2. THE PHYSICS OF DIFFRACTION GRATINGS

2.1. THE GRATING EQUATION

When monochromatic light is incident on a grating surface, it is diffracted into discrete directions. We can picture each grating groove as being a very small, slit-shaped source of diffracted light. The light diffracted by each groove combines to form a diffracted wavefront. The usefulness of a grating depends on the fact that there exists a unique set of discrete angles along which, for a given spacing \( d \) between grooves, the diffracted light from each facet is in phase with the light diffracted from any other facet, so they combine constructively.

Diffraction by a grating can be visualized from the geometry in Figure 2-1, which shows a light ray of wavelength \( \lambda \) incident at an angle \( \alpha \) and diffracted by a grating (of groove spacing \( d \), also called the pitch) along angles \( \beta_m \). These angles are measured from the grating normal, which is the dashed line perpendicular to the grating surface at its center. The sign convention for these angles depends on whether the light is diffracted on the same side or the opposite side of the grating as the incident light. In diagram (a), which shows a reflection grating, the angles \( \alpha > 0 \) and \( \beta_m > 0 \) (since they are measured counterclockwise from the grating normal) while the angles \( \beta_0 < 0 \) and \( \beta_{1} < 0 \) (since they are measured clockwise from the grating normal). Diagram (b) shows the case for a transmission grating.

By convention, angles of incidence and diffraction are measured from the grating normal to the beam. This is shown by arrows in the diagrams. In both diagrams, the sign convention for angles is shown by the plus and minus symbols located on either side of the grating normal. For either reflection or transmission gratings, the algebraic signs of two angles differ if they are measured from opposite sides of the grating normal. Other sign conventions...
Diffraction by a plane grating. A beam of monochromatic light of wavelength $\lambda$ is incident on a grating and diffracted along several discrete paths. The triangular grooves come out of the page; the rays lie in the plane of the page. The sign convention for the angles $\alpha$ and $\beta$ is shown by the + and – signs on either side of the grating normal. (a) A reflection grating: the incident and diffracted rays lie on the same side of the grating. (b) A transmission grating: the incident and diffracted rays lie on opposite sides of the grating.

Exist, so care must be taken in calculations to ensure that results are self-consistent. Another illustration of grating diffraction, using wavefronts (surfaces of constant phase), is shown in Figure 2-2. The geometrical path difference between light from adjacent grooves is seen to be $d \sin \alpha + d \sin \beta$. [Since $\beta < 0$, the latter term is actually negative.] The principle of interference dictates that only when this difference equals the wavelength $\lambda$ of the light, or some integral multiple thereof, will the light from adjacent grooves be in phase (leading to constructive interference). At all other angles $\beta$, there will be some measure of destructive interference between the wavelets originating from the groove facets.

These relationships are expressed by the grating equation

$$m\lambda = d (\sin \alpha + \sin \beta), \quad (2-1)$$

which governs the angles of diffraction from a grating of groove spacing $d$. Here $m$ is the diffraction order (or spectral order), which is an integer. For a particular wavelength $\lambda$, all values of $m$ for which $|m\lambda d| < 2$ correspond to physically realizable diffraction orders.

It is sometimes convenient to write the grating equation as

$$Gm\lambda = \sin \alpha + \sin \beta, \quad (2-1')$$

where $G = 1/d$ is the groove frequency or groove density, more commonly called "grooves per millimeter".

Eq (2-1) and its equivalent Eq. (2-1') are the common forms of the grating equation, but their
validity is restricted to cases in which the incident and diffracted rays are perpendicular to the grooves (at
the center of the grating). The vast majority of grating systems fall within this category, which is called
classical (or in-plane) diffraction. If the incident light beam is not perpendicular to the grooves, though, the
grating equation must be modified:

\[ Gm\lambda = \cos \varepsilon (\sin \alpha + \sin \beta), \quad (2-1'') \]

Here \( \varepsilon \) is the angle between the incident light path and the plane perpendicular to the grooves at the grating
center (the plane of the page in Figure 2-2). If the incident light lies in this plane, \( \varepsilon = 0 \) and Eq. (2-1'')
reduces to the more familiar Eq. (2-1'). In geometries for which \( \varepsilon \neq 0 \), the diffracted spectra lie on a cone
rather than in a plane, so such cases are termed conical diffraction.

For a grating of groove spacing \( d \), there is a purely mathematical relationship between the wavelength and
the angles of incidence and diffraction. In a given spectral order \( m \), the different wavelengths of
polychromatic wavefronts incident at angle \( \alpha \) are separated in angle:

\[ \beta (\lambda) = \arcsin (m\lambda d - \sin \alpha). \quad (2-2) \]

When \( m = 0 \), the grating acts as a mirror, and the wavelengths are not separated (\( \beta = -\alpha \) for all \( \lambda \)); this is
called specular reflection or simply the zero order.

---

Figure 2-2. Geometry of diffraction, for planar wavefronts. The terms in the path difference, \( d \sin \alpha \) and \( d \sin \beta \), are shown.

---

A special but common case is that in which the light is diffracted back toward the direction from which it
came (i.e., \( \alpha = \beta \)); this is called the Littrow configuration, for which the grating equation becomes

\[ m\lambda = 2d \sin \alpha, \quad \text{in Littrow.} \quad (2-3) \]

In many applications (such as constant-deviation monochromators), the wavelength \( \lambda \) is changed by
rotating the grating about the axis coincident with its central ruling, with the directions of incident and
diffracted light remaining unchanged. The deviation angle \( 2K \) between the incidence and diffraction
directions (also called the angular deviation) is

\[ 2K = \alpha - \beta = \text{constant} \quad (2-4) \]

while the scan angle \( \Phi \), which is measured from the grating normal to the bisector of the beams, is
2φ = α + β. \hfill (2-5)

Note that \( \Phi \) changes with \( \lambda \) (as do \( \alpha \) and \( \beta \)). In this case, the grating equation can be expressed in terms of \( \Phi \) and the half deviation angle \( K \) as

\[ m\lambda = 2d \cos K \sin \Phi. \] \hfill (2-6)

This version of the grating equation is useful for monochromator mounts (see Chapter 7). Eq. (2-6) shows that the wavelength diffracted by a grating in a monochromator mount is directly proportional to the sine of the angle \( \Phi \) through which the grating rotates, which is the basis for monochromator drives in which a sine bar rotates the grating to scan wavelengths (see Figure 2-3).

### 2.2. DIFFRACTION ORDERS

2.2.1. Existence of Diffraction Orders.

For a particular set of values of the groove spacing \( d \) and the angles \( \alpha \) and \( \beta \), the grating equation (2-1) is satisfied by more than one wavelength. In fact, subject to restrictions discussed below, there may be several discrete wavelengths which, when multiplied by successive integers \( m \), satisfy the condition for constructive interference. The physical significance of this is that the constructive reinforcement of wavelets diffracted by successive grooves merely requires that each ray be retarded (or advanced) in phase with every other; this phase difference must therefore correspond to a real distance (path difference) which equals an integral multiple of the wavelength. This happens, for example, when the path difference is one wavelength, in which case we speak of the positive first diffraction order \((m = 1)\) or the negative first diffraction order \((m = -1)\), depending on whether the rays are advanced or retarded as we move from groove to groove. Similarly, the second order \((m = 2)\) and negative second order \((m = -2)\) are those for which the path difference between rays diffracted from adjacent grooves equals two wavelengths.

The grating equation reveals that only those spectral orders for which \( |m\lambda/d| < 2 \) can exist; otherwise, \(|(m\lambda/d)| > 2\), which is physically meaningless. This restriction prevents light of wavelength \( \lambda \) from being diffracted in more than a finite number of orders. Specular reflection \((m = 0)\) is always possible; that is, the zero order always exists (it simply requires \( \beta = -\alpha \)). In most cases, the grating equation allows light of wavelength \( \lambda \) to be diffracted into both negative and positive orders as well. Explicitly, spectra of all orders \( m \) exist for which

\[-2d < m\lambda < 2d, \quad m \text{ an integer.}\] \hfill (2-7)

For \( \lambda/d << 1 \), a large number of diffracted orders will exist.

As seen from Eq. (2-1), the distinction between negative and positive spectral orders is that

\[ \beta > -\alpha \quad \text{for positive orders } (m > 0), \]

\[ \beta < -\alpha \quad \text{for negative orders } (m < 0). \]
This sign convention for $m$ requires that $m > 0$ if the diffracted ray lies to the left (the counter-clockwise side) of the zero order ($m = 0$), and $m < 0$ if the diffracted ray lies to the right (the clockwise side) of the zero order. This convention is shown graphically in Figure 2-4.

2.2.2. Overlapping of Diffracted Spectra.

The most troublesome aspect of multiple order behavior is that successive spectra overlap, as shown in Figure 2-5. It is evident from the grating equation that, for any grating instrument configuration, the light of wavelength $\lambda$ diffracted in the $m = 1$ order will coincide with the light of wavelength $\lambda/2$ diffracted in the $m = 2$ order, etc., for all $m$ satisfying inequality (2-7). In this example, the red light (600 nm) in the first spectral order will overlap the ultraviolet light (300 nm) in the second order. A detector sensitive at both wavelengths would see both simultaneously. This superposition of wavelengths, which would lead to ambiguous spectroscopic data, is inherent in the grating equation itself and must be prevented by suitable filtering (called order sorting), since the detector cannot generally distinguish between light of different wavelengths incident on it (within its range of sensitivity). [See also Section 2.7 below.]

2.3. DISPERSION [top]

The primary purpose of a diffraction grating is to disperse light spatially by wavelength. A beam of white light incident on a grating will be separated into its component colors upon diffraction from the grating, with each color diffracted along a different direction. Dispersion is a measure of the separation (either angular or spatial) between diffracted light of different wavelengths. Angular dispersion expresses the spectral range per unit angle, and linear resolution expresses the spectral range per unit length.

\[ \beta < -\alpha \quad \text{for negative orders} \ (m < 0), \]
\[ \beta = -\alpha \quad \text{for specular reflection} \ (m = 0), \]  

(2-8)

This sign convention for $m$ requires that $m > 0$ if the diffracted ray lies to the left (the counter-clockwise side) of the zero order ($m = 0$), and $m < 0$ if the diffracted ray lies to the right (the clockwise side) of the zero order. This convention is shown graphically in Figure 2-4.

**Figure 2-4.** Sign convention for the spectral order $m$. In this example $a$ is positive.
2.3.1. Angular dispersion

The angular spread $d\beta$ of a spectrum of order $m$ between the wavelength $\lambda$ and $\lambda + d\lambda$ can be obtained by differentiating the grating equation, assuming the incidence angle $\alpha$ to be constant. The change $D$ in diffraction angle per unit wavelength is therefore

$$D = \frac{\partial \beta}{\partial \lambda} = \frac{m}{d \cos \beta} = \frac{m}{d} \sec \beta = Gm \sec \beta$$  \hspace{1cm} (2-9)

where $\beta$ is given by Eq. (2-2). The ratio $D = d\beta / d\lambda$ is called the angular dispersion. As the groove frequency $G = 1/d$ increases, the angular dispersion increases (meaning that the angular separation between wavelengths increases for a given order $m$).

In Eq. (2-9), it is important to realize that the quantity $m/d$ is not a ratio which may be chosen independently of other parameters; substitution of the grating equation into Eq. (2-9) yields the following general equation for the angular dispersion:

$$D = \frac{\partial \beta}{\partial \lambda} = \frac{\sin \alpha + \sin \beta}{\lambda \cos \beta}$$  \hspace{1cm} (2-10)

For a given wavelength, this shows that the angular dispersion may be considered to be solely a function of the angles of incidence and diffraction. This becomes even more clear when we consider the Littrow configuration ($\alpha = \beta$), in which case Eq. (2-10) reduces to

$$D = \frac{\partial \beta}{\partial \lambda} = \frac{2}{\lambda} \tan \beta. \quad \text{in Littrow.} \hspace{1cm} (2-11)$$

When $\beta \mid$ increases from 10° to 63° in Littrow use, the angular dispersion increases by a factor of ten, regardless of the spectral order or wavelength under consideration. Once $\beta$ has been determined, the choice must be made whether a fine-pitch grating (small $d$) should be used in a low order, or a coarse-pitch grating (large $d$) such as an echelle grating should be used in a high order. [The fine-pitched grating, though, will provide a larger free spectral range; see Section 2.7 below.]
2.3.2. Linear dispersion

For a given diffracted wavelength \( \lambda \) in order \( m \) (which corresponds to an angle of diffraction \( \beta \)), the linear dispersion of a grating system is the product of the angular dispersion \( D \) and the effective focal length \( r'(\beta) \) of the system:

\[
r'D = r' \frac{\partial \beta}{\partial \lambda} = \frac{mr'}{d \cos \beta} = \frac{mr'}{d} \sec \beta = Gmr' \sec \beta. \tag{2-12}
\]

The quantity \( r' \, d\beta = d\lambda \) is the change in position along the spectrum (a real distance, rather than a wavelength). We have written \( r'(\beta) \) for the focal length to show explicitly that it may depend on the diffraction angle \( \beta \) (which, in turn, depends on \( \lambda \)).

The reciprocal linear dispersion, also called the plate factor \( P \), is more often considered; it is simply the reciprocal of \( r'D \), usually measured in nm/mm:

\[
P = \frac{d \cos \beta}{mr'}. \tag{2-12'}
\]

\( P \) is a measure of the change in wavelength (in nm) corresponding to a change in location along the spectrum (in mm). It should be noted that the terminology plate factor is used by some authors to represent the quantity \( 1/\sin \Phi \), where \( \Phi \) is the angle the spectrum makes with the line perpendicular to the diffracted rays (see Figure 2-6); in order to avoid confusion, we call the quantity \( 1/\sin \Phi \) the obliquity factor. When the image plane for a particular wavelength is not perpendicular to the diffracted rays (i.e., when \( \Phi \neq 90^\circ \)), \( P \) must be multiplied by the obliquity factor to obtain the correct reciprocal linear dispersion in the image plane.

2.4. RESOLVING POWER, SPECTRAL RESOLUTION, AND BANDPASS [top]

2.4.1. Resolving power

The resolving power \( R \) of a grating is a measure of its ability to separate adjacent spectral lines of average wavelength \( \lambda \). It is usually expressed as the dimensionless quantity

\[
R = \frac{\lambda}{\Delta \lambda}. \tag{2-13}
\]
Here $\Delta \lambda$ is the *limit of resolution*, the difference in wavelength between two lines of equal intensity that can be distinguished (that is, the peaks of two wavelengths $\lambda_1$ and $\lambda_2$ for which the separation $|\lambda_1 - \lambda_2| < \Delta \lambda$ will be ambiguous). The theoretical resolving power of a planar diffraction grating is given in elementary optics textbooks as

$$R = mN.$$  \hspace{1cm} (2-14)

where $m$ is the diffraction order and $N$ is the total number of grooves illuminated on the surface of the grating. For negative orders ($m < 0$), the absolute value of $R$ is considered.

A more meaningful expression for $R$ is derived below. The grating equation can be used to replace $m$ in Eq. (2-14):

$$R = \frac{N d (\sin \alpha + \sin \beta)}{\lambda}.$$  \hspace{1cm} (2-15)

If the groove spacing $d$ is uniform over the surface of the grating, and if the grating substrate is planar, the quantity $N d$ is simply the ruled width $W$ of the grating, so

$$R = \frac{W (\sin \alpha + \sin \beta)}{\lambda}.$$  \hspace{1cm} (2-16)

As expressed by Eq. (2-16), $R$ is not dependent explicitly on the spectral order or the number of grooves; these parameters are contained within the ruled width and the angles of incidence and diffraction. Since

$$|\sin \alpha + \sin \beta| < 2$$  \hspace{1cm} (2-17)

the maximum attainable resolving power is

$$R_{\text{MAX}} = \frac{2W}{\lambda}$$  \hspace{1cm} (2-18)

regardless of the order $m$ or number of grooves $N$. This maximum condition corresponds to the grazing Littrow configuration, i.e., $\alpha = \beta$ (Littrow), $|\alpha| = 90^\circ$ (grazing).

It is useful to consider the resolving power as being determined by the maximum phase retardation of the extreme rays diffracted from the grating. Measuring the difference in optical path lengths between the rays diffracted from opposite sides of the grating provides the maximum phase retardation; dividing this quantity by the wavelength $\lambda$ of the diffracted light gives the resolving power $R$.

The degree to which the theoretical resolving power is attained depends not only on the angles $\alpha$ and $\beta$, but also on the optical quality of the grating surface, the uniformity of the groove spacing, the quality of the associated optics, and the width of the slits and/or detector elements. Any departure of the diffracted wavefront greater than $\lambda/10$ from a plane (for a plane grating) or from a sphere (for a spherical grating) will result in a loss of resolving power due to aberrations at the image plane. The grating groove spacing must be kept constant to within about 1% of the wavelength at which theoretical performance is desired. Experimental details, such as slit width, air currents, and vibrations can seriously interfere with obtaining optimal results.

The practical resolving power is limited by the spectral half-width of the lines emitted by the source. This explains why systems with resolving powers greater than 500,000 are usually required only in the study of spectral line shapes, Zeeman effects, and line shifts, and are not needed for separating individual spectral lines.

A convenient test of resolving power is to examine the isotopic structure of the mercury emission line at 546.1 nm. Another test for resolving power is to examine the line profile generated in a spectrograph.
or scanning spectrometer when a single mode laser is used as the light source. Line width at half intensity (or other fractions as well) can be used as the criterion. Unfortunately, resolving power measurements are the convoluted result of all optical elements in the system, including the locations and dimensions of the entrance and exit slits and the auxiliary lenses and mirrors, as well as the quality of these optics. Their effects are necessarily superimposed on those of the grating.

2.4.2. Spectral resolution

While resolving power can be considered a characteristic of the grating and the angles at which it is used, the ability to resolve two wavelengths \( \lambda_1 \) and \( \lambda_2 = \lambda_1 + \Delta \lambda \) generally depends not only on the grating but on the dimensions and locations of the entrance and exit slits (or detector elements), the aberrations in the images, and the magnification of the images. The minimum wavelength difference \( \Delta \lambda \) (also called the limit of resolution, or simply resolution) between two wavelengths that can be resolved unambiguously can be determined by convoluting the image of the entrance aperture (at the image plane) with the exit aperture (or detector element). This measure of the ability of a grating system to resolve nearby wavelengths is arguably more relevant than is resolving power, since it takes into account the image effects of the system. While resolving power is a dimensionless quantity, resolution has spectral units (usually nanometers).

2.4.3. Bandpass

The bandpass \( B \) of a spectroscopic system is the wavelength interval of the light that passes through the exit slit (or falls onto a detector element). It is often defined as the difference in wavelengths between the points of half-maximum intensity on either side of an intensity maximum. An estimate for bandpass is the product of the exit slit width \( w' \) and the reciprocal linear dispersion \( P \):

\[
B \approx w' P. \quad (2-19)
\]

An instrument with smaller bandpass can resolve wavelengths that are closer together than an instrument with a larger bandpass. Bandpass can be reduced by decreasing the width of the exit slit (to a certain limit; see Chapter 8), but usually at the expense of decreasing light intensity as well.

Bandpass is sometimes called spectral bandwidth, though some authors assign distinct meanings to these terms.

2.4.4. Resolving power vs. resolution

In the literature, the terms resolving power and resolution are sometimes interchanged. While the word power has a very specific meaning (energy per unit time), the phrase resolving power does not involve power in this way; as suggested by Hutley, though, we may think of resolving power as 'ability to resolve'.

The comments above regarding resolving power and resolution pertain to planar classical gratings used in collimated light (plane waves). The situation is complicated for gratings on concave substrates or with groove patterns consisting of unequally spaced lines, which restrict the usefulness of the previously defined simple formulae, though they may still yield useful approximations. Even in these cases, though, the concept of maximum retardation is still a useful measure of the resolving power.

2.5. FOCAL LENGTH AND f/NUMBER [top]

For gratings (or grating systems) that image as well as diffract light, or disperse light that is not collimated, a focal length may be defined. If the beam diffracted from a grating of a given wavelength \( \lambda \) and order \( m \) converges to a focus, then the distance between this focus and the grating center is the focal length \( r'(\lambda) \). [If the diffracted light is collimated, and then focused by a mirror or lens, the focal length is that of the refocusing mirror or lens and not the distance to the grating.] If the diffracted light is diverging, the focal length may still be defined, although by convention we take it to be negative (indicating that there is a virtual image behind the grating). Similarly, the incident light may diverge toward the grating (so we define
the incidence or entrance slit distance \( r(\lambda) > 0 \) or it may converge toward a focus behind the grating (for which \( r(\lambda) < 0 \)). Usually gratings are used in configurations for which \( r \) does not depend on wavelength (though in such cases \( r' \) usually depends on \( \lambda \)).

In Figure 2-7, a typical concave grating configuration is shown; the monochromatic incident light (of wavelength \( \lambda \)) diverges from a point source at A and is diffracted toward B. Points A and B are distances \( r \) and \( r' \), respectively, from the grating center O. In this figure, both \( r \) and \( r' \) are positive.

Calling the width (or diameter) of the grating (in the dispersion plane) \( W \) allows the input and output \( f/\text{numbers} \) (also called focal ratios) to be defined:

\[
f_{\text{input}} = \frac{r}{W}, \quad f_{\text{output}} = \frac{r'(\lambda)}{W}
\]  

(2-20)

Usually the input \( f/\text{number} \) is matched to the \( f/\text{number} \) of the light cone leaving the entrance optics (e.g., an entrance slit or fiber) in order to use as much of the grating surface for diffraction as possible. This increases the amount of diffracted energy while not overfilling the grating (which would generally contribute to stray light).

For oblique incidence or diffraction, Eqs. (2-20) are often modified by replacing \( W \) with the projected width of the grating:

\[
f_{\text{input}} = \frac{r}{W \cos \alpha}, \quad f_{\text{output}} = \frac{r'(\lambda)}{W \cos \beta}
\]  

(2-21)

These equations account for the reduced width of the grating as seen by the entrance and exit slits; moving toward oblique angles (i.e., increasing \( |\alpha| \) or \( |\beta| \)) decreases the projected width and therefore increases the \( f/\text{number} \).

---

**Figure 2-7.** Geometry for focal distances and focal ratios (\( f/\text{numbers} \)). GN is the grating normal (perpendicular to the grating at its center, O), \( W \) is the width of the grating (its dimension perpendicular to the groove direction, which is out of the page), and A and B are the source and image points, respectively.

The focal length is an important parameter in the design and specification of grating spectrometers, since it governs the overall size of the optical system (unless folding mirrors are used). The ratio between the input and output focal lengths determines the projected width of the entrance slit that must be matched to the exit slit width or detector element size. The \( f/\text{number} \) is also important, as it is generally true that spectral aberrations decrease as \( f/\text{number} \) increases. Unfortunately, increasing the input \( f/\text{number} \) results in the grating subtending a smaller solid angle as seen from the entrance slit; this will reduce the amount of light energy the grating collects and consequently reduce the intensity of the diffracted beams. This trade-off prohibits the formulation of a simple rule for choosing the input and output \( f/\text{numbers} \), so sophisticated design procedures have been developed to minimize aberrations while maximizing collected energy. See Chapter 7 for a discussion of the imaging properties and Chapter 8 for a description of the efficiency...
characteristics of grating systems.

### 2.6. ANAMORPHIC MAGNIFICATION [top]

For a given wavelength $\lambda$, we may consider the ratio of the width of a collimated diffracted beam to that of a collimated incident beam to be a measure of the effective magnification of the grating (see Figure 2-8). From this figure we see that this ratio is

$$\frac{b}{a} = \frac{\cos \beta}{\cos \alpha}$$  \hspace{1cm} (2-22)

Since $\alpha$ and $\beta$ depend on $\lambda$ through the grating equation (2-1), this magnification will vary with wavelength. The ratio $b/a$ is called the anamorphic magnification; for a given wavelength $\lambda$, it depends only on the angular configuration in which the grating is used.

![Figure 2-8. Anamorphic magnification. The ratio b/a of the beam widths equals the anamorphic magnification.](image)

The magnification of an object not located at infinity (so that the incident rays are not collimated) is discussed in Chapter 8.

### 2.7. FREE SPECTRAL RANGE [top]

For a given set of incidence and diffraction angles, the grating equation is satisfied for a different wavelength for each integral diffraction order $m$. Thus light of several wavelengths (each in a different order) will be diffracted along the same direction: light of wavelength $\lambda$ in order $m$ is diffracted along the same direction as light of wavelength $\lambda/2$ in order $2m$, etc.

The range of wavelengths in a given spectral order for which superposition of light from adjacent orders does not occur is called the free spectral range $F\lambda$. It can be calculated directly from its definition: in order $m$, the wavelength of light that diffracts along the direction of $\lambda_j$ in order $m+1$ is $\lambda_j + \Delta\lambda$, where
The concept of free spectral range applies to all gratings capable of operation in more than one diffraction order, but it is particularly important in the case of echelles, because they operate in high orders with correspondingly short free spectral ranges.

Free spectral range and order sorting are intimately related, since grating systems with greater free spectral ranges may have less need for filters (or cross-dispersers) that absorb or diffract light from overlapping spectral orders. This is one reason why first-order applications are widely popular.

The distribution of incident field power of a given wavelength diffracted by a grating into the various spectral order depends on many parameters, including the power and polarization of the incident light, the angles of incidence and diffraction, the (complex) index of refraction of the metal (or glass or dielectric) of the grating, and the groove spacing. A complete treatment of grating efficiency requires the vector formalism of electromagnetic theory (i.e., Maxwell’s equations), which has been studied in detail over the past few decades. While the theory does not yield conclusions easily, certain rules of thumb can be useful in making approximate predictions. The topic of grating efficiency is addressed more fully in Chapter 9.

Recently, computer codes have become commercially available that accurately predict grating efficiency for a wide variety of groove profiles over wide spectral ranges.

2.9. SCATTERED AND STRAY LIGHT [top]

All light that reaches the image plane from anywhere other than the grating, by any means other than diffraction as governed by Eq. (2-1), is called stray light. All components in an optical system contribute stray light, as will any baffles, apertures, and partially reflecting surfaces. Unwanted light originating from the grating itself is often called scattered light.

2.9.1. Scattered light

Of the radiation incident on the surface of a diffraction grating, some will be diffracted according to Eq. (2-1) and some will be absorbed by the grating itself. The remainder is unwanted energy called scattered light. Scattered light may arise from several factors, including imperfections in the shape and spacing of the grooves and roughness on the surface of the grating.

Diffuse scattered light is scattered into the hemisphere in front of the grating surface. It is due mainly to grating surface microroughness. It is the primary cause of scattered light in interference gratings. For monochromatic light incident on a grating, the intensity of diffuse scattered light is higher near the diffraction orders for that wavelength than between the diffracted orders. M.C. Hutley (National Physical Laboratory) found this intensity to be proportional to slit area, and probably proportional to $1/\lambda^4$.

In-plane scatter is unwanted energy in the dispersion plane. Due primarily to random variations in the groove spacing or groove depth, its intensity is directly proportional to slit area and probably inversely proportional to the square of the wavelength.

Ghosts are caused by periodic errors in the groove spacing. Characteristic of ruled gratings,
interference gratings are free from ghosts when properly made.

2.9.2. Instrumental stray light

Stray light for which the grating cannot be blamed is called *instrumental stray light*. Most important is the ever-present light reflected into the zero order, which must be trapped so that it does not contribute to stray light. Light diffracted into other orders may also find its way to the detector and therefore constitute stray light. Diffraction from sharp edges and apertures causes light to propagate along directions other than those predicted by the grating equation. Reflection from instrument chamber walls and mounting hardware also contributes to the redirection of unwanted energy toward the image plane; generally, a smaller instrument chamber presents more significant stray light problems. Light incident on detector elements may be reflected back toward the grating and rediffracted; since the angle of incidence may now be different, light rediffracted along a given direction will generally be of a different wavelength than the light that originally diffracted along the same direction. Baffles, which trap diffracted energy outside the spectrum of interest, are intended to reduce the amount of light in other orders and in other wavelengths, but they may themselves diffract and reflect this light so that it ultimately reaches the image plane.

2.10. SIGNAL-TO-NOISE RATIO (SNR) [top]

The *signal-to-noise ratio* (SNR) is the ratio of diffracted energy to unwanted light energy. While we might be tempted to think that increasing diffraction efficiency will increase SNR, stray light usually plays the limiting role in the achievable SNR for a grating system.

Replicated gratings from ruled master gratings generally have quite high SNRs, though holographic gratings sometimes have even higher SNRs, since they have no ghosts due to periodic errors in groove location and lower interorder stray light.

As SNR is an instrument function, not a property of the grating only, there exist no clear rules of thumb regarding what type of grating will provide higher SNR.