Comprehensive Exam:

DESIGNING AND OPTIMIZING GRATINGS FOR SOFT X-RAY DIFFRACTION EFFICIENCY

By
Mark Boots

Abstract

For almost two hundred years, diffraction gratings have been – and still are – at the heart of many instruments responsible for breakthroughs in scientific understanding. Today, diffraction gratings are used in astronomy telescopes, chemistry spectrographs, spectrophotometers for life science, and in optics for material science experiments, over a range of wavelengths from the far infrared to soft x-rays. The sensitivity and speed of these instruments depends on the efficiency of gratings, i.e., the intensity of useful diffracted light compared to the incident light. In some cases – such as Peter Zeeman's famous discovery of energy level splitting in a magnetic field – improvements in grating efficiency make a previously-undetectable effect detectable.

To support these applications, one component of my graduate work has focussed on the diffraction efficiency of gratings. While this research happens to be applicable to a wide range of scenarios – both theoretical and applied, I have been working toward a very tangible application: designing and optimizing an innovative soft x-ray emission spectrometer for the REIXS beamline at the Canadian Light Source. Over the course of this project, I have created software tools to automate efficiency calculations, used these tools to understand efficiency trends, and applied them to the optical design of the spectrometer. I have also measured the efficiency of real-world gratings and extensively compared these measurements with calculations.

In this comprehensive report, I present the motivation for this project and some background on both gratings and soft x-ray spectroscopy. I also provide a history of developments in grating theory, explain how we selected the theoretical method I used for modelling grating efficiencies, and introduce this theory. (For the sake of length constraints, I set up the foundation of the problem, but omit some mathematical details of the algorithms for now.) In the last chapter, I briefly cover the work that I have already written into my thesis draft, and outline future work to be done.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Motivation: Why grating efficiency matters

In 1821, when Joseph von Fraunhofer first resolved the sodium doublet lines using a diffraction grating he fashioned out of metal wire stretched between the grooves of two screws, he probably would not have anticipated the full scientific impact of his invention. Immediately, these observations helped reinforce Fresnel's new wave theory of light [10]. More importantly, the diffraction grating quickly became the foundation of spectroscopy, superseding the prism as a wavelength-dispersive element with higher resolution, and applicable to radiation from the infrared to x-rays. Eventually it would enable a huge range of experiments and new discoveries in all fields of science:

- In astronomy, Fraunhofer himself was the first to conduct spectroscopic measurements on light from the sun, moon, planets, and stars. Shifts in the position of well-known absorption lines proved, using the Doppler Effect, that the universe was expanding. Today, the composition and temperature of galactic objects is routinely measured using grating-based spectroscopic techniques.[1]

- In physics, visible spectroscopy of the hydrogen emission lines (Balmer Series) provided the data for Niels Bohr's explanation of electron levels in the atom. Peter's discovery of the Zeeman Effect – the splitting of emission lines in a magnetic field – was accomplished using a 20-foot Rowland Circle spectrometer [42]; his results back up the modern version of quantum theory and the magnetic and spin quantum numbers.

- In chemistry, many elements (such as caesium and rubidium, identified by Kirchhoff and Bunsen in 1860) were first discovered in trace amounts using spectral analysis.

- Biologists and biochemists regularly use spectrophotometers to assay the concentration of a tagged reagent in a solution, making gratings a routine (and almost forgotten) tool in life sciences, pharmaceutical, and genetic research.

One common lamentation of the early spectroscopists was the faintness of the light leaving the grating; indeed, we can imagine them in darkened rooms, peering through telescope objectives, straining to make out the faintest lines by eye:

> *Some lines can be distinguished in the spectrum of* Procyon; *but they are seen with difficulty, and so indistinctly that their positions cannot be determined with certainty. I* think *I saw a line at the position D in the orange.*

> Joseph Fraunhofer, in *Prismatic and Diffraction Spectra* [11, p. 61]

Photographic film – and later, modern imaging devices like CCDs – have probably succeeded in removing at least the physical pain associated with spectroscopy. However, increases in grating efficiency are even more important and useful today as they would have been in 1830. At visible wavelengths, more efficient gratings have already enabled a range of compact spectrometers with very high sensitivity, such as the convenient USB-powered computer peripheral in Figure 1.1. Other wavelength ranges are more challenging; grating efficiency is especially critical to the variety of soft x-ray spectroscopy experiments now taking place at synchrotrons around the world, where the low reflectivity of optical materials makes it difficult to build highly efficient devices.

---

[1]In fact, the company that manufactured the gratings for this project also ruled the gratings used in the Hubble Space Telescope.

Figure 1.1: A rather sensitive, compact, visible light spectrometer (Ocean Optics USB4000-UV-VIS). It uses a blazed reflection grating optimized for 300nm light, and connects to a personal computer using USB. **Image credit:** Ocean Optics, Inc. [14]

Simply put, the *diffraction grating efficiency* is the fraction of useful diffracted light outgoing from a grating, relative to the amount of incoming light. (Chapter 3 offers a more formal definition.) While spectroscopy experiments vary in hardware, instrumentation, and purpose, we can make a pair of very general observations on why the grating efficiency is so important. From the point of view of an experimenter, it affects the amount of light available to their sample or detector, and therefore:

1. It affects the *speed* at which experiments can be done, by determining the exposure time required to record data of sufficient quality. In general, improvements in grating efficiency could reduce the amount of time for a given measurement – or equivalently, increase the number of measurements that could be taken in a certain time period. (For example, it takes anywhere from 3 minutes to several hours to measure a soft x-ray emission spectrum on the 8.0.1 beamline at the Advanced Light Source, depending on the concentration of the sample and the desired resolution. One could argue that a factor of 2 optimization in the grating efficiency could almost double the number of users or the scientific throughput of the beamline.)

2. It affects the *feasibility* of doing an experiment in the first place. In situations where the experimental light source is extremely weak (for example, spectral analysis of faint stars in astronomy, or emission line measurements of trace elements in material science) the grating efficiency must be sufficient to raise the signal level above the background noise level seen by the detector. This is no longer a question of patience; if the signal level is below the background noise, our unhappy experimentalist could accumulate detector readings all day (or indefinitely) to no avail.[2]

The motivation for this project was therefore to understand the factors affecting the efficiency of diffraction gratings, and to apply this knowledge to their optimization. To accomplish this, we sought the ability to model gratings numerically and calculate their efficiency – a useful outcome for all grating spectroscopy applications.

However, we had another specific, selfish goal, which was actually our primary objective. At the onset of this project, we were involved in the optical design of a soft x-ray emission spectrometer, destined for use on the REIXS beamline at the Canadian Light Source. Working with David Muir, whose studies on spectrometer resolution are published in his M.Sc. thesis [25], we attempted to simultaneously achieve both *world-class resolution* and *record efficiency* for this machine.[3] Therefore, although the calculation techniques presented in this thesis are general, our examination of trends in diffraction efficiency looks most closely at the types of gratings used in the soft x-ray regime.

Given our specific motivation, the following sections explain the ultimate goal of such a machine, and show how gratings are typically employed in soft x-ray spectroscopy.

---

[2]In fact, Zeeman mentions in the introduction to his paper that he was inspired by Faraday, who spent the last years of his life trying, "but in vain, to detect any change in the lines of the spectrum of a flame when the flame was acted on by a powerful magnet". Zeeman decided "it might be yet worth while to try the experiment again with the excellent auxiliaries of the spectroscopy of the present time..." [42], brought on by Henry Rowland's new mechanically-ruled reflection gratings.

[3]As it turns out, these two goals are implicitly in conflict; see Section 1.2.2.

1.1 Soft X-ray Spectroscopy Techniques

Soft x-rays are photons with energies in the range of approximately 100 to 10 000 eV (or wavelengths of about 10 to 0.1 nm). Unlike with hard x-rays, which are highly penetrating, soft x-ray energies correspond to the binding energies of core-level electrons in common, lightweight elements. This property is ironically responsible for both the experimental challenge of working with them – they are quickly absorbed by any matter over very short distances – as well as their inherent usefulness as a probe of the electronic structure in materials.

Soft x-ray experiments use this light – usually created by a tuneable source such as a synchrotron – and focus it onto a sample to be studied. Two techniques provide complementary information: absorption spectroscopy, and emission spectroscopy.

1.1.1 Absorption spectroscopy

Absorption spectroscopy measures the absorption rate of photons as a function of their wavelength (or energy). Experimentally, this is done by shining a monochromatic beam of light onto the sample and measuring the amount of light absorbed as the energy of the beam is changed.

Figure 1.2 shows the available absorption processes, on the left side in red. When the photon energy increases to become sufficient to excite an electronic transition in the material, dramatically more photons will be absorbed; this is known as an *absorption edge*. Near the edge, as photons are absorbed by exciting electrons from the core level into unoccupied levels, adjacent unoccupied levels will have different probabilities of experiencing a transition, according to the quantum mechanical nature of the bonding in the material. (Note that according to the selection rule for dipole radiation, only electron transitions with a change in orbital angular momentum quantum number $\Delta l = \pm 1$ are allowed, since momentum must be conserved and photons have an angular momentum [spin] of one unit.) The absorption will increase for energies where the transition probability is higher; therefore, the absorption spectrum is actually a measure of the *density of unoccupied states* for electrons in the material.[4]

One outstanding question concerns how the absorption rate is measured – how do we know how many photons were absorbed? Perfect absorption spectroscopy would shine the beam clear through the sample, and measure the intensity of the beam before and after to determine the fraction of light absorbed. While *transmission measurements* like this are feasible with hard x-rays, the short attenuation length of soft x-rays would require prohibitively thin samples for any beam to be left on the other side. Instead, different measurements are used as a proxy for the total absorption rate:

Total Electron Yield

For the excited electron, the most probable decay mechanism is to quickly relax into a lower-energy state, transferring its energy to a more loosely bound electron in the process. This secondary electron, known as an Auger electron, can then be ejected from the sample – assuming it is close enough to the surface. The **total electron yield** (TEY) method determines the absorption rate by measuring the electric current that must flow into the sample to neutralize the ejected electrons and keep the sample uncharged. (Experimentally, this is done by simply connecting a wire to the sample holder, and placing a sensitive ammeter along the path to a solid ground connection.)

TEY measurements are difficult for some samples, either because an insulating sample doesn't allow current to flow in to replenish the ejected electrons ("sample charging"), or because electrons ejected deep in the material are reabsorbed elsewhere. This makes TEY measurements most sensitive to absorption events near the surface ($\sim$2nm), and restricts them to conductive samples.

Total Fluorescence Yield

Total Fluorescence Yield (TFY) measurements overcome these problems by using a light-sensitive detector near the sample. Although several orders of magnitude less probable than the Auger decay process, excited

---

[4]To be accurate, we should say the '*partial* density of unoccupied states', since the vacancy left behind in the original electron state ("core-hole") will affect the energy of the unoccupied states.
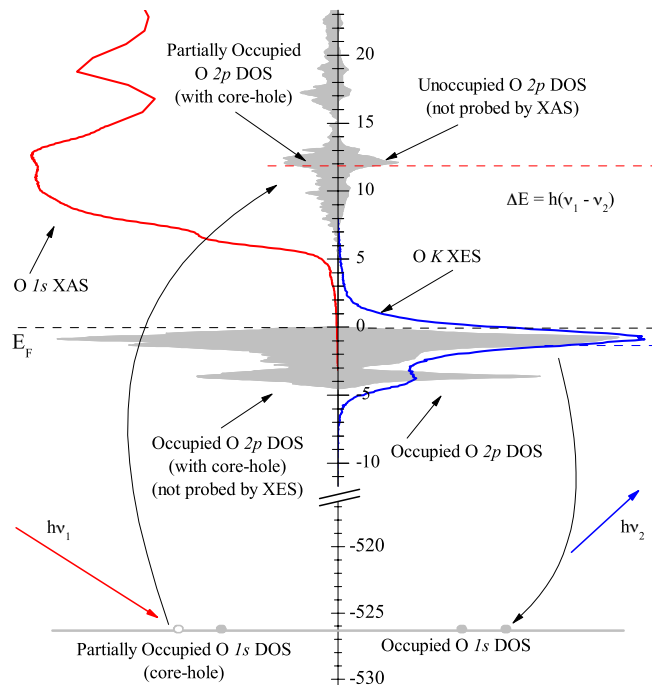
Figure 1.2: X-ray absorption spectroscopy (on the left, in red) probes the density of *unoccupied* electronic states, modified by the presence of a core-hole: a vacancy left behind by the excited electron. In this example for oxygen, the absorption edge starts at 6 eV above the Fermi level, labelled E_F. X-ray emission spectroscopy (on the right, in blue) probes the density of *occupied* valence states, as a valence electron decays by emitting a photon to fill the core-hole. In this diagram, the absorption and emission spectrum have been plotted vertically to line up with the schematic of the energy levels; they are conventionally plotted on a horizontal energy axis. Grey areas represent the computed density of states. **Image from:** Ref. [21]

states can also relax by emission of a photon. Instead of ejected electrons, TFY measures the intensity of all photons emitted during the decay, which makes it applicable to both insulating and non-insulating samples. Because photons have a greater escape depth than electrons, this technique is also able to probe deeper within a material than TEY can. Due to the low probability of fluorescence transitions compared to Auger transitions (several orders of magnitude), TFY measurements benefit greatly from concentrated samples and a light source which is capable of high intensity.

1.1.2   Emission spectroscopy

While absorption measurements provide information about the *unoccupied* states, we can also study what happens *after* the initial photon is absorbed. When a core-level electron is promoted by the absorption of a photon, it leaves behind a "core-hole", and the atom (or molecule, or crystal) is left in an excited state. While there are many ways for the system to collapse back to the ground state, there is a small probability that some valence-band electron will decay to fill the core-hole by the emission of another photon. Since the energy of the emitted photon will match the energy difference between that valence electron and the core level, the intensity distribution of *all emitted light* will correspond to the probability of finding electrons in the valence band at those energies. Therefore, if we could collect the *fluorescence* emitted from the sample and plot its intensity as a function of energy, we would have a measure of the *density of occupied states* for electrons in the material. In this way, emission spectroscopy provides information on the bound electronic states, which is not present in the absorption spectrum.

Experimentally, XES measurements are done by illuminating the sample with a fixed photon energy above the absorption edge. The fluorescence is captured using an energy- (or wavelength-)sensitive detector which

is tuned to the energy range just below the absorption edge. Over time, an intensity spectrum is built up from the collected photons (Figure 1.2, Figure 1.7). Since fluorescence transitions are a random process, and highly improbable compared to other decay mechanisms like Auger decay, XES measurements require a sufficient exposure time to build up good statistics. They also benefit greatly from a high-intensity beam source and an efficient detector – such as a sensitive spectrometer with high-efficiency gratings.

Resonant Inelastic X-ray Scattering (RIXS)

An advanced form of XES is known as RIXS (*Resonant Inelastic X-ray Scattering*). Instead of exciting a sample with a photon energy well above the absorption edge, the energy of the beam is tuned to match transitions previously identified in the absorption spectrum (or stepped incrementally through this range). This allows the experimenter to preferentially excite into specific electronic states: at *resonance*, the transition probability will be extremely high because the exciting photon energy exactly matches the transition energy and the emitted photon energy.

RIXS is a one-step process, but it can be explained mathematically as a simultaneous two-step process which combines absorption and emission: from the initial electronic state, an incoming photon is absorbed, creating a "core-hole" and an excited electronic configuration. This intermediate state decays by the emission of another photon into a final state, and this transition doesn't necessarily need to involve the original electron. The difference in energy of the emitted and incident photons provides an *energy-loss spectrum* describing the nature of excitations within the material.

According to this two-step explanation, RIXS can probe transitions that would be forbidden by the dipole selection rule. For example, a 2p electron could be excited into a 3d state, and another 3d electron with a slightly different energy could collapse to fill the 2p hole, thereby effectively creating of a 'd-d' excitation [6].

1.1.3   Importance of Soft X-ray Spectroscopy (SXS)

Soft x-ray spectroscopy is a valuable tool in material science for its ability to gain insight into the electron environment within a material. Since the electronic structure determines the bonding between atoms and thereby a material's mechanical, chemical, and physical properties, this is a big deal indeed. Additionally, x-ray spectroscopy has a few complementary advantages over other analytical techniques like neutron diffraction and photoemission spectroscopy:

- It is **element-specific**: for a material containing a number of elements, it is usually possible to find absorption edges for each element that do not overlap with the others, making it possible to independently probe the bonding environment of each. For example, in an organic sample containing carbon, nitrogen, and oxygen, one can probe the oxygen using the O 1s absorption at 543 eV, and then separately excite the nitrogen 1s electrons at 410 eV.

- More than being just element-specific, it is also **site-specific**: soft x-ray spectra make it possible to distinguish, for example, single-bonded carbon atoms at one location in a molecule from double-bonded atoms at another.

- Depending on the detection technique, it can be **surface-sensitive** or **bulk-sensitive** (i.e.: it can probe through surface contamination or oxidation, testing the nature of the material below).

Additionally, there are a few experimental considerations which make SXS desirable:

- It can be done **non-destructively** on whole samples, without having to crush them into powders, dilute them in solution, etc. It can also allow in-situ measurements of samples created directly within the vacuum chamber – for example, crystal samples created using sputtering or vapour deposition techniques.

- Because of the high brightness and small size of synchrotron beams, it can be done on **very small and thin samples**. Since the penetration depth of soft x-rays is so short, the interaction volume created by the beam will be tiny regardless of the sample thickness, making it a very useful characterization technique for thin film samples and even monolayers.

One obvious experimental disadvantage is that these techniques require access to synchrotron accelerators to produce the soft x-ray beam, which – for the foreseeable future – are not yet available in convenient desktop or bench-top models. Additionally, since soft x-ray experiments must be performed under ultra-high vacuum (UHV) conditions (see section 1.2.3), samples must either be UHV-compatible or carefully isolated from the vacuum environment behind thin windows.

Despite these limitations, SXS techniques have created some notable and very interesting discoveries in physics; some highlights are listed here:

- In 1990, de Groot et. al. presented the first comprehensive understanding of soft x-ray absorption in transition metal compounds. Not only did they produce some state-of-the-art experimental spectra for the time, but they also gave a thorough interpretation of them using crystal field atomic-multiplet theory [7].

- Butorin et. al. published the first RIXS studies of transition metals, and discovered d-d excitations in manganese oxide [6].

- Recently, Braicovich et. al. used RIXS to measure the dispersion of magnetic excitations in cuprate superconductors. This study also found a magnetic dispersion branch that had never been found before using neutron scattering, and found that these types of materials are in non-homogenous spin states, revealing a bit more about the mysterious nature of cuprate superconductors [4].

- With the proper experimental equipment, RIXS studies can also be done on gaseous samples. In 2011, Pietzsch et. al. measured very high resolution RIXS on oxygen gas (O_2) to observe the vibronic structure. The exciting result from this study was the presence of spatial quantum beats in their spectra – essentially, an observation of quantum mechanical interference like the famous double-slit experiment, but using excitations into different states instead of transmission through different slits [32].

1.2 Spectroscopy Instrumentation

The preceding sections make it clear that to perform soft x-ray spectroscopy experiments, we need three capabilities:

1. Obviously, one needs a source of soft x-rays – typically, the brighter, the better. This became possible with the advent of of synchrotron particle accelerator facilities, which emit broad-spectrum x-rays as relativistic electrons are forced to change their path by bending magnets.[5]

2. **For absorption and emission spectroscopy:**
   From this broad spectrum of light, one needs to produce a nearly monochromatic beam of light to shine onto the sample, and must be able to adjust the energy of this beam. This is the role of a *monochromator*, which takes a broad-spectrum light source and extracts a small range of wavelengths from it (Figure 1.3).

3. **Additionally, for emission spectroscopy:**
   One must be able to capture the light emitted from the sample and resolve it by wavelength. The end goal is to measure the relative intensity as a function of wavelength (or energy); this is accomplished using a *spectrometer* (Figure 1.5).

---

[5]Much more intense synchrotron light can be generated by "insertion devices" known as *undulators* and *wigglers*, in which an alternating array of magnets in a straight section of the accelerator forces the electron beam to bend many times over a distance of a few meters. Although light from insertion devices still contains of a range of wavelengths, its spectrum consists of sharp intensity peaks which can be adjusted to the desired wavelength by changing the strength of the magnetic field. We intentionally avoid going into too much detail on synchrotron physics in this thesis; more information can be found in Ref. [27].
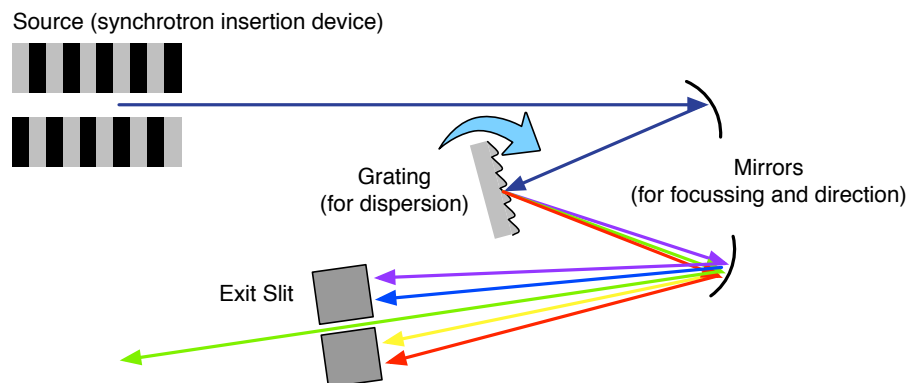
Figure 1.3: In this schematic of a grating monochromator, light from the source is focussed by mirrors and dispersed by the grating. An exit slit picks out the desired wavelength or energy range, and blocks the remaining light. Depending on the design, the output wavelength can be adjusted by changing the angle of the grating, the angle of the mirrors, and/or the position of the exit slit. (The resolution – the energy bandwidth of the outgoing light – depends on the dispersion of the grating, the geometry, and the size of the exit slit. An ideal monochromator would produce truly monochromatic light, but this would require an infinitely small slit.)
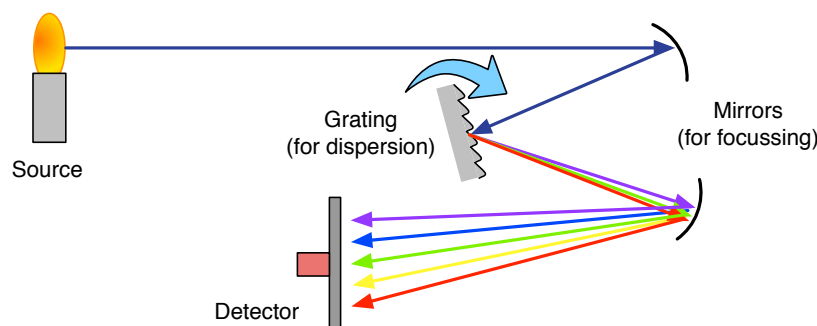
Figure 1.4: In this schematic of a grating spectrometer, mirrors are used to focus light from the source (or entrance slit) onto the detector, passing over a plane grating en-route to disperse the light by wavelength. (This is a *Czerny-Turner* arrangement, typical of compact visible-light designs like the one in Figure 1.1.) Since the grating diffracts different wavelengths at different angles, the intensity distribution across the detector surface creates a spectrum related to the wavelength (or photon energy).
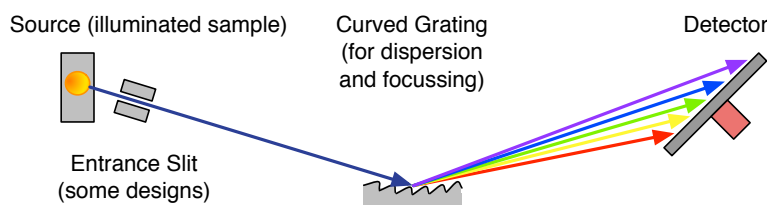
Figure 1.5: The spectrometer in this schematic uses a curved grating to both disperse light by wavelength, and focus it onto the detector. (This is typical of soft x-ray designs where the poor reflectivity of additional mirrors is usually avoided.) The intensity distribution along the surface of the detector can be converted into a spectrum with respect to wavelength or energy. Since the detector only captures a limited range of outgoing angles, it can be moved to pick out the desired wavelength window for each measurement.

1.2.1 Beamline optics: Monochromators and Spectrometers

Monochromators

The role of a monochromator (shown schematically in Figure 1.3) is to pick out a narrow range of wavelengths from a chromatic light source, and deliver the monochromatic beam to the experiment. For infrared light out to soft x-rays, diffraction gratings provide the most efficient way of separating the incoming beam based on wavelength.[6] When the incidence angle and mounting angle of the grating are chosen to direct a single ($n \neq 0$) diffraction order toward the exit slit, the wavelength term in the grating equation (equation 2.61)

$$n\lambda/d = \sin\theta_{2,n} - \sin\theta_2$$

creates a dependence on the sine of the outgoing angle $\theta_{2,n}$, so that shorter wavelengths leave more normal, and longer wavelengths leave at more grazing angles. Depending on the optical and mechanical design, the output wavelength can be selected by changing the angle of the grating, the angle of the mirrors (and hence the incidence angle onto the grating), and/or the position of the exit slit.

The resolution of the monochromator – ie: the bandwidth of wavelengths present in the output light – depends on the angular dispersion of the grating, the geometry, and the size of the exit slit; an ideal device would obviously produce a perfectly monochromatic beam, but this would require an infinitely-small exit slit. (In practice, the size of the exit slit is used to adjust the resolution, in an unavoidable trade-off against the amount of flux produced.) Resolution is measured as the energy bandwidth ΔE (full width at half-maximum) for a given central energy; for example: $\Delta E = 500$ meV at 1000 eV. Often it is more convenient to normalize it as the *resolving power* $RP = E/\Delta E$, which has the benefit of being identical when measured in wavelength as well:

$$RP = \frac{E}{\Delta E} = \frac{hc/\lambda}{\Delta\lambda \, dE/d\lambda} = \frac{hc/\lambda}{\Delta\lambda \, (-hc/\lambda^2)} = \frac{\lambda}{-\Delta\lambda} \rightarrow \frac{\lambda}{\Delta\lambda} \tag{1.1}$$

In addition to dispersing the light by wavelength, monochromators must also act to *focus* light from the source, typically onto the exit slit, or onto the focal point of a downstream mirror. The schematic shown in Figure 1.3 is a *plane grating monochromator (PGM)*, where the focussing is accomplished using the two curved mirrors. Other designs use either a curved grating (*spherical grating monochromators*), or subtly change the line spacing of the grooves (*variable line space (VLS) grating monochromators*) to create the required focussing effect and reduce the number of optical elements. We avoid going into detail on focussing and monochromator design here; a good reference is provided by Petersen in Ref. [27].

Our efficiency calculations in the rest of this thesis assume plane gratings and uniform line spacing; however, when curved gratings are used, the radius of curvature is typically so large – on the order of 10m – that the local changes over a grating surface (a few cm) do not affect the efficiency. With VLS gratings, the groove spacing may change by a few percent from end-to-end, and we have modelled these situations by averaging the results of multiple efficiency calculations using a set of representative points.

Many monochromator designs offer multiple switchable gratings to let the user optimize between efficiency and resolution, since higher line density gratings are required to maintain the resolving power at higher energies. (As we will show in Chapter 4 of my thesis draft, the grating efficiency declines as both the line density and energy increase.) As a representative example, the HE-PGM-3 monochromator at Bessy [28] became the starting point for many soft x-ray beamline designs; it uses a 366 line/mm grating to cover the energy range between 30 eV and 700 eV, and a 1221 line/mm grating to cover the energy range between 120 and 1900 eV.

Spectrometers

If the goal of a monochromator is to produce monochromatic light, the goal of a spectrometer is to produce a *spectrum* – ie: to resolve the frequency components that exist in an unknown light source and measure their relative intensity. The device in Figure 1.4 is identical to the monochromator shown in Figure 1.3,

---

[6]For hard x-rays, monochromators use "natural" diffraction gratings consisting of atomic planes in blocks of single crystals, since the inter-atomic spacing is comparable to the short wavelength of the light; these might be analyzed more appropriately using Bragg scattering theory than the electromagnetic approach we take for the man-made structures in this thesis.

except that the exit slit has been replaced by an area-sensitive detector. With the grating positioned so that an outgoing diffraction order (typically the 1st order, for best efficiency) lands on the detector, the angular dependence on wavelength puts short wavelengths onto the top of the detector, and long wavelengths onto the bottom of the detector. The intensity profile recorded across the detector surface is a spectrum, although some mathematical correction will need to be done to calibrate the energy axis, by mapping detector positions to diffraction angles, and diffraction angles to energy using the grating equation.

The spectrometer in Figure 1.5 is more representative of those used in soft x-ray applications, where the low initial levels of fluorescence from the sample and the poor reflectivity of mirrors make it desirable to eliminate as many optical components as possible. Just like monochromators, spectrometers must focus light from the entrance slit – or directly from the source, in the case of slit-less designs – onto the detector. This is accomplished by using spherical gratings and arranging the geometry to exploit the *Rowland Circle* focussing condition discovered by Henry Rowland [27, p. 169], or again by using VLS gratings to alter the shape of the focal curve. More information on spectrometer focussing can be found in the M.Sc. thesis by David Muir [25].

Like monochromators, spectrometer designs usually offer switchable gratings with different line densities and coatings, optimized for different energy ranges as we show in Chapter 5 of my thesis draft. Figure 1.6 shows a top and side view of the SXF endstation on Beamline 8.0.1 of the Advanced Light Source, a typical "workhorse" spectrometer, which balances moderate resolution with reasonable efficiency. It uses four gratings with groove densities of 600, 1000, and 1500 lines/mm to cover the energy range from 70 to 1200 eV, and uses a 40 mm wide multi-channel plate detector with an effective spatial resolution between 40 and 80 um. Figure 1.7 shows an image recorded by that detector, and the corresponding spectrum.
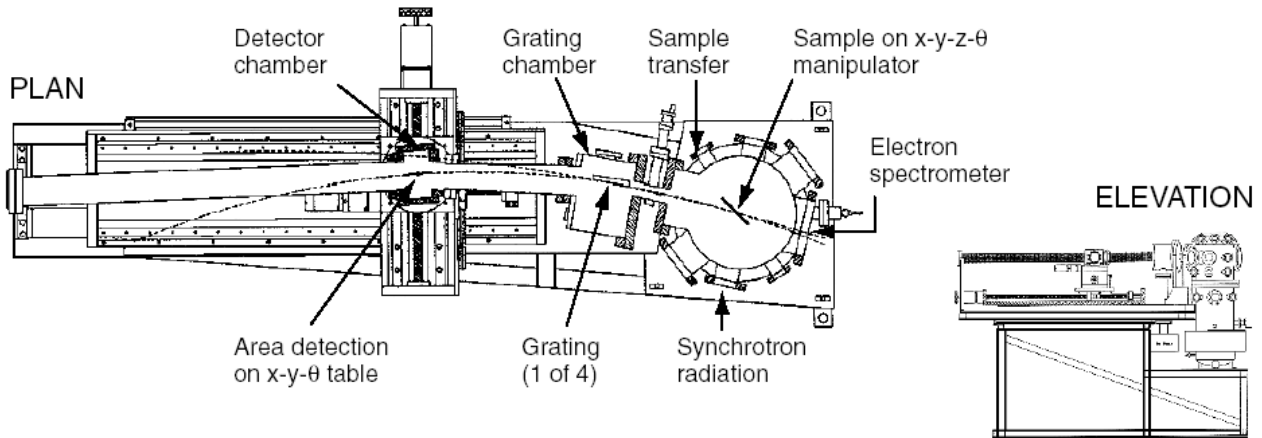

Figure 1.6: The SXF endstation spectrometer on Beamline 8.0.1 of the Advanced Light Source – a typical "workhorse" spectrometer. It uses spherical gratings in a Rowland Circle design; four gratings let users choose between higher resolution or higher efficiency at different energy ranges.

1.2.2 Goals for Soft X-ray Instruments

In the design of both spectrometers and monochromators, there are two goals we have already mentioned that are important from the experimenter's perspective:

- The **resolution** is important in order to see details in spectra and probe new science. On a soft x-ray beamline, the monochromator determines the resolution of an absorption scan, and the spectrometer determines the resolution of an emission scan. Both are involved in the experimental resolution of RIXS studies.

- The overall **efficiency** is important because it determines how much light is available, and indirectly how long it takes to generate enough photon interactions to record acceptable statistics. At a minimum, the efficiency must be high enough to produce a signal above the background noise level of the
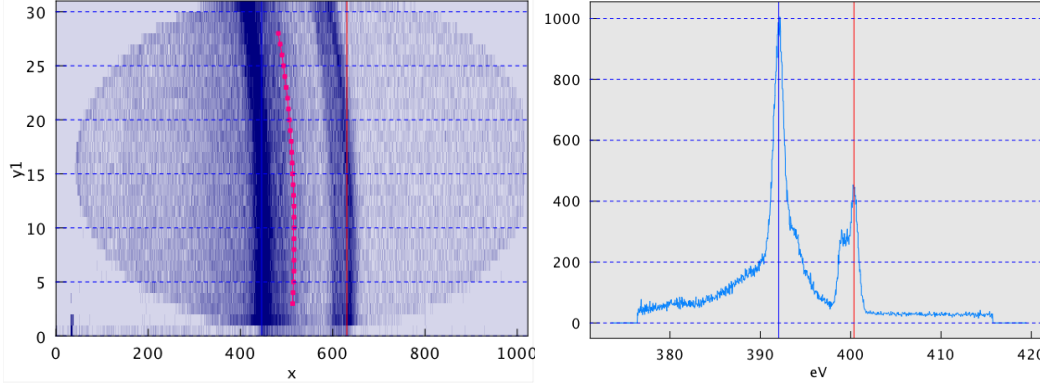
Figure 1.7: A detector image and corresponding spectrum produced by the SXF spectrometer in Figure 1.6. (This particular scan is of the nitrogen emission lines in nitrogen-doped zinc oxide.) The curvature of the spherical gratings provides focussing in the dispersion direction ($x-$axis), but unfortunately produces a curved image in the perpendicular direction ($y-$axis). This curvature is corrected by aligning rows along the red curve when the image is summed to produce the spectrum on the right.

detector. Beyond this threshold, increases in efficiency make it faster to do experiments and let the user accomplish more science in their shift.

Since the probability of an excited atom decaying via fluorescence transitions is so low compared to other decay methods, the initial amount of light available in emission experiments is extremely low. This is further compounded in experiments on doped or dilute samples, when trying to observe the emission lines of the dopant elements. Finally, the entrance slit of a spectrometer can only capture a tiny geometric fraction of all photons out of the sample. All of these reasons combine to make emission spectroscopy experiments extremely "photon hungry", which means that efficiency is especially critical here.

What makes these two goals challenging is that they are inherently in tension. To see why, we consider as an example a simple emission spectrometer like the one shown again in Figure 1.8.
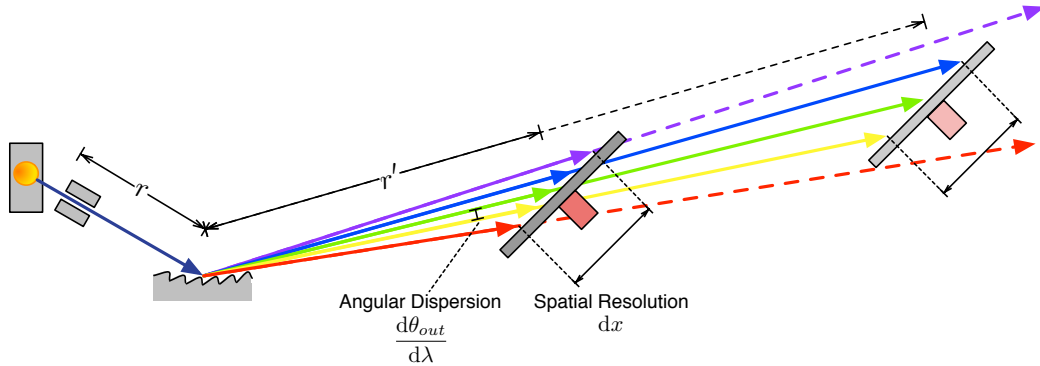

Figure 1.8: The detector has an effective spatial resolution $\mathrm{d}x$ which is the minimum distance required to resolve two adjacent incident rays. Except for the $n = 0$ order, the grating produces an angular dispersion $\mathrm{d}\theta_{2,n}/\mathrm{d}\lambda$, which is the angular separation between infinitesmally-adjacent wavelengths. To increase the spacing between adjacent wavelengths on the detector, we can either increase the grating-detector distance r', or increase the angular dispersion.

The immediate result of a spectroscopy experiment is the image produced along the detector surface. For all detectors – be they photographic film, multi-channel plate detectors, or CCDs – there is an *effective spatial resolution*: the minimum distance along the detector that is required for two adjacent rays to be

distinguished. For a given detector, if you wanted to increase the *energy resolution*, you would want to either

1. increase the angular dispersion of the grating – thereby increasing the angular separation between adjacent wavelengths, or

2. you would want to put the detector a long way from the grating, so that over distance the angular dispersion creates a large spatial separation between adjacent wavelengths.

Both of these solutions inherently reduce the *geometric efficiency* of the spectrometer: the fraction of all photons captured by the detector assuming perfect mirror and grating efficiency. In the dispersion direction (or *meridional* direction, vertical in Figure 1.8), the detector will now only capture photons from a smaller energy range; the corresponding reduction in light gathered is an unavoidable tradeoff no matter what scheme is used to increase resolution. However, the second method – increasing the distance from the grating to the detector – reduces the solid angle captured by the detector also in the perpendicular direction (or *saggital* direction, out of the page), further decreasing the geometric efficiency to an extent which depends on the nature of focussing in that direction.[7] The only way to increase resolution with only the minimum essential reduction in geometric efficiency is by using the first method: increasing the grating's angular dispersion. By differentiating the grating equation (2.61) with respect to wavelength (holding the incident angle θ_2 constant):

$$\frac{n\lambda}{d} = \sin\theta_{2,n} - \sin\theta_2 \tag{1.2}$$

$$\frac{n}{d} = \cos\theta_{2,n}\frac{\mathrm{d}\theta_{2,n}}{\mathrm{d}\lambda} \tag{1.3}$$

$$\frac{\mathrm{d}\theta_{2,n}}{\mathrm{d}\lambda} = \frac{n}{d\cos\theta_{2,n}} \tag{1.4}$$

we derive an expression for the angular dispersion; it is clear that to create more dispersion, we can either increase the groove density (decrease the groove spacing d), or use a higher diffraction order n. Unfortunately, as will be shown in Chapter 5 of my thesis draft, both of these actions substantially reduce the *grating efficiency*. As an additional challenge, we can see that the dispersion will be reduced for higher energies, which leave the grating at more normal angles, creating a larger $\cos\theta_{2,n}$.

No matter which method is used, resolution and overall efficiency are in conflict. Whether it is a hit to the grating or to the geometry, increasing the resolution demands a reduction in efficiency and vice-versa, the cleverness of the beamline designer being measured in his or her ability to negotiate this compromise.

While this example examined resolution as seen by a spectrometer detector, the same unfortunate principles apply to monochromators as well.

1.2.3 Challenges of Soft X-ray Applications

All grating spectrometers, regardless of their wavelength range, are subject to the unavoidable tradeoff between resolution and efficiency. However, the nature of soft x-rays adds an additional set of challenges within this regime, which combine to make efficiency even more important.

UHV Compatibility

Because the energy of soft x-ray photons matches the energy of core level electron transitions in all light elements, they are readily absorbed in matter over very short distances. (Just passing through 1 mm of nitrogen at atmospheric pressure is enough to attenuate a 200 eV beam by 63 percent!) For starters, this means that all SXS experiments must be done under *ultra-high vacuum* (UHV) conditions, where the beam path has been evacuated of air and other contaminants, ideally to a pressure lower than 10^{-8} torr. This is accomplished by using vacuum chambers and sealed beam pipes, pumped down using turbo-molecular

---

[7]A complete analysis of focussing criteria would show that, at least for Rowland Circle spectrometers, increasing the source-grating distance requires an increase in both the entrance slit-grating distance, *and* an increase in the grating radius; both of these actions further decrease the geometric efficiency, possibly as as function of the cube of the distance. One could compensate by using larger gratings, but the maximum grating size must be limited to manage spherical aberrations; see Ref. [25, p. 98].

pumps, ion pumps, and/or cryogenic pumps. At these low pressures, most common materials would "out-gas", boiling off contaminants into the vacuum environment. As a result, chambers and instrumentation must be built using a restricted set of UHV-compatible materials (certain grades of stainless steel and aluminum, copper, gold, and some special ceramics and high-temperature thermoplastics). Anything that comes into contact with the vacuum environment must be carefully cleaned prior to assembly, and it is often necessary to "bake out" the chambers by heating them to temporarily raise the vapour pressure while pumping in order to remove water and other contaminants adsorbed to the inside surfaces before eventually attaining UHV levels.

For the experimenter, UHV requirements also apply to samples. Liquid samples and non-UHV-compatible samples must be carefully enclosed and sealed behind thin windows to let the beam in. (Beryllium is often used in this application, due to its strength and relatively low absorption). Loading samples into the vacuum chamber must be done through an airlock, and remote actuators are required to adjust the sample position and the position of optical elements.

Low Reflectivity and Grazing Incidence

Beamline designers face a more significant challenge when choosing optical elements. The easy absorption of soft x-rays means that mirror and grating surfaces have extremely low reflectivity at these wavelengths – at least at normal incidence. For visible light, the reflectivity of a polished aluminum surface approaches 90%; at 200eV, the reflectivity of the same surface is less than 0.005% [9].

Two approaches are necessary to work around this challenge. Optical coatings need to be selected to avoid absorption edges in the region of interest. For example, carbon and nickel have relatively high peak reflectivities but very strong absorption features at 284 eV and 853 eV respectively; these coatings would only be appropriate for optical elements used within a narrow wavelength range away from those edges. In many-electron materials like gold and platinum, the most pronounced core-level absorption edges occur outside the soft x-ray region, and as result they have acceptable reflectivity over a wide bandwidth.[8]

Second, mirrors and gratings need to be aligned and used at *grazing incidence* – i.e., with the incident light striking them at glancing angles just a few degrees from parallel to their surface. At soft x-ray wavelengths, most materials have a refractive index with a real part *less than* one; the phase velocity of light is actually *faster* than it would be in a vacuum. This allows us to exploit the phenomenon of total internal reflection (TIR), except in this case it becomes a *total external reflection*, since the refractive index is higher in the grating material than it is in the vacuum above it. Just as with conventional TIR in glass and air, we can calculate a critical angle above which total external reflection will occur. For most metals, this turns out to be around 83 to 85 degrees from the surface normal.

1.3 REIXS spectrometer project

This thesis – and its motivation to establish a better understanding of diffraction grating efficiency – emerged from a primary engineering goal: to design and build a world-class emission spectrometer for the REIXS beamline at the Canadian Light Source. The REIXS (*Resonant Elastic and Inelastic X-ray Scattering*) beamline is optimized for material science experiments; it offers an elliptically-polarizing undulator to produce a high brightness beam of soft x-ray photons, and a high-resolution monochromator with a flux of $\sim 10^{13}$ photons/second and a resolving power greater than 5000. The beamline has room for two endstations; our task was to build the emission spectrometer, which will be used for inelastic scattering experiments, while a team from the University of British Columbia built the second endstation used for elastic scattering measurements.

In Chapter 5 of my thesis draft, I describe the design process and the resulting optical design of the spectrometer. Used in combination with the work of David Muir on spectrometer resolution [25], the ability to model grating efficiencies allowed us to optimize the design for an intelligent compromise between these two competing goals. In the process, we discovered an efficiency peak in the 3rd order – typically assumed

---

[8]Gold and platinum do have absorption edges in the soft x-ray range, but these correspond to outer shell electrons (for example, gold 4s electrons bound at 802 eV) instead of core-level electrons. The outer electrons see shielding by the inner ones, and therefore have a lower transition probability.

to be unusable – that prompted us to create an innovative design capable of reaching much higher resolution than would otherwise be possible given the space constraints of the beamline.

1.4   Summary: Why Grating Efficiency Matters

Figure 1.9 lists a variety of soft x-ray spectroscopy techniques, and shows the number of gratings in the beam path to the detector for each. Of these techniques, those in *italic text* are possible on the REIXS beamline; the remaining are possible on other beamlines at the CLS. For all of these techniques, the ability to create more efficient gratings would increase the speed of experiments, improve the quality of data, and increase the minimum concentration of samples that can be feasibly studied.

Conversely, during the design phase of these kinds of instruments, the unavoidable tradeoff between resolution and efficiency implies that accurate predictions of the grating efficiency could ironically be used improve *resolution*: by deliberately sacrificing efficiency, designers could push their resolving power to the limit once the efficiency is known to be "good enough".

	Photon In	Electron In
Electron Out	*Near-edge X-ray Absorption Fine Structure (NEXAFS): electron yield*	
	Extended Xray Absorption Fine Structure (EXAFS): electron yield	Electron Energy Loss Spectroscopy (EELS)
	Xray Photoelectron Spectroscopy (XPS)	
Photon Out	*Soft X-ray Emission Spectroscopy (SXE)*	
	Resonant Inelastic X-ray Spectroscopy (RIXS)	*Inverse Photoelectron Spectroscopy (IPES)*
	X-ray Excited Optical Luminescence (XEOL)	
	NEXAFS: fluorescence yield	

Figure 1.9: A variety of soft x-ray spectroscopy techniques, and the number of gratings required in the beam path to accomplish each one. Of these techniques, those in *italic text* are possible on the REIXS beamline using the emission spectrometer endstation. For all these techniques – and particularly those using two gratings – more efficient gratings would increase the speed of experiments, improve the quality of data, and increase the minimum concentration of samples that can be feasibly studied.

CHAPTER 2

THEORY: HOW TO CALCULATE GRATING EFFICIENCY
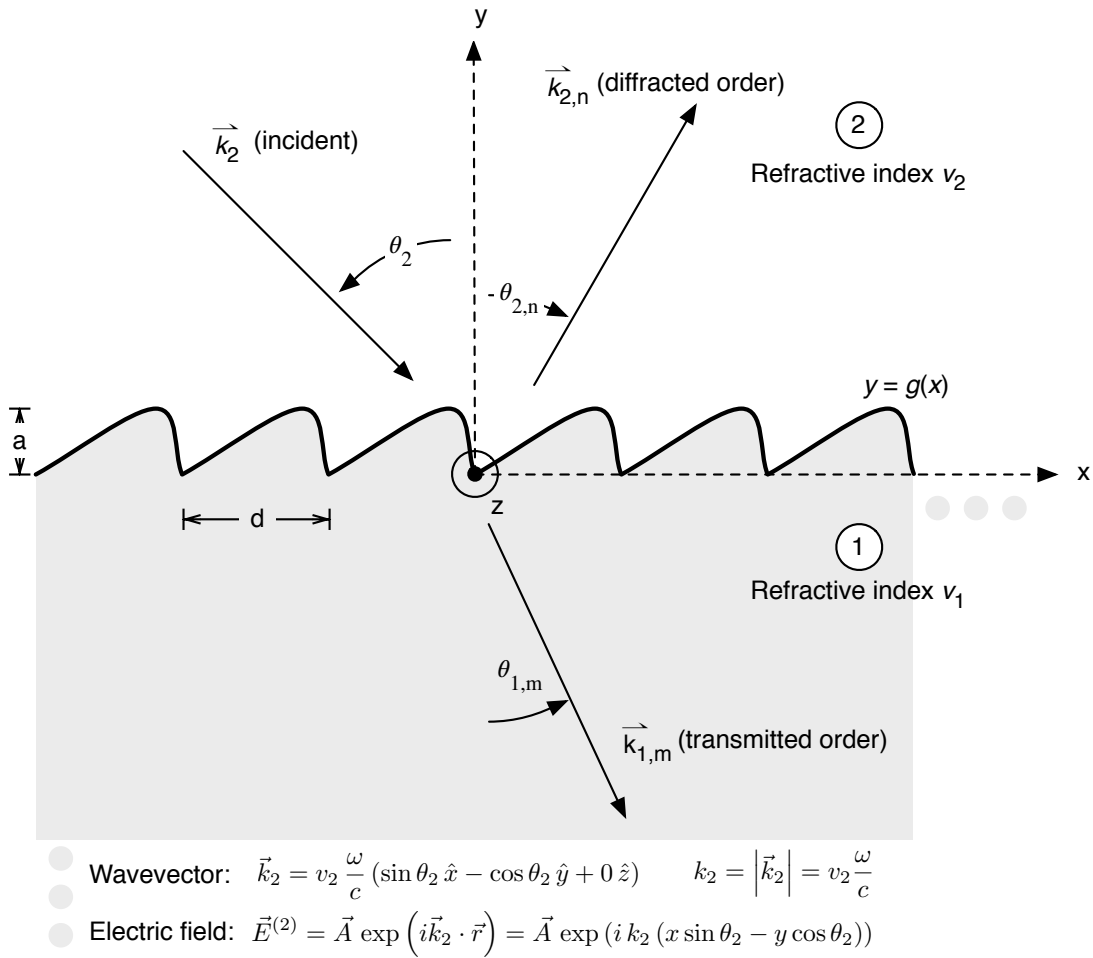
2.1   Introduction to grating theory

Figure 2.1: A one-dimensional grating with in-plane incidence.

For centuries, we have known that gratings reflect and transmit light at a set of discrete angles – called *orders* – where the angles depend on the wavelength (Figure 2.1). A variety of simple proofs have been used to explain this. In undergraduate books, the grating equation is introduced as a consequence of constructive interference from a linear array of coherent emitters. (This makes sense for a slotted transmission grating, but it is less obvious why an arbitrarily-shaped periodic reflecting structure like the one in Figure A.1 would

behave the same way.) Other proofs – such as those using Fermat's Principle to minimize an optical path function with a phase offset introduced by the grating [27, p. 93 – 99] – are general enough to explain the existence of diffraction orders, but they do not say anything about how much light goes into each.

This is unfortunate, because to determine grating efficiency, we must determine the intensity of the orders. A huge amount of research has gone into this question since Lord Rayleigh's pioneering work in 1907 [38], and some clever analytical methods were even discovered that simplify the problem – but only for certain special groove shapes. It is only within the last twenty years that mathematical techniques and computational power have advanced enough to accurately model the electromagnetic field of light within and near the surface of an arbitrary grating, allowing us to actually determine how energy is distributed between the orders, and how much is absorbed within the grating. In this chapter, we look at a history of the key developments in grating theory, give a comparison of the current methods, and then set up the grating problem using the most general of these methods.

2.1.1   A brief history of grating theory

The "grating problem" – as we define it – is the problem of determining the intensity of a grating's reflected and transmitted light everywhere. In general, this requires finding the solution to Maxwell's equations for the electric and magnetic fields of the light, in the presence of the grating's periodic boundary conditions, for a known incident field.

Prior to the availability of computers and numerical techniques, great attempts were made to solve this problem analytically. In 1907, Lord Rayleigh set up the problem[1] using a Fourier approach identical to the one we use later in this chapter [38]. However, without access to powerful numerical methods, he could only make coarse approximations, using expansions truncated at the first or second power of the groove height.[2] This led to some obviously non-physical results, such as predictions of infinite intensity in any order when the outgoing angle of a higher order reached exactly 90 degrees. From 1907 until 1960, advancements in grating theory amounted to finding clever but limited simplifications to these expansions, which would work in the case of one particular groove profile, incidence angle, or approximation limit.

The introduction of computer-based numerical methods in the 1960s spawned a renewal of research into general solutions to the problem. The first "rigorous" algorithm[3] was invented by several researchers in 1966:

The "integral" approach uses Maxwell's equations in integral form and the Kirchoff integral theorem [29][41][15]. It was generalized from perfectly conducting (i.e.: perfectly reflecting) gratings to finite-conductivity metallic gratings in an important paper in 1972 [20].

Alongside these "integral" methods, two other broad families of algorithms emerged based on Maxwell's equations in differential form:

The "modal" method relates the electromagnetic field to a Fourier expansion in the vertical (y) direction with unknown coefficients, called modes. The total field satisfying the boundary conditions is solved algebraically as a linear combination of the modes [2][3][1].

The "differential" methods exploit the periodicity of the grooves to express both the grating boundary and the fields as Fourier expansions in the x-direction. The total field satisfying the boundary conditions is determined by numerically integrating a finite set of differential equations. (We explore the details of this method later in this chapter.)

The modal methods are not actually general because in their application of the boundary conditions, they require the edges of the grooves to be vertical; this limits them to rectangular gratings. However, they are accurate in this special case, and fast compared to the differential method because they avoid numerical

---

[1]Actually, Rayleigh's first application of this technique was to the diffraction of sound waves [37]. It turns out that the wave equation for a pressure wave in compressible media is the same as for the fields in an electromagnetic wave, and the boundary conditions at interfaces are also analogous. In Ref. [38] he shows applications to both sound and light.

[2]And this, despite assuming a perfectly reflecting grating, exactly normal incidence, and only one term in the Fourier expansion for the groove shape – i.e., a sinusoidal grating.

[3]"Rigorous" meaning a method that does not introduce approximations into the theoretical equations, other than inevitably in the process of finite-precision numerical calculations.

integration. The other two approaches are general in representation, although the stability and accuracy of the integral approach is highly contingent on the groove geometry.

By 1980, grating theorists had devised a wide variety of implementations for each family. Reference [30] provides a comprehensive review of all the methods that were available by this date.[4] Of the three families, the differential approach had proven to be the most general way of setting up the problem, but it ran into two numerical issues – particularly when used to describe TM-polarized light on highly-conducting metallic gratings. One issue was strictly computational: round-off errors when doing finite-precision computer arithmetic would lead to ever-growing contamination during the numerical integration of exponential functions. This problem was solved by the introduction of the "S-matrix" propagation algorithm in 1996 [18][24]. The second problem was related to the mathematical implications of truncating an infinite Fourier series – unfortunately but obviously necessary for numerical calculations. This challenge was also finally resolved by the breakthrough work of Li [19] in 1996 on the factorization of truncated Fourier series, and has been applied extensively by Neviere and Popov both to the simple cases of one-dimensional gratings like those shown in Figure 2.1 [36], and to complex three-dimensional periodic structures of anisotropic materials [34][26].

Choosing a theoretical method

Appendix A presents just enough theory on each method to understand their limitations and areas of applicability with respect to groove geometry, material conductivity, polarization, numerical stability, and their ability to handle layers and coatings. Since we are primarily interested in the efficiency of soft x-ray gratings used in grazing incidence, these limitations provide clear guidance for choosing an appropriate method. At the frequency of soft x-ray radiation, metals behave as absorbing, weak dielectrics with a refractive index near $(0.999 - 0.001i)$; therefore, we do not need to worry about limitations on perfectly-conducting materials. At grazing incidence, polarization effects are also minimized; there turns out to be very little difference between measured and calculated results for TE and TM polarization. Finally, gratings optimized for grazing incidence tend to have shallow grooves, relative to the groove period, which frees up the limitation on deep gratings.

However, we must be able to model a variety of groove profiles, including rectangular but also triangular, trapezoidal, and sinusoidal shapes, with at least one coating layer. (The triangular, or "blazed" profile – featuring sharp vertices – will turn out to be particularly important for our spectrometer design.) Therefore, we ruled out the modal and RCW methods, and were left with a choice between the integral method and the full differential method. Given the necessity of modelling coatings and sharp profiles, we selected the differential method for its simplicity and robustness. We used an implementation of this method for all the grating efficiency calculations, optimizations, and comparisons shown later in the thesis.

2.1.2   Overview of the differential theory

The rest of this chapter lays the groundwork for a full electromagnetic theory describing the interaction of light waves with periodic structures. The conventions we use here will be applicable to either the classical differential or RCW methods, and the notation is consistent with the notation used by Neviere in [26]. However, before diving into the details of the mathematics, it might be useful to summarize the basic physical principles behind the solution:

1. The theory starts by using the Maxwell equations to derive a 2nd-order wave equation for the electric and magnetic field of the light. According to classical optics, the incoming light can be divided into two independent polarization components; the "Transverse Electric" (TE) component has its *electric* field vector always parallel to the $z-$axis in Figure 2.1, and the "Transverse Magnetic" (TM) component has its *magnetic* field vector always parallel to the $z-$axis. Due to the beautiful symmetry of the Maxwell equations, it turns out that the wave equation for the electric field in TE polarization is the same as the equation for the magnetic field in TM polarization, and for most of the theory we can use the same

---

[4]The same author provides a follow-up paper, written ten years later on new methods that differ completely from the fundamentals used in 1980 [31].

mathematics to work with both, referred to as the "general field" or just "the field".[5] Because the TE and TM polarizations are independent, we can solve the grating problem for both separately, and then combine the diffracted field solutions in proportion to the polarization of the incident light.

2. Because the grating is periodic, we use Fourier series expansions in the horizontal direction to express both the permittivity of the grating material, and the field itself. In the vertical direction, there are three distinct regions:

   - Region 2: above the grooves, where the refractive index is uniform ($v = v_2$);
   - Inside the grooves – the "modulated region" – where the refractive index changes as a function of the x and y position: it is either v_1 or v_2;
   - Region 1: below the grooves, inside the grating substrate, where the refractive index is again uniform ($v = v_1$).

3. **Above and below the grooves:**
   By using the grating as a periodic operator and applying boundary conditions for the field at infinite, we can prove that light is reflected and transmitted at discrete angles, and that a Fourier sum of plane waves known as the Rayleigh Expansion can be used to express the total field which satisfies the wave equation in this region. Each propagating plane wave corresponds to a diffraction "order", and once we know the angle of the incident light, we can determine the angles of the diffraction orders (Figure 2.4). At this point, the Rayleigh expansion still has unknown coefficients, and we need to determine these coefficients to find out how much energy is diffracted into each order.

4. **Within the grooves:**
   However, the field within the grooves depends on the exact shape of the groove profile and the interaction of light within the material. It cannot be simply represented as a sum of plane waves, and the boundary conditions are complicated. There are two leading methods used to handle this. Both express the wave equation using Fourier expansions for the field and for the grating permittivity. (These expansions would theoretically be infinite sums, therefore they need to be truncated to be approximated using computer calculations. Special rules apply for accurately calculating the products of truncated Fourier series when they have discontinuities – for example, the normal component of the electric field at the grating boundary; these were discovered by Li in Ref. [19].)

   (a) In the "Rigorous Coupled Wave" (RCW) approach [22] [23], the grating groove shape is sliced into thin layers (Figure A.1) and approximated with vertical walls between layers (the "staircase approximation"). This simplifies the boundary conditions because the normal and tangential field components are either entirely along the x- or y-axis; thus, at every layer the boundary value problem can be converted into a set of simultaneous linear equations and solved algebraically. Each layer is treated as a separate grating, and the effects of each layer on the up-going and down-going fields are propagated to the next using matrix methods. (As we describe in Appendix A, the RCW accuracy and convergence suffers in TM polarization because the staircase approximation introduces sharp corners and artificially large electric field components at the step boundaries [8].)

   (b) In the "Differential Method" approach [34], we numerically integrate the wave equation many times using different assumed initial values, to generate a complete orthogonal set of particular solutions. Then we use techniques of linear algebra to solve for the coefficients of the general solution that satisfy the boundary conditions along the grating interface.

5. Finally, most common grating structures consist of one (or often many) stacked layers.[6] (Typical and contrived examples are shown in Figures 2.2 and 2.3). In classical optics, the *reflection matrix* and *transmission matrix* are often used to propagate an incoming field through an optical layer. We can

---

[5]While the wave equation is the same for both the TE electric field and the TM magnetic field, the boundary conditions for the fields at the grating interface are different and need to be handled separately. This is responsible for the difference in efficiency dependent on polarization.

[6]For example, dielectric gratings with a metal coating used for soft x-ray beamlines, gratings with thin-film multilayer coatings, etc.

generalize this concept to gratings by defining a matrix which maps the Rayleigh coefficients of the incident field to the coefficients of the field that is reflected and transmitted by each layer. The matrices for each layer can then be multiplied together to find the effect of the complete grating; this is known as the "T-matrix" approach. While theoretically sound, the T-matrix becomes unstable in computer calculations because rounding errors introduce instability into diverging exponential functions. An alternative formulation called the "S-matrix" approach defines a matrix at each layer which represents the cumulative effect of *all layers below and including that layer* on the incident field; by definition, this matrix remains bounded and is safe to use for numerical calculations [18]. This matrix links the Rayleigh coefficients of the up-going and down-going waves between each layer with the more complicated functions within the grooves of each layer, where the Rayleigh expansion does not apply.

2.1.3 Simplifying assumptions

Before we tackle the mathematics of grating theory, this section defines the geometry and terminology we'll use throughout this text, and introduces some assumptions that simplify the problem. Figure 2.1 shows a side view of a reflection grating, in the following situation:

1. The mean surface of the grating is in the $x-z$ plane, with the grooves running parallel to the $z-$axis. The groove profile is periodic and repeats every distance d along the $x-$axis.

2. The grating is illuminated with a light ray travelling parallel to the $x-y$ plane (i.e.: within the plane of the page). This is referred to as "in-plane incidence", and is typical of how most monochromators and spectrometers are used. (When the incident ray has a component in the $z-$direction, the diffraction peaks end up dispersed over the surface of a three-dimensional cone, and this is referred to as "conical mounting".)

These first two simplifications reduce the grating problem to two dimensions, since the whole system is unchanged by translation along the $z-$axis.

Additionally, we assume that:

1. The incident light can be represented as a plane wave, infinite in extent.

2. The grating also stretches forever in the $x-$ and $z-$directions.

Typical gratings used in soft x-ray devices are much much larger (typically: 40 mm) than their groove spacing d (typically: a few μm). As long as they are illuminated with collimated light having a beam width much larger than the groove spacing, both of these assumptions seem reasonable.

Finally, to simplify the electromagnetic field calculations, we assume that

1. The grating material is non-magnetic, with a permeability of μ_0.

2. The dielectric constant ϵ of the grating material is the same in all directions.

(These last two assumptions are also necessary to keep us within a two-dimensional problem. If the material is non-isotropic, then the dielectric constant ϵ becomes a tensor $[[\epsilon]]$ and the problem requires a full 3D analysis.)

These assumptions make the grating theory much simpler to present in this chapter. However, it should be noted that at the expense of larger matrices and higher mathematical complexity, it is possible to express the same theory in the full three-dimensional case, which enables the analysis of exotic situations like conical mount gratings, crossed 2D gratings, and non-isotropic materials. [Ref: Neviere Ch. 5] provides a derivation of the full general 3D version.

As one final assumption, we assume that the grating surface exactly matches the periodic profile given by $y = g(x)$ in Figure 2.1. For real-world applications, this turns out to be the most limiting assumption, since actual gratings will vary from groove to groove due imperfections in the manufacturing process, and their ideally smooth surfaces will have an unavoidable amount of random microroughness. Unfortunately, this assumption is necessary for any mathematical treatment; in Section **??** of my thesis draft, we look at the impacts of real-world imperfections on grating efficiency.
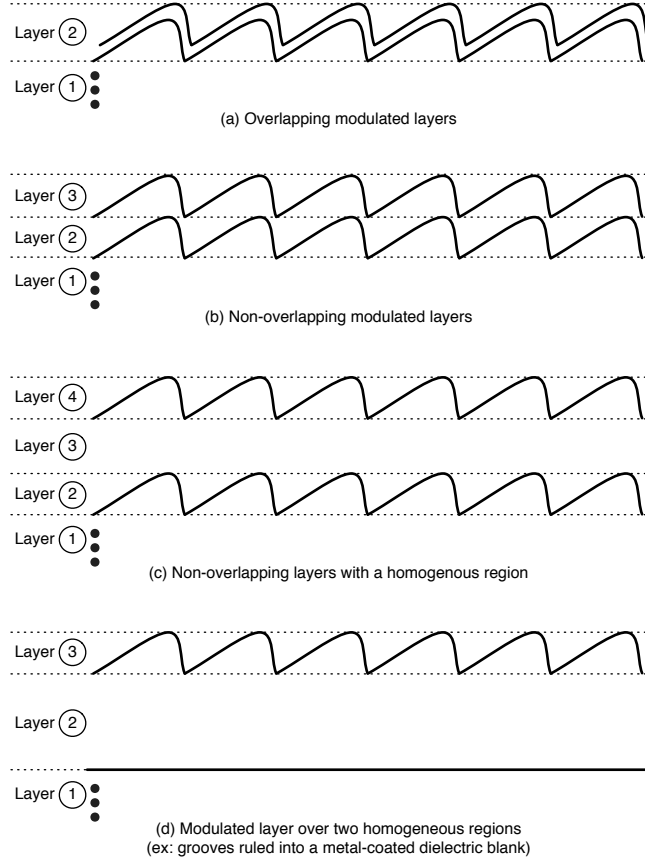
Figure 2.2: Arbitrarily-complicated structures can be handled by dividing the grating into layers, where each layer is either homogenous (constant refractive index), or modulated (with a refractive index that changes periodically as a function of x at any given height y).
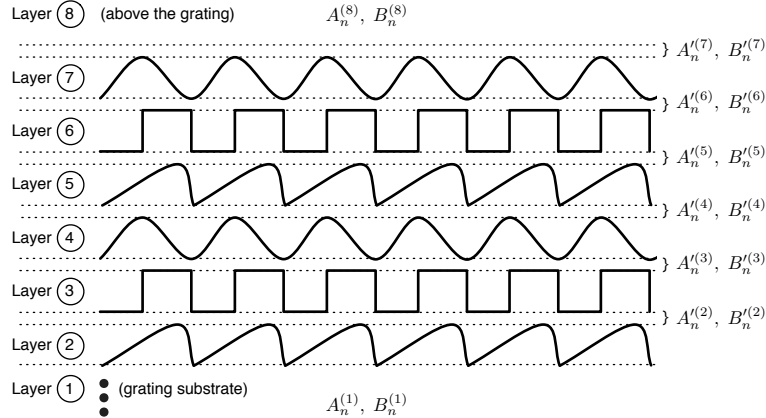

Figure 2.3: Modulated layers in a complicated stack of gratings. In between every layer, we can insert an imaginary, infinitely-thin homogenous layer where the Rayleigh expansion applies. The A'_n and B'_n expansion coefficients connect the boundary conditions between layers. Within each layer, the Rayleigh expansion does not apply – inside the grooves, the field cannot be represented as a simple sum of plane waves – and numerical methods are required to approximate it.

A few brief notes on notation:

- Quantities that change depending on the grating region are subscripted to indicate which region they are in. For example, k_2 and v_2 designate the wave vector and refractive index, and $\theta_{2,n}$ the angle of the n-th order ray, in Region 2 above the grating. Eventually, we will tackle layered gratings with many numbered regions, and the subscript convention will become very helpful.

- Angles are measured from the surface normal as shown in Figure 2.1. There are two different sign conventions in common use; we will use positive angles for both the incident ray (θ_2) and the diffracted ray ($\theta_{2,n}$) when they are on opposite sides of the surface normal. Transmitted rays ($\theta_{1,n}$) are also measured using positive angles from the $-y$ axis, as shown in Figure 2.1.[7]

2.1.4 Electromagnetic field and polarization

In Figure 2.1, incoming light strikes the grating along wavevector $\vec{k}_2$ at an angle θ_2 from perpendicular to the grating plane.

$$\vec{k}_2 \;\; = \;\; v_2 \frac{\omega}{c} \left(\sin\theta_2 \hat{i} - \cos\theta_2 \hat{j} + 0\hat{k} \right) \tag{2.1}$$

(We use v to designative the refractive index – in this case, in Region 2 above the grating, which is normally air or vacuum. $\omega = 2\pi f$ is the angular frequency of the light, and c is the speed of light in vacuum.)

The incoming light is a travelling electromagnetic plane wave with sinusoidal dependence on time. We can express the electric field vector using the complex exponential form:

$$\vec{E}_{incident} \;\; = \;\; \vec{A} \exp\left(i(\vec{k}_2\,\vec{r} - \omega t) \right) = \vec{A} \exp\left(i\,k_2(x\sin\theta_2 - y\cos\theta_2) \right) \exp\left(-i\omega t \right) \tag{2.2}$$

where the true (physical) field is contained in the real part. Since all fields will have the same harmonic dependence on time, we drop the $e^{-i\omega t}$ factor from here on. (The scalar k_2 is simply the magnitude of $\vec{k}_2$, i.e.: $\left| \vec{k}_2 \right| = v_2\,\omega/c$.)

The electric field vector is always perpendicular to the direction of the wave $\vec{k}_2$, but can have an arbitrary polarization determined by $\vec{A}$. This can always be expressed as a superposition of two orthogonal components:

- The TE (transverse electric) polarization represents the component of the electric field vector parallel to the z axis (i.e.: into the page). For a pure TE wave, the E_x, E_y, and B_z fields are always 0.

- The TM (transverse magnetic) polarization represents the component of the electric field vector in the $x - y$ plane (i.e.: the *magnetic* field vector is along z). For a pure TM wave, the E_z, B_x, and B_y fields are always zero.

In general, the grating problem will be solved by applying the Maxwell equations to the incident electric field, according to the boundary conditions imposed by the grating. By solving it separately for each polarization case (TE and TM), we can simplify these equations. In the end, we can superpose the results we calculate for the outgoing fields, weighted based on the actual polarization of the incident light.

2.1.5 Maxwell's equations for sinusoidal time-varying fields

Our assumptions in Section 2.1.3 mean that the gratings are not loaded with free charge, or carrying free currents. Therefore, two relevant Maxwell equations are:

$$\nabla \times \vec{E} \;\; = \;\; -\frac{\partial \vec{B}}{\partial t} \tag{2.3}$$

---

[7]The sign convention for diffraction angles affects the sign in the right-hand side of the grating equation (2.61). In North American literature it is more common to measure all angles positive counter-clockwise from the surface normal; this creates a '+' rather than a '−' in the grating equation.

$$\nabla \times \vec{B} \quad = \quad \mu\epsilon\frac{\partial \vec{E}}{\partial t} \tag{2.4}$$

For sinusoidal time-varying fields, the electric field $\vec{E}$ is proportional to $e^{-i\omega t}$, so the time derivatives reduce to:

$$\nabla \times \vec{E} \quad = \quad i\omega\mu\vec{H} \tag{2.5}$$

$$\nabla \times \vec{H} \quad = \quad -i\omega\epsilon\vec{E} \tag{2.6}$$

or, expressed in Cartesian coordinates:

$$\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} \quad = \quad i\omega\mu H_x \tag{2.7}$$

$$\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} \quad = \quad i\omega\mu H_y \tag{2.8}$$

$$\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} \quad = \quad i\omega\mu H_z \tag{2.9}$$

$$\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} \quad = \quad -i\omega\epsilon E_x \tag{2.10}$$

$$\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} \quad = \quad -i\omega\epsilon E_y \tag{2.11}$$

$$\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} \quad = \quad -i\omega\epsilon E_z \tag{2.12}$$

Since the grating and the incident light are uniform along the $z-$axis, all of the partial derivatives with respect to z are 0. For TE polarization, since E_x and H_z are 0, equations (2.7, 2.8, and 2.12) are decoupled into

$$\frac{\partial E_z}{\partial y} \quad = \quad i\omega\mu H_x \tag{2.13}$$

$$-\frac{\partial E_z}{\partial x} \quad = \quad i\omega\mu H_y \tag{2.14}$$

and

$$\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} \quad = \quad -i\omega\epsilon E_z \tag{2.15}$$

We can use (2.13) and (2.14) to eliminate H_x and H_y in (2.15):

$$\frac{i}{\omega\mu}\frac{\partial^2 E_z}{\partial x^2} + \frac{i}{\omega\mu}\frac{\partial^2 E_z}{\partial y^2} \quad = \quad -i\omega\epsilon E_z \tag{2.16}$$

Note that ϵ is a function of position $\epsilon(x, y)$, since it changes whether inside the grating or above the grating. Since

$$k^2(x, y) = v^2(x, y)\,\omega^2/c^2 = \omega^2\mu\epsilon(x, y) \tag{2.17}$$

we get a single second-order wave equation in E_z

$$\nabla^2 E_z + k^2 E_z = 0 \tag{2.18}$$

where $E_z = E_z(x, y)$ and $k^2 = k^2(x, y)$ are both functions of position.

For the case of TM polarization, the Maxwell equations (2.10, 2.11, and 2.9) reduce to

$$\frac{\partial H_z}{\partial y} = -i\omega\epsilon E_x \tag{2.19}$$

$$-\frac{\partial H_z}{\partial x} = -i\omega\epsilon E_y \tag{2.20}$$

$$\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = i\omega\mu H_z \tag{2.21}$$

and an identical procedure produces the wave equation in H_z:

$$\nabla\left[\frac{1}{k^2}\nabla H_z\right] + H_z = 0 \tag{2.22}$$

In general, k here is still a function of position $k(x, y)$ depending on whether we are inside or outside of a groove valley, or above or below the modulated region. However, in the homogenous space above and below the modulated region, k is constant and both Equation 2.18 and Equation 2.22 reduce to the Helmholtz equation:

$$\nabla^2 U_z + k^2 U_z = 0 \tag{2.23}$$

Due to the (near) symmetry of the Maxwell equations, the form of the wave equation in homogeneous space is the same for the $U_z = H_z$ field in TM polarization as it is for the $U_z = E_z$ field in TE polarization. To solve the grating problem, we will need to find the solution to this wave equation in the presence of the grating boundary conditions.[8]

2.1.6 Periodicity of gratings and fields (aka, pseudo-periodic functions and the Fourier basis)

The periodic nature of the grating grooves immediately hints that we could represent them conveniently using a Fourier series. But can a Fourier expansion also be used to represent the fields E_z and H_z?

Here we define $u = E_z$ when working in TE polarization, and $u = H_z$ when working in TM polarization; it represents the "generic field". The grating can be imagined as an operator $\mathbb{G}$ that transforms the incident field u_i into a diffracted field u:

$$u(x, y) = \mathbb{G}\, u_i(x, y) \tag{2.24}$$

Because the grating is periodic and extends forever, the operator $\mathbb{G}$ is invariant (does not change) under translation by a grating period: $x \to x + d$.

$$u(x + d, y) = \mathbb{G}\, u_i(x + d, y) \tag{2.25}$$

Since the incident light arrives at an angle θ_2, this translation will add an extra path distance $d\sin\theta_2$ to the incident wave u_i, for a phase change $e^{ik_2 d \sin\theta_2}$:

$$u_i(x + d, y) = e^{ik_2 d \sin\theta_2} u_i(x, y) = C\, u_i(x, y) \tag{2.26}$$

where the constant C has been matched to $e^{ik_2 d \sin\theta_2}$.

Because the set of coupled Maxwell partial differential equations (2.7 - 2.12) is linear, any solution multiplied by a constant C is still a solution:

$$u(x, y) = \mathbb{G}\left(u_i(x, y)\right) \tag{2.27}$$

$$C\, u(x, y) = \mathbb{G}\left(C\, u_i(x, y)\right) \tag{2.28}$$

$$C\, u(x, y) = \mathbb{G}\, u_i(x + d, y) \tag{2.29}$$

---

[8]Although the wave equation is the same, the boundary conditions for electric and magnetic fields are different at the grating boundary where ϵ and v change suddenly – for example, the normal component of the electric field is discontinuous, while the normal component of the magnetic field is continuous. This leads to a difference in the TE and TM efficiency.

but since $\mathbb{G} u_i(x+d, y) = u(x+d, y)$ also,

$$C u(x, y) = u(x+d, y) \tag{2.30}$$

In other words, the total field translated by one period d is equal to its untranslated self, multiplied by a complex constant:

$$u(x+d, y) = C u(x, y) = e^{ik_2 d \sin \theta_2} u(x, y) = e^{i\alpha_0 d} u(x, y) \tag{2.31}$$

where we have defined

$$\boxed{\alpha_0 \equiv k_2 \sin \theta_2} \tag{2.32}$$

This is known as a *pseudo-periodic* relationship:

$$u(x+d, y) = e^{i\alpha_0 d} u(x, y) \tag{2.33}$$

since a true (strictly) periodic relationship would have the form:

$$v(x+d, y) = v(x, y) \tag{2.34}$$

However, we can easily create such a function by defining $v \equiv e^{-i\alpha_0 x} u$. As a legitimately periodic function, v can be represented as a Fourier series expansion on the grating period d:

$$v(x, y) = e^{-i\alpha_0 x} u(x, y) \tag{2.35}$$

$$v(x, y) = \sum_{n=-\infty}^{\infty} u_n(y) e^{i2\pi n x/d} \tag{2.36}$$

$$u(x, y) = e^{i\alpha_0 x} \sum_{n=-\infty}^{\infty} u_n(y) e^{i2\pi n x/d} \tag{2.37}$$

If we define

$$\boxed{\alpha_n \equiv \alpha_0 + 2\pi n/d} \tag{2.38}$$

we can express the total field as something that looks *very close* to a Fourier series expansion:

$$u(x, y) = \sum_{n=-\infty}^{\infty} u_n(y) e^{i\alpha_n x} \tag{2.39}$$

This is the Fourier basis for the total field $u = E_z$ or H_z, with an infinite number of Fourier coefficients u_n. (Eventually, we will need to truncate this series to work with it numerically.)

2.1.7   Deriving the Grating Equation

Equipped with a Fourier expansion for the total field and a wave equation, we can attempt to solve for the field. Figure 2.1 shows the coordinate system sliced into three regions:

- Region 2: Above the grooves where $y > a$, the impedance $k(x, y)$ is constant and proportional to the refractive index of the air/vacuum: $k_2(x, y) = v_2 \omega/c$.

- Region 1: Below the grooves where $y < 0$, the impedance $k(x, y)$ is also constant and proportional to the grating's refractive index: $k_1(x, y) = v_1 \omega/c$.

- Inside the grooves, the impedance is changing as a function of position: whether inside or outside of a groove. We ignore this difficult region for now and try to work with the uniform regions as much as possible.

23

Where $k(x,y)$ is constant, the wave equations for both TE (2.18) and TM polarization (2.22) reduce to:

$$\nabla^2 u + k^2 u = 0 \tag{2.40}$$

where $k = k_1$ above the grating, and $k = k_2$ below it. If we insert the Fourier expansion for the field $u(x,y) = \sum u_n(y)e^{i\alpha_n x}$ into the wave equation:

$$\nabla^2 \left(\sum_{n=-\infty}^{\infty} u_n(y)e^{i2\pi nx/d}e^{i\alpha_0 x} \right) + k^2 \sum_{n=-\infty}^{\infty} u_n(y)e^{i2\pi nx/d}e^{i\alpha_0 x} = 0 \tag{2.41}$$

$$e^{i\alpha_0 x} \sum_{n=-\infty}^{\infty} \left(\frac{\partial^2}{\partial y^2} + (k^2 - (\alpha_0 + 2\pi n/d)^2) \right) u_n(y)e^{i2\pi nx/d} = 0 \tag{2.42}$$

For this Fourier sum to be equal to zero, all of the coefficients must be zero:

$$\left(\frac{\partial^2}{\partial y^2} + (k^2 - (\alpha_0 + 2\pi n/d)^2) \right) u_n(y) = 0 \tag{2.43}$$

$$\left(\frac{\partial^2}{\partial y^2} + (k^2 - \alpha_n^2) \right) u_n(y) = 0 \tag{2.44}$$

This is a standard differential equation with the solution

$$u_n(y) = A_n e^{-i\beta_n y} + B_n e^{i\beta_n y} \tag{2.45}$$

where

$$\beta_n = \sqrt{k^2 - \alpha_n^2} \tag{2.46}$$

and A_n and B_n are unknown constants to be determined by the boundary conditions.

Special attention must be paid to the square root for β_n, depending on whether we are above or below the grating.

Above the grating

Above the grating, we are likely inside air or vacuum ($k = k_2$), so the refractive index and k would be real. Since $\alpha_n = \alpha_0 + 2\pi n/d$ will increase with n, there will be a finite number of n near $n = 0$ where $(k^2 - \alpha_n^2)$ will be a positive number. However, there will be an infinite number of n (approaching $n \to \infty$ and $n \to -\infty$) where $(k_2 - \alpha_n^2) < 0$. This creates two possibilities for β_n (which we label $\beta_n^{(2)}$ because we are in Region 2):

$$\beta_n^{(2)} = \sqrt{k^2 - \alpha_n^2} \qquad (k^2 - \alpha_n^2) > 0 \qquad \text{(finite occurrences, } \beta_n^{(2)} \text{ real)} \tag{2.47}$$
$$\beta_n^{(2)} = i\sqrt{\alpha_n^2 - k^2} \qquad (k^2 - \alpha_n^2) < 0 \qquad \text{(infinite occurrences, } \beta_n^{(2)} \text{ complex)} \tag{2.48}$$

Using this solution for the Fourier coefficients u_n, the total field is:

$$u(x,y) = \sum_{n=-\infty}^{\infty} A_n^{(2)} e^{i\alpha_n x - i\beta_n^{(2)} y} + \sum_{n=-\infty}^{\infty} B_n^{(2)} e^{i\alpha_n x + i\beta_n^{(2)} y} \tag{2.49}$$

This result is remarkable. The total field created above the grating is a finite sum of propagating plane waves (for all n where $\beta_n^{(2)}$ is real) and an infinite sum of decaying (when $\beta_n^{(2)}$ is complex) plane waves.

To expand on this conclusion: The first sum (with $A_n^{(2)}$ coefficients) are waves travelling in the $-y$ direction, down toward the grating. For the finite set of n for which $\beta_n^{(2)}$ is real, these are non-decaying, normal waves. There is also an infinite number of exponentially growing waves that explode as $y \to +\infty$, for the remaining n that cause $\beta_n^{(2)}$ to be complex.

Similarly, the second sum (with $B_n^{(2)}$ coefficients) represents waves travelling away from the grating, in the $+y$ direction. There is a finite set of propagating waves, and an infinite set of exponentially decaying waves that tend to zero as $y \to +\infty$.

Above the grating:

$$U_z^{(2)} = A_0^{(2)} \exp\left(i\alpha_0 x - i\beta_0^{(2)} y\right) + \ldots$$

$$B_{-N}^{(2)} \exp\left(i\alpha_{-N} x + i\beta_{-N}^{(2)} y\right) + \ldots +$$

$$B_{-1}^{(2)} \exp\left(i\alpha_{-1} x + i\beta_{-1}^{(2)} y\right) +$$

$$B_0^{(2)} \exp\left(i\alpha_0 x + i\beta_0^{(2)} y\right) +$$

$$B_1^{(2)} \exp\left(i\alpha_1 x + i\beta_1^{(2)} y\right) +$$

$$\ldots + B_N^{(2)} \exp\left(i\alpha_N x + i\beta_N^{(2)} y\right)$$

(incident)

n = -N

n = -1

n = 0

n = 1

n = N

θ_2 $\theta_{2,0}$

② Refractive index $v_2 = 1$
(vacuum)

① Refractive index v_1
(possibly complex, lossy)

$$\boxed{\begin{array}{l} \alpha_0 = k_2 \sin\theta_2 \\ \alpha_n = \alpha_0 + \dfrac{2\pi n}{d} \end{array}}$$

$\theta_{1,0}$

n = N

n = 1

n = 0

n = -1

n = -N

$$\ldots + A_N^{(1)} \exp\left(i\alpha_N x + i\beta_N^{(1)} y\right)$$

$$A_1^{(1)} \exp\left(i\alpha_1 x + i\beta_1^{(1)} y\right) +$$

$$A_0^{(1)} \exp\left(i\alpha_0 x + i\beta_0^{(1)} y\right) +$$

$$A_{-1}^{(1)} \exp\left(i\alpha_{-1} x + i\beta_{-1}^{(1)} y\right) +$$

Inside the grating: $U_z^{(1)} = A_{-N}^{(1)} \exp\left(i\alpha_{-N} x + i\beta_{-N}^{(1)} y\right) + \ldots +$

$\theta_{2,n}$ $\vec{k}_{2,n}$

$k_{2,n,x} = k_2 \sin\theta_{2,n}$

Reflected wavevectors $\vec{k}_{2,n}$

$k_{1,n,x} = k_1 \sin\theta_{1,n}$

$\theta_{1,n}$ $\vec{k}_{1,n}$

Transmitted wavevectors $\vec{k}_{1,n}$

Figure 2.4: The Rayleigh expansion describes the electric field (TE polarization) or magnetic field (TM polarization) in homogenous media, above and below the grating. The terms in the expansion include a finite number of propagating plane waves – the diffraction orders – and an infinite number of decaying, or "evanescent" plane waves. From the geometry of the diffracted and transmitted wave vectors, we can derive the grating equation, but we need to solve the A_n and B_n coefficients to determine the efficiency of each order.

The total field will have a unique solution only when the incident field u_i is totally specified. From the grating setup in Figure 2.1, we know that we only have a single down-going wave, corresponding to the incident plane wave; all of the $A_n^{(2)}$ for the other propagating waves must be 0. (We label this incident wave as $n = 0$.) We also need to reject the non-physical waves that explode as $y \to +\infty$, therefore the expansion of the field *above the grating* simplifies to:

$$u(x, y) = A_0^{(2)} e^{i\alpha_0 x - i\beta_0^{(2)} y} + \sum_{n=-\infty}^{\infty} B_n^{(2)} e^{i\alpha_n x + i\beta_n^{(2)} y} \tag{2.50}$$

The diffraction grating's reflected orders appear out of this expansion as the finite set of n, $\beta_n^{(2)}$, and $B_n^{(2)}$ values that create propagating plane waves travelling away from the grating. At this point, n can now be properly identified with the *diffraction order*. This is known as the *Rayleigh Expansion* for the diffraction field. (Rayleigh assumed this solution, but did not prove it, in Ref. [38].) Figure 2.4 shows this visually and mathematically.

Below the grating

The field below the grating ($y < 0$) can be expanded using the same technique. One complication is that for grating materials that absorb energy, the refractive index (and hence $k = k_1$) will be complex.[9] In this situation there are two possibilities for the square root $\beta_n^{(1)} = \sqrt{k_1^2 - \alpha_n^2}$. The correct choice can be made by requiring that the diffracted waves remain bounded when $y \to -\infty$. The result provides an expansion for the transmitted field, corresponding to the transmitted orders. These are also shown visually in Figure 2.4:

$$u(x, y) = \sum_{n=-\infty}^{\infty} A_n^{(1)} e^{i\alpha_n x - i\beta_n^{(1)} y} \tag{2.51}$$

The grating equations

These expansions for the reflected and transmitted fields show that as soon as the direction of the incident wavevector is fixed, the outgoing directions of light are determined. For the *reflected* orders, Equation 2.50 gives the x- and y-components of the wavevectors:

$$B_n^{(2)} e^{i(k_x x + k_y y)} = B_n^{(2)} e^{i(\alpha_n x + \beta_n^{(2)} y)} \tag{2.52}$$

$$k_x = \alpha_n \tag{2.53}$$

$$= \alpha_0 + \frac{2\pi n}{d} \tag{2.54}$$

$$= k_2 \sin\theta_2 + \frac{2\pi n}{d} \tag{2.55}$$

From the geometry analyzed in Figure 2.4, the x-component of the outgoing wavevector at angle $\theta_{2,n}$ is $k_2 \sin\theta_{2,n}$. Therefore

$$k_x = k_2 \sin\theta_2 + \frac{2\pi n}{d} = k_2 \sin\theta_{2,n} \tag{2.56}$$

and, after replacing the magnitude of the wavevector above the grating $k_2 = v_2 \omega/c = v_2 2\pi f/c = 2\pi/\lambda$, the famous grating equation (for reflected orders) is finally:

$$\frac{n\lambda}{d} = \sin\theta_{2,n} - \sin\theta_2 \tag{2.57}$$

(where we have assumed that $v_2 = 1$ because we are in free space, and λ is also the free-space wavelength.) This is often called the *Fraunhofer Grating Equation*.

---

[9]In fact, this is the case for all materials at soft x-ray wavelengths.

When $n = 0$, the grating equation reverts to the law of reflection ($\theta_{2,0} = \theta_2$, i.e.: the angle of reflection is equal to the angle of incidence.) This wave corresponds to a classically-reflected wave from a normal surface. The reflected orders that fall *between* the incident wave and the $n = 0$ reflection are referred to as *inside orders*; using our sign convention for diffraction angles, they correspond to $n < 0$. The *outside orders* ($n > 0$, see Figure 2.4) are diffracted at angles beyond the $n = 0$ reflection.

We can also use the same technique to determine the angles of the transmitted orders. Equation 2.51 gives the x- and y-components of the transmitted wavevectors:

$$A_n^{(1)} e^{i(k_x x + k_y y)} = A_n^{(1)} e^{i(\alpha_n x + \beta_n^{(1)} y)} \tag{2.58}$$

$$k_x = \alpha_n = \alpha_0 + \frac{2\pi n}{d} = k_2 \sin \theta_2 + \frac{2\pi n}{d} \tag{2.59}$$

From the geometry analyzed in Figure 2.4, the x-component of the outgoing wavevector at angle $\theta_{1,n}$ is $k_1 \sin \theta_{1,n}$. Therefore

$$k_x = k_2 \sin \theta_2 + \frac{2\pi n}{d} = k_1 \sin \theta_{1,n} \tag{2.60}$$

and, after again replacing k_1, the magnitude of the wavevector below the grating $k_1 = v_1 \omega/c = v_1 2\pi f/c = v_1 2\pi/\lambda$, the transmission grating equation is:

$$\frac{n\lambda}{d} = v_1 \sin \theta_{1,n} - v_2 \sin \theta_2 \tag{2.61}$$

This time, instead of checking for the law of reflection, we can check that when $n = 0$, the transmission equation reverts to Snell's law of refraction ($v_1 \sin \theta_{1,0} = v_2 \sin \theta_2$).

Note: Simplifying β_n

Equations 2.56 and 2.60 provide a useful simplification for β_n. Since

$$k_2 \sin \theta_2 + \frac{2\pi n}{d} = k_2 \sin \theta_{2,n} = k_1 \sin \theta_{1,n} \tag{2.62}$$

we can go back to the expression for β_n, and easily show that

$$\beta_n^{(2)} = \sqrt{k_2^2 - \alpha_n^2} = \sqrt{k_2^2 - (\alpha_0 + 2\pi n/d)^2} \tag{2.63}$$

$$= \sqrt{k_2^2 - (k_2 \sin \theta_2 + 2\pi n/d)^2} \tag{2.64}$$

$$= \sqrt{k_2^2 - (k_2 \sin \theta_{2,n})^2} \tag{2.65}$$

$$= \sqrt{k_2^2 (1 - \sin^2 \theta_{2,n})} \tag{2.66}$$

$$= \sqrt{k_2^2 \cos^2 \theta_{2,n}} \tag{2.67}$$

$$= k_2 \cos \theta_{2,n} \tag{2.68}$$

Similarly, for the transmitted order,

$$\beta_n^{(1)} = k_1 \cos \theta_{1,n} \tag{2.69}$$

Finishing the grating problem

Based only on Maxwell's equations and an assumption of periodicity, we have shown that a grating reflects and transmits light into a set of discrete angles. Incredibly, this result is fully general – it does not depend at all on the shape or nature of the grating profile; all that's required is that it be periodic. Unfortunately, this impressive result still says nothing at all about the *grating efficiency*, or the *amount* of light diffracted into each order. We still do not know anything about the *amplitudes* B_n, A_n of the diffracted plane waves, and to determine these coefficients we will need to get down and dirty inside the grooves of the grating.

Within the grooves, the wave equations (2.18, 2.22) are much more difficult to solve due to the position-dependence of $k(x, y)$. The refractive index (and therefore the impedance k) will change whether inside or outside of a groove; if the grating shape is complicated, $k(x, y)$ will be a complicated function indeed. At this point, we need to apply the numerical integration techniques of the classical differential method, or the eigenvalue method used in the RCW approach. For brevity in this report, we omit those mathematical details here. Before wrapping up the theory section, however, we take a closer look at defining grating efficiency.

2.2   Defining Grating Efficiency

In spectroscopy applications, the experimenter typically illuminates a grating with light and uses a single outgoing order ($n \neq 0$) to resolve the light by wavelength. If they are concerned about optimizing the amount of light delivered to their experiment, the "efficiency" question that matters to them is, "*How much light do I get out of the grating* (in the useful order), *compared to how much light went in?*" Fundamentally, this depends on how much energy is absorbed in the grating itself, and how energy is distributed between orders.

We can define the grating efficiency for a single order quite rigorously in the same way. For electromagnetic plane waves, the *Poynting Vector* represents the energy flux (or energy per unit area, W/m^2) carried by the wave:

$$\vec{S} = \vec{E} \times \vec{H} \tag{2.70}$$

This gives the instantaneous energy flux, which will oscillate in time with the wave. The *time-averaged Poynting vector* gives the average flux delivered over a full period of the wave. For harmonic waves, this turns out to be

$$\bar{S} = \frac{1}{2} \mathbf{Re} \left(\vec{E} \times \vec{H}^* \right) \tag{2.71}$$

where $\vec{H}^*$ denotes the complex conjugate of $\vec{H}$.

Energy is conserved – between incidence, reflection, transmission, and absorption – over a constant grating area.[10] Therefore, we we define the grating efficiency formally as *the ratio of the total time-averaged Poynting flux* – **through a surface parallel to the mean grating plane** – *of the outgoing order* ($\bar{S}_n^{(2)}$) *relative to the incident wave* ($\bar{S}^{(2)}$). Figure 2.5 highlights a convenient surface Q_2 to use for reflected efficiencies ($e_n^{(r)}$), which spans one grating period d in the x-direction, and has unit length in the z-direction. (Due to the periodicity of the fields, any other surface spanning one or more complete grooves will give the same result.) For defining transmitted efficiencies $e_n^{(t)}$, we use a similar surface Q_1 at $y \leq 0$:

$$\textbf{Reflected order efficiency: } e_n^{(r)} \quad \equiv \quad \frac{\iint\limits_{Q_2} \bar{S}_n^{(2)} \, \hat{y} \, \mathrm{d}z \mathrm{d}x}{\iint\limits_{Q_2} \bar{S}^{(2)} \, \hat{y} \, \mathrm{d}z \mathrm{d}x} \tag{2.72}$$

$$\textbf{Transmitted order efficiency: } e_n^{(t)} \quad \equiv \quad \frac{\iint\limits_{Q_1} \bar{S}_n^{(1)} \, \hat{y} \, \mathrm{d}z \mathrm{d}x}{\iint\limits_{Q_1} \bar{S}^{(2)} \, \hat{y} \, \mathrm{d}z \mathrm{d}x} \tag{2.73}$$

It turns out that this definition for efficiency can be nicely expressed in terms of the coefficients in the Rayleigh expansion for the reflected and transmitted fields (2.50, 2.51), so that if we can solve for these coefficients, we'll have determined the grating efficiency:

---

[10]Because the orders propagate away a different angles, it would not be correct to compare intensities based on areas of equal wavefront; instead we need to use an area of constant grating surface.
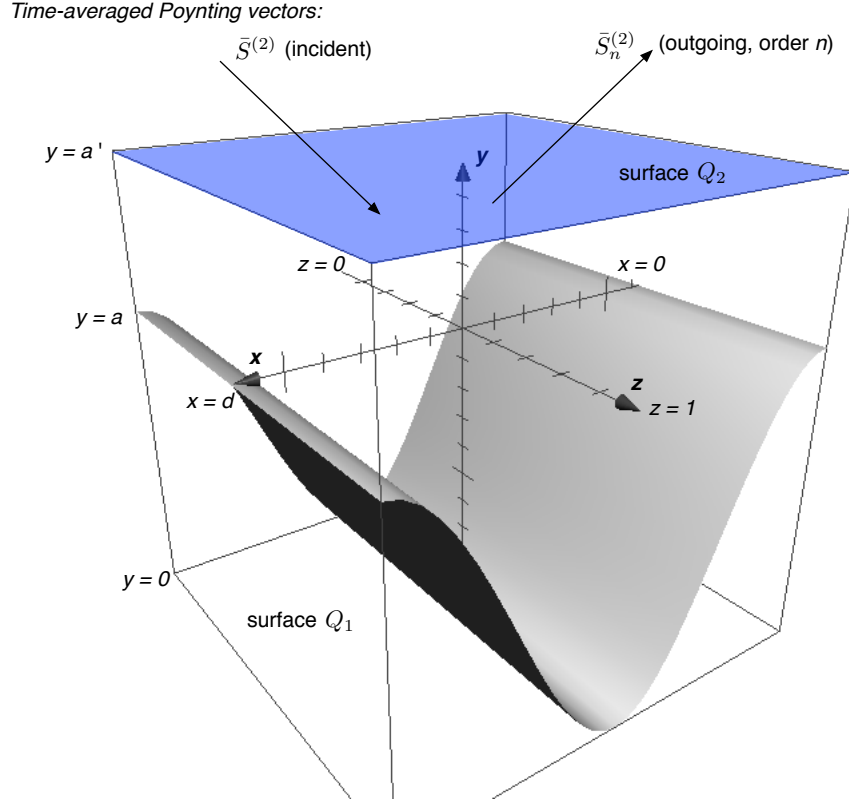
Figure 2.5: The total electromagnetic flux through this highlighted area (Q_2) is used to define the grating efficiency of a diffraction order n, as the ratio of the flux of the diffracted wave $\bar{S}_n^{(2)}$ compared to the incident wave $\bar{S}^{(2)}$.

Reflected Efficiencies

We know that the incident and diffracted orders are plane waves, so the magnetic field $\vec{H}$ is related to the electric field $\vec{E}$ as

$$\left|\vec{H}\right| = \frac{\left|\vec{E}\right|}{Z_2} \tag{2.74}$$

where $Z_2 = v_2 Z_0$ is the impedance of the space above the grating (usually free space, so $v_2 = 1$ and $Z_2 = Z_0 = 377\Omega$, the impedance of free space). Based on the Rayleigh expansion for the outgoing field (equation 2.50), the magnitude of the time-averaged Poynting vector for the *reflected order* is therefore

$$\left|\bar{S}\right| = B_n^{(2)} B_n^{(2)*} \eta_2 \tag{2.75}$$

where we have defined $\boxed{\eta_2 \equiv 1/(2Z_2)}$ in the case of TE polarization, and $\boxed{\eta_2 \equiv 2Z_2}$ for TM polarization. The direction of the vector is along the propagation direction, i.e.: at an angle $\theta_{2,n}$ to the y-axis, therefore the integrated flux through the surface is:

$$\iint_{Q_2} \bar{S}_n^{(2)} \, \hat{y} \, \mathrm{d}z\mathrm{d}x \quad = \quad \int_0^d \int_0^1 \bar{S}_n^{(2)} \, \hat{y} \, \mathrm{d}z\mathrm{d}x \tag{2.76}$$

$$= \quad B_n^{(2)} B_n^{(2)*} \eta_2 \, d \cos\theta_{2,n} \tag{2.77}$$

29

For the incident wave, the integrated flux is similarly

$$\iint\limits_{Q_2} \bar{S}^{(2)} \, \hat{y} \, \mathrm{d}z \mathrm{d}x = A_0^{(2)} \, A_0^{(2)*} \, \eta_2 \, d \cos\theta_2 \tag{2.78}$$

and so the reflected efficiency in order n is:

$$e_n^{(r)} \quad = \quad \frac{B_n^{(2)} \, B_n^{(2)*} \cos\theta_{2,n}}{A_0^{(2)} \, A_0^{(2)*} \cos\theta_2} \tag{2.79}$$

We can simplify this somewhat by choosing a unit amplitude for the incident field, i.e.: $A_0^{(2)} = 1\text{V/m}$ for the TE electric field, or 1A/m for the TM magnetic field. Also, since $\beta_n^{(2)} = k_2 \cos\theta_{2,n}$, we can simplify this to

$$e_n^{(r)} \quad = \quad B_n^{(2)} \, B_n^{(2)*} \, \frac{\beta_n^{(2)}}{\beta_0^{(2)}} \tag{2.80}$$

Transmitted Efficiencies

Gratings used in the soft x-ray regime are always used in reflection, due to the high absorption of materials at these wavelengths. However, since we're working on a general theory with application to all gratings, we can go through the same process for simplifying the transmitted efficiencies.

Again, we define $\boxed{\eta_1 \equiv 1/(2Z_1)}$ in TE polarization, and $\boxed{\eta_1 \equiv 2Z_1}$ in TM polarization, where $Z_1 = v_1 Z_0$ is the impedance of the grating substrate material. The same integrations over the surface Q_1 give the ratio between the total transmitted and incident fluxes:

$$e_n^{(t)} \quad = \quad \frac{A_n^{(1)} \, A_n^{(1)*} \cos\theta_{1,n} \, \eta_1}{A_0^{(2)} \, A_0^{(2)*} \cos\theta_2 \, \eta_2} \tag{2.81}$$

In this case, the difference in impedance above and below the grooves causes η_1 and η_2 to remain in the formula, so the final simplification depends on polarization:

$$e_n^{(t)} = \quad A_n^{(1)} \, A_n^{(1)*} \frac{\beta_n^{(1)}}{\beta_0^{(2)}} \qquad \text{(TE Polarization)} \tag{2.82}$$

$$= \quad A_n^{(1)} \, A_n^{(1)*} \frac{\beta_n^{(1)}}{\beta_0^{(2)}} \left(\frac{v_2}{v_1}\right)^2 \quad \text{(TM Polarization)} \tag{2.83}$$

CHAPTER 3

CURRENT AND FUTURE WORK

Work currently covered in my thesis draft

The rest of my thesis draft covers the implementation of this grating theory, and describes the software I wrote to make efficiency calculations more efficient. Next, I show a comprehensive survey of calculations that can be used by beamline designers to understand trends and factors that affect grating efficiency. After experimentally validating the calculations, I go on to explain the process I used – in collaboration with David Muir – to create an innovative optical design for the REIXS spectrometer. It turns out that for grating instruments like spectrometers and monochromators, resolution and efficiency are inherently in tension. By intelligently balancing the trade-offs between these two goals, and exploiting our understanding of efficiency trends and anomalies, we created a design that is predicted to approach world records in both areas (although obviously not simultaneously). This work was novel because in most cases of beamline design, the grating efficiency is hardly considered, or left to the manufacturer to worry about.

It is even more rare to actually test the gratings for their actual efficiency. The remainder of my thesis draft describes how I characterized the gratings that were manufactured for our spectrometer, and accounted for differences between the predicted and measured efficiency. There are very few comparisons of measured and calculated diffraction efficiencies published for the soft x-ray regime; I conjecture that this is because most beamline scientists are so eager upon receiving new gratings to start doing science on their own beamline, that they hesitate to spend time measuring the gratings on another beamline. In our case, my characterization of the REIXS spectrometer gratings using atomic force microscopy and reflectometer measurements of real-world efficiency turned out to be invaluable: In addition to contributing another set of experimental comparisons, we discovered a substantial manufacturing error in one of the gratings that nearly eliminated its ability to diffract in the useful orders. We also discovered substantial deviations from the nominal blaze angles of some of the gratings, and adjusted our usage of them accordingly. Finally, we confirmed the presence of a surface oxide layer on the nickel-coated gratings, and determined how this affects nickel's ideally high reflectivity at soft x-ray wavelengths. For all of these results, I was able to account for the differences in the real-world gratings using adjusted calculations.

Future work

By using a fitting routine to match calculated efficiencies to reflectometer measurements, I have recently discovered that I can use grating efficiency theory to predict the parameters of real-world gratings (such as groove geometry, density, and even non-ideal effects such as surface roughness and oxidation) based on their measured efficiency. For all of the gratings I studied, the predictions turn out to be consistent with AFM (atomic force microscopy) and PSD (power spectral density) characterization measurements of the actual grating parameters. This shows that grating efficiency calculations are not only useful in the design and optimization of soft x-ray instruments, but can also be used in the reverse direction as a powerful tool in the characterization of their performance.

Another area of future work focusses on improved numerical implementations of grating theory using massively parallel processing and cloud computing. All of the rigorous grating methods are computationally expensive, so it can be tedious to calculate efficiency curves over a range of energies, or conduct an optimization of grating parameters. (This is even more true when doing a large number of calculations to attempt to fit a measured efficiency curve.) The computation can be easily parallelized to take advantage of many processors, but no existing software codes make use of this. I intend to implement one – or ideally several – of the grating theories "from scratch" on a grid computing platform; by combining this computing engine with a web-based user interface, it would become possible for beamline designers around the world to benefit from quasi-instant access to grating efficiency results.

References

[1] J.R. Andrewartha, G.H. Derrick, and R.C. McPhedran. A general modal theory for reflection gratings. *Optica Acta: International Journal of Optics*, 28(11):1501–1516, 1981.

[2] I.C. Botten, M.S. Craig, R.C. McPhedran, J.L. Adams, and J.R. Andrewartha. The dielectric lamellar diffraction grating. *Optica Acta: International Journal of Optics*, 28(3):413–428, 1981.

[3] L.C. Botten, M.S. Craig, R.C. McPhedran, J.L. Adams, and J.R. Andrewartha. The finitely conducting lamellar diffraction grating. *Optica Acta: International Journal of Optics*, 28(8):1087–1102, 1981.

[4] L. Braicovich, J. van den Brink, V. Bisogni, M. Moretti Sala, L. J. P. Ament, N. B. Brookes, G. M. De Luca, M. Salluzzo, T. Schmitt, V. N. Strocov, and G. Ghiringhelli. Magnetic excitations and phase separation in the underdoped $La_{2-x}Sr_xCuO_4$ superconductor measured by resonant inelastic x-ray scattering. *Phys. Rev. Lett.*, 104:077002, Feb 2010.

[5] C. B. Burckhardt. Diffraction of a plane wave at a sinusoidally stratified dielectric grating. *J. Opt. Soc. Am.*, 56(11):1502–1508, Nov 1966.

[6] S. M. Butorin, J.-H. Guo, M. Magnuson, P. Kuiper, and J. Nordgren. Low-energy d - d excitations in MnO studied by resonant x-ray fluorescence spectroscopy. *Phys. Rev. B*, 54:4405–4408, Aug 1996.

[7] F. M. F. de Groot, J. C. Fuggle, B. T. Thole, and G. A. Sawatzky. $2p$ x-ray absorption of $3d$ transition-metal compounds: An atomic multiplet description including the crystal field. *Phys. Rev. B*, 42:5459–5468, Sep 1990.

[8] Boris Gralak Evgeny Popov, Michel Nevière and Gérard Tayeb. Staircase approximation validity for arbitrary-shaped gratings. *J. Opt. Soc. Am. A*, 19(1):33–42, Jan 2002.

[9] Center for X-Ray Optics. X-ray interactions with matter calculator. `http://henke.lbl.gov/optical_constants/`, October 2011.

[10] J. Fraunhofer. Kurzer bericht von den resultaten neuerer versuche über die gesetze des lichtes, und die theorie derselben (Short account of the results of new experiments on the laws of light, and their theory). *Annalen der Physik*, 74(8):337–378, 1823.

[11] J. Fraunhofer, J.S. Ames, and W.H. Wollaston. *Prismatic and diffraction spectra*. Harper's scientific memoirs. Harper & brothers, 1898.

[12] Leonid I. Goray and John F. Seely. Efficiencies of master, replica, and multilayer gratings for the soft-x-ray-extreme-ultraviolet range: Modeling based on the modified integral method and comparisons with measurements. *Appl. Opt.*, 41(7):1434–1445, Mar 2002.

[13] L.I. Goray and S. Yu. Sadov. Numerical modelling of gratings in sensitive cases. *OSA Trends in Optics and Photonics Series*, 75 (Diffractive Optics and Micro-Optics):365, 2002.

[14] Ocean Optics Inc. Usb4000-uv-vis miniature fiber optic spectrometer. `http://www.oceanoptics.com/Products/usb4000uvvis.asp`, October 2011.

[15] Pavageau J. and Bousquet M.J. Diffraction par un reseau conducteur nouvelle methode de resolution. *Journal of Modern Optics*, 17(6):469–478, 1970.

[16] G. Kirchhoff. Zur theorie der lichtstrahlen. *Annalen der Physik*, 254(4):663–695, 1883.

[17] L.I. and Goray. Numerical analysis of the efficiency of multilayer-coated gratings using integral method. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 536(1-2):211 – 221, 2005.

[18] Lifeng Li. Formulation and comparison of two recursive matrix algorithms for modeling layered diffraction gratings. *J. Opt. Soc. Am. A*, 13(5):1024–1035, May 1996.

[19] Lifeng Li. Use of Fourier series in the analysis of discontinuous periodic structures. *J. Opt. Soc. Am. A*, 13(9):1870–1876, Sep 1996.

[20] D. Maystre. Sur la diffraction d'une onde plane par un reseau metallique de conductivite finie. *Optics Communications*, 6(1):50 – 54, 1972.

[21] John Anderson McLeod. Exafs study of amorphous selenium. Master's thesis, University of Saskatchewan, April 2010.

[22] M. G. Moharam and T. K. Gaylord. Rigorous coupled-wave analysis of planar-grating diffraction. *J. Opt. Soc. Am.*, 71(7):811–818, Jul 1981.

[23] M. G. Moharam, Drew A. Pommet, Eric B. Grann, and T. K. Gaylord. Stable implementation of the rigorous coupled-wave analysis for surface-relief gratings: enhanced transmittance matrix approach. *J. Opt. Soc. Am. A*, 12(5):1077–1086, May 1995.

[24] F. Montiel, M. Neviere, and P. Peyrot. Waveguide confinement of cerenkov second-harmonic generation through a graded-index grating coupler: Electromagnetic optimization. *Journal of Modern Optics*, 45(10):2169–2186, 1998.

[25] David Muir. Design of a high performance soft x-ray emission spectrometer for the REIXS beamline at the canadian light source. Master's thesis, University of Saskatchewan, 2006.

[26] M. Nevière and E. Popov. *Light propagation in periodic media: differential theory and design*. Optical engineering. Marcel Dekker, 2003.

[27] W.B. Peatman. *Gratings, mirrors, and slits: beamline design for soft X-ray synchrotron radiation sources*. Gordon and Breach Science Publishers, 1997.

[28] H. Petersen, C. Jung, C. Hellwig, W. B. Peatman, and W. Gudat. Review of plane grating focusing for soft x-ray monochromators. *Review of Scientific Instruments*, 66(1):1–14, 1995.

[29] R. Petit. Contribution à l'étude de la diffraction par un réseau métallique. *Rev. Opt.*, 45:249–276, 1966.

[30] R. Petit and L.C. Botten. *Electromagnetic theory of gratings*. Topics in current physics. Springer-Verlag, 1980.

[31] R. Petit and M. Cadilhac. Electromagnetic theory of gratings: some advances and some comments on the use of the operator formalism. *J. Opt. Soc. Am. A*, 7(9):1666–1674, Sep 1990.

[32] A. Pietzsch, Y.-P. Sun, F. Hennies, Z. Rinkevicius, H. O. Karlsson, T. Schmitt, N. Strocov, V. J. Andersson, B. Kennedy, J. Schlappa, A. Föhlisch, J.-E. Rubensson, and F. Gel'mukhanov. Spatial quantum beats in vibrational resonant inelastic soft x-ray scattering at dissociating states in oxygen. *Phys. Rev. Lett.*, 106:153004, Apr 2011.

[33] A. Pomp. The integral method for coated gratings: Computational cost. *Journal of Modern Optics*, 38(1):109–120, 1991.

[34] Evgeni Popov and Michel Nevière. Grating theory: new equations in Fourier space leading to fast converging results for TM polarization. *J. Opt. Soc. Am. A*, 17(10):1773–1784, Oct 2000.

[35] Evgeny Popov, Boris Chernov, Michel Nevière, and Nicolas Bonod. Differential theory: application to highly conducting gratings. *J. Opt. Soc. Am. A*, 21(2):199–206, Feb 2004.

[36] Evgeny Popov and Michel Nevière. Maxwell equations in fourier space: fast-converging formulation for diffraction by arbitrary shaped, periodic, anisotropic media. *J. Opt. Soc. Am. A*, 18(11):2886–2894, Nov 2001.

[37] J.W.S. Rayleigh. *The theory of sound.* Number v. 2 in The Theory of Sound. Macmillan, 1896.

[38] Lord Rayleigh. On the dynamical theory of gratings. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 79(532):pp. 399–416, 1907.

[39] T. Tamir S. T. Peng and H. L. Bertoni. Theory of periodic dielectric waveguides. *IEEE Trans. Microwave Theory and Techn.*, MTT-23:123–123, 1975.

[40] John F. Seely, Yu. A. Uspenskii, Yu. P. Pershin, V. V. Kondratenko, and A. V. Vinogradov. Skylab 3600 groove /mm replica grating with a scandium-silicon multilayer coating and high normal-incidence efficiency at 38-nm wavelength. *Appl. Opt.*, 41(10):1846–1851, Apr 2002.

[41] A. Wirgin. Selected papers from the ursi symp. *Alta. Freq.*, 38:327–338, 1969.

[42] P. Zeeman. On the influence of magnetism on the nature of the light emitted by a substance. *Astrophysical Journal*, 5:332, 1897.

APPENDIX A

COMPARISON AND APPLICABILITY OF

GRATING THEORY FAMILIES

To choose an appropriate grating theory for a particular problem, we need to understand the assumptions and limitations of each method. As a "grating method user's guide", this appendix presents just enough theory to understand the areas of applicability of the integral method, the modal method, and two implementations of the differential method. This survey is not exhaustive; additional methods have been developed that are only applicable in specific cases. For example, for certain particular groove profiles, there exist coordinate transformations and "conformal mappings", which simplify the boundary conditions so that the grating problem can either be solved analytically, or reduced in complexity. We concentrate here on the methods that have been intended for general applications.

A.1 The Integral Method

For a perfectly conducting metal, when an electromagnetic is incident on the surface, it will induce a surface current $\mathbf{j}_S(\mathbf{M}')$ at each point $\mathbf{M}'$. This surface current radiates a field – the diffracted electric field $E(P)$ at a given point $\mathbf{P}$ above the surface; the field is related to the surface current by the Kirchoff Integral Theorem [16] using the Green's function $G(\mathbf{P}, \mathbf{M}')$:

$$\mathbf{E}(\mathbf{P}) = \int\limits_{\text{grating period}} G(\mathbf{P}, \mathbf{M}')\, \phi(\mathbf{M}')\, \mathrm{d}s' \tag{A.1}$$

where $\mathrm{d}s$ is a curvilinear coordinate along the grating profile, and $\phi(\mathbf{M}')$ is proportional to the surface current. Unfortunately, this becomes tricky, because the surface current is not only induced by the incident electric field, but also by the diffracted field radiated from all other points along the grating surface. This means that the surface current $\mathbf{j}_S(\mathbf{M}')$ at $\mathbf{M}'$ depends not on the incident field, but also on the current $\mathbf{j}_S(\mathbf{M}'')$ at all coordinates $\mathbf{M}''$. By substituting the corresponding $\phi(\mathbf{M}')$ into the equation above, this transforms it into not a differential equation, but an integral equation for $\phi(\mathbf{M}')$.

Limitations

In practice, the numerical techniques required to solve this integral equation can be complicated, depending on the shape of the profile. For imperfectly-conducting gratings, the problem becomes even more difficult, and even harder still when the grating is made up of one or many layers – such as over-coated soft x-ray gratings and multilayers. Although it is universal, the main limitation of the integral approach is its computational cost, and the programming difficulty of applying it correctly to every unique grating situation; numerical challenges abound, particularly regarding the discretization of the grating profile and how sharp corners and discontinuities are handled. Reference [33] presents an overview of integral implementations, as well as approximations for reducing the computation time and handling coating layers. The latest version – referred to as the "modified integral method" by Goray, has been the topic of many recent publications [12] [13] [40] [17].

A.2 Methods using Maxwell's Equations in differential form

The remaining two families (the modal method and the differential method) start from Maxwell's equations in differential form. Within the differential method, there are two variants: the "classical" differential method, and the "Rigorous Coupled Wave" (RCW) approach. We present all three here together because they share

35
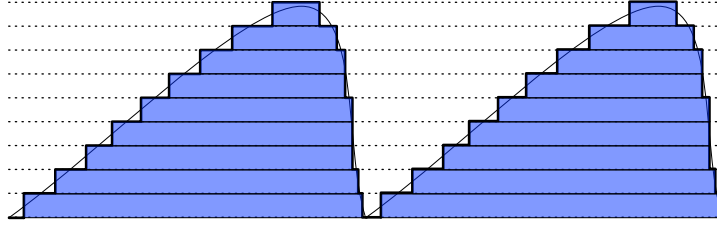
Figure A.1: The modal method and the RCW method approximate every real grating as a stack of rectangular gratings. This simplifies the boundary conditions (the normal and tangential field components line up either along the x- or y-axis), so that the problem can be reduced to an algebraic solution for the eigenvalues of the modes or the Fourier coefficients. Every layer is treated as a separate grating with its own effect on the incoming and outgoing fields; the total grating effect is propagated up the stack using matrix methods.

the same foundation. They all start by projecting the solution to Maxwell's Equations – i.e., the strength of the electric or magnetic field of the waves – onto some periodic basis in the x-direction, but with an unknown dependence on y. The task for each method is to determine the y-dependence by ensuring the total field satisfies the boundary conditions on the grating surface and at infinite. The modal method and the RCW method both assume a rectangular groove profile, so that the boundary conditions are simplified: this implies that the tangential and normal field components at the grating interface are purely along either the x or y-direction.

A.2.1 The modal method

The modal method assumes a basis of functions that are periodic in x, with "modes" F_m related to the field $u_m(x)$ as

$$F_m(x, y) = u_m(x) \exp\left(i\rho_m y\right) \tag{A.2}$$

where ρ_m is a "mode constant". The total field is assumed to be a linear combination of the modes with unknown amplitudes, and since this field must satisfy the boundary conditions on the surface, it is possible to create a set of linear algebraic equations which can be solved for the mode amplitudes.

Limitations

For the modal method, the number of modes and the mode constants ρ_m must be sufficient to fully describe the actual field. For highly conducting materials, finding appropriate mode constants is difficult, but a technique is available in Ref. [1].

The other obvious limitation of the modal method is to rectangular gratings with vertical groove boundaries. We can attempt to avoid this limitation by using stacks of thin rectangular gratings to represent an arbitrary shape, as shown in Figure A.1. This is called the "staircase approximation" for obvious reasons, and is also used in the RCW method. Unfortunately, no matter how thin the slices are made, this approximation creates sharp edges between the layers at the stair-step corners. Popov et. al. discovered that for TM-polarized light, this creates unrealistically strong electric fields at the corners, which cannot be physically representative of the true electric field near a smooth grating. Therefore, algorithms that use the staircase approximation require a larger number of basis components to represent this abruptly changing field, which slows down the computation. More importantly, even when there are enough components to let the solution converge, the results are not representative of reality. [8]

A.2.2 The differential method

The differential method's "classical solution" uses a periodic field similar to the modal method, except that the basis functions are Fourier terms in the x-direction, and the unknown functions to solve are along the

y-direction:

$$F(x, y) = \exp\left(ik \sin \theta_2 x\right) \sum_m F_m(y) \exp\left(\frac{2\pi i x}{d}\right) \tag{A.3}$$

By matching the total field to the boundary conditions at the top and bottom of the grooves, a set of coupled differential equations is created; this set is solved using a combination of numerical integration and linear algebra, which we describe later in this chapter.

Limitations

The advantage of the classical differential method is that it works for arbitrary groove profiles. However, for rectangular profiles, it is slower than the other two options because of the necessity to perform a numerical integration for each basis component. The pure differential method originally suffered from numerical convergence problems in TM polarization on conductive gratings, but these were resolved in 1995 [18] [19] [34]. The only remaining limitation is that the method assumes the material can be described by a well-behaved complex dielectric constant (or the related refractive index) both above and inside the grating material. This rules out "perfectly conducting" (i.e., perfectly reflecting) gratings, for which it is not reasonable to assign a dielectric constant. This limitation results in a numerical instability when trying to work with nearly perfectly conducting gratings – for example, gold and aluminum metallic gratings used with deep infrared or millimetre light – where the real part of the refractive index falls below 0.1. (A few suggestions have been proposed recently for extending the differential method to highly conducting gratings, such as by substituting a finite-thickness perfectly-conducting layer above an absorbing substrate. [35])

Because this method conducts a numerical integration using assumed starting values along the y-direction from the bottom to the top of the grooves, if the grooves are very deep, the long integration distance will increase the computation time. (Any numerical instabilities associated with the integration will also be increased.) Therefore, the speed and accuracy of the differential method are reduced for gratings with a very large depth-to-period ratio.[1]

A.2.3 The "Rigorous Coupled Wave" approach

The differential method's "Rigorous Coupled Wave" approach is a simplification that can be applied for rectangular gratings. As we mentioned in the case of the modal method, when the groove edges are vertical, the boundary conditions within the grooves are invariant along the y-direction. This means that an eigenvalue technique can be used instead of the full numerical integration along y, which reduces the problem to an algebraic solution, speeds up the calculation, and removes the numerical challenges associated with integrating growing exponentials. As soon as this simplification was proposed, various authors [5] [39] [22] assumed that it could be generalized to arbitrary groove profiles by approximating them with rectangular slices, which could be made as thin as required for a given accuracy; this is the same "staircase approximation" used in the modal method. The RCW approach was used extensively for thirty years, and has been shown to produce fast and accurate results for TE-polarized light on dielectric and absorbing gratings. However, as we mentioned above, the fundamental validity of the staircase approximation was challenged in Ref. [8].

Limitations

As a differential method, the RCW approach is limited to finitely-conducting materials. It is also limited to rectangular gratings, or (using the staircase approximation) to stacks of rectangular layers in TE polarization. It can also produce approximate results in TM polarization for dielectric gratings.

---

[1]This limitation is not strictly limited to the differential approach; it will affect the numerical integration of the integral method as well.

A.3 Comparison

Figure A.2 gives a visual comparison of the limitations and strengths of all four techniques. It shows that in general there is no clear "home-run" universal method; the techniques are complementary rather than supplementary. The optimal choice for a particular class of grating problems depends on the overlap between the grating characteristics and the limitations of each method.

Figure A.2: A visual comparison of the limitations and strengths of the main methods in grating theory: the integral approach, the modal method, the full (numerically integrated) differential method, and differential method's "RCW" staircase-approximation simplification.