The “ray” ray tracing package

Don Wells*

March 6, 1998

Abstract

The “ray” package and program rayMain trace sets of rays representing wavefronts through systems of rotationally-symmetric aspheric optical elements. The starting sets of rays can represent either plane or spherical wavefronts, with feedhorn tapering. The optical elements can be de-centered and/or tilted conic sections (planes, spheres, ellipsoids, paraboloids, hyperboloids) with additional superimposed radially-symmetric aspheric terms, and they can be mirrors as well as refracting surfaces. Both foci and nearly-plane wavefronts can be analyzed.

Contents

1 The ray library 2
1.1 Tracing Bundles of Rays through a System: rayTrace() 3
1.2 Defining the Surfaces of the System: rayAddSurface() 4
1.3 Generating Bundles of Rays: rayGenerator() 5
1.4 Analyzing foci and wavefronts: rayGetFoci() and rayGetPlanes() 8
1.5 Printing & plotting from lists (rayPrtPlanes(), rayPrtPS(), etc) 9

2 The ray-tracing main program rayMain 10
2.1 Example 1: Multiple plane waves imaged to GBT Gregorian foci 12
2.2 Example 2: Spherical waves from Q-band horns produce plane waves 15

A Appendices 18
A.1 Include file ray.h 18
A.2 List structures – initializing, inserting and deleting 20
A.3 Features expected to be added in future releases of ray 21

Bibliography 22

*mailto:dwells@nrao.edu

1Source code of the ray package is available at ftp://fits.cv.nrao.edu/pub/dwells_ray.tar.gz (140 KB). The ray package is supplied separately from other GBT code because it is potentially useful in a wide range of other applications.
1 The ray library

An unusual feature of this ray tracing package is that it is available as the library libray.a of ray-tracing routines in ANSI-C. The individual functions of this library can be called by custom application programs coded in C; a good example is the program srFocusTrackingTable1 which is used to verify that the GBT focus tracking algorithm [Wel98a] will produce high quality imaging as the GBT distorts due to gravity as a function of elevation. In order to do this, srFocusTrackingTable1 must call the structural model and BFP functions [WK95b, WK95a] in addition to the ray functions; no conventional standalone ray tracing program could do this.

For conventional ray tracing applications a user can use program rayMain (Section 2, p.10) to invoke the individual ray tracing functions from a script. Program rayMain is analogous to commercial standalone ray tracing programs. For example, in the analysis of the imaging properties of the GBT subreflector [Wel98b], the following terse script named srEllipsoidCase2.in commands rayMain to trace 29 spherical waves of 13 rays each through the off-axis ellipsoidal mirror:

```
Digits 4 0.00001
System GBT_Subreflector
  rayAddSurface "subreflector" -0.1331 10 0.528 0 0 -1
    [4.91667,0,0.0] \
    "cone"@[0,0,0]@[1,0,0] 1
  rayAddSurface "prime_plane" 0 0 0 0 1
    [0,0,0]@[0,0,0] 1
  rayGenerator "spherical"[-11,0,0]@[-0.302411,0.312393,0]0.261677 \n    0.02,3,6,2 0.261677,-13.0 2,13 "bundle"
  rayTrace
  rayGetFoci
  rayPrtFoci
  rayPltSystem
  rayPltPS [12,12] 0.15 [-11.0,0.0,0.] orthographic srEllipsoidCase2b.ps
  Quit
```

The commands shown in the example above define two surfaces, generate and trace rays representing a set of spherical wavefronts originating from a grid of 29 points near the second ellipsoid focal point, determine the locations of the foci near the first focal point, print the results to standard output and send a Postscript plot to file srEllipsoidCase2b.ps. These operations, which invoke the individual functions in libray.a, are documented in this report.

Copyright: The source files of library ray and program rayMain are being made available under a GNU "copyleft" license. Each of the source files begins with a copyright notice which is suppressed in the source file listings of this memo.

Acknowledgements: It is a pleasure to thank R. Fisher, R. Norrod and P. Napier for a number of valuable discussions and R.G. Tull (Univ. Texas) for supplying a listing of the original Fortran version of the tracing algorithm.

---

2 See file "GNU_GPL_2.txt" in the ray distribution kit.
3 "Copyright (C) 1995 Associated Universities, Inc. Washington DC, USA. This program is free software; you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation; either version 2 of the License, or (at your option) any later version. This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details. You should have received a copy of the GNU General Public License along with this program; if not, write to the Free Software Foundation, Inc., 675 Mass Ave, Cambridge, MA 02139, USA. Correspondence concerning GBT software should be addressed as follows: GBT Operations, National Radio Astronomy Observatory, P. O. Box 2, Green Bank, WV 24944-0002 USA".
1.1 Tracing Bundles of Rays through a System: rayTrace()

Function rayTrace() traces a list of lists of rays through a list of surfaces, returning a new list of lists of rays. It uses Feder’s algorithm [Fed51] for tracing a ray through a centered rotationally-symmetric aspheric surface,4 augmented by extensions [AS52] for de-centering and tilting of the surface. Feder’s technique for tracing aspheric surfaces has been extended to use the numerical eccentricity [SM62, p.675], so that arbitrary conic sections can be traced, including superimposed radially-symmetric polynomial deviations from the conic sections.5

The surfaces to be traced are conic sections plus polynomial terms, and are computed by rayTrace() as

\[
x = \frac{cs^2}{1 + \sqrt{1 - c^2s^2(1 - \varepsilon^2)}} + A_2s^2 + A_4s^4
\]

where \( s \) is the radius of ray intercept \( (s^2 = y^2 + z^2) \), \( c \) is the osculating curvature at the vertex \( (c = 1/r) \), \( \varepsilon \) is the numerical eccentricity of the conic and the \( A_i \) are the radially-symmetric polynomial coefficients.

Spheres are specified by setting \( \varepsilon = 0 \) and curvature \( c = 1/r \). The ray package follows Feder’s convention [Fed51] that the X axis is the main optical axis of a system, and the sign convention for \( c \) is that it is positive if the center-of-curvature is to the right of the vertex \( (x_{\text{con}} > x_{\text{vertex}}) \). Planes are specified as spheres with curvature \( c = 0 \) \( (r = \infty) \). Paraboloids with focal length \( f \) are specified as \( \varepsilon = 1 \) and \( c = 1/2f \). Ellipsoids with major axis \( a \) and minor axis \( b \) are specified as \( \varepsilon = \sqrt{a^2 - b^2}/a \) and \( c = a/b^2 \). The reader can verify by inspection that equation (1) computes a paraboloid \( (\varepsilon = 1) \) of focal length \( f = r/2 \) with \( A_2 = A_4 = 0 \) as \( x_{\text{paraboloid}} = s^2/2f \). A plane with \( \varepsilon = 0 \), \( A_2 = A_4 = 0 \) and \( c = 0 \) is computed as \( x_{\text{plane}} = 0 \). A simple sphere with radius \( r \) is computed as \( x_{\text{sphere}} = s^2/r(1 + \sqrt{1 - s^2/r^2}) \), where the sign of \( x_{\text{sphere}} \) varies with the sign of \( c = 1/r \) (i.e. positive if \( x_{\text{con}} > x_{\text{vertex}} \)).

The concept of the extensions for de-centering and tilting is that the coordinates of both surfaces and rays are defined relative to the same origin, and rays are transformed to local coordinates just before tracing them through a surface, and then are transformed back to the original coordinate system afterward.

Argument Segments is a list to which items of type Segment (see ray.h in Appendix A.1) are appended by the ray tracing algorithm for plotting by function rayPlotPS() (section 1.5, p.9); see Figure 3 for an example of plotted ray segments.

The local variable declarations within rayTrace() include temporary variables used by Feder in [Fed51]. \( \mathbf{M} \) is the vector from the vertex of the surface and perpendicular to the incident ray and having its terminus on the ray. \( \mathbf{M}_{1.2} = M_2^2 \), the square of the length of this vector, and \( \mathbf{M}_{ix} \) is its X-component \( M_{ix} \). \( \mathbf{xi} \_1 \_p \) \( (\xi_1) \) is the cosine of the angle of incidence, \( \mathbf{xi} \_1 \_p \) \( (\xi_1') \) is the cosine of the angle of refraction. \( \mathbf{L} \) is the length of the ray intercepted between the two surfaces.

---

4 If toric (non-rotationally-symmetric) surfaces, such as the tertiary which was proposed [Str91] for the GBT, need to be traced, this algorithm will need further extensions [SM62, p.675].

5 The basic algorithm used in rayTrace() was developed by the author during the summer of 1967 as part of the design work for the Cooke spectrograph of the McDonald Observatory 2.7-meter telescope [Tul69, Tul72], under the direction of Robert G. Tul (Univ. Texas), the designer of the spectrograph. The algorithm was embodied in a program named COUDE. In 1993, the author contacted Tul in order to obtain a 26-year-old line printer listing of COUDE, and then translated the code from Fortran-66 to ANSI-C while typing in the first version of function rayTrace(). Program COUDE was used to analyze a spectrograph which has optical complexity comparable to the GBT: the collimator is an off-axis paraboloid, the camera is a Schmidt with the diffraction grating at the entrance pupil, and a tilted Schmidt corrector in the collimated beam is used to generate the necessary elliptically-symmetric Schmidt-type wavefront correction to compensate the spherical aberration of the camera.

6 Equation (1) implies that a paraboloid can also be specified as \( c \equiv 0, \varepsilon \equiv 0, A_2 \equiv 1/2f \).
Function rayTrace() performs many of its calculations using macros defined in mathVectorMatrix.h, which is included by ray.h. Application programs which call rayTrace() and other functions of the library may also find these macros useful. The macros expand operations on three-element vectors and 3 x 3 matrices as inline scalar operations (no for-loops are used), in hopes that scalar optimization in compilers will improve performance.

1.2 Defining the Surfaces of the System: rayAddSurface()

```c
struct Node *rayAddSurface() /* returns ptr to new node */
struct Node *list, /* list of surfaces */
char surfname[], /* descriptive string */
double curvature, /* 1/r */
double epsilon, /* eccentricity of conic */
double A_2, /* deformation terms */
double A_4,
double index_ratio, /* N/N_1, -1 means mirror */
double S[], /* XYZ of vertex */
double E[], /* Euler angles of vertex tilt */
enum VignetteType
    vignet_type, /* VIGN_CYLINDER | VIGN_CONE */
double VO[], /* vignette origin XYZ */
double VW[], /* vignette direction cosines */
double VR) /* radius, linear | radians */
```

Argument epsilon is the numerical eccentricity ε; curvature is the inverse of the radius of the sphere which osculates to the surface at the vertex point. The index_ratio argument is the ratio of the indices of refraction μ_1 = N/N_1, where N is the index to the left of the surface and N_1 is the index to the right. For a focal plane or aperture stop, which neither refracts nor reflects, set μ_1 = 1; for a mirror, set μ_1 = -1.

The position of the vertex of the surface is given by argument S[], relative to an arbitrary origin chosen by the user. For example, in Table 3 the GBT’s prime focal point has been chosen as the origin, and so the vertex of the paraboloidal primary is at X = -60 meters. The vertex point of a plane is undefined, and the vertex point of a sphere can be varied by varying the tilt E[]; the user should supply whatever values seem reasonable in such cases.

Although the comment in the function prototype above asserts that argument E[] is the “Euler angles of vertex tilt” of the surface, this is not correct — the algorithm currently implemented in rayTrace() uses E[0] to rotate about the X axis, E[1] to rotate about the Y axis and E[2] to rotate about the Z axis, in that order. For example, in Table 3 (p.12) the tilt of the GBT’s subreflector is specified as E[2], rotation about Z, with value -0.097214 radians (−5.57°), and the vertex is specified to be (4.8935, -0.4772, 0.0) meters. It is likely that future versions of rayTrace() will implement a true Euler angle convention for E[].

The vignetting arguments vignet_type, VO[], VW[] and VR are provided to define the edge of the optical element, however the current implementation of function rayTrace() does not use them. When the vignetting logic is added to rayTrace() it will be necessary to implement an output list for the rays which are vignette; this will support “spillover” calculations in radio telescopes. The author’s present thinking is that this will be done by adding a pointer to that list of bundles of rays to the Surface struct; tools will be provided to sum the fluxes of rays in the various lists. The concept of the vignetting arguments is that two types of aperture stops will be supported: surfaces delimited by cylinders and surfaces delimited by cones. The GBT paraboloidal primary is an example of the cylindrical case and its ellipsoidal secondary is

---

7The ray package does not define the unit used for positions of surfaces, but it does provide arguments tolerance and d (number of digits after decimal point) in functions rayTrace(), rayPrtSystem(), etc. Users can consider the positions and tolerances to be inches, centimeters or meters as desired. Angle arguments are always radians.
an example of the cone case. The cylinder or cone is defined by a point, a vector and a radius. An example
of the use of these arguments in the cylinder case is shown in Table 3, for the case of the "main_mirror",
where an aperture stop of 50 meters radius is specified to be located at (-60,-54,0) meters, normal to the
X-axis of the system. In practice an optical system like the GBT can be traced with reasonable accuracy by
the present implementation of rayTrace(), because function rayGenerator() can produce bundles of rays
which will suffer negligible vignetting.

1.3 Generating Bundles of Rays: rayGenerator()

<table>
<thead>
<tr>
<th>steps</th>
<th>1-D</th>
<th>2-D</th>
<th>3-D</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>13</td>
<td>33</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>29</td>
<td>123</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>49</td>
<td>257</td>
</tr>
<tr>
<td>5</td>
<td>11</td>
<td>81</td>
<td>515</td>
</tr>
<tr>
<td>6</td>
<td>13</td>
<td>113</td>
<td>925</td>
</tr>
<tr>
<td>7</td>
<td>15</td>
<td>149</td>
<td>1419</td>
</tr>
<tr>
<td>8</td>
<td>17</td>
<td>197</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>19</td>
<td>253</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>21</td>
<td>317</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>23</td>
<td>377</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>25</td>
<td>441</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>27</td>
<td>529</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Items (rays,cases) as function of number of “steps”

The concept is that a “bundle” of rays can be generated and appended to an existing list of bundles. Each
bundle is intended to represent an individual wavefront case; sets of bundles can be traced, analyzed, printed
and plotted as sets, with single calls to functions, because the sets are represented as lists of lists of rays and
lists of computed results.

Two types of bundles can be generated, plane and spherical; character string argument wave_type[] specifies
this. A bundle of rays will originate at \texttt{wave point} and will proceed in \texttt{wave direction}. Actually, \texttt{rayGenerator()} appends a set of ray bundles with a single call; the arguments and algorithms provided are intended to support typical analytical problems in optical systems, such as tracing plane wavefronts with the angle off-axis being stepped, or tracing spherical wavefronts with the position of the feedhorn being stepped relative to the on-axis focal point. Argument \texttt{axis mask} specifies which axes will be stepped. Argument \texttt{case steps} specifies the number of steps \pm off-axis, and \texttt{case step} is the step size. For each step a bundle of rays will be generated of size \texttt{wave radius} and with \texttt{ray steps} rays \pm about the axis of the bundle. Function \texttt{rayGenerator()} can produce vast numbers of rays from innocuous-looking values of these few arguments! For example, in Table 3 (p. 12) the four arguments “0.0001388, 2, 2, 4” generate five sets of 49 rays each, 245 rays in total. This is a plane wave case, and so \texttt{case step}=0.0001388 is interpreted as 0.14 mr (29 arcsec). The \texttt{case steps} argument is 2, so five cases (-2, -1, 0, +1, +2) will be computed for each active axis. The \texttt{axis mask} is 2, which we see is the Y axis only, so only five cases will be generated. The \texttt{ray steps} argument is 4; the number of rays to be generated for this value is given as 49 in the “2-D” column of Table 1.\footnote{The number of items is simply }2n+1\text{ for the 1-D case, but for 2-D and 3-D it is limited by the \texttt{wave radius} argument; Figure 1 shows how the 49 rays fit into the circular area.} With a few keystrokes we could change this argument to 6, and thereby generate 113 rays, more than doubling the sampling of the wavefront. In practice, when debugging the setup of a new problem, it is convenient to start with a small value for \texttt{ray steps} so that the calculation will go rapidly, and then increase the value to improve accuracy when the problem appears to be operating correctly.

Arguments \texttt{taper angle} and \texttt{taper db} implement radio-style beam tapering (in optical systems aperture

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{ray_pattern}
\caption{Ray pattern in 100 meter aperture (\texttt{ray steps} = 4 \rightarrow 49 rays)}
\end{figure}
Figure 2: Power-weighting of rays ($q = 43.17, \theta_H = 15^\circ$)

tapering is called “apodization”). The relative E-field pattern of the feedhorn in decibels is

$$E(\theta) = 20 \log(\cos \theta)^q,$$

where $\theta$ is the angle off-axis [RSCWL81]. Parameter $q$ is specified implicitly by giving the value of $E(\theta)$ at the edge of the beam. If $\theta_H$ is the subtended half-angle of the subreflector, we have

$$E(\theta_H) = 20q \log \cos \theta_H,$$

and we can solve for the exponent $q$ as

$$q = \frac{E(\theta_H)}{20 \log \cos \theta_H}.$$

For example, in [Nor90b] we have $\theta_H = 15^\circ$ and $E(\theta_H) = -12$ dB, which implies $q = 43.17$. The relative power of the rays in the generated wave is computed as

$$I(r) = \cos^2\left(\frac{r}{r_H}\right),$$

where $r_H$ is the radius in the wavefront corresponding to angle $\theta_H$, and this intensity (power) is used in the `ray` library as a weighting factor for analyzing foci and wavefronts. For example, a taper of $-12$ dB is a weight of 0.0630 for marginal rays (see Figure 2). An un-apodized (un-tapered, uniformly weighted) optical system can be traced by setting the `taper_db` parameter to zero.

The three color arguments will assign color codes to individual rays or to individual bundles of rays. In Table 3 these three arguments are “2,13 "bundle"”; "bundle" means that the rays of the five cases (angle off-axis stepping by 29 arcsec) will have five different colors, starting with color code “2” (i.e., codes 2, 3, 4, 5 and 6 will be used). The colors assigned to the codes are arbitrary.
1.4 Analyzing foci and wavefronts: rayGetFoci() and rayGetPlanes()

Function `rayGetFoci()` uses a least-squares regression to solve for the XYZ coordinates of a focal point about which a nearly spherical wavefront converges to the smallest RMS volume. The regression also produces the pathlength of the wavefront for the focal point. The algorithm then selects a delta pathlength which is large (50x) compared to the 3-D RMS of the ray points on the wavefront about the focal point solution. It subtracts this delta pathlength from the pathlength produced by the regression and computes the XYZ coordinates of this pathlength for each ray, and then computes the difference between the delta pathlength and the Euclidean distance between these XYZ coordinates and the XYZ position of the focal point; i.e., it forms the radial residual between the ray points and a sphere centered on the focal point. The RMS of these residuals is the RMS phase deviation of the nearly spherical wavefront. The rays are weighted by their intensity (radio power), as discussed in Section 1.3. This calculation of the focal point location ignores wave interference effects, which will be significant at low frequencies (< 5GHz?) but are less and less important as frequency is increased.

Function `rayGetFoci()` adds segments to the segment list to draw a ±3σ 3-D box (sides 6σx, 6σy, 6σz) centered on each focal point that it computes (see Figure 4 for examples).

For the analysis of nearly-planar wavefronts, we need to fit the pathlengths of rays across the aperture with Zernike polynomials [BW59], and identify the terms which represent the primary (Seidel) aberrations. The argument `nord` to function `mathZernike()` in `rayGetPlanes()` is

\[ \text{int} \ nord[1]=4,3,2,-1; \]  

which will command `mathZernike()` to compute the nine terms shown in Table 2 (cf. Table 6).

During the fitting process the y, z coordinates of each ray intercepting the plane are converted to polar coordinates \( r, \theta \) relative to \( yz \text{ axis} \). The formulas are

\[
\begin{align*}
\rho &= \sqrt{\Delta y^2 + \Delta z^2}, \\
\cos \theta &= \Delta y / \rho,
\end{align*}
\]

\[ (6) \]

\[ (7) \]

The function prototype is:

```
int mathZernike()  
    /* returns #coeffs in u[] & ui[] */  
    double r,  
    /* radius in range 0--1.0 */  
    double cost,  
    /* cosine of theta */  
    double sint,  
    /* sine of theta */  
    int nord[],  
    /* order limits */  
    double u[],  
    /* Zernike coefficients return here */  
    int ui[]  
    /* indicies of coeffs return here */
```

It is likely that the details of this C function will be discussed in a future GBT Memo, as a part of the active surface servo development effort. A similar set of polynomials was used earlier in the GBT project [Nel90] to analyze the shapes of the GBT panels.
<table>
<thead>
<tr>
<th>index</th>
<th>Zernike term</th>
<th>n</th>
<th>m</th>
<th>Function</th>
<th>Seidel Abberation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>( A_{0000} )</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>Zero Point (Pathlength)</td>
</tr>
<tr>
<td>1</td>
<td>( A_{2000} )</td>
<td>2</td>
<td>0</td>
<td>( 2\rho^2 - 1 )</td>
<td>Curvature (Defocus)</td>
</tr>
<tr>
<td>2</td>
<td>( A_{4000} )</td>
<td>4</td>
<td>0</td>
<td>( 6\rho^4 - 6\rho^2 + 1 )</td>
<td>Spherical Abberation</td>
</tr>
<tr>
<td>3</td>
<td>( A_{1010} )</td>
<td>1</td>
<td>1</td>
<td>( \rho \cos \theta )</td>
<td>Tilt (Pointing Error)</td>
</tr>
<tr>
<td>4</td>
<td>( A_{1011} )</td>
<td>1</td>
<td>1</td>
<td>( \rho \sin \theta )</td>
<td>([\text{sine term}])</td>
</tr>
<tr>
<td>5</td>
<td>( A_{3010} )</td>
<td>3</td>
<td>1</td>
<td>( (3\rho^3 - 2\rho) \cos \theta )</td>
<td>Coma</td>
</tr>
<tr>
<td>6</td>
<td>( A_{3011} )</td>
<td>3</td>
<td>1</td>
<td>( (3\rho^3 - 2\rho) \sin \theta )</td>
<td>([\text{sine term}])</td>
</tr>
<tr>
<td>7</td>
<td>( A_{2020} )</td>
<td>2</td>
<td>2</td>
<td>( \rho^2 \cos 2\theta )</td>
<td>Astigmatism</td>
</tr>
<tr>
<td>8</td>
<td>( A_{2021} )</td>
<td>2</td>
<td>2</td>
<td>( \rho^2 \sin 2\theta )</td>
<td>([\text{sine term}])</td>
</tr>
</tbody>
</table>

Table 2: Seidel \((3^r)^d\) order terms fitted to wavefronts by \texttt{rayGetPlanes()} 

\[
\sin \theta = \Delta z/r \quad (8)
\]

and

\[
\rho = r/r_{\text{max}}, \quad (9)
\]

where \(r_{\text{max}}\) is argument \texttt{yz\_radius}. 

If argument \texttt{print\_mode} has value "\texttt{verbose}" , \texttt{rayGetPlanes()} will print a listing of the least squares solution for the coefficients, the RMS of fit and the correlation matrix, \textit{for the first bundle only}. This is mainly intended as a debug aid, because \texttt{rayPrtPlanes()} produces a more readable listing of the coefficients, and does it for all of the bundles which were traced.

1.5 Printing & plotting from lists (\texttt{rayPrtPlanes()}, \texttt{rayPrtPS()}, \textit{etc})

```c
void rayPrtSystem(struct Node *list, /* list of surfaces */
int d) /* digits after decimal point */
```

A good example of the output produced by this command is shown in the middle part of Table 4 (p.14). The curvatures are printed in “g” format with six significant figures, the eccentricities with five figures, the \(A_i\) with two figures and index ratios \(\mu_i\) with five figures. The positions \(S[i]\) are printed with \(d\) digits after the decimal point, while the \(E[i]\) angles are always printed with 5 decimal places \((10\mu r \approx 2\ \text{arcsec} \ \text{precision})\). The primary purpose of this listing is to facilitate debugging of optical prescriptions.

```c
void rayPrtFoci (struct Node *foci_list, /* list of Focus structs */
int d) /* digits after decimal point */
```

A good example of the output produced by this command is shown in the lower part of Table 4 (p.14). The focal point solution is printed for each bundle of rays that was traced. The “n” column is the number of rays in the \(i^{\text{th}}\) bundle. Columns \((x_c,y_c,z_c)\) are the positions of the phase centers of the converging nearly-spherical wavefronts, \(l_c\) is the path length to the phase center, \((x_s,y_s,z_s)\) are the mean errors \((\sigma_x,\sigma_y,\sigma_z)\) of the phase center coordinates and \(l_s\) is the mean error of the pathlengths \(l_c\). These values are printed with \(d\) digits precision.

```c
void rayPrtPlanes (struct Node *planes_list, /* list of Plane structs */
int d) /* digits after decimal point */
```

A good example of the output produced by this command is shown in the lower part of Table 6 (p.17), where the values of the Seidel terms are printed for each bundle of rays which was traced. The quantity being fitted
is the pathlength to the final aperture stop; the mean pathlength of the wavefront is given as \( A_{\text{ave}} \). Only those coefficients which are non-zero to \( d \) digits precision are printed—negligible values are printed as blanks. The values are printed with \( d \) digits after the decimal point. \texttt{rayPrtPlanes()} prints the values of terms \( A_{1011} \) and \( A_{101} \) (Wavefront Tilt, Pointing Error) in \textit{milliradians} rather than the wavefront amplitude; it does this by multiplying the amplitudes by 1000.0/\( yz\_\text{radius} \) (the latter is an argument supplied to \texttt{rayGetPlanes()}) which is transferred in \texttt{struct Plane}). \texttt{rayPrtPlanes()} prints a warning message if the RMS wavefront residual after fitting the \( 3^{rd} \)-order Zernike terms is greater than zero to \( d \) digits precision; \textit{i.e.}, the residual represents the \( 5^{th} \) order terms of the wavefront.\footnote{It is shown in [Wel96a] that the RMS wavefront residual (the combined \( 5^{th} \) order terms) of the GBT Gregorian optical system for \( E = 0^\circ \) and for \( E = 90^\circ \) is about 15\( \mu \), whereas the \( 3^{rd} \)-order terms have typical amplitudes of a millimeter, 50\( \times \) or more larger than the RMS residual. \textit{The GBT is a \( 3^{rd} \)-order system to high accuracy.}}

\begin{verbatim}
void rayPrtBundles(struct Node *set, /* list of lists of rays */
                  int d) /* digits after decimal point */

This is a debug utility.

void rayPrtSegments(struct Node *list, /* list of ray segments */
                    int d) /* digits after decimal point */

This is a debug utility.

void rayPltSystem( /* returns lines in segments list */
                   struct Node *system, /* list of optical elements */
                   struct Node *segments) /* list of line segments */

The current implementation of this function merely plots the axes of the optical system; a future version will walk through the list of surfaces \texttt{system} and will plot a mesh of lines for each of them, delimited by their respective vignetting parameters. Lines plotted are appended to the list of \texttt{segments}.

\begin{verbatim}
void rayPltPS( /* returns PS in psname [] */
              struct Node *segments, /* list of line segments */
              double height_cm, /* height of PS in centimeters */
              double width_cm, /* width of PS in centimeters */
              double width, /* width in units of System */
              double to_point[], /* XYZ of point in center of PS */
              char plt_mode[], /* "Orthographic","Perspective" */
              char psname[]) /* Postscript output file */

A view of the list of \texttt{segments} is computed as the Postscript file \texttt{psname []}. The size of the Postscript image in centimeters is \texttt{(width\_cm, height\_cm)}, and argument \texttt{width} specifies the width in the units of the optical system. The \texttt{to\_point[]} will appear in the center of the image. Although \texttt{plt\_mode} is provided, it is not implemented in the current version of \texttt{rayPltPS()}; the implementation always computes "Orthographic" projections with +\( X \) to the right and +\( Y \) pointing up, looking toward \( -Z \).

2 The ray-tracing main program \texttt{rayMain}

Program \texttt{rayMain} contains a simple command language parser which recognizes function names, parses argument values and passes the arguments to the functions. The algorithm of the parser is the following steps:
1. Read next input line and append it to the command line being accumulated
2. If the last character of the input line is the character “\”, remove it and control goes back to step 1.
3. If the first character of the command line is “#”, it is a comment; it is echoed to the standard output, the command line is cleared, and control goes back to step 1.
4. The command line is scanned for occurrences of any of the characters “,”(,)“@” and tab; they are changed to blanks.
5. The command line is scanned for occurrences of two or more blanks; they are changed to single blanks.
6. If the first character of the command line is a blank, the blank is deleted.
7. If the last character of the command line is a blank, the blank is deleted.
8. If the command line is empty (all blank), control goes back to step 1.
9. The transformed command line is echoed to standard output; if longer than 65 characters it is printed on more than one line in segments of 65 characters.
10. The command code is parsed as the initial string of characters up to the first blank. If the code is “Quit”, rayMain exits.
11. The command code is tested against the list of valid codes: Digits, SetName, System, rayAddSurface, rayGenerator, rayGetFoci, rayGetPlanes, rayPlotPS, rayPlotSystem, rayPrtBundles, rayPrtFoci, rayPrtPlanes, rayPrtSegments, rayPrtSystem and rayTrace. If the code is recognized, the blank-delimited arguments are parsed by scanf() formats in which the successive format codes are appropriate for the data types of the arguments of the functions to be invoked. Then the specified function is called with the arguments. After the function returns, control goes back to step 1.

The “Digits” command wants two arguments: the d value and the tolerance value to be used in calls to the functions. The print commands, whose functions take argument “d”, will get the number of digits from the “Digits” command, and rayTrace will use the tolerance.

The “System” command wants one argument, the name to be given to the optical system which will be defined by the rayAddSurface() calls.

The “SetName” command wants one argument, the name to be given to the bundles of rays which will be generated by subsequent rayGenerator() calls. A default name choice will be used if this command is not supplied. A good example of the use of this command is shown in five occurrences in the lower part of Table 5 (p.16), and the result of the operation is shown in the bundle.name column of the rayPrtPlanes output in the lower part of Table 6 (p.17).

All string arguments must be single strings—they cannot contain embedded blanks. Underscores can be used to separate words; good examples of this technique are the system name and three of the surface names in the command script shown in Table 3 (p.12). The double-quotes around the surface name main.mirror in the script are irrelevant—the command parser strips them, and would regard main and mirror as two arguments rather than one if a blank were present instead of the underscore. The quotes are used as punctuation in these cases because the author judged that they make the script more readable.

Other punctuation symbols and the backslashes are allowed to make scripts more readable, but they are irrelevant to command interpretation because rayMain strips them at steps 2 and 4 above. The author uses the punctuation characters to group arguments into logical, mnemonic patterns.

The function arguments for lists of bundles of rays, lists of surface elements and lists of line segments are not supplied with the commands to rayMain. These lists are created and managed by rayMain automatically. There is no command which will delete these lists and re-initialize them—rayMain is intended to used to execute one problem at a time. To calculate two problems, run rayMain twice with separate scripts.
2.1 Example 1: Multiple plane waves imaged to GBT Gregorian foci

The script shown in Table 3 defines the GBT surfaces for $E = 44^\circ$ (the rigging angle) in the order in which an incident plane wave from the sky encounters them; the subreflector geometry parameters are taken from [Nor90a]. The comments in this script give the details of calculation of a number of key quantities of interest which were used to compute the arguments to the ray tracing functions. The prime plane and greg plane surfaces are defined in order to make these important locations explicit in the optical prescription—they have $\mu_i = 1$, and so do not deflect the rays, even though function rayTrace does do the full tracing calculation on them. The origin is at the prime focal point, and the unit is meters, so the vertex of the main mirror is at $x = -60$ and the prime plane is at $x = 0$. The arguments to rayGenerator create five plane waves starting at the origin ($x = 0$), with tilts stepping by about 29 arcsec and with each wave represented as 49 rays. Because the main mirror is to the left of the origin, the wave moves to left initially.

Figure 3 shows how the five sets of rays proceed through the GBT (because rayPltSystem does not yet plot the surfaces, we see only the points where the rays change direction). The width of this scene is 100 meters. Figure 4 shows the region around the Gregorian focus greatly enlarged (width 100 mm, 1:1 scale). We see that the on-axis wave focuses perfectly but that the off-axis waves focus to caustic surfaces.
Figure 3: Overview of rays traced through the GBT (rayTestCase1a.ps)

Figure 4: Foci at the Gregorian focal plane (rayTestCase1b.ps)
Table 4: rayMain output for Foci example
2.2 Example 2: Spherical waves from Q-band horns produce plane waves

In this example we trace cones of rays (spherical waves) originating from the phase centers of the four horns proposed for the 40-50 GHz Q-band receiver system [WS95] and we analyze the resulting nearly-plane wavefronts at the GBT aperture plane. Figure 5 is a view of the Gregorian focal point with width 100 mm, reproduced at 1:1 scale (these feedhorns are tiny!). The upper and lower ray cones are actually two cones projected one on top of the other. The central cone of rays is the on-axis case which is computed as a reference. The input script shown in Table 5 defines the GBT at the rigging angle in the same manner as in the first example, but in opposite order, and it defines the prime plane as the last intercept point for the rays. Because the arrangement of the four feedhorns does not correspond to any combination of the rayGenerator() arguments, the five ray bundles are computed as five separate executions of rayGenerator with different bundle names.

The output produced in this example calculation is in Table 6. The interesting feature is the rayPrtPlanes output at the bottom of the table. The five nearly-plane wavefronts have been analyzed, and their 3rd-order Zernike coefficients, the Seidel aberrations, are tabulated. Most of the cells of the tabular listing are blank, meaning that those coefficients are zero in the 4th decimal place (note Digits 4 at the top of Table 6). The A_{0000} values are the pathlengths for the four horns in meters, differing by only 100µ. The non-zero A_{1010} and A_{1011} coefficients are the tilts of the four wavefronts produced by the off-axis feedhorns. The A_{1011} (sine) terms are wavefront tilt in azimuth, out of the GBT’s plane of symmetry, and they are all the same value, ±0.1388. The unit for these tilt values is milliradians, as was mentioned in Section 1.5 in the discussion of rayPrtPlanes(). So, four beams are off-axis in azimuth by ±0.1388 mr = ±28.63 arcsec. The A_{1010} tilts are in elevation, and they differ by about one percent—the scale factor between linear focal plane displacement and angle on the sky in elevation is varying significantly. Because only the tilt terms are non-blank in this listing, and rayPrtPlanes has not warned us about significant RMS wavefront errors (5th-order terms), we can conclude that these four 100 meter-wide wavefronts are planes to better than 100µ.
Table 5: Input file for Planes example
Table 6: `rayMain` output for Planes example
A  Appendices

A.1  Include file ray.h

The function prototypes which are at the end of ray.h have been suppressed in the following listing:

/* ray.h -- Include file for the "ray" (ray-tracing) package.
   D.Wells, MRAO-CY
   97-05-03 many name changes
   97-05-30 removed changes to rayGenerator() */

[GNU GPL copyright notice omitted]

#ifndef RAY_H
#define RAY_H

#include <stdio.h>
#include <stdlib.h>
#include <string.h>
#include "mathVectorMatrix.h"

#define TRUE 1
#define FALSE 0
#define MAXMEMAX 21
#define FORMMAX 100
#define AMNMAX 9

enum VignetteType {VIGN_CYLINDER, VIGN_CONE};
enum ColorType {COLOR_BUNDLE, COLORRAY};
enum PlotType {ORTHOGRAPHIC, ORTHOGRAPHIC_X, PERSPECTIVE};

struct Surface {
  char name[MAXMEMAX]; /* user-supplied descriptive string */
  double c_1; /* curvature c=1/r; positive if center of curvature is
                 to right of surface. c_1 for following surface */
  double eps; /* numerical eccentricity of the conic section. */
  double a_2; /* coefficients of radially-symmetric */
  double a_4; /* aspheric deformation series. */
  double mu_1; /* $\mu_1/\equiv \mu/18$, $\mu_1$ is index to right
                of following surface */
  double S[3]; /* absolute vector to vertex of this surface. */
  double E[3]; /* tilt of next vertex, 3 Euler Angles (see
                "General Ray-Tracing Procedure",
                G.H.Spencer and M.V.R.K.Murty, JOSA 52,
                pp.672-678, (June 1962)). */

  enum VignetteType vign_type;
  double vign_origin[3];
  double vign_vector[3];
  double vign_radius;
};

struct Ray {
  double T[3]; /* $T \equiv (x, y, z)$ is the vector from the
                 vertex of the first surface to the point of
                 incidence of the ray on this surface. */
  double Q[3]; /* $Q \equiv (X, Y, Z)$ is the unit vector along
the ray to the right of the first surface. X,Y,Z
are direction cosines; ray path is $T + 1 * Q$ */

double direction;  /* experimental, +1->+X, -1->-X */
double Length;  /* cumulative path length as ray traverses system */
double Intensity;  /* same as power for radio */
int ColorCode;  /* color code to use when plotting this ray */
}

struct Segment {
    double T1[3];  /* x,y,z position of starting point of ray */
    double T2[3];  /* x,y,z position of ending point of ray */
    int ColorCode;  /* color code to use when plotting this ray */
}

struct Focus {
    char name[NAMEMAX];  /* name_string for the ray bundle */
    int n;  /* number of rays in the bundle */
    double xyz[3];  /* position of focus */
    double xyzs[3];  /* std_err of focus position */
    double l1;  /* mean pathlength to focus */
    double l2;  /* spherical wave rms wrt focus */
}

struct Plane {
    char name[NAMEMAX];  /* name_string for the ray bundle */
    int n;  /* number of rays in the bundle */
    int nu;  /* number of Seidel terms */
    double Amni[AMNIMAX];  /* Seidel terms for pathlengths */
    double Asig[AMNIMAX];  /* Sigmas for Amni[AMNIMAX] */
    int mni[AMNIMAX];  /* indices of Seidel terms */
    double sig;  /* rms of (pathlength-sum(Amni[])) */
    double yz_radius;  /* Zernike radius in rayGetPlanes() */
}

#elif SEIDEL

struct Seidel_item {
    int mni;  /* A_mni index */
    char *name;  /* names of Seidel aberrations */
}

static struct Seidel_item Seidel_list[] = {
    {0000, "Zero_Pt"},
    {2000, "Defoc"},
    {4000, "Sph_Ab"},
    {1010, "Tilt"},
    {1011, "<sin>"},
    {3010, "Coma"},
    {3011, "<sin>"},
    {2020, "Astigmatism"},
    {2021, "<sin>"}
};

const static int Seidel_n = sizeof(Seidel_list) / sizeof(struct Seidel_item);
#endif

struct Node {
    void *item;  /* pointer to an "item" struct */
    struct Node *next;  /* pointer to next item in list */
};
A.2 List structures – initializing, inserting and deleting

The ray package deals with lists of objects, not with fixed-dimension arrays. A simple list processing package is sufficient. This one is based on algorithms given by Sedgewick [Sed90, p.20].

The lists are composed of “nodes”. Each Node structure consists of a pointer to another node plus a pointer to some data object (see definition in ray.h, Appendix A.1). A list consists of a chain of nodes, with the head node containing a pointer to the second node, which points to the third node, ..., an so forth until the tail node, whose “next” pointer points to itself. The head node always exists, and the data item associated with it is a descriptive character string supplied when the list is initialized by calling listInitialize(). The tail node does not point to a data object. The minimum (empty) list is a head node with its descriptive string and a tail node. The six function prototypes shown below are in ray.h.

```c
struct Node *listInitialize(char *name) /* descriptive string for list */
```

A pointer to the head of the new list is returned.

```c
struct Node *listInsertAfter ( /* returns ptr to new node */
void *newitem, /* ptr to new item struct */
struct Node *t) /* ptr to node in a list */
```

This function will malloc() a new instance of Node, set the item pointer of the new node to the pointer newitem, copy the pointer to the following node from the prior node into the new node and change the pointer of the prior node to point to the new node.

```c
struct Node *listAppend ( /* returns ptr to new node */
void *newitem, /* ptr to new item struct */
struct Node *list) /* ptr to existing list */
```

This function reads the list until it finds the node just before the tail node, then it calls listInsertAfter(). The terseness of the idioms for searching and manipulating these list structures is exemplified by this slightly simplified version of the function:

```c
struct Node *listAppend ( /* returns ptr to new node */
void *newitem, /* ptr to new item struct */
struct Node *list) { /* ptr to existing list */
struct Node *t;
t = list;
while (t->next->next != t->next) { t = t->next; }
return(listInsertAfter (newitem, t));
}
```

11 Use of list techniques in ray tracing is not merely a question of programming convenience and notational elegance — it is very nearly a matter of necessity, because the number of output rays produced for a given number of input rays is not predictable a priori. First, when rays are vignette by one of the optical elements of a system, they need not be traced through the remaining elements, and they need not (probably should not) even appear in the output list of rays. (Of course, in radio telescopes vignetting is “spillover”, and vignette radio rays are interesting in their own right.) Second, when a ray is traced through a refractive surface, it becomes two rays, one reflected and one refracted. Each of these rays can then encounter another refractive surface (e.g., by an internal reflection) at which it too can become two rays. In general, each individual input ray passing through a refractive system produces an infinite cascade of output rays. In practice, many of the rays in the cascade will be vigneted by the optical elements and others can be discarded when their intensity falls below some specified sensitivity level. The cascading of refracted and reflected rays has actually been implemented and used for tracing special radomes, in which rays can even be reflected from the inside of one part of the radome and then refracted through an entirely different part (in [PL06] these cases are called “flash lobes”). The ray package does not currently support tracing of rays reflected from refractive surfaces because the author only needs to trace reflective surfaces for the GBT analysis.
```c
int listDeleteNext(struct Node *t) /* ptr to node in a list */
```

This function does `free()` on the data object to which the node points, cuts the node out of the chain of pointers and then does `free()` on the `Node` structure itself.

```c
void listDeleteList (struct Node *list) /* list of items */
```

This function walks through the list doing `listDeleteNext()` on each node, then does `free()` on the tail node, the descriptive character string and the head node.

```c
void listDeleteListList (struct Node *list) /* list of lists of items */
```

The `item` pointer of a `Node` can contain a pointer to the head node of another list. This list-of-lists concept is very powerful, and is heavily used in the `ray` package. This function walks through the list of lists and calls `listDeleteList()` and `listDeleteNext()` to `free()` all storage associated with each node. It finishes the operation by deleting the tail node, descriptive string and head node of the list of lists.

### A.3 Features expected to be added in future releases of ray

- **rayTrace**
  - Implement vignetting functions. Is ray intercept point `inside` the CONE or CYLINDER? If test fails, append ray to bundle-set of vignette rays, marked by the vignetteing surface. This will support spillover calculations in radio telescopes.
  - Implement polarization of rays, including refraction/reflection calculations at surfaces. This will support cross-polarization calculations in radio telescopes.
  - Change to Euler angle rotations for surfaces

- **rayGenerator**
  - Implement “generalized spiral points” [SK97] for ray origins on spherical waves, and probably for plane waves too (use equal-area mapping from sphere to plane). This equal-area distribution will facilitate numerical integrations, such as for cross-polarization calculations in radio telescopes. The quasi-hexagonal pattern of the generalized spiral points will fit circular aperture stops better than does the square pattern generated by the present algorithm.

- **rayPltSystem**
  - Implement plotting of vignetting-bound surfaces, with axes.

- **rayPltPS**
  - Implement “Perspective” mode; may need to add `from_point` and `up_vector` arguments.
References


[Nel90] Jerry Nelson. Panel surfaces for the Green Bank Telescope. GBT Memo 31, National Radio Astronomy Observatory, January 1990. This memo uses a set of polynomials analogous to the Zernike polynomials [BW59] to analyze the GBT panels, and concludes that the panels can be approximated by quadratic surfaces. This memo is also Keck Observatory Technical Note 291.

[Nor90a] Roger D. Norrod. Increase in distance between large and small subreflectors. GBT Memo 40, National Radio Astronomy Observatory, March 1990. This memo quotes the formula for the \( \alpha, \beta \) and \( \theta_T \) angles of the subreflector and Gregorian feedhorn geometry. It discusses the cross-polarization justification for \( \beta \). Table 1 gives the official values for the parameters of the GBT geometry.

[Nor90b] Roger D. Norrod. Spillover blockage in the GBT design. GBT Memo 41, National Radio Astronomy Observatory, March 1990. This memo discusses feed spillover, spillover efficiency (feedhorn power patterns) and shields that could be deployed to reduce scattering.


[Wel98a] Don Wells. GBT Gregorian focus tracking in C. GBT Memo [tbd], NRAO, [in preparation] 1998. The GBT Gregorian subreflector, an off-axis portion of an ellipsoid, and the feedroom with the feedhorn move relative to the prime focal point of the main paraboloidal mirror. The subreflector must be maneuvered relative to the prime focal point and the feedhorn to maintain nearly stigmatic imaging (maximum gain, minimum sidelobes). An algorithm is described which computes the required actuator motions as a function of elevation, with no focus error, spherical aberration or coma, and with only 0.5 mm of residual astigmatism. The focus tracking algorithm is expressed in C.

[Wel98b] Don Wells. Imaging properties of the GBT subreflector. GBT Memo [tbd], NRAO, [in preparation] 1998. The subreflector images points in the neighborhood of its first focus onto points in the neighborhood of its second focus. nearly-stigmatic imaging, can be obtained for a variety of tilts and displacements. results computed by ray tracing. minimum phase error, if ellipsoid is tilted slightly. fitted with polynomials. expressed in C.

[WK95a] Don Wells and Lee King. GBT Best-Fitting Paraboloid [BFP] in C. GBT Memo 131, NRAO, June 1995. Abstract: The gravitational displacements of the GBT actuators have been fitted with a paraboloid. The parameters of the paraboloid for various elevations have been fitted with polynomials and expressed as C code which computes the parameters of this best-fitting-paraboloid [BFP] as a function of elevation. The BFP will be used by the control software modules for the pointing, focus-tracking and active-surface subsystems of the GBT. We give a description of this C-code version of the BFP and two examples of its application to practical problems. We also give a function in C which fetches node data from the structural model and transforms it to a coordinate system tied to the BFP. The predicted gravitational term of the GBT's traditional pointing model and the predicted prime focus focus-tracking formula of the GBT are given. See ftp://fits.cv.nrao.edu/pub/gbt_dwell_doc.tar.gz for the current revision of this memo (131.2 as of 1997-06-23).

[WK95b] Don Wells and Lee King. The GBT Tipping-Structure Model in C. GBT Memo 124, NRAO, March 1995. Abstract: The finite element model of the GBT tipping structure has been translated into executable code expressed in the C language, so that it can be used by the control software modules for the pointing, focus-tracking, quadrant detector, active-surface and laser-rangefinder subsystems of the GBT. We give a description of this C-code version of the tipping structure model and two examples of its application to practical problems. See ftp://fits.cv.nrao.edu/pub/gbt_dwell_doc.tar.gz for the current revision of this memo (124.3 as of 1997-06-23).