Procedure for simulating divergent-light halos

Lars Gislén

Divergent-light halos are halos produced by light from nearby light sources, like street lamps being scattered by small crystals of ice floating in the air. The use of "brute-force" Monte Carlo methods to simulate such halos is extremely inefficient, as most scattered rays will not hit the eye of the observer. I present a new procedure for Monte Carlo simulations of divergent-light halos. This procedure uses rotational symmetries to make a selected sampling of events that greatly improves the computational efficiency of the algorithm. We can typically generate a simulated halo display in minutes using a personal computer, several orders of magnitude more rapid than a simple brute-force method. The algorithm can also optionally generate three-dimensional pictures of divergent-light halo displays. © 2003 Optical Society of America

OCIS codes: 010.1290, 010.2940, 010.3920.

1. Introduction

Halo phenomena are generated when light, normally from the Sun or the moon, is reflected and/or refracted by small crystals of ice floating in the atmosphere. Such halos have been simulated extensively in the past with use of raytracing Monte Carlo techniques; one recent compilation of such simulations is by Tape.[1] Such Monte Carlo simulations for parallel-light halos have been extremely important to the development of both theory and observation in this field of optics. With these halos, the light source is the distance to the light source (the Sun, the Moon) is in practice infinite, and we can assume that the light rays hitting the crystals are parallel; these halo are *parallel-light halos.* This greatly simplifies the simulation procedure: The incident ray has a fixed direction determined by the altitude of the light source, and we only have to raytrace the incident ray through a suitably oriented crystal and determine the direction and intensity/probability of the scattered ray. The observer will perceive a spot of light in a direction that is the reverse of that of the scattered ray. This completely determines the halo display. In other words, the simulation has essentially only to handle the *directions* of the scattered rays. When light from a nearby light source is scattered by ice crystals in the atmosphere, the rays from the source will be divergent and will generate a *divergent-light halo.*

L. Gislén (larsg@thep.lu.se) is with the Department of Theoretical Physics, Lund University, Lund, Sweden.

Received 31 March 2003; revised manuscript received 30 June 2003.

0003-6935/03/336559-05$15.00/0

© 2003 Optical Society of America

Divergent-light halos can be observed in regions with a cold winter climate in connection with artificial light sources, such as street lamps. Their geometry often differs from that of parallel-light halos, and many of them give a three-dimensional impression. Unfortunately there are rather few good field observations of divergent-light halos. Being able to simulate and predict such halos in an easy way would certainly stimulate interest in making more and well-documented observations of these halo phenomena.

Divergent-light halos include an extra geometrical constraint that relates the location of the light source, the observer, and the scattering crystal. This constraint means that the precise location in space of the crystal is important and leads to computational complications. In Fig. 1 the difference of fixed angle scattering of parallel and divergent light by two crystals is shown.

Consider crystals where the ray is scattered a fixed angle. In Fig. 1 (a) we have incident parallel light, and all crystals along the line of sight are equivalent. In Fig. 1 (b), where we have a nearby source of light L crystals along a line of sight are not equivalent. The crystal at C_2 will not scatter light to the eye of the observer. To do so we have to locate the crystal at C_2'. A brute-force method of simulating a divergent-light halo would be to select a ray with random direction from the light source, randomly select a crystal along that ray path, raytrace the ray through the crystal, and then investigate if the scattered ray happens to pass through the pupil of the eye of the observer. Evidently in almost all cases this would not happen, and the event would be wasted. Thus this simulation method is extremely inefficient and typically needs computer execution times of the order of days, even on a mainframe computer. We are not aware of any pub-

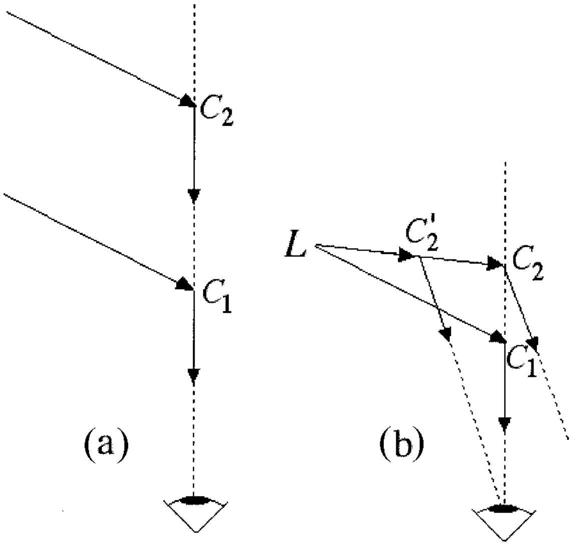

Fig. 1. (a) Fixed angle scattering of parallel light by two crystals. (b) The same scattering for a divergent-light source.

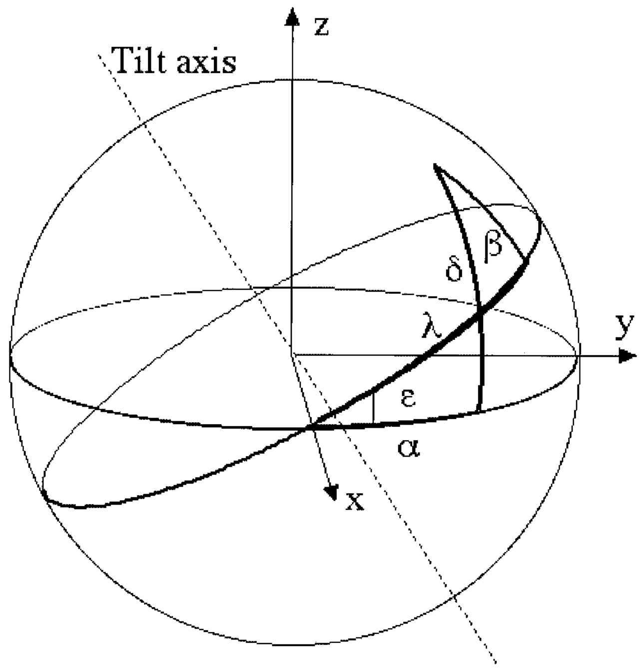

Fig. 2. Coordinate systems.

lished results using such brute-force methods. There are a few published attempts to describe divergent-light halos analytically, finding surfaces or regions in space of the crystals that generate specific halo arcs.[2–5]

The simulation algorithm, for which we present the theoretical background below, uses the fact that for almost all types of divergent-light halos there are two rotational symmetries involved. The appearance of the halo will not be affected if we rotate the observer around a vertical axis through the light source (assuming the emission from the light-source to be isotropic). Secondly, we study halos where the crystals are given random rotations around a symmetry axis; in the case of hexagonal ice crystals, this is the axis perpendicular to the hexagonal end faces. (We do not treat here the case of Parry-oriented crystals, although they can be included as a special case.) This means that we are free to rotate the crystal and with it the incident and scattered ray around this axis. Exploiting these symmetries makes it possible for the simulation algorithm to select in advance those crystals that are located in space such that a ray scattered by them will hit the eye of the observer. The increase in computational efficiency is quite dramatic; a typical divergent-light halo can be simulated in a matter of minutes with this method. The mathematical background to the new algorithm is developed in Section 2.

2. Theory

We represent the light ray emitted from the source by a vector. The scattered ray is represented by a second vector. The geometrical constraint is that the sum of these vectors has a constant length that is equal to the distance observer–source, which we normalize to 1. Assume that we have an ice crystal with a vertical symmetry axis. We tilt this axis in a mathematically positive direction around the horizontal x axis. Further assume that we have generated a ray with random direction $\mathbf{k}_1$ from the source that has been raytraced through the crystal and exits with direction $\mathbf{k}_2$. Both of these vectors are unit vectors. The initial and scattered ray can then be represented mathematically by

$$\begin{aligned} \mathbf{r}_1 &= \mathbf{k}_1 t_1, \quad t_1 > 0, \\ \mathbf{r}_2 &= \mathbf{k}_2 t_2, \quad t_2 > 0, \end{aligned} \tag{1}$$

where t_1 and t_2 are the *parametric distances,* t_1 being the distance source–crystal, t_2 being the distance crystal–observer, both measured in units of the distance source–observer.

Consider a sphere with unit radius, centered at the origin where we assume the source is located. We imagine two coordinate systems on the sphere, one equatorial (α, δ) and one ecliptical (λ, β). The vernal point of the two coordinate systems is on the x axis. The tilt can then be represented by the obliquity ε of the ecliptic, the crystal symmetry axis being perpendicular to the ecliptic plane (Fig. 2).

The vector from the source to a point where the scattered ray hits the surface of the sphere is

$$\mathbf{k} = \mathbf{k}_1 t_1 + \mathbf{k}_2 t_2, \ |\mathbf{k}|^2 = 1. \tag{2}$$

The hitpoint is assumed to have equatorial coordinates (α, δ) given by the relations

$$\begin{aligned} k_x &= \cos\delta \cos\alpha, \\ k_y &= \cos\delta \sin\alpha, \\ k_z &= \sin\delta. \end{aligned} \tag{3}$$

The hitpoint generally will not coincide with the position of the observer. Assume that the observer sits at $(0, \delta_0)$.

We now invoke the two rotational symmetries.

We are allowed to perform arbitrary rotations of $\mathbf{k}_1$ and $\mathbf{k}_2$ around two axes: a vertical axis and an axis parallel to the crystal symmetry axis, the tilt axis. We will try to determine these rotations and the lengths of the vectors $\mathbf{r}_1 = \mathbf{k}_1 t_1$ and $\mathbf{r}_2 = \mathbf{k}_2 t_2$ such that the scattered ray ends up in the eye of the observer. If we rotate the observer around the z axis, his latitude will vary in the interval $[\beta_{0,\max}, \beta_{0,\min}]$ where

$$\beta_{0,\max} = \min[\delta_0 + \varepsilon, \pi - (\delta_0 + \varepsilon)],$$
$$\beta_{0,\min} = \max[\delta_0 - \varepsilon, \pi - (\delta_0 - \varepsilon)]. \quad (4)$$

When we vary t_1 and t_2, the hitpoint will trace out a great circle on the sphere as the vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ span a plane through the center of the sphere. However, because of the geometrical constraint condition $|\mathbf{k}|^2 = 1$, t_1 and t_2 are not independent, and only part of the great circle will be covered. We have

$$\begin{aligned}|\mathbf{k}|^2 &= 1\\ &= (\mathbf{k}_1 t_1 + \mathbf{k}_2 t_2)^2\\ &= \mathbf{k}_1^2 t_1^2 + 2\mathbf{k}_1\cdot\mathbf{k}_2 t_1 t_2 + \mathbf{k}_2^2 t_2^2\\ &= t_1^2 + 2\mathbf{k}_1\cdot\mathbf{k}_2 t_1 t_2 + t_2^2, \end{aligned} \quad (5)$$

which gives

$$t_2 = -a t_1 \pm [1 - t_1^2(1 - a^2)]^{1/2}, \quad a = \mathbf{k}_1\cdot\mathbf{k}_2. \quad (6)$$

If $a \geq 0$ we have only the solution $t_2 = -a t_1 + [1 - t_1^2 (1 - a^2)]^{1/2}$. If $a < 0$ we have for $t_1 > 1$ both solutions in Eq. (6). Further we have

$$t_1 \in [0, t_{1,\max}],\ t_{1,\max} = \begin{cases} 1 & \text{for } a \geq 0 \\ 1/(1 - a^2)^{1/2} & \text{for } a < 0 \end{cases} \quad (7)$$

In the t_1, t_2 plane the solution traces out part of a tilted ellipse, centered at the origin and passing through the points (0, 1) and (1, 0). The ellipse degenerates into a straight line for $a = 1$ (see Fig. 3).

To get a single-valued relation between the variables, we introduce two new quantities q_1 and q_2 by

$$q_1 = t_1 - t_2, \quad q_2 = t_1 + t_2, \quad (8)$$

where

$$q_1 \in [-1, 1] \text{ and } q_2 = \left(\frac{2}{1 + a} - q_1^2 \frac{1 - a}{1 + a}\right)^{1/2}. \quad (9)$$

The hitpoint has ecliptic coordinates (λ, β) given by

$$\sin\beta = k_z \cos\varepsilon - k_y \sin\varepsilon,$$
$$\cos\beta \sin\lambda = k_z \sin\varepsilon + k_y \cos\varepsilon,$$
$$\cos\beta \cos\lambda = k_x. \quad (10)$$

Using Eqs. (2) and (8) we have

$$\mathbf{k} = \frac{1}{2}(\mathbf{k}_1 - \mathbf{k}_2) q_1 + \frac{1}{2}(\mathbf{k}_1 + \mathbf{k}_2) q_2. \quad (11)$$

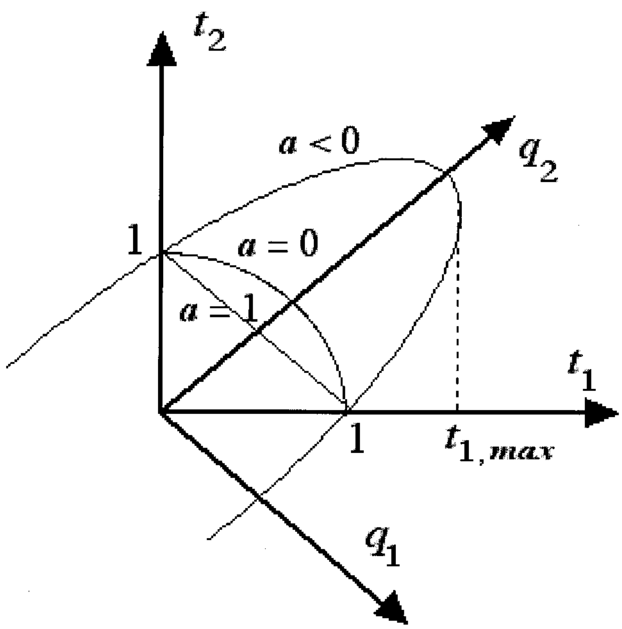

Fig. 3. q–t-plane.

The endpoints at the great circle are given by inserting $\mathbf{t} = (0, 1)$ and $\mathbf{t} = (t_{1,\max}, 0)$ corresponding to $\mathbf{q} = (-1, 1)$ and $\mathbf{q} = (1, 1)$, respectively. Then Eqs. (10) and (11) will give us a starting value β_{start} and ending value β_{end} for the latitude of the great circle. If β is a monotonic function of λ, we know that the great circle latitude is confined in the interval $[\beta_{\text{start}}, \beta_{\text{end}}]$. However, if β is not monotonic, there will be an extremal value of β given by

$$\frac{d\beta}{dq_1} = 0. \quad (12)$$

From Eqs. (10) and (11) we have

$$\begin{aligned}\sin\beta = {}& \frac{1}{2}[(k_{1z} + k_{2z})\cos\varepsilon - (k_{1y} + k_{2y})\sin\varepsilon] q_2\\ &+ \frac{1}{2}[(k_{1z} - k_{2z})\cos\varepsilon - (k_{1y} - k_{2y})\sin\varepsilon] q_1,\end{aligned} \quad (13)$$

or using Eq. (9),

$$\sin\beta = A(C - D q_1^2)^{1/2} + B q_1 \quad (14)$$

with

$$A = \frac{1}{2}[(k_{1z} + k_{2z})\cos\varepsilon - (k_{1y} + k_{2y})\sin\varepsilon],$$
$$B = \frac{1}{2}[(k_{1z} - k_{2z})\cos\varepsilon - (k_{1y} - k_{2y})\sin\varepsilon],$$
$$C = \frac{2}{1 + a},$$
$$D = \frac{1 - a}{1 + a}. \quad (15)$$

Then from Eq. (12) the extremal value of β is realized by

$$\hat{q}_1 = \pm\left[\frac{C}{D(1 + A^2D/B^2)}\right]^{1/2}. \quad (16)$$

There is an extremal value within the allowed part of the great circle if $|\hat{q}_1| \le 1$. $\hat{q}_1$ has the same sign as AB. From Eq. (14) we get the extremal value in latitude $\hat{\beta} = \beta(\hat{q}_1)$. Then the interval in latitude in which the hitpoint can vary will be the largest of the intervals $[\hat{\beta}, \beta_{\text{start}}]$ and $[\hat{\beta}, \beta_{\text{end}}]$.

We now find the intersection set Ω_β between the intervals of the latitude of the observer and the latitudes of the great circle. This can, in the nonmonotonic case, result in two disjoint intervals of β. If the intersection is zero, it is not possible to make the hitpoint coincide with the observer. Finding this intersection and implementing it in program code is an involved process. The intersection region in latitude β will correspond to a certain region,

$$\Omega_q = \{q_1;\ \beta(q_1) \in \Omega_\beta\}, \quad (17)$$

possibly disjoint, in q_1. The equation for solving q_1 as a function of β is Eq. (12), with solution

$$q_1 = \frac{B \sin\beta}{A^2D + B^2} \pm \left[\frac{B^2 \sin^2\beta}{(A^2D + B^2)^2} + \frac{A^2C - \sin^2\beta}{A^2D + B^2}\right]^{1/2}, \quad |q_1| \le 1. \quad (18)$$

If there is an extremum latitude within the allowed part of the great circle, there will be two solutions; otherwise, only one. We have to check the validity of the solutions by inserting them back in Eq. (14).

We now want to select a scattering crystal anywhere in space along the line

$$\mathbf{r} = \mathbf{k}_1 t_1, \qquad t_1 \in [0,\ t_{1,\max}], \quad (19)$$

and such that q_1 belongs to the allowed region Ω_q. However, we have a problem because the relation between q_1 and t_1 is not single valued for $\mathbf{k}_1 \cdot \mathbf{k}_2 = a < 0$. If we trace the possible corresponding regions in t_1, we get something quite complicated. We avoid this problem by using a single-valued variable $t_1{}'$ such that (Fig. 4)

$$t_1{}' = \begin{cases} t_1 & \text{for } q_1 \in [-1, Q] \\ 2\mathrm{t}_{1,\max} - t_1 & \text{for } q_1 \in [Q, 1], \end{cases} \quad (20)$$

where

$$Q = q_1(t_{1,\max}) = \begin{cases} 1 & \text{for } a \ge 0, \\ \left(\dfrac{1 + \mathrm{a}}{1 - \mathrm{a}}\right)^{1/2} & \text{for } a < 0. \end{cases} \quad (21)$$

This will fold out the region $t_1 \in [1, t_{1,\max}]$ and remove the ambiguity. The region Ω_q will now map one to one on a region

$$\Omega_t = \{t_1{}';\ q_1(t_1{}') \in \Omega_q\}. \quad (22)$$

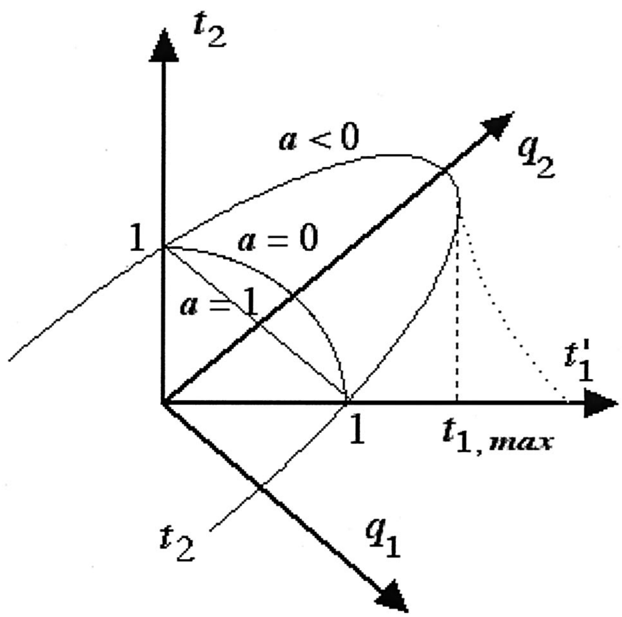

Fig. 4. Definition of $t_1{}'$.

We generate a uniform random number t' in Ω_t. Actually, the probability should be proportional to the square of the distance of the crystal from the origin, the number of crystals in a spherical shell being proportional to the square of this distance, but as the irradiance of the crystal is inversely proportional to the square of the same distance, we can combine the two and generate numbers uniformly, which is a great simplification. t' will correspond to a unique point $\mathbf{q}$ in the q plane. This will in turn by Eqs. (10) and (11) correspond to a unique hitpoint (λ, β).

We now find a rotation α of the observer around the z axis and a rotation Λ around the tilt axis such that the hitpoint coincides with the observer. The observer will then have equatorial coordinates (α, δ_0), and the hitpoint will have ecliptic coordinates $(\lambda + \Lambda, \beta)$. The condition is that these points coincide. Using well-known relations between an equatorial and an ecliptic system[6] we have the equations:

$$\sin\beta = \sin\delta_0 \cos\varepsilon - \cos\delta_0 \sin\alpha \sin\varepsilon,$$

$$\cos\beta \sin(\lambda + \Lambda) = \sin\delta_0 \sin\varepsilon + \cos\delta_0 \sin\alpha \cos\varepsilon,$$

$$\cos\beta \cos(\lambda + \Lambda) = \cos\delta_0 \cos\alpha. \quad (23)$$

We can solve for sin α from the first of these equations, where it is now possible to choose randomly between two possible values for the angle, α_1 and $\alpha_2 = \pi - \alpha_1$. Inserting the chosen value in the remaining equations gives a corresponding rotation angle Λ.

Finally we perform two rotations on $\mathbf{k}_2$, the direction of the scattered ray: first, a rotation Λ around the tilt axis, then a rotation −α around the z axis. The rotated vector will be the direction of a scattered ray that hits the observer.

3. Results

We have implemented the described routine in a computer program unit that can be run on a personal computer together with a standard raytracing program. The source code in PASCAL and in C++ of this program unit is freely available on the web at http://www.thep.lu.se/~larsg/ for anyone to use and modify. Simulation runs show that the program can rapidly and efficiently reproduce field observations of divergent-light halos. We have published these results in a separate paper.[7] There we make a detailed analysis of specific divergent-light halo phenomena and compare simulations using the present algorithm with field observations. In that paper we also present an "atlas" of divergent-light halos for different source elevations and crystal orientations and model the transformation of a halo from a divergent-light halo to a parallel-light halo as the distance between the observer and the light source increases. The agreement with existing observations is very good. This means that the simulation algorithm that we present in this paper together with a standard halo raytracing computer program can be used as an efficient research tool to predict the appearance of a divergent-light halo for different viewing conditions. This gives new and exciting possibilities in this field of optics.

4. Additional Remarks

Some divergent-light halo displays are strikingly three-dimensional. It is interesting that the vector $-\mathbf{k}_2 t_2$ will be the parametric location in space of the scattering crystal relative to the observer. Saving a file with such vectors makes it possible to reconstruct the appearance of a halo display in three dimensions using for instance stereo pictures. One example of such a simulation is given in fig. 5.

In most field situations, the ground will cut off rays coming from crystals that would be located below ground level. This can easily be implemented by excluding rays with $k_{1z} t_1$ less than some fixed value.

For tilts that are exactly zero, the program will not be able to find the rotations, as then only their direct sum can be determined. We avoid this problem by giving the tilt a very small nonzero value in this case.

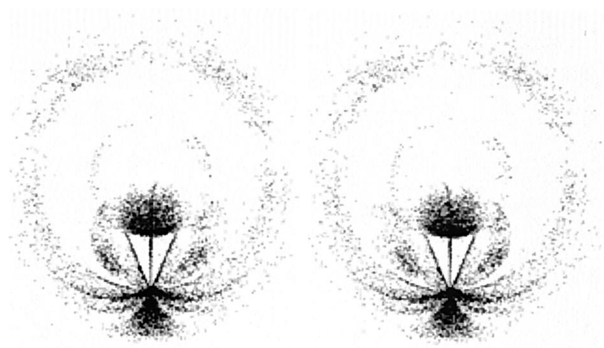

Fig. 5. Simulated stereo pair for a fisheye divergent-light halo display. Source elevation is 30°; the crystals are plate crystals with c/a ratio equal to 0.3 and an axis tilt of 0.5°. The c/a ratio is the ratio between the distance between the hexagonal end plates of the crystal and the distance between two opposite-end plates.

I have benefited greatly from stimulating discussions and critical comments from Jan O. Mattsson, Department of Physical Geography and Ecosystems Analysis, Lund University; Lund, Sweden; and Jarmo Moilanen and Marko Riikonen, Ursa Astronomical Association, Finland.

References and Note

1. W. Tape, *Atmospheric Halos* (American Geophysical Union, Washington, D.C., 1994).
2. A. J. Mallman, J. L. Hock, and R. G. Greenler, "Comparison of Sun pillars with light pillars from nearby light sources," Appl. Opt. **37,** 1441–1449 (1998).
3. K. Sassen, "Polarization and Brewster angle properties of light pillars," J. Opt. Soc. Am. A **4,** 570–580 (1987).
4. There is a figure made by Walter Tape displaying the surface of location of plate crystals that generate the parhelic circle in the Finnish journal Ursa Minor **11,** 1/2000 p. 11, edited by Tähtitieteellinen yhdistys Ursa in Helsinki. This figure also shows the position of the 120° parhelia.
5. J. O. Mattsson, L. Bärring, and E. Almqvist, "Experimenting with Minnaert's Cigar," Appl. Opt. **39,** 3604–3611 (2000).
6. P. K. Seidelmann, *Explanatory Supplement to the Astronomical Almanac,* (University Science, Sausalito, Calif., 1992).
7. L. Gislén and J. O. Mattsson, "Observations and simulations of some divergent-light halos," Appl. Opt. **42,** 4269–4279 (2003).