. A smoothing operator uses a few iterations of a classical scheme like Richardson, Jacobi, or Gauss-Seidel with the fine mesh operator. In many multigrid algorithms, a smoother is also often used before the coarse mesh solve to ensure the error is "smooth" before restriction; however, in this case the joint scattering and fission operator $\mathcal{S} + \frac{1}{k}\mathcal{X}\mathcal{F}^{\mathsf{T}}$ precedes the coarse mesh solve and can be expected to provide a suitable input vector.

For the coarse mesh diffusion preconditioner, smoothing requires the action of the fine mesh diffusion matrix. While the cost of producing this matrix may be somewhat large, that cost is much smaller than inverting the operator in a fine mesh preconditioner (and much, much smaller than application of the transport operator). The smoothing process used is the weighted Jacobi method, which for the generic problem $\mathcal{A}x = b$ is defined by the process

$$x^{(n+1)} = -\omega\mathcal{A}_D^{-1}(\mathcal{A}_L + \mathcal{A}_U)x^{(n)} + (1-\omega)\mathcal{A}_D^{-1}b\,, \tag{14.24}$$

where $\mathcal{A}_D$ represents the diagonal of $\mathcal{A}$, $\mathcal{A}_L$ and $\mathcal{A}_U$ represent the strictly lower and strictly upper triangular parts of $\mathcal{A}$, and ω is the weighting parameter. For completely solving the linear system, $\omega = 1$ is ideal (with $\omega > 1$ potentially yielding instability); however, for the purposes of damping high frequency errors, the optimum is typically smaller, with special 1-D and 2-D model cases having optima of $\omega = \frac{2}{3}$ and $\omega = \frac{4}{5}$, respectively [**?**]. Algorithm 3 provides a basic outline of the coarse mesh diffusion preconditioner.

14.4 Transport-Corrected Diffusion Preconditioners

Explicit Versus Matrix-Free Transport Operators

A final technique developed is to form a preconditioner that contains information from the transport operator rather than relying solely on the diffusion approximation. In theory, forming the transport operator defined by Eq. 13.14 explicitly is possible. With an explicit operator, various approximate factorizations become natural

Data: Input vector v_0, number of smoothing iterations N, weighting factor ω
Result: Preconditioned vector v_1
// Apply the scattering and fission operator
$x = (\mathcal{S} + \frac{1}{k}\mathcal{X}\mathcal{F}^{\mathsf{T}})v_0$
// Apply restriction
$x_H = \nabla x$
// Solve the coarse mesh diffusion problem
$y_H = \mathcal{C}_H^{-1} x_H$
// Apply prolongation
$z = \sqrt{y_H}$
// Optionally apply smoothing
for *i from 1 to* N **do**
  $z = -\omega\mathcal{C}_D^{-1}(\mathcal{C}_L + \mathcal{C}_U)z + (1-\omega)\mathcal{C}_D^{-1}x$
end
// Compute output vector
$v_1 = v_0 + z$

Algorithm 3: Two-Grid Coarse Mesh Multigroup Diffusion Preconditioner

preconditioner options. However, constructing such operators implies working with matrices with sizes proportional to the number of angles, at least implicitly, as part of the construction process. This can quickly lead to huge memory requirements, a complicated construction process, or both.

Construction of explicit transport operators have been studied for for the discrete ordinates method [36] and for the method of characteristics [**?**]. The former work treated the angular flux directly, and the analysis was limited to relatively small multigroup 1-D problems. The latter work noted the cost of constructing the transport operator was exceedingly large and developed a marginally successful parallel scheme limited to the within-group equations.

In the present work, the transport operator is always a "matrix free" operator, meaning that the action $y \leftarrow \mathcal{A}x$ is done by functions and not explicit matrix operations. Hence, while the action of $\mathcal{A}$ is available, $\mathcal{A}$ itself is not, so preconditioners based on approximate factorizations are not directly applicable. Ultimately, we restrict any preconditioner using a transport operator to be based on actions only.

Newton-based Correction

As an alternative, suppose we have a matrix inverse $\mathcal{M}^{-1}$ that represents an initial approximation of $\mathcal{A}^{-1} \in \mathbb{R}^{m\times m}$. For application to transport problems, we let $\mathcal{M}$ be the diffusion preconditioner $\mathcal{P}_{\text{MG-DSA}}$ or its coarse mesh equivalent. Our goal is

to improve this preconditioning process using the (possibly approximate) action of $\mathcal{A}^{-1}$.

To do this, we apply a method attributed by various sources to Shulz or Hotelling and Bodewig [**?**, **?**], but in fact which represents application of Newton's method for finding a matrix inverse. For brevity, we refer to it as the Newton-Shulz method. While a more rigorous analysis of the method can be found elsewhere [**?**], it is possible to motivate it by considering the simple problem $ax = 1$ following Ref. [**?**]. Given a, we seek x. One way to view this is as the nonlinear problem $f(x) = 1/x - a = 0$. Then $f'(x) = -1/x^2$, and Newton's method (see Section **??**) gives the process

$$\begin{aligned} x_n &= x_{n-1} + \frac{1}{x_{n-1}^{-2}} \left(\frac{1}{x_{n-1}} - a \right) \\ &= 2x_{n-1} - x_{n-1} a x_{n-1} \,. \end{aligned} \tag{14.25}$$

For the more general case of matrices, this suggests the process

$$\mathcal{X} = \mathcal{X}(2\mathcal{I} - \mathcal{A}\mathcal{X}) \,, \tag{14.26}$$

where $\mathcal{X} \approx \mathcal{A}^{-1}$.

As is always the case for Newton's method, convergence to the solution depends on the initial guess being sufficiently close to the solution. By using $\mathcal{P}_{\text{MG-DSA}}$ as the initial guess, convergence has always been achieved in our studies. Hence, a simple one step transport-corrected diffusion preconditioning process is defined by

$$\mathcal{P}^{-1}_{\text{TC-MG-DSA-1}} = \mathcal{P}^{-1}_{\text{MG-DSA}}(2\mathcal{I} - \mathcal{A}_{\text{MG}}\mathcal{P}^{-1}_{\text{MG-DSA}}) \,. \tag{14.27}$$

While Eq. 14.27 represents an improved preconditioner, it adds the cost of an additional space-angle-energy sweep, the number of which we are ultimately trying to minimize. As an alternative, we can substitute

$$\mathcal{A}_{\text{MG}} \approx \tilde{\mathcal{A}}_{\text{MG}} \,, \tag{14.28}$$

where the tilde indicates some approximation. Here, we let that imply a coarser angular quadrature. While this appears at first to be a multigrid method, note that the (approximate) transport operator is not being inverted but rather is being used to improve the inversion of the diffusion preconditioner. In the terminology of Newton methods, using the approximate operator results in an approximate Jacobian.

Because only a few iterations of Newton's method are needed to converge given an appropriate initial guess, using the approximate operator $\tilde{\mathcal{A}}_{\text{MG}}$ will lead to *its* approximate inverse and *not* that of $\mathcal{A}_{\text{MG}}$. However, $\tilde{\mathcal{A}}^{-1}_{\text{MG}}$ should be even closer to

$\mathcal{A}_{\text{MG}}^{-1}$ than the original diffusion preconditioning process, and so application of the full transport operator can be expected to yield a much more valuable improvement if used in one final Newton iteration.

Algorithm 4 provides the complete preconditioning process for application to an input vector based on the recursive application of the Newton-Schulz method defined in Algorithm 5. The true operator is only applied once in the last step, but the number of actions of $\tilde{\mathcal{A}}$ grows exponentially with the number of iterations: one for $k = 2$, three for $k = 3$, seven for $k = 4$, and so on. Hence, the cost of applying the preconditioner is likely to become excessive for more than two or three iterations unless the approximate operator is significantly less expensive than the true operator.

Data: transport operator $\mathcal{A}$, approximate transport operator $\tilde{\mathcal{A}}$, diffusion preconditioner $\mathcal{P}_{\text{MGDPC}}$, number of corrections k, input vector x
Result: output vector y
$y \leftarrow$ Newton-Shulz$(\tilde{\mathcal{A}}, \mathcal{P}_0, k-1, x)$
$y \leftarrow \mathcal{A}y$
$y \leftarrow 2x - y$
$y \leftarrow \mathcal{P}_0^{-1}y$

Algorithm 4: Transport-Corrected Diffusion Preconditioner

Data: operator $\mathcal{A}$, initial preconditioner $\mathcal{P}_0$, number of corrections k, input vector x
Result: output vector y
if $k > 1$ **then**
  // Apply earlier steps recursively
  $y \leftarrow$ Newton-Shulz$(\mathcal{A}, \mathcal{P}_0, k-1, x)$
  $y \leftarrow 2x - \mathcal{A}y$
  $y \leftarrow$ Newton-Shulz$(\mathcal{A}, \mathcal{P}_0, k-1, y)$
else
  // Apply the initial preconditioner
  $y \leftarrow (2\mathcal{I} - \mathcal{A}\mathcal{P}_0^{-1})x$
  $y \leftarrow \mathcal{P}_0^{-1}y$
end

Algorithm 5: Newton-Shulz

14.5 Summary

This chapter has presented the classic diffusion synthetic acceleration technique and its multigroup analog in the context of preconditioning for iterative transport solvers. Additionally, a coarse mesh variant was proposed based on two spatial grids and optional smoothing that represents in theory a much less computationally-intensive process than the fine mesh preconditioner. To incorporate more of the transport operator, the Newton-Shulz method was used to form a transport-corrected diffusion preconditioner that has the potential to yield better preconditioners than diffusion alone can provide.

Lecture 15 | Time-Dependent Transport

This lecture investigates approaches for solving the time-dependent neutron transport equation using the discrete ordinates method. Several important historical schemes are presented. Additionally, a generalized approach for solving the equations via generic solvers is presented, and several options and challenges are identified. Throughout, delayed neutron precursors are explicitly included.

15.1 The Time-Dependent Transport Equation

We begin with the time-dependent, multigroup transport equation (TDTE),

$$\begin{aligned}
\overbrace{\frac{1}{v_g}\frac{\partial \psi_g}{\partial t}}^{\text{time rate of change}} &= -(\overbrace{\hat{\Omega}\cdot\nabla\psi_g(\vec{r},\Omega,t)+\Sigma_{tg}(\vec{r},t)\psi_g(\vec{r},\Omega,t)}^{\text{rate of destruction}}) \\
&+ \overbrace{\frac{1}{4\pi}\int_{4\pi} d\Omega' \sum_{g'=1}^{G}\Sigma_{sg\leftarrow g'}(\vec{r},\Omega'\cdot\Omega,t)\psi_{g'}(\vec{r},\Omega',t)}^{\text{production rate from in-scatter}} \\
&+ \overbrace{\frac{1}{4\pi k_{\text{crit}}}\int_{4\pi} d\Omega' \left(\chi_{pg}(\vec{r})(1-\beta(\vec{r}))\sum_{g'=1}^{G}\nu\Sigma_{fg'}(\vec{r},t)\psi_{g'}(\vec{r},\Omega',t)\right)}^{\text{production rate from prompt fission}} \\
&+ \overbrace{\sum_i \lambda_i C_i(\vec{r},t)\chi_{ig}(\vec{r})}^{\text{production rate from delayed fission}} \\
&+ \overbrace{q_g(\vec{r},\hat{\Omega},t)}^{\text{production rate from external sources}} \quad , \quad g=1\ldots G
\end{aligned} \tag{15.1}$$

with a corresponding set of precursor equations,

$$\frac{\partial C_i}{\partial t} = \frac{\beta_i(\vec{r})}{k_{\text{crit}}} \int_{4\pi} d\Omega' \sum_{g'=1}^{G} \nu\Sigma_{fg'}(\vec{r},t)\psi_{g'}(\vec{r},\Omega',t) - \lambda_i C_i(\vec{r},t)\,, \quad i = 1 \ldots N\,. \tag{15.2}$$

While the notation is essentially standard, it does imply isotropic scattering in the LAB system. Furthermore, it assumes that a steady-state calculation has been performed for the critical k-eigenvalue and use of modern delayed neutron data such that a single set of decay constants λ_i is applicable to all materials.

Before proceeding, we adopt the operator notation developed in Lecture **??** to simplify somewhat the exposition to follow. For the time-dependent multigroup problem, this notation generalizes to

$$\frac{1}{v_g}\frac{\partial \psi_g}{\partial t} = -\mathcal{L}_g\psi_g + \mathcal{M}\sum_{g'=1}^{G}\left(\mathcal{S}_{gg'} + \mathcal{X}_{pg}\mathcal{F}_{g'}\right)\phi_{g'} + \mathcal{M}\sum_{i=1}^{I}\lambda_i\mathcal{X}_{dig}C_i\,, \tag{15.3}$$

and

$$\frac{\partial C_i}{\partial t} = \beta_i\sum_{g=1}^{G}\mathcal{F}_g\phi_g - \lambda_i C_i\,, \tag{15.4}$$

where $\mathcal{F}$ has absorbed the eigenvalue, and $\mathcal{X}_{pg}$ has absorbed the $1-\beta$ term.

15.2 Time-Integration Schemes

There appear to exist two broad categories of time integration approaches applicable to solving the TDTE. The first and historically dominant approach discretizes the equations in time in some (usually simple) manner, and introduces "synthetic" materials and sources to convert the time step into a fixed source problem. This approach is attractive because it can make use of an existing fixed source code without much additional code. Moreover, by consistently differencing the fluxes and precursors, it is possible to eliminate the precursors as an unknown, thus reducing the size of the system. A drawback to this approach is that for complex time integration schemes, the modified coefficients may become arbitrarily complex. Any change in a scheme would require different coefficients. Furthermore, coupling to other physics would be limited to operator splitting, possibly with an inner iteration to correct.

A second approach is to consider the TDTE as a generic ODE for use in general time integration schemes. This approach is attractive because arbitrarily high order integration schemes would be applicable. Moreover, it would be comparitively easier to couple other physics, yielding one consistent set of ODE's. A significant potential

downside to this approach is that the operators used to define the ODE may be more difficult to apply and may require significant new code. Interestingly, and perhaps telling of these challenges, it appears no production codes use this latter approach.

15.3 A Survey of Historical Time Integration Schemes

At least three variants of the first approach are found in production codes of the last several decades. Much of the discussion to follow is similar to the review given by Dishaw [16], though we have made corrections and our own observations where needed. In all cases, where applicable, the historical schemese make use of synthetic sources and materials in order to convert the time-dependent problem into a series of fixed-source problems.

TIMEX-1972

In the first incarnation of the Los Alamos TIMEX code, the time discretization is defined using a semi-implicit, first order discretization defined by

$$\frac{1}{\Delta_t v_g}\left(\psi_g^{n+1}-\psi_g^n\right)+\mathcal{L}\psi_g^{n+1}=\mathcal{M}\sum_{g'=1}^{G}\left(\mathcal{S}_{gg'}+\mathcal{X}_{pg}\mathcal{F}_{g'}\right)\phi_{g'}^n+\mathcal{M}\sum_{i=1}^{I}\lambda_i\mathcal{X}_{dig}C_i^n\,, \tag{15.5}$$

where n indexes the time step. Note that the latest time step $n+1$ occurs only on the left hand side; hence, the implicit component of this step is relatively small. Note, original documentation for this method was not found; hence, we rely on Dishaw's review [16].

Substituting

$$\tilde{\Sigma}_{tg}\equiv\Sigma_{tg}+\frac{1}{\Delta_t v_g}\,, \tag{15.6}$$

with a corresponding change to $\tilde{\mathcal{L}}$ and defining

$$\tilde{q}^{n+1}\equiv\mathcal{M}\sum_{g'=1}^{G}\left(\mathcal{S}_{gg'}+\mathcal{X}_{pg}\mathcal{F}_{g'}\right)\phi_{g'}^n+\mathcal{M}\sum_{i=1}^{I}\lambda_i\mathcal{X}_{dig}C_i^n-\frac{1}{\Delta_t v_g}\psi_g^n\,, \tag{15.7}$$

yields the equation

$$\tilde{\mathcal{L}}_g\psi_g^{n+1}=\tilde{q}_g^{n+1}\,, \tag{15.8}$$

which is a one group fixed source problem with pure absorption. The precursors are updated based on the new flux.

This method has two chief downsides. The first is its inherent instability for large time steps being only slightly implicit. The second is its inherent inaccuracy caused

by lagging all scattering and fission sources. What this implies is that the scattering and fission sources are never converged at a time step, a potentially huge source of error. The method was attractive because of its efficiency: only a single transport sweep is required per step.

TIMEX-1976

A second incarnation of TIMEX [23] improved its time discretization, using a modified semi-implicit scheme defined by

$$\frac{1}{\Delta_t v_g}\left(\psi_g^{n+1}-\psi_g^n\right)+\mathcal{L}\psi_g^{n+1}=\mathcal{M}\sum_{g'=1}^{g-1}\left(\mathcal{S}_{gg'}+\mathcal{X}_{pg}\mathcal{F}_{g'}\right)\phi_{g'}^{n+1} + \mathcal{M}\sum_{g'=g}^{G}\left(\mathcal{S}_{gg'}+\mathcal{X}_{pg}\mathcal{F}_{g'}\right)\phi_{g'}^{n}+\mathcal{M}\sum_{i=1}^{I}\lambda_i\mathcal{X}_{dig}C_i^n\,. \tag{15.9}$$

This updated scheme introduces something similar to Gauss-Seidel iteration for defining the scattering and fission sources, though it converges neither. Consequently, this scheme is still subject to gross inaccuracy. As done for TIMEX 1972, we can define synthetic materials and sources. Using $\tilde{\mathcal{L}}$ as above and defining

$$\tilde{q}_g=\mathcal{M}\sum_{g'=1}^{g-1}\left(\mathcal{S}_{gg'}+\mathcal{X}_{pg}\mathcal{F}_{g'}\right)\phi_{g'}^{n+1}+\mathcal{M}\sum_{g'=g}^{G}\left(\mathcal{S}_{gg'}+\mathcal{X}_{pg}\mathcal{F}_{g'}\right)\phi_{g'}^{n}+\mathcal{M}\sum_{i=1}^{I}\lambda_i\mathcal{X}_{dig}C_i^n\,, \tag{15.10}$$

we have the equation

$$\tilde{\mathcal{L}}\psi_g=\tilde{q}_g\,, \tag{15.11}$$

which, as expected, has the same form as above.

PARTISN

A successor to TIMEX is the Los Alamos code PARTISN [7], which remains a standard production code for many institutions today. The time discretization in PARTISN is equivalent to the the implicit midpoint method [6]. Consider the ODE

$$\frac{dy(t)}{dt}=f(t,y(t))\,. \tag{15.12}$$

The implicit midpoint method defines the step

$$y^{n+1}=y^n+\Delta_t f\left(t^n+\frac{\Delta_t}{2},\frac{1}{2}\left(y^{n+1}+y^n\right)\right)\,. \tag{15.13}$$

This can implemented as a sequence of two steps. Defining the midpoint value

$$y^{n+\frac{1}{2}} = \frac{1}{2}\left(y^{n+1} + y^n\right) , \tag{15.14}$$

we perform a backward Euler step,

$$y^{n+\frac{1}{2}} = y^n + \frac{\Delta_t}{2} f\left(t + \frac{\Delta_t}{2}, y^{n+\frac{1}{2}}\right) . \tag{15.15}$$

Using the computed $y^{n+\frac{1}{2}}$ and the midpoint definition, we then extrapolate to t^{n+1} using

$$y^{n+1} = 2y^{n+\frac{1}{2}} - y^n . \tag{15.16}$$

It can be shown this is a second order accurate method.

We can apply this method to the TDTE. The fully implicit backward Euler half-step is defined via

$$\frac{2}{\Delta_t v_g}\left(\psi_g^{n+\frac{1}{2}} - \psi_g^n\right) + \mathcal{L}\psi_g^{n+\frac{1}{2}} = \mathcal{M}\sum_{g'=1}^{G}\left(\mathcal{S}_{gg'} + \mathcal{X}_{pg}\mathcal{F}_{g'}\right)\phi_{g'}^{n+\frac{1}{2}} + \mathcal{M}\sum_{i=1}^{I}\lambda_i\mathcal{X}_{dig}C_i^{n+\frac{1}{2}} , \tag{15.17}$$

where the consistently-discretized precursors are governed by

$$\frac{2}{\Delta_t}\left(C_i^{n+\frac{1}{2}} - C_i^n\right) = \beta_i\sum_{g=1}^{G}\mathcal{F}_g\phi_g^{n+\frac{1}{2}} - \lambda_i C_i^{n+\frac{1}{2}} . \tag{15.18}$$

Taking these steps, the updated fluxes and precursors are defined via

$$\begin{aligned}\psi_g^{n+1} &= 2\psi_g^{n+\frac{1}{2}} + \psi_g^n \\ C_i^{n+1} &= 2C_i^{n+\frac{1}{2}} + C_i^n .\end{aligned} \tag{15.19}$$

As done for the TIMEX methods, we can cast Eqs. 15.17 and 15.18 as a fixed-source problem. First, we solve Eq. (15.18) for the updated concentrations,

$$C_i^{n+\frac{1}{2}} = \frac{\Delta_t'\beta_i}{1+\Delta_t'\lambda_i}\sum_{g=1}^{G}\mathcal{F}_g\phi_g^{n+\frac{1}{2}} + \frac{C_i^n}{1+\Delta_t'\lambda_i} , \tag{15.20}$$

where $\Delta_t' = \frac{\Delta_t}{2}$. Substituting this into Eq. (15.17) yields

$$\begin{aligned}\frac{1}{\Delta_t' v_g}\left(\psi_g^{n+\frac{1}{2}} - \psi_g^n\right) + \mathcal{L}_g\psi_g^{n+\frac{1}{2}} &= \mathcal{M}\sum_{g'=1}^{G}\left(\mathcal{S}_{gg'} + \mathcal{X}_{pg}\mathcal{F}_{g'}\right)\phi_{g'}^{n+\frac{1}{2}} \\ &+ \mathcal{M}\sum_{i=1}^{I}\lambda_i\mathcal{X}_{dig}\left(\frac{\Delta_t'\beta_i}{1+\Delta_t'\lambda_i}\sum_{g'=1}^{G}\mathcal{F}_{g'}\phi_{g'}^{n+\frac{1}{2}} + \frac{C_i^n}{1+\Delta_t'\lambda_i}\right) .\end{aligned} \tag{15.21}$$

Rearranging, we find

$$\left(\frac{1}{\Delta_t' v_g}+\mathcal{L}_g\right)\psi_g^{n+\frac{1}{2}} = \mathcal{M}\sum_{g'=1}^{G}\left(\mathcal{S}_{gg'}+\left(\mathcal{X}_{pg}+\sum_{i=1}^{I}\lambda_i\mathcal{X}_{dig}\left(\frac{\Delta_t'\beta_i}{1+\Delta_t'\lambda_i}\right)\right)\mathcal{F}_{g'}\right)\phi_{g'}^{n+\frac{1}{2}} + \mathcal{M}\sum_{i=1}^{I}\lambda_i\mathcal{X}_{dig}\left(\frac{C_i^n}{1+\Delta_t'\lambda_i}\right)+\frac{\psi_g^n}{\Delta_t' v_g}\,, \tag{15.22}$$

Defining

$$\tilde{\mathcal{L}}_g \equiv \frac{1}{\Delta_t' v_g}+\mathcal{L}_g\,, \tag{15.23}$$

$$\tilde{\chi}_g \equiv \left(\mathcal{X}_{pg}+\sum_{i=1}^{I}\lambda_i\mathcal{X}_{dig}\left(\frac{\Delta_t'\beta_i}{1+\Delta_t'\lambda_i}\right)\right)\,, \tag{15.24}$$

and

$$\tilde{q}_g^{n+\frac{1}{2}} \equiv \mathcal{M}\sum_{i=1}^{I}\lambda_i\mathcal{X}_{dig}\left(\frac{C_i^n}{1+\Delta_t'\lambda_i}\right)+\frac{\psi_g^n}{\Delta_t' v_g}\,, \tag{15.25}$$

we have

$$\tilde{\mathcal{L}}_g\psi_g^{n+\frac{1}{2}} = \mathcal{M}\sum_{g'=1}^{G}\left(\mathcal{S}_{gg'}+\tilde{\mathcal{X}}_g\mathcal{F}_{g'}\right)\phi_{g'}^{n+\frac{1}{2}}+\tilde{q}_g^{n+\frac{1}{2}}\,, \tag{15.26}$$

which is a standard fixed source multigroup problem.

The actual implementation in PARTISN is slightly different from that presented above. In PARTISN, the angular flux is defined for $t^{i-\frac{1}{2}}, t^{i+\frac{1}{2}}, \ldots$ while the scalar fluxes, and, hence, reaction rates are computed at the centered time. Moreover, it does not appear that PARTISN includes delayed neutrons.

Another modern code, TD-TORT, also employs a fully implicit approach [39], though the exact method was not described in the available literature. However, it appears the implementation used is limited to backward Euler and hence is only first order.

15.4 Higher-Order Synthetic Schemes

It was mentioned above that for all but the simplest time-stepping schemes, a time discretization of the TDTE becomes quite complex. There does exist a class of high order methods that does allow a rather simple discretization in terms only of synthetic sources and materials called the *backward differentiation formula* (BDF) methods

[29]. The first order BDF method is equivalent to the backward Euler scheme. The BDF methods of orders 1 through 3 are defined by

$$\begin{aligned} y^{n+1} - y^n &= \Delta_t f(t^{n+1}, y^{n+1}) \\ \frac{3}{2}y^{n+1} - \frac{4}{2}y^n + \frac{1}{2}y^{n-1} &= \Delta_t f(t^{n+1}, y^{n+1}) \\ \frac{11}{6}y^{n+1} - \frac{18}{6}y^n + \frac{9}{6}y^{n-1} - \frac{2}{6}y^{n-2} &= \Delta_t f(t^{n+1}, y^{n+1}) \,, \end{aligned} \tag{15.27}$$

though the methods are stable through sixth order. These can be written more compactly as

$$\sum_{i=0}^{m} a_i y^{n-i+1} = \Delta_t f(t^{n+1}, y^{n+1}) \,. \tag{15.28}$$

As done for backward Euler, we can define synthetic materials and sources. In this case, defining

$$\tilde{\mathcal{L}}_g \equiv \frac{a_0}{\Delta_t v_g} + \mathcal{L}_g \,, \tag{15.29}$$

$$\tilde{\chi}_g \equiv \left(\mathcal{X}_{pg} + \sum_{i=1}^{I} \lambda_i \mathcal{X}_{dig} \left(\frac{\Delta_t \beta_i}{a_0 + \Delta_t \lambda_i} \right) \right) \,, \tag{15.30}$$

and

$$\tilde{q}_g^{n+1} \equiv \mathcal{M} \sum_{i=1}^{I} \left(\frac{\mathcal{X}_{dig} \lambda_i}{a_0 + \lambda_i \Delta_t} \sum_{j=1}^{m} a_m C_i^{n-j+1} \right) - \sum_{j=1}^{m} \frac{a_j \psi_g^{n-j+1}}{\Delta_t v_g} \,, \tag{15.31}$$

yields the same standard multigroup form

$$\tilde{\mathcal{L}}_g \psi_g^{n+1} = \mathcal{M} \sum_{g'=1}^{G} \left(\mathcal{S}_{gg'} + \tilde{\mathcal{X}}_g \mathcal{F}_{g'} \right) \phi_{g'}^{n+1} + \tilde{q}_g^{n+1} \,. \tag{15.32}$$

Of course, using a a BDF method with order $m > 1$ requires that more than one previous state vector is stored. The cost associated with this higher memory use may or may not be outweighted by the larger time steps higher order methods allow.

It is worth noting that no mention of the BDF methods for neutron transport can be found in the literature (though admittedly the search was not great). Numerical studies will be needed to assess how valuable these schemes are for such problems.

15.5 Solution Via General Solvers

We can cast the TDTE into a form suitable for solution via general ODE time integration schemes. In particular, we transform the TDTE into an ODE of the form

$$\frac{du(t)}{dt} = \mathcal{A}\left(t, u(t)\right) u(t)\,. \tag{15.33}$$

Rearranging Eqs. 15.3 and 15.4, we have the following system of equations

$$\frac{d}{dt}\overbrace{\begin{bmatrix} \psi_1 \\ \vdots \\ \psi_G \\ C_1 \\ \vdots \\ C_I \end{bmatrix}}^{\mathbf{u}(t)} = \overbrace{\begin{bmatrix} v_1\mathcal{A}_{11}^{\psi\psi} & \cdots & v_1\mathcal{A}_{1G}^{\psi\psi} & v_1\mathcal{A}_{11}^{\psi C} & \cdots & v_1\mathcal{A}_{I1}^{\psi C} \\ & \ddots & & & \ddots & \\ v_G\mathcal{A}_{G1}^{\psi\psi} & \cdots & v_G\mathcal{A}_{GG}^{\psi\psi} & v_G\mathcal{A}_{1I}^{\psi C} & \cdots & v_G\mathcal{A}_{II}^{\psi C} \\ \mathcal{A}_{11}^{C\psi} & \cdots & \mathcal{A}_{11}^{C\psi} & -\lambda_1 & & \\ & \ddots & & & \ddots & \\ \mathcal{A}_{I1}^{C\psi} & \cdots & \mathcal{A}_{11}^{C\psi} & & & -\lambda_I \end{bmatrix}}^{\mathcal{A}(t)} \overbrace{\begin{bmatrix} \psi_1 \\ \vdots \\ \psi_G \\ C_1 \\ \vdots \\ C_I \end{bmatrix}}^{\mathbf{u}(t)} \tag{15.34}$$

where

$$\mathcal{A}_{gg'}^{\psi\psi} = -\mathcal{L}_g\delta_{gg'} + \mathcal{M}\left(\mathcal{S}_{gg'} + \mathcal{X}_{pg}\mathcal{F}_{g'}\right)\mathcal{D}\,, \tag{15.35}$$

$$\mathcal{A}_{ig}^{\psi C} = \lambda_i\mathcal{M}\mathcal{X}_{dig}\,, \tag{15.36}$$

and

$$\mathcal{A}_{ig}^{C\psi} = \beta_i\mathcal{F}_g\mathcal{D}\,. \tag{15.37}$$

The size of the operator $\mathcal{A}$ is quite large. Assuming one spatial unknown per cell, this size is

$$\text{size}(\mathcal{A}) = N_{\text{cells}} \times \left(N_{\text{groups}} \times N_{\text{angles}} + N_{\text{precursors}}\right)\,. \tag{15.38}$$

To put this in context, suppose we use an S_8 approximation in angle, 2 energy groups, 8 precursor groups, and a 1 cm mesh for a typical two-dimensional reactor benchmark problem such as the LRA problem. This leads to

$$\text{size}(\mathcal{A}) \approx 30000 \times (2 \times 40 + 8) \approx 3 \times 10^6 \text{ unknowns}\,. \tag{15.39}$$

While solving for millions of unknowns is not infeasible, doing so for several thousand time steps brings the total to billions, and so the problem does become quite challenging.

One drawback of the form of Eq. (15.34) is the challenge in defining the action of the operator $\mathcal{A}$. Typically, the linear systems resulting from implicit discretizations

would be solved via Krylov methods, and so the full matrix is not necessary. However, work has been done to construct this full operator for the case of no precursors [44]. The full explicit matrix is large and sparse; typical sparsity patterns are given in Figure **??** for the $\mathcal{A}_{gg'}^{\psi\psi}$ block of $\mathcal{A}$ [38].

An alternative to constructing the operator explicitly would be to define only its action. Most of the action $\mathcal{A}$ is straightforward; for example, a term such as $\mathcal{M}\mathcal{S}_{gg'}\mathcal{D}\psi_{g'}$ represents a typical source construction likely present in most codes in some form. However, the action $\mathcal{L}_g\psi_g$ is not likely found. This is because discrete ordinates codes solve the problem via sweeping across the domain for all angles. This sweep is represented by the *inverse* of the operator $\mathcal{L}_g$. To be more explicit, consider the equation

$$\mathcal{L}_g\psi_g = \mathcal{M}\mathcal{S}_{gg}\mathcal{D}\psi_g + q_g\,. \tag{15.40}$$

A common iteration scheme would solve this fixed source problem via the process

$$\psi_g^{k+1} = \mathcal{L}_g^{-1}\mathcal{M}\mathcal{S}_{gg}\mathcal{D}\psi_g^k + \mathcal{L}_g^{-1}q_g\,, \tag{15.41}$$

where the iteration index k is not to be confused with a time index. Such iteration is called *Richardson iteration* or more frequently in the context of neutron transport *source iteration*.

Hence, transport solvers typically have the action of $\mathcal{L}^{-1}$ available. To construct the action $y \leftarrow \mathcal{L}x$, we can solve the linear system

$$\mathcal{L}^{-1}y = x \tag{15.42}$$

via Krylov methods that require only the action $\mathcal{L}^{-1}$. Alternatively, the individual $\mathcal{L}_g$ for each group could be explicitly constructed, though of course such construction would potentially be repeated at every time step due to any material changes.

It is not clear whether explicitly constructing $\mathcal{A}$ or constructing its action using an explicit or approximate action of $\mathcal{L}_g$ is most economical. As with most such questions, the answer is likely highly dependent on the problem being solved and the machine on which the problem is solved. Worth noting is that either method that does not employ the action of $\mathcal{L}^{-1}$ can be considered a "sweepless" algorithm; such an algorithm is the focus of the work of Davidson and Larsen [14], though they discuss it in the context of a particular spatial discretization and solver.

Exercises

1. **Time-Dependent Diffusion**. Derive the time-dependent, multigroup diffusion equations including delayed neutron precursors.

Part II

Additional Topics

Lecture 16 | The Adjoint and Perturbation Theory

In this lecture, we introduce the *adjoint* form of the transport equation and describe what it represents physically. We then apply the adjoint equation in a useful technique known as *first order* or *linear perturbation theory*. In the next lecture, we make further use of the adjoint in the context of variational approximations.

The Adjoint Function

We define the *inner product* of two functions $\phi(x)$ and $\psi(x)$ to be

$$\langle \phi, \psi \rangle \equiv \int \phi(x)\psi(x)dx \,, \tag{16.1}$$

where ϕ and ψ satisfy appropriate continuity and boundary conditions. A *self-adjoint* operator M satisfies

$$\langle \phi, M\psi \rangle \equiv \langle M\phi, \psi \rangle \,. \tag{16.2}$$

As an example, the operator corresponding to one-speed diffusion can be shown to be self-adjoint, an exercise left to the reader. Note, self-adjoint and *hermitian* are synonymous. You may be familiar with the latter term from quantum physics, and you might recall that such operators (say the Hamiltonian) have real eigenvalues (like the energy) and orthogonal eigenfunctions (such as the nice sines and cosines of the infinite well).

If an operator L is not self-adjoint, it is possible to define an operator L^* that is adjoint to L. Then L^* will operate on adjoint functions $\psi^*(x)$ that may satisfy different boundary conditions than those of $\psi(x)$ on which L operates. The adjoint operator L^* must satisfy

$$\langle \psi^*, L\psi \rangle = \langle \psi, L^*\psi^* \rangle \,, \tag{16.3}$$

which we refer to as the *adjoint identity* (and which is actually a special case of a generalized Green's theorem).

Transport Operator

We now define the transport equation in operator form. Defining the operator

$$\begin{aligned} L\psi \equiv & -\hat{\Omega}\cdot\nabla\psi(\mathbf{r},\mathbf{\Omega},E) \\ & -\Sigma_t\psi + \int_0^\infty dE' \int_{4\pi} d\Omega'\Sigma_s(\mathbf{r},\mathbf{\Omega}\cdot\mathbf{\Omega}',E'\to E)\psi(\mathbf{r},\mathbf{\Omega}',E')\,, \end{aligned} \tag{16.4}$$

the transport equation is simply $L\psi = -Q$, for some source Q; note the sign of the right hand side. The transport operator L is *not self-adjoint*. Convince yourself that this is indeed the case by evaluating the adjoint identity and paying specific attention to the terms corresponding to streaming and scattering. For convenience, we assume ψ is also subject to vacuum conditions on all external surfaces, i.e. $\psi(\mathbf{r},\mathbf{\Omega},E) = 0$ for $\hat{n}\cdot\hat{\Omega} < 0$ where $\hat{n}$ is the outward normal vector.

The adjoint transport operator L^* will satisify $\langle\psi^*, L\psi\rangle = \langle\psi, L^*\psi^*\rangle$, if and only if we define it such that

$$\begin{aligned} L^*\psi^* \equiv & \hat{\Omega}\cdot\nabla\psi^*(\mathbf{r},\mathbf{\Omega},E) \\ & -\Sigma_t\psi^* + \int_0^\infty dE' \int_{4\pi} d\Omega'\Sigma_s(\mathbf{r},\mathbf{\Omega}\cdot\mathbf{\Omega}',E\to E')\psi^*(\mathbf{r},\mathbf{\Omega}',E')\,, \end{aligned} \tag{16.5}$$

with the further restriction that $\psi^*(\mathbf{r},\mathbf{\Omega},E) = 0$ for $\hat{n}\cdot\hat{\Omega} > 0$. It's worth noting we could have chosen conditions other than vacuum boundaries for ψ, which would yield different conditions for ψ^*; see the exercises for the case of reflecting boundaries.

Interpreting the Adjoint

Let's consider a subcritical, time-independent system containing an arbitrary source. Suppose we are interested in a detector response with an associated cross-section $\Sigma_d(\mathbf{r},E)$. Then we have the forward equation

$$\begin{aligned} & \hat{\Omega}\cdot\nabla\psi(\mathbf{r},\mathbf{\Omega},E) + \Sigma_t\psi = \\ & \int_0^\infty dE' \int_{4\pi} d\Omega'\Sigma_s(\mathbf{r},\mathbf{\Omega}\cdot\mathbf{\Omega}',E'\to E)\psi(\mathbf{r},\mathbf{\Omega}',E') + Q(\mathbf{r},\mathbf{\Omega},E)\,, \end{aligned} \tag{16.6}$$

subject to vacuum conditions, and the adjoint equation

$$\begin{aligned} & -\hat{\Omega}\cdot\nabla\psi^*(\mathbf{r},\mathbf{\Omega},E) + \Sigma_t\psi^* = \\ & \int_0^\infty dE' \int_{4\pi} d\Omega'\Sigma_s(\mathbf{r},\mathbf{\Omega}\cdot\mathbf{\Omega}',E\to E')\psi^*(\mathbf{r},\mathbf{\Omega}',E')\,, \end{aligned} \tag{16.7}$$

subject to the appropriate conditions. Here, the detector cross-section, which can be thought of as a "detector response function", is the adjoint source. Now, we multiply Eq. 16.6 by ψ^* and Eq. 16.7 by ψ, subtract the latter from the former, and integrate over all variables to get (operator and inner-product notation)

$$\langle \psi^*, L\psi \rangle - \langle \psi, L^*\psi^* \rangle = -\langle \psi^*, Q \rangle + \langle \psi, \Sigma_d \rangle \,, \tag{16.8}$$

but by the adjoint identity, the left hand side vanishes, and we are left with a most important result:

$$\langle \psi^*, Q \rangle = \langle \psi, \Sigma_d \rangle \,. \tag{16.9}$$

Suppose our forward source is a unit monoenergetic, unidirectional point source, i.e. $Q(\mathbf{r}, \mathbf{\Omega}, E) = \delta(\mathbf{r} - \mathbf{r}_0)\delta(\mathbf{\Omega} - \mathbf{\Omega}_0)\delta(E - E_0)$. From Eq. 16.9, we find

$$\psi^*(\mathbf{r}_0, \mathbf{\Omega}_0, E_0) = \int_V d^3r \int_E dE \int_{4\pi} d\Omega \Sigma_d(\mathbf{r}, E')\psi(\mathbf{r}, \mathbf{\Omega}, E) = R \,, \tag{16.10}$$

where R is the total detector response. In this case, $\psi^*(\mathbf{r}_0, \mathbf{\Omega}_0, E_0)$ is the expected contribution to the detector response due to a unit delta source located at $(\mathbf{r}_0, \mathbf{\Omega}_0, E_0)$ in phase space. More broadly, ψ^* represents the importance of neutrons in a particular region of phase space to a given detector response.

Perturbation Theory — Fixed Source

In this and the next section, we apply the adjoint in determining the change to an integral system parameter* due to a small perturbation in the system. We begin with a fixed source example and follow with an eigenvalue example.

Suppose we have a system described by†

$$L_0\psi_0 = Q_0 \tag{16.11}$$

with vacuum boundaries. Our goal is to evaluate a detector response that takes the form

$$R_0 = \int\int\int d^3r d\Omega dE \Sigma_d(\mathbf{r}, E)Q(\mathbf{r}, E, \mathbf{\Omega}) \,. \tag{16.12}$$

Suppose we now introduce some small change. Perturbation theory can allow us to determine the detector response, to first order accuracy, for a "small" perturbation to the system. Let us define a new, perturbed system to be

$$L\psi = Q \,, \tag{16.13}$$

*Here, an "integral" parameter is any integrated quantity, e.g. a reaction rate in a volume, or k_{eff} in a reactor.

†Note, the operators in this section have absorbed the minus sign of the right hand side sources above.

also subject to vacuum conditions, and where $L = L_0 + \delta L_0$, $Q = Q_0 + \delta Q_0$, and $\psi = \psi_0 + \delta\psi_0$. We assume we know L and Q (we're making the perturbation), and we want to know ψ (and eventually, its effect on R). Rewriting the system, we have

$$\begin{aligned}(L_0 + \delta L_0)(\psi_0 + \delta\psi_0) &= Q_0 + \delta Q_0 \\ L_0\psi_0 + \delta L_0\psi_0 + L_0\delta\psi_0 + \theta(\delta^2) &= Q_0 + \delta Q_0\,,\end{aligned} \tag{16.14}$$

but noting our original equation is contained, we are left with two separate equations,

$$L_0\psi_0 = Q_0\,, \tag{16.15}$$

and

$$\delta L_0\psi_0 + L_0\delta\psi_0 = \delta Q_0\,, \tag{16.16}$$

where we have neglected terms of order δ^2; hence the theory is known as "first order" or "linear" perturbation theory.

We now introduce the adjoint to the *original equation*, with the adjoint source being our detector response function Σ_d, i.e.

$$L_0^*\psi_0^* = \Sigma_d\,, \tag{16.17}$$

and as before, we have $\langle\psi_0^*, Q_0\rangle = \langle\psi_0, \Sigma_d\rangle = R_0$. Our goal is to determine δR_0, or $R - R_0 = \delta R_0 = \langle\delta\psi_0, \Sigma_d\rangle$. To do so, we multiply Eq. 16.16 by ψ_0^* and Eq. 16.17 by $\delta\psi_0$, subtract the latter from the former, and integrate over all phase space, yielding

$$\langle\psi_0^*, \delta L_0\psi_0\rangle + \langle\psi_0^*, L_0\delta\psi_0\rangle - \langle L_0^*\psi_0^*, \delta\psi_0\rangle = \langle\psi_0^*, \delta Q_0\rangle - \langle\delta\psi_0, \Sigma_d\rangle\,. \tag{16.18}$$

The second and third terms on the left hand side cancel by way of the adjoint identity, and the second term on the right hand side is δR_0. Thus, we have

$$\delta R_0 = \langle\psi_0^*, \delta Q_0\rangle - \langle\psi_0^*, \delta L_0\psi_0\rangle\,. \tag{16.19}$$

From Eq. 16.19, we see that an increased source gives rise to a *greater* response, and an increase in L, corresponding to greater leakage or interaction, produces a *lower* response, as is expected.

Perturbation Theory — Eigenvalue

As another example of perturbation theory, consider the unperturbed eigenvalue problem

$$H_0\psi_0 = \lambda_0 F_0\psi_0 \tag{16.20}$$

subject to vacuum conditions. As above, suppose we introduce some small change, and let a new, perturbed system be

$$H\psi = \lambda F\psi \,. \tag{16.21}$$

where $H = H_0 + \delta H_0$, $F = F_0 + \delta F_0$, $\psi = \psi_0 + \delta\psi_0$, and $\lambda = \lambda_0 + \delta\lambda_0$. Our goal is to find $\delta\lambda_0$ due to the perturbation.

We rewrite the perturbed system

$$(H_0 + \delta H_0)(\psi_0 + \delta\psi_0) = (\lambda_0 + \delta\lambda_0)(F_0 + \delta F_0)(\psi_0 + \delta\psi_0) \tag{16.22}$$

and expand

$$\begin{aligned} H_0\psi_0 + \delta H_0\psi_0 + H_0\delta\psi_0 + \theta(\delta^2) \\ = \lambda_0 F_0\psi_0 + \lambda_0 F_0\delta\psi_0 + \lambda_0\delta F_0\psi_0 + \delta\lambda_0 F_0\psi_0 + \theta(\delta^2)\,. \end{aligned} \tag{16.23}$$

Again, we recognize our original equation and a second equation with first-order perturbations,

$$(H_0 - \lambda_0 F_0)\delta\psi_0 = (\lambda_0\delta F_0\psi_0) - (\delta H_0 - \delta\lambda_0 F_0)\psi_0\,. \tag{16.24}$$

We define the adjoint problem

$$H_0^*\psi_0^* = \lambda_0 F_0^*\psi_0^* \tag{16.25}$$

subject to the appropriate boundary conditions. Similar to our treatment in the fixed source example, we multiply Eq. 16.24 by ψ_0^* and Eq. 16.25 by $\delta\psi_0$, subtract the latter from the former, and integrate over all phase space. After a bit of rearranging, we find

$$\delta\lambda_0 = \frac{\langle\psi_0^*, (\delta H_0 - \lambda_0\delta F_0)\psi_0\rangle}{\langle\psi_0^*, F_0\psi_0\rangle}\,. \tag{16.26}$$

Further Reading

Most of the development here follows that of Bell and Glasstone [9], Chapter 6. Duderstadt and Hamilton [17] develop the adjoint within the diffusion framework and apply it to problems of reactor physics in Chapters 5 and 7. A particularly appealing description of the physical interpretation of the adjoint, albeit with a reactor physics flavor, is given by Henry [22].

It's worth noting that the adjoint was developed first by Lagrange as a mathematical construct; however, its physical utility was first realized much later in the context of quantum mechanical perturbation theory, and later yet in reactor physics. This

history and more is to be found in Marchuk's treatise on adjoint methods [32]. That first application of the adjoint in reactor physics was due to the "father of nuclear engineering," Eugene Wigner [46].

The available literature on perturbation theory is quite large. One important recent effort has been to couple sensitivities defined by perturbation theory to cross-section uncertainties in order to estimate the uncertainty of integral system parameters including the eigenvalue [10] and various worths [48] due to the underlying data uncertainty.

Exercises

1. **Self-adjointness**. Prove the one-speed diffusion operator, i.e. $L\phi = D\phi_{xx} - \Sigma_a\phi(x) = -S$ is self-adjoint. You may assume a homogeneous medium with zero-flux boundary conditions, neglecting extrapolation distances.
2. **Adjoint Transport Equation**. Demonstrate that the adjoint operator L^* defined by Eq. 16.5 really is the adjoint to the forward operator L for the given boundary conditions.
3. **Adjoint Boundary Conditions**. (a) For the case that ψ satisfies vacuum conditions, we found that $\psi^*(\mathbf{r}, \mathbf{\Omega}, E) = 0$ for $\hat{n} \cdot \hat{\Omega} > 0$. What does this mean physically? (b) For the one-speed transport equation, derive the boundary conditions for ψ^* when ψ satisfies reflecting conditions.
4. **Using the Adjoint**. (a) Briefly describe the physical meaning of the adjoint flux. (b) Suppose we have a known shield with a known detector on one side. Suppose further that the particle source on the opposite side of the shield is not known *a priori* and can take widely varying forms. (An example of this might be a shielding analysis for a fusion reactor, where we think we have a good shield and then we try using it for several possible sources). How could the adjoint be used so that only one "transport" calculation would be needed to compute the detector dose given an arbitrary source?
5. **A Point Detector**. Repeat the process used to obtain $\psi^* = R$ but for the case of a point detector, $\Sigma_d = \delta(\mathbf{r} - \mathbf{r}_0)\delta(\mathbf{\Omega} - \mathbf{\Omega}_0)\delta(E - E_0)$. Clearly, one obtains an expression for $\psi(\mathbf{r}_0, \mathbf{\Omega}_0, E_0)$. What does $\psi^*(\mathbf{r}, \mathbf{\Omega}, E)$ represent in this case? This result is a generalization of reciprocity relations previously discussed for one-speed transport.
6. **The "Contributon" Flux**. In 1-d and one-speed, define the quantity

$$C(x, \mu) = \psi(x, \mu)\psi^*(x, \mu)\,,$$

where ψ and ψ^* are the forward and adjoint angular fluxes. Take the forward problem to have vacuum boundaries. (a) Mathematically, express the vacuum

boundary condition for ψ at a general external boundary x_b. (b) With your knowledge of the corresponding adjoint boundary condition, write down the mathematical expression for the boundary condition of $C(x, \mu)$. (c) Taking the adjoint source to be a flux-to-dose factor, what are the units of ψ, ψ^*, and C? Can you interpret C physically? (For more information on this mysterious quantity, see the paper by Williams [47].)

7. **Eigenvalue Perturbation**. Prove Eq. 16.26. Also, describe what possible changes in the system coincide with the perturbations δH_0 and δF_0 and how such changes impact the eigenvalue perturbation. Remember that k_{eff} is λ^{-1}.
8. **Eigenvalue Sensitivity**. Defining the sensitivity of k_{eff} to a cross-section Σ_x to be

$$S_{k,\Sigma_x} = \frac{\Sigma_x}{k_{eff}} \frac{\partial k_{eff}}{\partial \Sigma_x} ,$$

find an expression for S_{k,Σ_x} in terms of the partial derivatives of H_0 and F_0 with respect to Σ_x.

Lecture 17 | Variational Methods

In this lecture, we investigate use of *variational principles* in the context of nuclear engineering. An entire course could be devoted to this subject; here, we focus on the same two quantities analyzed in Lecture 16, namely the reaction rate for a fixed source problem and the eigenvalue of an eigenvalue problem. We demonstrate that variational principles can be used to estimate these quantities using approximate fluxes. We finish by showing how our first order perturbation theory can be derived directly from variational principles.

Variational Principles, Functionals, and Stationary Points

A variational principle casts a particular function, usually a problem's solution, as the stationary point of some *functional*. Often, the functional itself is called the variational principle. A functional is simply a function that takes another function as its argument and returns a scalar as its value. Consider a function $f(x) = Ax + B$. A possible functional would be $F[f(x)] = \int_{x_1}^{x_2} f(x)dx$, which certainly yields a scalar value. Typically, the value of the functional represents a quantity of interest, which in transport applications is often a reaction rate.

We see that functionals are quite like functions; how exactly then do we define a *stationary point*, and what does it mean in the context of a variational principle? Recall from elementary calculus that a stationary point is the value of the independent variables such that the function reaches a local extremum (or saddle-point), i.e. the function's derivative (or gradient) vanishes. The same idea applies to functionals. Defining the "weak derivative" of F at a point $f(x)$ in the direction $g(x)$ as

$$\delta F[f, g] = \lim_{\epsilon \to 0} \frac{F[f + \epsilon g] - F[f]}{\epsilon}, \tag{17.1}$$

the *first variation* of F is defined

$$\delta F[f, g] = \left(\frac{d}{d\epsilon} F[f + \epsilon g] \right) \Bigg|_{\epsilon=0}, \tag{17.2}$$

for arbitrary g. When $\delta F[\tilde{f}, g] = 0$ for all g, $\tilde{f}$ is called a *stationary point* of F, and the expression

$$\delta F[\tilde{f}, g] = 0\,, \quad \forall g \tag{17.3}$$

defines the variational principle for $\tilde{f}$. For $F[f]$ that represents a quantity of interest, $F[\tilde{f}]$ represents the true value for that quantity. Moreover, very near the stationary point, $\delta F \to 0$ by construction, and the errors in F (and hence the quantity of interest) are second order, which gives rise to the utility of variational approximations.

A Simple, Illustrative Example

It is easiest to understand these ideas through a simple example (unrelated to nuclear engineering). Suppose we wish to find the curve giving us the shortest difference between two points in a plane Of course, this is intuitive: the solution should be a line. We show this using variational techniques.

Let the curve be $f(x)$. From any calculus book, we know the differential arc length is then

$$dl = \sqrt{dx^2 + df^2} = \sqrt{1 + (f')^2}\,. \tag{17.4}$$

We take as our functional the arc length,

$$F[f] = \int_{x_1}^{x_2} \sqrt{1 + (f')^2} dx\,, \tag{17.5}$$

where (x_1, y_1) and (x_2, y_2) are the points of interest, as illustrated in Figure 17.1.

Taking the first variation of F,

$$\begin{aligned}
\delta F[f, g] &= \left(\frac{d}{d\epsilon} \int_{x_1}^{x_2} \sqrt{1 + (f'(x) + \epsilon g'(x))^2} dx \right) \Bigg|_{\epsilon=0} \\
&= \left(\int_{x_1}^{x_2} \frac{(f'(x) + \epsilon g'(x)) g'(x)}{\sqrt{1 + (f'(x) + \epsilon g'(x))^2}} \right) \Bigg|_{\epsilon=0} \\
&= \int_{x_1}^{x_2} \frac{f'(x) g'(x) dx}{\sqrt{1 + (f'(x))^2}}\,,
\end{aligned} \tag{17.6}$$

we find that the arc length is minimized when

$$\delta F[y, g] = \int_{x_1}^{x_2} \frac{f'(x) g'(x)}{\sqrt{1 + (f'(x))^2}} = 0\,. \tag{17.7}$$

In this form, we can't say anything about f or g. Instead, we use integration by parts to go from f' and g' to f'' and g, and Eq. 17.7 becomes

$$\delta F[f, g] = \left[\frac{f'(x) g(x)}{\sqrt{1 + (f'(x))^2}} \right]_{x_1}^{x_2} - \int_{x_1}^{x_2} \frac{f''(x) g(x) dx}{(1 + (f'(x))^2)^{\frac{3}{2}}} = 0\,. \tag{17.8}$$

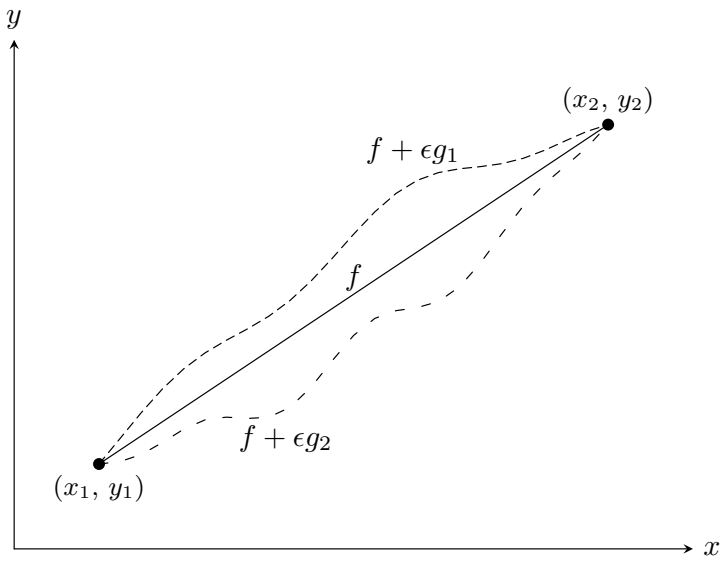

Figure 17.1: The shortest path between two points $(x_1,\ y_1)$ and $(x_2,\ y_2)$ is indicated by the solid curve f. Two possible curves for arbitrary g are also given.

While $g(x)$ is arbitrary *between* x_1 and x_2, we must have $g(x_1) = g(x_2) = 0$ to that $f(x_1) = y_1$ and $f(x_2) = y_2$. Thus, the first term of Eq. 17.8 vanishes. By inspection, the second term vanishes for arbitrary g only if $f'' = 0$. This relation is known as the *Euler equation*[*]. Of course, to satisfy $f'' = 0$ requires our solution to be of the form $Ax + B$, as expected.

A Variational Principle for Fixed Source Problems

Suppose we are interested in a linear functional of the flux, such as $G_{\text{fs}}[\psi] = \langle \psi, \Sigma_d \rangle$, a reaction rate. An appropriate variational principle is represented by the *generalized Roussopoulos functional*

$$F_{\text{fs}}[\psi, \psi^*] = G_{\text{fs}}[\psi] + \langle \psi^*, (Q - L\psi) \rangle \,. \tag{17.9}$$

Note, $\delta F_{\text{fs}}[\psi, \psi^*] = 0$ is a variational principle for G_{fs} if the corresponding solution ψ yields $F_{\text{fs}}[\psi, \psi^*] = G_{\text{fs}}[\psi]$. We see this is so when the second term of F vanishes when ψ solves $L\psi = Q$, i.e. when ψ is the solution to the transport equation.

[*]In general, setting the first variation to zero gives rise to a set of partial differential equations collectively known as the Euler equations.

To determine the variational principle, we take the first variation of F_{fs},

$$\begin{aligned}
&\delta F_{\text{fs}}[\phi, \phi^*] \\
&\quad = \left(\frac{d}{d\epsilon} \left(\langle (\phi + \epsilon\delta\phi)\Sigma_d \rangle + \langle (\phi^* + \epsilon\delta\phi^*), (Q - L(\phi + \epsilon\delta\phi)) \rangle \right) \right) \Bigg|_{\epsilon=0} \\
&\quad = \langle \delta\phi, \Sigma_d \rangle - \langle \phi^*, L\delta\phi \rangle + \langle \delta\phi^*, Q \rangle - \langle \delta\phi^*, L\phi \rangle + \mathcal{O}(\epsilon^2) \,.
\end{aligned} \tag{17.10}$$

Noting $\langle \phi^*, L\delta\phi \rangle = \langle L^*\phi^*, \delta\phi \rangle$, we have for our principle

$$0 = \langle \delta\phi, (\Sigma_d - L^*\phi^*) \rangle + \langle \delta\phi^*, (Q - L\phi) \rangle \,, \tag{17.11}$$

which is satisfied when $L\phi = Q$ and $L^*\phi^* = \Sigma_d$. These are the corresponding Euler equations, and we see they are just the original forward and adjoint transport equations.

The importance of this variational principle (and others in general) is that it gives us an estimate for G_{fs} accurate to second order for approximate fluxes ϕ and ϕ^*. To demonstrate this, suppose the true solutions to the Euler equations are ψ and ψ^*. Let $\phi = \psi + \delta\psi$ and $\phi^* = \psi^* + \delta\psi^*$. We subsitute these expressions into F and find

$$\begin{aligned}
F_{\text{fs}}[\phi, \phi^*] &= \langle \Sigma_d, (\psi + \delta\psi) \rangle + \langle (\psi^* + \delta\psi^*), (Q - L(\psi + \delta\psi^*) \rangle \\
&= \langle \Sigma_d, \psi \rangle + \langle \Sigma_d, \delta\psi \rangle + \langle \psi^*, Q \rangle + \langle \delta\psi, Q \rangle \\
&\quad - \langle \psi^*, L\psi \rangle - \langle \psi^*, L\delta\psi \rangle - \langle \delta\psi^*, L\psi \rangle - \langle \delta\psi^*, L\delta\psi \rangle \\
&= \langle \Sigma_d, \psi \rangle - \langle \delta\psi^*, L\delta\psi \rangle \\
&= G_{\text{fs}}[\psi] + \mathcal{O}(\delta^2) \,.
\end{aligned} \tag{17.12}$$

Hence, F provides a first order accurate (i.e. good through first order) estimate of the reaction rate given approximate forward and adjoint fluxes.

A Variational Principle for Eigenvalue Problems

For eigenvalue problems, an appropriate functional is the *Rayleigh quotient*

$$F_{\text{ev}}[\psi, \psi^*] = \frac{\langle \psi^*, L\psi \rangle}{\langle \psi^*, F\psi \rangle} \,. \tag{17.13}$$

Here, the quantity of interest is $G_{\text{ev}} = \lambda$, and, unlike the fixed source case, we have $G_{\text{ev}} = F_{\text{ev}}$. You should show that F_{ev} is in fact a valid expression for λ. The associated Euler equations are just the forward and adjoint eigenvalue equations, which can be shown by setting the first variation of F_{ev} to zero, an exercise left to the reader. For approximate fluxes ϕ and ϕ

$$F_{\text{ev}}[\phi, \phi^*] = \lambda + \mathcal{O}(\delta^2) \,, \tag{17.14}$$

the proof of which is also left as an exercise.

Perturbation Theory from Variational Principles

In general, it is possible to find first order perturbation estimates using the expression

$$\delta = \bar{F}[\psi, \psi^*] - F[\psi, \psi^*] \,, \tag{17.15}$$

where F is the appropriate functional with nominal operators and $\bar{F}$ is the function evaluated with perturbed operators. It must be stressed that ψ and ψ^* are assumed to be the exact fluxes for the unperturbed problem.

As an example, we re-derive the first order perturbation to a detector response. Suppose the perturbations to our system include $L + \delta L$, $Q + \delta Q$, and $\Sigma_d + \delta\Sigma_d$. Then Eq. 17.15 gives

$$\begin{aligned} \delta_{\text{fs}} &= \bar{F}[\psi, \psi^*] - F[\psi, \psi^*] \\ &= \langle \psi, (\Sigma_d + \delta\Sigma_d)\rangle + \langle \psi^*, (Q + \delta Q - (L + \delta L)\psi)\rangle - \langle \psi, \delta\Sigma_d\rangle \\ &= \langle \psi, \delta\Sigma_d\rangle + \langle \psi^*, \delta Q\rangle - \langle \psi^*, \delta L\psi\rangle \,. \end{aligned} \tag{17.16}$$

The last line is equivalent to Eq. 16.19 with the addition of the first term, which explicitly accounts for changes in the detector response function.

Eq. 17.15 can also be applied to the Rayleigh quotient, yielding Eq. 16.26, proof of which is left as an exercise.

Further Reading

Most of the material presented here is contained in Chapter 7 of Duderstadt and Martin [18] in addition to Chapter 1 of Stacey [42]. The former contains several examples and a variational derivation of the diffusion equation. The reader is also encouraged to look up Pomraning's large body of work on variational methods[†].

Exercises

1. **Second Order Accuracy**. Show that the Roussopoulos functional $F[\phi, \phi^*]$ gives a second order accurate value for $\langle \psi, \Sigma_d\rangle$ when ϕ and ϕ^* are approximate values of ψ and ψ^*, i.e. prove Eq. 17.12.
2. **Fixed Source Perturbation**. Prove Eq. 17.16.
3. **Rayleigh Quotient**.

[†]It is worth noting that Stacey and Pomraning, both with prolific work in variational methods, did their graduate work in this department.

a. Take the first variation of the Rayleigh quotient with respect to the forward and adjoint fluxes and show the stationarity conditions are just the forward and adjoint eigenvalue problems.

b. Demonstrate that the Rayleigh quotient is a second order estimator for λ.

4. **Losses-to-Gains**. Directly from the eigenvalue equation, we find

$$\lambda = (L\psi)/(F\psi)\,,$$

a compact way to define losses-to-gains.

a. Show why this is not a variational principle for λ.

b. Consider again Eq. 17.13. How does the adjoint change the physical interpretation of gains-to-losses?

5. **Eigenvalue Perturbation**. Prove that $\bar{F}_R[\psi, \psi^+] - F_R[\psi, \psi^+]$ yields the first order perturbation estimate for $\delta\lambda$ given by Eq. 16.26.

6. **Applying Roussopoulos**. Consider a 1-d, 1-speed diffusion problem in a slab of width $2a$, $\Sigma_a = 0.022$, and $D = 0.14$.

a. Solve for the exact scalar flux $\phi(x)$, using the conditions $\phi(\pm a) = 0$, and assuming a uniform source $Q(x) = 1$ and $a = 1$.

b. Compute the total absorption rate in the slab.

c. Now, approximate the solution as $\tilde{\phi}(x) = Ax^2 + Bx + C$ with the same maximum and the same boundary conditions. Compute the absorption rate using $\tilde{\phi}(x)$.

d. Finally, recalling the 1-d, 1-speed diffusion equation is self-adjoint, use the Roussopoulos principle to compute the absorption rate. What can you conclude?

7. **More Roussopoulos**. For the same problem, estimate the total absorption rate if something causes $\Sigma_a = 1.1$. Be careful to consider the impact of the change in Σ_a on both Σ_d and L.

8. **Approximating Escape from a Slab**. Recall the escape probability P_{esc} for a homogeneous 1-d slab, the subject of Chapter 5, Problem 8. Here, consider a discrete ordinates approximation using the two point Gauss-Legendre quadrature and Mark boundary conditions for a 10 cm slab with $\Sigma_t = 10$ [1/cm].

a. Solve for P_{esc} directly using the S_2 approximation.

b. Use a variational principle to improve your estimate of P_{esc}.

c. Compare both values to the exact value.

Lecture 18 | Computational Geometry for Transport

Quadratic Surfaces

A 2-D quadratic (or second-order) surface is defined implicitly as

$$f(x,y) = Ax^2 + By^2 + Cxy + Dx + Ey + F\,. \tag{18.1}$$

For points *on the surface*, $f = 0$, and for points outside/inside the surface, $f \gtrless 0$. Let

$$\mathbf{r} = [x, y, 1]^\mathsf{T} \quad \text{and} \quad \mathbf{M} = \begin{bmatrix} 2A & D & E \\ D & 2B & C \\ E & C & 2F\,. \end{bmatrix}$$

Then,

$$f(x,y) = \frac{1}{2}\mathbf{r}^\mathsf{T}\mathbf{M}\mathbf{r}\,. \tag{18.2}$$

Prove this to yourself!

Tracking Particles

Common problem: *where does a ray intersect a surface?*

Let a ray $\mathbf{r}$ be defined as

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{d} \tag{18.3}$$

where $\mathbf{r}_0$ is some starting point, $\mathbf{d}$ is some direction (so $|\mathbf{d}| = 1$), and t is the distance from the starting point along the direction.

Substitute this ray into Eq. (**??**) to find where the ray intersects the surface (and, hence, $f = 0$). The result is (*show this!*)

$$t^2 \overbrace{\mathbf{d}^\mathsf{T}\mathbf{M}\mathbf{d}}^{a} + t \overbrace{\mathbf{r_0}^\mathsf{T}\mathbf{M}\mathbf{d}}^{b} + \overbrace{\mathbf{r_0}^\mathsf{T}\mathbf{M}\mathbf{r_0}}^{c} = 0\,, \tag{18.4}$$

which is quadratic in t. If

$$\begin{array}{ll} b^2 > 4ac & \text{there are two intersections} \\ b^2 = 4ac & \text{there is one intersection (tangent)} \\ b^2 < 4ac & \text{there are no intersections} \end{array} .$$

Creating Geometries

Detran's computational geometry consists of *regions* defined by combinations of *nodes*.

A *node* can be a surface (called a *primitive*), an *operator* between two nodes (e.g., *union* and *intersection*), or a *translation* of another node.

It would be useful to simplify what is presented here using Python (but still without function implementations). Ultimately, all of the note presented so far will be typeset as formal lectures, and I'd like one solid set of Python routines to form the core.

Lecture 19 | Criticality Safety

In this lecture, we discuss the topic of *criticality safety*, one of the most important physics-oriented aspects of nuclear engineering not specifically dealing with core physics, beginning with a brief overview. Thereafter, the chief physical concerns related to criticality safety are described and illustrated with a simple example, and a few notable accidents over the past several decades are described. Finally, computational aspects of criticality safety are discussed, and the topic of burnup credit is used as a case study to illustrate recent efforts in criticality safety analysis.

Overview

Nuclear criticality safety, as defined by the American National Standard ANSI/ANS 8.1 [**?**], is the "protection against the consequences of a criticality safety accident, preferably by prevention of the accident," and it defines a *criticality accident* as "the release of energy as a result of accidental production of a self-sustaining or divergent neutron chain reaction."

ANSI/ANS 8.1 provides general guidance for nuclear criticality safety applied to fissionable materials *outside* of reactors. In addition to ANSI/ANS 8.1, a number of more specialized standards exist, a few of which are summarized in Table 19.1. The interested student can request these standards from the library, but they tend to be expensive (and, admittedly, rather dry reading). ANSI/ANS 8.1 has been put on the 22.106 Stellar site.

The Nuclear Regulatory Commission (NRC) also offers guidance with respect to criticality safety. NRC documents are most often found in the NUREG series, published by the NRC itself, by contractors, or through international agreements.

Additionally, the NRC provides *regulation* through relevant portions of the Code of Federal Regulations (CFR). Both DOE and NRC share regulations under title 10, *i.e.* those regulations beginning with 10CFR.* The parts of 10CFR most relevant to

*That DOE and NRC share the same title is most likely due their common origin: the Atomic Energy Commission.

Table 19.1: Several ANSI/ANS Standards Applicable to Criticality Safety.

Number-Revised	Title
8.1-1998	Nuclear Criticality Safety in Operations with Fissionable Materials Outside Reactors
8.3-1997	Criticality Accident Alarm System
8.5-1996	Guide for Nuclear Criticality Safety in the Storage of Fissile Materials
8.17-2004	Criticality Safety Criteria for the Handling, Storage, and Transportation of LWR Fuel Outside Reactors
8.24-2007	Validation of Neutron Transport Methods for Nuclear Criticality Safety Calculations
8.27-2008	Burnup Credit for LWR Fuel

criticality safety are 10CFR-0, 1, 2, 20, and several between 50 and 75. As an exercise, the student is encouraged to find one or more of these regulations (or NUREG documents) related to criticality safety and explain its relevance.

As a historical note, nuclear criticality safety as a discipline is about as old as other nuclear engineering areas; it began with the large scale chemical processing at the K-25 gaseous diffusion plant in Oak Ridge, TN and the Hanford, WA site, both major components of the Manhattan Project. Of course, the initial work at Los Alamos generated much knowledge before the larger scale projects were begun, and in fact, it was a young Richard Feynman who carried much of that experience from Los Alamos to Oak Ridge in 1940's at the bequest of Oppenheimer [**?**]. In Feynman's own words, he was told by Oppenheimer to tell the Oak Ridge folks (stubborn military types), "Los Alamos cannot accept the responsibility for the safety of the Oak Ridge plant unless they are full informed as to how it works." He delivered the message, and when they agreed to listen, he "told them all about neutrons, how they worked, da da, ta ta ta, there are too many neutrons together, you've got to keep the material apart, cadmium absorbs . . . " and so forth. The Oak Ridge folks went back to the design board, and a "practice" was born.

Physical Concerns

When we analyze a system for criticality safety, what are the important characteristics the system? An analyst must understand *what* to analyze before computational techniques become useful.

Several features can be intuited by any student of reactor physics: the more fissile material one has, the easier it is to achieve criticality. That means increased *mass*, *concentration*, *enrichment*, or *volume* of a fissile material should bring a system closer to criticality (or past it, unfortunately).

But there are other factors: what neutrons do we like in our typical light water reactors? Thermal ones, of course, and to get those, we need *moderation*. Moreover, those same reactors feature a layer of water around the periphery, which *reflect* neutrons back into the core. To reduce power in a reactor (or to shut it down), we insert some level of *absorption* via control rods, which limits the *interaction* of fissile assemblies with one another.

These key characteristics, easily remembered via the acronym MAGICMERV [25], are equally applicable to nuclear criticality safety. To illustrate several of the characteristics, consider Figure 19.1, which shows the critical Pu-239 mass in water solution as a function of Pu concentration. As an exercise, describe why the critical mass curve behaves as it does and what characteristics described above are involved.

Accidents

What exactly *is* a criticality accident? In the simplest terms, it is an excursion, an uncontrolled chain reaction. To help understand accidents both qualitatively and quantitatively, let's review a few ideas from reactor kinetics. Recall that *reactivity* measures a departure from criticality, and is most often expressed via

$$\rho = \frac{k-1}{k}\,. \tag{19.1}$$

Positive ρ denotes a supercritical state, negative ρ a subcritical state, and $\rho = 0$ is equivalent to $k = 1$ or a critical state. We can further differentiate between a *delayed critical* state and a *prompt critical* state. The latter occurs when $\rho > \beta$, where β is the delayed neutron fraction. This situation is to be avoided in all but a few specific experimental situations, for in the prompt critical regime, the growth of the neutron population occurs on timescales close to the prompt neutron lifetime (say 10^{-8} to 10^{-4} seconds in many systems of interest).

To get an idea of the orders of magnitude we deal with in an accident situation, consider a critical system with a constant neutron population of just one neutron. Suddenly, a perturbation brings the system into a prompt critical state for a short time Δt. During this time, the population grows as $n \propto n_o e^{(\rho-\beta)t/\Lambda}$. Assuming $\Lambda = 10^{-5}$ seconds, $\rho - \beta = 0.002$, and the perturbation lasts two tenths of second[†],

[†]the average human reaction time is about 200 milliseconds. See `http://www.humanbenchmark.com/tests/reactiontime/stats.php`

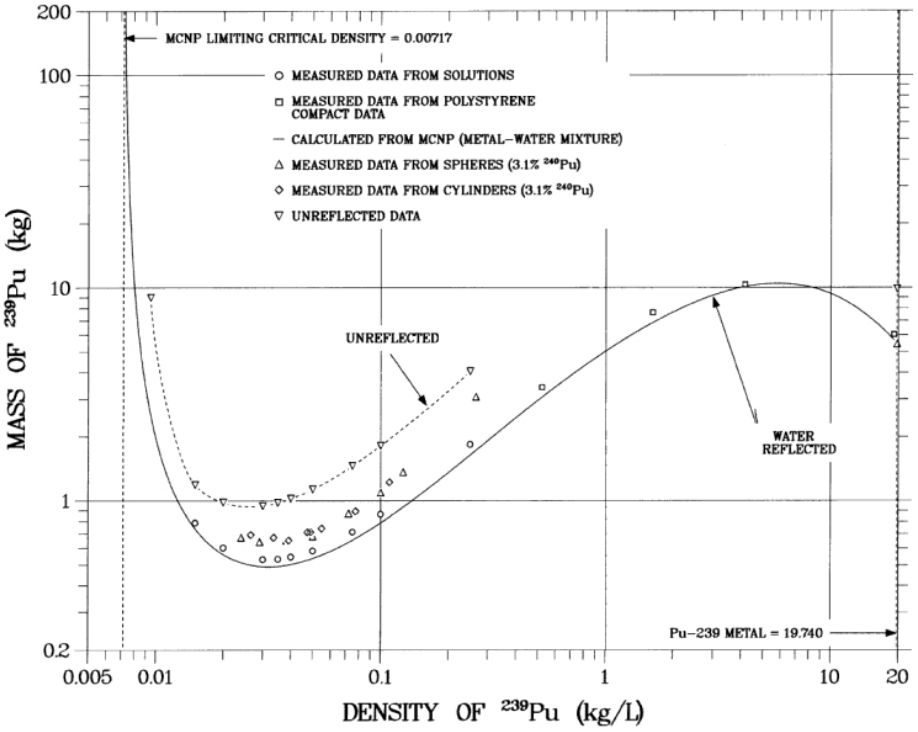

Figure 19.1: Pu-239 solution critical mass. (Borrowed from Martin, J. Physics for Radiation Protection: A Handbook, Wiley (2016).

we estimate that the number of neutrons produced is roughly

$$\begin{aligned} N &= \int_0^{0.5} dt e^{(\rho-\beta)t/\Lambda} \\ &= \frac{10^{-5}}{0.002}(e^{40}-1) \\ &\approx 1 \cdot 10^{15}\,. \end{aligned} \tag{19.2}$$

Suppose a worker is about a half a meter from the neutron source, so that the fluence is $\Phi \approx 10^{15}/(4 \cdot \pi \cdot 50^2)$ n/cm^2. Looking up neutron fluence-to-dose equivalent factors (see 10CFR-20 [‡]) suggests that a high energy (1 MeV) fission fluence of $27 \cdot 10^6$ n/cm^2 corresponds to a dose of 1 rem. Hence, our excursion yields a dose of

[‡] http://www.nrc.gov/reading-rm/doc-collections/cfr/part020/part020-1004.html

around 1200 rem (12 Sv). For perspective, a dose of 4 Sv is lethal roughly half the time. Of course, this is a crude example, but it gives a very clear picture of the issues at hand.

In fact, the example above is not too far different than one of the first documented criticality accidents from August of 1945. Figure 19.2 shows a 6.2 kg plutonium sphere coated with nickel and surrounded by blocks of tungsten carbide for reflection. In the accident, an experimenter was placing blocks to achieve criticality. While placing the last block, the detector reading suggested that the block would actual produce a supercritical state. Unfortunately, the experimenter dropped the brick onto the assembly, yielding a prompt supercritical excursion of 10^{16} fissions before he was able to remove the brick. A later study estimated the resulting dose to be 510 rem, which proved to be fatal some 28 days later. The same assembly would be involved in a second fatal accident just months later, where the experimenter handled one of two beryllium hemisphere reflectors with his thumb (in a hole in the hemisphere) and a screwdriver, a procedure Feynman dubbed "tickling the dragon's tail." The screwdriver, holding the reflectors apart, slipped and caused an excursion that led to the experimenter's death 9 days later and significant doses to observers; see Figure 19.3 for a recreation of the experiment.

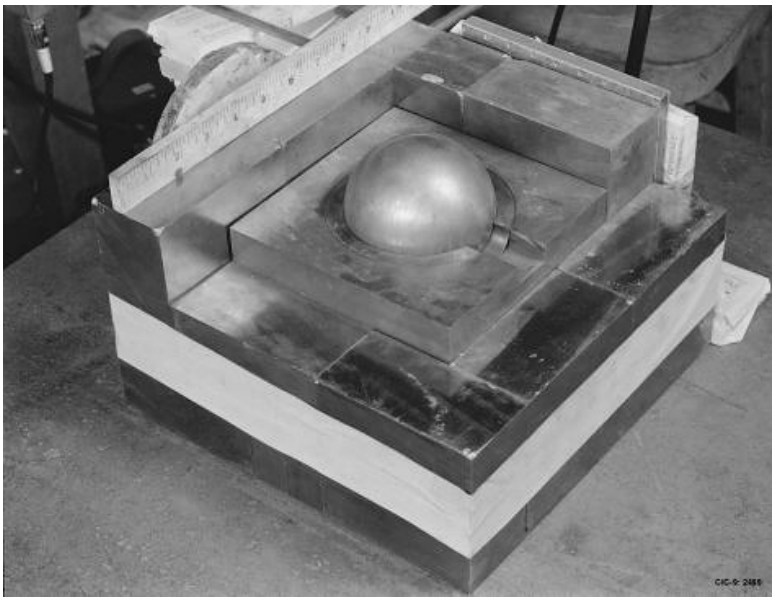

Figure 19.2: Plutonium sphere reflected by tungsten-carbide bricks.

A second type of accident involves processing operations in which fissile materi-

Figure 19.3: Recreation of the “tickling the dragon’s tail” experiment.

als are in solution. The first domestic processing accident occurred at the Y-12 plant in Oak Ridge, TN. Y-12 produces parts from HEU for use in nuclear weapons.

The accident occurred in a portion of the plant used to recover HEU from waste material and place it in tanks of favorable geometry. These tanks were to be emptied, cleaned, and leak-tested with water. Before the leak-testing began, however, uranium solution had leaked from the process stream into the piping below the storage tanks (and actually into one of the tanks, as its release valve had been let open). When the other tanks, full of water, were emptied, the uranium solution and water accumulated in a 55 gallon drum. A nearby worker first noticed dark yellow fumes followed by a blue flash. Eight workers received significant doses, though none died as a result of acute effects.

Unfortunately, criticality accidents are not a thing of the past. A recent example happened in Tokai-mura, Japan at a uranium conversion plant. The accident occurred when workers placed 16.6 kg of 18.8% enriched uranium into a tank designed to contain no more than 2.4 kg of uranium at such high enrichments. The accident was the first of its kind in Japan, and the first resulting in casualties. The accident is somewhat unique in that criticality was maintained intermittently for roughly 20 hours before effective action was taken.

A rather complete history of criticality accidents in the U.S. and around the world is contained in the latest edition of *A Review of Criticality Accidents* from

Los Alamos. This document is a really great piece of nuclear history, and students are highly encouraged to skim through the many accidents covered.

Computational Aspects

The intent of this section to provide the reader with an overview of criticality safety analysis validation. A brief review of regulations and guidance pertinent to criticality safety of fissile materials outside reactors is given, with a particular emphasis on requirements for ensuring subcritical conditions. The traditional approach to bias determination is discussed, and one specific method used in practice is described and demonstrated in an illustrative example. Subsequently, modern S/U-based validation techniques are discussed, and an illustrative example is provided and related to the traditional approach.

Subcritical Limits

As noted above, ANSI/ANS-8.1 provides guidance for avoiding criticality accidents during handling of fissionable material outside of reactors [2]. The standard provides basic safety principles in addition to several limits for simple systems of single isotopes. Specifically, the standard defines a *subcritical limit* to be "the limiting value assigned to a controlled parameter that results in a subcritical system under specified conditions. The parameter limit allows for uncertainties in the calculations and experimental data used in the derivation but not for contingencies."

These limits are *absolute maxima*, and since they do not include contingencies, in practice, administrative margins are employed. A typical value, as we'll see below, tends to be 5% on k_{eff} .

Table 19.2 gives several examples of single control parameters used in criticality safety analysis. Note, while ANSI/ANS 8.1 specifies limits in terms of a single parameter, in some cases, multiple parameter limits are also employed. The standard suggests a few examples, and cites the technical literature for further guidance. However, while less conservative, multiple parameter limits require additional administrative margins and may be harder to use in validation.

Criticality Safety Analysis Validation

While the single parameter limits provide useful guidance, two questions naturally arise. First, how does one actually establish these limits? And second, how does one ensure subcriticality in systems that are much more complex than, for example, a sphere? The answer in both cases is by using computational methods validated against experiment.

Table 19.2: Example single parameters and subcritical limits.

parameter	example limit
fissile mass	0.78 kg ^{235}U in UO_2NO_3
dimension (width, volume, etc.)	6.2 L ^{235}U in UO_2NO_3
concentration	11.6 g/L^{235}U in UO_2NO_3
enrichment	5% ^{235}U in UO_2
fissile mass	20.1 kg ^{235}U in UO_2 (w/ $\rho \leq 18.81$ g/cc)

In addition to the single parameter limits, ANSI/ANS-8.1 provides requirements for ensuring computational methods used in criticality safety analyses are both valid and within the "area of applicability" for specific applications. The standard defines the area of applicability as "limiting ranges of material compositions, geometric arrangements, neutron energy spectra, and other relevant parameters ... within which the bias of the calculational method is established" [2]. In other words, an *application* is the system of interest, such as a spent fuel canister for shipment and disposal. A computational method is *applicable* if it is verified against a set of *experiments* that are similar to the application with respect to composition, geometry, and so forth.

A *computational bias* is the systematic discrepancy between and experimental data and calculated results. Biases have associated uncertainties, which quantify the accuracy and precision of calculated values and the uncertainty in measured data.

When computational tools are used in criticality safety analyses, the standard requires that the bias be established. Qualitatively, the bias must be determined via correlating data from critical experiments to computational models of the same experiments using the tool to be validated. Typically, the measured and computed values of k_{eff} are related, but other physical quantities may also be used. The bias is used to normalize a particular code within its area of applicability so that its results, after applying the bias, may be used to predict criticality within the bias uncertainty.

Another American National Standard, ANSI/ANS-8-17, further explicates use of computational tools by defining specific criteria for establishing subcriticality [1]. Whenever computational methods are used in criticality safety analyses, the standard requires that the calculated application multiplication factor k_a must be less than or equal to the allowable multiplication factor.

The easiest way to think of this is to note that the largest, "worst case" value for k_a must be below the smallest, least conservative computed estimate. That is to say, the maximum expected application multiplication factor (*i.e.* $k_a + \Delta k_a$) must be less than the minimum expected computed value (*i.e.* $k_c - \Delta k_c$) for an applicable critical

experiment, *i.e.* a real, measured system whose k_{eff} is unity, or an average computed k_{eff} for several such experiments. Mathematically, this can be written as

$$k_a + \Delta k_a \leq k_c - \Delta k_c \,, \tag{19.3}$$

or

$$k_a \leq k_c - \Delta k_a - \Delta k_c \,. \tag{19.4}$$

For added conservatism, the standard further requires that

$$k_a \leq k_c - \Delta k_a - \Delta k_c - \Delta k_m \,, \tag{19.5}$$

where the standard defines:

k_a is the calculated k_{eff} of the application system for all normal or credible abnormal conditions;

k_c is the average k_{eff} from the calculation of the benchmark criticality experiments. The experiments used should be neutronically similar to the application system. If the application system has parameters beyond the area of applicability established by the benchmark experiments, then the area of applicability may be extended using trends in the calculated values of k_c as functions of those parameters;

Δk_a is an allowance for

- statistical or convergence uncertainties in the computed k_a;
- material and fabrication tolerances;
- uncertainties due to geometry or material simplifications and approximations;

Δk_c is a margin for uncertainty in k_c that includes allowances for

- uncertainties in the critical experiments;
- statistical or convergence uncertainties in the computed k_c;
- uncertainties due to extrapolation of k_c outside the experimental data range;
- uncertainties due to geometry or material simplifications and approximations;

Δk_m is an arbitrary "administrative" margin to ensure the subcriticality of k_a. Eq. 19.5 can be rewritten as

$$k_a + \Delta k_a + \Delta k_m - \beta + \Delta\beta \leq 1\,, \tag{19.6}$$

where $\beta = k_c - 1$ is the bias and $\Delta\beta = \Delta k_c$ is the uncertainty in the bias. The definition of β is based on the fact that critical experiments, by their definition, have k_{eff} = 1, and hence the bias is just the difference between the computed value and unity. By convention, the bias is defined such that a *negative* β indicates an *underestimated* k_{eff}, which is undesirable.

To ensure subcriticality in the application system, an *upper subcritical limit* is defined as

$$USL = 1 - \Delta k_m + \beta - \Delta\beta\,, \tag{19.7}$$

and from Eq. 19.6, it is apparent that $k_a + \Delta k_a \leq USL$ [31]. Thus, the USL is the maximum value an application k_{eff} (plus any uncertainties) may have for which the application can, with a high degree of confidence, be considered subcritical. The value $1 - USL$ is the mathematical definition of the subcritical margin. In the event the bias β is positive, *i.e.* the computed value k_c overestimates k_{eff}, current practice is to set the bias to zero rather than reduce the subcritical margin.

Traditional Bias Determination

In the United States, biases and associated uncertainties and USL's have often been found through use of *trending analysis*. A suite of critical experiments is selected for use in a specific safety analysis based on the similarity of the experiments to the safety analysis system. Traditionally, this similarity is based on physical parameters that include the fissile material, hydrogen-to-fissile atom ratio (H/X), average neutron energy group causing fission, and energy of average lethargy causing fission (EALF), among others [10].

While various methods using trending analysis exist for determining the biases and USL's, one common approach, denoted USL_1 in the literature [10], is discussed here to provide at least some background of current practice. The following description is largely adapted from the technical report in which it was first developed [31].

The method computes $k_c(x)$ as a function of a trending parameter x using linear regression. The bias, $\beta(x)$, is just $k_c(x) - 1$. .

A lower confidence band $w(x)$ is statistically computed using current experiments and uncertainties as well as a specified confidence. This confidence band defines the value below which an additional computed k_{eff} value (*i.e.* not included in the analysis) must be for the additional system to be considered subcritical with a high degree of confidence. Equivalently, for a given value of the trending parameter

x, $w(x)$ represents the value *above* which the additional computed value k_c will be *if* the system in question is critical—and this consequently implies any negative $\beta(x)$ is no worse than $w(x) - 1$. Hence, if our aim is to ensure the actual k_{eff} of a system is subcritical, then its computed value should be *less* than the appropriate confidence band value.

To simplify analyses, the limiting width, W, of the confidence band is used, which (like $w(x)$) takes into account all uncertainties associated with the experiments (*e.g.* experimental, stochastic, etc), and consequently, is a statistical measure of $\Delta\beta$. The width W of the confidence band is defined

$$W = \max\Big\{ w(x) | x_{\min}, x_{\max} \Big\} , \tag{19.8}$$

where

$$w(x) = t_{1-\gamma} s_p \left(1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \right)^{1/2} , \tag{19.9}$$

and n is the number of critical experiments included in the analysis, $t_{1-\gamma}$ is the student-t distribution statistic for $1-\gamma$ and $n-2$ degrees of freedom (γ is the desired confidence level), $\bar{x}$ is the mean value of x in the set of experiments, and s_p is the pooled standard deviation for the set of calculations.

The pooled standard deviation s_p is defined by

$$s_p^2 = s_{k(x)}^2 + s_w^2 , \tag{19.10}$$

where $s_{k(x)}^2$ is the mean-square error of the linear regression, defined as

$$s_{k(x)}^2 = \frac{1}{n-2} \left(\sum_{i=1}^{n} (k_i - \bar{k})^2 - \frac{\Big(\sum_{i=1}^{n} (x_i - \bar{x})(k_i - \bar{k}) \Big)^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \right) , \tag{19.11}$$

and s_w^2 is the variance of the data, defined as

$$s_w^2 = \frac{1}{n} \sum_{i=1}^{n} \sigma_i^2 , \tag{19.12}$$

where σ_i is the uncertainty in the ith calculated value, accounting for method uncertainty (*e.g.* Monte Carlo statistics) and estimated uncertainty due to cross-section uncertainty.

The width of the confidence W is chosen to be the *maximum* value of w at the limits of the range of x corresponding to the experiments to be conservative, and moreover, it serves to simplify the definition of USL. Current guidance is to define W at the 95% confidence level, *i.e.* choose $\gamma = 0.95$. Below, we provide an illustrative example of the USL_1 methodology.

S/U-Techniques

Unfortunately, the proper selection of parameters over which to trend, and the experiments with which to trend, is largely based on expert judgment. As experiments continue to become more expensive, use of computational methods will grow. Furthermore, new applications will continue to extend beyond the areas of applicability of current experimental data, and consequently, methods to extend beyond these areas are needed.

For the past several years, ORNL has worked with the support of the NRC and the Department of Energy's (DOE) Nuclear Criticality Safety Program to develop "a rigorous physics-based approach for the determination of system similarity" for determining areas of applicability [10]. Additionally, their efforts aimed to develop the methodology and computational tools needed for "determination of biases and uncertainties due to experimental descriptions, computational methods, and nuclear data."

As an alternative to traditional trending analysis, work was done to develop S/U-based, easily-quantifiable parameters to measure the similarity of systems. It is beyond our scope to describe the methods in detail. Interested readers should see the exercises of Lecture **??** for some basic theory needed to derive the results, and Ref. [10] for greater detail.

Skipping the theory, what we end up with is the sensitivity of k_{eff} to the various underlying nuclear data,

$$S_{k,\sigma_x} = \frac{\sigma_x}{k}\frac{\partial k}{\partial \sigma_x} \tag{19.13}$$

where x denotes some reaction of interest. We express the sensitivity of k_{eff} to all nuclides in a system as the vector $\mathbf{S}_k$, which can implicitly represent dependence on multigroup data.

Sensitivity vectors can be used to compare both the qualitative and quantitative similarity between two systems with respect to single or several nuclide-reactions. Let the entire set of group-wise, nuclide-reaction cross-sections be denoted $\boldsymbol{\sigma} \equiv \sigma_i, \quad n = 1,\ 2, \ldots, N$, where N is the total number of nuclide-reactions multiplied by the number of energy groups. The $N \times N$ correlation (*i.e.* relative covariance) matrix of $\boldsymbol{\sigma}$ is defined

$$\mathbf{C}_{\sigma\sigma} \equiv \left[\frac{\text{COV}(\sigma_i, \sigma_j)}{\sigma_i \sigma_j}\right], \tag{19.14}$$

where i and j range from 1 to N, and

$$\text{COV} = \langle (\sigma_i - \bar{\sigma}_i)(\sigma_j - \bar{\sigma}_j) \rangle = \langle \delta\sigma_i \delta\sigma_j \rangle\,, \tag{19.15}$$

where $\bar{\sigma}_i$ represents the expected value of the data σ_i.

Because the components of $\mathbf{C}_{\sigma\sigma}$ represent relative uncertainties, and because from above, we know the k_{eff} sensitivity represents the relative change in k_{eff} due to relative changes in some nuclide-reaction data, the relative uncertainty in k_{eff} is therefore

$$\frac{\delta k}{k} = \sqrt{\mathbf{S}_k \mathbf{C}_{\sigma\sigma} \mathbf{S}_k^T}\,, \tag{19.16}$$

where T denotes the matrix transpose. If several systems are of interest, then an $N \times I$ matrix of sensitivity vectors $\bar{\mathbf{S}}_k$ can be defined, where I is the number of systems. By folding $\bar{\mathbf{S}}_k$ with $\mathbf{C}_{\sigma\sigma}$, one obtains

$$\mathbf{C}_{kk} = \bar{\mathbf{S}}_k \mathbf{C}_{\sigma\sigma} \bar{\mathbf{S}}_k^T\,, \tag{19.17}$$

an $I \times I$ matrix consisting of each system's relative variance in k_{eff} (*i.e.* $(\Delta k/k)^2$) due to data uncertainties (the diagonal terms, α_{ii}) and the relative covariances in k_{eff} between systems (the off-diagonal terms, α_{ij}). These off-diagonal terms represent "shared variance" or "shared uncertainty" between systems.

The correlation coefficient between system i and j is defined

$$c_k = \frac{\alpha_{ij}^2}{\alpha_i \alpha_j}\,, \tag{19.18}$$

which, as is expected, takes on values between -1 (completely anti-correlated) and 1 (completely correlated). A value $c_k = 0$ indicates no correlation between systems.

The correlation of two systems is greatest if they share basic characteristics, *e.g.* fuel type, moderator, other materials. Two water-moderated thermal UO_2 systems would be expected to have a higher degree of correlation than would such a thermal system and a molten salt fast reactor. Use of c_k as a trending parameter in the USL_1 method is illustrated below.

Example Analyses

To illustrate the USL_1 method using both traditional parameters and c_k, 25 thermal LEU systems were chosen for example analyses. Two additional systems were selected as "applications" for which the β, $\Delta\beta$, and the USL are determined[§]. The experiments range from 2.35% to 5% enrichment, and have EALF values ranging from 0.017 to 2.24 eV. All experiments are included in the International Handbook of Evaluated Criticality Benchmark Experiments [4] and are low enriched uranium, thermal assemblies. The LCT-079 and LCT-050 are experiments of use to burnup credit, a topic discussed below; however, their use here is purely for example.

[§]The minimum recommended number of experiments for trending with the USL_1 method is 25; here, just the minimum was used for these example cases.

Figures 19.4-19.7 show the trending analysis for using EALF, enrichment, and c_k. For both EALF and enrichment, only one analysis was needed for the applications since neither parameter depends on the application. However, separate cases were run using c_k values specific to the given application. The experiments are the black dots, and the applications are red shapes. For all cases, the uncertainty is computed using Eq. 19.17, where the uncertainty for the system is its associated diagonal element of $\mathbf{C}_{kk}$, *i.e.* the uncertainty in cross-sections propagated to k_{eff} via use of sensitivity profiles. This is in line with previous analyses [10]. Note, for the USL, an administrative margin of $\Delta k_m = 0.05$ was used.

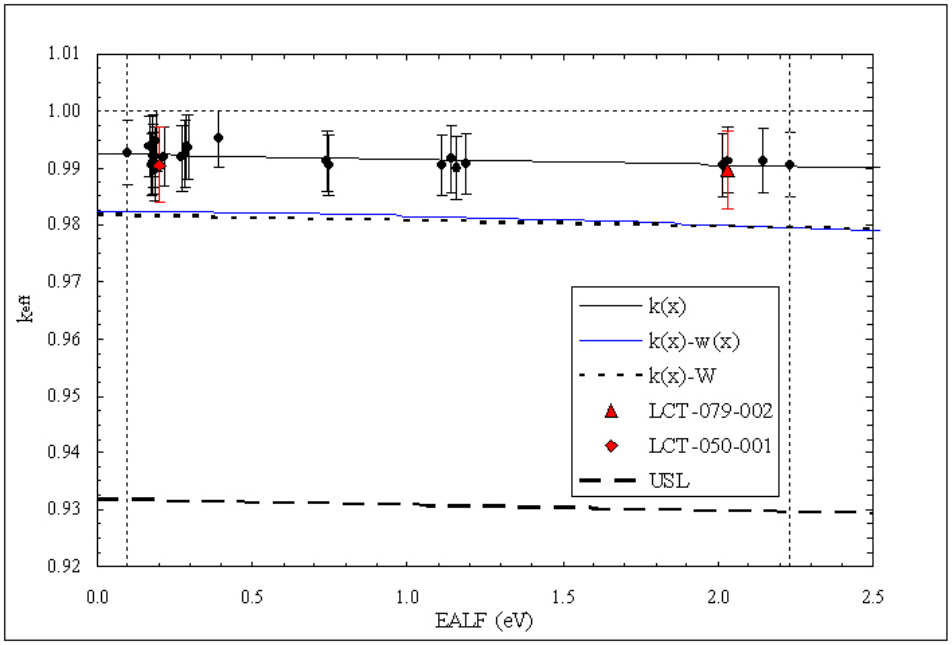

Figure 19.4: Example trending analysis using EALF.

From the trends, the bias $k(x) - 1$ (where x is the parameter value for the application) and associated bias uncertainty (*i.e.* W from Eq. 19.8) can be computed in addition to the USL. Since the "applications" are known critical experiments, it is useful to compare the observed bias (*i.e.* $k_{\text{calc}} - 1$) to the bias as predicted via trending. Table 19.3 provides the observed and predicted biases (with uncertainty) and the USL for both applications. For each application, all three methods yield very similar USL's, and all the biases are slightly underpredicted but well within one standard deviation of the observed biases.

With such a large Δk_m, it is easy to wonder why we care about $\Delta\beta$. However,

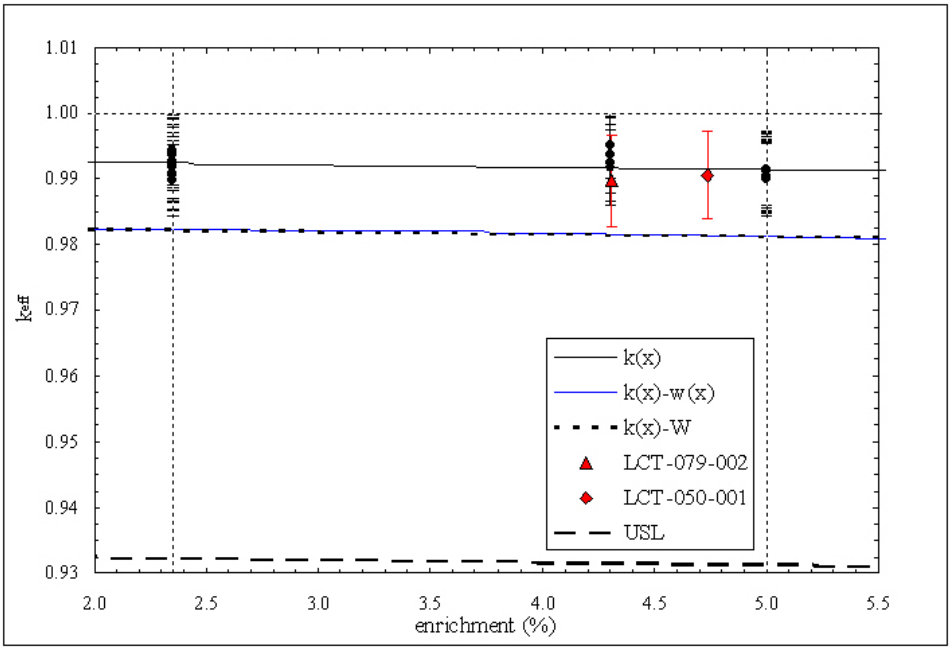

Figure 19.5: Example trending analysis using enrichment.

if we interpret subtracting Δk_m from k_{eff} as simply shifting the definition of critical, then it becomes clearer that $\Delta\beta$ is still important.

Table 19.3: Observed and predicted biases (β_{ob} and β), $\Delta\beta$, all in percent, and USL using each trending technique.

	LCT-079-002				LCT-050-001			
	β_{ob}	β	$\Delta\beta$	USL	β_{ob}	β	$\Delta\beta$	USL
EALF	-1.08	-0.95	1.08	0.9298	-0.94	-0.77	1.08	0.9316
Enrich.	-1.08	-0.84	1.02	0.9314	-0.94	-0.85	1.02	0.9313
c_k	-1.08	-0.74	1.16	0.9306	-0.94	-0.71	1.07	0.9322

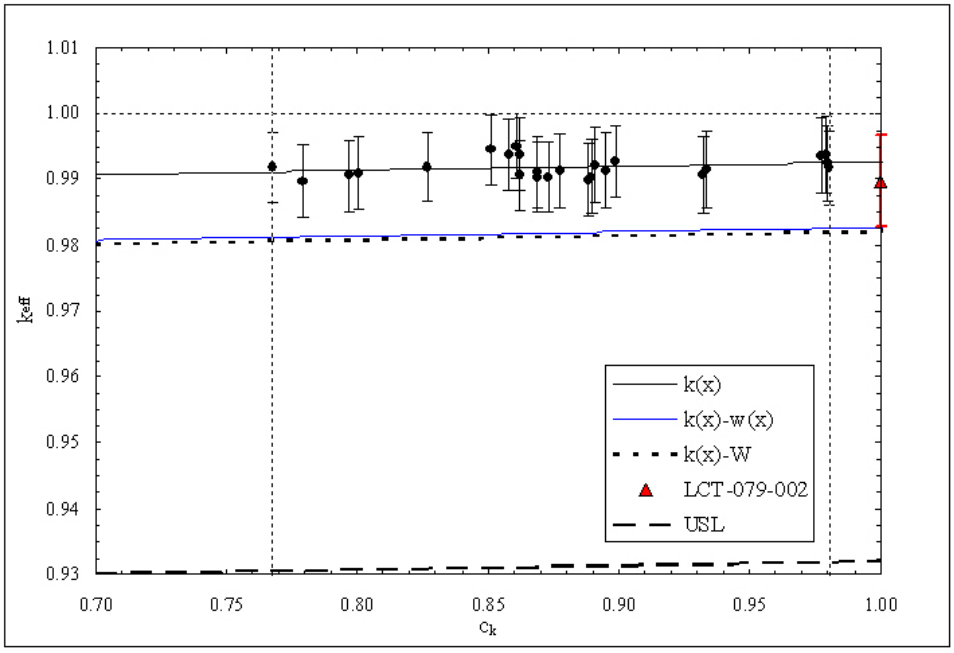

Figure 19.6: Example trending analysis using c_k (for LCT-079-002).

Case Study: Burnup Credit

Overview and Motivation

The transportation and storage of used nuclear fuel is an integral component of any nuclear fuel cycle. During handling of used nuclear fuel, strict attention must be paid to criticality safety. The chief concern of criticality safety is to ensure the effective multiplication factor k_{eff} of the system in question is below unity at all times, i.e. to ensure subcriticality. For the case of used nuclear fuel, several factors affect the subcritical margin, i.e. by how much a system is subcritical. Typical, unirradiated light water reactor (LWR) fuel consists of uranium-oxide (UO2) enriched to 3-5% ^{235}U. During its time in the reactor, nuclear fuel undergoes significant compositional changes, a process referred to as burnup. Most importantly, the fissile isotope 235 U is depleted, generating various fission products (FP's), many of which are parasitic neutron absorbers. Simultaneously, ^{235}U breeds ^{239}Pu which also undergoes fission and produces FP's. The net effect of these compositional changes is to decrease the k_{eff} of the fuel with increased burnup. Accounting for this decrease in k_{eff} (or equivalently, a decrease in reactivity) in subcritical margins is often referred to as

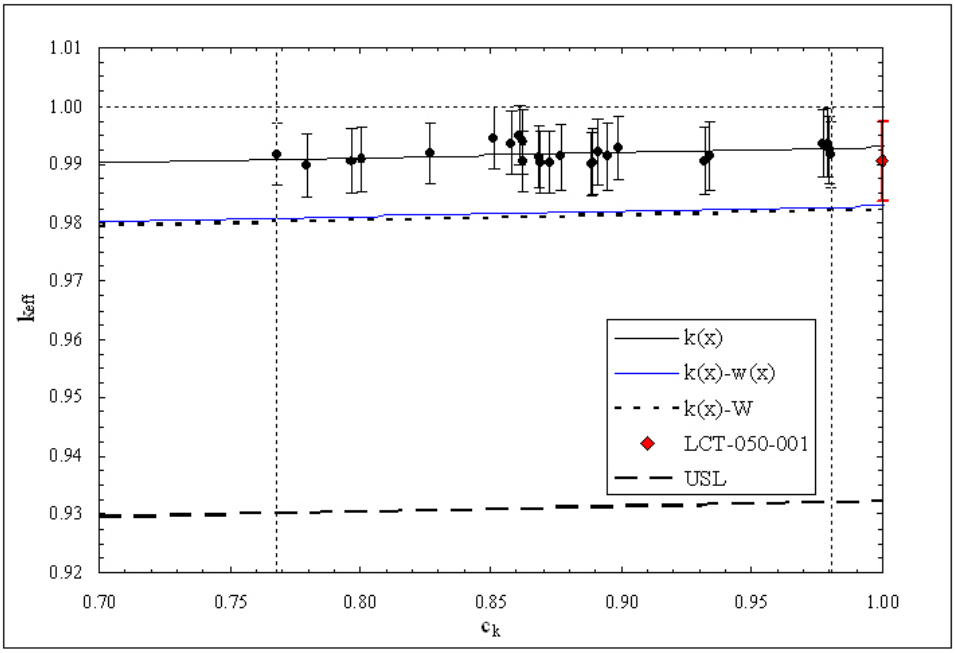

Figure 19.7: Example trending analysis using c_k (for LCT-050-001).

burnup credit.

Historically, the so-called *fresh fuel assumption* was used as a conservative bound in criticality safety analysis of used nuclear fuel [1]. More recently, the Nuclear Regulatory Commission (NRC) offered guidance for crediting the major (fissile) actinides in such analyses in its Interim Staff Guidance - 8, revision 2 (ISG-8R2) report [2]. However, even this results in a very conservative estimate of the subcritical margin of used nuclear fuel.

Changes in the major actinides account for about 66-75% of the net reduction in reactivity; FP's account for the remainder, the percentages depending on burnup. The FP's relevant to burnup credit, roughly in order of importance, include SM.

Why do we care? Naturally, assuming a canister of some number of burned assemblies contains only fresh fuel is quite conservative. Figure 19.8 shows loading curves for a generic used fuel canister with 32 WH 17x17 assemblies of varying initial enrichments and burnups. Configurations below a given line do not meet subcritical limits under the given assumptions. The reference case assumes fresh fuel, and each subsequence curve relaxes the assumptions.

Considering that much of the current U.S. inventory of used fuel lies below the reference curve, a 32-assembly canister could not be widely used (and instead, canis-

ters such as the 21-assembly Transportation, Aging, and Disposal (TAD) canister [3] intended for ultimate disposition in Yucca Mountain would be required). The cost of producing, loading, shipping, and disposing canisters is not negligible, and a study by Parks and Wagner suggest that crediting the FP's in criticality safety analysis, thus allowing the high capacity canisters, would potentially save upwards of $200 million dollars in total disposal costs [35].

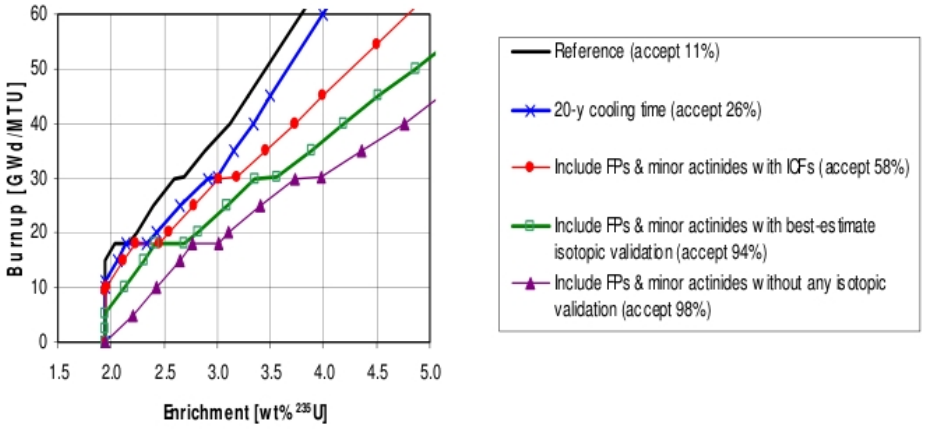

Figure 19.8: Effect of calculational assumptions on loading curves for the GBC-32 and WH 17x17 assemblies.

Current Research

To take into account the reduction in reactivity due to fission products, called *full burnup credit*, requires adequate knowledge of the isotopic content of the spent fuel (via validated depletion methods), and, our focus here, *methods* to determine criticality and *experiments* to validate those methods.

Current work has expanded on the S/U-based techniques outlined above to develop methods for determining biases of individual nuclides. While the details are beyond our scope, we give a brief description and several references for interested readers.

The basis of the work is the generalized linear least-square method (GLLSM). GLLSM takes as input a set of nuclear data and covariances, the sensitivities of experiment models to nuclear data, the model k_{eff} and uncertainty, the experiment k_{eff} and uncertainty (and any experiment correlation), and *adjusts the nuclear data* so

that the the experiment and model k_{eff} match. It does so in a way that minimizes the variation of all parameters (data and k_{eff}) in terms of the standard deviations of those parameters. In other words, for all parameters p, the method forces the experiment and model k_{eff} to match while minimizing $\sum_i (\Delta p_i / \sigma_{p_i})^2$. The adjustments can be propagated to an application's k_{eff} via its model sensitivities, and the resulting difference is the bias.

To find biases of individual nuclides, the GLLSM method can be applied to so-called "replacement experiments". These experiments consists of a single reference case, and one or more cases that introduce some small perturbation to the system, such as small foils of a fission product between fuel pellets or small concentrations of a fission product in a bulk moderator. GLLSM can then be applied to the reactivity difference between a reference-perturbation pair of models and experiments (rather than on the eigenvalue), since the sensitivities of the reactivity difference should be small for all but the perturbation material. Consequently, the corresponding adjustment to the data should primarily be due to the test material, and as above, the adjustment can be propagated to the application to define a nuclide-specific bias.

Limitations

Two significant limitations apply to the methods under development. First, the methods require that relevant experiments are available. However, only a handful of experiments have been performed relevant to burnup credit, and of those, the difference between reference and perturbation cases may be too high to extract partial biases. Moreover, little data is available for the correlation between experiments. For the the reactivity difference method described above, the resulting biases are extremely sensitive to the correlation between experiments. This suggests that future experiments should be designed with the various S/U methods as guidance.

A second, perhaps more significant issue is the general lack of reliable nuclear covariance data. Only for the most important isotopes does credible data exist (such as uranium isotopes). For isotopes of generally less importance (like many fission products), little if any covariance data exists. Until reliable covariance data exists, the methods described above are of limited value.

19.1 Further Reading

Much of the content in this lecture was inspired by Knief's *Nuclear Criticality Safety: Theory and Practice* [24]. Of course, in one lecture, all that material cannot be covered, and the student is directed to that book for more info most of the topics addressed here. A more succinct overview of some of the topics is given by Pevey in a chapter of the *Nuclear Engineering Handbook* [25].

For an overview of the application of sensitivity and uncertainty analysis to criticality safety, see Broadhead et. al [10], and for the latest work on adjustment techniques applicable to full burnup credit, see Rearden et al. [37].

Part III

Bibliography and Appendices

Bibliography

[1] "'American National Standard for Criticality Safety Criteria for the Handling, Storage, and Transportation of LWR Fuel Outside Reactors," ANSI/ANS-8.17-2004, American Nuclear Society (2004). (ANSI/ANS-8.17-2004).

[2] "American National Standard for Nuclear Criticality Safety in Operations with Fissionable Material Outside Reactors," ANSI/ANS-8.1-1998, American Nuclear Society (1998). (ANSI/ANS-8.1-1998).

[3] *Transportation, Aging, and Disposal Canister System Performance Specification*. U.S. Dept. of Energy, Office of Civilian Radioactive Waste Management.

[4] *International Handbook of Evaluated Criticality Safety Benchmark Experiments*. NEA/NSC/DOC(95)03/I-VIII, OECD-NEA, September 2007.

[5] M.L. Adams and E.W. Larsen. Fast Iterative Methods for Discrete-Ordinates Particle Transport Calculations. *Progress in Nuclear Energy*, 40(1):3–159, 2002.

[6] R.E. Alcouffe and R.S. Baker. Time-dependent deterministic transport on parallel architectures using partisn. Technical Report LA-UR-98-275, Los Alamos National Laboratory.

[7] R.E. Alcouffe, R.S. Baker, J.A. Dahl, S.A. Turner, and R. Ward. Partisn: A time-dependent, parallel neutral particle transport code system. Technical Report LA-UR-05-3925, Los Alamos National Laboratory, 2005.

[8] G.B. Arfken and H.J. Weber. *Mathematical Methods for Physicists*. Academic Press, 1995.

[9] G.I. Bell and S. Glasstone. *Nuclear Reactor Theory*. Van Nostrand-Reinhold, New York, 1970.

[10] B.L. Broadhead, B.T. Rearden, C.M. Hopper, J.J. Wagschal, and C.V. Parks. Sensitivity-and Uncertainty-Based Criticality Safety Validation Techniques. *Nucl. Sci. Eng*, 146(3):340–366, 2004.

[11] K.M. Case and P.F. Zweifel. *Linear Transport Theory*. Addison-Wesley, Reading,MA, 1967.

[12] G.R. Cefus and E.W. Larsen. Stability Analysis of Coarse-Mesh Rebalance. *Nucl. Sci. Eng*, 105:31, 1990.

[13] S. Chandrasekhar. *Radiative Transfer*. Oxford University Press, 1950.

[14] G. Davidson and E.W. Larsen. Sweepless Time-Dependent Transport Simulations Using the Staggered-Block Jacobi Method. In *International Conference on Mathematics, Computational Methods & Reactor Physics (M&C 2009), Saratoga Springs, USA, May*.

[15] B. Davison. *Neutron Transport Theory*. Oxford, 1957.

[16] J.R. Dishaw. *Time Dependent Discrete Ordinates Neutron Transport Using Distribution Iteration in XYZ Geometry*. PhD thesis, 2007.

[17] J.J. Duderstadt and L.J. Hamilton. *Nuclear Reactor Analysis*. John Wiley and Sons, Inc., New York, 1976.

[18] J.J. Duderstadt and W.R. Martin. *Transport Theory*. John Wiley and Sons, Inc., New York, 1979.

[19] T.M. Evans, G.G. Davidson, and R.N. Slaybaugh. Three-Dimensional Full Core Power Calculations for Pressurized Water Reactors. In *Journal of Physics: Conference Series, accepted*, 2010.

[20] T.M. Evans, A.S. Stafford, R.N. Slaybaugh, and K.T. Clarno. Denovo: A New Three-Dimensional Parallel Discrete Ordinates Code in SCALE. *Nuclear Technology*, 171(2):171–200, 2010.

[21] A. Hébert. *Applied Reactor Physics*. Presses Internationales Polytechnique, 2009.

[22] A.F. Henry. *Nuclear Reactor Analysis*. MIT Press, Cambridge, MA, 1975.

[23] T.R. Hill and W.H. Reed. TIMEX: A Time-Dependent Explicit Discrete Ordinates Program for the Solution of Multigroup Transport Equations with Delayed Neutrons. Technical Report LA-6201-MS, Los Alamos National Laboratory, 1976.

[24] R.A. Knief. *Nuclear Criticality Safety*. American Nuclear Society, 1991.

[25] R.E. Knott. Neutronics. In K.D. Kok, editor, *Nuclear Engineering Handbook*, pages 575–608. CRC, 2009.

[26] K.R. Koch, R.S. Baker, and R.E. Alcouffe. Solution of the First-Order Form of the 3-D Discrete Ordinates Equation on a Massively Parallel Processor. *Trans. Amer. Nucl. Soc.*, 65(108):198–199, 1992.

[27] E.W. Larsen. Asymptotic Analysis in Nuclear Reactor Theory (invited). *Transactions of the American Nuclear Society*, 101:711–713, 2009.

[28] E.W. Larsen and J.E. Morel. Advances in Discrete-Ordinates Methodology.

[29] R.J. LeVeque. *Finite Difference Methods for Ordinary and Partial Differential Equations*. SIAM, 2007.

[30] E.E. Lewis and W.F. Miller Jr. *Computational Methods of Neutron Transport*. American Nuclear Society, Lagrange Park, IL, 1993.

[31] J.J. Lichtenwalter, S.M. Bowman, M.D. DeHart, and C.M. Hopper. Criticality Benchmark Guide for Light-Water-Reactor Fuel in Transportation and Storage Packages. Technical report, NUREG/CR-6361 (ORNL/TM-13211), US Nuclear Regulatory Commission, Oak Ridge National Laboratory, 1997.

[32] G.I. Marchuk. *Adjoint Equations and Analysis of Complex Systems*. Springer, 1995.

[33] W.F. Miller Jr and E.E. Lewis. Nonlinear Response Matrix Methods for Radiative Transfer. In *International Topical Meeting on Advances in Reactor Physics, Mathematics and Computation*, Paris, France, April 1987. American Nuclear Society.

[34] Y.R. Park and N.Z. Cho. Coarse-Mesh Angular Dependent Rebalance Acceleration of the Discrete Ordinates Transport Calculations. *Nuclear Science and Engineering*, 148(3), 2004.

[35] C.V. Parks and J.C. Wagner. Current Status and Potential Benefits of Burnup Credit for Spent Fuel Transportation. In *14th Pacific Basin Nuclear Conference, Honolulu, HI, Mar. 21-25, 2004*, 1997.

[36] B.W. Patton and J.P. Holloway. Application of Preconditioned GMRES to the Numerical Solution of the Neutron Transport Equation. *Annals of Nuclear Energy*, 29(2):109–136, 2002.

[37] B.T. Rearden, M.L. Williams, M.A. Jessee, D.E. Mueller, and D.A. Wiarda. Sensitivity and uncertainty analysis capabilities and data in scale. *Nucl. Tech.*, 174(2):236–288, 2011.

[38] J.A. Roberts. Direct Solution of the Discrete Ordinates Equations. *MIT Course 18.086 Class Project*, 2010.

[39] A. Seubert, A. Pautz, M. Becker, and R. Dagan. Time-dependent Anisotropic distributed Source Capability in Transient 3-D Discrete Ordinates Code TORT-TD. In *International Conference on Mathematics, Computational Methods & Reactor Physics (M&C 2009), Saratoga Springs, USA, May*.

[40] K.S. Smith. *Spatial Homogenization Methods for Light Water Reactor Analysis*. PhD thesis, 1980.

[41] K.S. Smith. Nodal Method Storage Reduction by Nonlinear Iteration. *Trans. Am. Nucl. Soc.*, 44, 1983.

[42] W.M. Stacey. *Variational Methods in Nuclear Reactor Analysis*. Academic Press, 1974.

[43] R.J.J. Stamm'ler and M.J. Abbate. *Methods of Steady-State Reactors Physics in Nuclear Design*. Academic Press, New York, 1983.

[44] F. Swesty. The Solution of the Time–Dependent S N Equations on Parallel Architectures. *Computational Methods in Transport*, pages 469–486, 2006.

[45] J.S. Warsa, T.A. Wareing, and J.E. Morel. Krylov Iterative Methods and the Degraded Effectiveness of Diffusion Synthetic Acceleration for Multidimensional SN Calculations in Problems with Material Discontinuities. *Nuclear Science and Engineering*, 147(3):218–248, 2004.

[46] E.P. Wigner. Effect of Small Perturbations on Pile Period. *Chicago Report CP-G-3048; see also The Collected Works of Eugene Paul Wigner, Part A.*, 1945.

[47] M.L. Williams. Generalized Contributon Response Theory. *Nuclear Science and Engineering*, 108:355–383, 1991.

[48] M.L. Williams. Sensitivity and Uncertainty Analysis for Eigenvalue-Difference Responses. *Nucl. Sci. Eng.*, 155(1):18–36, 2007.

[49] Z. Zhong, T.J. Downar, Y. Xu, M.D. DEHART, and K.T. CLARNO. Implementation of Two-Level Coarse-Mesh Finite Difference Acceleration in an Arbitrary Geometry, Two-Dimensional Discrete Ordinates Transport Method. *Nuclear Science and Engineering*, 158(3):289–298, 2008.