Radiative Transfer Across Disciplines
Part A : Introduction and overview

A. Jerkstrand, with contributions from U. Noebauer and C. Vogl + inviting contributions from YO

March 25, 2024

Contents

Introduction

In the view of the great philosopher G.M., there are three main domains of computational astrophysics; gravity, hydrodynamics, and radiative transfer (RT). One may also delineate three main categories of numeric approaches; Smoothed Particle Hydrodynamics (SPH), Partial Differential Equations (PDEs), and Monte Carlo (MC) methods. In this course we will focus on radiative transfer, and treatment of this with both PDE and MC methods. We will study actual tools and methods used at the current research forefront in various astrophysical discplines, and get hands-on experience with these.

RT Literature

- Mihalas (1970): "Stellar Atmospheres." Also a second edition (1978) exists.
- Mihalas & Mihalas (1984), "Foundations of Radiation Hydrodynamics" (book).
- Rob Rutten 2003 compendium "Radiative Transfer in Stellar Atmospheres". Limited to plane-parallel RT.
- Hubeny & Mihalas (2014), "Theory of Stellar Atmospheres" (book).

This writeup is to a large extent based on the fundamental theory presented in Mihalas & Mihalas (1984).

1 A survey of open-source RT software

1.1 Astro

- SKIRT. Galaxies, circumnuclear disks, star-forming regions. 3D MC. C++, open-MP and MPI parallelisation.
- MAGRITTE. Stellar winds from evolved stars. Ray-tracing. C++.
- RADMC-3D. Protoplanetary disks, circumstellar envelopes, dusty molecular clouds, dusty tori around AGN and models of galaxies. 3D MC. Fortran 90, but user can do everything with a Python interface.
- HYPERION. Dust continuum RT. 3D MC. Fortran with MPI parallelization and Python wrapper.
- TARDIS. Supernova photospheric spectra. 1D MC. Python with C extensions.
- petitRADTRANS. Exoplanets. Plane-parallel RTE, LTE, with scattering.
- CMFGEN. Stellar winds and supernovae. 1D RTE.

In description of basic operation of these and other modern codes, one finds expressions like "Feautrier method, with ALI of Olson et al. 1986, and Ng acceleration" (petitRADTRANS) and "Λ-iteration with Variable Eddington Factor, acceleration with approximate Newton-Raphson solver" (CMFGEN). Thus, we will need to investigate such concepts to understand how they work.

1.2 Non-astro

- Open-MC. MC for neutron and photon transport. Romano et al. (2015).
- MCARATS. Earth atmosphere. Iwabuchi (2006).
- LBLRTM - Line by line radiative transfer module. Earth atmosphere. Clough et al. (2005).
- SpectralCalc.

2 The radiation field

The radiation field is, at any given time, a 6-dimensional quantity, varying over 3 spatial dimensions, 2 angle dimensions, and frequency. If the number of points per spatial dimension is N_s, the number of angular points N_a, and number of frequency points N_f, just storing the radiation field takes

$$N = 10^{11} \left(\frac{N_s}{100}\right)^3 \left(\frac{N_a}{10}\right)^2 \left(\frac{N_f}{1000}\right), \tag{1}$$

i.e. 100 billion numbers, much more than what can be stored in a modern computer ($\sim$ 1 billion numbers).

Do we need to store the radiation field? Discuss in pairs.

Storage requirement is often one factor for needing to simplify. Much RT work is limited to spherical symmetry, where 2 space variables and 1 angle variable are dropped, reducing the dimensionality from 6 to 3. Note that going from 3D to 2D still requires two angle variables, so dimensionality goes from 6 to 5 only. In the simplest problems we can conceive of we drop also the frequency variable (giving "gray transport"), reducing the dimensionality to 2 (radius and angle).

If the chief objective is to obtain the dynamic effects of the radiation field, much simplification and reduced dimensionality and resolution is often motivated. On the other hand, if the objective is to predict the emergent radiation in high detail, larger problems need to be considered.

2.1 Spatial gridding (N_s)

A code may operate with spatial coordinates defining the grid, or with optical depth defining the grid. In principle, the radiation field at a given wavelength can change over length scales corresponding to the mean-free-path (the average length a photon travels before interaction, see more below) at that wavelength. This translates to, for full resolution, a number of spatial points $N_s = max\,[\tau_\lambda]$, if optical depth is used as independent varible. From this follows

- Reduction of the spatial domain is always a priority in 3D, because the N_s factor has the highest power in Eq. 1. Note that in 2D, there is already an equal power (2) by another term (the angles).
- For spatial grids, it is beneficial to have grids with variable cell sizes, becoming larger where mean-free-paths are longer. One example is to use a spherical coordinate system instead of Cartesian, if density decreases outwards.

2.2 Angular gridding (N_a)

If there is structure in the spatial domain on scale size R_s, the radiation field at a given point may change on angular scales R_s/R where R is the domain scale. Thus, to simulate a region in 3D that is say clumped by a factor of 10, a factor $\sim 10^5$ ($N_s^3 N_a^2$) finer space-frequency grid is needed. This helps us to understand why models of clumped stellar winds, for example, have only quite recently been computed. When relativistic effects are important, higher angular resolution becomes necessary by a factor $\sim \Gamma$ (the beaming factor).

2.3 Frequency gridding (N_f)

The needed frequency resolution is determined by the frequency scale over which the emission and absorption coefficients vary. If bound-bound transitions are involved, these become the limiting factor due to their narrowness. The thermal line width of an atom of mass A (units of m_p) is given by:

$$v_{th} \sim \sqrt{\frac{2kT}{Am_p}} = 12 \text{ km s}^{-1} \left(\frac{T}{10^4 \text{ K}}\right)^{1/2} A^{-1/2}. \quad (2)$$

A line at rest wavelength $\lambda_0 = 5000$ Å has from this formula a thermal width of $\lambda_0 \times v_{th}/c = 0.2$ Å (for $T/A = 10^4$ K), and to resolve it one would need a frequency spacing of $\ll 0.2$ Å. To model the whole optical/NIR spectrum ($\sim$ 3000-25,000 Å), we would then need $\gg 110,000$ frequency points if a linear grid is used. From this we comment:

- A logarithmic wavelength spacing is beneficial, because the line widths scale with λ. This also reduces N_f - in the example above the range is covered by ∼53,000 points instead of ∼110,000 if a log-grid is used.
- Modelling of cold and/or heavy gas gas requires many more frequency points than hot/light gas. How much more expensive is it to resolve r-process line formation in a kilonova ejecta at 1000 K compared to H-line formation in a planetary nebula at 20,000 K?

2.4 Discussion

- To what extent have you thought about (or tested) the impact of different griddings and resolution for the radiation field in your current RT application?

3 Radiation field quantities

As particles can be treated as ensembles in statistical mechanics and thermodynamics, photons can be treated in a macroscopic field framework. For this to work, over any scale in length, angle, or frequency over which physical conditions change, the number of photons must be large. This is fulfilled in the vast majority of astrophysical situations encountered, but exceptions exist (e.g.

single-photon heating of dust grains).

The most commonly used description of the radiation field in the macroscopic picture is the **specific intensity** $I_\nu(x, y, z, \theta, \phi, t)$. This describes the energy flux per unit area at position $\mathbf{x} \equiv [x, y, z] \equiv [r, \Theta, \Phi]$, at frequency ν, into direction $\mathbf{n} = f(\theta, \phi) =$ "ω", per steradian. In general we will take θ to be the angle relative to the radial direction, and ϕ to be the orthonormal (azimuthal) angle. The unit of I_ν in cgs is erg cm^{-2} s^{-1} Hz^{-1} ster^{-1}. An important property of I_ν is that in the absence of sources or sinks, it stays constant along a ray.

An equivalent quantity is the **photon number density** ψ_ν:

$$\psi_\nu = \frac{1}{ch\nu} I_\nu, \tag{3}$$

which has unit cm^{-3} Hz^{-1} ster^{-1}. A third is the **photon distribution function** $f_{R,\nu}$:

$$f_{R,\nu} = \frac{c^2}{h^4\nu^3} I_\nu, \tag{4}$$

which has units cm^{-4} g^{-1} s ster^{-1} and describes the number density in phase space (space and momentum). A description of the 6-D phase space distribution is equivalent to the distribution over space, frequency, and angles. Of these quantities, $f_{R,\nu}$ is the relativistically invariant distribution. This means that to transform I_ν between frames, Eq. 4 tells us there is a $(\nu/\nu')^3$ conversion factor, where ν and ν' are the frequencies in the two frames.

3.1 Zeroth moments

The **mean intensity** J_ν is the angle-average of I_ν:

$$J_\nu = \frac{1}{4\pi} \int I_\nu d\omega, \tag{5}$$

which is also called the *zeroth moment* of the radiation field. It has the same units as I_ν. Here $\int d\omega$ denotes integration over all solid angles.

An equivalent quantity is the **monochromatic radiation energy density** E_ν:

$$E_\nu \equiv h\nu \int \psi_\nu d\omega, \tag{6}$$

which has units erg cm^{-3} Hz^{-1}. Combining the last two equations gives the relation $E_\nu = 4\pi/cJ_\nu$. This is often used in MC codes in which photon number densities ψ_ν are easier to tally up than I_ν (because the latter needs definition of surfaces whereas ψ_ν does not), but J_ν is the final quantity needed.

3.2 First moments

The **monochromatic radiation flux** $\mathbf{F}_\nu$ is a vector defined by

$$\mathbf{F}_\nu \equiv \int I_\nu \mathbf{n}(\omega) d\omega, \tag{7}$$

which is also called the *first moment* of the radiation field. It has unit erg s^{-1} cm^{-2} Hz^{-1}, and represents the net flux of energy through a unit surface (so in spherical symmetry the spectral luminosity at radius r is $4\pi r^2 F_\nu(r)$).

Sometimes an alternative first moment $\mathbf{H}_\nu$, called the Eddington flux, is used:

$$\mathbf{H}_\nu \equiv \frac{1}{4\pi} \mathbf{F}_\nu. \tag{8}$$

A third equivalent quantity is the **monochromatic radiation momentum density** $\mathbf{g}_\nu$:

$$\mathbf{g}_\nu \equiv c^{-2} \mathbf{F}_\nu, \tag{9}$$

which has units erg s cm^{-4} Hz^{-1}.

In spherical symmetry, $\mathbf{F}_\nu$ (and $\mathbf{H}_\nu$ and $\mathbf{g}_\nu$) are always in the radial direction and their magnitudes F_ν, H_ν, g_ν are used for notation, then simply called "the flux".

3.3 Second moments

The (3×3) **radiation stress tensor** $\mathbf{P}_\nu^{ij}$ is

$$\mathbf{P}_\nu^{ij} \equiv \int \psi_\nu(\omega) h\nu \left(\mathbf{n}^i \cdot \mathbf{n}(\omega)\right) \left(\mathbf{n}^j \cdot \mathbf{n}(\omega)\right) d\omega, \tag{10}$$

and is also called the *second moment of the radiation field*. It represents the net rate of transport of momentum in direction i across a surface with normal in direction j. It has dimension dyne cm^{-2} Hz^{-1}, and is symmetric. From the definitions of $\mathbf{P}_\nu^{ij}$ and E_ν directly follows the general relation

$$\sum_{i=1,2,3} \mathbf{P}_\nu^{ii} = E_\nu, \tag{11}$$

between the trace of the second-moment tensor and the zeroth moment.

In spherical symmetry $\mathbf{P}_\nu^{ij}$ is a diagonal matrix. If we let the three direction axes be $(\hat{\theta}, \hat{\phi}, \hat{r})$, then (MM84 Eq. 66.6):

$$\mathbf{P}_\nu^{ij} = \begin{pmatrix} (E_\nu - P_\nu)/2 & 0 & 0 \\ 0 & (E_\nu - P_\nu)/2 & 0 \\ 0 & 0 & P_\nu, \end{pmatrix} \tag{12}$$

where

$$P_\nu \equiv P_\nu^{\hat{r}\hat{r}} = \int \psi_\nu(\omega) h\nu \left(\hat{\mathbf{r}} \cdot \mathbf{n}(\omega)\right) \left(\hat{\mathbf{r}} \cdot \mathbf{n}(\omega)\right) d\omega = \frac{2\pi}{c} \int_{-1}^{1} I_\nu \mu^2 d\mu, \tag{13}$$

where we have introduced $\mu \equiv \hat{\mathbf{r}} \cdot \mathbf{n}$. For isotropic radiation fields ($I_\nu(\mu) = I_\nu = J_\nu$), we get $P_\nu = 4\pi/3cJ_\nu = E_\nu/3$ and all entries on the diagonal equal P_ν.

Sometimes the alternative quantities $\mathbf{K}_\nu^{ij}$ and K_ν (called "the K-integral") are used:

$$\mathbf{K}_\nu^{ij} = \frac{c}{4\pi} \mathbf{P}_\nu^{ij} \tag{14}$$

$$K_\nu \equiv \frac{c}{4\pi} P_\nu = \frac{1}{2} \int_{-1}^{1} I_\nu \mu^2 d\mu \quad \text{(last relation for spherical symmetry).} \tag{15}$$

The **mean radiation pressure** $\bar{P}_\nu$ is defined as

$$\bar{P}_\nu \equiv \frac{1}{3} \sum_{i=1,2,3} P_\nu^{ii} = E_\nu/3. \tag{16}$$

3.4 Other quantities

Sometimes **normalized moments** are better suited to characterize the radiation field than the ordinary moments.

The normalized first moment $F_\nu/(cE_\nu) \equiv H_\nu/J_\nu$ is also called the **flux factor** and it measures how much net transport of radiation energy takes place. Its absolute value is bounded by 0 and 1. Values close to zero are indicative of conditions close to isotropy (in which no net energy is transported in any direction), while values close to 1 arise in a free-streaming regime where most radiation moves along the same direction and therefore transport all their energy in that direction.

The **Eddington factor** f_ν is the ratio of the second moment (in the radial direction) and the zeroth moment:

$$f_\nu \equiv \frac{P_\nu}{E_\nu} = \frac{K_\nu}{J_\nu}. \tag{17}$$

Its value also lies between 0 and 1. Similarly to the flux factor, the Eddington factor measures the degree of anisotropy of the radiation field, while going from isotropic diffusion to free-streaming conditions, f_ν increases from 1/3 to 1. As a geometric measure, it can usually be obtained to satisfactory accuracy by solving (fast) approximations to the transfer equation (more later on).
Some methods deploy even higher order integrals of I_ν, but most use some combination of those described here.

Quantity	General	Radial component
Radiation field	I_ν, ψ_ν, $f_{\nu,R}$	-
Zeroth moment	J_ν, E_ν	-
First moment	$\mathbf{F}_\nu$,$\mathbf{H}_\nu$, $\mathbf{g}_\nu$	F_ν, H_ν, g_ν
Second moment	$\mathbf{P}_\nu^{ij}$,$\mathbf{K}_\nu^{ij}$	P_ν, K_ν

Table 1: Radiation field quantities.

4 Emission and absorption coefficients

To set up a transfer equation, we need terms specifying the creation and destruction of radiation; these are the emission and absorption coefficients.

The **emission coefficient** (or **emissivity**) is often denoted $\eta_\nu(x, y, z, \theta, \phi)$ and is in general a 6-D quantity, with unit erg s^{-1} cm^{-3} Hz^{-1} ster^{-1}. The product $\eta_\nu ds$, where ds is a path segment, gives the amount of specific intensity dI_ν injected into the beam over the segment length. In the fluid rest frame (comoving frame), η_ν is in general still a 6-D quantity as the emissivity arising from scattering can be angle-dependent. If there is no scattering, or if the scattering is isotropic, it reduces to an (isotropic) 4-D quantity (3 space + frequency).

The **absorption coefficient** is often denoted $\chi_\nu(x, y, z, \theta, \phi)$[1] and is also a 6-D quantity, with unit cm^{-1}. The product $I_\nu \chi_\nu ds$ gives the amount of specific intensity dI_ν lost from the beam over the path length. In the fluid rest frame, χ_ν is isotropic so reduces to a 4-D quantity (3 space + frequency).

The **mean-free path** λ_ν is

$$\lambda_\nu \equiv \frac{1}{\chi_\nu}, \tag{18}$$

and has unit cm. It specifies how far a photon travels on average before it has some kind of interaction.

The **optical depth** between two points is the integral of the absorption coefficient between the points

$$\tau_\nu(\mathbf{x}, \mathbf{x}') = \int_{\mathbf{x}}^{\mathbf{x}'} \chi_\nu(\mathbf{x} + \mathbf{n}s) ds. \tag{19}$$

Absorption can sometimes be divided into the processes of scattering and thermalization. By **scattering** we mean that the photon is reemitted with the same energy in the frame of the scattering agent, but in a new direction. This happens for example in Thomson scattering by free electrons, and in resonance line scattering by atoms. By **thermalization** we mean that the photon energy is transformed to thermal kinetic energy of the gas (and possibly also to potential energy). This happens for instance when a line absorption is followed by collisional deexcitation, or in photoionization (here also potential energy is created along with thermal energy).

One should note that, in the general case, one does not know a-priori what will happen following e.g. a line absorption; the atom can do a myriad of things. Thus, this division between scattering and thermalization does not really exist in the most generalized description - but for many simplified ones used in practise, not least as a conceptual aid.

An often used nomencalture is that thermal absorption is represented by κ_ν and scattering by σ_ν, so

$$\chi_\nu = \kappa_\nu + \sigma_\nu. \tag{20}$$

The emissivity can likewise be divided into thermal and scattering components, often denoted as

$$\eta_\nu = \eta_\nu^t + \eta_\nu^s. \tag{21}$$

[1]The absorption coefficient is sometimes called the **opacity** (e.g. in Mihalas) but we will here use opacity for cross section per unit gram.

The **source function** is the ratio between emissivity and absorption coefficients

$$S_\nu \equiv \frac{\eta_\nu}{\chi_\nu}. \tag{22}$$

It has the same units as I_ν.

4.1 Scattering

In general, a scattering emissivity function may be written (e.g. Eq 1-23 in Mihalas 1978)

$$\eta^s_\nu = \frac{1}{4\pi}\int_0^\infty \int I_{\nu'}(d\omega')R(\nu', \mathbf{n}', \nu, \mathbf{n})d\nu' d\omega', \tag{23}$$

where R is the **redistribution function**, a 9-D quantity (3 space, 2 incoming angle, 2 outgoing angle, 1 incoming frequency, 1 outgoing frequency). Its high dimensionality means in virtually all applications simplifying assumptions about it are necessary.

4.1.1 Complete coherence limit

In this limit one assumes perfectly coherent scattering ($\nu = \nu'$). This is usually the treatment for continuum scattering, and sometimes also line scattering (Mihalas 1978). We can then write

$$R = g(\mathbf{n}', \mathbf{n})\sigma_\nu \psi(\nu')\delta(\nu - \nu'), \tag{24}$$

where ψ is the (normalized) line profile function. If we also assume isotropy, then $g = 1$, and

$$\eta^s_\nu = \sigma_\nu \frac{1}{4\pi}\int I_\nu d\omega' = \sigma_\nu J_\nu. \tag{25}$$

For dipole scattering, $g = 3/4(1 + \cos(\mathbf{n}\cdot\mathbf{n}')^2)$.

4.1.2 Complete incoherence (=complete redistribution) limit

In this limit one instead assumes that there is no correlation between ν' and ν. Then...TBD...

4.1.3 Partial redistribution

TBD..

5 Transfer equation methods vs Monte Carlo methods

There are two general categories of methods for solving RT problems. The first one is the "classic" one, where the bookkeeping equation for the radiation field is solved by discretizing the ODEs, PDEs, or integro-differential equations (depending on problem) using finite differences for derivatives, and/or numeric integration techniques. Sometimes part of the problem can be solved with analytic steps as well. We will call these methods "RTE methods". One may outline two branches of RTE methods; *grid discretization* and *ray-tracing*. In ray-tracing we always solve a differential equation for the specific intensity along some beam, which can be done if emission and absorption coefficients are known along the path. This is also called a formal solution. When either emission or absorption depends on the radiation field (to be computed) itself, we have coupling between beams and we have to co-solve for them.

The second category is Monte Carlo simulations. While the continuum-description of the radiation field lends itself to differential-equation solving that works well for the majority of applications, its fundamental constituency of discrete particles (photons) also opens up possibilities to simulate the RT directly, photon by photon (or photon packet by photon packet). Because the quantum physics governing matter-radiation interactions is random by nature, the random sampling capabilities of modern computers can mimic the true physics in an intuitive and correct way.

There are various pros and cons of these two approaches, and which works best varies with application. While historically RTE methods were almost exclusively, MC methods have seen a strong growth over the past 10-20 years. Their main weakness - the need to simulate a large number of photon trajectories and interactions to get an accurate solution - made them hardly feasible until the extraordinaty power of modern parallel computing systems became widely available in the early 2000s.

Here are some pros and cons of MC methods:

- + Intuitive and quite easy to code up.
- + Well suited to large-scale parallel HPC, which took off in the early 2000s as the fundamental clock speed of individual processors was reached.
- + No significant increase in code complexity when the domain dimensionality increases (1D,2D,3D) or microphysics (e.g. anisotropic instead of isotropic scattering) gets more complex.
- + Allow focussing the computing power on "where the energy is".
- + Allow informative looks "inside the machinery", can tag packets with information about their origin, trajectories and interactions.
- - When RTE methods are an alternative, they are usually faster - sometimes significantly so.
- - The MC noise associated with the random sampling introduces issues with testing and reproducability. It can also lead to issues with convergence criteria. It is expensive to reduce the levels of MC noise ($\propto 1/\sqrt{N_{packets}}$).

6 The transfer equation

The transfer equation corresponds to the Boltzmann equation for photons. The non-relativistic transfer equation in the observer frame, along a path s is (e.g. Eq. 76.3 in Mihalas & Mihalas 1984)

$$\left[\frac{1}{c}\frac{\partial}{\partial t} + \frac{\partial}{\partial s}\right] I_\nu(\mathbf{x},\mathbf{n},t) = \eta_\nu(\mathbf{x},\mathbf{n},t) - \chi_\nu(\mathbf{x},\mathbf{n},t) I_\nu(\mathbf{x},\mathbf{n},t). \tag{26}$$

Depending on the problem and algorithm, η_ν and χ_ν may be explicitly known, or they may be functions of I_ν (in which case the problem is much more difficult). In the first case we have to solve a PDE in two variables (t and s), in the second case an integro-partial-differential equation.

The $\partial/\partial s$ differential corresponds in generic notation to $\mathbf{n}\cdot\nabla$. In Cartesian coordinates:

$$\frac{\partial I}{\partial s} = (\mathbf{n}\cdot\nabla) I = \frac{\partial I}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial I}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial I}{\partial z}\frac{\partial z}{\partial s} = n_x\frac{\partial I}{\partial x} + n_y\frac{\partial I}{\partial y} + n_z\frac{\partial I}{\partial z}. \tag{27}$$

In curvilinear coordinates, moving on a straight line path ds corresponds to continuous rotation of the basis vectors (e.g., in spherical symmetry the angle to the normal θ changes along a ray (unless $\theta = 0$.)) In **spherical symmetry**,

$$\frac{\partial}{\partial s} = \frac{\partial r}{\partial s}\frac{\partial}{\partial r} + \frac{\partial \theta}{\partial s}\frac{\partial}{\partial \theta}, \tag{28}$$

where we have used $dr = \cos\theta ds$ and $rd\theta = -\sin\theta ds$. Eq. 26 then becomes

$$\left[\frac{1}{c}\frac{\partial}{\partial t} + \mu\frac{\partial}{\partial r} + \frac{1-\mu^2}{r}\frac{\partial}{\partial \mu}\right] I_\nu(\mathbf{x},\mathbf{n},t) = \eta_\nu(\mathbf{x},\mathbf{n},t) - \chi_\nu(\mathbf{x},\mathbf{n},t) I_\nu(\mathbf{x},\mathbf{n},t). \tag{29}$$

6.1 Zeroth moment of the transfer equation

Integrating the transfer equation over angle, we get (Eq. 78-2 in MM84)

$$\frac{1}{c}\frac{\partial J_\nu}{\partial t} + \frac{1}{4\pi}\nabla\cdot\mathbf{F}_\nu = \frac{1}{4\pi}\int (\eta_\nu - \chi_\nu I_\nu)\, d\omega. \tag{30}$$

Here the $\mathbf{n}$ in the second term has moved inside the integral, giving a first-moment of the radiation field term ($\mathbf{F}_\nu$). In the static case, χ_ν is isotropic (while η_ν may not be), which means that the RHS will contain only $\int I_\nu d\omega = J_\nu$, not I_ν. Therefore, only moments of I_ν remain:

$$\frac{1}{c}\frac{\partial J_\nu}{\partial t} + \frac{1}{4\pi}\nabla \cdot \mathbf{F}_\nu = \frac{1}{4\pi}\int \eta_\nu d\omega - \chi_\nu J_\nu \quad \text{(static)}. \tag{31}$$

In spherical symmetry, this becomes (compare Eq. 78.5 in MM84):

$$\frac{1}{c}\frac{\partial J_\nu}{\partial t} + \frac{1}{4\pi r^2}\frac{\partial \left(r^2 F_\nu\right)}{\partial r} = \frac{1}{2}\int_{-1}^{1} \eta_\nu d\mu - \chi_\nu J_\nu \quad \text{(static)}. \tag{32}$$

By substituting the expression relating J_ν to E_ν, one can show that the first moment equation (Eq. 30) expresses energy conservation : the rate-of-change for the radiative energy density is the net creation minus the net outflow (see Eq. 78.3 in Mihalas & Mihalas 1984).

Note that even for anisotropic scattering, the RHS does still not depend on I_ν because we integrate over outgoing angle.

Radiative equilibrium refers to the state that the gas emits as much as it absorbs:

$$\int_\nu \int_\omega \left(\eta_\nu - \chi_\nu I_\nu\right) d\omega d\nu = 0. \tag{33}$$

This must hold in any static situation. Since such a situation also has $\partial J_\nu/\partial t = 0$, it follows from Eq. 30 that $\nabla \cdot \mathbf{H} = 0$, which gives L = constant. Thus, in a stellar atmosphere, for example, the luminosity is the same at each depth. In the interior of the star radiative equilibrium does not hold because of energy generation by fusion, and a gradient in L is established.

6.2 First moment of the transfer equation

Multiply now the transfer equation (Eq. 26) by $\mathbf{n}$ and integrate over angle. This gives

$$\left[\frac{1}{c}\frac{\partial}{\partial t} + \mathbf{n}\cdot\nabla\right]\int I_\nu \mathbf{n}(\omega) d\omega = \frac{1}{4\pi}\int \left(\eta_\nu - \chi_\nu I_\nu\right)\mathbf{n}(\omega) d\omega. \tag{34}$$

This can be written

$$\frac{1}{4\pi c}\frac{\partial \mathbf{F}_\nu}{\partial t} + \nabla \cdot \int I_\nu n^2 d\omega = \frac{1}{4\pi}\int \left(\eta_\nu - \chi_\nu I_\nu\right)\mathbf{n}(\omega) d\omega, \tag{35}$$

or equivalently (Mihalas & Mihalas (1984, Eq 78.9))

$$\frac{1}{c}\frac{\partial \mathbf{F}_\nu}{\partial t} + c\nabla \cdot \mathbf{P}_\nu^{ij} = \int \left(\eta_\nu - \chi_\nu I_\nu\right)\mathbf{n}(\omega) d\omega. \tag{36}$$

For an isotropic η_ν, the first term on the RHS will become zero and we can write (e.g. Eq. 4.8 in Rutten)

$$\frac{1}{c}\frac{\partial \mathbf{F}_\nu}{\partial t} + c\nabla \cdot \mathbf{P}_\nu^{ij} = -\chi_\nu \mathbf{F}_\nu. \tag{37}$$

This equation expresses momentum conservation of the radiation field. The time rate of change of momentum is the transport of momentum across the boundary surface (second term on LHS), plus addition of momentum by emission (first term on RHS), minus removal of momentum by absorption (second term on RHS). Then, the radiative force due to absorption must equal this last term

$$\mathbf{f}_R = \frac{1}{c}\int\int \chi_\nu I_\nu \mathbf{n}(\omega) d\omega d\nu. \tag{38}$$

In spherical symmetry, the first moment equation (with isotropic scattering) becomes

$$\frac{1}{c^2}\frac{\partial F_\nu}{\partial t} + \frac{\partial P_\nu}{\partial r} + \frac{3P_\nu - E_\nu}{r} = -\chi_\nu F_\nu \tag{39}$$

6.3 Closure

These moment equations do not contain I_ν, but only angle-integrated quantities of I_ν ; J_ν, $\mathbf{F}_\nu$ and $\mathbf{P}_\nu{}^{ij}$. Thus, the dimensionality of the problem is reduced by 2 in multi-D problems (both 2D and 3D require two angles), and 1 in spherical symmetry. This is the main motivation for seeking solutions to them instead of to the full transfer equation.

However, we always have at least one unknown moment (possibly non-scalar) more than the number of equations. This is called the closure problem. We cannot arbitrarily simplify our problem by just integrating the equations. All the angular details of I_ν need still to be known to define an exact closure relation.

A closure relation can be obtained by various means. One is **variable Eddington factor methods** (often used in non-hydro RT), where an iteration with a formal solver of the regular transfer equation for I_ν (with source function fixed) is carried out. Note there is some subtility here : if we anyway have to solve the transfer equation many times, why don't we just take J etc from its solution directly? The answer is that convergence is faster if we determine J from the moment equations, where we can couple it in implicitly in scattering cases (more later).

Another approach is to use **analytic formulae**, basically fixed approximations for the closure estimated from the geometry (often used in radiation hydro). Examples are "FLD" (flux-limited diffusion) and "M1". This will be covered at depth in the lectures by Evan O'Connor.

Note that in multi-D, we have 4 equations but 10 (J_ν,$\mathbf{F}_\nu$,$\mathbf{P}^{ij}_\nu$) unknowns ($\mathbf{P}^{ij}_\nu$ is symmetric). Thus we need not one closure relation but 6 in each cell. In spherical symmetry we have 2 equations (no 32, 39) and 3 unknowns (J_ν, F_ν, P_ν), so a single closure relation is sufficient.

7 Time-dependence

A central question for many problems is whether the time-derivative term in the transfer equation is needed or not. A good discussion of this is given in Mihalas & Mihalas (Section 6.5 in 1984).

Define a **optically thin radiation flow timescale t_R:**

$$t_R = \frac{l}{c}, \tag{40}$$

where l is a length scale of the problem. If the distance l is optically thick ($l/\lambda = \tau \gg 1$), it takes instead a **diffusion time scale** t_d

$$t_d = \frac{l^2}{c\lambda} = \tau t_R \tag{41}$$

for a photon to travel the distance ($\tau^2 = l^2/\lambda^2$ scatterings, each taking a time λ/c).

Define also a **fluid flow timescale**

$$t_f = \frac{l}{v}, \tag{42}$$

where v is a fluid velocity scale.

Regimes for the transfer:

1. **(Effectively) optically thin regime.** If the gas is optically thin, $t_R/t_f = v/c$.

   Mihalas & Mihalas (1984): "If v/c $\ll$ 1, the radiation field at any position adjusts essentially instantaneously to changes in physical conditions".

   For example, if the radiation field changes because the density is getting lower in an optically thin outflow with $v/c \ll 1$, the time-derivative can be dropped.

   If we care about the effect of radiation on the fluid flow, we need to resolve the fluid flow timescale t_f. If $v/c \ll 1$, $t_R/t_f \ll 1$, and a change in fluid conditions changes the radiation field on a time-scale much shorter than the fluid flow time scale.

   Note that this line of reasoning assumes that changes in emission and absorption are related to changes in flow, i.e. occur on time scale t_f.

Mihalas & Mihalas (1984) denotes this regime **quasi-static**.

Note that if the region is optically thick, but τ is small enough and/or v small enough that $t_d/t_f = \tau v/c < 1$ (i.e. $1 \ll \tau \ll c/v$). Then we come to the same regime as the optically thin case.

2. **Optically thick, frozen diffusion regime**. Now assume the region is optically thick and $t_d/t_f \gg 1$. Then the formal diffusion time is long compared to the flow time, and it would appear we need to solve for the radiation transport time-dependently.

   However, if the optical depth is high, for many situations the radiation will interact with the matter in a thermalizing manner regularly and that 'resets' conditions. Then, the 'effective' time that photons travel without being destroyed is not t_d but something smaller.

3. **Optically thick, dynamic diffusion regime.** Assume now τ is large enough that $t_d/t_f = \tau v/c > 1$, but not so large as to $t_d/t_f \gg 1$, i.e. $t_d \sim t_f$.

   Mihalas & Mihalas (1984) states that here we must include the time-derivative term, and solve on the fluid time scale.

4. **External power source.** If some time-varying external power source, such as a neutrino source or radioactivity, governs the radiation field, the radiation field can change on the time-scale t_s of that source, which may be shorter than the flow time scale t_f. Thus, if we take $t_R = R/c = v/ct$, we need $t_R \ll t_s$ for a time-independent solution, or

$$t \ll \frac{c}{v} t_s. \tag{43}$$

   For the optically thick, dynamic diffusion situation (No 3 above), we have to solve time-dependently on the shorter time-scale of t_s and t_f.

Discussion : discuss the treatment if time your current RT application.

8 Transfer equation in a moving medium

8.1 Reference Frames

There are three reference frames which are of fundamental importance in radiative transfer theory.

1. **Atom frame ("AF")**. The frame in which the individual atom (or electron) undergoing an interaction has zero velocity.
2. **Lab frame ("LF","observer-frame")**. The frame in which observer, or sometimes star (or equivalent system) as a whole, is at rest.
3. **Comoving frame ("CMF", "Lagrangian frame")**. Frame at (**x**,t) in which the integral of velocity vectors in a small volume around **x** is zero. Since the matter seen as "bulk" is at rest in this frame, the opacity, and often also the emissivity, are isotropic. This is the key property that makes the CMF attractive to solve RT equations in.

Note that the CMF is in general not an inertial frame, as the matter bundle, and therefore the frame, can accelerate. The CMF of a given bundle should be understood as a time-sequence of inertial frames, each moving with the instantaneous velocity of the fluid element under consideration.

Combining energy and momentum gives a four-vector and Lorentz transforming it yields two important special relativistic effects in radiative transfer, namely the **Doppler shift**:

$$\nu_0 = \gamma\nu\,(1 - \mathbf{n}\cdot\mathbf{v}/c)\,, \tag{44}$$

$$\nu = \gamma\nu_0\,(1 + \mathbf{n}_0\cdot\mathbf{v}/c)\,, \tag{45}$$

and the **aberration**

$$\mathbf{n}_0 = (\nu/\nu_0)\,(\mathbf{n} - \gamma\mathbf{v}/c)\left[1 - \frac{\gamma\mathbf{n}\cdot\mathbf{v}/c}{\gamma+1}\right], \tag{46}$$

$$\mathbf{n} = (\nu_0/\nu)\,(\mathbf{n}_0 + \gamma\mathbf{v}/c)\left[1 + \frac{\gamma\mathbf{n}_0\cdot\mathbf{v}/c}{\gamma+1}\right]. \tag{47}$$

Here, we have adopted the nomenclature of Mihalas & Mihalas (1984) and denote quantities evaluated in the CMF with a subscripted '0'. In 1D, the aberration formulae simplify to

$$\mu_0 = \frac{\mu - v/c}{1 - v/c\mu}, \tag{48}$$

$$\mu = \frac{\mu_0 + v/c}{1 + v/c\mu_0}. \tag{49}$$

Using a series of simple "Gedankenexperiments", Thomas (1930) derived the transformation laws for the specific intensity

$$I(\nu, \mu) = \left(\frac{\nu}{\nu_0}\right)^3 I_0(\nu_0, \mu_0), \tag{50}$$

$$\eta(\nu, \mu) = \left(\frac{\nu}{\nu_0}\right)^2 \eta_0(\nu), \tag{51}$$

$$\chi(\nu, \mu) = \left(\frac{\nu}{\nu_0}\right)^{-1} \chi_0\left(\nu_0\right). \tag{52}$$

8.1.1 Moments

TBD...

8.2 Comoving frame formulations

When the fluid is accelerating (as e.g. during a SN explosion) or there is velocity gradient (as in the coasting phase of a supernova), it is difficult to solve the transfer equation in the observer frame, because the motions cause η_ν and χ_ν to become anisotropic. One approach is to simplify the RT equation with first-order expansions for η_ν and χ_ν. This method is described in section 93 in Mihalas & Mihalas (1984). The limitation of this method is when lines are important; the first-order expansions are then insufficient. One then normally works in the comoving frame, described in section 95 in Mihalas & Mihalas (1984). The main advantage of the comoving frame is the isotropy of χ_ν (and sometimes η_ν) and easier matter-radiation interaction calculations. The drawback is a more complex transfer equation. The equation here not only has coupling in space and time, but also in frequency and angle.

The fully relativistic comoving transfer equation is, even in spherical symmetry, a very long expression (Eq. 95.9 in Mihalas & Mihalas 1984). The moment equations are also lengthy (Eqs. 95.11 and 95.12). If we retain only terms to order v/c, and also ignore the fluid acceleration term a, we get in spherical symmetry (Eq. 95.17 MM84):

$$\begin{aligned}\frac{1}{c}\frac{DI_0}{Dt} + \frac{\mu}{r^2}\frac{\partial}{\partial r}\left[r^2 I_0\right] + \frac{\partial}{\partial \mu_0}\left(\left(1-\mu_0^2\right)\left[\frac{1}{r} + \frac{\mu_0}{c}\left(\frac{v}{r} - \frac{\partial v}{\partial r}\right)\right] I_0\right) \\ - \frac{\partial}{\partial \nu_0}\left(\nu_0\left[\left(1-\mu_0^2\right)\frac{v}{cr} + \frac{\mu_0^2}{c}\frac{\partial v}{\partial r}\right] I_0\right) + \left[\left(3-\mu_0^2\right)\frac{v}{cr} + \frac{\left(1+\mu_0^2\right)}{c}\frac{\partial v}{\partial r}\right] I_0 \\ = \eta_0 - \chi_0 I_0\end{aligned}$$

where $D/Dt = \partial/\partial t + v\partial/\partial r$. The many terms have their origin from several effects, all related to the fact that the comoving frame is not an inertial frame; a Lagrangian path segment implies changes in the comoving angle and frequency. Compare e.g. to the equation solved by the `CMFGEN` supernova code (Hillier & Dessart 2012), which uses $v/c \ll 1$ and assumes homology ($dv/dr = v/r$):

$$\frac{1}{c}\frac{\partial I_0}{\partial t} + \frac{\mu_0 c + v}{c}\frac{\partial I_0}{\partial r} + \frac{\left(1-\mu_0^2\right)}{r}\frac{\partial I_0}{\partial \mu_0} - \frac{v\nu_0}{rc}\frac{\partial I_0}{\partial \nu_0} + \frac{3v}{rc} I_0 = \eta_0 - \chi_0 I_0. \tag{53}$$

9 RTE solutions with scattering

If emission and absorption coefficients are numerically known, numeric solution to the transfer equation in any of its formats is straightforward - not necessarily easy but in principle not problematic. Such solutions are called *formal solutions*. Some very simple problems fall in this category, and the challenge lies mainly in designing as fast as possible algorithms. Formal solutions are discussed in the next chapter.

Normally, however, η and χ depend on the radiation field in some way. One solution approach is then iteration; alternate formal solutions with recomputation of η and χ. This is called Λ-iteration, and in some problems such brute-force split-up of the equation system and solution with iteration works well, in particularly when optical depths are low or moderate. Indeed, with modern computing power this approach should probably be the first one considered. This approach may not be particularly fast. But many alternative 'clever' approaches invented in the RT literature since 1960s were sought mainly because computers were too slow to do this iteration in a reasonable time. This may not at all be the case today, and little is computationally gained from studying and implementing certain alternative techniques (although they may still provide physical insight).

However, when optical depths are high, Λ-iteration can converge so slowly, that for practical purposes it can become a non-convergent method. One of the main problems is how to assess the convergence (see Fig. 1). When this is so, this equation system splitting has to be avoided or reduced.

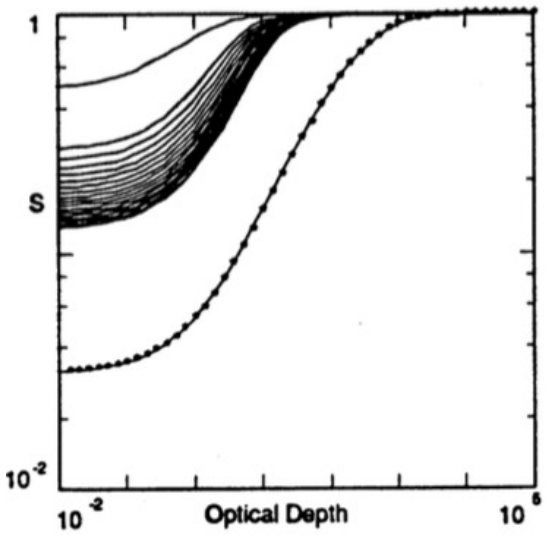


Figure 1: Illustration of convergence problems in Λ-iteration, from Auer (1991). The true solution of S is the dotted line. The solid lines show successive iterations. At $i = 20$, changes are so slow that an **apparent false convergence** is obtained. The true solution is eventually obtained : but only after about 1000 iterations (iterations 21-999 not plotted).

Another approach is to solve explicit scattering problems by brute force : the equation system is of size $N_s N_a N_f$, and we may attempt matrix inversion, which costs $N_s^3 N_a^3 N_f^3$. But if we have say 100 points in each of these three dimensions, we need to store a matrix of size $10^6 \times 10^6$, which is not possible. Even if it was, inversion would cost $\sim 10^{18}$ flops.

Instead, better solution approaches are one of three kinds; the **Feautrier method**, the **Rybicki method**, and the **Variable Eddington Factor** method. Some confusion easily arises, because variants of the Feautrier method are sometimes also applied as part of Variable Eddington Factor methods (as formal solvers). Different articles and books can appear contradictory or inconsistent because authors can differ in their exact definition of the term 'Feautrier method'.

9.1 The Feautrier method

For illustration of this method, we first consider a problem in planar geometry:

$$\mu \frac{dI_\nu}{d\tau_\nu} = I_\nu(\mu) - S_\nu(\mu). \tag{54}$$

For rays going inward (-) and outward (+):

$$\mu \frac{dI_\nu^+}{d\tau_\nu} = I_\nu^+(\mu) - S_\nu^+(\mu), \tag{55}$$

$$-\mu \frac{dI_\nu^-}{d\tau_\nu} = I_\nu^-(\mu) - S_\nu^-(\mu). \tag{56}$$

Adding these gives:

$$\mu \left[\frac{dI_\nu^+}{d\tau_\nu} - \frac{dI_\nu^-}{d\tau_\nu} \right] = \left[I_\nu^+(\mu) + I_\nu^-(\mu) \right] - \left[S_\nu^+(\mu) + S_\nu^-(\mu) \right]. \tag{57}$$

Subtracting them gives:

$$\mu \left[\frac{dI_\nu^+}{d\tau_\nu} + \frac{dI_\nu^-}{d\tau_\nu} \right] = \left[I_\nu^+(\mu) - I_\nu^-(\mu) \right] - \left[S_\nu^+(\mu) - S_\nu^-(\mu) \right]. \tag{58}$$

The last equation can also be written

$$\left[I_\nu^+(\mu) - I_\nu^-(\mu) \right] = \mu \left[\frac{dI_\nu^+}{d\tau_\nu} + \frac{dI_\nu^-}{d\tau_\nu} \right] + \left[S_\nu^+(\mu) - S_\nu^-(\mu) \right], \tag{59}$$

which when put back into Eq. 57, combined with an assumption that S is isotropic (so $S^+(\mu) + S^-(\mu) = S_\nu$ and $S^+(\mu) - S^-(\mu) = 0$) gives

$$\mu^2 \frac{\partial \bar{I}_\nu(z,\mu)}{\partial \tau_\nu^2} = \bar{I}_\nu(z,\mu) - S_\nu(z), \tag{60}$$

where

$$\bar{I}_\nu(\mu) \equiv \frac{1}{2} \left[I_\nu(\mu) + I_\nu(-\mu) \right]. \quad (0 \le \mu \le 1). \tag{61}$$

Writing out the source function (isotropy not enforced):

$$\mu^2 \frac{\partial \bar{I}_\nu(z,\mu)}{\partial \tau_\nu^2} = \bar{I}_\nu(z,\mu) - S_\nu^t(z) + \frac{\int_0^\infty \int_0^1 \sigma(z, \mu' - \mu, \nu' - \nu) \bar{I}_{\nu'}(z,\mu') d\mu' d\nu'}{\chi_\nu(z,\mu)} \quad (0 \le \mu \le 1). \tag{62}$$

Note that the demand of scattering isotropy already removed lab-frame differential velocity field problems, although some tricks salvage some such situations (Rutten, page 118).

At first glance it looks like each equation constains $\bar{I}_{\nu'}(\mu')$ for all angles and frequencies, which we cannot know unless solving for all angles and frequencies at once, e.g. a full coupling, and full inversion is needed. But Feautrier (1964) came up with a method that is significantly cheaper. The essence lies in that the matrix system is for the right ordering *block tridiagonal* : for each angle and frequency combination, there is only coupling in space between three depth points (any second order spatial derivative has this property). Then, one can come up with a backsubstitution algorithm that is cheaper than brute force matrix inversion. Thus, Feautrier's method is in essence a pure 'math trick'. But the method (or very similar methods) are used also in alternative approaches such as VEF : it is therefore quite broad in its application and important to understand.

The solution details are given in Appendix A. The essence is that each matrix inversion is now limited to matrixes of size $N_a N_f$. The method requires four such inversions at each depth point, so in total a computational inversion cost of $4N_s N_a^3 N_f^3$ instead of $N_s^3 N_a^3 N_f^3$ in the fully coupled brute force method, a gain by factor $N_s^2/4$, or 2500 for $N_s = 100$. With samplings of order 100 points in each dimension, CPU time (1 GHz) is about 24 hours, but drops to seconds in the gray or coherent scattering case.

According to Castor (2004), the Feautrier method has been used "in the vast majority of slab geometry work since 1964". One if its advantages is that it works with the second order form of the transfer equation, which is second order accurate and numerically more benign than the first order form (Castor 2004).A natural application area is when we want to avoid Lambda iteration (high optical depth) and explicitly include scattering directly in the solution. Rutten (page 122) describes that also when S is numerically known, the Feautrier method is a very efficient Lambda operator, faster than integral solutions (e.g. using exponential functions). Thus, the method finds use also in Lambda iterations.

The method scales with $N_a^3 N_f^3$; thus when very high resolution is needed in these dimensions, the method may become expensive. The Rybicki method (next section) is an alternative similar method which is preferred in certain problems where N_f in particular is large.

9.1.1 Spherical symmetry

A similar second-order equation can be derived in spherical symmetry (see more below), so the method generalizes to spherical symmetry. This statement holds strictly for isotropic scattering, Mihalas 1986 tells us 'Observer-frame partial redistribution (non-isotropic) calculations in spherical symmetry have never been carried out using Feautrier variables.' (but one goes to comoving frame).

9.1.2 Multi-D

A generalization of the Feautrier method to 2D (and sketching for 3D) is presented in Cannon 1970. One of the spatial dimensions can be 'recursed' away, but matrices of size $N_{d1} N_a N_f$ now need to be inverted, becoming $N_{d1} N_{d2} N_a N_f$ in 3D. A statement from Mihalas 1986 is that 'the approach is quite costly in 2 dimensions'.

9.1.3 Comoving frame

The Feautrier method (or similar derivatives) can be applied for comoving frame formulations, see e.g. Noerdlinger & Rybicki (1974) (plane parallell case), Mihalas et al. (1976) (spherical symmetry). Noerdlinger & Rybicki (1974) uses the two first order moment equations, rather than combining them into a second-order one.

9.2 The Rybicki method

The Rybicki method switches the inner and outer ordering of the Feautrier method, defining vectors for each angle-frequency point instead of for each depth point. This replaces the $N_d N_a^3 N_f^3$ cost of Feautrier with $N_d^2 N_a N_f + N_d^3$ (see MM84 page 377 for details), and is thus preferential for problems with many angle-frequency points. However, it is reported in the literature that there are some subtle caveats with this method for certain problems when coupling to other physical constraints (MM84 page 380, M86).

9.2.1 Spherical symmetry

TBD..

9.2.2 Multi-D

A generalization to 2D is presented in Mihalas et al. (1978).

9.2.3 Comoving frame

TBD..

9.3 The Variable Eddington Factor method

Assume we have isotropic scattering. In the Feautrier method, this means only that the off-diagonal elements (with the same frequency) in the **B** matrix become identical (see appendix A). This does not translate to any computational speedup, and cost remains at $4 N_s N_a^3 N_f^3$. A different approach, where the moment equations are iterated with formal solutions to determine closure, then typically

becomes faster : this is the class of Variable Eddington Factor (VEF) methods.

Integration over angle of Eq. 62 gives ($1/2\int_0^1 \bar{I}_\nu \mu^2 d\mu = 1/2\int_{-1}^1 I_\nu \mu^2 d\mu = K_\nu$):

$$\frac{\partial^2 K_\nu}{\partial \tau_\nu^2} = J_\nu - S_\nu^t - \frac{\int R(\nu,\nu') J_{\nu'} d\nu'}{\chi_\nu}. \tag{63}$$

We have defined earlier the *Eddington factor* $f_\nu \equiv K_\nu / J_\nu$, so we can write this also as

$$\frac{\partial^2 f_\nu J_\nu}{\partial \tau_\nu^2} = J_\nu - S_\nu^t - \frac{\int R(\nu,\nu') J_{\nu'} d\nu'}{\chi_\nu}. \tag{64}$$

This equation has the same form as Eq. 62, except that the variable is now J_ν instead of (the angle-dependent) $\bar{I}_\nu$. There may still be off-diagonal entries in **B** due to frequency redistribution. Thus, Eq. 64 can be solved with a similar back-substitution scheme as the Feautrier method. We have got rid of the angles, bringing each iteration down from $4N_d N_a^3 N_f^3$ to $4N_d N_f^3$. However, iteration is now needed. The total gain factor is then N_a^3/N_{iter}, assuming neglegible computation time for the formal solutions. In many applications a few iterations is sufficient. This benign property comes from the fact that the Eddington factor measures the degree of asymmetry of the radiation field; an integral quantity that is robustly recovered in a few iterations.

9.3.1 Spherical symmetry

An equation can be derived by combination of the zeroth and first moment equations (Eq. 83.74 in MM84):

$$\frac{1}{q_\nu}\frac{\partial}{\partial \tau_\nu}\left[\frac{r^2}{q_\nu}\frac{\partial (f_\nu q_\nu J_\nu)}{\partial \tau_\nu}\right] = \frac{r^2}{q_\nu}\left(J_\nu - S_\nu\right), \tag{65}$$

where q_ν is called the *sphericality factor*, and is fully specified by f_ν:

$$\ln q_\nu = \int_{rc}^{r} \left[(3f_\nu - 1)/r' f)\nu\right] dr'. \tag{66}$$

Eq. 65 has again a second order spatial derivative and can be discretized and solved by e.g. the Feautrier scheme.

9.3.2 Multi-D

TBD..

9.3.3 Comoving frame

In the comoving frame in spherical symmetry, two Eddington factors are needed instead of one (e.g. Mihalas 1986-II).

10 Formal solutions

Formal solutions, for numerically known emission and absorption coefficients (or, equivalently, the source function), need to be carried out either in simple problems where these are known (and fixed) from the outset, or in iterative methods such as Λ-iteration or Variable Eddington Factor methods. This is generally quite straightforward and costs of order $N_d N_a N_f$ in 1D as matrix inversion is generally not needed.

Hubeny & Mihalas (2014) outline three main classes of formal solvers:

1. **Feautrier method.**
2. **Short/long characteristics.** First order transfer integrals on set of discrete ray segments.
3. **Discontinuous finite element method.** Unlike short characteristics, here the source function is allowed to have discontinuities (still polynomial). Does not require exponentials.

10.1 Feautrier method as formal solver

With a known source function, the Feautrier method does only scalar operations, and scales as $N_d N_a N_f$, i.e. as fast as possible (Hubeny & Mihalas 2014). Some improvements regarding accuracy and convergence (e.g. getting a 4th order accurate 'Hermite' method) are discussed in Sec 12.4 of Hubeny & Mihalas (2014).

10.2 Short characteristics

This method was first introduced by Olson & Kunasz (1987). The idea is to make analytic solutions between discrete grid points - cell by cell.[2]

In 1D the standard analytic solutions are of the form

$$I(\tau_d, \mu) = I(\tau_{d-1}, \mu) e^{-(\tau_d - \tau_{d-1})\mu} + \int_{\tau(d-1)}^{\tau(d)} S(t) e^{-(\tau_d - t)/(-\mu)} dt/\mu \quad \mu < 0 \tag{67}$$

$$I(\tau_d, \mu) = I(\tau_{d+1}, \mu) e^{-(\tau_{d+1} - \tau_d)/\mu} + \int_{\tau(d)}^{\tau(d+1)} S(t) e^{-(t - \tau_d)/\mu} dt/\mu \quad \mu > 0 \tag{68}$$

Normally one then lets $S(\tau)$ be represented by some polynomial between the grid points, to allow analytic solutions of the integrals.

- **Linear**. Is sometimes not accurate enough.
- **Quadratic.** The three coefficients are chosen by fitting to three neighboring points.
- **Advanced interpolations** like "Bezier interpolants" or "monotonized splines". Jaime is an expert.

E.g. with a linear interpolation, one gets

$$I(\tau_d, \mu) = I(\tau_{d-1}, \mu) e^{-\Delta\tau_{d-1/2}} + \lambda^-_{d,d} S_d + \lambda^-_{d,d-1} S_{d-1} \quad \mu < 0 \tag{69}$$

where

$$\lambda^-_{d,d} = 1 - \frac{1}{\Delta\tau_{d-1/2}} + \frac{e^{-\Delta\tau_{d-1/2}}}{\Delta\tau}. \tag{70}$$

Note that, the analytic solution of the integral means accuracy is retained also if the optical depth between grid points becomes large - as long as the analytic representation is accurate enough. Here, an advantage of working in optical depth units instead of say spatial units clarifies : the source function is a more slowly varying function than the emission and absorption coefficients (e.g. Rutten 2003, page 14), and thus more accurately represented by a polynomial fit.

Hubeny & Mihalas (2014) do some comparisons between Feautrier method, short characteristics etc, and all methods are basically within a factor 2 of each other for standard test problems. Any formal solver should need of order $N_d N_a N_f$ operations.

10.2.1 Spherical symmetry

Two main approaches (page 380 in MM84).

Tangent ray. Cost is $\sim N_s^2 N_f$ in 1D (M86). A drawback of the tangent ray method is that it cannot be made fully consistent with the moment equations (M86); this is possible instead in the discrete space method (although normally more expensive).

Discrete space. Cost is $\sim N_s N_a^3 N_f$ in 1D (M86), making it preferable when a small number of angle points is needed.

An example : CMFGEN Section 3.3 in Hillier 2012.

[2] Hubeny & Mihalas (2014) remark that the name 'short characteristics' is not ideally chosen, as 'characteristics' normally refers to rays in PDE problems, but in plane parallell static atmospheres there is just an ODE to solve.

10.2.2 Multi-D

TBD..

10.3 Discontinuous finite element method

TBD..

11 Acceleration of Lambda iteration

Rutten section 5.3 and notes to J. Leenarts Master course lab are good.

11.1 Classic Lambda iteration

Classic Lambda iteration refers to a scheme where we alternate solutions to the radiation field with source function fixed (formal solutions) and to the source function (i.e. the gas state). For a given source function, any method to calculate J_ν can be written in operator form (assuming for now S_ν to be isotropic)

$$J_\nu = \Lambda_\nu[S_\nu]. \tag{71}$$

We may also talk about a Lambda operation as calculating I and not J; we then add a μ as second subscript

$$I_{\nu\mu} = \Lambda_{\nu\mu}[S_{\nu\mu}]. \tag{72}$$

For example, if our method is to do a formal integral solution, $\Lambda_\nu[S_\nu]$ performs integrals in various directions, sampling S_ν from each cell. But it may also be a computer program that implements the (differential) Feautrier scheme (fixed source function variant). Some formal solution methods may correspond to a matrix multiplication with the source function vector, but some involve multi-step procedures that cannot be written in such a way.

The work flow is then to initially guess $J_\nu^{n=0}$ (for all frequencies and depths), then loop over the steps

1. Compute $S_\nu^n = f(J_\nu^n)$ (all frequencies).
2. Compute $J_\nu^{n+1} = \Lambda_\nu[S_\nu^n]$.

The idea is that this is faster and more manageable than solving the full coupled system, which for a test problem of $S = \varepsilon B + (1-\varepsilon)J$ can be written

$$\mathbf{S} = [\mathbf{I} - (1-\varepsilon)\mathbf{\Lambda}]^{-1}[\varepsilon\mathbf{B}]. \tag{73}$$

For example, a problem with 10^4 depth points is too large for direct solution by such inversion, and for complex problems with angle and/or frequency coupling it is not possible even for small number of depth points (there are limitations from both CPU time and RAM memory).

11.1.1 Convergence issues

Classic Lambda iteration is a perfectly good, and often the best, technique for low and moderate optical depths (e.g. Rutten). For high optical depth, however, convergence of classic Lambda iteration is poor. This is because in a pure iteration, information propagates one optical depth at a time. For an optical depth of τ, photons carry out of order τ^2 scatterings, so as τ gets high the number of iterations needed becomes very large.

This is, however, not the worst part of the story. The more malicious property of the iteration scheme is that correction steps start decreasing with each iteration long before convergence is achieved, giving rise to 'false convergence impression'. As the true solution is (obviously) not known, we cannot easily probe how 'good' the current very slowly changing solution is.

Consider the example of LTE gas with coherent scattering. In the iteration scheme this becomes (assume T is known for now)

$$S_\nu^{n+1} = \varepsilon B_\nu(T) + (1-\varepsilon)\Lambda_\nu[S_\nu^n] \tag{74}$$

Information must propagate to the thermalization depth, which is $\tau_t \sim 1/\sqrt{\varepsilon}$, which requires at least $1/\varepsilon$ iterations.

Significance of ϵ. Does convergence improve of ϵ is large? It does if there is no strong coupling between radiation field and gas state. But if there is such coupling, it is less clear how important ϵ is.

A general iterative scheme can be written as

$$x^{n+1} = Fx^n + b'. \tag{75}$$

The F matrix is called the **amplification matrix**, and its properties determines convergence properties. One may show

$$e^{n+1} = Fe^n, \tag{76}$$

where e is the error vector. Requiring that $||e^{n+1}||/||e^n|| < 1$ requires $||F|| < 1$. Hubeny & Mihalas states 'there are several possible choices for norm, we choose the spectral radius which equals the largest eigenvalue σ". Thus, one gets divergence for $\sigma > 1$, fast convergence for $\sigma \ll 1$, and slow convergence for $\sigma \sim 1$.

In classic Lambda iteration $F = (1 - \varepsilon)\Lambda$. In the optically thick limit $\Lambda \approx 1$ on the diagonal, and if $\varepsilon \ll 1$, $\sigma \sim 1$, predicting bad convergence.

11.2 Accelerated Lambda iteration

We can view full coupling and direct inversion (Eq. 73) as one extreme solution approach, and classic Lambda iteration as the other extreme, in which the equation system blocks for gas state and RT are completely decoupled in each iteration.

"Accelerated Lambda iteration" refers to schemes that are positioned somewhere in between these extremes; there is *some* coupling of the source function to the *current* (to be calculated) radiation field. This coupling is selected in such a way that each iteration is much faster than in the full coupling regime, and each iteration brings a bigger step towards the solution than in classic Lambda iteration (because of the partial coupling).

This class of methods are sometimes also called *approximate lambda iteration* or *operator perturbation* methods. The first paper introducing them for astrophysics was Cannon (1973), although the philosophy did not take off until the papers by (Scharmer 1981, 1983) and Olson et al. (1986, "OAB"). Note that these techniques were known from other areas of science long before Cannon brought them into the RT field.

The basic idea is to split the Lambda operator into two parts:

$$\Lambda = \Lambda^* + (\Lambda - \Lambda^*). \tag{77}$$

Here Λ^* is called the approximate Lambda operator. Then, the Λ^* part is ("implicitly") linked to act on the *next* S estimate, so change the iteration strategy to

$$S_\nu^{n+1} = \epsilon B_\nu(T) + (1 - \epsilon)\Lambda^*[S_\nu^{n+1}] + (1 - \epsilon)\,(\Lambda - \Lambda^*)\,[S_\nu^n]. \tag{78}$$

We have now made a "semi-implicit" setup where part of the operator acts on the current solution, and the other part acts on the previous solution. It is clear that *if* we achieve convergence, this is identical to Eq 74, so the solution will be the correct one.

We can rewrite, or 'solve', Eq. 78 for S_ν^{n+1}: (we should combine the equations for each depth, for each of which the Lambda operators define a particular row in a matrix), now also letting Λ specifically be a matrix

$$\begin{aligned} \mathbf{S}_\nu^{n+1} &= [\mathbf{I} - (1 - \varepsilon)\mathbf{\Lambda}^*]^{-1}\,[\varepsilon\mathbf{B} + (1 - \varepsilon)\mathbf{\Lambda}[\mathbf{S}_\nu]^n - (1 - \varepsilon)\,\mathbf{\Lambda}^*[\mathbf{S}_\nu^n]] && (79)\\ &= [\mathbf{I} - (1 - \varepsilon)\mathbf{\Lambda}^*]^{-1}[\mathbf{S}_\nu^{FS,n} - (1 - \varepsilon)\mathbf{\Lambda}^*[\mathbf{S}_\nu^n]]. && (80) \end{aligned}$$

Solution requires inversion of a matrix that is the unity matrix minus a (scaled) $\mathbf{\Lambda}^*$ matrix. If we e.g. choose $\mathbf{\Lambda}^*$ to be diagonal, this will be much faster than inverting the quantity in Eq. 73. Note that we also have to perform 'perturbed formal solutions' $\Lambda^*[S^n]$ in each iteration.

If Λ^* carries enough of the full physics of Λ, we will now get bigger steps towards the true solution in each iteration compared to formal solutions because of the implicit coupling. To make any gain compared to the fully implicit approach, Eq. 78 must also be solvable with less CPU time and/or memory than Eq. 73 (otherwise we might as well do the fully implicit approach). Eq. 80 shows that Λ^* needs to be more easily inverted than Λ.

Some insight of the acceleration gained can be seen by writing

$$S^{FS,n} = (1-\epsilon)\Lambda[S^n] + \epsilon B, \tag{81}$$

so

$$(S^{n+1} - S^n)_{classic} = S^{FS,n} - S^n. \tag{82}$$

With an approximate operator, on the other hand

$$(S^{n+1} - S^n)_{ALI} = \left[1 - (1-\epsilon)\Lambda^*\right]^{-1} \left[S^{FS,n} - S^n\right]. \tag{83}$$

For an heuristic $\Lambda^* = 1$ (e.g. large optical depth $J \approx S$), we get an acceleration of order $1/\epsilon$ (e.g. Eq 5.45 in Rutten) by having coupling. **Discuss: works only for Λ close to 1.**

Note that from Eq. 83, iteration may preceed by calculating normal formal solutions, then letting $\left[1 - (1-\epsilon)\Lambda^*\right]^{-1}$ operate on $S^{FS} - S^n$, and subtract S^n (so no need to let Λ^* operate directly on anything).

11.2.1 Choices of Λ^*

The approximate operator Λ^* should have the desired properties

- Represent much of the basic physical properties of Λ, to make sure a significant part of the problem gets 'implicitly treated'.
- Be fast and/or require less memory for construction and inversion than Λ.

Needless to say, this wishlist opens up a more or less infinite zoo of possible operators, each better or worse for particular applications, and there is a vast literature on this topic.

In the classic literature terms like "diagonal" and "tridiagonal" dominate; the example problem is almost always a mono-frequency 1D stellar atmosphere, so what varied along the array was just the depth point. "Diagonal" then means only the local cell is coupled, while "tridiagonal" also the neighboring shells/sheets are coupled. In a more generalized context, one may talk about the size of the local region for which implicit coupling is expressed, "diagonal" then corresponds to the local cell, "tridiagonal" to nearest neighbors (e.g. 6 cells in a Cartesian 3D grid) etc.

Diagonal (local) operator. In this approach Λ^* contains only contributions by emission in the local cell to create J. In the 1D gray case that means the diagonal elements of the matrix are the only non-zero ones, and at each depth point there is a simple scalar division. This method was initially presented by Olson et al. (1986), and is also called the "OAB method". The original scheme uses Feautrier variables and the 2nd order form of the transfer equation (Olson & Kunasz 1987). The formal solution is done with the 'Hermite (differential) scheme' of Auer (1976) which is a variant of the Feautrier method. Local operators can be easily generalized to multi-D (OAB, Rutten).

What may this actually look like in practice? Note that a diagonal operator does not have to be the exact diagonal of Λ. We are free to choose any approximate formula we want. Some examples are

- **Escape probability (OAB).** $\Lambda^* = 1 - \beta p(\tau)$, where β is a free parameter and $p(\tau)$ is an escape probability.

- **Approximate second order solution (AOB86)**. From

$$u''(\mu,\tau) = u(\mu,\tau) - S(\tau) \tag{84}$$

Combine with the approximations

$$u_{i-1} = u_i e^{-\Delta\tau_{i,i-1}/\mu}, \tag{85}$$
$$u_{i+1} = u_i e^{-\Delta\tau_{i,i+1}/\mu} \tag{86}$$

To get

$$u_i = S_i \frac{1}{\Delta_1\Delta_2}\left[-e^{-\Delta_1} - e^{-\Delta_2} + 2 + \Delta_1\Delta_2\right]^{-1}. \tag{87}$$

Then

$$\Lambda_{ii} = \int \phi_x \int u_i d\mu. \tag{88}$$

- **Gray case.** Here can we use $E_1(0)$ from $J = \int SE_1(t)dt$.

$$J_\nu(\tau_\nu) = \int S_\nu(t_\nu)E_1(|t - t_\nu|)dt_\nu \tag{89}$$

- **Short-characteristics** (Olson & Kunasz 1987). Used also in multi-D (Hubeny 2003).

The formal solution using first-order (linear interpolation) short characteristics is

$$I(\tau_d,\mu) = I(\tau_{d+1},\mu)e^{-\Delta\tau_{d+1/2}} + \lambda^+_{d,d}S_d + \lambda^-_{d+1,d}S_{d+1} \quad \mu > 0 \tag{90}$$
$$I(\tau_d,\mu) = I(\tau_{d-1},\mu)e^{-\Delta\tau_{d-1/2}} + \lambda^-_{d,d}S_d + \lambda^-_{d,d-1}S_{d-1} \quad \mu < 0 \tag{91}$$
$$\tag{92}$$

where

$$\lambda^+_{d,d} = 1 - \frac{1}{\Delta\tau_{d+1/2}} + \frac{e^{-\Delta\tau_{d+1/2}}}{\Delta\tau_{d+1/2}} \tag{93}$$
$$\lambda^-_{d,d} = 1 - \frac{1}{\Delta\tau_{d-1/2}} + \frac{e^{-\Delta\tau_{d-1/2}}}{\Delta\tau_{d-1/2}} \tag{94}$$
$$\tag{95}$$

where $\Delta\tau_{d-1/2} = (\tau_d - \tau_{d-1})/\mu$. Thus, keeping only the local term, and constructing J from I:

$$\Lambda^*_\nu = 1 - \frac{1}{2}\int_0^1 \left(\frac{1 - e^{-\Delta\tau_{d-1/2}}}{\Delta\tau_{d-1/2}} + \frac{1 - e^{-\Delta\tau_{d+1/2}}}{\Delta\tau_{d+1/2}}\right) d\mu. \tag{96}$$

Thus, to make the iteration we just multiply $S^{FS} - S^n$ with a scalar correction factor that corresponds to this integral calculation, that for a fixed grid is constant through the iterations.

- **Feautrier scheme** (Rybicki & Hummer 1991). A drawback of the short characteristic method is the costly evaluation of exponentials. Rybicki & Hummer (1991) showed that one can obtain the exact diagonal of the Λ matrix in a cheap way by a Feautrier-like approach. For a given frequency and angle, the solution to u can in general be written

$$u = \mathbf{T}^{-1}\mathbf{S}, \tag{97}$$

where u is the symmetric average of I, $\mathbf{T}$ is tri-diagonal, and there are N_d equations. From u we get J from

$$\Lambda = \int \mathbf{T}^{-1} d\mu. \tag{98}$$

Thus we need the diagonal elements of $\mathbf{T}^{-1}$. Rybicki & Hummer (1991) showed how these can be found through a recursive sweep algorithm with order N_d operations (see e.g. Hubeny and Mihalas page 438 for details).

With a diagonal operator, convergence decreases with increased resolution (OAB).

Tridiagonal/one neighbour operator. This method is less sensitive to grid cell size than the diagonal one. It was introduced by Olson & Kunasz (1987), who suggested to use short characteristics to find the coefficients.

Pentadiagonal and higher-order operator. See e.g. Hauschildt 1992, 1994, MacFarline1992. From extensive trial and error experiments reported in the literature, the benefits of fewer iterations obtained by going to higher bands than tridiagonal has not turned out to outweigh the larger computational costs, and Hubeny and Mihalas state that these methods are rarely used.

Upper tridiagonal operator. Scharmer 1981 ("core saturation" method). Eddington-Barbier. Trujillo Bueno 1995. Upper diagonal does not extend to multi-D (Olson et al. 1986). Olson & Kunaz 1987, using short characteristics.

- **Core saturation operator**. This method relates to transfer through lines. At high optical depth we know that $J_\nu \approx S_\nu$, so with a one-quadrature formula (Scharmer 1981), we simply get $\Lambda^*_{ii} = 1$. For optically thin case one may instead use the Eddington-Barbier approximation, $J = 1/2 S_\nu(\tau_\nu = \gamma)$.What is counted as optically thick can be controlled with a parameter γ.

- **Scharmer operator**. The Scharmer operator relates to calculation of I, not J. Scharmer set out to find functional forms that have the Eddington-Barbier form ($I_\nu = S_\nu(f(\mu))$

$$\Lambda^*_{\nu\mu} = W^{\pm}_{\nu\mu} S_{\nu\mu}(f_{\nu\mu}(\tau_{\nu\mu})) \tag{99}$$

  Eddington-Barbier approximation: Consider the transfer equation

$$\frac{dI_{\nu\mu}}{d\tau_{\nu\mu}} = I_{\nu\mu} - S_\nu \tag{100}$$

  Solutions are

$$I^+_{\nu\mu} = \int_\tau^\infty S_\nu(t) e^{-(t-\tau)} dt \tag{101}$$

$$I^-_{\nu\mu} = \int_0^\tau S_\nu(t) e^{-(\tau-t)} dt \tag{102}$$

  Now linearise the source function, $S_\nu = a + b\tau_{\nu\mu}$. This gives

$$I^+_{\nu\mu} = a + b + b\tau_{\nu\mu} = 1 \times S_\nu(\tau = \tau_{\nu\mu} + 1) \tag{103}$$

$$I^-_{\nu\mu} = a - b + b\tau - (a-b)e^{-\tau} = \left(1 - e^{-\tau}\right) \times \left(a - b + \frac{b\tau}{1 - e^{-\tau}}\right) \tag{104}$$

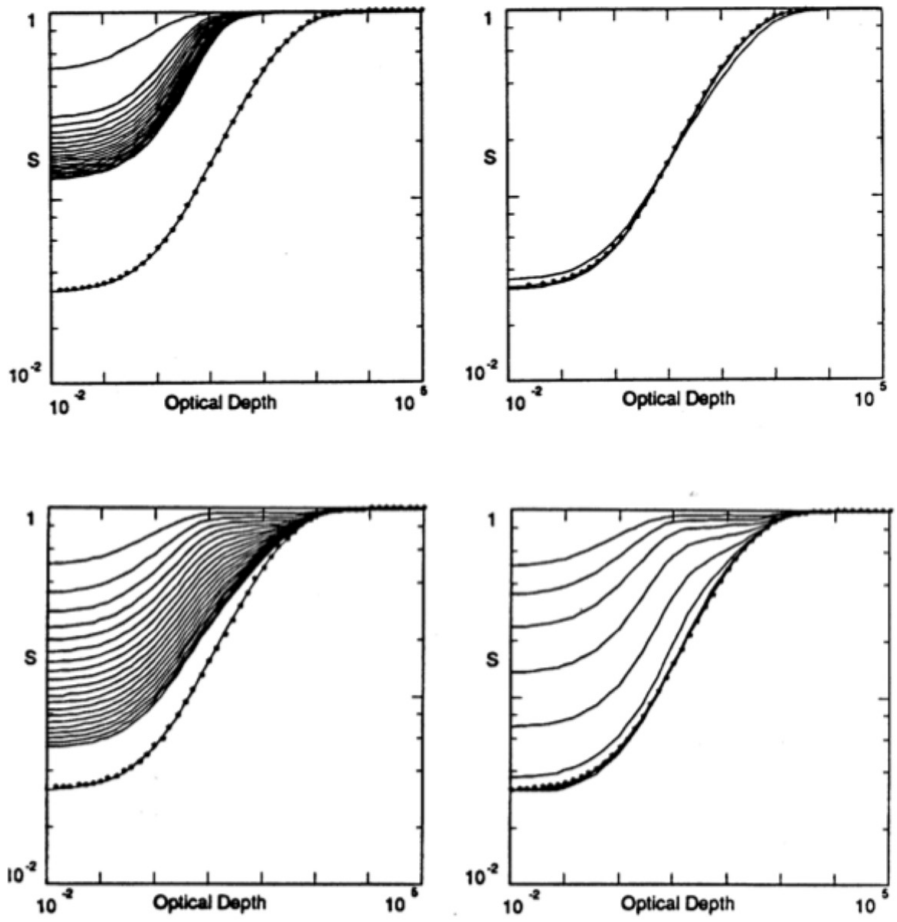
$$= \left(1 - e^{-\tau}\right) S_\nu(\tau = \frac{b\tau}{1 - e^{-\tau}} - 1) \tag{105}$$

  One here uses the property that for a linear source function, the exact solution corresponds to S at one particular distance away. Thus, the matrix elements in Λ^* have values $W^+ = 1$ for outgoing beams and $W^- = 1 - e^{-\tau}$ for ingoing beams, and their locations are given by $\tau + 1$ for outgoing and $\frac{b\tau}{1-e^{-\tau}} - 1$ for ingoing. Each row in the Λ^* matrix has exactly one entry.

  This operator was implemented together with a complete linearization scheme for level populations in Scharmer and Carlsson 1985. An advantage of the Scharmer operator is that it allows for anisotropy in S (Rutten). However it does not generalize to multiD.

11.2.2 Spherical symmetry

Hamann 1985, 1986, 1987, Hempe & Schoenberg 1986, Schoenberg and Hempe 1986, Puls & Herrero: spherical case (sometimes velocity gradient). Diagonal operators.
Hillier 1990 (CMFGEN) : spherical, velocity gradient, multilevel. Tridiagonal, pentadiagonal operator, but note that there are some subtle differences between Hilliers method and classic ALI, which leads Hubeny and Mihalas to classify them as "approximate Newton Raphson" rather than ALI.



11.2.3 NLTE multilevel case

The more generic multi-level problem can be stated

$$J_\nu = \Lambda[S_\nu] \tag{106}$$
$$\mathbf{n} = f_1(J) \tag{107}$$
$$S_\nu = f_2(\mathbf{n}) \tag{108}$$

These equations may also be supplemented by an energy equation (radiative equilibrium for stellar atm.) and hydrodynamic equations (hydrostatic equilibrium).

The level populations are normally considered separate for each species I (all ion. stages), and when there is coupling (e.g. with charge transfer) one can typically use the last iterate for the partner populations (Hubeny & Mihalas Ch 18). Treating also n_e as fixed, the steady state rate equations form a linear system

$$\mathbf{A}_I \mathbf{n}_I = \mathbf{b}_I \tag{109}$$

One of these equations has to be replaced by a number conservation equation. This choice is arbitrary. The set of equation blocks for each species I is complemented with a charge neutrality equations, matching n_e with the sum of all ions.

Eqs. 108 provide a large set of coupled, non-linear equations (e.g. in rate equations level populations are multiplied with J). "Complete linearization" is the term RT theorists use for solving this full system by Newton-Raphson. All variables can be put together into a "state vector" in each cell

$$\psi_d = \{J_1, .., J_{NF}, T, n_e, n_1, .., n_{NL}\} \tag{110}$$

of size $NF + NL + NC$, where NC is number of constraints (1-3).

All partial derivatives for all equations are computed, which forms the Jacobi matrix, and iteration preceeds by

$$\Delta\psi = \mathbf{J}^{-1}\mathbf{e} \tag{111}$$

where **e** is the current error matrix. In practise we can arrange the equations so to obtain a block tridiagonal Jacobian, so inversion takes $N_d \left(N_{NF} + N_{NL}\right)^3$ operations (using a Feautrier-like technique).

In the old literature problems where $NF \sim 100$, $NL \sim 10$ were treated, which meant that equation system of size ~ 1000 were inverted. However, in modern applications $NF \sim NL \sim 10^5$, and simultaneous solution of the whole equation system (e.g. complete linearization) is ruled out. **Hubeny & Mihalas : This may perhaps be done on large parallell systems..discuss**

In the general case its impossible to explicitly link J back into the RT equation; S depends on the solution of a set of non-linear equations involving J. But with ALI we may use Λ^* for simple relations between J and S that can be worked in analytically or by compact matrix generation code

Rutten describes Auer & Mihalas 1969 as first paper outlining complete linearization, i.e. coupled solutions of statistical equilibrium and RT equations. Here level population correction are replaced by sums over radiation field corrections, and these are solved for using the transfer equation (Feautrier scheme in Auer and Mihalas). No real explanations in Rutten.

Scharmer & Carlsson 1985, on the other hand, use an approximate transport equation when iterating the level populations. Note that the final answer is always the output of the 'classic' lambda iteration, which can be seen as doing the approximate step plus the correction step.

The problem is really fully specified by all level populations at all depths. Thus, if we can do ALI in the J domain, it must be equivalent to doing it in the n domain, where some degree of neighbour implicit coupling in each iteration, benefits convergence. This must be alternated with the 'exact' solution using the last guess for neighbors. Thus, one must do something clever/coupled, alternated with standard, local, (linear) equation system solutions. Both MULTI and CMFGEN do this (but the CMFGEN papers dont explicitly state this second correction, or lambda iteration, step).

11.2.4 Multi-D

A diagonal operator is broadly considered the only practical choice for multi-D cases (Hubeny 2003). Tridiagonal and higher methods extend with difficulty to multi-level problems. See also OAB86.

12 External accelerations

12.1 Ng acceleration

In Ng acceleration (Ng 1974), we aim to minimize the difference between a new estimate and the next iteration based on that estimate (at convergence there would be no difference). If we have a solution x^n for iteration n, let an extrapolated solution $\hat{x}^n$ be some linear combination of the last 3 solutions (one may also use more/less but this is the common choice)

$$\tilde{x}^n = c_0 x^n + c_1 x^{n-1} + c_2 x^{n-2} \tag{112}$$

under the constraint $c_0 + c_1 + c_1 = 1$. Now lets say we want to minimize $F[\tilde{x}^n] - \tilde{x}^n$ (i.e. the next correction to our extrapolation), or rather the norm $|F[\tilde{x}^n] - \tilde{x}^n|$. This is a least-squares problem (e.g. Hubeny and Mihalas), fitting two values c_1 and c_2 to minimize a sum of residuals. We are basically saying, there is already a lot of converged or quasi-converged information in our current estimate x^n, but scalings are off and we want to correct those using c_1 and c_2.

Now, if F is linear, $F[\tilde{x}^n] = c_0 F[x^n] + c_1 F[x^{n-1}] + c_2 F[x^{n-2}] = c_0 x^{n+1} + c_1 x^n + c_2 x^{n-1}$. Thus with $x^{n+1}, x^n, x^{n-1}, x^{n-2}$ known, its an algebraic sum involving c_0, c_1, c_2, easily calculable.

The solution is (Hubeny & Mihalas, see also OAB 1986)

$$c_1 = (\delta_{01}\delta_{22} - \delta_{02}\delta_{21})/(\delta_{11}\delta_{22} - \delta_{12}\delta_{21}) \tag{113}$$

$$c_2 = (\delta_{02}\delta_{11} - \delta_{01}\delta_{21})/(\delta_{11}\delta_{22} - \delta_{12}\delta_{21}) \tag{114}$$

where

$$\delta_{ij} = \left(\Delta x^n - \Delta x^{n-i}\right) \cdot \left(\Delta x^n - \Delta x^{n-j}\right) \tag{115}$$

and

$$\Delta x^n = x^n - x^{n-1} \tag{116}$$

Note that we need 4 consecutive solutions to make one acceleration. In practise one minimizes a residual that uses weights W_i that can be chosen by various arguments (e.g. OAB86 uses $W_i = 1/x_i$):

$$\Omega = \sum_i \left[\tilde{x}_i - F(\tilde{x}_i)\right] W_i \tag{117}$$

One may optimize over c coefficients locally (let them vary for each depth), or globally (one and the same for each depth).

Ng accelaration can be done earliest in the 4th iteration, although one normally waits somewhat longer as initial iterates are still far from convergence.

Ng acceleration is particularly useful in combination with ALI methods that are only 'moderately' fast in convergence on their own, such as local operators (OAB86).

12.2 ORTHOMIN

TBD..

12.3 Chebyshev acceleration

See Olson 1987. Hubeny & Mihalas : Not clear if can be used as a general RT method.

12.4 Broyden method

Instead of using Jacobian matrix, use Broyden matrix, which correspond to secant method in 1D. Each inverse Broyden matrix can be calculated from the previous inverse without new inversion.

12.5 Kantorovich method

This method is simply to hold the Jacobian constant once one approaches convergence ("keep tangent fixed"). Method has been shown to be surprisingly robust in RT context. One performs a few normal NR iterations and then swicthes to constant Jacobian. A refreshment may be done at some point.

12.6 Solution guiding

Any iterative scheme needs a good starting guess (even complete linearization), and sometimes active guidance also in later steps. For example, one may "switch on" more and more detailed physics in a a series of iterations such as

LTE-gray → LTE → NLTE-continua → NLTE-few lines → NLTE-many lines This may bring about a solution in fewer steps than doing the full NLTE-lines calculations from scratch.

12.7 Successive overrelaxation

E.g. Trujillo Bueno 1995.

12.8 Problem simplification

To reduce the number of frequencies, levels and depth points that are implicitly coupled, one may consider

- Implicit coupling only for some selected (judged important) frequencies and levels.
- Divide levels into *superlevels*, where LTE is assumed within each superlevel. A related idea is *level grouping*, where corrections within a group of lines is constant.
- Reduce N_D. Most algorithms are linear in N_D so no big gain. Note that AMR methods have not yet been used in RT codes, and this may be one of the most important remaining improvements (Hubeny & Mihalas).

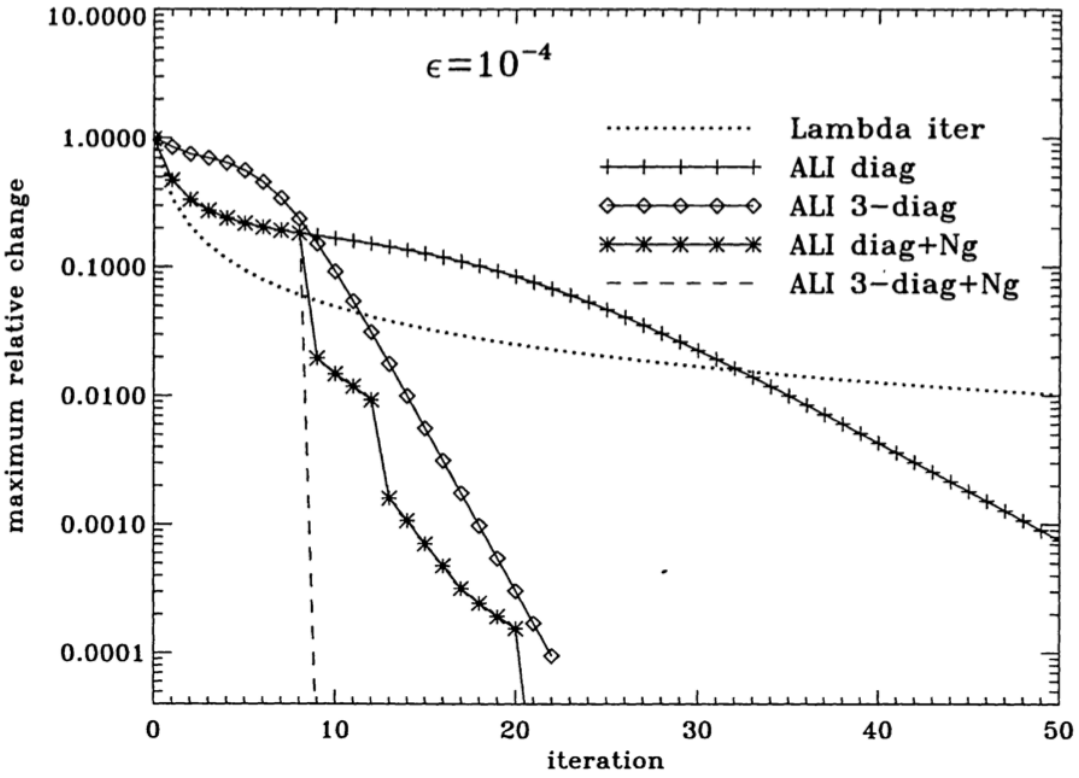


Figure 2: Comparison of iteration schemes with and without Ng acceleration, from Hubeny 2003.

A Feautrier method

Solution method. For each depth point d, create vectors $\bar{\mathbf{I}}_{d+1/2}$ than contain $\bar{I}$ (see Eq. 61) for each angle-frequency combination at that depth (so the $\bar{\mathbf{I}}_{d+1/2}$ vectors are of length $N_a N_f$). We have, at depth d, for angle-frequency point number i (Eq 83.17 in MM84)

$$\mu_i^2 \frac{1}{\Delta\tau_{d+1/2,i}} \left[\frac{1}{\Delta\tau_{d+1,i}} \bar{I}_{d+3/2,i} - \left(\frac{1}{\Delta\tau_{d,i}} + \frac{1}{\Delta\tau_{d+1,i}} \right) \bar{I}_{d+1/2,i} + \frac{1}{\Delta\tau_{d,i}} \bar{I}_{d-1/2,i} \right] \tag{118}$$

$$= \bar{I}_{d+1/2,i} - \frac{\eta^t_{d+1/2} + \sum_{i'} \sigma_{d+1/2,i,i'} \bar{I}_{d+1/2,i'}}{\chi_{d+1/2,i}} \tag{119}$$

Here quantities centred at cell mid-points have indeces of half (e.g. 3/2 for the first cell) and quantities centred at cell edges have integer indeces (e.g. 1 for the first cell). For each depth, there is spatial derivative coupling between $\bar{\mathbf{I}}_{d+1/2}$ and $\bar{\mathbf{I}}_{d-1/2}$ and $\bar{\mathbf{I}}_{d+3/2}$ for a given i. Then, we can create (diagonal) matrices at each depth, $\mathbf{A_{d+1/2}}$ and $\mathbf{C_{d+1/2}}$, containing the discretized derivative factors for $\bar{I}_{d-1/2,i}$ (third term on LHS) and $\bar{I}_{d+3/2,i}$ (first term on LHS), respectively:

$$A^{ii}_{d+1/2} = \mu_i^2 \frac{1}{\Delta\tau_{d+1/2,i}} \frac{1}{\Delta\tau_{d,i}} \tag{120}$$

$$C^{ii}_{d+1/2} = \mu_i^2 \frac{1}{\Delta\tau_{d+1/2,i}} \frac{1}{\Delta\tau_{d+1,i}} \tag{121}$$

Create also a (full) matrix $\mathbf{B_{d+1/2}}$, which on the diagonal has derivative operators for $\bar{\mathbf{I}}_{d+1/2}$ (second term on LHS), a -1 (first term on RHS), and a coherent scattering term (for no deflection coherent scattering $i = i'$ component of last summation term on RHS). The rest of the matrix is filled with terms from discretization of the scattering term where angle and/or frequency changes

(so the matrix is full).

$$B^{ii}_{d+1/2} = -\mu_i^2 \frac{1}{\Delta\tau_{d+1/2,i}} \left[\frac{1}{\Delta\tau_{d,i}} + \frac{1}{\Delta\tau_{d+1,i}} \right] - 1 + \frac{\sigma_{d+1/2,i,i}}{\chi_{d+1/2}}, \tag{122}$$

$$B^{ij}_{d+1/2} = \frac{\sigma_{d+1/2,i,j}}{\chi_{d+1/2}}. \tag{123}$$

On the RHS we obtain a vector $\mathbf{L}_{d+1/2}$ with the fixed thermal part of the source term (numerically known):

$$L^i_{d+1/2} = -S^t_{d+1/2,i} = -\frac{\eta^t_{d+1/2,i}}{\chi_{d+1/2,i}}. \tag{124}$$

Note that for the same frequency, each entry in $\mathbf{L}$ is the same for isotropy in χ. The transfer equations for each angle and frequency combination, for each depth d, is now a matrix system:

$$-\mathbf{A}_{d+1/2}\bar{\mathbf{I}}_{d-1/2} + \mathbf{B}_{d+1/2}\bar{\mathbf{I}}_{d+1/2} - \mathbf{C}_{d+1/2}\bar{\mathbf{I}}_{d+3/2} = \mathbf{L}_{d+1/2} \tag{125}$$

Lower boundary condition. Specify an outgoing intensity $I^+_\nu(\tau_{max}, \mu)$ from the lower boundary $d = D$. One may show (page 370 in Mihalas)

$$\mu \frac{\partial \bar{I}_\nu}{\partial \tau_\nu}_{taumax} = I^+_\nu(\tau_{max}, \mu) - \bar{I}_\nu(\tau_{max}, \mu). \tag{126}$$

Discretize:

$$\mu_i \frac{\bar{I}_{D-1/2,i} - \bar{I}_{D+1/2,i}}{\Delta\tau_{D+1/2,i}} = I^+_{D+1/2,i} - \bar{I}_{D+1/2,i}, \tag{127}$$

giving

$$A^{ii}_{D+1/2} = -\frac{\mu_i}{\Delta\tau_{D+1/2,i}} \tag{128}$$

$$C^{ii}_{D+1/2} = 0 \tag{129}$$

$$B^{ii}_{D+1/2} = \frac{\mu_i}{\Delta\tau_{D+1/2,i}} - 1 \tag{130}$$

$$B^{ij}_{D+1/2} = 0 \tag{131}$$

$$L^{ii}_{D+1/2} = I^+_{D+1/2,i} \tag{132}$$

Upper boundary condition. The upper (outer) boundary condition is normally $I^- = 0$. Then one can show (Eq. 83.20 in MM84):

$$\mu \frac{\partial \bar{I}_\nu(\mu)}{\partial \tau_\nu} = \bar{I}_\nu(\mu), \tag{133}$$

which discretized becomes (Eq. 83.36 in MM84)

$$\mu_i \frac{\bar{I}_{5/2,i} - \bar{I}_{3/2,i}}{\Delta\tau_{2,i}} = \bar{I}_{3/2,i} \left[1 + 1/2\Delta\tau_{3/2.i}/\mu_i\right]^{-1} + \left[\Delta\tau_{3/2,i}/\mu_i\right] \left[\bar{I}_{3/2,i} - S_{3/2,i}\right] \tag{134}$$

so (to be completed)

$$A^{ii}_{3/2} = 0 \tag{135}$$

$$C^{ii}_{3/2} = \frac{\mu_i}{\Delta\tau_{3/2,i}} \tag{136}$$

$$B^{ii}_{3/2} = \frac{\mu_i}{\Delta\tau_{2,i}} - 1 \tag{137}$$

$$L^{ii}_{3/2} = ... \tag{138}$$

Solution technique. For the first depth point ($d = 3/2$), we get from Eq. 125

$$\bar{\mathbf{I}}_{3/2} = \mathbf{B}_{3/2}^{-1}\mathbf{C}_{3/2}\bar{\mathbf{I}}_{5/2} + \mathbf{B}_{3/2}^{-1}\mathbf{L}_{3/2} \tag{139}$$

Define $\mathbf{D}_{3/2} = \mathbf{B}_{3/2}^{-1}\mathbf{C}_{3/2}$ and $\mathbf{X}_{3/2} = \mathbf{B}_{3/2}^{-1}\mathbf{L}_{3/2}$. By substituting Eq. 139 into Eq. 125 for $d = 5/2$, we get

$$\bar{\mathbf{I}}_{5/2} = \mathbf{D}_{5/2}\bar{\mathbf{I}}_{7/2} + \mathbf{X}_{5/2} \tag{140}$$

where $\mathbf{D}_{5/2}$ can be calculated from the known $\mathbf{A}_{5/2}, \mathbf{B}_{5/2}, \mathbf{C}_{5/2}$ and $\mathbf{D}_{3/2}$. The method is to perform a forward-backward sweep to calculate all the $\mathbf{D}$ matrices, starting at the upper boundary $d = 1$:

1. Compute all the $\mathbf{A_{d+1/2}}$, $\mathbf{B_{d+1/2}}$, $\mathbf{C_{d+1/2}}$ matrices at each depth d.
2. Compute $\mathbf{D}_{3/2} = \mathbf{B}_{3/2}^{-1}\mathbf{C}_{3/2}$.
3. Compute $\mathbf{X}_{3/2} = \mathbf{B}_{3/2}^{-1}\mathbf{L}_{3/2}$.
4. Compute next $\mathbf{D}_{d+1/2} = \left[\mathbf{B}_{d+1/2} - \mathbf{A}_{d+1/2}\mathbf{D}_{d-1/2}\right]^{-1}\mathbf{C}_{d+1/2}$.
5. Compute next $\mathbf{X}_{d+1/2} = \left[\mathbf{B}_{d+1/2} - \mathbf{A}_{d+1/2}\mathbf{D}_{d-1/2}\right]^{-1}\left[\mathbf{L}_{d+1/2} + \mathbf{A}_{d+1/2}\mathbf{X}_{d-1/2}\right]$.
6. When you arrive at last depth point D, solve $\bar{\mathbf{I}}_{d+1/2} = \mathbf{X}_{D+1/2}$.
7. Determine all other $\bar{\mathbf{I}}_{d+1/2} = \mathbf{D}_{d+1/2}\bar{\mathbf{I}}_{d+3/2} + \mathbf{X}_{d+1/2}$.

B Analytic solutions

B.1 The time-independent formal solution

Described in section 79 of Mihalas & Mihalas (1984). In plane-parallel geometry, the time-independent radiative transfer equation is

$$\mu\frac{\partial I_\nu}{\partial \tau_\nu} = I_\nu - S_\nu. \tag{141}$$

This can be rewritten as

$$\frac{\partial[I_\nu \exp(-\tau_\nu/\mu)]}{\partial \tau_\nu} = -\frac{S_\nu \exp(-\tau_\nu/\mu)}{\mu}. \tag{142}$$

If the source function S_ν is known, we can integrate Eq. (142):

$$I(\tau_1, \mu, \nu) = I(\tau_2, \mu, \nu)e^{-(\tau_2-\tau_1)/\mu} + \mu^{-1}\int_{\tau_1}^{\tau_2} S_\nu(t)e^{-(t-\tau_1)/\mu}\,\mathrm{d}t. \tag{143}$$

Eq. (143) is called a formal solution of the transfer equation. For the outgoing intensity ($\mu \geq 0$) in a semi-infinite medium $\tau_1 = \tau_\nu$ and $\tau_2 = \infty$. The formal solution is then given by

$$I(\tau_\nu, \mu, \nu) = \int_{\tau_\nu}^{\infty} S_\nu(t)e^{-(t-\tau_\nu)/\mu}\,\mathrm{d}t/\mu, \qquad (0 \leq \mu \leq 1) \tag{144}$$

If we assume that no radiation is entering through the outer boundary ($I^-(0) = 0$), we obtain

$$I(\tau_\nu, \mu, \nu) = \int_0^{\tau_\nu} S_\nu(t)e^{-(\tau_\nu-t)/(-\mu)}\,\mathrm{d}t/(-\mu), \qquad (-1 \leq \mu \leq 0). \tag{145}$$

for the incoming radiation field ($\mu \leq 0$).

Schwarzschild-Milne Equations

Using Eqs. (144) and (145) we can derive the following expression for the mean intensity

$$J_\nu(\tau_\nu) = \frac{1}{2}\left[\int_0^1 \mathrm{d}\mu \int_{\tau_\nu}^\infty S_\nu(t) e^{-(t-\tau_\nu)/\mu}\,\mathrm{d}t/\mu + \int_{-1}^0 \mathrm{d}\mu S_\nu(t) \int_0^{\tau_\nu} e^{-(\tau_\nu - t)/(-\mu)}\,\mathrm{d}t/(-\mu)\right]. \quad (146)$$

To simplify the problem, we substitute $w = \pm 1/\mu$ and change the order of integration:

$$J_\nu(\tau_\nu) = \frac{1}{2}\left[\int_{\tau_\nu}^\infty \mathrm{d}t S_\nu(t) \int_1^\infty \mathrm{d}w \frac{e^{-w(t-\tau_\nu)}}{w} + \int_0^{\tau_\nu} \mathrm{d}t S_\nu(t) \int_1^\infty \mathrm{d}w \frac{e^{-w(\tau_\nu - t)}}{w}\right]. \quad (147)$$

By identifying the integrals that appear in Eq. (147) as the first exponential integral, we arrive at a concise expression for the mean intensity

$$J_\nu(\tau_\nu) = \frac{1}{2}\int_0^\infty S_\nu(t_\nu) E_1|t_\nu - \tau_\nu|\,\mathrm{d}t_\nu. \quad (148)$$

The exponential integral is defined by

$$E_n(x) \equiv \int_1^\infty t^{-n} e^{-xt}\,\mathrm{d}t = x^{n-1}\int_x^\infty t^{-n} e^{-t}\,\mathrm{d}t. \quad (149)$$

Similarly, expressions for the next two moments can be derived:

$$F_\nu(\tau_\nu) = 2\pi\int_{\tau_\nu}^\infty S_\nu(t_\nu) E_2(t_\nu - \tau_\nu)\,\mathrm{d}t_\nu - 2\pi\int_0^{\tau_\nu} S_\nu E_2(\tau_\nu - t_\nu), \quad (150)$$

and

$$K_\nu(\tau_\nu) = \frac{1}{2}\int_0^\infty S_\nu(t_\nu) E_3|t_\nu - \tau_\nu|\,\mathrm{d}t_\nu. \quad (151)$$

Due to the central importance of Eq. (148) in radiative transfer theory an abbreviated operator notation has been introduced. Following this notation the mean intensity

$$J_\nu(\tau_\nu) = \Lambda_{\tau_\nu}[S(\tau_\nu)] \quad (152)$$

is obtained by applying the so-called **lambda operator**:

$$\Lambda_\tau[f(t)] \equiv \frac{1}{2}\int_0^\infty f(t) E_1|t - \tau|\,\mathrm{d}t \quad (153)$$

to the source function. Over time, however, the term 'lambda operator' has taken on a broader meaning, being used nowadays to describe any procedure (including non-analytic) to obtain J from S.

B.2 Wave Limit

In a vacuum ($\chi_\nu = \eta_\nu = 0$) the transfer equation (26) reduces to

$$\left[\frac{1}{c}\frac{\partial}{\partial t} + \frac{\partial}{\partial s}\right] I_\nu(\mathbf{x}, \mathbf{n}, t) = 0. \quad (154)$$

Introducing

$$I^+ \equiv I_\nu(\mathbf{x}, \mathbf{n}, t) \quad (155)$$

and

$$I^- \equiv I_\nu(\mathbf{x}, -\mathbf{n}, t) \quad (156)$$

we obtain the transfer equations

$$\frac{\partial I^+}{\partial t} + c\frac{\partial I^+}{\partial s} = 0 \quad (157)$$

and

$$\frac{\partial I^-}{\partial t} - c\frac{\partial I^-}{\partial s} = 0 \quad (158)$$

from Eq. (154). Defining the mean-intensity-like quantity

$$j \equiv \frac{1}{2}(I^+ + I^-) \tag{159}$$

and the flux-like quantity

$$h \equiv \frac{1}{2}(I^+ - I^-) \tag{160}$$

addition and subtraction of Eqs. (157) and (158) yields

$$\frac{\partial j}{\partial t} + c\frac{\partial h}{\partial s} = 0 \tag{161}$$

and

$$\frac{\partial h}{\partial t} + c\frac{\partial j}{\partial s} = 0. \tag{162}$$

The quantities h and j will appear again later in the lecture when we will derive the Feautrier Equations. Taking the partial derivative of Eq. (161) with respect to time and inserting Eq. (162) in the resulting equation yields

$$\frac{\partial^2 j}{\partial t^2} = c^2\frac{\partial^2 j}{\partial s^2}. \tag{163}$$

Similarly

$$\frac{\partial^2 h}{\partial t^2} = c^2\frac{\partial^2 h}{\partial s^2} \tag{164}$$

can be obtained. Eqs. (163) and (164) are wave equations with the solutions

$$j(s,t) = A_1 f_1(s - ct) + A_2 f_2(s + ct) \tag{165}$$

and

$$h(s,t) = B_1 f_1(s - ct) + B_2 f_2(s + ct). \tag{166}$$

Here A_1, A_2, B_1, B_2 are constants that are determined by the initial and boundary conditions. One possible solution of Eqs. (165) and (166) is a monochromatic, plane wave:

$$I(\mathbf{x}, t; \mathbf{n}', \nu') = I_0\delta(s - ct)\delta(\mathbf{n}' - \mathbf{n})\delta(\nu' - \nu) \tag{167}$$

In this case $J_\nu = H_\nu = K_\nu$ and the Eddington factor $f_\nu = 1$.

B.3 Diffusion limit

The discussion below assumes a static medium and that the radiation field and the medium are in thermal equilibrium. Radiative diffusion in moving media and in nonequilibrium are discussed in section 97 of Mihalas & Mihalas (1984).

Static, LTE atmosphere

Described in section 80 of Mihalas & Mihalas (1984). For $\tau_\nu \gg 1$ the source function S_ν (see Eq. (22)) approaches the Planck function B_ν. It is thus possible to write the source function at any optical depth t_ν as a Taylor expansion around some reference optical depth τ_ν:

$$S_\nu(t_\nu) = \sum_{n=0}^{\infty} \frac{\partial^n B_\nu}{\partial \tau_\nu^n}(t_\nu - \tau_\nu)^n/n! \tag{168}$$

Recall Eq. (144) for the intensity of outgoing radiation in a planar, static medium:

$$I(\tau_\nu, \mu, \nu) = \int_{\tau_\nu}^{\infty} S_\nu(\tau'_\nu) e^{-(\tau'_\nu - \tau_\nu)/\mu}\, \mathrm{d}\tau'_\nu/\mu, \qquad (0 \le \mu \le 1). \tag{169}$$

The assumption of plane-parallel geometry is justified due to the small photon mean free paths. Inserting the Taylor expansion of the source function into Eq. (169) yields

$$I(\tau_\nu, \nu) = B_\nu(\tau_\nu) + \mu\frac{\partial B_\nu(\tau_\nu)}{\partial \tau_\nu} + \mu^2\frac{\partial^2 B_\nu(\tau_\nu)}{\partial \tau_\nu^2} + \ldots \tag{170}$$

For $-1 \leq \mu \leq 0$ the intensity can be calculated in a similar fashion using Eq. (145). The result is identical to Eq. (170) apart from terms of the order of $\exp(-\tau/\mu)$, which vanish for $\tau \to \infty$. From the defining Eqs. (5), (8) and (15) we find

$$J_\nu = B_\nu(\tau_\nu) + \frac{1}{3}\frac{\partial^2 B_\nu(\tau_\nu)}{\partial \tau_\nu^2} + \dots \tag{171}$$

$$H_\nu = \frac{1}{3}\frac{\partial B_\nu(\tau_\nu)}{\partial \tau_\nu} + \frac{1}{5}\frac{\partial^3 B_\nu(\tau_\nu)}{\partial \tau_\nu^3} + \dots \tag{172}$$

$$K_\nu = \frac{1}{3}B_\nu(\tau_\nu) + \frac{1}{5}\frac{\partial^2 B_\nu(\tau_\nu)}{\partial \tau_\nu^2} + \dots \tag{173}$$

for the moments of the specific intensity. Replacing derivatives by difference quotients (i.e. $\partial B_\nu^n(\tau_\nu)/\partial \tau_\nu^n \to B_\nu/\tau_\nu^n$), we see that the ratio of successive terms is of $O(1/\tau_\nu^2) = O(\lambda_\nu^2/l^2)$. Here λ_ν denotes the photon mean free path and l is a characteristic structural length (e.g. a pressure scale height in a stellar envelope.) Since the ratio λ_ν/l is small (typical values in the sun range from 10^{-7} to 10^{-10}) it is sufficient to retain only the first terms in Eqs. (171) to (173):

$$J_\nu(\tau_\nu) = 3K_\nu(\tau_\nu) = B_\nu(\tau_\nu) \tag{174}$$

Thus both the mean intensity J_ν and the radiation pressure K_ν have their equilibrium values. In contrast to equilibrium the flux

$$H_\nu = \frac{1}{3}\frac{\partial B_\nu(\tau_\nu)}{\partial \tau_\nu} = -\frac{1}{3\chi_\nu}\frac{\partial B_\nu}{\partial T}\frac{\mathrm{d}T}{\mathrm{d}r} \tag{175}$$

is nonzero. By integrating this expression over frequency, we obtain the total flux

$$F = -(4\pi/3)\left(\int_0^\infty \frac{1}{\chi_\nu}\frac{\partial B_\nu(\tau_\nu)}{\partial T}\,\mathrm{d}\nu\right)\frac{\mathrm{d}T}{\mathrm{d}r}. \tag{176}$$

This equation is formally identical to Fourier's Law for heat conduction. Introducing the so called **Rosseland mean opacity**

$$\chi_\mathrm{R}^{-1} = \int_0^\infty \frac{1}{\chi_\nu}\frac{\partial B_\nu(\tau_\nu)}{\partial T}\,\mathrm{d}\nu \Big/ \int_0^\infty \frac{\partial B_\nu(\tau_\nu)}{\partial T} \tag{177}$$

we can define a **radiative conductivity**

$$K_R = \frac{4\pi}{3\chi_R}\frac{\mathrm{d}B}{\mathrm{d}T} = \frac{4}{3}c\lambda_R a_R T^3 \tag{178}$$

in analogy to the thermal conductivity.

In conclusion, at large optical depths the transfer problem can be described by the single equation Eq. (175), which behaves like a diffusion equation. The dimensionality of the problem has been reduced from six to one!

B.4 The Grey Atmosphere

For a grey material the opacity is independent of frequency i.e. $\chi_\nu = \chi$. As a consequence the radiation field becomes independent of the state of the material.

Applications

- Starting point in the calculation of more complex models. For stellar atmospheres it is typical to proceed through a series of intermediate models with increasing physical complexity e.g. LTE-gray → LTE → NLTE
- Test problem for numerical methods
- Provides boundary conditions for stellar structure calculations
- Neutron transport in heavy-water nuclear reactors

Basic Results

Frequency integration of the time independent, planar transfer equation yields:

$$\mu \frac{\partial I}{\partial \tau} = I - S \tag{179}$$

Here $I \equiv \int_0^\infty I_\nu \, \mathrm{d}\nu$ and $S \equiv \int_0^\infty S_\nu \, \mathrm{d}\nu$ denote the total intensity and source function respectively. For a grey material the radiative equilibrium condition

$$\int_0^\infty \chi_\nu J_\nu \, \mathrm{d}\nu = \int_0^\infty \chi_\nu S_\nu \, \mathrm{d}\nu \tag{180}$$

reduces to the simple requirement $J = S$. Using this result Eq. (179) can be simplified as follows:

$$\mu \frac{\partial I}{\partial \tau} = I - J \tag{181}$$

This is the transfer equation for a plane-parallel, grey atmosphere in radiative equilibrium. From the first moment of the transfer equation

$$\frac{\mathrm{d}H}{\mathrm{d}\tau} = J - S = 0 \tag{182}$$

we see that the flux is constant throughout the atmosphere. This is a general result for static, planar atmospheres in radiative equilibrium. The second moment equation is given by

$$\frac{\mathrm{d}K}{\mathrm{d}\tau} = H, \tag{183}$$

which has the solution

$$K(\tau) = H\tau + c = \frac{1}{4\pi} F\tau + c. \tag{184}$$

Recall that for large optical depths the specific intensity can be approximated as $I(\mu) = I_0 + I_1\mu$, which implies that $K(\tau) = 1/3 J(\tau)$. From this and Eq. (184) we infer that

$$J(\tau) \to \frac{3}{4\pi} F\tau \qquad (\tau \gg 1). \tag{185}$$

The general solution can then be written as

$$J(\tau) = \frac{3}{4\pi} F(\tau + q(\tau)), \tag{186}$$

where $q(\tau)$ denotes the as of yet undetermined Hopf function. It is possible to connect the constant c in Eq. (188) for the second moment of the radiation field to the newly introduced Hopf function by taking the limit of large optical depths:

$$\lim_{\tau\to\infty} \left[\frac{1}{3} J(\tau) - K(\tau)\right] = \frac{1}{4\pi} F \lim_{\tau\to\infty} \left[\tau + q(\tau) - \tau - c\right] = 0 \tag{187}$$

Thus $c = q(\infty)$ and the second moment is given by

$$K(\tau) = H\tau + c = \frac{1}{4\pi} F\tau + q(\infty). \tag{188}$$

LTE

If LTE is assumed (i.e. $S_\nu = B_\nu$) it is possible to assign a temperature T to the radiation field via the radiative equilibrium equation:

$$J(\tau) = S(\tau) = B[T(\tau)] = \sigma_\mathrm{R} T^4/\pi \tag{189}$$

Defining the effective temperature T_eff as the temperature a black body would have to reproduce the emergent flux (i.e $F = \sigma_\mathrm{R} T_\mathrm{eff}^4$), we can rewrite Eq. (186) in terms of T

$$T^4 = \frac{3}{4} T_\mathrm{eff}^4 [\tau + q(\tau)] \tag{190}$$

Mean Opacities

See section 3-2 in MH78. The goal is to define mean opacities in such a way that the general equations

$$\mu(\partial I_\nu/\partial z) = \chi_\nu(S_\nu - I_\nu) \tag{191}$$

$$\partial H_\nu/\partial z = \chi_\nu(S_\nu - J_\nu) \tag{192}$$

$$\partial K_\nu/\partial z = -\chi_\nu H_\nu \tag{193}$$

can be reduced to their grey counterparts

$$\mu(\partial I/\partial z) = \chi(S - I) \tag{194}$$

$$\mathrm{d}H/\,\mathrm{d}z = 0 \tag{195}$$

$$\mathrm{d}K/\,\mathrm{d}z = -\chi H. \tag{196}$$

It is impossible to achieve a complete correspondence between the grey and nongrey problem. However, suitable choices of the mean opacity can establish one to one correspondences for selected quantities.

Rosseland Means

If the goal is to reproduce the correct integrated energy flux H, the mean opacity must be defined as follows:

$$-\int_0^\infty \chi_\nu^{-1}(\partial K_\nu/\partial z)\,\mathrm{d}\nu = \int_0^\infty H_\nu\,\mathrm{d}\nu = H = -\bar{\chi}^{-1}(\mathrm{d}K/\,\mathrm{d}z). \tag{197}$$

$$\bar{\chi}^{-1} = \int_0^\infty \chi_\nu^{-1}(\partial K_\nu/\partial z)\,\mathrm{d}\nu \Big/ \int_0^\infty (\partial K_\nu/\partial z)\,\mathrm{d}\nu \tag{198}$$

Since K_ν is not known a priori, it is necessary to find approximations in order to evaluate the opacity. At high optical depths $K_\nu \to {}^1\!/_3\,J_\nu$, $J_\nu \to B_\nu$ and we can write $\partial K_\nu/\partial z = {}^1\!/_3\,(\partial B_\nu/\partial T)(\mathrm{d}T/\,\mathrm{d}z)$. With these simplifications Eq. (198) can be written as

$$\frac{1}{\bar{\chi}_\mathrm{R}} = \frac{\int_0^\infty \frac{1}{\chi_\nu}\frac{\partial B_\nu}{\partial T}\,\mathrm{d}\nu}{\int_0^\infty \frac{\partial B_\nu}{\partial T}\,\mathrm{d}\nu}. \tag{199}$$

The assumptions used to derive the Rosseland mean are the same as those used in the diffusion approximation. It is thus appropriate to use Rosseland means to describe radiative diffusion at high optical depths. This allows the determination of the thermal structure of the atmosphere at great depths via

$$T^4 = \frac{3}{4}T_\mathrm{eff}^4(\bar{\tau}_\mathrm{R} + q(\bar{\tau}_\mathrm{R})). \tag{200}$$

Flux-Weighted Mean

To transform the nongrey equation for the second moment K_ν [Eq. (193)] into the grey equation (196) the mean opacity must be defined as follows

$$\bar{\chi}_\mathrm{F} \equiv H^{-1}\int_0^\infty \chi_\nu H_\nu\,\mathrm{d}\nu. \tag{201}$$

This opacity is the flux-weighted mean of the frequency dependent opacity χ_ν. We can verify that the definition in Eq. (201) has the desired properties by integrating Eq. (193) over frequency:

$$-\,(\mathrm{d}K/\,\mathrm{d}z) = \int_0^\infty \chi_\nu H_\nu\,\mathrm{d}\nu = \bar{\chi}_\mathrm{F}H \tag{202}$$

Thus $K(\bar{\tau}) = H\bar{\tau} + c$ applies as in the grey case. This guarantees that the correct values for the radiation pressure and radiation force are recovered. This is of relevance for the calculation of the density structure of the atmospheres of early-type stars.

Planck and Absorption Means

Whereas the Rosseland mean is the appropriate average for optically thick systems the Planck mean is suitable for optically thin systems. For details see section 3-2 in MH78.

Eddington approximation

We know that at high optical depths $J = 3K$ holds. In the Eddington approximation this condition is applied throughout the entire atmosphere. In combination with $K = 1/(4\pi) \cdot F\tau + c$ this simplifying assumption leads to the following expression for the mean intensity

$$J_{\mathrm{E}} = \frac{3}{4\pi} F\tau + c'. \tag{203}$$

Recall the formal solution for the flux at optical depth τ_ν

$$F_\nu(\tau_\nu) = 2\pi \int_{\tau_\nu}^{\infty} S_\nu(t_\nu) E_2(t_\nu - \tau_\nu)\, \mathrm{d}t_\nu - 2\pi \int_0^{\tau_\nu} S_\nu E_2(\tau_\nu - t_\nu). \tag{204}$$

The flux at the outer boundary is then given by

$$F(0) = 2\pi \int_0^{\infty} \left(\frac{3}{4\pi} F\tau + c' \right) E_2(\tau)\, \mathrm{d}\tau = 2\pi c' E_3(0) + \frac{3}{4} F \left[\frac{4}{3} - 2E_4(0) \right]. \tag{205}$$

From $F(0) = F$ follows that

$$c' = \frac{3}{4\pi} \frac{E_4(0)}{E_3(0)} F. \tag{206}$$

Using the relation $E_n(0) = 1/(n-1)$ we find that $c' = F/(2\pi)$. This implies that the mean intensity is given by

$$J_{\mathrm{e}} = \frac{3}{4\pi} F \left(\tau + \frac{2}{3} \right). \tag{207}$$

Thus the Hopf function in the Eddington approximation is $q(\tau) = 2/3$. In LTE the temperature structure is determined by

$$T^4 = \frac{3}{4} T_{\mathrm{eff}}^4 \left(\tau + \frac{2}{3} \right). \tag{208}$$

Since $T = T_{\mathrm{eff}}$ for $\tau = 2/3$, this optical depth is commonly identified as the effective depth of continuum formation. Despite the simplyfing Eddington approximation Eq. (208) provides a quite accurate description of the thermal structure of a grey atmosphere. We expect the greatest departures from the analytic solution to occur close to the boundary. However, the ratio of the boundary temperature T_0 to the effective temperature in the Eddington approximation $T_0/T_{\mathrm{eff}} = 0.841$ still agrees fairly well with the analytic solution $T_0/T_{\mathrm{eff}} = 0.8114$.

Solution with discrete ordinates

For an in-depth discussion see section 3-2 of MH78. Both approximate and exact solution can be obtained by replacing the integrals in the transfer equation

$$\mu[\partial I(\tau,\mu)/\partial\mu] = I(\tau,\mu) - \frac{1}{2} \int_{-1}^{1} I(\tau,\mu)\, \mathrm{d}\mu \tag{209}$$

with quadrature sums i.e.

$$\frac{1}{2} \int_{-1}^{1} I(\tau,\mu)\, \mathrm{d}\mu \approx \frac{1}{2} \sum_{j=-n}^{n} a_j I_j(\mu_j) \tag{210}$$

This reduces the integro-differential equation to a system of $2n$ coupled ordinary differential equations. In the limit of $n \to \infty$ this approximation becomes exact and allows the derivation of the analytic solution.

Spherical Geometry

A discussion of grey spherical atmospheres in LTE and radiative equilibrium can be found in section 7-6 of Mihalas (1978).

References

Auer, L. 1976, J. Quant. Spec. Radiat. Transf., 16, 931

Auer, L. 1991, in NATO Advanced Study Institute (ASI) Series C, Vol. 341, Stellar Atmospheres - Beyond Classical Models, ed. L. Crivellari, I. Hubeny, & D. G. Hummer, 9

Cannon, C. J. 1973, ApJ, 185, 621

Castor, J. I. 2004, Radiation Hydrodynamics

Clough, S. A., Shephard, M. W., Mlawer, E. J., et al. 2005, J. Quant. Spec. Radiat. Transf., 91, 233

Feautrier, P. 1964, Comptes Rendus Academie des Sciences (serie non specifiee), 258, 3189

Hillier, D. J. & Dessart, L. 2012, MNRAS, 424, 252

Hubeny, I. & Mihalas, D. 2014, Theory of Stellar Atmospheres

Iwabuchi, H. 2006, Journal of the Atmospheric Sciences, 63, 2324

Mihalas, D. 1970, Stellar atmospheres

Mihalas, D. 1978, Stellar Atmospheres, 2nd edn. (San Francisco: W. H. Freeman and Co)

Mihalas, D., Auer, L. H., & Mihalas, B. R. 1978, ApJ, 220, 1001

Mihalas, D., Kunasz, P. B., & Hummer, D. G. 1976, ApJ, 210, 419

Mihalas, D. & Mihalas, B. W. 1984, Foundations of Radiation Hydrodynamics (New York: Oxford University Press)

Noerdlinger, P. D. & Rybicki, G. B. 1974, ApJ, 193, 651

Olson, G. L., Auer, L. H., & Buchler, J. R. 1986, J. Quant. Spec. Radiat. Transf., 35, 431

Olson, G. L. & Kunasz, P. B. 1987, J. Quant. Spec. Radiat. Transf., 38, 325

Romano, P. K., Horelik, N. E., Herman, B. R., et al. 2015, Annals of Nuclear Energy, 82, 90, joint International Conference on Supercomputing in Nuclear Applications and Monte Carlo 2013, SNA + MC 2013. Pluri- and Trans-disciplinarity, Towards New Modeling and Numerical Simulation Paradigms

Rutten, R. J. 2003, Radiative Transfer in Stellar Atmospheres, 8th edn. (Utrecht University lecture notes)

Rybicki, G. B. & Hummer, D. G. 1991, A&A, 245, 171

Scharmer, G. B. 1981, ApJ, 249, 720

Scharmer, G. B. 1983, A&A, 117, 83

Thomas, L. H. 1930, The Quarterly Journal of Mathematics, os-1, 239