

Radiative transfer with finite elements

I. Basic method and tests

S. Richling¹, E. Meinköhn^{1,2}, N. Kryzhevoi^{1,3}, and G. Kanschat^{2,3}

¹ Institut für Theoretische Astrophysik, Universität Heidelberg, Tiergartenstr. 15, 69121 Heidelberg, Germany

² Institut für Angewandte Mathematik, Universität Heidelberg, Im Neuenheimer Feld 294, 69120 Heidelberg, Germany

³ Interdisziplinäres Zentrum für Wissenschaftliches Rechnen, Universität Heidelberg, Im Neuenheimer Feld 368, 69120 Heidelberg, Germany

Received 16 May 2001 / Accepted 4 October 2001

Abstract. A finite element method for solving the monochromatic radiation transfer equation including scattering in three dimensions is presented. The algorithm employs unstructured grids which are adaptively refined. Adaptivity as well as ordinate parallelization reduce memory requirements and execution time and make it possible to calculate the radiation field across several length scales for objects with strong opacity gradients. An a posteriori error estimate for one particular quantity is obtained by solving the dual problem. The application to a sample of test problems reveals the properties of the implementation.

Key words. radiative transfer – scattering – methods: numerical

1. Introduction

Radiative transfer calculations are an important tool for the interpretation of spectra and images of astrophysical objects. A three-dimensional treatment is required for objects with complex geometry and kinematics. Intruding examples are molecular clouds consisting of clumpy filaments (Wiseman & Ho 1998), circumstellar disks and jets accompanying the formation of multiple stellar systems (Zinnecker et al. 1999), planetary nebulae with precessing bipolar outflows (Miranda et al. 1999), quasars with infrared spectra indicating a clumpy nature of the torus (Haas et al. 2000), and high-redshift radio galaxies with jet-induced star forming regions (Bicknell et al. 2000).

Three-dimensional radiative transfer codes frequently employ the Monte Carlo method. This simple concept of following the path of photon packages works in principle for any problem. But Monte Carlo calculations are very time-consuming, especially in three dimensions, because the error of the results decreases very slowly. When N is the number of simulated photons, the rate of convergence is $\propto N^{-1/2}$ in the worst case. Nevertheless, rates of convergence $\propto N^{-1}$ or even better can be obtained (e.g. Haber 1970). Monte Carlo methods are implemented in codes which consider scattering of photons by dust in different environments (e.g. Witt & Gordon 1996; Wolf et al. 1998; Városi & Dwek 1999; Wood & Reynolds 1999) and

in codes which calculate line spectra in clumpy molecular clouds (Juvela 1998; Park & Hong 1998). However, the line transfer problem in three dimensions is also widely treated with the method of short characteristics (Väth 1994; Papkalla 1995; Fabiani Benedicho et al. 1997; Folini 1998).

Alternatively, we will introduce a finite element method for the solution of the radiative transfer equation. Finite element methods are still very infrequently used for the modeling of astrophysical systems. However, Meier (1999) proposes a finite element code suitable for a wide range of astrophysically interesting differential equations. Radiative transfer calculations employing finite element methods were already performed by Dykema et al. (1996). They solve the zeroth and first moments of the transfer equation in two dimensions and use a finite element technique on fixed quadrilateral meshes for the computation of the Eddington tensor. In contrast, we solve the radiative transfer equation directly in three dimensions for discrete ordinates on a hexahedral mesh which is adaptively refined. Therefore, it is possible to calculate the radiation field across several length scales even for objects with strong opacity gradients. In addition, we apply ordinate parallelization for the integral operator via the Message Passing Interface (MPI) which permits the use of computer clusters.

This work provides an extensive description of the basic monochromatic radiative transfer code originally developed by Kanschat (1996) and of some tests which reveal

Send offprint requests to: S. Richling,
e-mail: richling@ita.uni-heidelberg.de

the main properties of the finite element implementation. The series will be continued with the presentation and application of extensions to frequency-coupled radiative transfer: a line transfer version which considers the kinematics of the material and complete frequency redistribution and a continuum transfer version which utilizes the energy balance equation to determine self-consistently a temperature distribution. The applications will focus on the radiation fields of high-redshift galaxies.

In Sect. 2, we describe the application of the finite element method to the monochromatic radiative transfer equation and the details of our implementation, i.e. discretization, grid adaptation, structure and solution of the resulting matrix, and parallelization strategy. In Sects. 3, 4 and 5 we present results for three test problems: the searchlight test problem, the angular distribution of the specific intensity escaping a plane-parallel layer, and the propagation of radiation emitted from an extended region embedded in a scattering halo. We summarize our results in Sect. 6.

2. The finite element method

2.1. The radiative transfer problem

Our aim is the calculation of the radiation field in diffuse matter in space. Assuming the matter is surrounded by a vacuum, we will only consider a convex domain containing the area of interest. Radiation leaving this domain will not enter it again. Inside this three-dimensional domain $\Omega \subset \mathbb{R}^3$, the specific intensity \mathcal{I} satisfies the monochromatic radiative transfer equation

$$\begin{aligned} \mathbf{n} \cdot \nabla_x \mathcal{I}(\mathbf{x}, \mathbf{n}) + \kappa(\mathbf{x}) \mathcal{I}(\mathbf{x}, \mathbf{n}) \\ + \sigma(\mathbf{x}) \left(\mathcal{I}(\mathbf{x}, \mathbf{n}) - \frac{1}{4\pi} \int_{S^2} P(\mathbf{n}', \mathbf{n}) \mathcal{I}(\mathbf{x}, \mathbf{n}') d\omega' \right) = f(\mathbf{x}), \end{aligned} \quad (1)$$

where $\mathbf{x} \in \Omega$ is the space variable and \mathbf{n} the direction pointing to solid angle $d\omega$ of the unit sphere S^2 . The optical properties of the matter are given by the absorption coefficient $\kappa(\mathbf{x})$ and the scattering coefficient $\sigma(\mathbf{x})$. The angular phase function P occurring in the scattering integral is normalized, such that $\frac{1}{4\pi} \int P(\mathbf{n}', \mathbf{n}) d\omega' = 1$. The source term

$$f(\mathbf{x}) = \kappa(\mathbf{x})B(T(\mathbf{x})) + \epsilon(\mathbf{x}) \quad (2)$$

consists of thermal emission depending on a temperature distribution $T(\mathbf{x})$ and an additional emissivity $\epsilon(\mathbf{x})$ of a point source or an extended object. B is the Planck-function. To be able to solve Eq. (1), boundary conditions of the form

$$\mathcal{I}(\mathbf{x}, \mathbf{n}) = g(\mathbf{x}, \mathbf{n})$$

must be imposed on the “inflow boundary”

$$\Gamma_- = \{(\mathbf{x}, \mathbf{n}) \in \Gamma \mid \mathbf{n}_\Gamma \cdot \mathbf{n} < 0\},$$

where \mathbf{n}_Γ is the unit vector perpendicular to the boundary surface Γ . The sign of the product $\mathbf{n}_\Gamma \cdot \mathbf{n}$ describes the “flow direction” of the photons across the boundary. If there are no light sources outside the modeled domain, the function g will be zero everywhere.

The left hand side of Eq. (1) will be abbreviated as an operator \mathcal{A} applied to the intensity function \mathcal{I} , yielding the very compact operator form of the radiative transfer equation

$$\mathcal{A}\mathcal{I}(\mathbf{x}, \mathbf{n}) = f(\mathbf{x}). \quad (3)$$

This equation is five-dimensional in a 3D domain. In addition, the numerical solution is complicated by rapidly changing values of \mathcal{I} in very small parts of the domain and smooth transport in larger regions. Using a reasonable resolution of $h = 1/1000$ in space and 1000 directions or *ordinates* already results in 10^{12} unknowns in the 3D case. Since this amount of data cannot be handled even on most advanced supercomputers, the application of efficient error estimation and grid adaptation techniques is necessary to obtain reliable quantitative results. Finite element methods, in particular so-called Galerkin methods, are most suitable for these techniques.

2.2. Finite element discretization

Equation (1) is analyzed in Dautray & Lions (1993) and the natural space for finding solutions is

$$W = \{\mathcal{I} \in L^2(\Omega \times S^2) \mid \mathbf{n} \cdot \nabla_x \mathcal{I} \in L^2(\Omega \times S^2)\},$$

where L^2 is the Lebesgue space of the square integrable functions. If we consider homogeneous vacuum boundary condition, i.e. $g(\mathbf{x}, \mathbf{n}) = 0$, the solution space is

$$W_0 = \{\mathcal{I} \in W \mid \mathcal{I} = 0 \text{ on } \Gamma_-\}.$$

In order to apply a finite element method, we have to use a weak formulation of (1). Therefore, we multiply both sides of Eq. (1) by a trial function $\varphi(\mathbf{x}, \mathbf{n})$ and integrate over the whole domain $\Omega \times S^2$. Thus, the weak formulation reads: Find $\mathcal{I} \in W_0$, such that $\forall \varphi \in W_0$

$$\begin{aligned} \int_{\Omega} \int_{S^2} f(\mathbf{x}) \varphi(\mathbf{x}, \mathbf{n}) d\omega d^3x = \\ \int_{\Omega} \int_{S^2} \mathbf{n} \cdot \nabla_x \mathcal{I}(\mathbf{x}, \mathbf{n}) \varphi(\mathbf{x}, \mathbf{n}) d\omega d^3x \\ + \int_{\Omega} \int_{S^2} (\kappa(\mathbf{x}) + \sigma(\mathbf{x})) \mathcal{I}(\mathbf{x}, \mathbf{n}) \varphi(\mathbf{x}, \mathbf{n}) d\omega d^3x \\ - \frac{1}{4\pi} \int_{\Omega} \int_{S^2} \int_{S^2} \sigma(\mathbf{x}) P(\mathbf{n}', \mathbf{n}) \mathcal{I}(\mathbf{x}, \mathbf{n}') \varphi(\mathbf{x}, \mathbf{n}) d\omega' d\omega d^3x. \end{aligned} \quad (4)$$

By extending the definition of the L^2 -scalar product we introduce the abbreviation

$$(\mathcal{I}, \varphi) = (\mathcal{I}, \varphi)_{\Omega \times S^2} = \int_{\Omega} \int_{S^2} \mathcal{I} \varphi d\omega d^3x.$$

Thus, the weak formulation of the operator form (3) is

$$(\mathcal{A}\mathcal{I}, \varphi) = (f, \varphi) \quad \forall \varphi \in W_0. \quad (5)$$

If there is no scattering, i.e. $\sigma(\mathbf{x}) = 0$ on Ω , the problem decouples to a system of convection equations on Ω . These equations are hyperbolic. If the solutions are not smooth, standard finite element techniques applied to this type of equations are known to produce spurious oscillations. We can achieve stability by applying the streamline diffusion modification:

$$(\mathcal{A}\mathcal{I}, \varphi + \delta \mathbf{n} \cdot \nabla_x \varphi) = (f, \varphi + \delta \mathbf{n} \cdot \nabla_x \varphi) \quad \forall \varphi \in W_0. \quad (6)$$

The cell-wise constant parameter function δ depends on the local mesh width and the coefficients κ and σ . Note that the solution of (5) solves (6), too. No additional consistency error is induced by the stabilization. In the following, the streamline diffusion discretization term will be omitted but implicitly assumed.

Applying standard Galerkin finite elements to solve (5), we choose a finite dimensional subspace W_h of W consisting of functions that are piecewise polynomials with respect to a subdivision or *triangulation* \mathbb{T}_h of $\Omega \times S^2$. The mesh size h is the piecewise constant function defined on each triangulation cell K by $h|_K = h_K = \text{diam } K$. The discrete analogue of (5) is finding $\mathcal{I}_h \in W_h$, such that

$$(\mathcal{A}\mathcal{I}_h, \varphi_h) = (f, \varphi_h) \quad \forall \varphi_h \in W_h. \quad (7)$$

The construction of the subspace W_h needs some further consideration (see also Kanschat 1996). The discretized domain is a tensor product of two sets of completely different length scales: While Ω represents a domain in physical space, S^2 is the unit sphere in the Euclidean space \mathbb{R}^3 . Therefore, we use a tensor product splitting of the five-dimensional domain $\Omega \times S^2$, such that a grid cell of the five-dimensional grid will be the tensor product of a two-dimensional cell K_ω and a three-dimensional cell K_x . Accordingly, the mesh sizes with respect to \mathbf{x} and ω will be different.

On S^2 we use fixed discretizations based on a refined icosahedron as illustrated in Fig. 1. Quadrature points are the cell centers projected on S^2 . This projection method guarantees equally distributed quadrature points, each associated with a solid angle $d\omega$ of the same size. In comparison to other numerical quadrature methods, additional symmetry conditions are redundant and discretization artifacts at the poles are avoided. Furthermore, we use piecewise constant trial functions. That way, the seven-dimensional integration of the scattering term in the weak formulation (4) can be calculated very efficiently, since integration weights are not necessary. For example, the integration of the intensity over the whole unit sphere is simply replaced by a sum divided by the number of discrete ordinates M

$$\frac{1}{4\pi} \int_{S^2} \mathcal{I}(\mathbf{n}) d\omega \rightarrow \frac{1}{M} \sum_{j=1}^M \mathcal{I}_j.$$

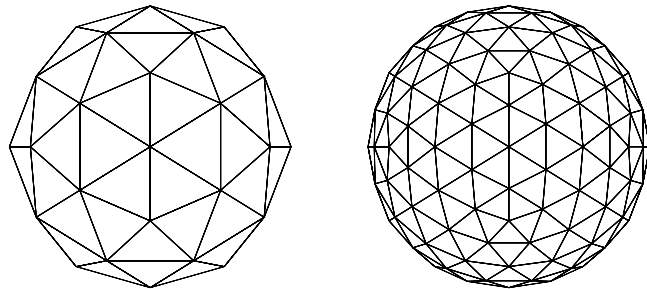


Fig. 1. Refined icosahedron with 80 and 320 cells. The cell centers projected on a unit sphere are equally distributed.

This is crucial, since for a longitude-latitude discretization a simple Gauss-formula accurate for quadratic polynomials involves 72 integration points. Nevertheless, our scheme is second order accurate in the evaluated ordinate points due to super-convergence.

For the space domain Ω we use locally refined hexahedral meshes. The mesh size with respect to the space variable will be denoted by h in the course of this publication. Since the boundaries are arbitrary for our astrophysical application, we can choose a unit cube for Ω and do not have to worry about boundary approximation. We use continuous piecewise trilinear trial functions in space.

2.3. Error estimation and adaptivity

The calculation of complex radiation fields in astrophysics often requires high resolution in parts of the domain, for example in regions with strong opacity gradients. Then, reliable error bounds are necessary to rule out numerical errors. Due to the high dimension of the computational domain, a well suited method for error estimation and grid adaptation is necessary to achieve results of sufficient accuracy even on parallel computers.

In addition, computational goals in astrophysics are often more specific. The result of a simulation is to be compared with observations to develop a physical model for an object. For instance, in the case of a distant unresolved object, the total flux leaving the domain Ω in *one* particular direction is of interest. Generally, a measured quantity like the total flux can be expressed as a linear functional $\mathcal{M}(\mathcal{I})$. By linearity, the error of the measured quantity is $\mathcal{M}(\mathcal{I}) - \mathcal{M}(\mathcal{I}_h) = \mathcal{M}(e)$, where $e = \mathcal{I} - \mathcal{I}_h$. In the following, we will show, that it is possible to obtain an a posteriori estimate for $\mathcal{M}(e)$, even if the exact solution \mathcal{I} is unknown.

Suppose, $z(\mathbf{x}, \mathbf{n})$ is the solution of the dual problem

$$\mathcal{M}(\varphi) = (\varphi, \mathcal{A}^* z) \quad \forall \varphi \in W_0, \quad (8)$$

where the dual radiative transfer operator is defined by

$$\begin{aligned} \mathcal{A}^* z(\mathbf{x}, \mathbf{n}) = & -\mathbf{n} \cdot \nabla_x z(\mathbf{x}, \mathbf{n}) + (\kappa(\mathbf{x}) + \sigma(\mathbf{x}))z(\mathbf{x}, \mathbf{n}) \\ & - \sigma(\mathbf{x}) \frac{1}{4\pi} \int_{S^2} P(\mathbf{n}', \mathbf{n}) z(\mathbf{x}, \mathbf{n}') d\omega'. \end{aligned}$$

The boundary conditions for the dual problem are complementary to those in the primal problem, i.e. $\mathcal{I} = 0$ on the “outflow boundary” $\Gamma_+ = \{(\mathbf{x}, \mathbf{n}) \in \Gamma \mid \mathbf{n}_\Gamma \cdot \mathbf{n} > 0\}$. Then, we get the formal error representation

$$\begin{aligned} \mathcal{M}(e) &= (e, \mathcal{A}^* z) \\ &= (\mathcal{A}e, z) \\ &= (\mathcal{A}e, z - z_i) \\ &= \sum_{K \in \mathbb{T}_h} (f - \mathcal{A}\mathcal{I}_h, z - z_i)_K \end{aligned} \quad (9)$$

for arbitrary $z_i \in W_h$. In the third line of Eq. (9), we use a characteristic feature of finite element methods, the Galerkin orthogonality

$$(\mathcal{A}\mathcal{I} - \mathcal{A}\mathcal{I}_h, \varphi_h) = 0 \quad \forall \varphi_h \in W_h. \quad (10)$$

Since the dual solution z is unknown, it is a usual approach to apply Hölder’s inequality and standard approximation estimates of finite element spaces to obtain the estimate

$$\begin{aligned} \mathcal{M}(e) \leq \eta &= \sum_{K \in \mathbb{T}_h} \eta_K \\ \eta_K &= C_K h_K^2 \|\varrho\|_K \|\nabla^2 z\|_K, \end{aligned} \quad (11)$$

where the constant C_K is determined by local approximation properties of W_h . The residual function ϱ of \mathcal{I}_h is defined by $\varrho = f - \mathcal{A}\mathcal{I}_h$. Since the dual solution z is not available analytically, it is usually replaced by the finite element solution z_h to the dual problem (8). This involves a second solution step of the same structure as the primal problem. It is clear, that by this replacement the error estimate (11) is not strictly true anymore. Experience shows that the additional error is small and the estimate may be used if multiplied with a small security factor larger than one.

A first approach in the development of strategies for grid refinement based on a posteriori error estimates is the control of the global energy or L^2 -error involving only local residuals of the computed solution (Führer & Kanschat 1997). Formally, using the functional $\mathcal{M}(\varphi) = \|e\|^{-1}(e, \varphi)$ in the dual problem (8), we obtain $\mathcal{M}(e) = \|e\|^{-1}(e, e) = \|e\|$ for the left hand side of (9). Estimating the right hand side of (9) yields

$$\begin{aligned} \|e\| \leq \tilde{\eta} &= \sqrt{\sum_{K \in \mathbb{T}_h} \eta_{L^2}^2} \\ \eta_{L^2} &= C_K C_s h_K^2 \|\varrho\|_K, \end{aligned} \quad (12)$$

where C_s only depends on the shape and size of Ω . This a posteriori bound for the L^2 -error is well-known (see e.g. Erikson et al. 1995 or Verfürth 1996). Estimates of this kind are always sub-optimal in the case of inhomogeneous coefficients and especially with dominant scattering. Still, these L^2 -indicators η_{L^2} provide a good refinement criterion to study the qualitative behavior of the solution \mathcal{I} everywhere in Ω . See also Wehrse et al. (1999) for an illustration of this approach.

The grid refinement process on the basis of an a posteriori error estimate is organized in the following way: Suppose that some error tolerance TOL is given. The aim is to find the most economical grid \mathbb{T}_h on which

$$|\mathcal{M}(e)| \leq \hat{\eta}(\mathbb{T}_h) = \sqrt{\sum_{K \in \mathbb{T}_h} \eta_K^2(I_h)} \leq TOL, \quad (13)$$

with *local refinement indicators* η_K taken from (11). Having computed the solution on a coarse grid, we apply the so-called *fixed fraction* grid refinement strategy (Kanschat 1996; Becker & Rannacher 2001): The cells are ordered according to the size of η_K and a fixed portion ν (say 30%) of the cells with largest η_K is refined. This guarantees, that in each refinement cycle a sufficient large number of cells is refined. Then, a solution is computed on the new grid and the process is continued until the prescribed tolerance is achieved. This algorithm is especially valuable, if a computation “as accurate as possible” is desired, that is, the tolerance is not reached, but computer memory is exhausted. Then, the parameter ν has to be determined by the remaining memory resources.

2.4. Resulting matrix structure

Given a discretization with N degrees of freedom (=number of *vertices* of the hexahedral mesh) in Ω and M ordinates in S^2 , the discrete system has the form

$$\mathbf{A} \cdot \mathbf{u} = \mathbf{f}, \quad (14)$$

with the vector \mathbf{u} containing the discrete specific intensities and the vector \mathbf{f} the values of the source term. Both vectors are of length $(N \cdot M)$. \mathbf{A} is a $(N \cdot M) \times (N \cdot M)$ matrix. Applying the tensor product structure proposed above, we may write

$$\mathbf{A} = \mathbf{T} + \mathbf{K} + \mathbf{S}$$

with the block structure

$$\begin{aligned} \mathbf{T} &= \begin{pmatrix} \mathfrak{T}_1 & \mathbf{0} \\ & \ddots \\ \mathbf{0} & \mathfrak{T}_M \end{pmatrix}, \\ \mathbf{K} &= \begin{pmatrix} \mathfrak{K}_1 & \mathbf{0} \\ & \ddots \\ \mathbf{0} & \mathfrak{K}_M \end{pmatrix}, \\ \mathbf{S} &= \begin{pmatrix} \omega_{11}\mathfrak{S}_1 & \dots & \omega_{1M}\mathfrak{S}_1 \\ \vdots & \ddots & \vdots \\ \omega_{M1}\mathfrak{S}_M & \dots & \omega_{MM}\mathfrak{S}_M \end{pmatrix}, \end{aligned}$$

where $\omega_{il} = P(\mathbf{n}_i, \mathbf{n}_l)/M$. Considering the finite element streamline diffusion discretization, the entries of the $N \times N$ blocks are defined by

$$\begin{aligned} \mathfrak{T}_i^{jk} &= (\varphi_j + \delta \mathbf{n}_i \cdot \nabla_x \varphi_j, \mathbf{n}_i \cdot \nabla_x \varphi_k) \\ \mathfrak{K}_i^{jk} &= (\varphi_j + \delta \mathbf{n}_i \cdot \nabla_x \varphi_j, (\kappa + \sigma) \varphi_k) \\ \mathfrak{S}_i^{jk} &= (\varphi_j + \delta \mathbf{n}_i \cdot \nabla_x \varphi_j, -\sigma \varphi_k), \end{aligned}$$

where $j = 1, \dots, N$ and $k = 1, \dots, N$.

Table 1. Test I: for each refinement step are given the minimum intensity \mathcal{I}_{\min} , the relative number of vertices with negative intensity values n_{low} , the maximum intensity \mathcal{I}_{\max} , the relative number of vertices with intensity values greater than unity n_{high} , the total number of vertices N , and the total number of required vertices N_{uni} for a grid with uniformly distributed cells of the same resolution.

refinement step	\mathcal{I}_{\min}	n_{low} in %	\mathcal{I}_{\max}	n_{high} in %	N	N_{uni}
0	-0.0457	35	1.000	0.37	289	289
1	-0.0805	39	1.076	0.32	631	1 089
2	-0.0833	38	1.104	1.7	1 395	4 225
3	-0.0963	34	1.109	7.6	2 995	16 641
4	-0.1018	36	1.171	14	6 312	66 049
5	-0.0833	36	1.106	24	13 084	263 169
6	-0.0833	34	1.080	24	28 005	1 050 625
7	-0.0929	32	1.095	25	57 606	4 198 401
8	-0.0934	29	1.093	23	118 179	16 785 409

2.5. Iterative methods

The linear system of equations resulting from the discretization described above is large (10^7 unknowns at least), sparse, and strongly coupled due to the integral operator.

The standard algorithm used in astrophysics to solve these systems is the so-called Λ -iteration. Considering the whole discrete system, it is a Richardson method with nearly block-Jacobi-preconditioning. Using a full Jacobi preconditioner is a first step to better convergence rates (Turek 1993). Since the transport operator is inverted explicitly, these methods converge very fast for transport dominated problems. Exploiting the triangular matrix structure of upwind-discretizations, the inversion is indeed very cheap (one matrix-vector-multiplication).

Unfortunately, in the interesting case of scattering dominance this method – like other stationary iterations – breaks down, since the condition number of the iteration matrix becomes very large. Since the convergence rate of preconditioned Richardson iteration methods only depends on the condition number, these are not suited for the scattering dominated case.

The eigenvalue distribution of the scattering matrix \mathcal{S} is clustered, where usually one of them is zero and the others are close to unity or at least bounded away from zero. We exploit this structure by applying a Krylov space method like GMRES or bi-CGSTAB (see e.g. Sleijpen & Fokkema 1993 and references therein) to the solution. Unfortunately, the scattering operator is perturbed by the transport operator. As a differential operator, it requires preconditioning by multi-level methods to obtain convergence independent of the mesh size.

2.6. Parallelization

Transport dominated problems differ in one specific point from elliptic problems: there is a distinct direction of information flow. This has to be considered in the development of parallelization strategies. While a domain decomposition for Poisson’s equation should minimize the length of

interior edges, this does not yield an efficient method for transport equations.

In Kanschat (2000), it was concluded that parallelization strategies for transport equations should not divide the domain across the transport direction. The solution of the radiative transfer equation consists of a bunch of transport inversions for different directions. The construction of an efficient domain decomposition method would require a direction dependent splitting of the domain Ω . Since this causes immense implementation problems, we decided to use ordinate parallelization.

This strategy distributes the ordinate space S^2 of the radiative transfer equation. Since we use discontinuous shape functions for the ordinate variable and there is no local coupling due to the integral operator, this results in a true non-overlapping parallelization. Clearly, it has disadvantages too: As the integral is a global operator, ordinate parallelization involves global communication. Additionally, the resolution in space is restricted by the per node memory and not by the total memory of the machine.

3. Test I: Searchlight beam test

The searchlight beam test is a standard test for the transport operator: a narrow beam of radiation is introduced into the computational grid at a certain angle. The beam should cross a vacuum ($\kappa = 0, \sigma = 0, f = 0$) without dispersion.

For this test we started with a two-dimensional, globally pre-refined grid with 16×16 cells which covers the domain $x = [-1, 1]$ and $y = [-1, 1]$. This grid consists of 17×17 vertices and has a resolution of $\Delta x = \Delta y = 0.125$. The beam of radiation with intensity $\mathcal{I}_0 = 1.0$ was launched at $x = -1$ between $y = [-0.875, -0.750]$ at an angle of approximately 45° . We calculated the radiation field for the basic grid and 8 successively refined grids, where in each step 25% of the cells with the worst L^2 -indicator η_{L^2} (see Sect. 2.3) were refined.

Figure 2 shows the intensity distribution and the grid structure for step 2, 4 and 8. The minimal cell size of

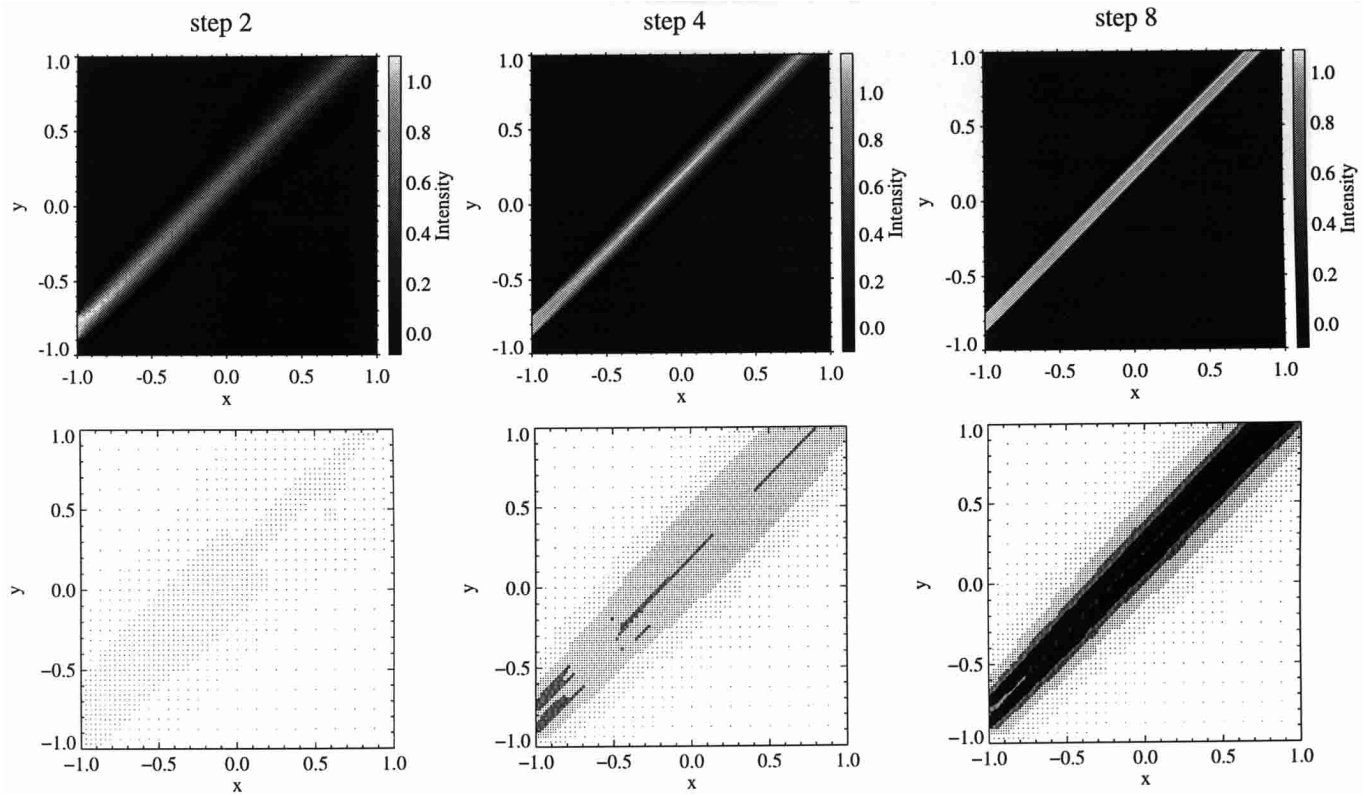


Fig. 2. Test I: intensity distribution (top) and structure of the grid (bottom) for refinement steps 2, 4 and 8. Each dot marks the location of a vertex. The incoming beam is introduced at $x = -1.0$ between $y = [-0.875, -0.750]$.

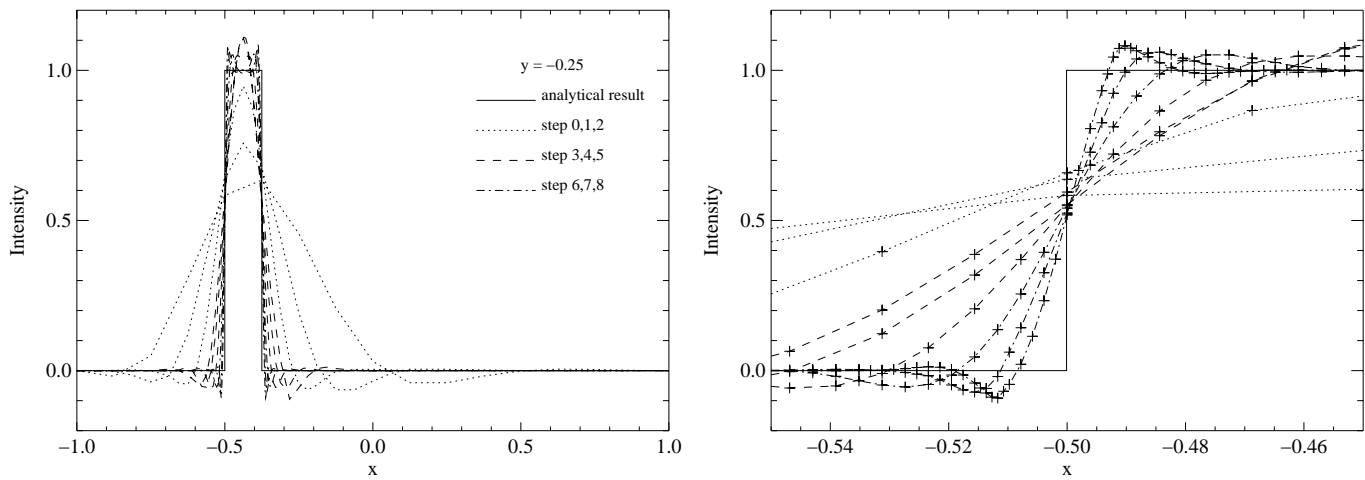


Fig. 3. Test I: intensity distribution at $y = -0.25$ for different adaptive steps as indicated. The right hand side gives a more closer view to the range $x = [-0.55, -0.45]$. In addition, the crosses mark the location of the vertices.

the grid at step 2 corresponds to an equidistant grid with 64×64 cells. The dispersion of the beam at this step is qualitatively comparable to the results of the short characteristic method by Kunasz & Auer (1988). The following refinement steps resolve at first the beam and later mainly the fringes of the beam which is most clearly visible near the entrance of the beam into the grid. At step 8 the intensity distribution shows a straight beam with negligible dispersion. This result is obtained with about 10^5 vertices

which is less than a hundredth of the vertices required for an equidistant grid with the same resolution (see last two rows of Table 1).

The functioning of the code is demonstrated in detail in Fig. 3. Here the intensity distribution at $y = -0.25$ is shown for all refinement steps. The right hand side of this figure magnifies the left fringe of the beam. The location of the vertices are indicated by crosses. As already expected from the grid structure, the first steps begin to resolve

the beam itself and the following steps try to reproduce the intensity distribution at the sharp edges of the beam. Since the streamline diffusion discretization is a second order method, we encounter the problem of negative intensities and overshooting. Table 1 shows this behavior more quantitatively. For all steps the deviation of the minimum intensity \mathcal{I}_{\min} and the maximum intensity \mathcal{I}_{\max} from the analytical solution is about 10% of \mathcal{I}_0 . With increasing resolution the fraction of vertices with negative intensities n_{low} is slightly declining from 40% to 30%, whereas the fraction of vertices with intensity values greater than unity n_{high} is continuously increasing to a fraction of $\sim 25\%$. The minimum intensities are somewhat better than those obtained with the parabolic upwind interpolation scheme by Kunasz & Auer (1988), who find $\mathcal{I}_{\min} \simeq -0.15$ for equally spaced grids as well as for logarithmic grids. Unsurprisingly, their relative number of cells with negative intensities $n_{\text{low}} \simeq 20\%$ is smaller, because the finite element code just favors the generation of finer cells in regions of discontinuities where negative intensity values occur. Negative intensity values cause problems in cases where even the mean intensity J turns out to be negative and therefore prevent the determination of important physical quantities. A proper treatment of these cases, instead of simply adopting $\mathcal{I} = 0$ as a lower limit, would require expensive code development involving the tracing of discontinuities in order to switch between second order and first order methods. Refer to *shock capturing* methods described in the literature, e.g. Eriksson & Johnson (1993).

4. Test II: Radiation field of a plane-parallel layer

In order to test the scattering operator we calculate the radiation field of an infinitely extended plane-parallel layer. In this case, we compare the results with those of an analytical method, the separable representation (SR) method, and a finite difference (FD) code. Both methods are briefly described in the appendix.

The test starts with a homogeneous plane-parallel layer with thickness Δ . We assume zero boundary conditions and a constant source term $f = \chi(1 - \gamma)$ with extinction coefficient $\chi = \kappa + \sigma$ and albedo $\gamma = \sigma/\chi$. We calculate the radiation field for optical depth $\tau = \chi\Delta = 0.4, 2, 6, 20$ and albedo $\gamma = 0.02, 0.8, 0.98$ and investigate the angular dependence of the specific intensity escaping the plane-parallel layer.

For the calculations with the FE code we approximate the infinitely extended layer with a three-dimensional anisotropic grid with initially 4^3 cells which have an aspect ratio of 100:100:1. For each optical depth and albedo we perform up to a maximum of seven refinement steps and tap the outgoing intensity at the middle of the surface of the layer. We use local refinement criterions η_K taken from (11) for a functional $\mathcal{M}(e) = e(\mathbf{x}_0, \mathbf{n}_0)$, i.e. only the intensity leaving the domain Ω from *one* point \mathbf{x}_0 in *one* particular direction \mathbf{n}_0 is of interest. An aspect ratio of 100:100:1 means we get reliable results for $\mu \geq 0.02$,

where μ is the angle-cosine of ray direction \mathbf{n} relative to the vertical.

The FD code is also a three-dimensional code (Appendix B). Here, we use a numerical grid with 32^3 cells with the same aspect ratio of 100:100:1. The iterative procedure for the mean intensity J is carried out until the relative change of J is smaller than 10^{-5} .

The SR method (Appendix A) enables us to obtain the solution in terms of a function which is an infinite series. An approximation of this series by means of a finite sum is possible. A very good agreement of the approximated function with the exact one can already be obtained in the approximation with $N = 5$ for a wide range of parameters, where N is the number of terms in the finite sum. For small μ the exact and the approximated function show a different behavior which causes wrong results in this range. But the higher the order of the approximation N the better are the results for small μ . We calculated the specific intensity for all cases with $N = 10$.

Figure 4 shows the angular distribution of the specific intensity escaping from a plane-parallel layer obtained with the three methods. In general, the three codes yield the same results. Excellent agreement is obtained between the SR method and the FE method. The FE code needed 2 to 3 refinement steps for low optical depth and 5 refinement steps for $\tau = 20.0$ to match the solution of the SR method very well. The resolution of refinement step 5 corresponds to a globally refined grid with 128^3 cells. The FD code is not able to reproduce the results of the other methods for high optical depth and large albedo. In these cases the results of the FD code are too low. The results can be improved by using a grid with higher spatial resolution, but even a grid with 128^3 cells fails to produce equally good results as the FE code. In order to accomplish the results of the SR method, the required accuracy of J and the number of ordinates M play a minor role. Since the FD code only uses a first order discretization, the very good results of the FE code must be attributed to the second order streamline diffusion discretization.

The general form of the intensity distributions displayed in Fig. 4 becomes clear when considering the change of \mathcal{I} along the distance ds

$$\frac{d\mathcal{I}}{ds} = -\chi [\mathcal{I} - \gamma J - (1 - \gamma)]. \quad (15)$$

The solution of this equation for a given J and zero boundary condition is

$$\mathcal{I}(\mu) = \int_0^{\tau/\mu} (\gamma J + (1 - \gamma)) \exp(-\tau') d\tau', \quad (16)$$

where $d\tau' = \chi ds$. For $\gamma \sim 0$ the intensity distribution is

$$\mathcal{I}(\mu) = 1 - \exp(-\tau/\mu), \quad (17)$$

which explains the decrease of the intensity with increasing μ in the optical thin case and the flat distribution $\mathcal{I}(\mu) = 1$, for higher optical depth. For $\gamma \sim 1$ the intensity distribution is

$$\mathcal{I}(\mu) = \int_0^{\tau/\mu} J \exp(-\tau') d\tau', \quad (18)$$

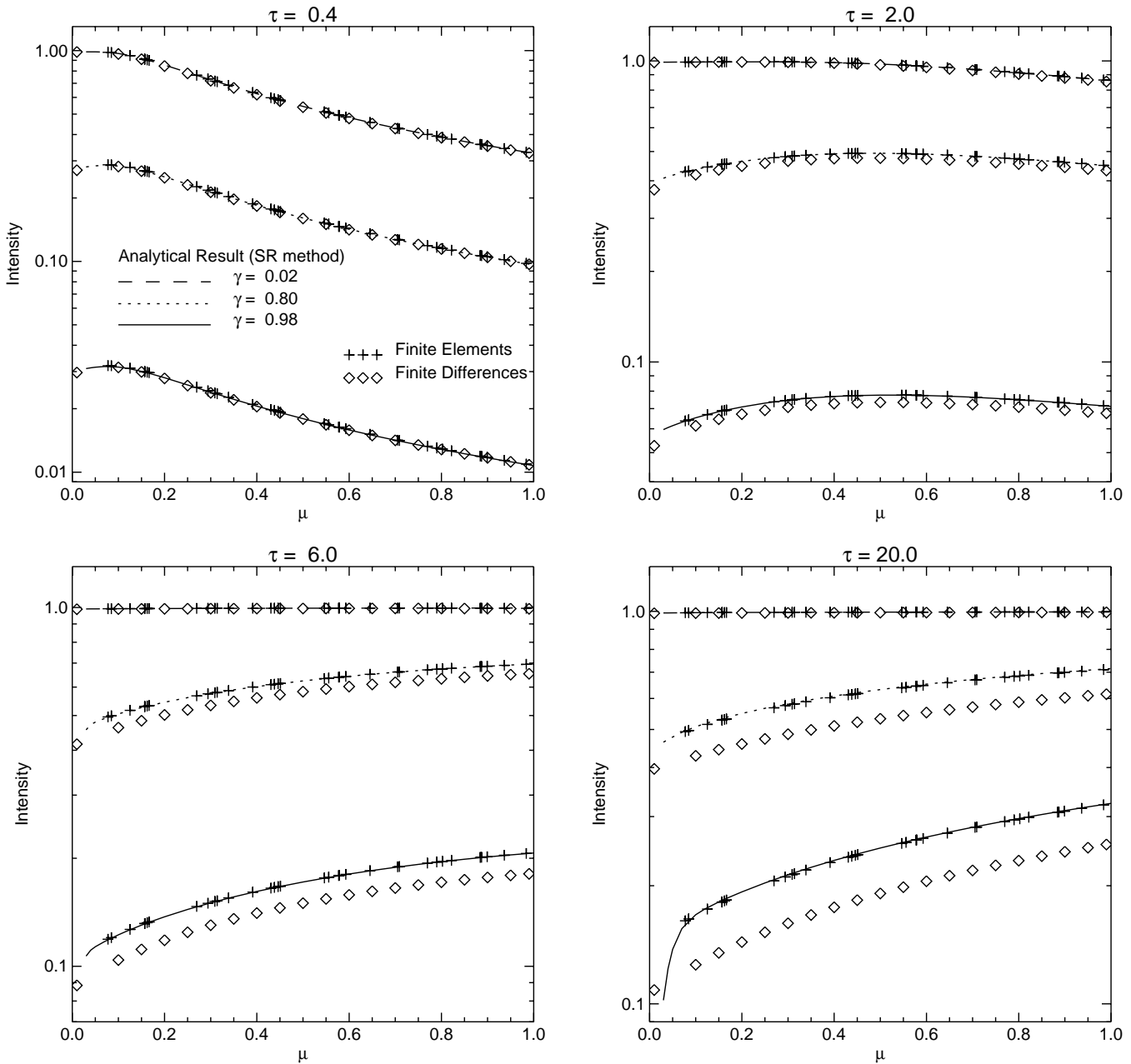


Fig. 4. Test II: angular distribution of the intensity for different total optical depth τ and albedo γ .

and depends on the spatial distribution of J within the layer. In the optical thin case we find a nearly homogeneous distribution for J . Hence, $\mathcal{I}(\mu)$ shows the same behavior for low and high albedo. At high optical depth J is stronger in the center of the layer and falls off to the edges. Since the contributions to the intensity along the path are only significant until $\tau' \sim 1$, the intensity has its maximum at $\mu = 1$.

We use this test problem to point out the advantage of employing an adaptive grid refinement process as described in Sect. 2.3 compared with the commonly used structured grids. Therefore, we calculate the specific intensity in a particular direction $\mu = 0.705$ on a variety of grids. The optical depth τ and albedo γ are 20 and 0.80, respectively. Table 2 lists the number of vertices and the

computed value of the specific intensity escaping the layer from the point \mathbf{x}_0 . The FE method uses grids which are generated by using local refinement criteria as described in Sect. 2.3 (DUAL, L2) and a global refinement criterion where all cells are refined in every step (GLOBAL). L2 indicates the control of the L^2 -error and DUAL employs the functional for a point value as described at the beginning of this section. The FD code is restricted to the use of structured grids only. First, we discuss the quality of the L2 and the DUAL grid compared with the quality of the structured grid used by the FE code. The quality of a grid is related to the question of how economic it is. This is defined by the number of vertices needed to produce an accurate solution. Table 2 shows that the L2 refinement strategy already improves the quality of the

Table 2. Test II: investigation of the solution of the specific intensity in a particular direction $\mu = 0.705$ for a variety of adaptively refined grids (L2, DUAL) and a structured grid (GLOBAL). The optical depth and the albedo is $\tau = 20$ and $\gamma = 0.80$, respectively.

		FE				FD	
DUAL		L2		GLOBAL		GLOBAL	
vertices	value	vertices	value	vertices	value	vertices	value
125	0.4274	125	0.4274	125	0.4274	35 937	0.5760
501	0.5960	579	0.5250	729	0.5912	274 625	0.6162
2 213	0.6609	2 474	0.6397	4 913	0.6420	2 146 689	0.6417
8 103	0.6723	10 935	0.6616	35 937	0.6612		
14 953	0.6740	86 903	0.6687	274 625	0.6686		

Table 3. Test II: comparison of the memory and CPU requirements of the three-dimensional codes for modeling an optical thick and an optical thin plan-parallel layer. The albedo is $\gamma = 0.80$ in both cases. The FE method employs the DUAL grid refinement strategy. The final grid has 670 and 4263 vertices for the optical thin and thick case, respectively. The FD code uses a structured grid (GLOBAL) with 65^3 vertices for $\tau = 0.4$ and 129^3 vertices for $\tau = 20$. In the latter case, the resolution of the spatial grid is too low to obtain a solution as accurate as for the FE code (see Table 2).

	$\tau = 0.4$						$\tau = 20$					
	FE			FD			FE			FD		
processors	1	2	4	8	16	1	1	2	4	8	16	1
memory [MB]	31.9	16.7	9.0	5.2	3.3	13	196	99.6	52.2	28.7	15.1	99
time [s]	846.5	371.8	121.4	56.2	48.2	1275.0	3167.2	1364.2	428.7	207.1	136.1	9878.9

generated grid by a factor of 3 to 4. A dramatic enhancement is obtained by applying the DUAL refinement process, since approximately 20 times less vertices are needed to produce a similar solution as in the GLOBAL case. The DUAL local refinement criterions η_K taken from (11) generate a grid that is much more appropriate for our test problem. Due to the local weights $\|\nabla^2 z\|_K$ the indicators η_K forces the refinement of the cells close to the boundary point where the outgoing intensity is measured. Finally, we compare the solution of the FE code (for the GLOBAL grid) with the solution of the FD method in Table 2. To compute a comparable solution, the FD code requires 129^3 vertices, whereas for the FE method 17^3 are sufficient. Again, this reflects the higher order of the FE streamline diffusion discretization.

When modeling three-dimensional radiative transfer problems, accuracy up to say 1% is not the one and only criterion for the usage of a particular code. Reasonable memory and CPU requirements are at least equally important aspects. Table 2 displays that using a higher order discretization and efficient grid refinement strategies reduce the computational costs. A more efficient solver for the linear system of equations and the application of parallelization techniques are additional tools for saving resources. In Table 3 we compare the memory and CPU requirements of the three-dimensional codes for the two extreme cases of an optical thin ($\tau = 0.4$) and an optical thick ($\tau = 20$) layer. The albedo is $\gamma = 0.80$.

The FD code uses a structured grid (GLOBAL) with 65^3 vertices for $\tau = 0.4$ and 129^3 vertices for $\tau = 20$. The FE code starts with a pre-refined grid with 5^3 vertices and stops the refinement process when the result is almost similar to the analytical result. The computations were performed on the Pentium III 650 MHz cluster of the Interdisciplinary Center for Scientific Computing (IWR) in Heidelberg. The cluster consists of 16 single nodes with 1 GB memory for each processor. If a single processor is used the FE code roughly needs twice as much memory as the FD code. However, the execution time of the FE code is for both cases of τ shorter. Taking advantage of the parallelization feasibility, the memory demand of the FE algorithm can be matched to the requirements of the FD method by employing two processors. This halves the CPU time of the FE code. If we use all nodes of the parallel cluster, the computational costs can be reduced even more. Especially, the execution time of the FE algorithm can be reduced by approximately two orders of magnitude for the optical thick case. The execution time difference between the two codes is even more dramatic, since a finer grid with at least 257^3 vertices is needed for the FD method to obtain results with the same accuracy as achieved with the FE code (see Table 2). Shorter computational times are very important, since the extension to time- or frequency-dependent problems requires the solution of the monochromatic radiative transfer equation for each discrete frequency point. When modeling a line

Table 4. Test III: flux in 10^{-3} of a scattering dominated halo for different optical depth τ , number of ordinates M and refinement steps.

τ	0.1				1.0				10.0			
M	20	80	320	1280	20	80	320	1280	20	80	320	1280
step 0	9.48	9.48	9.48	9.48	9.48	9.48	9.48	9.48	9.48	9.48	9.48	9.48
step 1	8.04	8.04	8.04	8.04	8.04	8.04	8.04	8.04	8.04	8.04	8.04	8.04
step 2	8.19	8.19	8.04	–	8.19	8.22	8.04	–	8.19	8.19	8.19	–
step 3	8.19	8.19	–	–	8.19	8.19	–	–	8.20	8.20	–	–
step 4	8.20	–	–	–	8.20	–	–	–	8.18	8.20	–	–

profile, approximately 50–70 monochromatic problems are to be solved. The problem is even more time-consuming, if frequency redistribution or complex velocity fields are considered. In these cases, the monochromatic radiative transfer equation must be solved several times for each frequency point. This extension of the FE code will be presented in a forthcoming paper. Comparing the memory and CPU requirements of these two codes, one has to bear in mind, that the FE method makes use of all the modern numerical tools described above.

5. Test III: Propagation of radiation in a scattering halo

The setup for the last test problem is as follows: The 3D computational domain ($x = [-1, 1]$, $y = [-1, 1]$, $z = [-1, 1]$) contains a spherically symmetric cloud with radius $r_c = 0.9$ which is illuminated by an extended emission region with radius $r_s = 0.125$. The source term is given by

$$f(r) = \begin{cases} 1 & \text{for } r \leq r_s \\ 0 & \text{for } r > r_s \end{cases}. \quad (19)$$

We chose a Lorentz profile for the radial distribution of the extinction coefficient

$$\chi(r) = \begin{cases} \chi_0/(1 + 100r^2) & \text{for } r \leq r_c \\ \chi(r_c)/100 & \text{for } r > r_c \end{cases}, \quad (20)$$

where the constant χ_0 is calculated from a given optical depth $\tau = \int_0^{r_c} \chi(r) dr$. In this way, the variation of $\chi(r)$ is almost two orders of magnitude between r_s and r_c .

We designed this test problem as a very simple model for Lyman- α halos associated with high-redshift galaxies. Lyman- α halos are diffuse and clumpy emission line regions which are more than ten times larger than the corresponding radio emission (van Ojik et al. 1996, 1997; Steidel et al. 2000). The Lyman- α line profiles suggest a complex velocity field of the material surrounding star-forming galaxies. Hereafter, we intend to model Lyman- α halos with the line transfer version of the finite element code. This test problem will help us to determine the demands on spatial resolution and the required number of ordinates.

Since the Lyman- α line is a resonance line, we assume $\chi = \sigma$ everywhere. In this case, the zeroth moment

equation of the radiative transfer equation reduces to

$$\nabla \cdot \mathbf{F} = \frac{1}{r^2} \frac{d}{dr} r^2 F_r = 4\pi f, \quad (21)$$

where F_r is the radial component of the radiative flux

$$\mathbf{F}(\mathbf{x}) = \int_{S^2} \mathbf{n} \mathcal{I}(\mathbf{x}, \mathbf{n}) d\omega.$$

The solution of Eq. (21) is

$$F_r(r) = \begin{cases} \frac{4}{3}\pi r & \text{for } r \leq r_s \\ \frac{4}{3}\pi \frac{r_s^3}{r^2} & \text{for } r > r_s \end{cases}. \quad (22)$$

Independent on τ , the analytical value for the flux from the computational box in every direction is $F_r(1.0) = 8.18 \times 10^{-3}$.

The numerical calculations start with a pre-refined grid with 16^3 cells. Thus, the size of the cells at step zero is equal to r_s . The grid refinement process is based on the global L^2 -indicators η_{L^2} taken from (12). We perform the calculations for $\tau = 0.1, 1, 10$ and $M = 20, 80, 320, 1280$. Due to limited memory resources, the number of feasible refinement steps critically depends on M . In Table 4 we show for all calculations the values for $F_r(1.0)$ averaged over all directions. In comparison to the analytical result, step zero overestimates and step one underestimates the radiative flux. With refinement step two the emission region is resolved sufficiently well and the relative error of the numerical value for the radiative flux is of order 10^{-3} for all τ and M . This results proves the global flux-conserving property of the FE implementation. On the contrary, the FD code is not flux-conserving. The radiative flux obtained for the same problem with the FD code and a global spatial resolution comparable to step 2 is about 6% smaller than the analytical result.

Table 4 also shows, that refinement step 2 is not sufficient for calculations with $\tau \leq 1.0$ and $M \geq 320$. In these cases, the error estimator prefers to refine cells in the outer regions of the cloud. Since we now know the minimal required resolution for the emission region, this effect can be avoided with a proper local pre-refinement of the central region at step 0.

For the spherically symmetric test case, it is sufficient to use only 20 ordinates to get a very good result for the radiative flux. But the required number of ordinates is

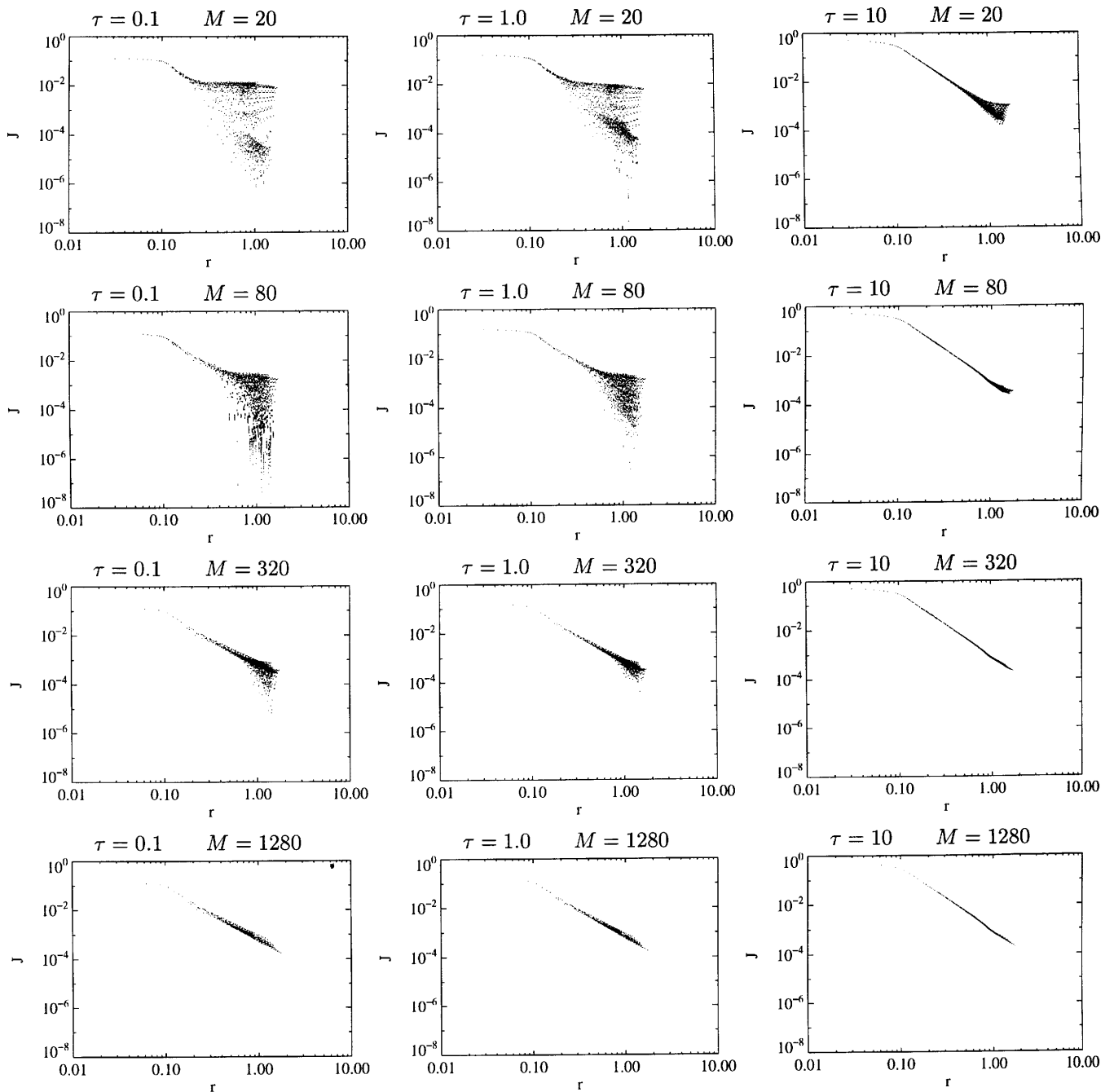


Fig. 5. Test III: dependence of the mean intensity J on the distance from the center r for different τ and M .

completely different, if one is interested in local quantities, for example the radiative flux from isolated regions for comparison with spatially resolved spectroscopic observations or the mean intensity in order to determine a temperature distribution. In Fig. 5 we show the distribution of the mean intensity $J(r)$ at all vertices for step 2, except for $M = 1280$, where we show step 1. Instead of a smooth distribution we find a large scatter of about three orders of magnitudes for small τ and $M = 20$. The scatter is typical for discrete ordinates methods and decreases with higher optical depth and with increasing number of ordinates. At high optical depth it seems sufficient to use

$M = 80$, but at small optical depth one has to use at least $M = 320$.

6. Conclusions

We have presented a finite element code for the solution of the monochromatic radiative transfer equation. The use of adaptively refined grids and ordinate parallelization enables us to compute the radiation fields of complex three-dimensional configurations.

The searchlight beam test showed the capability of the code to resolve discontinuities within a few adaptive steps.

An equidistant grid would require about 100 times more cells to achieve the same resolution. Due to the second order streamline discretization, negative intensities and overshooting occur in the vicinity of discontinuities.

The radiation field escaping from an infinite plane-parallel layer was used as a test problem for the scattering operator. We compared the angular distribution of the intensity with the results of an analytical method and of a simple finite difference technique for various values of optical depth and albedo. The results of the finite element code are in close agreement with the analytical results. No matter how many ordinates are considered, the finite element code needs 2–5 adaptive steps to reach convergence. The finite difference code obtains fairly good results for low optical depth. But for high optical depth and high albedo the analytical results could only be reproduced with extraordinary high grid resolution. This reflects the lower order of the finite difference discretization. Due to efficient grid adaptation and the application of parallelization techniques, the finite element code can obtain more accurate results about 20–70 times faster than the finite difference code. The total memory requirement of the two codes is of the same order of magnitude. The good performance of the finite element code is crucial for the planned extension to 3D frequency-dependent radiative transfer problems.

Finally, we examined the propagation of radiation emitted from an extended object embedded in a scattering, spherically symmetric halo. This three-dimensional problem is the basic setup for future line transfer calculations of Ly α halos associated with high-redshift galaxies. The results show that the total flux is already obtained after a few adaptive steps independent of the number of ordinates. But if one is interested in the spatial distribution of a quantity – for example the mean intensity – a high number of ordinates is required in order to obtain reliable results.

Acknowledgements. We thank Rainer Wehrse, Garii Efimov and Rolf Rannacher for their collaboration. This work was supported by the Deutsche Forschungsgemeinschaft (DFG) within the SFB 359 “Reaktive Strömungen, Diffusion und Transport” and SFB 439 “Galaxien im jungen Universum”. The calculations were mainly performed on the computer clusters of the Institute for Applied Mathematics and the Interdisciplinary Center for Scientific Computing (IWR) in Heidelberg.

Appendix A: Separable representation method

The separable representation method developed by Efimov et al. (1995, 1997) allows to obtain an analytical solution of the plane-parallel radiative transfer equation

$$\mu \frac{d\mathcal{I}}{d\tau} = -\mathcal{I} + \frac{\gamma}{2} \int_{-1}^1 \mathcal{I} d\mu' + (1 - \gamma), \quad (\text{A.1})$$

in the following form

$$\mathcal{I}_{\text{out}}(\mu) = \frac{1 - \gamma}{\gamma} \frac{2}{\sqrt{\mu}} \frac{1}{w^{(0)}(\mu, \mu') - w^{(2)}(\mu, \mu')} \frac{E(\mu'^2)}{\sqrt{\mu'}}, \quad (\text{A.2})$$

where the integration over μ' from 0 to 1 is implied, and the operators $w^{(0)}(\mu, \mu')$, $w^{(2)}(\mu, \mu')$ and the function $E(\mu^2)$ are given by

$$E(\mu^2) = \frac{4\gamma}{\tau} \sum_{m=0}^{\infty} \frac{1}{C_m} \cdot \frac{\mu^2}{1 + y_m^2 \mu^2},$$

$$w^{(0)}(\mu, \mu') = \left(1 + \tanh \left(\frac{\tau}{2\mu} \right) \right) \delta(\mu - \mu'),$$

$$w^{(2)}(\mu, \mu') = \frac{4\gamma}{\tau} \sum_{m=0}^{\infty} \frac{y_m^2}{C_m} \cdot \frac{\mu^{3/2}}{1 + y_m^2 \mu^2} \cdot \frac{\mu'^{3/2}}{1 + y_m^2 \mu'^2},$$

where

$$y_m = \frac{2\pi}{\tau} \left(m + \frac{1}{2} \right), \quad C_m = 1 - \gamma \frac{\arctan(y_m)}{y_m}.$$

γ and τ are albedo and the total optical thickness of the slab, respectively.

Because of the very slow convergence of the sums (several thousands terms must be taken in order to achieve the required accuracy), this solution can hardly be applied. Appropriate approximations of $E(\mu^2)$ and $w^{(2)}(\mu, \mu')$ are necessary. The Stieltjes-Markov theory of orthogonal functions (Perron 1957) allows to represent the infinite sums by means of finite ones in the following way:

$$E(\mu^2) \approx E_N(\mu^2) = \sum_{n=1}^N \frac{a_n \mu^2}{1 + A_n \mu^2},$$

where a_n and A_n depend on τ and γ . The operator $w^{(2)}(\mu, \mu')$ can be expressed through the function $E(\mu^2)$ as

$$w^{(2)}(\mu, \mu') = \mu^{3/2} \cdot \frac{\frac{E(\mu^2)}{\mu^2} - \frac{E(\mu'^2)}{\mu'^2}}{\mu'^2 - \mu^2} \cdot \mu'^{3/2},$$

that leads to the separable representation of $w_N^{(2)}(\mu, \mu')$:

$$w_N^{(2)}(\mu, \mu') \approx \sum_{n=1}^N \frac{\mu^{3/2}}{1 + A_n \mu^2} \cdot a_n A_n \cdot \frac{\mu'^{3/2}}{1 + A_n \mu'^2}.$$

Thus, the final solution becomes

$$\mathcal{I}_{\text{out}}(\mu) = \frac{1 - \gamma}{\gamma} \left(1 + e^{-\tau/\mu} \right) \times \left(\frac{1}{\mu} E_N(\mu^2) + \sum_{nn'} \frac{\mu}{(1 + A_n \mu^2)} S_{nn'} K_{n'} \right), \quad (\text{A.3})$$

where $S_{nn'}$ is a matrix and $K_{n'}$ a vector which are calculated only once for given parameters τ and γ .

Appendix B: Finite difference method

The finite difference code is based on an implicit discretization of the transfer equation, Eq. (1), in cartesian coordinates. The resulting linear system of equations is solved by a combination of recursion and iteration (Stenholm et al. 1991).

When $\mathbf{n} = (n_x, n_y, n_z)$ defines a specific direction and Δx , Δy and Δz are the spatial resolutions, the radiative transfer equation is discretized in a first order upwind scheme

$$\begin{aligned} & \frac{n_x}{\Delta x} (\mathcal{I}_{i,j,k} - \mathcal{I}_{i-1,j,k}) + \frac{n_y}{\Delta y} (\mathcal{I}_{i,j,k} - \mathcal{I}_{i,j-1,k}) \\ & + \frac{n_z}{\Delta z} (\mathcal{I}_{i,j,k} - \mathcal{I}_{i,j,k-1}) = -\chi_{i,j,k} (\mathcal{I}_{i,j,k} - S_{i,j,k}) \end{aligned} \quad (\text{B.1})$$

and recursively solved for

$$\begin{aligned} \mathcal{I}_{i,j,k} = & \left(\frac{n_x}{\Delta x} \mathcal{I}_{i-1,j,k} + \frac{n_y}{\Delta y} \mathcal{I}_{i,j-1,k} + \frac{n_z}{\Delta z} \mathcal{I}_{i,j,k-1} \right. \\ & \left. + \chi_{i,j,k} S_{i,j,k} \right) / \left(\frac{n_x}{\Delta x} + \frac{n_y}{\Delta y} + \frac{n_z}{\Delta z} + \chi_{i,j,k} \right). \end{aligned} \quad (\text{B.2})$$

With extinction coefficient $\chi = \kappa + \sigma$ and albedo $\gamma = \sigma/\chi$ the source function $S(\mathbf{x})$ is given by

$$S(\mathbf{x}) = \frac{f(\mathbf{x})}{\chi(\mathbf{x})} + \gamma(\mathbf{x}) \frac{1}{4\pi} \int_{S^2} P(\mathbf{n}', \mathbf{n}) I(\mathbf{x}, \mathbf{n}') d\omega', \quad (\text{B.3})$$

and determined in an iterative procedure using a simultaneous overrelaxation (SOR) method. The iteration is continued until $\max |(J^n - J^{n-1})/J^n| < \epsilon$, where J^n is the mean intensity of the n th iteration step and ϵ a very small number. This relatively simple solution method is stable for all optical depths and guarantees positive intensities everywhere. But it gets into difficulties with steep gradients and shows slow convergence in case of high albedo.

References

- Becker, R., & Rannacher, R. 2001, *Acta Numerica*, in press, <http://www.iwr.uni-heidelberg.de/sfb359/Preprints2001.html>
- Bicknell, G. V., Sutherland, R. S., van Breugel, W. J. M., et al. 2000, *ApJ*, 540, 678
- Dautray, R., & Lions, J.-L. 1993, *Math. Anal. and Num. Meth. for Sci. and Tech.*, Vol. 6 (Springer)
- Dykema, P. G., Klein, R. I., & Castor, J. I. 1996, *ApJ*, 457, 892
- Efimov, G. V., Waldenfels, W., von, & Wehrse, R. 1995, *J. Quant. Spectrosc. Rad. Transf.*, 53, 59
- Efimov, G. V., Waldenfels, W., von, & Wehrse, R. 1997, *J. Quant. Spectrosc. Rad. Transf.*, 58, 355
- Eriksson, K., & Johnson, C. 1993, *Math. Comput.*, 60/201, 167
- Eriksson, K., Estep, D., Hansbo, P., & Johnson, C. 1995, *Acta Numerica 1995* (Cambridge University Press), 105
- Fabiani Benedicho, P., Trujillo Bueno, J., & Auer, L. 1997, *A&A*, 324, 161
- Folini, D. 1998, Ph.D. Thesis, ETH Zürich, <http://www.astro.phys.ethz.ch/staff/folini/folini.html>
- Führer, C., & Kanschat, G. 1997, *Computing*, 58(4), 317
- Haas, M., Müller, S. A. H., Chini, R., et al. 2000, *A&A*, 354, 453
- Haber, S. 1970, *SIAM Rev.*, 12, 481
- Juvela, M. 1998, *A&A*, 322, 943
- Kanschat, G. 1996, Ph.D. Thesis, Univ. of Heidelberg, <http://www.iwr.uni-heidelberg.de/sfb359/Preprints1996.html>
- Kanschat, G. 2000, in *Parallel Solution of PDE*, IMA Vol. in Math. and its Appl. 120, ed. P. Björstad, & M. Luskin (New York, Springer)
- Kunasz, P., & Auer, L. H. 1988, *J. Quant. Spectrosc. Rad. Transf.*, 39, 67
- Meier, D. L. 1999, *ApJ*, 518, 788
- Miranda, L. F., Guerrero, M. A., & Torrelles, J. M. 1999, *ApJ*, 117, 1421
- Papkalla, R. 1995, *A&A*, 295, 551
- Park, Y.-S., & Hong, S. S. 1998, *ApJ*, 494, 605
- Perron, O. 1957, *Die Lehre von den Kettenbrüchen*, Band II (B.G. Teubner Verlagsgesellschaft, Stuttgart)
- Sleijpen, G. L. G., & Fokkema, D. R. 1993, *Electronic Transactions on Numerical Analysis* 1, 11 <http://etna.mcs.kent.edu>
- Steidel, C. C., Adelberger, K. L., Sharply, A. E., et al. 2000, *ApJ*, 532, 170
- Stenholm, L. G., Störzer, H., & Wehrse, R. 1991, *J. Quant. Spectrosc. Rad. Transf.*, 45, 47
- Turek, S. 1993, *Imp. Comp. Sci. Eng.*, 5, 201
- van Ojik, R., Röttgering, H. J. A., Carilli, C. L., et al. 1996, *A&A*, 313, 25
- van Ojik, R., Röttgering, H. J. A., Miley, G. K., & Hunstead, R. W. 1997, *A&A*, 317, 358
- Väth, H. M. 1994, *A&A*, 284, 319
- Városi, F., & Dwek, E. 1999, *ApJ*, 523, 265
- Verfürth, R. 1996, *A Review of A Posteriori Error Estimation and Adaptive Mesh Refinement Techniques* (Wiley-Teubner, New York-Stuttgart)
- Wehrse, R., Meinköhn, E., & Kanschat, G. 1999, in ed. P. Stee, *Forum du GRETA, Transfert de Rayonnement en Astrophysique*, CNRS-INSU-ASPS, 82 <http://www.iwr.uni-heidelberg.de/sfb359/Preprints1999.html>
- Wiseman, J. J., & Ho, P. T. P. 1998, *ApJ*, 502, 676
- Witt, A. N., & Gordon, K. D. 1996, *ApJ*, 463, 681
- Wolf, S., Fischer, O., & Pfau, W. 1998, *A&A*, 340, 103
- Wood, K., & Reynolds, R. J. 1999, *ApJ*, 525, 799
- Zinnecker, H., Krabbe, A., McCaughrean, M. J., et al. 1999, *A&A*, 352, L73