Astron. Astrophys. 245, 171–181 (1991)

ASTRONOMY AND ASTROPHYSICS

An accelerated lambda iteration method for multilevel radiative transfer

I. Non-overlapping lines with background continuum

G.B. Rybicki[1] and D.G. Hummer[2,★]

[1] Harvard-Smithsonian Center for Astrophysics, 60 Garden St., Cambridge, MA 02138, USA
[2] Joint Institute for Laboratory Astrophysics, University of Colorado and National Institute of Standards and Technology, Boulder, CO 80309, USA

Received August 17, accepted October 4, 1990

Abstract. A method is presented for solving multilevel transfer problems with non-overlapping lines and with background continuum (but no active continuum transfer). This method is based on the use of an approximate lambda operator, which is either the diagonal or a finite band of the "true" numerical lambda operator. Linear, "preconditioned" equations of statistical equilibrium are derived, the coefficients of which are found efficiently using a new fast method for finding the diagonal elements (or a band) of the "true" numerical lambda operator. The preconditioned equations are used iteratively with the formal solution of the transfer equation, so that the entire iteration scheme involves solving only linear equations based on one previous iteration. Applications of the method are made to several multilevel problems, including a model problem of Avrett & Loeser (1987) and an eleven-level neutral helium atom. Convergence properties of the method are systematically investigated for these problems, using diagonal and tridiagonal approximate lambda operators, with and without the acceleration technique of Ng (1974), and the method has proved to be fast and reliable for line transfer problems of this type. An appendix gives an improved version of the Feautrier scheme that has much better numerical conditioning for small optical depths.

Key words: radiative transfer – non-LTE – stellar atmospheres

1. Introduction

In recent years, a very powerful group of closely related techniques have emerged for the solution of radiative transfer and statistical equilibrium problems; this general approach is known collectively as Approximate (or Accelerated) Lambda Iteration (ALI). The monochromatic lambda operator $\Lambda_{\mu\nu}$ along a ray with direction μ is defined by

$$I_{\mu\nu} = \Lambda_{\mu\nu}[S_{\mu\nu}], \tag{1.1}$$

where $I_{\mu\nu}$ is the mean intensity and $S_{\mu\nu}$ is the source function, which is expressed in term of level populations. The common thread among ALI methods is the use of approximate lambda operators, which can be inverted with relative ease, and then corrected iteratively.

As is well known, the process of iterating between the atomic level populations and the radiation field (known as *lambda iteration*) in optically thick systems converges extremely slowly. Each cycle of the iteration corresponds to photons moving about one mean free path in the medium. In the cores of strong lines photons have very short mean free paths, and this implies that many iterations are required to move them any substantial distance. In a multilevel system, photons can also become "trapped" in cyclic groups of transitions. With a properly chosen approximate lambda operator, photons are propagated on a scale appopriate to the medium as a whole, in a sense to be made clear below.

These difficulties were overcome in Rybicki (1972, 1984) with a modified form of the transfer problem in which only those photons in the line wing were explicitly treated, while those in the core were regarded as passive and were eliminated by systematically preconditioning the statistical equilibrium equations. An important feature of this preconditioning is that all of the equations remain linear. This so-called "core saturation" method converged very much more rapidly than normal lambda iteration. As originally derived in Rybicki (1972) the method was intrinsically a multilevel one, and an analytic example of a simple multilevel case was treated in that paper. The first extensive numerical applications of core saturation to multilevel systems were those of Flannery et al. (1979, 1980), who treated UV pumping in N II and Si II.

These two ideas: 1) solving the radiative transfer equation approximately with a simplified operator, and 2) solving some form of the modified statistical equilibrium equations at every step of the iteration cycle, form the basis of all ALI methods. Cannon (1973) later introduced a more formal approach for solving a two-level system with reduced effort, by use of the identity $\Lambda_{\mu\nu} = \lambda^*_{\mu\nu} + (\Lambda_{\mu\nu} - \Lambda^*_{\mu\nu})$, where $\Lambda_{\mu\nu}$ and $\Lambda^*_{\mu\nu}$ are the exact and approximate operators, the latter with a simple inverse. This is used as the basis of an iteration technique by writing

$$I_{\mu\nu} = \Lambda^*_{\mu\nu}[S_{\mu\nu}] + (\Lambda_{\mu\nu} - \Lambda^*_{\mu\nu})[S^\dagger_{\mu\nu}], \tag{1.2}$$

where $S^\dagger_{\mu\nu}$ is the source function from the previous iteration. This is inserted into the equations of statistical equilibrium and solved

Send offprint requests to: G.B. Rybicki

* Staff member, Quantum Physics Division, National Institute of Standards and Technology

for the updated source function $S_{\mu\nu}$. Although this equation is only approximate at each stage of the iteration, it is clear that it becomes exact for the converged solution, where $S_{\mu\nu} = S^{\dagger}_{\mu\nu}$. Thus no error is introduced into the solution by using Eq. (1.2), so long as convergence is actually achieved. Roots of the iteration method (1.2) go back at least to Jacobi, and it is well known in numerical analysis as "operator splitting" (for a discussion of its history, see Varga 1962). Because it was introduced into the field of radiative transfer by Cannon (1973), it is most commonly known as *Cannon's method.*

As Cannon's original work was confined to two-level systems, it provides useful, but limited guidance to the construction of multilevel methods. Scharmer (1981, 1984) expanded on the ideas of Rybicki & Cannon and developed a well implemented multilevel code (Scharmer & Carlsson 1985) which has found wide application. Scharmer's work appears to have triggered wide interest among stellar astrophysicists which in turn has lead to further important developments. The central issue in these developments was the choice of the approximate operator. Although Rybicki's original choice was simply a diagonal or tridiagonal operator, Scharmer introduced two additional, more complicated, nonlocal operators that yielded a higher rate of convergence. However the application of these operators to multilevel systems led to considerable complexity and contained an arbitrary parameter (or two, in some applications by other workers) that had to be determined by trial and error for optimal convergence.

Werner & Husfeld (1985), in an important paper which has led to several important generalizations and applications, investigated diagonal and upper-diagonal operators that simplified drastically the amount of effort needed to linearize the resultant system. A crucial step towards the definition of an optimum operator was then taken by Olson et al. (1986, hereafter OAB), who showed that the diagonal part of the true lambda operator provided a very efficient approximate operator with no arbitrary parameter. Subsequently, Olson & Kunasz (1987) showed that the tridiagonal and pentadiagonal parts of the lambda operator converged even more rapidly.

Although we confine ourselves to planar models, we should mention that spherical atmospheres with rapid gas flow have been treated by Hamann (1985, 1986, 1987), using a parameterized diagonal operator, and by Hempe & Schoenberg (1986) and Schoenberg & Hempe (1986), who introduced a parameter-free diagonal operator. Puls & Herrero (1988) have adapted the OAB operator for spherical geometry, and Hillier (1990) has generalized the Hempe-Schoenberg operator to tridiagonal and pentadiagonal forms with significant improvements in convergence.

Although the convergence of ALI methods with well-chosen approximate operators is much more rapid than in the original lambda iteration procedure, it is possible to obtain a significant additional increase in the convergence rate by the use of purely numerical techniques. Buchler & Auer (1985) introduced into ALI methods an acceleration technique developed by Ng (1974); a convenient description of this method is given by OAB. More recently Klein *et al.* (1989) found an acceleration procedure due to Vinsome (1976) to be even more efficient. Auer (1987) reviews acceleration procedures for ALI methods.

Our work differs from that described above in several important respects. In most ALI methods so far developed, some form of linearization of the statistical equilibrium equation must be employed at every step in the iteration process; an exception is Pauldrach & Herrero (1988), who, however, achieve linearity at the cost of having to include *two* previous iterations in the iteration procedure. We avoid this computionally expensive calculation by "preconditioning" the statistical equilibrium equations, which automatically maintains their linearity. This preconditioning analytically eliminates the effect of photons trapped in the core of the optically-thick lines and consequently improves the convergence of the method. In addition, we can construct the diagonal part of the "true" numerical lambda operator in the order of N_d operations, where N_d is the number of depth points, and still maintain consistency with the method used to solve the approximate transfer equation. Finally, we introduce the psi operator, which is analogous to the lambda operator, but operates on the emissivity rather than on the source function; the utility of this operator is discussed below. In the present paper, we present the analysis and sample results only for non-overlapping lines with a background continuum, that is, a continuum with specified opacity and source function, that does not change with iteration. The analysis when continua are included is significantly more complicated, because the preconditioning becomes complicated when transitions overlap. These problems will be treated in subsequent papers.

The formulation of our method is given in Sect 2, and a description of the computer code, together with results for an eleven level helium atom, as well as for a simple three level problem defined by Avrett & Loeser (1987), are contained in Sect 3. Future work is outlined briefly in Sect 4. Two appendices present important technical and computational developments. Appendix A describes an improvement to Feautrier's method that makes it much more numerically robust, especially when very small optical depth steps are encountered. Appendix B presents a fast solution for the diagonal elements (or wider bands) of the inverse of a tridiagonal matrix, which is useful for constructing approximate lambda operators.

The analysis and computer code presented here is very useful for application to molecular band problems, and has already been used to investigate the formation of the 4.3 μm band of CO_2 in planetary atmospheres by Kutepov et al. (1991).

2. Multilevel formulation

The multilevel line problem is equivalent to the joint solution of the equations of statistical equilibrium for the populations of the levels and the equations of radiative transfer for the radiation field. As discussed in Sect. 1, in this paper we treat the simple case of the bound-bound transitions within a single ion in a plane-parallel medium. The temperature and electron density are assumed to be prescribed as a function of depth; we do not consider temperature corrections or shifts in ionization balance. The continuum opacity and emissivity at line frequencies are also prescribed as a function of depth. We include velocity fields using an observer's frame formulation; this allows for the treatment of velocities not too much larger than thermal.

In plane parallel geometry the transfer equation for the specific intensity $I_{\mu\nu}$ is

$$\mu \frac{\partial I_{\mu\nu}}{\partial z} = -\chi_{\mu\nu} I_{\mu\nu} + \eta_{\mu\nu}, \tag{2.1}$$

where $\chi_{\mu\nu}$ and $\eta_{\mu\nu}$ are the total opacity and emissivity at frequency ν and angle μ.

The ion is considered to have a number of levels, denoted by the indices $l, l', \ldots$. In addition to the population n_l, each level is characterized by its statistical weight g_l and its energy E_l. We abbreviate the energy relations $E_l < E_{l'}$ and $E_l > E_{l'} > E_{l'}$ by $l \prec l'$ and $l \succ l'$, respectively. The line radiative properties of the ion are completely characterized by the emissivity $\eta_{ll'}$, the opacity $\chi_{ll'}$, between each pair of levels for which a line transition is to be considered. These quantities can depend on μ when velocity fields are present. For $l \succ l'$, they are given by

$$\eta_{ll'}(\mu, \nu) = \frac{h\nu}{4\pi} n_l A_{ll'} \varphi_{ll'}(\mu, \nu),$$
$$\chi_{ll'}(\mu, \nu) = \frac{h\nu}{4\pi} (n_{l'} B_{l'l} - n_l B_{ll'}) \varphi_{ll'}(\mu, \nu), \tag{2.2}$$

where $A_{ll'}$, $B_{ll'}$, and $B_{l'l}$ are the Einstein coefficients and $\varphi_{ll'}(\mu, \nu)$ is the normalized line profile function. If the outward velocity in the medium is $v(z)$, then

$$\varphi_{ll'}(\mu, \nu) = \tilde{\varphi}_{ll'}(\nu - \nu_{ll'} - \nu_{ll'} \mu v(z)/c), \tag{2.3}$$

where $\nu_{ll'}$ is the line center frequency and $\tilde{\varphi}_{ll'}$ is the normalized profile function in the comoving frame of the material.

The line source function for the transition ll', $l \succ l'$, is defined by

$$S_{ll'} = \frac{n_l A_{ll'}}{n_{l'} B_{l'l} - n_l B_{ll'}} \tag{2.4}$$

For convenience we define the line source function to be symmetric in the indices l and l', so that $S_{l'l} = S_{ll'}$. The source function $S_{ll'}$ is frequency-independent as a result of assuming the equality of emission and absorption profiles, (the assumption of *complete redistribution*) and is independent of μ by assuming isotropic scattering.

The total emissivity and opacity are given in terms of the above line quantities by

$$\eta_{\mu\nu} = \sum_{l \succ l'} \eta_{ll'}(\mu, \nu) + \eta_c(\nu),$$
$$\chi_{\mu\nu} = \sum_{l \succ l'} \chi_{ll'}(\mu, \nu) + \chi_c(\nu), \tag{2.5}$$

where $\eta_c(\nu)$ and $\chi_c(\nu)$ are the *background* emissivity and opacity, which are prescribed and do not change during the course of solution (typically velocity fields do not affect these quantities, so they are independent of μ). The total source function is then given by

$$S_{\mu\nu} = \eta_{\mu\nu}/\chi_{\mu\nu}. \tag{2.6}$$

The equations of statistical equilibrium for the ion populations may be written,

$$n_l \sum_{l'} (R_{ll'} + C_{ll'}) = \sum_{l'} n_{l'} (R_{l'l} + C_{l'l}), \tag{2.7}$$

where $C_{ll'}$ are the collisional rate coefficients and $R_{ll'}$ are the radiative rate coefficients, given by

$$R_{ll'} = A_{ll'} + B_{ll'} \bar{J}_{ll'}, \quad l \succ l',$$
$$= B_{ll'} \bar{J}_{ll'}, \quad l \prec l'. \tag{2.8}$$

Also, $\bar{J}_{ll'}$ is the integrated mean intensity, defined by

$$\bar{J}_{ll'} = \frac{1}{4\pi} \int d\Omega \int d\nu \, \varphi_{ll'}(\mu, \nu) I_{\mu\nu} \tag{2.9}$$

Thus, we may write Eq. (2.7) in the convenient form,

$$\sum_{l' \prec l} \left[n_l A_{ll'} - (n_{l'} B_{l'l} - n_l B_{ll'}) \bar{J}_{ll'} \right]$$
$$- \sum_{l' \succ l} \left[n_{l'} A_{l'l} - (n_l B_{ll'} - n_{l'} B_{l'l}) \bar{J}_{ll'} \right]$$
$$+ \sum_{l'} (n_l C_{ll'} - n_{l'} C_{l'l}) = 0. \tag{2.10}$$

The ALI scheme to be introduced here is based on the operator splitting technique given in Eq. (1.2), which gives $I_{\mu\nu}$ in terms of the populations. It is important to note that $S^\dagger_{\mu\nu}$ as well as the operators $\Lambda_{\mu\nu}$ and $\Lambda^*_{\mu\nu}$ are constructed from the "old" populations $n^\dagger_l$. (In general, a dagger denotes quantities evaluated using "old" variables, from the previous iteration.) The "new" populations n_l enter only through the source function $S_{\mu\nu}$, which is expressed in terms of them by way of Eqs. (2.6), (2.5), and (2.2). If this form for $I_{\mu\nu}$ is substituted into Eq. (2.9) and the resulting form for $\bar{J}_{ll'}$ is substituted into Eq. (2.10), then the statistical equilibrium equations are expressed solely in terms of known quantities and the new populations n_l.

The iterative method now proceeds as follows: An initial choice for the old populations $n^\dagger_l$ is made. This allows one to set up the equations of statistical equilibrium (2.10). These equations are then solved for the new populations n_l. Regarding these to be the old populations another cycle of iteration can be made, and this process is continued until convergence is obtained.

Equation (1.2) describes a whole class of methods, each particular method being specified by its choice of approximate lambda operator $\Lambda^*_{\mu\nu}$. We now discuss the choices to be considered here. First of all, we note that in practice the so-called exact lambda operator $\Lambda_{\mu\nu}$ is itself an approximation based on some discretization of the problem in space, and it appears as a matrix operator acting on the values of the source function at the chosen discrete spatial grid. One of the simplest choices for an approximate operator is to take the *diagonal* part of this full matrix operator; this is the choice made by Olson et al. (1986) in their treatment of the two-level problem. As a better approximation one might take the approximate operator to be the *tridiagonal part* of the full matrix operator, or an even wider band approximation.

The principal advantage of the diagonal approximation is that the equations of statistical equilibrium remain completely *local*, whereas more sophisticated band approximations, such as tridiagonal, introduce nonlocalness into these equations, which are then harder to solve, and may be more unstable. On the other hand, for certain types of problems a tridiagonal approximation may be desirable, such as those involving radiative equilibrium, where a diffusion-like approximation is appropriate at large depths. Our development here is general enough to emcompass all of these possibilities.

We now derive linear preconditioned equations of statistical equilibrium for a multilevel line problem with background continuum. In doing this we are motivated to preserve the linearity of the original equations of statistical equilibrium, for reasons of simplicity. We shall find that there are two primary impediments to achieving such linear equations: First, the quantity $(n_l B_{ll'} - n_{l'} B_{l'l})$ multiplies the mean intensity in Eq. (2.10), and second, the source function $S_{\mu\nu}$ in Eq. (1.2) is inversely proportional to the opacity, which is linear in the populations. Because of these difficulties, it is worthwhile to illustrate the general procedure for two simple cases using a local approximate operator,

the first with no background continuum and the second with background continuum. Each of these simple cases adds some insight into the method. After this we treat the most general case of nonlocal operator with background continuum.

2.1. Local operator with no background continuum

Since the lines are assumed non-overlapping, in the neighborhood of line ll' we have $S_{\mu\nu} = S_{ll'}$, which is frequency-independent. Using this in Eq. (1.2) we obtain, for frequencies near the line,

$$I_{\mu\nu} = \Lambda^*_{\mu\nu} S_{ll'} + I^{\rm eff}_{\mu\nu}, \tag{2.11}$$

where

$$I^{\rm eff}_{\mu\nu} = \Lambda_{\mu\nu}[S^\dagger_{ll'}] - \Lambda^*_{\mu\nu} S^\dagger_{ll'} = I^\dagger_{\mu\nu} - \Lambda^*_{\mu\nu} S^\dagger_{ll'}. \tag{2.12}$$

The quantity $I^\dagger_{\mu\nu}$ is the radiation field that one gets from the formal solution with the old populations. Since the approximate lambda operator here is local (i.e., diagonal), it acts like an ordinary multiplication in its operation on the source function. Thus, brackets [...] are omitted on $\Lambda^*_{\mu\nu}$.

Substituting (2.11) into (2.9), we have,

$$\bar{J}_{ll'} = \bar{\Lambda}^*_{ll'} S_{ll'} + \bar{J}^{\rm eff}_{ll'}, \tag{2.13}$$

where,

$$\bar{\Lambda}^*_{ll'} = \int d\Omega \int d\nu\, \varphi_{ll'} \Lambda^*_{\mu\nu}, \tag{2.14}$$

and

$$\bar{J}^{\rm eff}_{ll'} = \int d\Omega \int d\nu\, \varphi_{ll'} I^{\rm eff}_{\mu\nu} = \bar{J}^\dagger_{ll'} - \bar{\Lambda}^*_{ll'} S^\dagger_{ll'}, \tag{2.15}$$

are angle and frequency averages of $\Lambda_{\mu\nu}$ and $I^{\rm eff}_{\mu\nu}$ using $\varphi_{ll'}$ as a weighting function. The quantity

$$\bar{J}^\dagger_{ll'} = \int d\Omega \int d\nu\, \varphi_{ll'} I^\dagger_{\mu\nu}, \tag{2.16}$$

is the value of the integrated mean intensity obtained by integrating over the "old" radiation field.

Substituting these results in Eq. (2.10), with the use of Eq. (2.4), we obtain the preconditioned equations of statistical equilibrium,

$$\begin{aligned}&\sum_{l'<l} [n_l A_{ll'}(1 - \bar{\Lambda}^*_{ll'}) - (n_{l'} B_{l'l} - n_l B_{ll'}) \bar{J}^{\rm eff}_{ll'}] \\ &\quad - \sum_{l'>l} [n_{l'} A_{l'l}(1 - \bar{\Lambda}^*_{ll'}) - (n_l B_{ll'} - n_{l'} B_{l'l}) \bar{J}^{\rm eff}_{ll'}] \\ &\quad + \sum_{l'} (n_l C_{ll'} - n_{l'} C_{l'l}) = 0. \end{aligned} \tag{2.17}$$

We see that the result of these substitutions is to leave the form of statistical equilibrium equations the same as before, except that the Einstein A-coefficient has been multiplied by the factor $(1 - \bar{\Lambda}^*_{ll'})$, and the integrated mean intensity $\bar{J}_{ll'}$ is now replaced by $\bar{J}^{\rm eff}_{ll'}$. These are the same types of alterations that appear in the core saturation method [Rybicki 1984; Eq. (5.14)], but where now $\bar{\Lambda}^*_{ll'}$ replaces the core normalization N_c. The conditioning of these equations is now improved, because much of the transfer in the "core" of the line (described by the local part of the lambda operator) has cancelled out analytically. These preconditioned statistical equilibrium equations are clearly still linear in the ion populations. Remarkably, for this simple case, the two impediments to achieving linearity discussed above have cancelled out, so that no additional assumptions or approximations have to be made.

Another desirable feature of these modified equations is that they automatically guarantee non-negative solutions for the new populations. This property follows from the non-negativity of the modified rate coefficients. The non-negativity of $\bar{J}^{\rm eff}_{ll'}$ follows directly from its definition (2.15) as an average over $I^{\rm eff}_{\mu\nu}$, which itself is non-negative, since it is defined in Eq. (2.12) as $(\Lambda_{\mu\nu} - \Lambda^*_{\mu\nu})$ (a non-negative operator) acting on the old source function (a non-negative quantity). The non-negativity of $1 - \bar{\Lambda}^*_{ll'}$ can be demonstrated as a special case of the preceding, namely, when the source function is everywhere constant and equal to unity, $S^\dagger_{ll'} \equiv 1$. Then the corresponding $\bar{J}^{\rm eff}_{ll'} \geq 0$, and Eq. (2.13) implies $\bar{\Lambda}^*_{ll'} \leq \bar{J}^\dagger_{ll'}$. But also $\bar{J}^\dagger_{ll'} \leq 1$, since lambda operators act as averaging operators. Thus $1 - \bar{\Lambda}^*_{ll'} \geq 0$. Note that it is important for the preceding argument that the approximate operator be the diagonal of the "exact" $\Lambda_{\mu\nu}$; otherwise one cannot be certain that the operator $(\Lambda_{\mu\nu} - \Lambda^*_{\mu\nu})$ will be entirely non-negative.

Having completed this simple case, we now turn our attention to the case of multilevel transfer with a background continuum, but still using a *local* approximate lambda operator.

2.2. Local operator with background continuum

For non-overlapping lines, the total emissity and opacity in the neighborhood of the the line ll' are given by $\eta_{\mu\nu} = \eta_{ll'} + \eta_{cll'}$ and $\chi_{\mu\nu} = \chi_{ll'} + \chi_{cll'}$, where $\eta_{cll'}$ and $\chi_{cll'}$ are the values of η_c and χ_c evaluated at the line frequency $\nu_{ll'}$. Then from (2.6),

$$S_{\mu\nu} = r_{ll'} S_{ll'} + (1 - r_{ll'}) S_{cll'}, \tag{2.18}$$

in the neighborhood of the ll' transition, where,

$$r_{ll'} = \frac{\chi_{ll'}(\mu, \nu)}{\chi_{ll'}(\mu, \nu) + \chi_{cll'}}, \tag{2.19}$$

It can already be seen that the remarkable cancellation of factors that led to linear equations in the preceding section cannot occur here, because the denominator is not simply proportional to $(n_l B_{ll'} - n_{l'} B_{l'l})$. One way to overcome this difficulty is to write, instead of Eq. (2.18),

$$S_{\mu\nu} = r^\dagger_{ll'} S_{ll'} + (1 - r^\dagger_{ll'}) S_{cll'}, \tag{2.20}$$

where we choose to evaluate $r^\dagger_{ll'}$ using "old" quantities from the previous iteration, that is,

$$r^\dagger_{ll'} = \frac{\chi^\dagger_{ll'}(\mu, \nu)}{\chi^\dagger_{ll'}(\mu, \nu) + \chi_{cll'}}. \tag{2.21}$$

This alteration does not affect the converged solution, since then $r^\dagger_{ll'} = r_{ll'}$. Note that a dagger does not need to appear on $\chi_{cll'}$, since the background quantities do not change with iteration.

Using Eq. (2.20) in Eq. (1.2), we express $I_{\mu\nu}$ as

$$I_{\mu\nu} = \Lambda^*_{\mu\nu} r^\dagger_{ll'} S_{ll'} + \check{I}^{\rm eff}_{\mu\nu}, \tag{2.22}$$

where now,

$$\check{I}^{\rm eff}_{\mu\nu} = \Lambda_{\mu\nu}[S^\dagger_{\mu\nu}] - \Lambda^*_{\mu\nu} r^\dagger_{ll'} S^\dagger_{ll'} = I^\dagger_{\mu\nu} - \Lambda^*_{\mu\nu} r^\dagger_{ll'} S^\dagger_{ll'}. \tag{2.23}$$

Note the terms containing $\Lambda^*_{ll'} S_{cll'}$ have cancelled in Eqs. (2.22) and (2.23), because background quantities do not change with iteration. Of course, the continuum source function still affects the solution through its appearance in $S^\dagger_{\mu\nu}$.

Substituting (2.22) into (2.9), we have,

$$\bar{J}_{ll'} = \check{\Lambda}^*_{ll'} S_{ll'} + \check{J}^{\rm eff}_{ll'}, \tag{2.24}$$

where,

$$\check{\Lambda}^*_{ll'} = \int d\Omega \int d\nu\, \varphi_{ll'} \Lambda^*_{\mu\nu} r^\dagger_{ll'}, \tag{2.25}$$

and

$$\check{J}^{\rm eff}_{ll'} = \int d\Omega \int d\nu\, \varphi_{ll'} I^{\rm eff}_{\mu\nu} = \bar{J}^\dagger_{ll'} - \check{\Lambda}^*_{ll'} S^\dagger_{ll'}. \tag{2.26}$$

$\bar{J}^\dagger_{ll'}$ is given, as before, by Eq. (2.15). The quantity $\check{\Lambda}^*_{ll'}$ differs from $\bar{\Lambda}^*_{ll'}$, defined in Eq. (2.14), by the additional factor $r^\dagger_{ll'}$ in the integrations over angle and frequency. It is still a local operator, and thus also acts as a scalar when applied to the source function $S_{ll'}$.

Substitution of these results into the equations of statistical equilibrium (2.10) yields,

$$\sum_{l'<l}\left[n_l A_{ll'}(1-\check{\Lambda}^*_{ll'}) - (n_{l'}B_{l'l} - n_l B_{ll'})\check{J}^{\rm eff}_{ll'}\right] - \sum_{l'>l}\left[n_{l'}A_{l'l}(1-\check{\Lambda}^*_{ll'}) - (n_l B_{ll'} - n_{l'}B_{l'l})\check{J}^{\rm eff}_{ll'}\right] + \sum_{l'}(n_l C_{ll'} - n_{l'}C_{l'l}) = 0. \tag{2.27}$$

Being linear in the populations, Eq. (2.27) again leads to a simple lambda iteration scheme. The arguments of the last section can be easily extended to show that the new populations are guaranteed to be non-negative, since the extra $r^\dagger_{ll'}$ factors satisfy $0 \le r^\dagger_{ll'} \le 1$.

One can see that the replacement of $r_{ll'}$ by $r^\dagger_{ll'}$ played a very important role in achieving the linearity of Eq. (2.27); otherwise the populations would appear in the denominator of $r_{ll'}$ through the dependence of $\chi_{ll'}$ on n_l and $n_{l'}$. In the next section we make a somewhat different choice to achieve linearity.

2.3. *Nonlocal operator with background continuum*

Nonlocal approximate lambda operators do not in general act like scalars, and one must be careful not to commute depth-dependent quantities with them. In this regard, we note that the commuting of the factor $(n_l B_{ll'} - n_{l'}B_{l'l})$ through $\bar{\Lambda}^*_{ll'}$ and $\check{\Lambda}^*_{ll'}$, where it could cancel another factor, played a large role in achieving the linearity of Eqs. (2.17) and (2.27). This simplification will no longer be available.

For the purposes of this section it is convenient to write the total source function as

$$S_{\mu\nu} = r^*_{ll'} S_{ll'} + (1 - r^*_{ll'}) S_{{\rm c}ll'}, \tag{2.28}$$

where,

$$r^*_{ll'} = \frac{\chi_{ll'}(\mu,\nu)}{\chi^\dagger_{ll'}(\mu,\nu) + \chi_{{\rm c}ll'}}. \tag{2.29}$$

This differs from Eq. (2.20) in that the numerator is evaluated with the new opacity $\chi_{ll'}$ rather than the old one. The denominator is still evaluated with the old opacity. Again, this does not affect the converged solution, where $r^*_{ll'} = r_{ll'}$. Using Eq. (2.29) in Eq. (2.9), we obtain,

$$I_{\mu\nu} = \Lambda^*_{\mu\nu}[r^*_{ll'} S_{ll'}] + \tilde{I}^{\rm eff}_{\mu\nu}, \tag{2.30}$$

where now,

$$\tilde{I}^{\rm eff}_{\mu\nu} = \Lambda_{\mu\nu}[S^\dagger_{\mu\nu}] - \Lambda^*_{\mu\nu}[r^*_{ll'} S^\dagger_{ll'}] = I^\dagger_{\mu\nu} - \Lambda^*_{\mu\nu}[r^*_{ll'} S^\dagger_{ll'}]. \tag{2.31}$$

Substituting (2.30) into (2.9), we have,

$$\bar{J}_{ll'} = \tilde{\Lambda}^*_{ll'}[S_{ll'}] + \tilde{J}^{\rm eff}_{ll'}, \tag{2.32}$$

where,

$$\tilde{\Lambda}^*_{ll'}[\cdots] = \int d\Omega \int d\nu\, \varphi_{ll'} \Lambda^*_{\mu\nu}[r^*_{ll'}\cdots], \tag{2.33}$$

and

$$\tilde{J}^{\rm eff}_{ll'} = \int d\Omega \int d\nu\, \varphi_{ll'} \tilde{I}^{\rm eff}_{\mu\nu} = \bar{J}^\dagger_{ll'} - \tilde{\Lambda}^*_{ll'}[S^\dagger_{ll'}]. \tag{2.34}$$

Note in Eqs. (2.30)–(2.34) that brackets are used with $\Lambda^*_{ll'}$ and $\tilde{\Lambda}^*_{ll'}$ in order to emphasize their nonlocal nature.

Substitution of these results into the equations of statistical equilibrium (2.10) yields,

$$\sum_{l'<l}\left[n_l A_{ll'} - (n_{l'}B_{l'l} - n_l B_{ll'})\tilde{\Lambda}^*_{ll'}[S_{ll'}] - (n_{l'}B_{l'l} - n_l B_{ll'})\tilde{J}^{\rm eff}_{ll'}\right] - \sum_{l'>l}\left[n_{l'}A_{l'l} - (n_l B_{ll'} - n_{l'}B_{l'l})\tilde{\Lambda}^*_{ll'}[S_{ll'}] - (n_l B_{ll'} - n_{l'}B_{l'l})\tilde{J}^{\rm eff}_{ll'}\right] + \sum_{l'}(n_l C_{ll'} - n_{l'}C_{l'l}) = 0. \tag{2.35}$$

We note that

$$\tilde{\Lambda}^*_{ll'}[S_{ll'}] = A_{ll'}\frac{h\nu}{4\pi}\int d\Omega \int d\nu\, \varphi_{ll'}\Lambda^*_{\mu\nu}\left[\frac{\varphi_{ll'}}{\chi^\dagger_{ll'} + \chi_{{\rm c}ll'}} n_l\right], \tag{2.36}$$

is linear in the new populations n_l. However, because of the factor $(n_l B_{ll'} - n_{l'}B_{l'l})$ in front of this term in Eq. (2.35), the overall linearity of the statistical equilibrium equations is spoiled. In order to get linearity here, we choose to evaluate this factor using the old populations, that is, we write,

$$\sum_{l'<l}\left[n_l A_{ll'} - (n^\dagger_{l'}B_{l'l} - n^\dagger_l B_{ll'})\tilde{\Lambda}^*_{ll'}[S_{ll'}] - (n_{l'}B_{l'l} - n_l B_{ll'})\tilde{J}^{\rm eff}_{ll'}\right] - \sum_{l'>l}\left[n_{l'}A_{l'l} - (n^\dagger_l B_{ll'} - n^\dagger_{l'}B_{l'l})\tilde{\Lambda}^*_{ll'}[S_{ll'}] - (n_l B_{ll'} - n_{l'}B_{l'l})\tilde{J}^{\rm eff}_{ll'}\right] + \sum_{l'}(n_l C_{ll'} - n_{l'}C_{l'l}) = 0. \tag{2.37}$$

This equation may be put into a more convenient form, if we define the operator $\hat{\Lambda}^*_{ll'}$ by the relation

$$(n^\dagger_{l'}B_{l'l} - n^\dagger_l B_{ll'})\tilde{\Lambda}^*_{ll'}[S_{ll'}] = A_{ll'}\hat{\Lambda}^*_{ll'}[n_l] \tag{2.38}$$

It follows that

$$\hat{\Lambda}^*_{ll'}[\cdots] = \kappa^\dagger_{ll'}\int d\Omega \int d\nu\, \varphi_{ll'}\Lambda^*_{\mu\nu}[r^\dagger_{ll'}(\kappa^\dagger_{ll'})^{-1}\cdots], \tag{2.39}$$

where

$$\kappa^\dagger_{ll'} = \frac{h\nu}{4\pi}(n^\dagger_{l'}B_{l'l} - n^\dagger_l B_{ll'}), \tag{2.40}$$

is the integrated line opacity in the ll' transition, evaluated with the old populations.

Equation (2.37) can then be put into the form,

$$\sum_{l'<l}\left[A_{ll'}(\mathbf{1} - \hat{\Lambda}^*_{ll'})[n_l] - (n_{l'}B_{l'l} - n_l B_{ll'})\tilde{J}^{\rm eff}_{ll'}\right] - \sum_{l'>l}\left[A_{l'l}(\mathbf{1} - \hat{\Lambda}^*_{l'l})[n_{l'}] - (n_l B_{ll'} - n_{l'}B_{l'l})\tilde{J}^{\rm eff}_{ll'}\right] + \sum_{l'}(n_l C_{ll'} - n_{l'}C_{l'l}) = 0, \tag{2.41}$$

where $\mathbf{1}$ is the identity operator, for which $\mathbf{1}n_l = n_l$.

Equation (2.41) is the fundamental equation of the present paper. It is clearly a linear equation in the populations and can be used as a nonlocal ALI method for multilevel line problems with background continuum. It was designed to be used with nonlocal operators, but of course can be used with local operators as well. However, it turns out to be completely equivalent to the method given in Sect. 2.2 in the local case, since the $\kappa^{\dagger}_{ll'}$ factors cancel owing to commutivity in Eq. (2.39), which then becomes identical to Eq. (2.25).

When used as a nonlocal tridiagonal method, Eq. (2.41) appears as a block tridiagonal matrix equation for the populations at all depths. This can be solved by standard Gaussian elimination, analogously to the Feautrier method.

As we shall show, the convergence of the nonlocal method is usually better than that of the local method. However, the nonlocal method may fail where the local one would not; for example, it is no longer true for the nonlocal method that the populations are always guaranteed to be non-negative. In difficult cases it may be necessary to use a local method, or perhaps to start the iterations with a local method, switching to the nonlocal method to obtain final convergence.

At several crucial stages in deriving Eq. (2.41), we chose to evaluate certain quantities using old populations rather than new ones. Admittedly, our particular choices are not unique, and there are a number of other ways that could have been used to achieve linear equations. To a large extent we have been guided by the desire to obtain equations most like the simple ones of Sect. 2.1. In treating more complex cases (such as continua arising from the same ion and overlapping lines), there arise even more possibilities for seemingly arbitrary choices of this type.

However, as we shall show in subsequent papers, there are fairly straightforward ways of limiting and organizing these choices. For instance, the choices can be made much more natural by introducing the *psi operator*, which is defined by

$$I_{\mu\nu} = \Psi_{\mu\nu}[\eta_{\mu\nu}], \tag{2.39}$$

that is, it gives the formal solution for the intensity as an operator acting on the emissivity rather than the source function. Because the operators $\Psi_{\mu\nu}$ and $\Lambda_{\mu\nu}$ are each constructed from old variables, Eq. (2.39) is *not* exactly equivalent to Eq. (1.1), except in the convergence limit. Using this operator instead of the lambda operator, the equivalent of the representation (2.28), with its particular definition of $r^{*}_{ll'}$, would have been made automatically.

In subsequent papers we shall make extensive use of the psi operator. It will become apparent then that this operator, rather than the lambda operator, is the more natural one to use for deriving analogous simple iterative methods for multilevel problems. Basically this is because the emissivities functions are linear in the populations while the source functions are not. We shall show that refinements necessary to treat true continua and overlapping lines can be done with straightforward extensions of the concepts already developed in this paper.

It is perhaps well to point out here that the preconditioning schemes used in this paper not only leads to *linear* equations of statistical equilibrium, but it leads to *homogeneous* equations as well. This is consistent with our aim to obtain modified equations that are as close as possible to the original ones. It also insures the positivity of the populations, at least for diagonal approximate operators. However, there may be compelling reasons to introduce preconditioned equations that are nonhomogeneous in the populations; this will be discussed in subsequent papers.

3. Computational features and results

We have incorporated the analysis given above into a Fortran program called MALI-L (to distinguish it from MALI-C, which includes bound-free transitions). This code allows for an arbitrary atomic model, specified in terms of energy levels and statistical weights. Radiative transition probabilities are specified in terms of Einstein A-coefficients, which are input, while collisional rate coefficients are specified as functions of temperature by subroutines. After initialization, the code works entirely in terms of transitions, which are classed as "active" or "passive", depending on whether or not the transfer problem is solved. Passive transitions include forbidden transitions, or allowed ones for which the transfer problem is regarded as unimportant; the radiative transitions are included in the equations of statistical equilibrium. Transitions can also be set in detailed balance. Each line is specified by the choice of a Doppler or Voigt profile function, and by a background opacity. An arbitrary velocity field can be specified, but, since the code uses observer's frame frequencies, it is only useful for velocities that are at most a few times the mean thermal speed.

The iterations can be initialized in a number of ways: 1) by asssuming LTE populations; 2) by using static escape probabilities expressed in terms of the K_2-function; 3) by assuming a "coronal" approximation of optical thinness in all lines; and 4) by reading in an initialization file. In difficult cases one can move to the case of interest by treating a succession of cases (with decreasing electron density, for example), using each result as the initialization of the next one.

The depth discretization is defined in terms of a column mass density grid m_d, $d = 1, 2, \ldots, N_D$. For the preliminary investigations reported here, we used a semi-logarithmic column mass density grid, specified by the total column mass density M and the number of depth points per decade p (in the logarithmic portion of the grid, at large depths). One can choose between two free boundaries, on one hand, and one free and one reflecting boundary, for symmetrical atmospheres, on the other. In the first case,

$$m_d = \frac{M \sinh ad}{2 \sinh (aN_D/2) \cosh a(d - N_D/2)}, \tag{3.1}$$

where $a = \ln(10)/2p$. In the second case, we have

$$m_d = \frac{M \sinh ad}{2 \sinh (aN_D) \cosh a(d - N_D)}, \tag{3.2}$$

where now M is the total mass column density to the reflecting surface at the midpoint. These forms are too restrictive for practical cases, but are ideal for investigations of the dependence of the convergence rate on the grid spacing, of the sort done by OAB.

Frequency and angle grids with corresponding quadrature weights are defined. For each angle-frequency point the monochromatic optical depths along the ray can be calculated to yield an optical depth grid. In this way the formal solution of the transfer problem is reduced to the solution of the Feautrier equations as given in Appendix A, which is significantly more robust than the original form when dealing with very small optical depths. In matrix form these equations take the form

$$Tu = S, \tag{3.3}$$

where T is a tridiagonal matrix constructed from the coefficients A_d, B_d, and C_d, S is the vector of source functions, and u is the solution vector of Feautrier variables. The solution can in principle be written as

$$u = T^{-1}S, \tag{3.4}$$

where T^{-1} is the matrix inverse of T. For static problems this implies that

$$T^{-1} = (\Lambda_{\mu\nu} + \Lambda_{-\mu\nu})/2. \tag{3.5}$$

(With appropriate modifications, this also applies to cases with velocity fields.) Thus the inverse of T is simply related to the lambda operator. In the local approximation, the approximate lambda operator $\Lambda^*_{\mu\nu}$ (actually the average of this operator for μ and $-\mu$) is simply the diagonal part of the inverse of T. In the nonlocal case, a band around the diagonal of the desired width is used. We emphasize that the full inverse T^{-1} never needs to be constructed; the vector u is found by the method of Appendix A and the diagonal elements (or a band) of T^{-1} is found quickly by the method of Appendix B.

Starting with a set of old populations the iterations proceed as follows: For each angle-frequency point the formal solution of the transfer problem is solved, along with the diagonal elements (or a tridiagonal band) of T^{-1}. These quantities are accumulated, with the appropriate weights, to eventually yield $\bar{J}^{\dagger}_{ll'}$ for each transition and the elements of the operator $\bar{\Lambda}^*_{ll'}$, and thus also the operator $\hat{\Lambda}^*_{ll'}$. From these, the quantity $\bar{J}^{\rm eff}_{ll'}$ can also be constructed. We thus obtain the coefficients of the preconditioned equations of statistical equilibrium (2.41), which are either completely local, or may themselves be tridiagonal in depth. These equations are then solved for the new populations, completing the iteration cycle.

The convergence of MALI-L is monitored by forming the relative changes in all populations at all depths in the last iteration. From these is found the relative change of maximum absolute size; this is called C_k, where k is the iteration number. When $|C_k|$ becomes less than a certain prescribed value, the method is deemed to have converged. For all the cases treated in this paper, we have also calculated the "true" relative error of maximum absolute size E_k at the k-th iteration, based on the "exact" solution (found by a preliminary long run with the fastest method).

Convergence is accelerated by application of Ng's (1974) procedure, applied to all the level populations at all depths. Two control parameters are specified, the iteration number at which the acceleration method is first applied, and the number of previous iterations that are used. The behavior of Ng's method and the optimal values of the control parameters will be discussed below.

We have also included a re-gridding procedure, so that an approximate solution can be obtained with a coarse grid, which leads to rapid convergence, and then all level populations are interpolated onto a finer grid.

3.1. Comparison with a three-level model calculation

Avrett & Loeser (1987) presented results for a simplified three-level hydrogen atom with lines but no continuum. The semi-infinite medium was assumed to be isothermal at a temperature of 5000 K. They treated two cases: 1) with constant collisional rates; and 2) with collisional rates that fall rapidly in a narrow layer near the surface. These results are well suited for comparisons with other codes, since their model was specified completely, including atomic data.

We solved both Avrett-Loeser cases 1 and 2 using MALI-L. The initial populations were chosen to be those appropriate to LTE. Four cases were run, using a diagonal (D) operator or the tridiagonal (T) operator, and using no acceleration (N) or Ng acceleration (A); the cases are thus denoted D/N, D/A, T/N, and T/A.

Considerable experimentation with Ng's method showed that if it was turned on too soon, it converged slowly or not at all. It appears to be important to wait until the changes in the successive iterations become small enough to be roughly represented as linear perturbations ($|C_k| \lesssim 1$). For the present case, good results were obtained by turning the acceleration on at $k = 20$ using 7 previous iterations for the diagonal (D) case, and at $k = 10$ using 5 previous iterations for the tridiagonal (T) case.

The detailed convergence properties of the runs for Avrett-Loeser case 1 are given in Fig. 1. The solid curves are for $|C_k|$ and the dotted curves are for $|E_k|$. For the non-accelerated cases, it can be seen that at large k the two curves become straight lines of the same slope on this semi-log plot, with $|E_k|$ larger than $|C_k|$ by a certain constant factor. With no acceleration, a true accuracy of about 1% is obtained after about 80 iterations with the diagonal method (D/N) and about 25 iterations with the tridiagonal method (T/N).

When Ng's acceleration method is used, the convergence is greatly improved. Now a true accuracy of 0.01% is obtained after about 25 iterations with the diagonal method (D/A) and about 12 iterations with the tridiagonal method (T/A). Furthermore, after the acceleration is begun, the quantity $|C_k|$, which can easily be computed during the course of iteration, becomes a fairly good

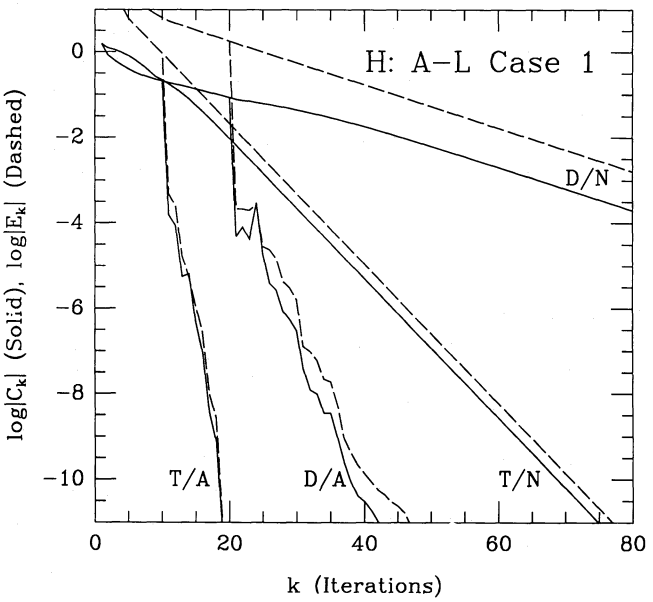

Fig. 1. Convergence plots for Avrett-Loeser case 1. $|C_k|$ is the maximum absolute relative change (solid curve), and $|E_k|$ is the maximum absolute relative true error (dashed curve), in the populations at iteration k. The labels on the curves refer to different runs, using either a diagonal (D) or tridiagonal (T) approximate operator, and using either no acceleration (N) or Ng's acceleration method (A)

estimator of the true error $|E_k|$. In any case, the rate of convergence is so high after acceleration that one can afford to set a more stringent convergence criterion than one really needs.

These convergence results were found to be relatively insensitive to the order of the frequency and angular quadrature. However, as expected from the discussion of OAB, the rate of convergence did depend on p, the number of depth points per decade. In order to reproduce the detailed results of Avrett-Loeser for the source functions as functions of optical depth to about 1%, it was necessary to use about 20 frequency points and 3 angle points. It was also necessary to use about 96 depth points ($\sim$6 points per decade) with the usual second-order difference scheme for the Feautrier method; with the fourth-order difference method of Auer (1976) only 70 depth points ($\sim$4 points per decade) were required.

We emphasize that a number of iterations found here (even for runs D/N) compares very favorably with those quoted by Avrett & Loeser ($\sim$10), since the iterations here are very much faster and simpler. Furthermore, the scaling for large problems is much more favorable. The only full matix operations required in each iteration for the diagonal method are those to solve the equations of statistical equilibrium for each depth, so the timing scales as $\sim N_D N_L^3$. The tridiagonal method requires the solution of a block tridiagonal system for the populations, but the timing for this also scales as $\sim N_D N_L^3$ (but with larger coefficient).

We also solved the Avrett-Loeser case 2 using exactly the same method. We found that the convergence properties of this case were virtually identical to those of case 1. Despite the rapid change in collision rates near the surface, we encountered no difficulties of any kind, unlike the experiences reported by Avrett & Loeser with their method.

We also solved a model six-level hydrogen problem, including levels through $n = 3$. The medium was taken as semi-infinite and isothermal at $T_e = 10^4$ K. Collisional rates were computed by fits to theoretical calculations as described in Hummer & Storey (1987). For densities larger than about 10^{10} cm^{-3}, the convergence results were very similar to the Avrett-Loeser cases. For lower electron densities difficulties arose due to population inversions between some of the upper states. We did not pursue this model further, because of its physical inadequacies: a lack of any couplings to a continuum and neglect of line overlapping within multiplets.

3.2. *An eleven-level neutral helium model*

We also applied MALI-L to a model He I atom with eleven levels (singlets and triplets through $n = 3$). Two cases were treated: 1) a semi-infinite medium with constant $N_e = 10^{14}$ cm^{-3}; and 2) a finite medium (mass column density of 3 10^{-6} g-cm^{-2}) with constant $N_e = 10^{10}$ cm^{-3}. Both cases were isothermal at $T_e = 2$ 10^4 K. Radiative data was taken from Wiese et al. (1966) and collisional rates were fitted to Chebyshev polynomials by Hummer & Storey (unpublished) from the cross sections calculated by Berrington & Kingston (1987). Each of these cases was solved, as before, with four runs using the diagonal or tridiagonal method, and using acceleration or no acceleration. Here the Ng acceleration was begun at $k = 25$ for the diagonal method and at $k = 15$ for the tridiagonal method.

Case 1 was initialized assuming LTE. The convergence of the methods is shown in Fig. 2, where the notation is the same as for Fig. 1. The numbers of iterations now are somewhat larger than for the Avrett-Loeser case, but the overall behavior is remarkably similar.

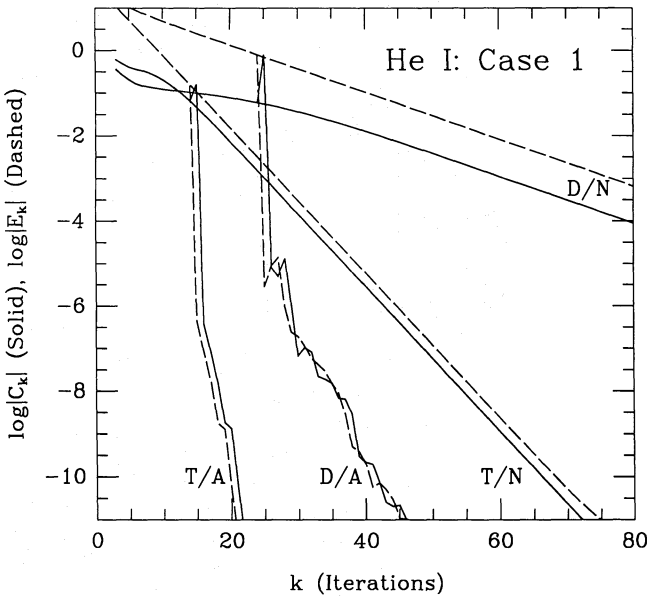

Fig. 2. Convergence plots for the eleven-level neutral helium model case 1, semi-infinite medium with $N_e = 10^{14}$ cm^{-3}

Case 2 could not be solved with the LTE initialization, since population inversions appeared almost immediately. However, using the escape probability initialization, the iterations did converge, as shown in Fig. 3. Here the numbers of iterations are considerably larger, and the pattern of the error curves is noticeably different than the case in Fig. 2; note especially the break in the $|C_k|$ curve at about $k = 50$. Clearly the more severe non-LTE character of the problem at lower electron densities causes the method to converge more slowly. Even so, convergence to a maximum relative error of 10^{-4} was achieved in only 17 iterations with the tridiagonal method and Ng's acceleration method.

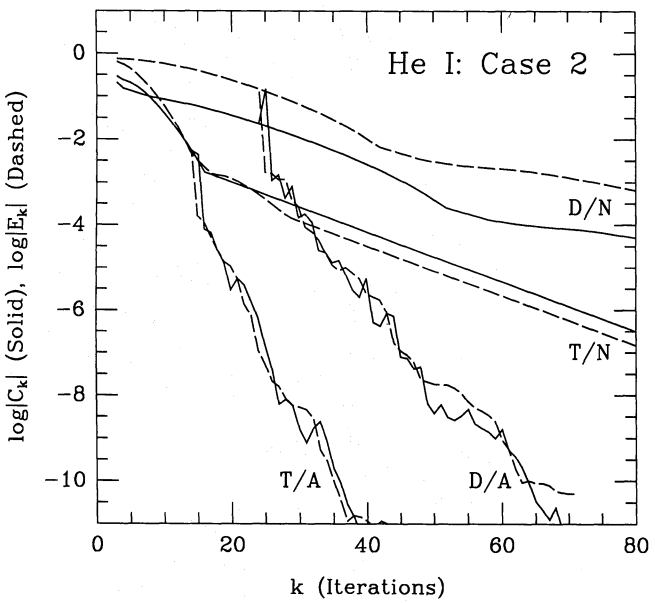

Fig. 3. Convergence plots for the eleven-level neutral helium model case 2, finite medium with $N_e = 10^{10}$ cm^{-3}

As a further test of MALI-L, we also applied it to a model nineteen-level neutral helium atom (levels through $n = 4$), with model parameters the same as the above helium case 1. Collisional rates here were augmented by unpublished data for transitions involving $n = 4$. The convergence properties of this nineteen-level case was virtually identical to the eleven-level case 1. This behavior, plus the analogous behavior seen in the six-level hydrogen model, suggests that the convergence of complicated, multilevel problems is probably no worse (nor better) than simple models with just a moderate number of the lower levels.

4. Final remarks

The methods of this paper have been shown to provide a fast and reliable method for solving multilevel transfer problems involving non-overlapping lines with background continuum. In further papers of this series, we shall show that these methods can be generalized to treat more complications, including active continua and overlapping lines. The preconditioned equations derived in this paper form very useful prototypes to guide the derivations of similar equations in problems with greater physical complexity.

Acknowledgements. We thank John Hillier and Wolf-Ranier Hamann for helpful comments on the manuscript and Peter Storey for his help with the collisional rates. This work and the research programs that have motivated it are supported in part by the National Science Foundation through Grant AST88-02937 and in part by the National Aeronautics and Space Administration through Grant NAGW-766. One of us (DGH) gratefully acknowledges support from the Smithsonian Institution Visitors Program that helped facilitate this work.

Appendix A: improved Feautrier solution

In static, plane parallel media, the transfer equation can easily be put into the Feautrier form

$$\frac{d^2u}{d\tau^2} = u - S, \tag{A1}$$

where S is the source function, $u = (I_{\mu\nu} + I_{-\mu\nu})/2$, and τ is the monochromatic optical depth along a ray at angle μ (Feautrier 1964). This can be generalized to spherical and moving media as well (see, e.g. Mihalas 1978). Straightforward discretization of this equation for an optical depth grid τ_d, $d = 1, \ldots, N_D$, leads to the tridiagonal system

$$-A_d u_{d-1} + B_d u_d - C_d u_{d+1} = S_d. \tag{A2}$$

The convention for signs in this equation reflects the fact that A_d, B_d, C_d, and S_d are typically all positive. For the simplest, second-order difference scheme, the quantities A_d, B_d, and C_d are given for $2 \le d \le N_D - 1$ by

$$A_d = \frac{2}{\Delta\tau_{d-1}(\Delta\tau_{d-1} + \Delta\tau_d)}, \quad B_d = 1 + \frac{2}{\Delta\tau_d \Delta\tau_{d-1}},$$
$$C_d = \frac{2}{\Delta\tau_d(\Delta\tau_{d-1} + \Delta\tau_d)}, \tag{A3}$$

where $\Delta\tau_d = \tau_{d+1} - \tau_d$. The coefficients for $d = 1$ and $d = N_D$ are special and depend on the boundary conditions chosen; these can also be made second-order accurate using the method of Auer (1967).

The linear system (A2) is solved by Gaussian elimination. Auxilliary variables D_d and Z_d are introduced by the recurrence relations

$$D_d = (B_d - A_d D_{d-1})^{-1} C_d, \qquad D_1 = B_1^{-1} C_1,$$
$$Z_d = (B_d - A_d D_{d-1})^{-1}(S_d + A_d Z_{d-1}), \qquad Z_1 = B_1^{-1} S_1, \tag{A4}$$

after which one finds,

$$u_d = D_d u_{d+1} + Z_d, \quad u_{N_D+1} = 0. \tag{A5}$$

The Feautrier method has proved to be an extremely useful technique for solving transfer problems. However, some difficulties with the method occur when dealing with small optical depth increments. We present a method that overcomes this difficulty in two ways: first, a particular cancellation between the coefficients is taken into account; and second, a new set of variables is introduced in the elimination scheme that lead to better numerical conditioning.

Small optical depth increments $\Delta\tau$ cause problems in the B_d coefficient when the first term (unity) becomes lost to machine precision when compared to the second term, which is order $(\Delta\tau)^{-2}$. This difficulty can be alleviated by analytically eliminating B_d in favor of the linear combination of coefficients.

$$H_d \equiv -A_d + B_d - C_d, \tag{A6}$$

with special cases $H_1 = B_1 - C_1$ and $H_{N_D} = -A_{N_D} + B_{N_D}$. For the particular set of coefficients given in (A3), we have $H_d \equiv 1$ for $2 \le d \le N_D - 1$, but we shall keep the notation H_d for generality.

Problems also occur in the elimination scheme, since the values of D_d can become very close to unity, and the difference can be lost in the numerical precision. Our solution to this problem is to introduce the new auxilliary variable

$$F_d = D_d^{-1} - 1, \tag{A7}$$

so that $D_d = (1 + F_d)^{-1}$.

With these new variables, the elimination scheme (A4) becomes,

$$F_d = \left(H_d + \frac{A_d F_{d-1}}{1 + F_{d-1}}\right) C_d^{-1}, \qquad F_1 = H_1/C_1,$$
$$Z_d = \frac{S_d + A_d Z_{d-1}}{C_d(1 + F_d)}, \qquad Z_1 = B_1^{-1} S_1, \tag{A8}$$

followed by

$$u_d = (1 + F_d)^{-1} u_{d+1} + Z_d, \quad u_{N_D+1} = 0. \tag{A9}$$

An indication of the good numerical conditioning of these equations is that only positive signs appear in them. This implies firstly that the auxilliary quantities introduced here are all positive, and secondly that extreme cancellation problems simply cannot occur.

These equations have been used to solve test problems with very small optical depth increments, of order $\Delta\tau \sim 10^{-15}$, using 32 bit arithmetic with no difficulties. It is important to understand, however, that while this method is not much limited by machine *precision*, it is limited by the machine *exponent range*. In particular, the method cannot treat cases where the optical depth increments actually vanish.

Appendix B: fast solution for the diagonal elements of the inverse of a tridiagonal matrix

Let T be an $N \times N$ tridiagonal matrix and let its inverse be $\lambda \equiv T^{-1}$. We shall show here how the entire set of diagonal elements λ_{ii}, $1 \leq i \leq N$, can be found in order N operations.

The equation for the inverse can be written $T\lambda = 1$, or, in component form,

$$-A_i\lambda_{i-1,j} + B_i\lambda_{ij} - C_i\lambda_{i+1,j} = \delta_{ij}. \tag{B1}$$

For any fixed value of j this equation can be solved by one of two forms of Gaussian elimination. In the usual implementation the elimination proceeds from $i = 1$ to $i = N$, followed by back-substitution from $i = N$ to $i = 1$,

$$D_i = (B_i - A_iD_{i-1})^{-1}C_i, \qquad D_0 \equiv 0, \tag{B2}$$

$$Z_{ij} = (B_i - A_iD_{i-1})^{-1}(\delta_{ij} + A_iZ_{i-1,j}), \qquad Z_{0j} \equiv 0, \tag{B3}$$

$$\lambda_{ij} = D_i\lambda_{i+1,j} + Z_{ij}, \qquad \lambda_{N+1,j} \equiv 0. \tag{B4}$$

It is also possible to implement the method using the reverse order,

$$E_i = (B_i - C_iE_{i+1})^{-1}A_i, \qquad E_{N+1} \equiv 0, \tag{B5}$$

$$W_{ij} = (B_i - C_iE_{i+1})^{-1}(\delta_{ij} + C_iW_{i+1,j}), \qquad W_{N+1,j} \equiv 0, \tag{B6}$$

$$\lambda_{ij} = E_i\lambda_{i-1,j} + W_{ij}, \qquad \lambda_{0j} \equiv 0. \tag{B7}$$

The crucial idea of the present method is to use parts of *both* of these implementations to find the diagonal elements λ_{ii}.

Since $\delta_{ij} = 0$ for $i \neq j$, it follows from Eqs. (B3) and (B6) that

$$Z_{ij} = 0, \quad \text{for } i < j, \tag{B8}$$

$$W_{ij} = 0, \quad \text{for } i > j. \tag{B9}$$

Thus, from Eqs. (B3) and (B4) we obtain, for special choices of i and j,

$$Z_{ii} = (B_i - A_iD_{i-1})^{-1}, \tag{B10}$$

$$\lambda_{ii} = D_i\lambda_{i+1,i} + Z_{ii}, \tag{B11}$$

$$\lambda_{i-1,i} = D_{i-1}\lambda_{ii}, \tag{B12}$$

Similarly, from Eqs. (B6) and (B7),

$$W_{ii} = (B_i - C_iE_{i+1})^{-1}, \tag{B13}$$

$$\lambda_{ii} = E_i\lambda_{i-1,i} + W_{ii}, \tag{B14}$$

$$\lambda_{i+1,i} = E_{i+1}\lambda_{ii}. \tag{B15}$$

From Eqs. (B10), (B11), and (B15) we eliminate Z_{ii} and $\lambda_{i+1,i}$ to obtain

$$\lambda_{ii} = (1 - D_iE_{i+1})^{-1}(B_i - A_iD_{i-1})^{-1}. \tag{B16}$$

The right hand side now depends only on the single-index quantities A_i and B_i, which are given, and D_i and E_i, which can be found by two passes through the depth grid, using the recursion relations (B2) and (B5). Thus all λ_{ii} can be found in order N operations.

If one is already performing a formal solution of the transfer equation based on the matrix T, there is very little extra work involved in determining the diagonal elements λ_{ii}, since the quantities A_i, B_i, C_i, and D_i are common to both problems. One needs only to include the recursion relation (B5) as part of the back-substitution to find the E_i. The quantities $(B_i - A_iD_{i-1})^{-1}$ are required for the recursion relations (B2) and (B3), so they are most conveniently stored, rather than recomputed, for use in equation (B16).

An alternate set of equations for obtaining λ_{ii} can be found by using Equations (B12), (B13), and (B14), eliminating W_{ii} and $\lambda_{i-1,i}$ to obtain

$$\lambda_{ii} = (1 - E_iD_{i-1})^{-1}(B_i - C_iE_{i+1})^{-1}. \tag{B17}$$

This form is more convenient if the associated formal solution of the transfer equation is being done in reverse order, starting with recursions from $i = N$ to $i = 1$.

Once the diagonal elements λ_{ii} have been found, off-diagonal elements can be found from the recursions relations,

$$\lambda_{ij} = D_i\lambda_{i+1,j}, \quad \text{for } i < j, \tag{B18}$$

$$\lambda_{ij} = E_i\lambda_{i-1,j}, \quad \text{for } i > j, \tag{B19}$$

which follow from Eqs. (B4) and (B8), and from Eqs. (B7) and (B9), respectively. Thus a band of width M about the diagonal can be found in order MN operations. The entire inverse could in principle be constructed in this way in order N^2 operations; however, this is not an obvious improvement over the usual method, which also requires of order N^2 operations.

We have implicitly assumed that the elements of the tridiagonal matrix T and its inverse λ are scalars, and thus so are all the various auxilliary quantities introduced here. However, it should be noted that all formulas of this appendix have been written in forms that will apply as well to block tridiagonal matrices. In that case the term "operations" must be interpreted to mean "block matrix operations," so that the method would require of order NF^3 scalar operations for blocks of size $F \times F$.

The method of this appendix can also be formulated in terms of the improved Feautrier solution of appendix A, thus giving it the advantage of better numerical conditioning. Such a formulation has been incorporated into a FORTRAN subroutine, which simultaneously solves the formal solution and solves for a band of the inverse operator. This routine has been extensively used in the code MALI-L and has proven itself to be reliable and accurate.

References

Auer L.H., 1967, ApJ 150, L53

Auer L.H., 1976, J. Quant. Spectrosc. Radiat. Transfer 16, 931

Auer L.H., 1987, in: Numerical Radiative Transfer, ed. W. Kalkofen, Cambridge University Press, Cambridge, p. 101

Avrett E.H., Loeser R., 1987, in: Numerical Radiative Transfer, ed. W. Kalkofen, Cambridge University Press, Cambridge, p. 135

Berrington K.A., Kingston A.E., 1987, J. Phys. B. 20, 6631

Buchler J.H., Auer L.H., 1985, in: Proc. 2nd Intern. Conf. and Workshop on Radiat. Properties of Hot Dense Matter, eds. J. Davis, C. Hooper, R. Lee, A. Merts, B. Rozsnyai, World Scientific, Singapore, p. 58

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

Feautrier P., 1964, C. R. Acad. Sci. Paris 258, 3189

Flannery B.P., Rybicki G.B., Sarazin C.L., 1979, ApJ 229, 1057

Flannery B.P., Rybicki G.B., Sarazin C.L., 1980, ApJS 44, 539

Hamann W.-R., 1985, A&A 148, 364

Hamann W.-R., 1986, A&A 160, 347

Hamann W.-R., 1987, in: Numerical Radiative Transfer, ed. W. Kalkofen, Cambridge University Press, Cambridge, p. 35
Hempe K., Schoenberg, K., 1986, A&A 160, 141
Hillier D.J., 1990, A&A 231, 116
Hummer D.G., Storey P.J., 1987, MNRAS 214, 800
Klein R.I., Castor J.I., Greenbaum A., Taylor D., Dykema P., 1989, J. Quant. Spectrosc. Radiat. Transfer 41, 199
Kutepov A., Hummer D.G., Rybicki G.B., 1991, J. Quant. Spectrosc. Radiat. Transfer (in preparation)
Mihalas D., 1978, Stellar Atmospheres, Freeman, San Francisco
Ng K.C., 1974, J. Chem. Phys. 61, 2680
Olson G.L., Auer L.H., Buchler J.R., 1986, J. Quant. Spectrosc. Radiat. Transfer 35, 431 (OAB)
Olson G.L., Kunasz P.B., 1987, J. Quant. Spectrosc. Radiat. Transfer 38, 325
Pauldrach A., Herrero A., 1988, A&A 199, 262
Puls J., Herrero A., 1988, A&A 204, 219
Rybicki G.B., 1972, in: Line Formation in the Presence of Magnetic Fields, eds. R.G. Athay, L.L. House, G. Newkirk Jr., High Altitude Observatory, Boulder, p. 145
Rybicki G.B., 1984, in: Methods in Radiative Transfer, ed. W. Kalkofen, Cambridge University Press, Cambridge, p. 21
Scharmer G.B., 1984, in: Methods in Radiative Transfer, ed. W. Kalkofen, Cambridge University Press, Cambridge, p. 173
Scharmer G.B., 1981, ApJ 249, 720
Scharmer G.B., Carlsson, M., 1985, J. Computational Physics 59, 56
Schoenberg K., Hempe K., 1986, A&A 163, 151
Varga R.S., 1962, Matrix Iterative Analysis, Englewood, Prentice-Hall
Vinsome P.K.W., 1976, in: Proc. of the 4th Symp. on Reservoir Simulation, Boulder, Soc. of Petroleum Engineers, p. 149
Werner K., Husfeld D., 1985, A&A 148, 417
Weise W.L., Smith M.W., Glennon B.M., 1966, Atomic Transition Probabilities, Vol. 1, NSRDS-NBS 4

Note added in proof: It has come to our attention that the idea of improving the Feautrier method by using the sum of the Feautrier coefficients [our Eq. (A6)] had previously appeared in Nordlund (1982, A&A 107, 1), where it is credited to R.F. Stein (private communication, 1979). Our choice of the auxilliary variable F_d and the resulting elimination scheme [Eqs. (A7)°(A9)], however, appear to be new.