Ferronematics in Magnetic and Electric Fields

P. B. SUNIL KUMAR and G. S. RANGANATH
Raman Research Institute, Bangalore 560080, India

(Received November 4, 1988; in final form May 25, 1989)

We consider the effect of a magnetic field on a ferronematic. It is shown that first order Freedericksz transitions are possible in such systems as a consequence of elastic and diamagnetic anisotropies. Interestingly the possibility exists for a first order transition from a splay-bend distortion to a pure twist distortion or vice versa at a second threshold. A super-imposed electric field can also result in a first order Freedericksz transition.

Keywords: ferronematics, Freedericksz transition, tricritical behavior

INTRODUCTION

Ferronematics are nematic phases with preferentially aligned magnetic grains. Brochard and de Gennes\(^1\) were the first to consider theoretically the behavior of ferronematics in magnetic fields. They assumed the magnetization of the medium to be large enough for diamagnetic effects to be ignored (magnetization \(1\ G\)) and got a second order Freedericksz transition. Such ferronematic phases were first prepared in the Laboratory with needle-like grains by Rault et al.\(^2\) and later by others.\(^3\)\(^\text{-}^5\) Plate-like grains in the nematic phase have also been investigated.\(^6\) In these systems the magnetization appears to be very much smaller (of the order of \(10^{-4}\ G\)). In addition, Shen and Amer\(^5\) have reported in a ferronematic, with magnetization \(M\) perpendicular to the nematic director, the classical Freedericksz transition in the homeotropic geometry with applied field parallel to the magnetization. These results indicate that diamagnetic effects cannot be ignored. Further the full implications of elastic anisotropy have also not been worked out so far.

In this paper we consider the effects of both diamagnetic and elastic anisotropy in ferronematics in classical Freedericksz geometries. We find interestingly the transition to be first order below a critical magnetization or a critical ratio of the elastic constants and second order above it, thus exhibiting tricritical behavior. It should be mentioned that a first order transition is possible in the case of a twisted
nematic.\textsuperscript{7} Also more recently it has been shown\textsuperscript{8,9} that a classical Freedericksz transition can be made first order with the optical field of a laser.

In the homogeneous geometry we find at a higher field a first order transition from one type of distortion to another, in the presence elastic anisotropy. We have also worked out the consequences of an electric field acting along the imposed magnetic field. In this case also we find tricritical behavior.

**THEORY**

The free energy density for a ferronematic is given by:

\[
F = \frac{k_1}{2} (\nabla \cdot n)^2 + \frac{k_2}{2} (n \cdot \nabla \times n)^2 + \frac{k_3}{2} (n \times (\nabla \times n))^2
\]

\[- \frac{\chi_n}{2} (n \cdot H)^2 - M \cdot H - \frac{\epsilon_n}{8\pi} (n \cdot E)^2 \quad (1)
\]

where

\begin{align*}
M &= \text{Magnetization} \\
\chi_n &= \text{Positive diamagnetic anisotropy} \\
\epsilon_n &= \text{Positive dielectric anisotropy} \\
H &= \text{Magnetic field} \\
E &= \text{Electric field} \\
k_1, k_2, k_3 &= \text{are the splay, twist and bend elastic constants respectively.}
\end{align*}

We consider only ferronematics with $M$ along the director $n$. Throughout this paper we assume both electric and magnetic fields to be parallel to undistorted $n$. Further we assume the magnetic field to be antiparallel to the magnetization. The plates are at $z = 0$ and $z = d$ and are parallel to the $x$-$y$ plane. We consider only fields very close to the threshold.
HOMEOTROPIC ALIGNMENT

In this case the distortion is given by \( n = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta) \). This has a cylindrical symmetry about \( z \)-axis. In fact we get the familiar umbilic structure. Just above the threshold we can assume

\[
\theta = \theta_n(r) \sin \pi z/d
\]
\[
\phi = \pm \tan^{-1} \frac{y}{x}
\]

where \( d \) is the sample thickness.

Considering regions only far away from the core of the umbilic, i.e., \( r >> \xi \), we get the free energy density after averaging over sample thickness, as:

\[
\bar{F} = F_0 + \alpha/2 \theta_n + \beta/4 \theta_n^4 + \gamma/6 \theta_n^6
\]

(2)

where

\[
\alpha = (1/2)[k_3 \pi^2/d^2 - mH + \chi_u H^2 + \epsilon_u E^2/4\pi]
\]
\[
\beta = (1/4)[(k_1 - k_3) \pi^2/d^2 + mH/4 - \chi_u H^2 + \epsilon_u E^2/4\pi]
\]
\[
\gamma = (1/144)[(k_3 - k_1) \pi^2/d^2 + \chi_u H^2 - \epsilon_u E^2/4\pi - mH/16]
\]

We notice \( \bar{F} \) to be very similar to the Landau free energy density. We have included the sixth power term to allow for negative values of \( \beta \). Minimization of free energy density yields the equilibrium value of \( \theta_n \).

HOMOGENEOUS GEOMETRY

Here also we have both \( \theta \) and \( \phi \) distortions described by \( n = (\cos \theta \cos \phi, \cos \theta \sin \phi, \sin \theta) \). Just above the threshold we can assume \( \theta \) and \( \phi \) to be given by:

\[
\theta = \theta_n \sin \pi z/d
\]
\[
\phi = \phi_n \sin \pi z/d
\]

As before after averaging over the sample thickness we get

\[
\bar{F} = F_0 + \frac{\alpha_1}{2} \theta_n^2 + \frac{\alpha_2}{2} \phi_n^2 + \frac{\beta_1}{4} \theta_n^4 + \frac{\beta_2}{4} \phi_n^4 + \frac{\delta}{2} \theta_n^2 \phi_n^2
\]
Here

\[ \alpha_1 = \frac{1}{2} \left[ \frac{k_1 \pi^2}{d^2} - MH + \chi_a H^2 + \varepsilon_a E^2/4\pi \right] \]

\[ \alpha_2 = \frac{1}{2} \left[ \frac{k_2 \pi^2}{d^2} - MH + \chi_a H^2 + \varepsilon_a E^2/4\pi \right] \]

\[ \beta_1 = \frac{1}{4} \left[ \frac{(k_3 - k_1)\pi^2}{d^2} + MH/4 - \chi_a H^2 - \varepsilon_a E^2/4\pi \right] \]

\[ \beta_2 = \frac{1}{4} \left[ \frac{MH}{4} - \chi_a H^2 - \varepsilon_a E^2/4\pi \right] \]

\[ \delta = \frac{1}{8} \left[ \frac{(k_3 - 2k_2)\pi^2}{d^2} + 3 \left( \frac{MH}{2} - \chi_a H^2 - \varepsilon_a E^2/4\pi \right) \right] \]

This is similar to the generalized Lifshitz's expression for a two order parameter system.

When \( \beta_1 \) (or \( \beta_2 \)) is negative, we will also have to add the sixth power term in \( \theta \) (or \( \phi \)). Coefficient of this term for \( \theta \) and \( \phi \), respectively are:

\[ \gamma_1/6 = 1/96 \left[ \frac{(k_1 - k_3)\pi^2}{d^2} - \frac{MH}{24} + 2/3[\chi_a H^2 + \varepsilon_a E^2/4\pi] \right] \]

\[ \gamma_2/6 = 1/96 \left[ 2/3(\chi_a H^2 + \varepsilon_a E^2/4\pi) - MH/24 \right] \]

**RESULTS**

**Magnetic field effects**

In the case of homeotropic geometry even in the one constant approximation we find the important result of \( \beta \) changing sign as \( M \) varies. For a value of \( M \) less than the critical value \( M_c = 4\pi/d \ (k_\chi/3)^{1/2} \), \( \beta \) is negative resulting in a first order transition. And for higher values, \( \beta \) is positive giving a second order transition. Therefore there exists a tricritical behavior at \( M_c \). Figure 1 shows the variation of \( \theta \) with \( H \) in the first order transition. It should be emphasized that the harmonic solution that we have assumed is not strictly valid when the sixth power term is considered. Hence our \( \theta \) variation is only approximate. Figure 2 gives the variation of \( \Delta \theta \), the jump at transition as a function of \( M \). The first order transition takes place at \( \alpha = \beta^2/4\gamma \). Even when diamagnetic effects are ignored, we find tricritical behavior when elastic anisotropy is included. In this geometry we find the transition to be first order or second order depending upon whether \( k_3 > 4k_1/3 \) or \( k_3 < 4k_1/3 \). We find the Bruchard-de Gennes solution of a second order transition only when
neither the diamagnetic nor elastic anisotropies are included. However, when one or both of them are present we have tricritical behavior.

When the transition is second order we get an umbilic with a continuous core, i.e., $\theta$ smoothly goes to zero as $r \to 0$. The core is very similar to the one discussed by Rapini. In the case of first order Freedericksz transition the core would not be continuous and we can expect a first order jump in $\theta$ to zero as $r \to 0$.

In the present problem the vanishing of $\alpha$ yields two values of $H$, viz., $H_1$ and $H_2$ and $\alpha$ is negative between $H_1$ & $H_2$. The transition discussed above occurs at or below $H_1$. Also we can expect a reverse transition to the undistorted state for $H \gg H_2$. However the nature of this transition is difficult to work out due to the large value of $\theta$ the system has, near this field. In all probability this will be a first order transition since $\alpha > 0$ and $\beta > 0$ for $H \gg H_2$.

![Figure 1](image1.png)

**FIGURE 1** Onset of first order Freedericksz transition in the homeotropic geometry. $d = 20 \mu m$, $k_1 = k_2 = k_3 = 0.5 \times 10^{-6}$ dyne, $M = 1.8 \times 10^{-3}$ Gauss, $\chi_a = 0.5 \times 10^{-6}$ cgs units.

![Figure 2](image2.png)

**FIGURE 2** Jump $\Delta \theta$ at the transition as a function of $M$ for the homeotropic geometry and for parameters given in Figure 1.
However, for $M < 2\pi/d(k \chi \omega)^{1/2}$ the coefficient $\alpha$ is positive for all fields and we get a first order transition only when $\beta$ is negative.

In the case of homogeneously aligned sample in the one constant approximation the free energy density goes over to the Lifshitz expression for two order parameters.\textsuperscript{11} Lifshitz's theory predicts four possible states for the system:

(a) $\theta = 0, \phi = 0$ when $\alpha > 0$ (undistorted state)
(b) $\theta = 0, \phi \neq 0$ when $\alpha < 0, \beta < |\delta|
(c) $\theta \neq 0, \phi = 0$ when $\alpha < 0, \beta > |\delta|
(d) $\theta \neq 0, \phi \neq 0$

We always find only (b) and (c) solutions and never (d). Also (b) and (c) solution are equally energetic. This transition to either $\theta$ or $\phi$ alone can be first or second order, depending on $M$, just as in the case of homeotropic geometry. In this case it is possible to think of a new type of wall connecting these two distortions $\theta$ and $\phi$. This will be quite different from the familiar Brochard walls which connect only degenerate $\theta$ or $\phi$. These walls will be smooth when the transition is second order but will have discontinuities for a first order transition.

As before we again find interesting results even when diamagnetic terms are ignored. For example when $k_1 < k_2$ we always get solution (c) ($\theta \neq 0, \phi = 0$). This transition is first order for $k_3 < 4/3 k_1$ and second order for $k_3 > 4/3 k_1$. For $k_2 < k_1$ we always get solution (b) ($\phi \neq 0, \theta = 0$) whenever $k_3 \simeq k_1$ and this transition is always second order. However one can even get solution (c) i.e., ($\phi = 0 \theta \neq 0$), when $k_2 < k_1$ provides $k_3$ is very much smaller than $k_1$ and the transition to this state is always first order.

When diamagnetic effects are also included we get solution (c) ($\phi = 0, \theta \neq 0$) for $k_1 < k_2$ and $k_3 \neq k_1$. This transition will be first order or second order depending upon the magnetization. However, one can even get solution (b) (i.e., $\phi \neq 0, \theta = 0$) for $k_1 < k_2$ provided $k_3$ is very much larger than $k_1$. This will be first order. But diamagnetic effect does not drastically change the behavior when $k_2 < k_1$.

In this geometry we find another curious behavior. We have the possibility of a second threshold at which the system goes from solution (b) to solution (c) or vice versa under certain conditions. It goes from $\phi = 0, \theta \neq 0$ to $\phi \neq 0, \theta = 0$ through a first order transition when $k_1 < k_2$ and $k_1 < k_3$. Similarly it goes from $\theta = 0, \phi \neq 0$ to $\theta \neq 0, \phi = 0$ again through a first order transition when $k_1 > k_2$ and $k_1 > k_3$ (see Figure 3).

**COMBINED EFFECT OF ELECTRIC AND MAGNETIC FIELDS**

In view of many interesting effects associated with Freedericksz transition in combined electric and magnetic fields,\textsuperscript{12} it is worth looking at it in ferroelectricity also. We have considered the effects of an electric field applied parallel to the magnetic field. For simplicity we have assumed $k_1 = k_2 = k_3$. Up to a critical value of the electric field $E_c = 2\pi/d(k \pi/3\varepsilon \omega)^{1/2}$ the transition will continue to be second order. But at higher fields the transition will be first order thus exhibiting tricritical behavior as in the case of $M$. Figure 4 illustrates this feature.
FIGURE 3 The classical second order Freedericksz transition with the second instability in the homogeneous geometry. The second transition is of first order and it is from splay-bend to a twist configuration. $d = 20 \, \mu \text{m}$, $k_1 = 0.55 \times 10^{-6} \text{ dyne}$, $k_2 = 0.6 \times 10^{-6} \text{ dyne}$, $k_3 = 1.5 \times 10^{-6} \text{ dyne}$, $M = 8 \times 10^{-3} \text{ Gauss}$, $\chi_u = 10^{-7} \text{ cgs units}$.

FIGURE 4 Tricritical behavior in parallel electric and magnetic fields in the homeotropic geometry. $d = 5 \, \mu \text{m}$, $k_1 = k_2 = k_3 = 10^{-6} \text{ dyne}$. $M = 0.1 \text{ Gauss}$, $\epsilon_u = 0.1 \text{ cgs units}$. The dashed line represents the second order transition and the full line the first order transition.
NEMATICS WITH NEGATIVE DIAMAGNETIC ANISOTROPY

Here one always gets a second order Freedericksz transition in the one constant approximation. Also we find the elastic anisotropy to result in a first order transition from $0$ to $\phi$ or $\phi$ to $0$ at a second threshold.

Acknowledgments

Our thanks are due to Prof. S. Chandrasekhar and Prof. N. V. Madhusudana for discussions and to the referee for suggestions.

References

On Certain Liquid Crystal Defects in a Magnetic Field

P. B. SUNIL KUMAR and G. S. RANGANATH
Raman Research Institute, Bangalore 560080, India

(Received June 1, 1989; in final form July 10, 1989)

We find some interesting defect configurations in nematic, smectic and discotic liquid crystals in the presence of a magnetic field. We consider both positive and negative anisotropy materials subjected to a uniform or an all circular magnetic field. We have undertaken an energy analysis of many of these configurations.

Keywords: Defects, Disclinations, Dislocations, Defects in Magnetic Fields

INTRODUCTION

In a magnetic field a liquid crystal exhibits many interesting defect configurations arising due to the diamagnetic anisotropy of the molecules. The structure and properties of these defects have been studied in detail in nematics.\textsuperscript{1-5} But there does not appear to be such a thorough investigation in smectics and discotics. In this paper we look at some of the possible defect states in nematics, smectics and discotics in the presence of a magnetic field. We have considered not only the effect of a uniform magnetic field but also that of an all circular field generated by a linear current element. In addition we have treated the cases of positive as well as negative diamagnetic anisotropy.

Some interesting new results have emerged. For example, in nematics we find a new interaction law between disclinations when a uniform magnetic field is applied normal to the disclination lines. They interact with a distance independent force. A method of generating Poincaré half defects has been suggested. We also find that bubble domains can exist as natural states in an all circular magnetic field. In smectic A we find that screw dislocations in the presence of a uniform magnetic field interact like crystal screw dislocations. There also appears a structural instability resulting in screw dislocations in the presence of a magnetic field. The $s = +1$ all radial configuration develops a spiral dislocation in an all circular field above a threshold. In smectic C, however, we find a dispiration under a similar field. In columnar discotics, in addition to the smectic like spiral distortion we also get a helical columnar configuration in an all circular field.
INTERACTION BETWEEN DISCLINATIONS

When a magnetic field is applied normal to an half integral disclination line the resultant structure is a planar soliton terminating in a $\pm 1/2$ disclination. Interaction between two such defects is considered below.

Consider two $1/2$ wedge disclinations of opposite sign separated by a large distance. The splay soliton obtained in a magnetic field $H$ acting perpendicular to the line connecting them is shown in Figure 1(a). Here in the absence of the field the director at large distances is perpendicular to the line connecting the defects. It is

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{A pair of unlike disclinations of strength $1/2$ in a magnetic field acting normal to the disclination lines (a) Field perpendicular to the line joining the disclinations. (b) Field along the line joining disclinations.}
\end{figure}
clear from the figure that most of the splay distortion is confined to the region between the disclinations and within the width of the soliton. We see that by moving the two disclinations towards each other the director distortion in the regions indicated by dashed circles (with the radii of the order of magnetic coherence length) and in the far off regions are not much affected. But this process reduces the energy of the configuration in the central region. To a good approximation this is $\Delta E = 2H (k \chi_a)^{1/2} \Delta d$, with $\chi_a$ as the diamagnetic anisotropy and $k$ as the elastic constant in the one constant approximation. This is proportional to the change $\Delta d$ in the distance of separation between the disclinations. Hence we find a distance independent force as opposed to the familiar $1/d$ law.

Interestingly, unlike disclinations can also repel one another under certain conditions. This is depicted in Figure 1(b). Here the magnetic field is parallel to the line joining the disclinations. We get planar bend soliton going on either side to infinity away from the defects. The distortion in the region indicated by the dashed circles is again unaltered by changes in $d$. But in the central, nearly distortion-free region, energy decreases with increase of separation. Thus the field favours repulsion. The change in energy in this case is

$$\Delta E = - \frac{\chi_a H^2}{2} \Delta d.$$  

Thus we see that two disclinations of opposite strength will attract or repel with a distance independent force. The magnitude of the attractive and repulsive forces are, however, different. In the same way interaction between like disclinations can also be repulsive or attractive with a force independent of distance of separation. The same arguments can be extended to twist disclinations as well.

It must be remarked that the standard elastic interaction between disclinations is not totally absent in the presence of the magnetic field. But exact calculation of the net interaction is not easy. What we can do is to undertake and approximate analysis. We know that the elastic free energy density varies as $k (\nabla \theta)^2$ [$\theta$ is the director orientation] where as the magnetic energy density varies as $\chi_a H^2 \sin^2 \theta$. Hence over distances less than the coherence length $\xi = (k/\chi_a H^2)^{1/2}$, we can, to a good approximation ignore the magnetic term. Thus we can argue that when the separation $d$ between the unlike (like) defects is much less than $\xi$ the elastic attractive (repulsive) interaction dominates. The distance independent law is valid for $d \gg \xi$. Therefore $d$ of the order of $\xi$ is probably the region where we go from the $1/d$ law to the distance independent law of interaction.

**Poincaré structures**

It is known\(^7\) that a line singularity of strength $\pm 1$ can end in a half disclination point. This implies that a line singularity can be terminated with an unlike pair of 1/2 disclination points. We suggest here a method of getting these defect states which appear not to have been seen experimentally so far.

If a magnetic field is applied parallel to the director of a homeotropically aligned nematic with negative diamagnetic anisotropy, it will undergo a Freedericksz tran-
sition at a critical field. The resulting structure will be non-singular as shown in Figure 2(a). But in a central region the director still is opposing the magnetic torque. Hence at fields much higher than this critical field this structure can break down to the one shown in Figure 2(b). Here a $s = +1$ line singularity has ended in a pair of unlike half Poincaré point singularities. This is the analogue of the pinching effect in Brochard Walls.

This phenomena can be expected in nematic discotics as they usually have negative diamagnetic anisotropy. In rod-like nematics it is easier to get systems with negative dielectric anisotropy. Here the above arguments are valid *mutatis mutandis* in the presence of an electric field.

**Bubble domains**

Symmetry of a nematic liquid crystal permits one to consider, in a uniform magnetic field, cylindrical shell structures or bubble domains$^5$ separating the inside and the outside regions by a 180° twist or bend distortion. We now investigate the possibility of such a bubble domain in the presence of an all circular magnetic field $H_\alpha =$

![Diagram](image)

**FIGURE 2** A nematic with negative diamagnetic anisotropy subjected to a magnetic field along the director, in the homeotropic geometry. (a) Just above the Freericksz threshold. (b) At much higher fields.
A/r acting on a nematic with negative diamagnetic anisotropy. The director \( n \) defined in cylindrical polars are:

\[
n = [\sin \theta \cos (\phi - \alpha), \, \sin \theta \sin (\phi - \alpha), \, \cos \theta]
\]

The free energy density is given by

\[
F = \frac{k}{2} [(\nabla \theta)^2 + \sin^2 \theta (\nabla \phi)^2] + \frac{\chi_a A^2}{2r^2} \sin \theta \sin^2 (\phi - \alpha)
\]

The differential equation obtained by minimization of this free energy are

\[
k[(\nabla^2 \theta) - \sin \theta \cos \theta (\nabla \phi)^2] - \frac{\chi_a A^2}{r^2} \sin \theta \cos \theta \sin^2 (\phi - \alpha) = 0
\]

\[
k(\nabla^2 \phi) - \frac{\chi_a A^2}{r^2} \sin (\phi - \alpha) \cos (\phi - \alpha) = 0
\]

The solutions satisfying the boundary conditions \( \theta = 0 \) at \( r = 0 \) and \( \theta = \pi \) at \( r = \infty \) are:

\[
\theta = 2 \tan^{-1} \left( r/r_o \right)^n \text{ and } \phi = \alpha + \pi/2
\]

where

\[
\eta = [1 + \chi_a A^2/k]^{1/2}
\]

Here \( r_o \) is the point at which \( \theta \) becomes \( \pi/2 \). This represents a Bloch bubble domain as shown in Figure 3. The variation of \( \theta \) with respect to \( r/r_o \) is shown in Figure 4 for different values of \( \eta \). We see that the width of the domain wall decreases as the field increases. Thus in such a field a bubble domain is a natural soliton solution. The total energy of this structure per unit length is found to be \( 4\pi k \eta \). Interestingly the energy is independent of its radius \( r_o \). At high fields \( \eta = [\chi_a A^2/k]^{1/2} \) and the energy is \( 4\pi A (k\chi_a)^{1/2} \), which is \( 2\pi r_o \) times the surface tension of the planar soliton obtained in uniform fields.

Bubble domains exist also in diamagnetically positive materials but here the field should be such that \( \chi_a A^2 > k \). At lower fields we find a collapsed \(+1\) all circular disclination.\(^9\) This becomes a planar singular structure at the critical value of \( A = (k/\chi_a)^{1/2} \). In this all circular planar structure we can now construct a twist-bubble
or an inplane bend-bubble domain. It is easy to analyze the twist-bubble where
the boundary condition are

\[ \text{at } r = 0 \quad \theta = -\pi/2 \quad \text{and} \quad \text{at } r = \infty \quad \theta = +\pi/2 \]

from the equation of equilibrium we get

\[ \phi = \alpha + \pi/2 \]

and

\[ \theta = 2 \tan^{-1} \left[ \left( r/r_o \right)^{N} - 1 \right] \]

where

\[ \eta = \left[ \chi_{u}A^{2}/k - 1 \right]^{1/2} \]

and the extra energy due to the bubble domain is again \( 4\pi k \eta \) per unit length.
FIGURE 4 The director tilt $\theta$ in a twist-bubble domain of radius $r_o$, as a function of the distance from the centre. (a) In the absence of the field, i.e., $\eta = 1$. (b) In a field with $\eta = 4$ and (c) with $\eta = 6$.

**SMECTICS**

**Interaction between screw dislocations in smectic A**

In the linear elastic theory screw dislocations in smectic A have neither self energy nor do they interact. We shall consider this problem in a uniform field acting along z-axis. The free energy density for $\chi_a > 0$ is

$$F = \frac{B}{2} \left( \frac{\partial u}{\partial z} \right)^2 + \frac{k_{11}}{2} \left[ \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right]^2 + \chi_a H^2 \left[ \left( \frac{\partial u}{\partial x} \right)^2 + \left( \frac{\partial u}{\partial y} \right)^2 \right]$$

where $u =$ The layer displacement

$B =$ The elastic constant for lattice dilation

$k_{11} =$ the elastic constant for layer curvature
Minimization of energy results in the following equation of equilibrium:

\[ B \left( \frac{\partial^2 u}{\partial z^2} \right) - k_{11} \left[ \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right] u + \chi_a H^2 \left[ \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right] = 0 \]

This permits the following solution

\[ u = \frac{b}{2\pi} \tan^{-1} \left( \frac{y}{x} \right) \]

This is the familiar screw dislocation. Its strength \( b \) is an integral multiple of \( a_o \) the layer spacing. This has an energy per unit length given by

\[ \frac{b^2}{4\pi} \chi_a H^2 \ln \left( \frac{R}{r_c} \right) \]

Here \( R \) is the sample size and \( r_c \) is the core radius. This energy is purely magnetic with no elastic contribution whatever.

A linear combination of such solutions is also valid. This results in an interaction energy per unit length

\[ \varepsilon_t = \left( \frac{b_1 b_2}{2\pi} \right) \chi_a H^2 \ln \left( \frac{2R}{d} \right) \]

with \( d \) as the distance of separation between the dislocations.

Hence an interaction exists in a magnetic field. Like dislocations repel and unlike dislocations attract with a force proportional to \( 1/d \). This interaction is exactly the same as that between screw dislocations in crystals. The same interaction law is also obtained in a compressed smectic due to the non-linear term of the form

\[ \frac{B}{2} \left( \frac{\partial u}{\partial z} \right) \left[ \left( \frac{\partial u}{\partial x} \right)^2 + \left( \frac{\partial u}{\partial y} \right)^2 \right] . \]

Due to its similarities with the magnetic term we find the force to be proportional to the applied compressive strain. It may be mentioned in passing that a very similar interaction law has also been proposed by Pleiner, again arising from a non-linearity but of a different nature. In his analysis the dislocations interact even in a stress-free smectic.
Field induced defects

(1) H\textsubscript{z} field Consider a homeotropically aligned sample with negative diamagnetic anisotropy and with layers in the x-y plane. The free energy density in presence of a field parallel to z-axis, for small layer distortions \( u \) is given by

\[
F = \frac{B}{2} \left( \frac{\partial u}{\partial z} \right)^2 + \frac{k_{11}}{2} \left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right)^2 - \frac{\chi_a H^2}{2} \left[ \left( \frac{\partial u}{\partial x} \right)^2 + \left( \frac{\partial u}{\partial y} \right)^2 \right]
\]

It can be easily verified that \( u = \frac{b}{2\pi} \tan^{-1} (y/x) \) which represents a screw dislocation, is again a solution of the equation of equilibrium and the total energy per unit length is given by

\[
E = -\frac{\chi_a H^2 b^2}{4\pi} \ln \left( \frac{R}{r_c} \right)
\]

Since this energy is always negative the structure develops an instability through a proliferation of screw dislocations. Solutions with contributions from the elastic term will be of higher energy and a threshold will be needed to excite them.

We can expect the same solution in the case of materials with negative dielectric anisotropy, in the presence of an electric field along the layer normal.

(2) H\textsubscript{a} field Here we will investigate the same geometry as in (1) but with \( \chi_a > 0 \) and in the presence of an all circular field acting parallel to the layers. We expect a perturbation of the form

\[
n_r = 0, \quad n_\alpha = \frac{-1}{r} \left( \frac{\partial u}{\partial \alpha} \right), \quad n_z = 1 - \frac{1}{2r^2} \left( \frac{\partial u}{\partial \alpha} \right)^2
\]

The free energy density can be written as

\[
F = \frac{k_{11}}{2r^4} \left( \frac{\partial^2 u}{\partial \alpha^2} \right)^2 - \frac{\chi_a A^2}{2r^4} \left( \frac{\partial u}{\partial \alpha} \right)^2
\]

Minimization gives

\[
\frac{k_{11}}{r^4} \frac{\partial^4 u}{\partial \alpha^4} + \frac{\chi_a A^2}{r^4} \frac{\partial^2 u}{\partial \alpha^2} = 0
\]

This permits again the screw dislocation solution \( u = b/2\pi \alpha \) which has an energy per unit length, given by

\[
E = -\frac{\chi_a A^2 b^2}{8\pi} \left[ \frac{1}{r_c^2} - \frac{1}{R^2} \right]
\]
As $R > r_c$ this energy is always negative. Hence the system is again destabilised through screw dislocation creation. This screw dislocation is different from the classical one in the sense that its energy is finite even for an infinite sample.

**Spiral instability in $H_\alpha$ field**

Consider the wedge +1 disclination along the z axis with an all radial director configuration. In an all circular magnetic field of the form $(0, A/r, 0)$ acting parallel to the smectic layers with molecules having positive diamagnetic anisotropy we can expect a perturbation of the form

$$ n_r = 1 - \frac{1}{2r^2} \left( \frac{\partial u}{\partial \alpha} \right)^2, \quad n_\alpha = -\frac{1}{r} \left( \frac{\partial u}{\partial \alpha} \right), \quad n_z = 0 $$

Here $u$ is the layer displacement along the radial direction. The free energy density upto second order in $u$ is given by

$$ F = \frac{k_{11}}{2} \left[ \frac{1}{r^2} + \frac{1}{r^4} \left( \frac{\partial u}{\partial \alpha} \right)^2 \right] - \frac{\chi_\alpha A^2}{2r^2} \left[ -\frac{1}{r} \frac{\partial u}{\partial \alpha} \right] $$

minimisation of the free energy gives the solution

$$ u = \frac{a_\alpha N}{2\pi} \alpha $$

where $a_\alpha$ is the layer spacing and the $N$ an integer. This represents a spiral distortion, resulting in a helical wrapping of smectic layers. The change in free energy due to this distortion is

$$ \delta F = [k_{11} - \chi_\alpha A^2] \frac{1}{2r^2} \left( \frac{\partial u}{\partial \alpha} \right)^2 $$

Thus, for values of $A$ greater than $(k_{11}/\chi_\alpha)^{1/2}$ the system is unstable against such distortions. Such a structure was first described by Kleman and Parodi.\(^{10}\) This is a kind of dislocation. As indicated by them this structure can even have a weak twist in the $z$-direction. This solution is very different from the spiral structures discussed by Bouligand\(^ {11}\) for smectic A with layers ending on a cylindrical boundary.

**Smectic C**

In a magnetic field $H_z$ acting normal to the layers and with only layer displacements we find again the same solutions as those obtained earlier for smectic A. However, in a field $(O, H_\alpha, O)$ acting parallel to the smectic layer, we get a different solution. Here we can get as in smectic A a screw dislocation. In addition even a disclination
in the C -director can be expected. For a small perturbation \( u \), the free energy density with \( \chi_\alpha > 0 \) becomes

\[
F = \frac{k}{2} \theta^2 (\nabla \phi)^2 - \frac{\chi_\alpha A^2}{2r^2} \left[ \frac{1}{r} \left( \frac{\partial u}{\partial x} + \theta \sin (\phi - \alpha) \right) \right]^2 + \frac{k_{11}}{2r^4} \left[ \frac{\partial^2 u}{\partial \alpha^2} \right]^2
\]

where \( u \) = layer displacement

\( \theta \) = tilt angle

\( \phi \) = azimuth of the C director

\( k \) = the elastic constant for bend or splay in the C-director

Minimization yields solution:

the disclination \( \phi = \alpha + \pi/2 \) and

the screw dislocation \( u = a_\alpha N/2\pi \alpha \)

The free energy density for this solution is

\[
F = \frac{\theta^2}{2r^2} [k - \chi_\alpha A^2] - \frac{\chi_\alpha A^2}{2r^2} \left[ \left( \frac{b}{2\pi r} \right)^2 + \frac{b\theta}{2\pi r} \right]
\]

The system will have a negative energy due to the screw dislocation alone. At fields higher than the critical field \( A_c = (k/\chi_\alpha)^{1/2} \) this energy is further lowered by the creation of a disclination in the C director. The resultant distortion has both dislocation and disclination characteristics. In other words we end up with a dispiration as the low energy defect for field \( A > A_c \).

**COLUMNAR DISCOTICS**

The simplest of the columnar discotics has hexagonal symmetry. One possible defect state has columns bent around the z-axis into concentric circles,\(^{11} \) i.e., \( n_\alpha = 1 \). In an all circular field with \( \chi_\alpha < 0 \) (this is usually the case for discotics) this defect state gets perturbed. We can assume the perturbation to be of the form:

\[
n_r = - \frac{1}{r} \left( \frac{\partial u}{\partial \alpha} \right), \quad n_\alpha = 1 - \frac{1}{2r^2} \left( \frac{\partial u}{\partial \alpha} \right)^2, \quad n_z = 0
\]
As $R > r_c$ this energy is always negative. Hence the system is again destabilised through screw dislocation creation. This screw dislocation is different from the classical one in the sense that its energy is finite even for an infinite sample.

**Spiral Instability in $H_\alpha$ field**

Consider the wedge $+1$ disclination along the $z$ axis with an all radial director configuration. In an all circular magnetic field of the form $(0, A/r, 0)$ acting parallel to the smectic layers with molecules having positive diamagnetic anisotropy we can expect a perturbation of the form

$$n_r = 1 - \frac{1}{2r^2} \left( \frac{\partial u}{\partial \alpha} \right)^2, \quad n_\alpha = -\frac{1}{r} \left( \frac{\partial u}{\partial \alpha} \right), \quad n_z = 0$$

Here $u$ is the layer displacement along the radial direction. The free energy density up to second order in $u$ is given by

$$F = \frac{k_{11}}{2} \left[ \frac{1}{r^2} + \frac{1}{r^4} \left( \frac{\partial u}{\partial \alpha} \right)^2 \right] - \frac{\chi_\alpha A^2}{2r^2} \left[ \frac{-1}{r} \frac{\partial u}{\partial \alpha} \right]^2$$

Minimisation of the free energy gives the solution

$$u = \frac{a_\alpha N}{2\pi} \alpha$$

where $a_\alpha$ is the layer spacing and the $N$ an integer. This represents a spiral distortion, resulting in a helical wrapping of smectic layers. The change in free energy due to this distortion is

$$\delta F = \left[ k_{11} - \chi_\alpha A^2 \right] \frac{1}{2r^4} \left( \frac{\partial u}{\partial \alpha} \right)^2$$

Thus, for values of $A$ greater than $(k_{11}/\chi_\alpha)^{1/2}$ the system is unstable against such distortions. Such a structure was first described by Kleman and Parodi. This is a kind of dislocation. As indicated by them this structure can even have a weak twist in the $z$-direction. This solution is very different from the spiral structures discussed by Bouligand for smectic A with layers ending on a cylindrical boundary.

**Smectic C**

In a magnetic field $H_z$ acting normal to the layers and with only layer displacements we find again the same solutions as those obtained earlier for smectic A. However, in a field $(O, H_\alpha, O)$ acting parallel to the smectic layer, we get a different solution. Here we can get as in smectic A a screw dislocation. In addition even a disclination
DEFECTS IN MAGNETIC FIELDS

in the C-director can be expected. For a small perturbation \( u \), the free energy density with \( \chi_a > 0 \) becomes

\[
F = \frac{k}{2} \theta^2 (\nabla \phi)^2 - \frac{\chi_a A^2}{2r^2} \left[ \frac{1}{r} \left( \frac{\partial u}{\partial \alpha} \right) + \theta \sin (\phi - \alpha) \right]^2 + \frac{k_{11}}{2r^4} \left[ \frac{\partial^2 u}{\partial \alpha^2} \right]^2
\]

where \( u \) = layer displacement

\( \theta \) = tilt angle

\( \phi \) = azimuth of the C director

\( k \) = the elastic constant for bend or splay in the C-director

Minimization yields solution:

the disclination \( \phi = \alpha + \pi/2 \) and

the screw dislocation \( u = a_o N/2\pi \alpha \)

The free energy density for this solution is

\[
F = \frac{\theta^2}{2r^2} [k - \chi_a A^2] - \frac{\chi_a A^2}{2r^2} \left[ \left( \frac{b}{2\pi r} \right)^2 + \frac{b\theta}{2\pi r} \right]
\]

The system will have a negative energy due to the screw dislocation alone. At fields higher than the critical field \( A_c = (k/\chi_a)^{1/2} \) this energy is further lowered by the creation of a disclination in the C director. The resultant distortion has both dislocation and disclination characteristics. In other words we end up with a dispiration as the low energy defect for field \( A > A_c \).

COLUMNAR DISCOTICS

The simplest of the columnar discotics has hexagonal symmetry. One possible defect state has columns bent around the z-axis into concentric circles, i.e., \( n_\alpha = 1 \). In an all circular field with \( \chi_a < 0 \) (this is usually the case for discotics) this defect state gets perturbed. We can assume the perturbation to be of the form:

\[
n_r = -\frac{1}{r} \left( \frac{\partial u}{\partial \alpha} \right), \quad n_\alpha = 1 - \frac{1}{2r^2} \left( \frac{\partial u}{\partial \alpha} \right)^2, \quad n_z = 0
\]
where \( u \) is the displacement of the columns perpendicular to the director \( n \). The various columnar circles in any given plane get interconnected to form a spiral. The change in free energy density due to this distortion is

\[
\delta F = \frac{1}{2r^4} \left( \frac{\partial u}{\partial \alpha} \right)^2 (k_{33} - \chi_\alpha a^2)
\]

\( k_{33} \) = The bend elastic constant

Here to a good approximation we can neglect the lattice distortions. The minimisation of the free energy permits the solution in \( u = (b_\perp/2\pi)\alpha \), where \( b_\perp \) is an integral multiple of the spacing of the columns perpendicular to the \( z \)-axis. This will exist above a threshold given by \( A_c = [k_{33}/\chi_\alpha]^{1/2} \). Thus a smectic-like spiral dislocation exists in addition to the disclination in the director \( n \). We have a dispiration configuration in which a classical disclination is associated with a spiral dislocation.

Another possible mode of distortion in the same field is of the form

\[
\begin{align*}
n_x &= 0, \quad n_\alpha = 1 - \frac{1}{2r^2} \left( \frac{\partial u}{\partial \alpha} \right)^2, \\
n_z &= -\frac{1}{r} \left( \frac{\partial u}{\partial \alpha} \right)
\end{align*}
\]

Here also neglecting the lattice distortions we can write the extra free energy density as

\[
\delta F = \frac{1}{2r^4} \left( \frac{\partial u}{\partial \alpha} \right)^2 (2k_{33} - \chi_\alpha a^2)
\]

The critical field here is \( A_c = [2k_{33}/\chi_\alpha]^{1/2} \), above which we get \( u = (b_\parallel/2\pi)\alpha \). Here \( b_\parallel \) is an integral multiple of the layer spacing parallel to \( z \)-axis. This results in a helical connection between the circular column, i.e., we get coaxial helices.

These two possible modes are very different from the spiral columns around cylinder and a helical distortion around a helix proposed by Kleman\textsuperscript{12} and Bouligand\textsuperscript{11} for columnar discotics, from the theory of developable domains.

**SPECULATIONS ON THE CORE NEAR A-C TRANSITION**

Consider a \( \pm 1 \) disclination in smectic C near A-C transition. Very near the transition point the order parameter \( \theta \) is small. Using the Landau theory of A-C
transition, we get the Ginsburg-Pitaevskii equation for the core in the presence of a $H_z$ field for $\chi_a > 0$.

\[
\frac{1}{\xi} \frac{\partial}{\partial \xi} \left( \xi \frac{\partial f}{\partial \xi} \right) - f/\xi^2 + f(1 - f^2) = 0
\]

where

\[
\xi = \frac{r}{\xi_0}, \quad \xi_0 = (-k/\alpha')^{1/2}
\]

\[
f = \frac{\theta}{\theta_0}, \quad \theta_0 = \text{Tilt angle at } r \to \infty = (\alpha'/\beta')^{1/2}
\]

\[
\alpha' = \alpha + \chi_a H^2, \quad \beta' = \beta - \frac{2}{3} \chi_a H^2
\]

\[
\alpha = \alpha_0(T - T_c) \quad \beta > 0
\]

$\alpha_0$ and $\beta$ are thermodynamic parameters of the classical Landau expansion. For $\beta' > 0$ the tilt angle drops to zero at the centre of the singularity. Most of the variation in $\theta$ takes place over a distance $\xi_n$. As $H$ increases $\xi_n$ increases slowly sweeping whole area. At the critical field $H = [|\alpha/\chi_a|]^{1/2}$ tilt angle becomes zero everywhere, i.e., we get a smectic A state.

It is possible for $\beta'$ to be negative depending on the thermodynamic parameter $\beta$ and the field strength $H$. When this happens we can speculate on the possibility of a first order transition to the smectic A state.

On the smectic A side of the A-C transition with $\chi_a > 0$ and an all circular field acting parallel to the smectic layers the free energy density is

\[
F = \frac{k}{2} [((\nabla \theta)^2 + \theta^2/r^2] + \frac{\alpha \theta^2}{2} + \frac{\beta \theta^4}{2} + \frac{\chi_a A^2}{2 r^4} (1 - \theta^2 + \theta^4/3)
\]

From this we conclude that the coefficient of $\theta^2$ (which is always positive in the absence of the field) has a fair chance of becoming negative at very small distances. Within this range we get a smectic C-like disclination. A detailed analysis of this again leads to a Ginsburg-Pitaevskii type equation.

**Acknowledgments**

Our thanks are due to Professor S. Chandrasekhar and Professor N. V. Madhusudana for discussions and for the referee for comments.
References

On Some Topological Solitons in Ferronematics

P. B. SUNIL KUMAR and G. S. RANGANATH
Raman Research Institute, Bangalore 560080, India

(Received March 11, 1990)

We have considered the structure and energetics of planar and linear soliton states in Ferronematics in a magnetic field. The $2\pi$ planar soliton in a ferronematic with positive diamagnetic anisotropy becomes unstable above a threshold field, splitting into two $\pi$ planar solitons. When diamagnetic anisotropy is negative, we get the classical linear soliton below a threshold field. This becomes unstable above the threshold field leading to two new types of linear solitons. The structures of these solitons appear to be very sensitive to elastic anisotropy. An extension of this to nematics having solute molecules has also been considered.

Keywords: Topological defects; ferronematics; defect instability.

INTRODUCTION

Ferronematics (FN) and ferrocholesterics (FC) have small elongated magnetic particles suspended in the nematic and the cholesteric phases respectively. Such ferro systems have been prepared and studied experimentally. Generally the grains are of width of about 100 Å with an aspect ratio of about 10. There appears to be a very good mechanical coupling between the grains and the host. Recently even a ferroelastatic phase has been reported.

Their behaviour in an external magnetic field was first discussed theoretically by Brochard and de Gennes and has gained lot of attention recently. In this paper we deal with the topological solitons that can exist in FN in the presence of a magnetic field.

Diamagnetic anisotropy of the nematic host and the magnetisation of the grains strongly influence the structure of these defects. We have considered both positive and negative diamagnetically anisotropic host systems. The effects of grain segregation at higher fields have also been worked out. Throughout our discussion the magnetisation of the grains is considered to be parallel to the director. We find in the diamagnetically positive FN a $2\pi$ planar soliton to become unstable above a threshold field splitting into two $\pi$ solitons. In diamagnetically negative FN, we get two new types of three dimensional non singular defects again above a threshold. The structure and energy of these defects are found to depend markedly on the elastic anisotropy.
FERRONEMATICS WITH POSITIVE DIAMAGNETIC ANISOTropy

In this case in an infinite sample, we find a $2\pi$ wall (planar soliton) as a permitted solution. Far away from a given plane $z = 0$, the director will be aligned along magnetic field $H$, with $m$ the magnetisation parallel to $H$. And most of the $2\pi$ distortion is confined to a narrow region of space. The free energy density in the one constant approximation for a dilute solution of grains in the nematic is

$$F = \frac{k}{2} (\theta, z)^2 - \frac{\chi_d H^2}{2} \sin^2 \theta - fmH \sin \theta + \frac{f k_B T}{V} \ln (f)$$

where

$H$ = magnetic field along the $x$-axis
$k$ = elastic constant
$\chi_d$ = diamagnetic anisotropy
$m$ = magnetisation of grains
$V$ = volume of the individual grain
$f$ = volume fraction of the grains
$\theta$ = angle w.r.t. $y$-axis
$k_B$ = Boltzmann constant
$T$ = temperature

Minimising Equation (1) with respect to $f$ and $\theta$, we get

$$-\rho_0 H \sin \theta + \ln f + 1 = 0$$

and

$$\theta = \frac{\sin(2\theta)}{2\xi_1^2} - \frac{\varphi \cos(\theta)}{\xi_2^2}$$

where

$$\rho_0 = m \frac{V}{k_B T}$$
$$\varphi = \frac{\bar{f}}{f}$$
$$\xi_1^2 = k / (\chi_d H^2)$$
$$\xi_2^2 = k / (m\bar{f}H)$$

$\bar{f}$ = average volume fraction of the grains in the undistorted state.

The parameter $\rho_0$ can be considered as a measure of the degree of segregation of the grains. Solving Equation (2) with boundary condition $f = \bar{f}$ at $z = \pm \infty$, where $\theta = \text{odd multiple of } \pi/2$, we get

$$\varphi = \exp[\rho_0 (\sin \theta - 1)H]$$
Here it is assumed that when the grains are expelled out of the central region to the outer regions, the average grain concentration remains unaltered at $z \to \pm \infty$.

Equations (3) and (4) have been solved numerically with boundary conditions

$$\theta, z = 0 \text{ at } z = \pm \infty, \quad \theta = \pi/z \text{ at } z = +\infty \text{ and } \theta = -3\pi/2 \text{ at } z = -\infty.$$  

We find the interesting result that only below a critical field $H_c$ given by

$$H_c = \frac{n\tilde{g}}{\chi_a} \exp (-2 \rho_0 H_c)$$

Equation (3) allows a $2\pi$ planar soliton solution. But at fields $H > H_c$ this splits into two $\pi$ solitons. In Figure 1 we show this effect along with the segregation of grains. As is to be expected the magnetic grains have moved out of the central region. This segregation increases with increasing field. Also $H_c$ is found to decrease rapidly with increasing $\rho_0$. This variation is depicted in Figure 2. We also find that though the splitting, i.e., distance of separation between two $\pi$ walls increases with $\rho_0$ it is far less sensitive to increases in field strength.

It should be emphasized that the symmetry of the system also permits us to construct linear soliton of the Mineev-Volovik type.\textsuperscript{9,10} Here $m$ is parallel to $H$ far away from a given line. But its structure and energetics cannot be worked out without the higher gradient terms in $F$.\textsuperscript{9,11}

**FERRONEMATICS WITH NEGATIVE DIAMAGNETIC ANISOTROPY**

A non-singular defect, cylindrically symmetric about the magnetic field, is one of the allowed topological solutions. The distortion in this case is given by $n = (\sin \theta \cos(\psi - \alpha), \sin \theta \sin(\psi - \alpha), \cos \theta) \theta$ being angle w.r.t. to $H$ and $\psi$ is the azimuthal angle. Minimising the free energy w.r.t. $\psi$ we get the following differential equation in cylindrical polars.

$$[(k_3 \sin^2 \theta + k_2 \cos^2 \theta) \sin^2(\psi - \alpha) + k_1 \cos^2(\psi - \alpha)] \sin^2 \theta \frac{\dot{\psi},\alpha\alpha}{r^2}$$

$$\left\{[-k_1 \cos^2 \theta + k_3 \sin^2 \theta] \theta, r^2 + [k_3 \sin^2 \theta + k_2 \cos^2 \theta - k_1] \right\}$$

$$\sin^2 \theta \frac{(2 - \psi,\alpha)}{r^2} \psi,\alpha + [k_2 - k_1] \sin(2\theta) \frac{\theta, r \psi,\alpha}{r} \left\{ \frac{\sin(2(\psi - \alpha))}{2} \right\}$$

where $k_1, k_2, k_3$ are respectively the splay, twist and bend elastic constants. We consider the following situations, which are permitted by the above equation.
FIGURE 1 Director profile, tilt \( \theta \) and concentration \( f / \bar{f} \) for 2\( \pi \) twist planar soliton for \( k = 0.5 \times 10^{-6} \) dynes, \( |\chi_0| = 10^{-6} \) cgs units, \( \rho_n = 0.02 \) Gauss\(^{-1} \), \( f = 10^{-3} \) and \( H_c = 78 \) Gauss, for (a) \( H = 70 \) Gauss, (b) \( H = 1000 \) Gauss.
FIGURE 2 Variation of the critical field $H_c$ with the segregation parameter $\rho_\alpha$.

(i) All radial configuration ($\psi = \alpha$). In this case $\theta$ is given by

$$2[1 + \varepsilon' \cos(2\theta)]\theta,rr = \frac{(1 + \varepsilon') \sin(2\theta)}{r^2} - \frac{2(1 + \varepsilon' \cos(2\theta))}{r} \theta,rr$$

$$+ \frac{\sin(2\theta)}{\xi_1^2} + \frac{2 \sin\theta}{\xi_2^2} + 2 \varepsilon' \sin(2\theta) (0,r)^2$$

where

$$\xi_1^2 = \frac{k'}{\chi_a H^2} \hspace{1cm} \xi_2^2 = \frac{k'}{m f H}$$

$$\varepsilon' = \frac{k_1 - k_3}{k_1 + k_3} \hspace{1cm} k' = \frac{k_1 + k_3}{2}$$

(ii) All circular structure ($\psi = \alpha + \pi/2$). Here for $\theta$ we have

$$2(1 + \varepsilon'')\theta,rr = \frac{\sin(2\theta)}{r^2} (1 - \varepsilon'' + 2 \varepsilon'' \cos(2\theta))$$

$$\frac{\sin(2\theta)}{\xi_1^2} + \frac{2 \sin\theta}{\xi_2^2} - \frac{2(1 + \varepsilon'')}{r} \theta,rr$$
FIGURE 3  (a) N Flower and (b) W Flower configurations in ferronematics in a magnetic field $H$ 3000 Gauss.
SOLITONS IN FERRONEMATICS

It can be seen that the orientation at $r \to \infty$ can be obtained by considering the magnetic torque alone. This permits the director orientation to be given by

$$\sin \theta = 0$$

or

$$\cos \theta = (-m/\chi_a H)$$

If $H < m/\chi_a$, $\theta$ can be 0 or $N\pi$, resulting in either the uniform state or a linear soliton. Above $H > Hc = m/\chi_a$ the director will be aligned at an angle $\theta_m = \cos^{-1} (-m/\chi_a H)$. Thus the previous states are not permitted. However, when magnetic grains are absent, i.e., pure $\chi_a$ negative materials at every field $\theta = 0$ at $r = 0$ and $\theta = \pi/2$ at $r \to \infty$, resulting in what can be termed as “U Flower”.

In a field $H > Hc$ it is possible to have two cylindrically symmetric structures

![Figure 4](image-url)
FIGURE 5 Schematic representation of possible transformations of an Hedge Hog in a magnetic field. (a) Classical linear soliton for $H < H_c$; (b) N and W Flowers for $H > H_c$ and (c) U Flowers for $H \to \infty$. 
as shown in Figure 3. One of these structures is rich in splay distortion. It is structurally close to the umbilic obtained for diamagnetically negative nematics. The other solution is rich in bend distortion. These are three dimensional analogues of the ‘N’ and ‘W’ walls discussed by Dmitrienko and Belyakov\textsuperscript{11} in smectic C\* in an electric field. We call them “N flower” (Figure 3a) and “W flower” (Figure 3b) respectively. Just as the N wall is of lower energy compared to the W wall the N flower has a lower energy compared to the W flower. However, there is an important difference. The N and W planar solitons (walls) have the same thickness while the thickness of N flower is much greater than that of W flower. As $H \to Hc$ the N flower goes over to the uniform state while the W flower becomes the familiar linear soliton. At the other extreme of $H \to \infty$ both N and W flowers go over to a U Flower.

These “flowers” can exist either in an all radial (bend-splay) or in all circular (bend-twist) configuration depending upon the elastic anisotropy. The bend energy is different in the two cases. We find the interesting result that even when $k_1 < k_3$, a bend-twist configuration can be of lower energy compared to splay-bend, when $k_2$ is small enough. Likewise, when $k_1 > k_3$ it is possible to have a splay-bend structure to be energetically favourable provided $k_2$ is large enough.

All the above results have been obtained by ignoring grain segregation. Grain segregation does not alter very much the above conclusions pertaining to energetics. However, it changes the grain distribution which is given by

$$\varphi = f\tilde{f} = \exp[\rho_0(\cos\theta - \cos\theta_m)H],$$
and the orientation profile. We shown in Figure 4, the effect of grain segregation on N flower. As we increase the field above $H_c$ the core thickness decreases in the beginning and later increases. However, the core of W flower is virtually unaltered by segregation over the same range of field. As the field is increased the grains move out of the core of W flower, while they move into the core of N flower.

From symmetry arguments we can also construct flowers having $\psi = -\alpha$. In the one constant approximation, their behaviour will be exactly like that of N and W flowers. We can call them hyperbolic flowers as they lack cylindrical symmetry but possess hyperbolic symmetry. It is also possible to have topologically a $2\pi$ planar soliton at $H < H_c$. This $2\pi$ soliton becomes a W wall above $H_c$ and goes to a $\pi$ wall as $H \rightarrow \infty$. Above $H_c$ we can construct even an N wall. We find the Dmitrienko and Belyakov solutions of N and W walls having the same thickness if and only if grain segregation is absent. When grain segregation is included the two types of walls are of different thickness with N wall being thicker than W wall. Also, while the thickness of N wall increases with $p_0$ that of W wall remains more or less unaltered. With increase of $H$, N wall thickness decreases to start with and increases later, thus behaving like the N flower. However, the thickness of the W wall decreases with increase in $H$ while exactly the opposite is true for W flower.

It has already been remarked that the classical Frank energy expression is inadequate to workout the structure and energetics of a linear soliton.\textsuperscript{9,10} However, we find, interestingly, that when the orientation at $r = \infty$ is not $\pm \pi$ this difficulty does not arise and the problem is completely solvable.

POINT DEFECTS

In a magnetic field a $\pm 1$ point defect becomes a linear soliton terminating at the singular point.\textsuperscript{8,9} The above calculations appear to indicate that in the case of a diamagnetically negative FN this linear soliton will split into N flower and W flower solitons above a threshold field. The sequence of possible structural changes with increasing $H$ are shown in Figure 5 for a $+1$ point.

TWO COMPONENT SYSTEMS

A simple extension of the above model permits us, to a good approximation, to write the free energy density of a dilute solution (suspension) of solute molecules (grains) with magnetic anisotropy $x_a^m$ in a host nematic of anisotropy $x_n^m$ as

$$F = \frac{k}{2} (\nabla \theta)^2 + \left( x_a^m + f x_a^m \right) \frac{H^2}{2} \cos^2 \theta + \frac{f k_B T}{V} \ln f.$$  

Thus in a magnetic field the solute molecules (grains) get redistributed in the distorted configuration. In Figure 6 is shown the calculated segregation for the U flower configuration.
REFERENCES

Optical diffraction in twisted liquid-crystalline media—phase grating mode

by K. A. SURESH*, P. B. SUNIL KUMAR and G. S. RANGANATH
Raman Research Institute, Bangalore 560080, India

(Received 21 June 1991; accepted 22 July 1991)

We have calculated the diffraction of light perpendicular to the twist axis in a chiral smectic C liquid crystal. In contrast to a cholesteric liquid crystal, in a chiral smectic C liquid crystal we find extra orders which form the odd orders in the diffraction pattern. For an incident linearly polarized light, at a general azimuth, these odd orders are linearly polarized and the even orders are elliptically polarized. The intensities of the odd orders are always independent of the azimuth of the incident light, while this is possible for even orders only at a particular tilt angle of the chiral smectic C liquid crystal. Also, for the incident vibration parallel or perpendicular to the twist axis the odd orders are polarized in the orthogonal linear state with respect to incident vibration, while the even orders are in the same linear state.

1. Introduction

Cholesteric and chiral smectic C liquid crystals exhibit strong reflection bands in the Bragg reflection mode [1]. However, Oldano [2] showed that the chiral smectic C liquid crystal has features very different from that of a classical cholesteric. These new features are not lost even in the uniaxial approximation for the local index ellipsoid. He has shown that the polarization states of the waves propagating through the medium are generally smooth functions of the tilt angle except at an angle $\theta_{\text{rev}}$ at which drastic changes occur. In view of this result, we can except a chiral smectic C liquid crystal to behave very differently from classical cholesterics even in the phase grating mode: propagation perpendicular to the twist axis.

Optical diffraction in heterogeneous media can be understood as a consequence of phase fluctuations. Raman and Nath were the first to solve [3] this problem in the context of ultrasonic diffraction of light. Cholesteric and chiral smectic C liquid crystals are also examples of such media. These can be looked upon as spontaneously twisted nematic and smectic C phases respectively. In cholesterics, the experimentally observed optical diffraction [4] in a direction normal to the helical axis can be treated in an analogous manner [5] to diffraction in a phase grating. This is possible only when the electric vector of the incident light has a component perpendicular to the cholesteric twist axis. On the other hand in the chiral smectic C the medium acts as a phase grating for any azimuth of the incident light and results in a diffraction pattern. To our knowledge, this problem does not appear to have been addressed in the literature. In this paper we consider optical diffraction normal to the twist axis in cholesteric and chiral smectic C liquid crystals. We have also worked out the implications of absorption.

* Author for correspondence.
Our treatment shows that the Raman and Nath theory predictions concerning intensities and phases [6], which are implied in non-absorbing cholesterics [5], are valid only at very low values of layer birefringence. However, for realistic values of birefringence, the intensities and phases of the different orders are very different. Introducing linear dichroism alters the phases and the strengths of the diffraction peaks.

A chiral smectic C liquid crystal behaves very differently in the phase grating mode. Here we get extra orders of diffraction compared to cholesterics which arise due to the pitch, \( P \), being the optical period instead of \( P/2 \) as in cholesterics. These extra orders are surprisingly always linearly polarized and at an angle \((\pi/2) - \phi\) for any azimuth \( \phi \) (with respect to the twist axis) of the incident linearly polarized light. Also, interestingly, the intensities of these orders are independent of \( \phi \). The even orders are generally elliptically polarized and they become linearly polarized only when the incident light is linearly polarized parallel or perpendicular to the twist axis. The intensities of the even orders are sensitive functions of \( \phi \). However, at a particular value of the tilt angle equal to \( \theta_e \), the intensities of these even orders are independent of \( \phi \). The introduction of linear dichroism alters the intensities of the various orders. Here the odd orders continue to be linearly polarized, but the azimuths and ellipticities of the even orders are affected.

2. Theory

2.1. Cholesterics

Considering incident light of amplitude \( A_0 \) linearly polarized at an arbitrary azimuth \( \phi \) with respect to the twist axis, this will have a component \( A_0 \sin \phi \) of the electric vector perpendicular to the cholesteric twist axis. For this component, the refractive index of the medium is a periodic function of the position. In such a case an incident plane wavefront emerges as a periodically corrugated wavefront. As in the Raman and Nath theory we assume the wavelength of the phase fluctuation to be large compared to its amplitude and the wavelength of light. Hence local refractions in the light rays are neglected. Earlier treatment of this problem [5] assumed the refractive index variation to be sinusoidal giving rise to a sinusoidally corrugated wavefront. Under such an approximation the results of the Raman and Nath theory of ultrasonic diffraction of light [3] can be directly employed to obtain the diffraction pattern. However, in reality, the periodic refractive index variation is more complicated and one has to obtain the diffraction pattern by finding the Fourier transform of the periodically corrugated emergent wavefront.

The refractive index \( n_z \) for the vibration perpendicular to the twist axis at any point \( z \) is given by

\[
\frac{1}{n_z^2} = \frac{\sin^2 \alpha}{n_1^2} + \frac{\cos^2 \alpha}{n_2^2}
\]

where

\[
\alpha = \frac{2\pi}{P} z.
\]

\( n_1 \) and \( n_2 \) are the principal refractive indices parallel and perpendicular to the local director, respectively. The emergent wavefront is described by

\[
U(z) = A_0 \sin \phi \exp (i2\pi n_z t/\lambda),
\]
where \( t \) is the thickness of the sample. The diffraction pattern is given by the Fourier transform of \( U(z) \)

\[
F(K) = \int_{-\infty}^{\infty} U(z) \exp(-iKz) \, dz,
\]

where

\[
K = \frac{2\pi \sin \Theta}{\lambda},
\]

\( \Theta \) being the angle of diffraction.

We make use of the fact that \( U(z) \) is a periodic function that can be obtained as a convolution of the periodic array of \( \delta \) functions with \( V(z) \) representing the emergent wavefront for one pitch. Therefore the diffraction pattern \( F(K) \) is nothing but a product of the Fourier transform of the \( \delta \) array with that of \( V(z) \).

For the component parallel to the twist axis, the emergent wavefront continues to be planar suffering an uniform phase retardation of \( 2n_2 t/\lambda \). This does not result in any diffraction and contributes only to the central order which consequently will be elliptically polarized. By taking \( \tilde{n}_1 = n_1 + ik_1 \) and \( \tilde{n}_2 = n_2 + ik_2 \), the effects of the linear dichroism can be easily worked out.

### 2.2. Chiral smectic C

In a chiral smectic C we have a spontaneously twisted smectic C. Here the index ellipsoid spirals about the twist axis at a constant angle \( \Theta \), the tilt angle of the smectic C. This is always a triaxial ellipsoid. In our treatment the incident plane wavefront is assumed to be linearly polarized.

For local biaxiality we use Oldano’s model [2]: one of the principle directions is along the long axis of the molecule. The second one is along the local two-fold axis and the third is perpendicular to these two.

In this case the index tensor \([a] = [e]^{-1}\) is given by

\[
[a] = [a_{ij}],
\]

where

\[
\begin{align*}
    a_{11} &= b_1 \cos^2 \alpha + \sin^2 \alpha (b_2 \sin^2 \theta + b_3 \cos^2 \theta), \\
    a_{12} &= a_{21} = \frac{1}{2} (b_2 - b_3) \sin \alpha \sin 2\theta, \\
    a_{13} &= a_{31} = \frac{1}{2} [b_1 - (b_2 \sin^2 \theta + b_3 \cos^2 \theta)] \sin 2\alpha, \\
    a_{22} &= b_2 \cos^2 \theta + b_3 \sin^2 \theta, \\
    a_{23} &= a_{32} = \frac{1}{2} (b_3 - b_2) \cos \alpha \sin 2\theta, \\
    a_{33} &= b_1 \sin^2 \alpha + (b_2 \sin^2 \theta + b_3 \cos^2 \theta) \cos^2 \alpha,
\end{align*}
\]

where

\[
\begin{align*}
    b_1 &= \frac{1}{n_1^2}, & b_2 &= \frac{1}{n_2^2}, \quad \text{and} \quad b_3 &= \frac{1}{n_3^2},
\end{align*}
\]

\( n_1, n_2 \) and \( n_3 \) are the refractive indices along the local principal axes.

At any layer, the central elliptic section of the index ellipsoid perpendicular to the direction of propagation (x axis) gives the refractive indices \( \mu_1 \) and \( \mu_2 \) for the two
permitted linear orthogonal vibrations along the major and minor axes. These are given by

\[ \mu_1 = [a_{22} \cos^2 \psi + a_{23} \sin 2\psi + a_{33} \sin^2 \psi]^{-1/2}, \]
\[ \mu_2 = [a_{22} \sin^2 \psi - a_{23} \sin 2\psi + a_{33} \cos^2 \psi]^{-1/2}, \]

where \( \psi \) is the angle between one of the principal axes of the central elliptic section and the twist axis (z axis). It is given by

\[ \psi = \frac{1}{2} \tan^{-1} \left( \frac{a_{23}}{(a_{22} - a_{33})} \right). \]

As light travels along any layer it gets resolved into two orthogonal linear vibrations along the major and minor axes of the elliptic section of the layer. However, the orientations of these principal vibrations will periodically vary along the twist axis. When light emerges from the system, we have from each smectic layer, two linear orthogonal vibrations. The azimuths and amplitudes of these vibrations vary periodically along the twist axis. To calculate the diffraction pattern in such problems [7] where the emergent wavefront has polarization fluctuations, we mathematically resolve at every point on the wavefront, the two emergent linear vibrations, along the two chosen orthogonal linear states. Therefore we end up with two periodically corrugated but orthogonally polarized wavefronts given by

\[ U_\parallel(z) = A_\parallel \exp [i\Phi_\parallel(z)], \]
\[ U_\perp(z) = A_\perp \exp [i\Phi_\perp(z)], \]

where \( \Phi_\parallel(z) \) and \( \Phi_\perp(z) \) are the phase fluctuations respectively in the two wavefronts. They result in two diffraction patterns described respectively by their individual Fourier transforms. These are evaluated in the same manner as described previously in cholesterics. The net vibration in any direction is obtained by adding coherently these two diffraction patterns.

If \( \Delta \) is the phase difference between these two vibrations in any order and \( I_1, I_2 \) are their intensities then the ellipticity \( \omega \) and azimuth \( \lambda \) of the resultant vibration are given by the standard formulae [8]

\[ \tan 2\lambda = \cos \Delta \tan A, \]
\[ \tan 2\omega = \tan \Delta \sin 2\lambda, \]

where \( A \) is given by \( \tan A/2 = \sqrt{I_1/I_2} \).

3. Results
3.1. Cholesterics

The earlier study [5] on the diffraction in these systems assumed a sinusoidal fluctuation in the refractive index \( n_z \). This results in the well known Raman and Nath diffraction pattern [3] regarding intensity and phases. This is valid only for small values of layer birefringence. Our treatment yields the following new results:

(i) the intensity ratios do not follow the Bessel function law predicted by Raman and Nath theory;
(ii) the phases of the odd and even orders of diffraction are no longer equal to even and odd multiples of \( \pi \) (or \( \pi/2 \)) respectively [6]; this is due to the non-sinusoidal nature of the corrugated wavefront;
(iii) for high birefringence values, the lower orders are extremely weak in intensity and most of the intensity appears in the higher orders.

We have also worked out the diffraction pattern in absorbing cholesterics ($k$ and $\Delta k \approx 10^{-3}$). Here we find that:

(i) linear dichroism, as can be expected, decreases the overall intensity as compared with that of the non-absorbing case; however, it relatively suppresses certain of the strong diffraction peaks and enhances some of the weak ones;

(ii) at the same mean absorption, an increase in linear dichroism increases the intensities of the various orders.

3.2. Chiral smectic C

Our results very clearly bring out the differences between the chiral smectic C and the cholesteric diffraction patterns.

3.2.1. Odd orders (extra orders)

Since the optical period in a chiral smectic C liquid crystal corresponds to a $2\pi$ rotation of the local director, the diffraction pattern has extra orders compared to cholesterics. These extra orders form the odd orders of diffraction. For a tilt angle $\theta$, in the range $0 < \theta < \pi/2$ these orders are always linearly polarized. For $\phi = 0$ and $\phi = \pi/2$ their azimuths are orthogonal to that of the incident linearly polarized light. For any other value of $\phi$, their azimuths are at $\lambda = (\pi/2) - \phi$ to that of the incident light. In view of this, at $\phi = \pi/4$ the azimuth of the incident linearly polarized light and those of the odd orders are parallel.

We have carried out our calculations both for local uniaxial ($n_1 = n_2 \neq n_3$) and local biaxial ($n_1 \neq n_2 \neq n_3$) symmetries. Figures 1 and 2 depict the diffraction pattern for the uniaxial approximation while figures 3 and 4 show the results for the biaxial approximation. We can clearly see that in both approximations the computed intensities of the different odd orders do not change on changing the azimuth $\phi$ of the incident light. They depend only on the tilt angle $\theta$ of the chiral smectic C phase.

3.2.2. Even orders

These are generally elliptically polarized. However, for $\phi = 0$ or $\phi = \pi/2$, they are linearly polarized parallel to the incident vibration. As shown in figures 1 to 4, the $\phi = 0$ polarization mode incident on the structure with $\theta = \pi/2$ (i.e. cholesterics) does not yield any diffraction pattern in contrast to $\theta$ in the range $0 < \theta < \pi/2$ (i.e. chiral smectic C) which yields a rich diffraction pattern. For a given $\theta$ the different even orders have different intensities, azimuths and ellipticities which are very sensitive to the value of $\phi$. However, we get an intriguing result at a particular tilt angle equal to $\theta_e$ for the chiral smectic C. At this value of $\theta$, the intensities of the various even orders (computed up to the eighth order) are independent of $\phi$. This is depicted in figure 4, second row. For the set of parameters that we used $\theta_e$ turns out to be very nearly $\pi/8$. Generally $\theta_e$ happens to be less than $\pi/4$ and hence experimentally realizable. Our calculations indicate that $\theta_e$ is non-zero for the biaxial approximation and tends to zero for the uniaxial approximation. For any given value of $n_3$, there is an upper limit for $(n_1 - n_2)$ beyond which such a phenomenon does not exist. Also, at a fixed value of $(n_1 - n_2)$, $\theta_e$ increases with a decrease in the value of $n_3$ and vice versa.
Figure 1. Intensities (I) of the various diffraction orders indicated by the integers on the x axis as a function of the tilt angle (θ) of the uniaxial structure and the azimuthal angle (φ) of the incident linearly polarized light computed for \( n_1 = n_2 = 1.535, \ n_3 = 1.605, \ P = 10 \mu m, \ t = 20 \mu m \) and \( \lambda = 633 \) nm. Cholesterics (i.e. \( \theta = \pi/2 \)) exhibit diffraction with only even orders which are linearly polarized perpendicular to the twist axis. However, the central order is linearly polarized only for \( \phi = \pi/2 \). In the chiral smectic C (i.e. \( 0 < \theta < \pi/2 \)) the odd orders are always linearly polarized with the even orders in the elliptic state.
Figure 2. Chiral smectic C diffraction orders indicated by the integers on the x axis (for $n_1 = n_2 = 1.535$, $n_3 = 1.605$, $P = 10 \mu m$, $t = 20 \mu m$ and $\lambda = 633 nm$) for smaller angles of tilt angle $\theta$. 

Optical diffraction in twisted liquid-crystalline media
Figure 3. Intensities (I) of the various diffraction orders shown as integers on the x axis as a function of the tilt angle (θ) of the biaxial structure and the azimuthal angle φ of the incident linearly polarized light computed for n₁ = 1.535, n₂ = 1.545, n₃ = 1.605, P = 10 μm, t = 20 μm and λ = 633 nm.
Figure 4. Chiral smectic C diffraction orders shown as the integers along the x axis (for $n_1 = 1.535, n_2 = 1.545, n_3 = 1.605, P = 10 \mu m, t = 20 \mu m$ and $\lambda = 633$ nm) for smaller angles of tilt angle $\theta$. Notice that for $\theta = \pi/8 (\approx \theta_0)$ the intensities (I) of the even orders are independent of $\phi$. However, as one can see from figures 1 to 4, for any $\theta$, the intensities of the various odd orders are independent of $\phi$. 
3.2.3. Effect of wavelength

The variation in birefringence $\Delta n$ affects the intensities of the various orders. However, variation in the wavelength $\lambda$ affects both the intensities and positions of the different orders.

3.2.4. Effect of absorption

Introduction of linear dichroism ($k$ and $\Delta k \approx 10^{-3}$) redistributes the intensities of the different orders with respect to the non-absorbing case. However, the odd orders continue to be in the same linear polarization state, while those of the even orders are altered. For $\phi = 0$ or $\pi/2$, the even orders are in the same polarization state as the incident vibration and the odd orders are in the orthogonal state.

References


Geometrical theory of diffraction—A historical perspective

P. B. Sunil Kumar and G. S. Ranganath

The geometrical theory of diffraction is a very convenient and easy method of calculating diffraction patterns, and an elegant approach to the problems of Fresnel diffraction at apertures and obstacles. In spite of its long and chequered history, it has not found sufficient emphasis in the standard literature on optical diffraction. Between Young (1802) and Keller (1962), the Indian school (1917-45) led by Raman was active in the field.

In recent years, there has been a revival of interest in the so-called geometrical theory of the diffraction of light\(^1\)\(^{-3}\), consequent on the systematic work of Keller\(^4\). who in the late 1950s developed the detailed methodology for this approach. This concept of treating optical diffraction using geometrical ideas was first introduced by Thomas Young in 1802. In the intervening years, much progress was made and many salient features of the recent ideas in geometric theory were anticipated during the three decades ending in 1945 by the active school led by C. V. Raman. Unfortunately, this work has gone unnoticed in the subsequent, modern literature. It is our intention in this article to give a connected account of the development of the geometric theory bridging the gap between Young and Keller by presenting the work of Gouy, Sommerfeld, Rubinowicz and Raman.

The Helmholtz-Kirchhoff theory of scalar waves presents many computational difficulties in the theoretical calculation of a general diffraction pattern. The procedures become inordinately complex even in the cases of apertures and obstacles having standard geometrical shapes. Over the years, attempts at improving this technique have met with very limited success. It is in this context that the oldest and perhaps the simplest theory, viz. the geometrical theory becomes relevant.

**Diffraction at edges and apertures**

Thomas Young\(^5\) was the first to propose that when light falls on a straight edge, the edge 'reflects' the light into space and the associated interference between the 'edge wave' and the geometrically transmitted wave gives rise to the observed diffraction effect. Gouy\(^6\) in 1886 gave reality to Young's edge waves when he observed that the sharp metallic edge held in a pencil of light appears luminous and the strongly polarized light is diffracted through large angles. Maggi\(^7\) later elaborated Young's model and showed mathematically that the diffraction integral over an aperture can be reduced to a line integral on the boundary of the aperture and a contribution due to the geometrically transmitted light. Sommerfeld\(^8\), who was apparently unaware of this work, independently solved the problem exactly for a straight edge. This theory of electromagnetic diffraction at a straight edge made of perfectly conducting material leads to an interesting result. The field at any point can be looked upon as a sum of the transmitted wave and the wave that appears to emanate from the edge. This edge wave is given by the asymptotic formula

\[
u(r, \phi) = \nu(r, \phi - \phi_0) \pm \nu(r, \phi + \phi_0),
\]

where

\[

v(r, \phi) = \left[ \frac{1 + i}{4\sqrt{k}} \right] \left[ \frac{\sin \phi}{\sqrt{r}} \right] \cos \frac{\phi}{2}

\]

The + or - sign is taken according as the electric vector is parallel or perpendicular to the edge and \(\phi_0\) is the angle of the incident ray and \(\phi\) that of the diffracted ray as measured from the plane of the diffracting screen (see Figure 1). Along the shadow boundary \(r\) diverges since \(\phi = \pi\). (The so-called uniform geometrical theory of diffraction overcomes this lacuna\(^9\), but we do not discuss it further in this article.)

The geometrical theory resolved, for the first time, the apparent puzzle associated with the concept of edge

![Figure 1. Diffraction geometry: parallel rays incident on a straight edge at a glancing angle of \(\phi_0\).](image-url)
diffraction, that the intensity of the bright fringe remains almost constant as the observational plane recedes from the edge. Although the edge wave is cylindrical, its amplitude is dependent on the angle of diffraction. The amplitude increases as the shadow boundary is approached (by a factor $\sec \theta/2$ and the $1/\sqrt{r}$ law for amplitude diminution exactly compensates this.

Sommerfeld's theory also agrees with the Fresnel scalar theory of diffraction for a straight edge. But, it fails to account for the observed diffraction pattern by metallic edges. Raman and Krishnan$^9$ pointed out that this failure is due to the assumption that the material is perfectly conducting. Instead, by incorporating the complex metallic reflection coefficient in the second term of eq. (1) these authors neatly accounted for the experimentally observed features$^9,10$.

Rubinowicz$^{11}$ many years later rediscovered Maggi's result that the Kirchhoff's surface integral over the diffracting aperture in the limit of short wavelengths for an incident spherical wave could be reduced to a line integral over the aperture. He also obtained the result that the diffracted field at any point is made up of two components—(i) the familiar geometrical optical field, and (ii) a wave emitted by the boundary of the aperture. Lau$^{12}$ showed that in Fraunhofer diffraction also, a similar transformation from surface integral to line integral along the boundary is possible. Raman$^{13}$ showed the integral transformation to be a far simpler procedure if one makes the justifiable approximation of ignoring the obliquity factor. It must be remarked that all these procedures are valid only when the size of the diffracting object is large compared to the wavelength of light.

Another important aspect of Rubinowicz's work is that the contour integral can be reduced by the stationary phase method to contributions from a finite number of points on the boundary whose locations depend on the point of observation in the diffraction field. Further, these special points for normal incidence can be easily obtained by drawing perpendiculars to the diffracting boundary from the point of observation. At these points, the incident light ray and the diffracted ray reaching the point of observation satisfy a reflection condition. According to this, when the incident rays are parallel and normal to the diffracting edge, the diffracted light rays reaching any given point of interest are also normal to the edge. Clearly, this answer yields the cylindrical boundary waves for straight edges as shown in Figure 2.b. The feet of the perpendiculars mentioned earlier are the special points which seem to be the source of radiation. Based on Raman's model, Ramachandran$^{14,15}$ also obtained the same result.

From a given point of observation only these points—poles—should be visible and this was experimentally demonstrated by Raman$^{13,16}$, who showed that only a finite number of luminous points are visible on the boundary when viewed from the shadow region. For these special points the total optical path from the source to the point of observation via these points is an extremum. This principle is very reminiscent of the well known Fermat's principle in geometrical optics. For this reason, it has been referred to as Fermat's principle for edge diffraction by Keller. When the incident light rays are parallel but are incident on the edge at an angle then Fermat's principle of diffraction will result in diffracted rays travelling on a cone symmetrical about the local tangent to the edge. Thus, one gets diffraction wavefronts to be parallel cones with the edge as their common axis. This has been depicted in Figure 2.a.

Kathavate$^{17}$ stated that, when dealing with sharp corners of apertures and obstacles, the sharp corners should be taken as additional point sources of light emitting spherical waves. These are in addition to the poles already considered. A decade later Keller$^4$ also suggested the same procedure. However, these workers did not work out the diffraction coefficient. Independently, around the same year, Miyamoto and Wolf$^8$ not only came to the same conclusion but also worked out the corner-diffraction coefficient.

**Diffraction within the shadow**

The geometrical theory of diffraction clearly indicates that the shadow region of an obstacle gets light only from the edge wave. These edge waves will have to be added at any point within the shadow to get the net optical field there. From what has been said in the previous section it follows that we need to take only two types of contributions: (i) from the poles obtained with respect to the point of observation, and (ii) from the corners. The question of the phase of the radiation from the pole was considered by Ramachandran$^{14}$. He showed using the Cornu spiral method that the radiation received at the point of observation from the
regions neighbouring the special points (poles) resulted in a phase advance or a phase lag of n/4 depending upon whether it is one of maximum or minimum optical path from the source to the point of observation via the edge. The same result was obtained many years later by Miyamoto and Wolf. 

**Surface diffraction**

So far we have considered diffraction only at sharp edges. But in reality edges are never perfectly sharp, but are rounded. An extreme example of this was considered by Raman and Krishnan—the Fresnel diffraction by a spherical object. One might naively regard this as equivalent to diffraction by a circular disc. However, in the case of diffraction around metallic spheres, they noticed that the intensity of the central bright spot is always lower than what one observes for a circular disc of equal diameter. Also, the intensity of the central spot is found to be a very sensitive function of the distance from the centre of the sphere. In fact, it exponentially decays below the intensity of the disc spot as the point of observation approaches the sphere. They accounted for this by suggesting that light actually creeps over the spherical surface and the light reaching any point of observation emanates from the circular boundary along the tangent cone drawn to the sphere from the point of observation. They used the exponential law derived by Riemann–Weber for electromagnetic wave propagation around the earth and got a beautiful fit with experimentally observed data.

The more recent work on the geometrical theory of diffraction at smooth surfaces is based on essentially the same mechanism. The detailed theory gives a series expansion for the attenuation coefficient which turns out to be different for the electric vector parallel or perpendicular to the surface. Raman and Krishnan's theory has only the leading term of this series. This is sufficient to account for the experimental data. But one feature which is important in this process of creeping is that the attenuation coefficient is a complex number. Hence when the waves interfere after creeping they have additional phase differences over and above that due to the actual path travelled by light. Keller invokes the generalized Fermat's principle, whereby the actual path from the source to the observer via the surface should be an extremum. This is only possible when light 'creeps' on the surface, travelling along a geodesic on the surface. In fact, for oblique incidence on a cylindrical surface the light creeps along a helix.

**Implications of the theory**

**Diffraction caustics**

An important implication of the geometrical theory of diffraction was stressed by Raman as early as 1919. He argued that for normal incidence the diffracted 'rays' will be predominantly proceeding in the direction of the local normals to the edge of the aperture. Hence, there will be a concentration of light along the evolute (the envelope of the normals to a given curve is defined as its evolute) to the diffracting boundary. Raman also demonstrated this experimentally and called these the diffraction caustics (Figure 3). Of course, the diffraction caustic degenerates into a point in the case of a circular disc, leading to the familiar Poisson spot. A few years later, Coulson and Becknell (for a disc) and Nienhuis did similar experiments and obtained the same results.

**Slits and gratings**

In the case of multiple straight edges as in a slit or an array of slits, the standard procedure is to employ the Fresnel integral or Cornu spiral to work out the diffraction pattern. In the geometrical theory of diffraction, as Raman showed, the diffraction pattern can be obtained by adding the various cylindrical edge waves. This model leads to the well-known answers for a slit or a grating in the Fraunhofer diffraction limit. Keller's recipe to deal with these situations is also essentially the same.

**Semitrasparent edge**

Anathanarayanan studied the diffraction at straight edges of thin films of metals coated on glass. When the metallic coatings were thin enough, he saw fringes of high visibility in the shadow region behind the metallic film. But when the film was thick, he observed in this region, the familiar gradual decay in intensity. He explained this fringe system in the shadow region as a consequence of the interference between the cylindrical edge wave and the wave weakly transmitted by the thin metallic film. When both these waves are of nearly comparable amplitudes the fringe system had a high visibility or contrast.

![Figure 3. Diffraction caustic of an elliptical aperture. (After Raman)](image)
**Fraunhofer diffraction**

A special mention may be made of Raman's studies on Fraunhofer diffraction by triangular and semicircular apertures. Here, the boundary of the diffracting object is replaced by a set of points. Fraunhofer diffraction of an equilateral triangular aperture has a six-fold symmetry and is obtainable from an interference of radiation from three point sources placed at the vertices of the triangle. In the case of a semicircular aperture, Raman argued that in effect we can replace the boundary by three points. One lies on the curved edge its position given by the foot of the perpendicular from the point of observation to the curved edge, and two more respectively at the two corners. This leads to the observed higher symmetry in the Fraunhofer pattern than that of the object.

Again in all their studies on apertures, Raman's school made a special experimental study of pattern transformation as one went from the Fresnel diffraction limit to the Fraunhofer diffraction limit. This is important in view of the fact that Fraunhofer diffraction is centrosymmetric while Fresnel diffraction is not. In Figure 4 we have shown this phenomenon.

**Shadow patterns**

On the experimental side, Kathavate, using objects of a few millimetres and 10 to 50 hours of exposure, got beautiful and intricate diffraction patterns in the case of apertures and discs of various shapes. He came up with a simple and an elegant geometrical procedure, based on the geometrical theory of diffraction, to work out the positions of diffraction maxima (or minima). The whole geometrical construction is carried out on the plane of observation on to which we project the obstacle and the rays from the special points. It is easy to convince oneself that to get the projection of the poles, we just draw perpendiculars to the boundary of the shadow from the point observation. Light from the feet of these perpendiculars must be considered while calculating the positions of maxima or minima. In Figure 5, we show his theoretical calculation along with the observed diffraction pattern for the square disc. This work appears to have gone unnoticed in the literature. Recently English and George have reported the same result. The shadow patterns from elliptical discs of different eccentricities are shown in Figure 6.

**Some new results**

**The Poisson spot**

The bright central Poisson spot in the case of a circular disc is the brightest region of the diffraction pattern in the shadow and it is due to the constructive interference of radiation from the entire boundary. At other points of observation we have only two boundary waves emanating from diametrically opposite poles, leading to periodic weak maxima (and minima). Even in the case of other obstacles like elliptic, square, triangular and rectangular discs, we get such a central spot. It is easy to work out the features associated with this Poisson spot in the language of the geometrical theory of diffraction. For example, in the case of an elliptic disc, the centre of the pattern gets light from four poles—two poles at the ends of the major axis and two poles at the ends of the minor axis of the ellipse. Radiations from the poles of the major axis (or minor axis) are in phase at the centre. But radiations from a pole associated with the major axis may not always be in phase with the radiations from a pole associated with the minor axis. In fact, as we recede from the diffracting plane along the central axis these pole radiations will be successively in and out of phase, giving rise to a brightness fluctuation in the Poisson spot. For a square obstacle, the fluctuations are due to the pole and the corner radiation being in and out of phase. Since corner radiations are weaker as their intensity falls as $1/r^2$, these fluctuations will not be prominent.

The Poisson spot associated with an elliptic disc is, in many ways, different from the one associated with a rectangular obstacle whose length and breadth are respectively equal to the major and minor diameters. If we ignore the corner radiation, then we get four poles...
as in an elliptic disc. Yet the net intensity at the Poisson spot will be different for the elliptic disc due to the curvature at the boundary. This arises owing to the focusing effect of a curved wavefront emitted by a curved boundary. In fact, Keller shows from the geometrical theory of diffraction that a curved boundary contributes more than a straight boundary when the point of observation is towards the centre of curvature. Hence poles of the rectangle make a weaker contribution to the Poisson spot than the poles of an ellipse of 'equal' size. To our knowledge these interesting consequences of geometrical theory of diffraction have not been emphasized in the literature.

**Diffraction at a strip and at a cylinder**

In the shadow region at a finite distance from an opaque strip or a cylinder, experimentally one observes a fringe system. Careful investigations show that this fringe system is strictly not a set of equidistant bright and dark fringes. Nor is the visibility of the fringe system the same all over. In the language of the geometrical theory of diffraction the fringe pattern in the two situations is due to entirely different processes. For a strip it is the interference between the two cylindrical waves from the two straight edges. On the other hand, to reach any point in the shadow of a cylinder, light will have to creep from both sides along the boundary. Thus the interference pattern in general will be different in the two cases. The same arguments are valid even in the case of diffraction at a circular disc and a sphere. However, calculations of the diffraction pattern are easy for the case of a strip and cylinder. At extremely large distances there is very little creeping and the two patterns can be expected to be nearly the
same. In Figure 7 we have compared the calculated diffraction pattern due to a strip and that due to a cylinder of diameter equal to the width of the strip.

**Diffraction symmetry**

We have already touched upon the symmetry of a diffraction pattern in relation to aperture symmetry. Diffraction symmetries are also strongly influenced by polarization. Implications of the geometrical theory of diffraction in this regard will be briefly considered here. In the case of scalar wave diffraction, a circle or a sphere gives rise to a pattern with circular symmetry while a square yields a pattern with a four-fold symmetry. In polarized light, however, we can come to some interesting conclusions concerning these symmetries. Let us say that a linearly polarized wave is incident normally on a square aperture or an obstacle with its electric vector parallel to one pair of edges. Then from eq. (1) we conclude that the cylindrical waves emitted by this pair of edges are not identical to the ones emitted by the other pair, since for this second pair the electric vector is normal to the diffracting edge. Thus the diffraction pattern will not have four-fold symmetry but a two-fold symmetry. However, when the electric vector is parallel to the diagonal of the square the pattern will have four-fold symmetry. Only when the incident vibration is circularly polarized do we have the symmetry obtained for scalar waves. Calculations based on the Keller's theory indicate that these features are noticeable only at a short distance from the screen. At larger distances, i.e. in the paraxial approximation, this symmetry is lost owing to smallness of the second term in eq. (1).

**Multiple-edge radiations**

In another respect Keller improved the geometrical theory of diffraction. We shall illustrate Keller's correction with the example of single-slit diffraction. It was argued earlier that in this case we have two cylindrical waves diverging from the two edges. A wave from one such edge will reach the other edge and will result in a second cylindrical wave. This process could go on endlessly, indicating that each edge gives rise to a multiplicity of edge rays. These have been termed by Keller as second, third, etc. diffracted edge rays. In principle, a complete solution must include the effect of these multiply diffracted rays. However, in practice, these extra effects do not appear to be all that significant, since the strength of the diffracted ray decreases considerably with increasing order.

Acknowledgements. It is pleasure to acknowledge Prof. S. Ramaseshan for introducing us to the geometrical theory of diffraction and for many of his suggestions and hints in our discussions with him. We also thank Prof. Rajaram Nityananda for useful comments.

12. Laue, Von, M., Berliner Sitzungberichte, 1936, 89; see also Jaime, R. W., The Optical Principles of the Diffraction of X-rays, G. Bell and Sons Ltd, 1967.
13. Raman, C. V., Sayaji Rao Gokhawar Found Lectures in Physical Optics; The Indian Academy Sciences, Bangalore, 1959
Geometrical theory of diffraction

P B SUNIL KUMAR and G S RANGANATH
Raman Research Institute, Bangalore 560080, India

MS received 25 September 1991: revised 14 October 1991

Abstract. Geometrical theory of diffraction (GTD) is an alternative model of diffraction propounded first by Thomas Young in 1802. GTD has a long history of nearly 150 years over which many eminent people enriched this model which has now become an accepted tool in the calculation of diffraction patterns. In the conventional Helmholtz-Kirchhoff theory the diffracted field is obtained by computing the net effect of the waves emitted by all points within the area of the aperture. But GTD reduces this problem to one of computing the net effect of waves from a few points on the boundary of the aperture or obstacle, thus simplifying considerably the labour involved in computations. Also the theory can easily be modified to include polarization effects. This has been done specifically by Keller (1962) who exploited the Sommerfeld solution of diffraction of electromagnetic waves at a half plane, making the theory more versatile than the Kirchhoff scalar wave theory. Interestingly the geometry of diffracted rays is predictable from a generalized Fermat principle. According to this the total path chosen by light from the source to the point of observation via the diffracting boundary is an extremum. Historically it should be stated that many of the salient features of GTD were established by a school led by Raman which was active from 1919-1945. Later when Keller (1962) revived GTD independently, he and others who followed him rediscovered many of the results of the Raman school. We have stressed wherever necessary the contributions of the Indian School. We have also discussed certain geometries where GTD can be effectively used. We get some new and interesting results, which can be easily understood on GTD, but are difficult to interpret on the conventional theory of diffraction.

Keywords. Diffraction; Fresnel diffraction; boundary waves.

PACS Nos 42.10; 42.20; 42.30

Table of contents

1. Introduction
2. Boundary wave theory of diffraction
   2.1 Diffraction by a straight wedge
   2.2 Diffraction by a thin edge of arbitrary shape
   2.3 Phase of pole radiations
   2.4 Corner radiation
   2.5 Surface diffraction
3. Generalized Fermat principle
4. Applications of GTD
   4.1 Transparent and semitransparent laminae
   4.2 Apertures and obstacles with straight edges
   4.3 Apertures and obstacles with curved edges
   4.4 Poisson spot
   4.5 Surfaces versus edge diffraction
1. Introduction

Maxwell's electromagnetic theory is undoubtedly the most important contribution to our understanding of light. At one stroke it could easily explain the phenomena of interference, diffraction and polarization of light. In particular, the study of polarization and diffraction got a new impetus. Since light obeys Maxwell's equations, many great physicists from Lord Rayleigh to Sommerfeld tried to get everything about light from these equations. In these attempts both the power as well as the limitations of this approach became quickly obvious. For all their mathematical beauty, the partial differential equations of Maxwell were not easy to solve in every given problem. The practical difficulties forced many to think of other ways of solving some of the formidable problems particularly in the domain of optical diffraction.

The fact that light in many ways behaved like sound was exploited by a few to break new ground in this field. For example, the first mathematical theory of light diffraction is due to Fresnel. He reaffirmed the wave nature of light and gave a theory of half-period zones with which he could calculate diffraction patterns in simple geometries. This gave remarkable agreement with experiments.

Kirchhoff's theory of diffraction of sound waves was extended to the domain of light. This exercise was not without its failures. It leads to many computational difficulties when applied to even slightly complicated geometries. Using Kirchhoff's integral we can in principle calculate the diffraction patterns of apertures and obstacles. The procedure is tedious and time-consuming. In figure 1 we show the diffraction patterns obtained on a computer with this technique, for equilateral triangular apertures at fixed distance from the screen. It clearly brings out the experimentally observed features. Also we do recognize the gradual transition from a Fresnel pattern of three-fold symmetry to a Fraunhofer pattern of six-fold symmetry. But purists were constantly besieged by the fact that it was a scalar wave theory and was thus strictly speaking incapable of handling electromagnetic waves.

It is these technical difficulties that make diffraction problems hard to solve. And wherever one got answers, a physical understanding was still elusive. This became apparent even in one of the earliest problems, viz., diffraction at a circular disk. Though Poisson predicted from Fresnel's theory a bright spot at the centre of the shadow of a circular disk, he refused to believe it and used the result as an argument to reject Fresnel's theory of diffraction. Fortunately, Arago's experimental demonstration of this phenomenon saved Fresnel's theory.
Geometrical theory of diffraction

Figure 1.
Only against such a background we can appreciate the alternative model of diffraction, which was first propounded by Thomas Young (1802). Young's theory evolved out an interesting experimental observation. Even from deep within the shadow, a straight edge appears as a luminous line, and diffraction fringes associated with edges are seen only outside the shadow region. Young argued that these two facts can be explained by assuming that the edge emits waves in all directions. Outside the shadow boundary it interferes with the main geometrically transmitted wave, resulting in a fringe pattern and inside the shadow one gets a monotonic decrease in intensity. Though Young laid the foundations of a new theory of diffraction, it took many decades to work out the details that were necessary for its application to objects of arbitrary shape. It also became very important to make contacts with the Kirchhoff theory in view of its successes in the early years of diffraction theory. In addition, the inclusion of the polarization of the electromagnetic waves was neither easy nor straightforward. One had to appeal to typical solutions of Maxwell's equations to incorporate polarization effects. All this ultimately lead to the establishment of the geometrical theory of diffraction (GTD). In essence it is a recapitulation of Young's idea, viz., diffraction is a manifestation of interference between the directly transmitted light waves and waves emitted by the boundary of the diffracting object.

From the point of view of mathematical computations this theory is quite simple and straightforward. Also it is possible to easily and physically account for many strange and peculiar facts associated with diffraction patterns. For example, the Poisson spot associated with a circular disk is due to a constructive interference at the centre of the waves emanating from the circular boundary. In spite of all these
attractive features, GTD is still an approximate theory and it does not go over to the Fraunhofer diffraction limit. Its power lies in its utility as a tool to get quickly and to a reasonable accuracy the Fresnel diffraction patterns of objects and apertures of arbitrary shapes. It should be emphasized that calculations based on the Kirchhoff theory are very cumbersome since they involve oscillatory integrals. This results in substantial cancellations to the net effect. In GTD these cancellations isolate the points that make the important contributions to the diffraction pattern.

GTD has played a very important role in the computation of fields diffracted from macroscopic objects due to incident radio waves. In many radio engineering problems like antenna design we need to know the strength and structure of the scattered field. Since a solution based on electrodynamic equations is not easily obtainable, GTD techniques come in handy. This has recently been dealt with exhaustively by McNamara et al (1990). In this article, however, we shall not dwell upon these engineering problems but will confine ourselves to situations that arise in the region of wavelengths small compared with the size of the diffracting screen. Also we present some interesting results that have emerged out of the application of GTD to certain optical problems.

It is also of historical importance to notice that many of the essential steps of GTD were rediscovered by many succeeding investigators. In this context particular mention should be made of the contributions of the school lead by Raman. Over a period of nearly thirty years they worked out the important aspects of GTD. But all these investigations appear to have gone unnoticed by the later workers who revived Young’s theory of diffraction. We have emphasized elsewhere the work of the Indian School in this field (Sunil Kumar and Ranganath 1991).

2. Boundary wave theory of diffraction

Young’s model of diffraction proposed in 1802 lay dormant for nearly the next hundred years. This may partly be due to Fresnel’s criticism that Young’s theory was not amenable to quantitative analysis. But in 1888 Maggi showed that the Kirchhoff diffraction integral (which is a superior mathematical formulation of diffraction to Fresnel’s theory of half period zones) may be reduced to a sum of (i) a wave propagating according to the laws of geometrical optics called the geometrical wave, and (ii) a wave originating from every point on the boundary of the obstacle or aperture – called the boundary wave. This is reminiscent of Young’s theory. In 1896, Sommerfeld, without being aware of Maggi’s work, solved the problem of electromagnetic diffraction at a perfectly conducting half plane in the framework of Kirchhoff theory. Interestingly he got a solution similar to the one obtained by Maggi. The field at any point can be looked upon as the sum of a geometrically transmitted wave and a cylindrical wave diverging from the straight edge. Sommerfeld gave a new life to Young’s theory and many followed him in elaborating on this boundary wave theory of diffraction. In view of their significance we shall briefly present here the various important problems solved by different workers.

2.1 Diffraction by a straight wedge

Let Z axis be along the edge of a wedge. The wedge is described by a sector of interior angle \((2\pi - \beta)\) (see figure 2). The incident ray is along \(\phi = \phi_0\). The geometrically
reflected and the directly transmitted rays are respectively along $\phi_r = \pi - \phi_0$ and $\phi_t = \pi + \phi_0$.

Then the electromagnetic field in all space is given by

$$u(r, \phi) = v(r, \phi - \phi_0) \mp v(r, \phi + \phi_0)$$

with negative sign for the electric vector parallel to the edge and the positive sign for the electric vector perpendicular to the edge. The function $v$ can be shown to be

$$v(r, \theta) = v_s(r, \theta) + v_d(r, \theta).$$

Here $v_s$ represents the incident or the reflected wave as given by geometrical optics and $v_d$ represents the diffracted wave originating from the edge of the wedge. At distances $r$ (from the edge) large compared to the wavelength it is given by

$$v_d(r, \theta) = \frac{\pi}{\beta(2\pi kr)^{1/2}} \exp[i(\pi + \pi/4) - \pi/2] \frac{\sin(\pi^2/\beta)}{\cos(\pi^2/\beta) - \cos(\pi\theta/\beta)}$$

provided

$$\left[ \cos\left(\frac{\pi^2}{\beta}\right) - \cos\left(\frac{\pi\theta}{\beta}\right) \right]^2 \gg \frac{1}{kr}. $$

Hence for a half plane ($\beta = 2\pi$)

$$v_d = -\frac{1}{2(2\pi kr)^{1/2}} \exp[i(\pi + \pi/4)] \frac{1}{\cos(\theta/2)} \quad \text{for } \cos\theta/2 \gg \frac{1}{(kr)^{1/2}}$$
showing that the field is continuous at $r = 0$.

We can also calculate using Maxwell’s equations the magnetic field of the diffracted wave. For $kr \gg 1$ we find $H_x = -E_x$ and $H_y = 0$ while for $kr \ll 1$, $H_x$ and $H_y$ diverge like $1/\sqrt{r}$ excepting at $\phi = \pi$ at which $H_x = -\sin \phi_0 \exp(ikr \cos \phi_0)$ and at $\phi = 0, 2\pi$ when $H_y = 0$.

Also it can be shown that for the electric vector parallel to the edge the diffraction field vanishes on the two surfaces of the conducting wedge, i.e., at $\phi = 0$ and $\phi = \beta$.

We can also get directly from Kirchhoff’s integral the structure of the edge radiation diffracted at angle of $\theta$ for a straight edge. This turns out to be

$$v_d = \frac{\sqrt{2}}{(\pi k)^{1/2}} \frac{\cos \theta \exp(ikr)}{\tan \theta \sqrt{r}},$$

which is not the same as the Sommerfeld solution.

The Sommerfeld solution, however, is not valid in the close neighbourhood of the incident shadow boundary and reflection shadow boundary, i.e., $\phi = \pi \pm \phi_0$ in view of the divergence in the diffraction field $v_d$. These difficulties of edge diffraction were overcome in the uniform theory of diffraction which was developed in 1974 by Kouyoumjian and Pathak. They found by an asymptotic analysis that by multiplying $v_d$ by a transition function the diffraction fields can be bounded across the boundaries $\phi = \pi \pm \phi_0$. For a half plane we get

$$v_d(r, \phi) = \frac{\exp(-i\pi/4)}{2(2\pi k)^{1/2}} \left\{ \frac{F[kra(\phi - \phi_0)]}{\cos(\phi - \phi_0)/2} \pm \frac{F[kra(\phi + \phi_0)]}{\cos(\phi + \phi_0)/2} \right\} \left\{ \frac{\exp(ikr)}{\sqrt{r}} \right\}$$

Here

$$a(\phi \pm \phi_0) = 2\cos^2(\phi + \phi_0)/2$$

$$F(x) = \text{Transition function} = 2i\sqrt{x} \exp(ix) \int_{\sqrt{x}}^{\infty} \exp(-i\omega^2)d\omega.$$
Here $\hat{R}$ is the complex reflection coefficient of the metal of which the diffracting screen is made. Interestingly this simple correction completely accounted for the experimentally observed polarization features. Later experimental work of Savornin (1939) further established the validity of this modification.

2.2 Diffraction by a thin edge of arbitrary shape

A similar electrodynamic boundary wave model of diffraction at boundaries of arbitrary shape is yet to be developed. However, much progress has been made in the case of scalar wave diffraction at such boundaries. This is largely due to the fact that the starting point for such models has been the Kirchhoff's theory which can be used for an aperture or an obstacle of any shape. Maggi (1888) was the first to realize that the Kirchhoff's surface integral over the aperture could be reduced to a line integral over the boundary of the aperture together with a contribution corresponding to the geometrically transmitted wave. In fact this work precedes the Sommerfeld analysis but unfortunately remained unnoticed by majority of later workers. Rubinowicz (1917, 1924) independently came to the same conclusions many years later.

With incident spherical waves the Kirchhoff surface integral becomes:

$$v = \frac{1}{4\pi} \int_0^a \left\{ \frac{\exp(ikr)}{r} \frac{\partial \exp(ik\rho)}{\partial n} \frac{\exp(ik\rho)}{\rho} \frac{\partial \exp(ikr)}{\partial n} \right\} d\sigma$$

where $\mathbf{n}$ is the outward surface normal to $\sigma$ the surface spanning the aperture, $r$ and $\rho$ are the distances of the point of integration from the point of observation $P$ and the light source $P_0$ respectively. The surface $\sigma$ together with the diffracting screen will separate the region containing $P$ from that containing $P_0$.

One point to be noticed in Kirchhoff's theory is that the surface $\sigma$ can be entirely arbitrary excepting for the condition that it is limited by the curve 's' which forms the boundary of the aperture. Hence the above integration depends only on 's' and not on $\sigma$. We now consider the cone formed by the rays emitted by the source and the boundary of the aperture. Let $\Sigma$ be this conical surface with surface element $d\Sigma$. Then it is easy to show that for points within the volume defined by $\sigma$ and $\Sigma$

$$v = \frac{\exp(ik\rho)}{\rho} - \frac{1}{4\pi} \int L' d\Sigma$$

and for exterior points

$$v = -\frac{1}{4\pi} \int L' d\Sigma$$

$L'$ is the integrand in the Kirchhoff's integral.

Further the surface integral on $\Sigma$ can be simplified to a line integral on 's' the edge of the aperture. This is given by

$$\frac{1}{4\pi} \int_s \left\{ \frac{\exp(ik\rho)}{\rho} \frac{\exp(ikr)}{r} \frac{\cos \theta_1}{1 + \cos \theta_3} \sin \theta_2 \right\} ds.$$
The first factor gives the wave incident on the edge and the second factor corresponds
to the spherical wave reflected by the edge. Also

\[ \theta_1 = \text{Angle of reflection at the cone surface,} \]
\[ \theta_2 = \text{Angle of incidence at the curved element of the edge,} \]
\[ \theta_3 = \text{Angle of reflection at the edge.} \]

Rubinowicz (1917, 1924) further simplified the contour integral by the method of
stationary phases. A substantial contribution to the line integral comes from only
those points on the contour at which the phase is stationary with respect to a movement
on the curve 's'. The phase factor \([k(r + \rho)]\) remains constant on the contour only when

\[ \frac{dp}{ds} = - \frac{dr}{ds} \]

i.e.,

\[ \cos \theta_2 = - \cos \theta_3. \]

This is often referred to as the reflection condition. Each of such points at which phase
is stationary contributes considerable intensity to the point of observation. The rays
that are diffracted by the edge obeying the above reflection condition have been
depicted in figure 3, for normal and oblique incidence of incident rays. The cone of

![Figure 3](image-url)

Figure 3. Diffracted rays in Rubinowicz reflection condition. (a) Diffracted rays will be on
a cone symmetric about the edge for an oblique incidence, (b) diffracted rays will be on a
disk perpendicular to the edge for normal incidence.
diffracted rays in the case of oblique incidence has been seen experimentally by Maey (1893).

In this context it is important to mention that Raman (1941) obtained essentially the same result by a far simpler procedure. By ignoring the obliquity factor of Kirchhoff's theory, an approximation quite valid in real situations, we get

\[ v = \frac{A}{\lambda} \int_{\sigma} \frac{\exp[i(kr)]}{r} \, d\sigma. \]

Then it can be shown that

\[ v = -\frac{A}{2\pi} \int_{r} \frac{\sin \phi \exp(ikr) \, ds}{r \sin \theta} + \frac{\exp(ik\rho)}{\rho}. \]

Here \( \phi \) is the angle that the line element \( ds \) at \( B \) makes with the plane \( P_0BP \) and \( \theta \) is the angle between the incident ray reaching the line element at \( B \) and the diffracted ray reaching the point of observation from this line element as shown in figure 4. The boundary disturbance can be looked upon as being due to light sources of strength \( (A \sin \phi)/(2\pi \sin \theta) \) having a phase opposite to that of direct light at the boundary. The maximum contribution to \( v \) comes from the line elements for which \( \sin \phi \) is a maximum. Therefore the diffracted radiations principally originate at points on the edge where the line element \( ds \) is perpendicular to the rays reaching the points of observation.

Hence this analysis indicates that at any point of observation we get spherical waves from a finite number of points on the boundary of an aperture. This fact was experimentally demonstrated by Raman (1919). It is necessary to point out that radiations from these special points which Raman referred to as poles are added with appropriate phases at the point of observation. In addition we need also add the geometrically transmitted ray if it also happens to reach the point of observation.

Thus Young’s model of diffraction can be theoretically justified within the framework of the Helmholtz–Kirchhoff scalar wave model. It should be mentioned in

Figure 4. Geometry of diffraction at an aperture.
passing that a similar exercise is also valid for Fraunhofer diffraction. Laue (1936) showed that the Fourier transform integral over the aperture can be reduced to an integral taken round the boundary of the aperture. The Fraunhofer diffraction amplitude is given by

\[ v = \int_{-\infty}^{+\infty} f(x, y) \exp[i(k_x x + k_y y)] \, dx \, dy. \]

If \( f(x, y) = 1 \) over the aperture then

\[ v = \int_{-\infty}^{+\infty} \exp[i(xk_x + yk_y)] \, dx \, dy = \int_{-\infty}^{+\infty} \psi \, dx \, dy. \]

It can be easily shown that

\[ \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + (k_x^2 + k_y^2) \psi = 0. \]

Hence

\[ v = -\frac{1}{k^2} \int_{-\infty}^{+\infty} \nabla^2 \psi \, dx \, dy \]

where

\[ k = (k_x^2 + k_y^2)^{1/2} \]

From Green's theorem in two dimensions, we get

\[ v = -\frac{1}{k^2} \int_{-\infty}^{+\infty} \nabla \psi \cdot ds \]

\[ = -\frac{1}{k^2} \int_{-\infty}^{+\infty} k_n \exp(ik \cdot r) \, ds \]

Here \( k_n \) is the component of \( k \) that lies in the surface and is normal to the line element \( ds \) of the boundary of the aperture.

Laue (1936) used this result to easily account for the observed Fraunhofer diffraction patterns of triangular and hexagonal apertures. Ramachandran (1944) using Raman's method arrived at the same answer. He using the stationary phase approximation further simplified the contour integral to contributions from a finite number of points on the boundary. Using a very similar procedure, Mitra (1920) worked out the Fraunhofer pattern for a semicircular aperture.

2.3 Phases of pole radiations

We may get an impression that the diffraction pattern can be worked out completely once poles where the phases are stationary have been located. But this is far from true. It was pointed out by Ramachandran (1945) that in principle contributions from the points in the immediate neighbourhood of these poles should also be considered. He employed the Cornu spiral method to work out the net contribution from all such neighbouring points. He got a very important result as regards the phases of these pole radiations. Radiation from a pole will have a phase advance of \( \pi/4 \) or
Figure 5. Diffraction at an aperture of an arbitrary shape. The vector $p$ is in the direction of incident light. At point $Q'$ where $\hat{p} = -p$, the vector potential $W$ becomes singular.

phase lag of $\pi/4$ (all these extra phases are with reference to the phase of pole radiation calculated previously) depending upon whether the pole is one of maximum or minimum path with respect to the point of observation.

Miyamoto and Wolf (1962) interestingly arrived at the same conclusion using Kirchhoff's integral as explained below. The geometry is depicted in figure 5. The Kirchhoff's integrand for any incident vibration $u$ is given by

$$I = \frac{1}{4\pi} \left[ u \nabla \left( \frac{\exp(ik\rho)}{\rho} \right) - \frac{\exp(ik\rho)}{\rho} \nabla u \right].$$

Since $\nabla \cdot I$ vanishes we can write $I$ as a curl of a vector potential $W$. For any incident field $u(r') = A(r') \exp[ik\psi(r')]$, we get

$$W(r', r) = \frac{\exp(ik\rho)}{4\pi\rho} \frac{\hat{p} \times p}{1 + \hat{p} \cdot p}$$

where $\hat{p}$ is a unit vector along $\rho$ and $p = \nabla'\psi$ and $r'$ and $r$ are the position vectors of a point on the aperture and the point of observation. Then

$$v_\rho = u^\rho(r) + \Sigma_j F_j(r)$$

with

$$u^\rho(r) = \int W \cdot ds$$
and

\[ F_j(r) = l t \int_{C_j} W \cdot ds \]

\( C_j \)'s being small circles surrounding singularities of \( W \) on the aperture.

It can be shown that \( u^B(r) \) gives the boundary radiation and \( F_j \)'s contribute only when the point of observation is in the illuminated region where they give the direct beam.

When \( u^B(r) \) is evaluated using the stationary phase method we get in the asymptotic approximation

\[ u^B(r) = \frac{\exp(\pm i\pi/4)}{\sqrt{k}} \sum_j u(r_j) \exp(ikp_j) \frac{2\pi}{\partial^2(\rho + \psi)/\partial s^2} \left[ \frac{1}{1 + (\hat{\rho} \cdot \rho) - \frac{1}{2}(\hat{\rho} \times \rho) \cdot ds} \right]^{1/2} \]

We take the positive or the negative sign according as \( \frac{\partial^2(\rho + \psi)}{\partial s^2} \) is positive or negative. This gives essentially the Ramachandran phase \( \pm \pi/4 \) for the poles.

2.4 Corner radiation

In the case of obstacles and apertures having sharp corners it is not enough to work out the pole contribution. The procedure used for locating poles, i.e., the method of stationary phases fails when there are sudden changes in curvature. But contributions from such points cannot be ignored either. Khatavate (1945) concluded on experimental grounds that such sharp points act as point sources emitting spherical waves. It should be remarked that many years later Miyamoto and Wolf (1962) arrived at the same conclusion by using the Kirchhoff's integral. In the neighbourhood of such points we find the corner radiation to be given by*

\[ u_c = \frac{i\lambda \sin \psi \cos(n, r)}{4\pi r} \exp(-ikr) \frac{z^2}{d_1d_2} \]

\( \psi = \) angle between the two local tangents to the boundary at the corner
\( z = \) the perpendicular distance from the point of observations to the source
\( d_1, d_2 = \) components of the vector joining the foot of the perpendicular to the corner.

2.5 Surface diffraction

So far we have considered diffraction at infinitely thin apertures and obstacles. But in reality the lateral thickness cannot be ignored. In examples involving three-dimensional objects like cylinders and spheres this lateral dimension will very significantly alter the diffraction phenomena. Raman and Krishnan (1926) were probably the first to draw attention to the fact that spheres and discs behave differently in diffraction. Similarly diffraction at cylinders is very different from diffraction at strips of equal width.

* We are thankful to Prof. R. Nityananda for discussion on corner radiation.
It is well known that in the shadow of a circular disk we get a bright central spot—the Poisson spot. A very similar bright spot is also seen behind a sphere. But Raman and Krishnan found that the central bright spot obtained in the case of the sphere is always less intense than Poisson spot of a circular disk of equal diameter. Only at very large distances the two intensities are nearly equal. These authors accounted for this observed feature by postulating that light rays creep around the spherical surface, always leaving the surface along the local tangent. Thus at any point on the axis behind the sphere we get light from a circle on the sphere which is the rim of the tangent cone to the sphere from the point of observation. The ratio of the intensity of the central spot of a disk to that of the sphere of equal diameter decreases exponentially to unity as we recede from the sphere. By employing the Riemann–Weber (1927) analysis of electromagnetic waves propagating around the earth, these authors quantitatively explained this exponential decrease in intensity. According to this model the amplitude damping of the electromagnetic wave after it has a creep angle of $\theta$ is given by $\exp[-(0.70\theta)(2\pi R/\lambda)^{1/3}]$.

Many years later Keller (1956) independently attacked the same problem and got a very similar answer. The amplitude damping term is in the form of a series given by

$$\sum_{n=1} (D_n)^2 \exp(-\alpha_n R \theta).$$

With

$$\alpha_n = \frac{q_n}{R} \left( \frac{\pi R}{\lambda} \right)^{1/3} \exp(i\pi/6)$$

$$(D_n)^2 = \left( \frac{\lambda}{4\pi^2} \right)^{1/2} \left( \frac{\pi R}{\lambda} \right)^{1/3} \exp[-i(\pi/12)] \frac{1}{[A'(q_n)]^2}$$

for the electric vector parallel to the surface and for the electric vector perpendicular to the surface

$$\alpha_n = \frac{q_n}{R} \left( \frac{\pi R}{\lambda} \right)^{1/3} \exp(i\pi/6)$$

$$(D_n)^2 = \sqrt{\frac{\lambda}{4\pi^2}} \left( \frac{\pi R}{\lambda} \right)^{1/3} \exp(-i\pi/12) q_n[A(-q_n)]^2.$$  

Here $q_n$ is a zero of the Airy function $Ai(x)$ and $q'_n$ is zeros of the derivative of the Airy function $Ai'(x)$. We easily recognize the Raman and Krishnan solution obtainable as the leading term of the series.

3. Generalized Fermat principle

Keller (1962) suggested an interesting way of unifying all the above different cases. In geometrical optics, the light ray obeys Fermat’s principle choosing in refraction or reflection a path for which $\int \text{d}x$ is an extremum. In a similar way Keller accounted for the above results by postulating that the light ray is always so diffracted that the total path from the source to the diffracting boundary and from there to the point of observation is an extremum. The path of the diffracted ray can thus be easily
worked out. This generalization of Fermat principle naturally leads to cylindrical or conical edge waves at straight edges. At sharp corners it predicts spherical waves. Even light creeping at surfaces is a consequence of this principle. Also this process predicts that on a cylindrical object light will creep on an helix for oblique incidence.

But a complete solution of the problem of diffraction not only needs the paths taken by the diffracted rays but also their amplitudes. Keller appeals to the results of the previous section, in particular the Sommerfeld edge diffraction, to get in each case these strengths of diffracted rays. For example the diffracted rays for a parallel beam incident at an angle \( \alpha \), on a straight edge are given by

\[
U = U_0 D_0 \frac{\exp(ikr)}{\sqrt{r}}
\]

with

\[
D_0 = -\frac{\exp(-i\pi/4)}{2(2\pi k)^{1/2} \sin \alpha} \left\{ \frac{1}{\cos[(\phi - \phi_0)/2]} + \frac{1}{\cos[(\phi + \phi_0)/2]} \right\}
\]

and \( U_0 \) is the incident wave.

The diffracted rays will be travelling on the surface of a cone symmetric about the edge (figure 3). Keller also worked out the diffraction when the edge is not straight but has finite curvature. He argued that a local cylindrical wave will be emitted by the boundary. Thus the emerging wave front will also be locally curved parallel to the edge. From light flux conservation Keller showed that the diffracted rays in such a case are given by

\[
U = D_0 U_0 \frac{\exp(ikr)}{\left[ r \left( 1 + \frac{r}{\rho} \right) \right]^{1/2}}
\]

where \( \rho \) denotes the distance from the edge to the caustic of diffracted rays measured negatively in the direction of propagation. Hence as we cross the caustic \( U \) gets an extra phase of \( \pi/2 \), and the expression itself is not valid at and near the caustic.

We can also easily understand why only a few selected points of the boundary should contribute to the net diffraction field at any given point. The path from the source to the point of observation via the boundary will be an extremum only at these points on the boundary. Hence the generalized Fermat’s principle naturally leads to the notion of pole radiations. It is for these reasons the boundary wave theory of diffraction has come to be known as the Geometrical Theory of Diffraction (GTD).

Thus the problem of diffraction reduces to finding the extremum paths for the diffracted rays and then incorporating the amplitude for each such diffracted rays. At any point of observation we add with appropriate phases all these diffracted rays to get the net vibration.

4. Applications of GTD

It was said earlier that the Kirchhoff-Helmholtz diffraction is not amenable to calculations even in simple geometries. It is in this context that the GTD acquires a new significance. Though in the beginning GTD was considered as a different model
of diffraction, later developments and refinements evolved out of the realization that GTD could be an effective and powerful method in calculating diffraction patterns. However here we will be using GTD not to highlight these features but to account for peculiar features associated with certain diffraction patterns and to suggest solutions in situations where the Kirchoff theory is virtually powerless.

4.1 Transparent and semitransparent laminae

We have already mentioned the work of Raman and Krishnan (1927) on metallic screens wherein they generalized the Sommerfeld solution. Later Raman and Rao (1927) extended the same theory to account for diffraction at thin transparent laminae. This problem naturally arose out of earlier experimental work of Raman and Ghosh (1918) on the diffraction at boundaries of mica and of Raman and Banerji (1921) on the colours of mixed plates. They considered the geometry where light fell normally on the diffracting edge. There will be three regions in the light field: 1. Region of light transmitted by lamina alone; 2. Region of incident light; 3. Region of superposition of incident and reflected light.

They constructed a solution which not only satisfied the wave equation but also smoothly and asymptotically reached the solutions at the previously mentioned regions. Their solutions for the electric vector parallel and perpendicular to the edge respectively are given by

\[
E_z = \frac{i^{3/2} \exp(-i kr)}{4\pi(r/\lambda)^{1/2}} \left[ 1 - (C + iD) \right] \frac{1 - (C + iD)}{\cos(\phi - \phi_0)/2 - \cos(\phi + \phi_0)/2} - \frac{A + iB}{\cos(\phi - \phi_0)/2 - \cos(\phi + \phi_0)/2}
\]

\[
H_z = \frac{i^{3/2} \exp(-i kr)}{4\pi(r/\lambda)^{1/2}} \left[ 1 - (C' + iD') \right] \frac{1 - (C' + iD')}{\cos(\phi - \phi_0)/2 + \cos(\phi + \phi_0)/2} + \frac{A' + iB'}{\cos(\phi - \phi_0)/2 + \cos(\phi + \phi_0)/2}
\]

Here \((A + iB)\) and \((C + iD)\) are the reflection and transmission coefficients of the lamina. In calculating these quantities multiple reflections are taken into account. They are given by (for normal incidence)

\[
A + iB = A' + iB' = \frac{i(\mu^2 - 1)\sin kt}{i(\mu^2 + 1)\sin kt + 2\mu \cos kt}
\]

\[
C + iD = C' + iD' = \frac{2\mu(\cos k_0 t + i\sin k_0 t)}{i(\mu^2 + 1)\sin kt + 2\mu \cos kt}
\]

\(\mu\) = refractive index, \(t\) = lamina thickness, \(k = 2\pi\mu/\lambda, k_0 = 2\pi/\lambda\).

Interestingly the very same solution was arrived at very recently by Burnside and Bergener (1983), who were unaware of the earlier work.

Ananthanarayanan extended the above theory to metallic films by making \(\mu\) complex (Ananthanarayanan 1940). He used this to account for his experimental observations on the diffraction at thin metallic films. In this case he not only found a fringe system in the shadow region, but they were also of better clarity compared to fringes in the light region. He explained this observation by arguing that the cylindrical edge wave emanating from the straight boundary interferences with the wave weakly transmitted by the metallic film. The higher contrast is due to the fact that these two waves are nearly of comparable intensity. On the illuminated side the main beam is far more intense compared to the intensity of the cylindrical edge wave resulting in fringes of low visibility.
4.2 Apertures and obstacles with straight edges

We have already solved the diffraction problem at a straight edge. We now extrapolate the results obtained there to apertures and obstacles bounded by straight edges. When a plane wavefront, i.e., parallel beam of light falls on such an object, each of the straight edges will emit cylindrical waves, i.e., rays radiating out in all directions but normal to the edge. Any point within the shadow gets these cylindrical waves. They will have to be added with strengths (i.e., the diffraction coefficient) given by \( D_0 \) and with appropriate phases to get the net vibration. Since the diffracted rays are normal to the straight edges, any point of observation gets edge radiation only from a few points, i.e., poles on the boundary of the object. These poles are located at the feet of the perpendiculars drawn to the straight edges, from the point of observation. As the point of observation changes the location of the poles also will change. For points in the illuminated region we have to have in addition the main wave as well.

As Kathavate (1945) pointed out, this geometrical exercise can be further simplified to work out the diffraction pattern in these problems. On the plane of observation we locate the geometrical shadow boundary of the object. From the point of observation, which is on this plane, we drop perpendiculars to the shadow boundaries. The feet of these perpendiculars on this projected plane correspond to the projection of poles on the shadow boundary. It is a matter of elementary algebra to work out the path differences between the various pole radiations reaching any point of observation. We get some interesting results in the diffraction of a plane bonded by two straight edges, but at an angle. When this angle is acute every point in the shadow regions gets two pole radiations one from each edge. Hence we have hyperbolic fringes in the area of the geometrical shadow. On the other hand, if the angle between the edges is obtuse, then only in a smaller sector of an acute angle (which is complementary to the obtuse angle) within the shadow we get at any point, radiations from both the edges. Outside this region, but within the shadow boundaries, we have only one side sending out radiations to the point of observation. Hence in such obtuse angle geometries we get hyperbolic fringes inside the inner acute angle sector surrounded by weak field. In the case of a parallelogram, however, we find a small inner parallelogram (defined by the area between the perpendiculars drawn to the four boundaries) within which all the four sides, i.e., four poles, contribute. Then there is a region with only two pole contributions. In addition, near the tips of obtuse angular edges we have only one pole radiation. Hence the central parallelogram shines out with a network of fringes. Kathavate (1945) beautifully demonstrated these features of diffraction in such geometries.

All such objects which are bounded by straight edges, also have sharp corners. We have already mentioned that such corners act independently as sources of spherical waves diverging in all directions. Hence we have to add the cylindrical waves from poles and spherical waves from corners. But in certain regions of the diffraction pattern, we can clearly see the effect of corner waves. In the case of a square sector, near an edge, the pole radiations and corner raditations interfere to give a fringe pattern perpendicular to the edge. In fact Kathavate, who experimentally found these features, suggested corner spherical waves to account for his observations. In figure 6 we give Kathavate’s construction for a rectangular sector. This accounts for the appearance of fringes perpendicular to an edge. Also in certain illuminated regions we find the corner wave to yield circular fringes due to its interference with the main wave. In these regions we have no pole radiations.
Figure 6. Kathavate's geometrical construction for a rectangular sector in the shadow region. The lines parallel to the edges are the lines of equal phase difference for the cylindrical waves from the edges. The circles are the lines of equal phase difference for the spherical wave from the corner. The broken lines perpendicular to the edge indicate the position of the dark fringes due to the interference between the corner waves and the edge waves.

We show in figure 7a the diffraction pattern computed using GTD for an equilateral triangular aperture, with pole radiations only. In figure 7b we show the same but with effects due to corner waves.

4.3 Apertures and obstacles with curved edges

In these geometries an incident plane wavefront falling normally on the object results in a local cylindrical boundary wave which is also laterally curved due to the curvature of the edge. Here also the edge diffracted waves will diverge in directions that are normal to the local tangent. Hence there will be concentration of light along the evolute to the boundary to which these boundary diffracted rays are tangential. In the diffraction pattern this evolute will stand out conspicuously. Raman (1919, 1941) who was probably the first to study them, called them Diffraction Caustics in view of the similarity they bear with normal caustic formation in geometrical optics. Mitra (1919) made a detailed study of these caustics in different cases like apertures and disks with corrugated boundaries. In each case he accounted for their existence by invoking boundary radiations, and argued that the diffraction caustics are to be found along the evolute to the boundary. Coulson and Becknell (1922) arrived at the same
Figure 7. The simulated diffraction pattern for an equilateral triangular aperture of side 0.05 cm at a distance of 100 cm. (a) with pole radiations only, (b) with pole and corner radiations. Sharp steps in intensity in this and the later pictures are an artifact of intensity steps selected during the calculations. In reality, however, there is always a smooth variation in intensity.
conclusions by studying opaque disks of various shapes. Many years later Nienhuis (1948) undertook a similar study for apertures of different shapes and found diffraction caustics to result from boundary radiations. It may be remarked that for circular boundaries the evolute degenerates to a bright point at the centre, first predicted by Poisson.

In addition to these bright lines there are other features associated with the diffraction pattern which can again be easily accounted for by GTD. For example, in the case of a circular disk at points away from the centre but within the shadow, we find two boundary radiations emanating from the diametrically opposite poles on either side. They have to be added with their proper phase differences. While doing this addition we must also include the Ramachandran phase of $\pm \pi/4$ at each pole. Thus we can expect a system of concentric bright circular fringes. On any one such circular fringe the two pole radiations are in phase. In the case of an elliptic disk we find 4 poles to contribute to any point within the evolute while for points beyond the evolute but within a shadow boundary we have two pole radiations. Using Keller's diffraction coefficient $D$ we can get the diffraction pattern for any object with a curved boundary. In figures 8 and 9 we have presented our calculated diffraction pattern using GTD for an elliptic disk and elliptic aperture. We find a network of fringes within the evolute. Our calculations are in qualitative agreement with the observations of Kathavate (1945).

All the above arguments are valid if and only if the curvature changes smoothly over the boundary of the obstacle. If however there are sudden changes in curvature then such points act like additional sources of light emitting spherical waves. The diffraction coefficient is again given by the formula worked out earlier for corners. Sharp points where curved edges meet have the value of the diffraction coefficient same as that obtained for a corner with straight edges.

4.4 Poisson spot

We have already mentioned the existence of a bright central spot in the case of circular disks. This arises due to the constructive interference of the boundary waves at the centre. With the diffraction coefficient $D$ given previously for curved edges we get an infinite intensity at the centre. Keller (1962) has overcome this difficulty by correcting
D which gives good agreement with the exact theory in the far field, i.e., Fraunhofer limit.

In other geometries also we get a central spot whose features can be worked out using GTD. Firstly we get an infinite number of rays reaching such a spot only for circular boundaries. For any other shape of the boundary we find a finite number of points on the boundary contributing to the centre. Hence the central spot in every
other geometry will be weaker than the classical Poisson spot. This accounts for the experimental observation of English and George (1988) who found the central spot in the case of a square disk to be far weaker than what one finds for a circular disk. In this case four poles situated respectively at the centres of the four edges send out cylindrical waves to the centre. In addition, we also have four corner radiations. Thus only eight points on the boundary contribute to the central spot. It is also important to realize that the four pole radiations are in phase at the centre and the four corner radiations are also in phase at the centre. But the pole and corner radiations need not be in phase. The exact phase difference depends upon where we are on the central axis. In principle they will successively be in and out of phase as we recede from the plane of the square disk, thus resulting in fluctuations in the central spot intensity. In practice, however, this may not be conspicuous due to the fact that the corner radiations fall off as $1/r^2$ in intensity while the pole radiations fall off as $1/r$ in intensity.

However, an interesting possibility exists for rectangular and elliptic disks. In both the cases we have four pole radiations contributing to the centre. In the case of the rectangle we also have the corner radiations. But their contributions can be ignored due to their weak strengths. Thus in both the geometries we have four pole radiations. Two of the opposite pole radiations are in phase at the centre. Similarly the other
set of opposite pole radiations are also in phase. But these two pair of pole radiations will be in and out of phase as we recede from the plane of the object resulting in central spot intensity fluctuation. The same phenomenon will be observable at any given point as the ratio of sides (or axes) changes continuously. We now consider a rectangular disk and an elliptic disk of equal dimensions [i.e., length (breadth) of the rectangle is equal to the major (minor) diameter of the ellipse]. Similar poles in the two cases are situated at the same distances from the central axis. Hence we may conclude that the central spot intensity should be identical in the two cases at any given point on the axis. But a careful study of diffraction coefficients will show that this is not the case. For the rectangle we use the diffraction coefficient \( D_0 \) while for the elliptic disk we have to use \( D \) which includes the local radius of curvature. Hence the Poisson spot intensities are quite different in the two cases. Also for the elliptic disk one pair of pole radiations will have \( \pi/2 \) extra phase compared to the other pair of pole radiations since for this pair the central spot will be beyond the radius of curvature at the pole. Thus in many respects the Poisson spots are different in the two cases. In figure 10 we have given calculations pertaining to this comparison between elliptic and rectangular disks.
4.5 Surface versus edge diffraction

Raman and Krishnan (1926) were probably the first to undertake a systematic study of comparing edge waves and waves that creep on surfaces. Their studies were, however, confined to the discussion of Poisson spot behind a circular disk and a sphere of equal diameter. We should not forget that we have in addition in each case a system of circular fringes. In what respects they are different should also be equally interesting. This question can be answered within the framework of GTD.

As an illustration of the phenomena we have taken a strip and a cylinder of equal width as an example for comparison. In both the cases we assume a plane wave front to fall normally on the diffracting object. Experimentally we find in both cases a system of parallel straight fringes inside the geometric shadow. But their origins are very different. In the case of the strip we have two cylindrical edge waves diverging with a strength given by $D_0$ from the two edges. Interference between them results in a fringe pattern. In the case of the cylinder, however, we have waves creeping from both sides and then tangentially leaving the surface to reach the shadow region. Thus any point in the shadow gets two waves that have crept from the two sides of the cylinder. These two waves would have in general crept through different extents.

Figure 11. Fringe system for a cylinder (broken line) of radius 0.1 cm and strip (continuous line) of the same width at a distance 2 cm. Abscissa shows the distance from the centre of the pattern as we move across the shadow. Ordinate depicts the intensity scaled with respect to the intensity at the centre.
Hence the interference pattern obtained in this case would be quite different from that of a strip.

We show in figure 11 results of our calculation of the fringe pattern for the incident electric vector parallel to the edges. We find many interesting differences between the two cases. Firstly fringe spacing and fringe intensities are very different in the two cases. Secondly the visibility of the fringe system is not a constant over the pattern. Also the visibility of the pattern in the case of a cylinder as compared to that of the strip behaves very unexpectedly with distance. This is shown in figure 12. We see the visibility curves to cross over indicating a reversal in the clarity of fringes of a cylinder compared to that in a strip. It should be quite clear from these results that the usual textbook argument that the fringe system in the case of cylinder is due to two line sources placed at its outermost edges is a oversimplified picture. Only at large distances do the two fringe patterns agree.

5. Diffraction symmetry

The geometrical symmetry of an aperture or an obstacle influences strongly the symmetry of the diffraction pattern. In addition, the conditions under which one
studies the diffraction pattern also influences the diffraction symmetry. For example symmetry in the Fraunhofer limit is higher than that one gets in Fresnel limit and is obtained by adding a centre of symmetry. The well known example is that of an equilateral triangular aperture. In the Fresnel limit the three fold symmetry of the aperture gets imposed on the diffraction pattern. Thus the Fresnel pattern is non-centrosymmetric. However, in the Fraunhofer limit the same equilateral triangle results in a diffraction pattern of six-fold symmetry which is centrosymmetric. In fact as a general rule, following from the Fourier transform technique, we can say that the Fraunhofer diffraction is always centrosymmetric. In many examples Raman and his students studied this transition from Fresnel to Fraunhofer pattern as the aperture dimension is decreased. This smooth transition from Fresnel to Fraunhofer pattern can also be qualitatively demonstrated using GTD. In figure 13 we have given our calculated diffraction pattern for an equilateral aperture. The transition from 3-fold to 6-fold is clearly evident. However, it should be emphasized that GTD is strictly not valid in the Fraunhofer limit.

Interestingly the diffraction symmetry is also strongly influenced by the polarization of the incident radiation. This special feature of diffraction patterns is easily obtainable from GTD while a rigorous theoretical calculation could be mind boggling. We shall take a square disk as an example to illustrate this point. When the incident wave is linearly polarized parallel to one pair of opposite edges it will be orthogonal to the other pair of edges. From our expression for $D_0$ we can immediately conclude that the strength $D_0$ is different for the two pairs of opposite edges. Hence the pattern is not strictly 4-fold symmetric but is 2-fold symmetric. This polarization asymmetry is due to the second term in $D_0$ whose sign depends on the polarization of incident vibration relative to the edge. In reality, this term contributes significantly only at very small distances from the diffracting screen. Hence the 2-fold symmetry is realizable only at such distances. It must be remarked that the diffraction symmetry is also sensitive to the azimuth of the incident linear vibration. When it is along the diagonal all the four sides will become equivalent resulting in a 4-fold symmetric diffraction pattern. However, for incident unpolarized light the diffraction pattern will have polarization features which will have a 4-fold symmetry for a square lamina. These features are again calculable from GTD.

6. Multiple edge diffraction

The GTD as presented so far is incomplete. The incident wave results in edge waves and corner waves. These waves again travel towards the other edges and corners to result in a second diffraction. This process could go on and on. Hence from each edge or a corner we have a multiplicity of waves. Keller (1957) has calculated these higher order diffraction coefficients. Calculations based on these higher orders indicate that in most cases of interest these higher order diffraction coefficients are unimportant.

7. Diffraction at thick screens

So far we have discussed diffraction at obstacles and apertures of infinitesimally small thickness. In such cases, in principle, we can calculate the diffraction pattern using
Figure 13. Transition from Fresnel to Fraunhofer pattern for an equilateral triangular aperture as the side decreases. (a) 0.25 cm, (b) 0.2 cm, (c) 0.01 cm at a distance of 100 cm from the aperture.
the Kirchhoff theory. But it is not within the scope of Kirchhoff's theory to yield answers when the screens are of finite thickness. Interestingly, GTD can be effectively used to work out the implications of this lateral thickness of the screen. We shall illustrate this with a few examples.

7.1 A knife edge

In this geometry the diffracting edge has a lateral dimension as shown in figure 14a. It has two sharp straight edges seen as A and B in cross section. When a plane wavefront falls normally on the knife edge with A turned away from the light rays as in figure 14b, then the edge B acts as a source of cylindrical waves. This cylindrical wave will also reach the other edge A which after receiving this wave will act as a secondary source of cylindrical waves. These secondary waves will be of weaker strength. Hence there are three regions in the diffraction pattern. The region beyond the plane AB has only the wave from A while the region between AB and the normal at B to the front surface has waves from A and B. This results in an interference pattern very much like that in a strip. In the region of geometrically transmitted wave, we have in addition to the cylindrical wave from B, a secondary cylindrical wave from A. Thus the diffraction pattern is truly quite complex yet amenable to analysis by GTD. It must be remarked that if the edge A were to face light as in figure 14a the waves from A will not reach the classical geometric shadow. Secondly B will be a source of two cylindrical waves one due to the direct beam and another cylindrical wave excited by the cylindrical wave from A. Similarly A is also a source of a secondary cylindrical wave. Hence region II of the previous geometry is absent here. Also the shadow region will be brighter since two cylindrical waves both starting from B reaches it. In the region of geometrically transmitted light we have two further regions. One as in the previous geometry with two cylindrical waves in addition to the main beam. Second one has in addition to these three contributions another from the wave reflected by the face AB.
The excitation of the secondary cylindrical wave is a very sensitive function of the polarization of incident light. We can clearly see from the Sommerfeld diffraction coefficient $D_0$ that the amplitude of the wave vanishes along AB for the electric vector parallel to the edge. Hence for this polarization we do not get the secondary cylindrical wave.

### 7.2 Round edges

In figure 15 we show a straight edge with a round lateral extension which smoothly goes over to another plane face parallel to the front face of the thick screen. When a plane wave point falls on the face with the flat face towards light, then A acts as a source of cylindrical wave. This wave creeps along the smooth surface and sends light at all angles of diffraction. Hence diffraction pattern will be very similar to that of a thin straight edge excepting for the fact that the shadow region gets light through a creeping of the cylindrical wave. Hence it will be less luminous.

If the object is turned around so that the smooth surface is turned towards light then we have only cylindrical waves diverging from A. It will be exactly like that of a straight edge in the shadow region. But in the region of light we have the direct light, the cylindrical wave from A, reflection by the smooth surface. The pattern will therefore be very complicated but can be easily computed using GTD. For oblique incidence the cylindrical wave at A arises due to creeping of light.

We can use similar arguments in other geometries. The procedure is analogous to the multiple diffraction technique of Keller (1957) which we discussed previously.

### 8. White or polychromatic light diffraction

To get diffraction patterns in white light all that we need to do is to add the intensitites of the diffraction patterns due to the individual components of the incident light. In view of its simplicity one can do this far more quickly with GTD than with Kirchhoff's theory. Also the effects of polarization of incident light can be easily incorporated using GTD.

In figure 16 we show the calculated diffraction pattern of a square disk, with a polychromatic source. It has been compared with the pattern obtained for monochromatic light. In many respects we find a good agreement with the experimental observations of Kathavate (1945) on white light diffraction.

### 9. Effect of finite conductivity

Raman and Krishnan (1927) introduced in a simple way finite conductivity of the diffracting screen. The second term of the Sommerfeld solution gets multiplied by a complex reflection coefficient. This simple modification neatly accounts for the experimental observations in the case of straight edges. The implications of this modification on the nature of the diffraction pattern will be briefly discussed in this section.
Figure 15. Geometry for the diffraction at a thick screen with a smooth surface and a sharp edge.
9.1 Straight edge

As Raman and Krishnan (1927) pointed out, the diffracted light is strongly polarized in this situation. And for a general azimuth of incident linear vibration, diffracted light will be in general elliptically polarized. Another interesting aspect of diffraction pertains to fringe formation. In principle according to this theory the phase of the Sommerfeld wave for the electric vector parallel to the edge is not the same as its phase when the electric vector is perpendicular to the edge. As a consequence the fringe positions will be different in the two cases.

9.2 Poisson spot

In the case of a square disk we pointed out that the central maximum, i.e., the Poisson Spot can exhibit fluctuations in intensity as we recede from the screen. But it was also emphasized that these fluctuations would generally be too weak to be observable due to the weak contribution from corners. If finite conductivity is taken into account, then with a linear vibration parallel to one pair of edges, we find the two pairs of edges wave to diverge out with different phases. Hence due to this alone one must have intensity fluctuations in the central spot.

9.3 Thick screens

In problems involving thick screens we meet situations where the Sommerfeld wave undergoes a second diffraction at another edge or corner. When conductivity is infinite such a processes would be absent for the electric vector parallel to the first edge at which the primary wave was generated. However, in the case of finite conductivity this is not so. Hence the diffraction patterns are in generally different in this case.

10. Conclusions

We have reviewed the basic concepts and results of the geometrical theory of diffraction. This approach to problems involving diffraction is not widely known or
appreciated even though it has many technical advantages over the conventional theories. In addition, it offers simple and elegant explanations for many of the interesting features associated with diffraction patterns. We have applied this technique to a few geometries to bring out the power of this method.

Acknowledgements

We are thankful to Prof. S Ramaseshan for many stimulating discussions we had with him. We also acknowledge Prof. R. Nityananda for helpful suggestions and comments.

References

Ananthanarayanan N 1940 Proc. Indian Acad. Sci. A10 477
Banerji S K 1919 Philos. Mag. 37 112
Burnside W D and Bergener K W 1983 IEEE Trans. 31 104
Coupland J and Becknell G G 1922 Phys. Rev. 20 594
Kulatschikow 1912 J. Russ. Phys. Soc. 44 133
Keller J B 1957 J. Appl. Phys. 28 426
Lau C V 1936 Berliner Sitzungsberichte p. 89 (see also Jame R W 1967 The optical principles of the diffraction of X-rays (G. Bell and Sons Ltd.)
Macy F 1893 Ann. Phys. (Lpz) 49 93
Maggi G A 1888 Ann. di. Mat. 11a 16
McNamara D A, Pistorius C W I and Malherbe J A G 1990 Introduction to uniform geometrical theory of diffraction (Artech House, Norwood) and references therein
Mitra S K 1919 Philos. Mag. 38 289
Mitra S K 1920 Proc. Indian Assoc. Cultiv. Sci. 6 1
Nienhuis K 1948 Thesis, Gröningen
Raman C V 1919 Phys. Rev. 13 259
Raman C V 1941 Sayaji Rao Gaekwar foundation lectures on physical optics (Indian Academy of Sciences)
   Published in 1959
Raman C V and Banerji B N 1921 Philos. Mag. 41 338
Raman C V and Ghosh P N 1918 Nature (London) 102 205
Riemann-Weber 1927 Differential Gleichungen Phys. 2 594
Rubinowicz A 1917 Ann. Phys. 53 257
Rubinowicz A 1924 Ann. Phys. 73 339
Savornin J 1939 Ann. Phys. 11 129
Sunil Kumar P B and Ranganath G S 1991 Curr. Sci. 61 22
Young T 1802 Philos. Trans. R. Soc. (London) 20 26