LUND UNIVERSITY

Optical Studies For Synchrotron Radiation Beamlines

Grizolli, Walan

2015

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Grizolli, W. (2015). *Optical Studies For Synchrotron Radiation Beamlines*. [Doctoral Thesis (compilation), MAX IV Laboratory]. MAX IV Laboratory, Lund University.

Total number of authors:
1

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Optical Studies for Synchrotron Radiation Beamlines

Synchrotron radiation is a powerful tool for many scientific application and it has greatly contributed to the progress in many fields of science. Storage rings produce synchrotron radiation with outstanding properties and dedicated optics is necessary to exploit these properties. In this work several concepts of optics are studied in order to better use the synchrotron radiation produced at storage rings

Printed by Media-Tryck, Lund University, Sweden 2015

WALAN GRIZOLLI

Optical Studies for Synchrotron Radiation Beamlines

Optical Studies for Synchrotron Radiation Beamlines

Ray Optics, Polarization and Wave Optics

WALAN GRIZOLLI | MAX IV LABORATORY | LUND UNIVERSITY

LUND UNIVERSITY
MAX IV LABORATORY
ISBN 978-91-7623-417-4

Optical studies for synchrotron radiation beamlines

Ray Optics, Polarization and Wave Optics

Walan Grizolli

Doctoral Thesis
2015

Lund University

OPTICAL STUDIES FOR SYNCHROTRON RADIATION BEAMLINES

© 2015 Walan Grizolli
All rights reserved
Printed in Sweden by Media-Tryck, Lund, 2015

Cover: Picture of MAX IV Laboratory, December 2014

MAX IV Laboratory, Lund University
P.O. Box 118
SE–221 00 Lund
Sweden

http://www.maxiv.se

ISBN 978-91-7623-417-4 (printed version)
ISBN 978-91-7623-418-1 (electronic version)

[...] all our science, measured against reality,
is primitive and childlike—and yet it is the most
precious thing we have.
Albert Einstein

ABSTRACT

Synchrotron radiation sources have greatly contributed to the progress in many fields of science. The development of the storage ring technologies has made possible to obtain very low emittance electron beams, which together with the use of undulators allow guiding a very high photon flux into a very small spot required by the experiments.

Development of the sources has been accompanied with equally strong progress in beamline optics, improving further experimental conditions and opening new possibilities in science. This follows from the fact that beamlines have to transport the photon beam from the source to the experiment while conserving the beam quality and the photon flux of the source as good as possible.

This work uses basic concepts of X-ray optics to develop beamlines and beamline instrumentation. First, a beamline design that uses astigmatism is discussed. This project takes advantage of the low vertical emittance of the MAX II and MAX IV storage rings to improve the flexibility of soft X-ray beamlines. Secondly a polarimeter is introduced, a device that characterizes the polarization of the light at the experimental station. In this part a new method is presented for analyzing the polarimeter data. Finally, applying wave optics to X-ray optics is presented and discussed. It is shown how new tools were added into wave propagation modeling, taking into account surface errors of the optical components used in grazing incidence.

For these studies the source properties, X-ray interaction with matter together with many concepts in optics are needed and are also presented here.

POPULÄRVETENSKAPLIG SAMMANFATTNING

Bland alla olika forskningsexperiment som utförs är ett vanligt sätt att man låter en ljusstråle passera genom ett prov, och sedan studerar man resultaten av denna växelverkan mellan ljus och prov. Många ljuskällor har använts i detta syfte, till exempel glödlampor och laserljus. En modern ljuskälla som används för detta ändamål är de så kallade lagringsringarna. I lagringsringarna accelereras elektroner som sedan lagras, så att de kan cirkulera med en hastighet nära ljusets egen. När elektronerna färdas i dessa hastigheter i en krökt bana avger de en speciell typ av ljus - synkrotronljus. Den viktigaste egenskapen hos synkrotronljuset är att det är mycket koncentrerat (ett högt flöde i ett litet område) och att det täcker mycket höga energier, från ultraviolett ljus (UV) till röntgenstrålning (UV- och röntgenstrålning är till sin natur samma som synligt ljus, men har högre energi).

Innan UV-ljuset eller röntgenstrålningen når provet som skall studeras måste deras egenskaper i de flesta fall ändras för att uppfylla kraven i experimenten. Detta görs med optiska element såsom speglar, linser, gitter och kristaller. Grundtanken liknar det som görs i experiment med synligt ljus, men UV-ljus och röntgenstrålning har vissa egenskaper som ställer speciella krav på de optiska enheterna. Exempelvis måste speglar för röntgenstrålning fungera vid extremt små vinklar mellan ljusstrålen och spegelytan.

Detta arbete behandlar speciallösningar där man använder UV - och röntgenoptik utvecklade för synkrotronljus. En av dessa är inriktad på en optisk konstruktion som utnyttjar den höga kvaliteten hos MAX II och MAX IV ljuskällor för att förbättra flexibiliteten i experimenten. Nästa del av arbetet beskriver en anordning för att mäta polarisationen hos ljus (en av ljusets egenskaper), som passerar genom provet. I den sista delen undersöks hur vågegenskaper hos röntgenstrålning påverkar optikens prestanda.

POPULAR SCIENTIFIC INTRODUCTION

Among the large variety of experiments performed in science, a common one is to pass a light beam through a sample, the object to be studied, and to study the outcomes of the light-sample interaction. Many light sources have been used for this purpose, like light bulbs and lasers. A modern light source used nowadays is the so-called storage ring. In the storage rings electrons are accelerated and then stored, so that they circulate at a speed close to the speed of light. At these speeds the electrons emit a special kind of light when they travel in a bent path: the synchrotron light. The main property of the synchrotron light is that it is very concentrated (a high flux in a small area) and that it covers very high energies, from the ultraviolet light (UV) to the X-rays (UV and X-rays are in their nature just like visible light, but with higher energy).

Before reaching the sample under study, the properties of the UV light and of the X-rays must be modified to meet the requirements of the experiments. This is done with optical elements like mirrors, lenses, gratings and crystals. The concept is similar to what is done in experiments with visible light, but UV light and X-rays have some particularities which require operating the optical devices in different manners. For instance, mirrors for X-ray need to operate at very small angles between the light beam and the mirror surface.

This work discusses special solutions using UV and X-ray optics developed for synchrotron light. One of them focuses in an optical design that exploits the high quality of the light sources in MAX II and in MAX IV in order to improve the flexibility of the experiments. The second part of the work describes a device for measuring the polarization of light (one of its properties) that passes through the sample. The last part studies how the wave properties of light affects the performance of the optics.

ACKNOWLEDGEMENTS

During my career I've been fortunate to meet great people who have supported my work in many manners. I take this opportunity to express my gratitude to them.

First and foremost, I want to express my deep and sincere gratitude to my supervisor Dr. Rami Sankari for his invaluable guidance during my studies. His supervision and knowledge have guided me in becoming a better professional. Most of all, his respect and trust in my work and opinions (even the bad ones) always make me feel comfortable in my work (and in expressing bad opinions) and encourage me in being autonomous. I would also like to thank my assistant supervisor Prof. Ralf Nyholm, who has been very supportive in all stages of this work.

I am indebted to Dr. Johannes Bahrdt and Dr. Uwe Flechsig, respectively from BESSY and Swiss Light Source. Their invaluable guidances were of foremost importance in developing the wave optics part of this project. I am very grateful to Dr. Joakim Laksman for a productive collaboration in the commissioning of the polarimeters. I am also grateful to Christian Stråhlman, not only for his friendship but also for helping with less scientific parts of the projects.

I am very grateful by the support received from MAX-lab personnel and MAX N-fak department. It has been a privilege to be part of these teams. In particular, I would like to express my gratitude to the beamline support teams who assisted the experiments: Mats Leandersson at beamline I3; Margit Andersson, Conny Såthe and Samuli Urpelainen at SPECIES beamline; Maxim Tchaplyguine and Mikko-Heikki Mikkelä at I411; and Gunnar Öhrwall at beamline I1011. Their precious support was essential for the success of the experiments. A special thanks to Gunnar for evaluating this thesis in a very short notice.

This work had financial support from several institutions, and I hereby acknowledge their valuable support: Swedish National Infrastructure for Computing for providing high-performance computational resources at LUNARC; Lund University Natural Science Faculty and SPIE (International Society for Optical Engineering) for providing travel grants; and DAAD (Deutscher Akademischer

Austauschdienst) for funding my research visit at BESSY.

Finally I want to express my deepest gratitude to the support given by my family and friends. Your support and motivation were of foremost importance in pursuing my PhD and in my personal projects.

Expresso aqui minha profunda gratidão ao suporte recebido de minha família e amigos. O suporte e a motivação recebida de todos foram de extrema importância durante meu doutorado e em meus projetos pessoais.

List of Publications

This thesis is based on the following papers, which will be referred to by their Roman numerals in the text.

I **Use of astigmatic re-focusing at HP-XPS end-station**
W. Grizolli, F. Hennies, J. Knudsen, R. Nyholm, R. Sankari , J Schnadt.
Journal of Physics: Conference Serie **15**, 152005 (2013).

II **The SPECIES beamline at MAX-lab: a facility for soft-X-ray RIXS and HP-XPS experiments**
Såthe Conny; Urpelainen Samuli; Grizolli Walan; Head Ashley; Andersson Margit; Huang Shih-Wen; Jensen Brian; Wallén Erik; Sankari Rami; Nyholm Ralf; Lindberg Mirjam; Sjöblom Peter; Johansson Niclas; Reinecke Benjamin; Knudsen Jan; Schnadt Joachim; Hennies Franz.
Manuscript, in preparation.

III **Simplification of data analysis and experimental error propagation for a VUV polarimeter**
W. Grizolli, R. Nyholm, R. Sankari.
Manuscript, in preparation.

IV **Multilayer based EUV polarimeter at MAX IV laboratory**
W. Grizolli, J. Laksman, F. Hennies, B. Jensen, R. Nyholm, R. Sankari.
Manuscript, in preparation.

V **Propagation of coherent light pulses with PHASE**
J. Bahrdt, U. Flechsig, W. Grizolli, and F. Siewert.
Proc. SPIE 9209 **920908**, doi:10.1117/12.2065228.

VI **Design for a coherent imaging beamline SoftiMAX at MAX IV Laboratory**
R. Sankari, K. Schulte, S. Eisebitt, K. Klementiev, W. Grizolli, R. Nyholm.
Manuscript, in preparation.

ABBREVIATIONS

DLSR	Diffraction Limited Storage Ring
EM	Electromagnetic
EPU	Elliptically Polarizing Undulators
FEL	Free electron laser
FFT	Fast Fourier Transform
FWHM	full width at half maximum
LINAC	Linear accelerator
NSLS	National Synchrotron Light Source
PGM	Plane Grating Monochromator
rms	root mean square
SGM	Spherical Grating Monochromator
SR	Synchrotron Radiation
SRW	Synchrotron Radiation Workshop
TE	Transverse electric (EM waves)
TGM	Toroidal Grating Monochromator
TM	Transverse Magnetic (EM waves)
VUV	Vacuum Ultraviolet

CONTENTS

1 Introduction **1**

2 Ray optics applied to PGM beamlines **5**
2.1 Undulators 7
2.2 Petersen geometry 17
2.3 cPGM 25
2.4 cPGM with astigmatic re-focusing 27

3 Polarimetry **37**
3.1 Polarized light at synchrotron radiation sources 38
3.2 Optical properties of metals 47
3.3 Stokes Parameters and Mueller Matrices 50
3.4 Polarimeter 53
3.5 Determination of the polarization state of the light 58

4 Wave Optics Applied to SR **61**
4.1 Diffraction Limited Storage Rings 61
4.2 Wave Optics 65
4.3 Diffraction Theory 68
4.4 Statistical Optics and Coherence 79

5 Conclusion **89**

Appendix A - Stokes parameters and partially polarized light **91**

Appendix B - Mueller Matrix applied for the three mirror unit **95**

Appendix C - Special Cases for Polarimetry **97**

References **107**

Comments on the Papers **113**

Papers

I	**Use of astigmatic re-focusing at HP-XPS end-station**	**117**
II	**The SPECIES beamline at MAX-lab: a facility for soft-X-ray RIXS and HP-XPS experiments**	**123**
III	**Simplification of data analysis and experimental error propagation for a VUV polarimeter**	**131**
IV	**Multilayer based EUV polarimeter at MAX IV laboratory**	**137**
V	**Propagation of coherent light pulses with PHASE**	**147**
VI	**Design for a coherent imaging beamline SoftiMAX at MAX IV Laboratory**	**167**

CHAPTER 1

INTRODUCTION

Synchrotron Radiation from the Storage Rings

The unique properties of synchrotron radiation makes it a powerful tool for a wide variety of experiments in science. Synchrotron radiation (SR) is produced by charged particles at relativistic speed when they are accelerated [Jackson, 1997]. Several light sources have been developed in order to exploit this phenomenon, for instance storage rings, free electron lasers (FEL) and energy recovery linear accelerators [Als-Nielsen and McMorrow, 2001].

In a typical synchrotron light source, electrons are first accelerated up to relativist speed by a linear accelerator (LINAC) and then stored in a circular accelerator, the storage ring (see Fig. 1.1). The stored electrons produce light when they are forced to oscillate ("*wiggle*") around their main orbit by magnetic structures, since change of direction is equivalent to acceleration. These magnetic structures are based in a periodic assembly of magnetic poles which are called wigglers and undulators, where the assignment depends on the strength of the magnetic field of the poles. Light is also emitted when the electrons pass by the dipoles magnets which keep the electrons in the circular orbit of the storage ring. Light sources based on the light emitted by the dipole magnets are called second generation light sources, while the third generation sources are based on wigglers and undulators.

Third generation storage rings distinguish themselves when compared to other kind of light sources (*e.g.*lasers and discharge lamps) by the production of very high brilliance photon beams, *i.e.*, a very intense beam of very small dimensions and divergences (typically $\approx 10^{15}$ photons in a cross section of $\approx 100\mu m^2$ with a divergence $\approx 100\mu rad^2$). In addition, the light generated in the storage rings covers a broad energy range from ultraviolet light up to hard X-rays (*i.e.*, between ≈10eV and ≈100KeV). Further advances in storage ring and undulator technologies, like the development of low emittance stor-

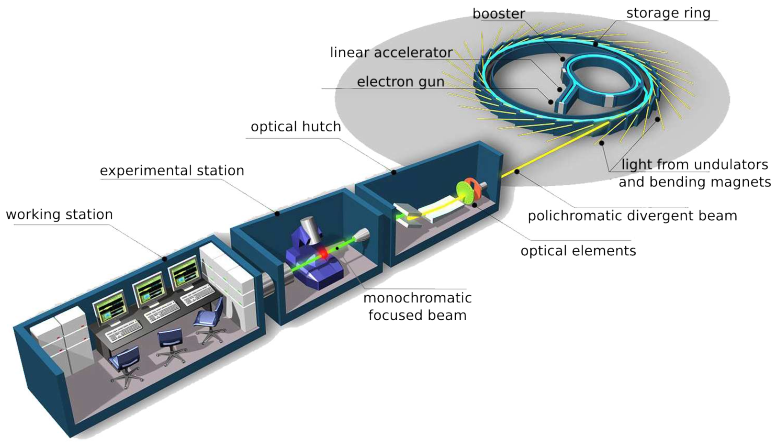


Figure 1.1. *General representation of a storage ring and of a beamline. Electrons are first generated in the electron gun and then accelerated by a linear accelerator. Many facilities use an intermediary circular accelerator, the booster, to inject the electrons at the correct energy in the storage ring. The electrons are then kept circulating in the storage ring, where they emit radiation when they pass trough undulators and bending magnets (not shown in the figure). The beamline optics collects and transports the light to the experimental station. During this process the light is made monochromatic and refocused to a small spot. In the experimental station the light interacts with the sample to be studied and dedicated instruments detect the outcome of the light-sample interaction. Credit of the image: Synchrotron SOLEIL.*

age rings and the development of elliptically polarizing undulators, added new possibilities like even higher brilliance and control of the polarization of the light.

More recently new developments are made possible by a new class of storage rings, so-called diffraction-limited storage rings (DLSR) [Eriksson et al., 2014]. The light from the DLSRs has a high degree of coherence, a feature that was earlier possible to obtain only with FELs for the X-rays regime[1]. This new class of storage rings relies on the multi-bend achromat lattice concept [Einfeld et al., 2014], and MAX IV 3GeV storage ring is the first to be built based on such concept [Tavares et al., 2014].

Optics for Synchrotron Radiation

To exploit the possibilities offered by synchrotron radiation from the storage rings, a dedicated optical setup is needed to modify the light

[1] Although FELs and storage rings cover the same energy range, they produce light with different characteristics. In the case of coherent flux, FELs have a higher peak value while storage rings have a higher integrated flux. It is out of the scope of this work to make a systematic comparison and it is enough to note that although similar, the differences make them suitable to slightly different scientific applications.

so that requirements set by experiments are met, while keeping the good properties of the radiation from source. Synchrotron light, although well collimated, has nevertheless some natural divergence and covers a broad energy range. Therefore, the primary functions of the optical setup are to refocus the light and to select a specific photon energy (*i.e.*, to make the light more monochromatic). These optical setups are called beamlines (also seen in Fig. 1.1) and consist of optical devices for X-ray, for instance grazing incidence mirrors, grating and crystal-based monochromators, compound refractive lenses, multilayer mirrors and zone plates, among others.

The variety of scientific applications of synchrotron radiation means that designing specific beamlines will focus in different aspects and properties of light. To list few examples, photoelectron spectroscopy has high requirements regarding energy resolution; diffraction and scattering techniques have stringent limits for beam divergence; and X-ray magnetic circular dichroism requires circularly polarized light. More recently, DLSRs have made possible the advent of many techniques that profit from the coherence of the X-ray beam, like coherent X-ray imaging. High brilliance is advantageous for the majority of techniques, for instance by making possible measurements at small regions (micro and nano probes).

The different properties of light are addresses by various topics of optics. One can say that theory of optics is formed by few layers as shown in Fig. 1.2. The outermost layer presents the most complete optical model and includes all the other, less complete models. As we move to inner layers, approximations are used and as we move to the outermost layer the complexity of the models increases. Being more specific, the division between quantum optics and electromagnetic (EM) optics is set by the limit between classical and quantum optics. EM optics assumes light as EM waves subject to Maxwell's equation. It describes to a large extent the interaction of light with matter[2] and the vectorial properties of the EM wave, which in turn gives rise to the description of polarization of light. In the cases where these vectorial properties are not relevant, the vectorial wave can be approximated by a scalar wave and this results in wave optics. Wave optics describes the wave properties of light and some phenomena like diffraction and interference. If the wavelength of light is much smaller than the dimensions of the optical setups, light can be treated as rays and the problem is treated within ray (geometric) optics. An additional and important topic is statistical optics, which can actually be described as part of wave optics: it describes the light generated by an ensemble of sources (like the individual electrons in the storage ring) and is necessary for modeling partially coherent light, which is in fact the case of DLSRs at some photon energies. It is important to note that there is actually no hard limits between the layers and many problems can be discussed across them, for in-

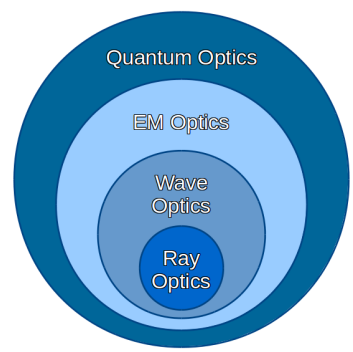


Figure 1.2. *Schematic representation of optic theory and its sub topics.*

[2] In the cases where a quantum physical description is not required.

stance considering rays with EM properties.

This work discusses how these different layers of optics are applied to beamline optics and beamline instrumentation at MAX II and MAX IV storage rings. More specifically, ray optics, polarization (EM optics) and wave optics are used and the text is divided in three chapters accordingly: first ray optics is used to model a soft X-ray beamlines, secondly EM optics is used to understand and measure polarization, and finally wave optics tools are used to study beamlines at DLSR in the coherent and partially coherent regime.

CHAPTER 2

RAY OPTICS APPLIED TO PGM BEAMLINES

The geometric properties of the synchrotron radiation have a dependence in the photon energy and as a consequence light has different properties at the opposite ends of the energy spectrum. For hard X-rays the wave effects due to the wave properties of light are very often negligible. For this case geometric optics can be used to model very accurately the performance of a beamline. For soft X-rays the wave effects are more evident and they are indeed observed, but for many applications they have only minor effects in the performance of the beamline and they can be neglected. The cases where wave effects are relevant will be discussed in Ch. 4. This chapter shows that many soft X-ray beamlines can be modeled by ignoring wave effects and by taking advantage of simpler and more efficient calculations provided by ray optics.

The features of beamline optics are dictated by the properties of the interaction of the X-rays with matter. In the soft X-ray regime this has two main consequences: first, it is not practical to use optical devices like lenses and gratings based on transmission of the radiation due to the high absorption of materials at these photon energies. Therefore, the most common optical components in a soft X-ray beamlines are based on reflection, like mirrors and reflecting diffraction gratings. Secondly, the refractive index in this energy range is lower than but very close to unity, which requires mirrors and gratings to be operated in grazing incidence. However, even then the efficiency is still relatively small, about 10% for the gratings and about 80% for the mirrors[Petersen et al., 1995].

Very often geometric optics makes use of paraxial approximation [Born and Wolf, 1999], which in general assumes a small angle between the incoming light and the normal of the optical surfaces. This approximation does not apply for grazing incidence optics, that is, to X-ray optics. Most of the the results in literature apply for paraxial optics (for instance Born and Wolf [1999]; Fowles [1975]; Saleh and Teich [2007]), which means that some caution is required when using

these references for X-ray optics. However ray optics for grazing incidence is a well developed topic, and analytical results are described for instance in [Peatman, 1997]. In addition, computer simulations based on *ray tracing* method make use of geometric optics without assuming paraxial approximation. For these reasons ray tracing has been a valuable tool in modeling beamlines for synchrotron radiation, with very accurate results[Cerrina and del Rio, 2010; Peatman, 1997]. Ray tracing simulations are extensively used in this work and the method is described in more details in pg. 30.

Gratings have been used for more than 200 years to disperse spatially the photons of different wavelengths[Loewen and Popov, 1997] and nowadays they have a central role in synchrotron radiation beamlines as a fundamental part of a soft X-ray monochromator. A typical beamlines in the soft X-ray regime are based on spherical, toroidal or plane gratings monochromators[Peatman, 1997] (referred as SGM, TGM and PGM respectively), used in a variety of geometries (*e.g.* Rowland and Petersen geometry)[Padmore, 1989].

Among the aforementioned monochromators, the PGM based on the Petersen geometry[Petersen, 1982] has been very successful compared to others because its ease of operation. This is due to a fixed exit slit position and even more due to the possibility to cover a huge energy range with a single grating. Moreover, high quality surfaces are crucial to reach high energy resolution, and planar surfaces can be manufactured to much higher quality than spherical or toroidal surfaces required for SGM and TGM beamlines. Therefore, it is not surprising that PGM beamlines in Petersen geometry have been extensively used in almost every synchrotron radiation facility in the world, including MAX-lab.

As an extension to the PGM in Petersen geometry, the plane grating can be illuminated with collimated light instead of the naturally divergent light from the synchrotron source, resulting in improvements in flexibility. The original Petersen geometry has some constraints that fix the operation mode of the beamline, but the PGM with collimated light (cPGM) in turn is free of such constraints, making possible to change the operation mode of the beamline at any time during the operation. This makes possible to operate for instance in mode for high energy resolution, high grating efficiency or effective suppression of higher order light[Follath, 1997].

In this chapter also the main features of the Petersen type plane grating monochromator are presented and discussed. Then the evolution of this concept up to a new geometry is discussed: the cPGM with astigmatic focus. This new concept was first proposed in Paper **I** and implemented at SPECIES beamline, which is described in Paper **II**. The chapter begins with introducing the geometric properties of the source, since they will play an important role in the beam profile at the experimental station.

2.1 Undulators

In the past the high emittance of the electron beam in storage rings resulted that the geometric properties of the radiation were mostly defined by the properties of the electron beam. However, at low emittance storage rings, like the ones at the forthcoming MAX IV, the characteristic dimensions of the radiation can be bigger than the electron beam dimensions. This is due to the diffraction limit and this affects the apparent source dimensions especially at the soft X-ray and lower energies (the properties of diffraction limited sources will be discussed in more details in Ch. 4.1). Therefore good understanding of the emission process is necessary in order to evaluate the performance of soft X-ray beamlines at low emittance storage rings.

Third generation synchrotron radiation sources are based on insertion devices (undulators and wigglers) placed in storage rings [Als-Nielsen and McMorrow, 2001]. Insertion devices are magnetic structures that force the relativistic electrons in a storage ring to oscillate around the main orbit, resulting in the emission of synchrotron radiation. They play an important role as high brilliance SR sources, since it is always the source that is imaged to the sample location.

Due to their importance, undulators have been widely studied and the literature is extensive ranging from general descriptions [as in Als-Nielsen and McMorrow, 2001; Willmott, 2011] to detailed presentations [like Clark, 2004; Onuki and Elleaume, 2003; Wiedemann, 2003]. The analytical description of the emitted radiation from an undulator leads to complicated results, which will not be completely described in this work. However, the main concepts will be discussed here making use of numerical calculations to demonstrate important properties of undulators radiation.

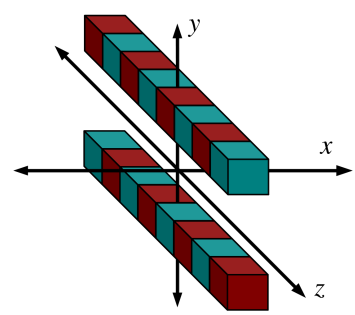


Figure 2.1. *Reference frame related to the undulator. The electrons propagate on the direction +z.*

General Aspects

Different kinds of undulators force the electrons to various oscillatory trajectories. For instance, in planar undulators the electron trajectory oscillates in a single plane (*e.g.* xz plane, coordinate system shown in Figure 2.1) whilst in an elliptical undulator the electron trajectory in the transverse xy plane forms an ellipse.

In order to fully describe the emitted radiation it is necessary to know the electron beam trajectory passing through the undulator, which in turn requires the full knowledge of the magnetic structure of the undulator. However, the geometric properties and the resonance energies of the radiation are a consequence only of the periodicity of the magnetic field in the undulator and no other information is required [Onuki and Elleaume, 2003]. This implies that the various undulator types have common properties that can be addressed before analyzing the specific characteristics in each case. These common properties are presented in this section.

The starting point is to consider the simplest case: radiation emitted by a filament electron beam. In such case, the angular spectral flux of the radiation in the direction θ_x, θ_y with frequency ω is given by [Onuki and Elleaume, 2003]:

$$\frac{d\Phi_n}{d\Omega}(\theta_x,\theta_y,\omega,\hat{u}) = \alpha\frac{I}{e}N^2\left|\vec{h}_n(\theta_x,\theta_y)\hat{u}\right|^2\left[\frac{\sin(\pi N(\omega/\omega_R - n))}{\pi N(\omega/\omega_R - n)}\right]^2 \tag{2.1}$$

where the angles θ_x and θ_y are the angles related to the $\hat{z}$ direction in the planes xz and yz, respectively. n is the harmonic number and the total angular spectral flux is the sum over the harmonics:

$$\frac{d\Phi}{d\Omega}(\theta_x,\theta_y,\omega,\hat{u}) = \sum_{n=1}^{\infty}\frac{d\Phi_n}{d\Omega}(\theta_x,\theta_y,\omega,\hat{u}). \tag{2.2}$$

In these equations N is the number of magnetic periods in the undulator, α the fine structure constant, I the electron beam current in the storage ring and e the electron charge. The complex unit vector $\hat{u}$ is perpendicular to the direction of light propagation and defines the polarization state.

Although the vector term $\vec{h}_n$ in equation (2.1) depends on the harmonic number, the equation shows that actually all the energy dependence is in the last term, inside the square brackets. This is the well known *sinc* function of the form $\mathrm{sinc}(x) = \frac{\sin x}{x}$ that has a maximum value equal to 1 for frequencies $\omega = n\omega_R$(see Figure 2.2 and Figure 2.3). ω_R is called the resonance frequency and $\omega_n = n\omega_R$ is the harmonic frequency of order n. The resonance frequency is given by:

$$\boxed{\hbar\omega_R(\theta_x,\theta_y) = \frac{1.24\cdot 10^{-6}}{\frac{\lambda_U}{2\gamma^2}\left[1+\frac{K_x^2}{2}+\frac{K_y^2}{2}+\gamma^2\theta_x^2+\gamma^2\theta_y^2\right]}[eV].} \tag{2.3}$$

where λ_U is the undulator period, c the speed of light and γ the Lorentz factor of the electrons. The K_x and K_y parameters depend on the magnetic field and on the undulator period[1], and by controlling these parameters it is possible to tune one of the harmonics to the desired energy. An important aspect is that the resonance frequency ω_R depends on the angles of observation. In terms of the resonance wavelength equation (2.3) becomes:

$$\lambda_R(\theta_x,\theta_y) = \frac{\lambda_U}{2\gamma^2}\left[1+\frac{K_x^2}{2}+\frac{K_y^2}{2}+\gamma^2\theta_x^2+\gamma^2\theta_y^2\right]. \tag{2.4}$$

[1]At this point, with no details about the magnetic field, it is not possible to define the K parameters in terms of other undulator parameters in the well know dependence of the magnetic field described later in page 39. However, for this discussion of the general aspects this is not necessary.

An additional property of sinc function is that the value of the function drops to zero when $\frac{\omega}{\omega_R} = n \pm \frac{1}{N}$ (Figures 2.2 and 2.3) . Considering that $\text{sinc}^2(0.442\pi) = \frac{1}{2}$, the FWHM (full width at half maximum) of the distribution can be derived as

$$\frac{\Delta\omega}{\omega_R} = \frac{0.884}{N} \approx \frac{1}{N}. \tag{2.5}$$

Such behavior can be observed comparing numerical examples: the width of the distribution for $N = 2$ is $\Delta\omega/\omega_R \approx 1/2 = 0.5$ in Figure 2.2 and $\Delta\omega/\omega_R \approx 1/20 = 0.05$ in Figure 2.3 when $N = 20$. Typical number of periods in an undulator is $N = 100$, which results in a value of FWHM equal to $\Delta\omega/\omega_R = 0.01$.

Regarding the term $\vec{h}_n(\theta_x, \theta_y)$, it does not have any dependence with photon energy, though it does depend on the harmonic number n, the electron energy, the K parameters and the angles of observation. Therefore, for a single harmonic and fixed operation mode of an undulator (fixed γ and K parameters), this vector depends only on the observation angle through θ_x and θ_y and has no effect over the spectral profile of the harmonic.

Due to the geometry of the radiation process (emission on axis), one should expect that eventually the resulting flux has peaks at the on-axis resonance frequencies $\omega_o = \omega_R(\theta_x = 0, \theta_y = 0)$ and its harmonics $\omega_{no} = n\omega_o$. However, due to the angular dependence of ω_R (equation (2.3)), the contribution of the out of axis resonance frequencies results in a maximum photon flux at energies different of the harmonics $\hbar\omega_{no}$. Since the off-axis energies are smaller than on-axis, the resulting maximum is also smaller. To evaluate this effect we need to study the energy dependence of the angle integrated flux, as discussed next.

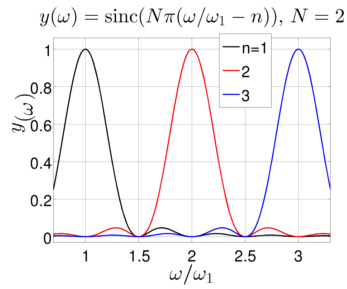


Figure 2.2. *Plot of* sinc *function for N=2.*

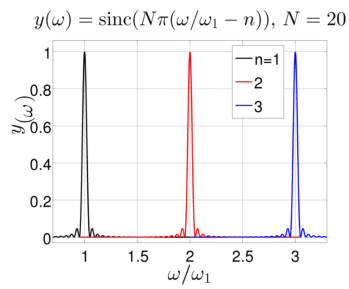


Figure 2.3. *Plot of* sinc *function for N=20.*

Angle integrated spectral flux

Following the discussion of the previous section, we must evaluate the angular integral of equation (2.1). As discussed, the spectral properties depend only on the term of the sinc function on equation (2.1). On the other hand, the angular beam profile depends both on the sinc term and on the vector term $\vec{h}_n(\theta_x, \theta_y)$. However, since the sinc term depends on the angles of observation through ω_R (eq. 2.3) and the value of N is in general large, the variation of the sinc term with (θ_x, θ_y) is much faster than the variation of the vector term. Therefore, the vector term can be assumed constant while studying the angular profile of the radiation, leaving only the sinc term to be considered. This assumption makes possible to approximate the spectral flux in equation (2.1) for small angles θ_x and θ_y as

$$\frac{d\Phi_n}{d\Omega}(\theta_x, \theta_y, \omega, \hat{u}) = \alpha\frac{I}{e}N^2\left|\vec{h}(0,0)\hat{u}\right|^2 \text{sinc}^2\Gamma \tag{2.6}$$

where

$$\Gamma = \pi N \left[\left(\frac{\omega}{\omega_R(\theta_x, \theta_y)} \right) - n \right]. \tag{2.7}$$

Some additional algebraic manipulation allows to rewrite Γ as

$$\Gamma = \frac{\pi L}{2\lambda}\theta^2 + \pi\beta \tag{2.8}$$

where L is the undulator length and λ the radiation wavelength; θ^2 is equal to $\theta_x^2 + \theta_y^2$; β is the detuning parameter defined through

$$\frac{\omega}{\omega_{no}} = \frac{\lambda_{no}}{\lambda} = 1 + \frac{\beta}{nN}. \tag{2.9}$$

For instance, for $\beta = 0$ the observed photon frequency ω is the on axis harmonics ω_{no}.

The result of this approximation is that all the angular dependence is now in the $\text{sinc}^2\Gamma$ function, and then the angle integrated flux $\Phi(\omega, \hat{u})$ is given by

$$\begin{aligned} \Phi(\omega, \hat{u}) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{d\Phi}{d\Omega}(\theta_x, \theta_y, \omega, \hat{u}) d\theta_x d\theta_y, \\ &= \alpha \frac{I}{e} N^2 \left|\vec{h}(0,0)\hat{u}\right|^2 \int_0^{\infty} \text{sinc}^2\Gamma d\Omega. \end{aligned}$$

This means that the calculation of the integrated flux reduces to solving the angular integral of $\text{sinc}^2\Gamma$. Figure 2.4 shows the integral of the sinc term of equation (2.6) as a function of the detuning parameter β. We can note that for $\beta < 0$ the integral value is higher than that for $\beta = 0$. Moreover, at $\beta = -1$ the integral value is roughly twice the value when $\beta = 0$. Regarding the integrated flux, it means that the flux at $\beta = -1$ is two times the flux at the resonant energy ω_n at $\beta = 0$. Mathematically we have:

$$\Phi(\omega', \hat{u}) \approx 2\Phi(\omega_n, \hat{u}). \tag{2.10}$$

where ω' is the frequency at $\beta = -1$ and given by

$$\omega' = \omega_{no}\left(\frac{nN-1}{nN}\right) \tag{2.11}$$

As a result, to obtain the maximum flux at certain frequency ω' the undulator must be tuned (through the K parameter) to a resonance frequency ω_n slightly higher than ω'. For instance, equation (2.11) shows that to generate radiation with $\hbar\omega' = 100eV$ in the first harmonic, an undulator with $N = 39$ periods must be tuned to $\hbar\omega_n = 102.56eV$.

This effect is illustrated in Figure 2.5 considering the undulator recently installed for SPECIES beamline at MAX-lab [Schnadt et al.,

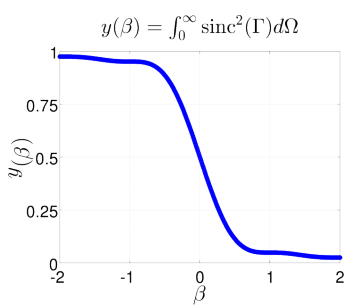


Figure 2.4. *Angular integral of the term* sinc *of the equation (2.6).*

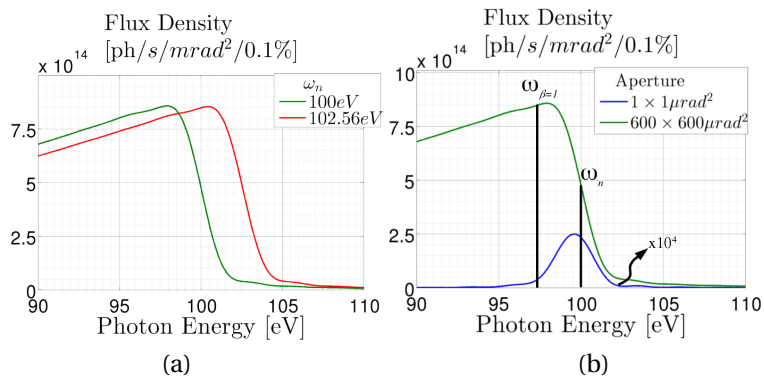


Figure 2.5. *Calculated Flux for SPECIES beamline. (a) Flux through an aperture of* $600x600\mu rad^2$ *for two different tunes of the undulator, at* $100eV$ *and* $102.56eV$*. It can be noted that the flux at* $100eV$ *is higher when the undulator is tuned to* $\hbar\omega_o = 102.56eV$*. (b) Flux at two different apertures when the undulator is tuned to* $\hbar\omega_o = 100eV$ *. Small apertures have a peak flux very close to the on-axis resonance frequency* ω_o*, while at bigger apertures the peak of flux is at a lower energy. Parameters of the undulator listed at Table 2.1.*

Period λ_p	61mm
Number of Periods N_p	39
Ring Current	200mA
Electrons Energy	1.5GeV
Minimum Gap	17mm
Natural Emmitance	$9.726 \cdot 10^{-9}$ m.rad
Coupling Constant	0.1

Table 2.1. *Undulator parameters for SPECIES beamline.*

2012][Paper **I**]. Figure 2.5.a shows the flux when the undulator is tuned to 100eV and to 102.56eV. It is clear that the flux at 100eV is higher when the undulator is tuned to 102.56eV. Figure 2.5.b shows that the peak flux is close to the resonant energy only for very small apertures.

Angular Beam Profile

The knowledge of the angular profile of the radiation has two main applications. Firstly, to determine the beam divergence, which is an important parameter while designing the size of optical components. Secondly, to determine the spatial distribution of the radiation out of the focus (*e.g.* on the optical elements), since at these positions the beam profile is mostly defined by the angular profile.

By using again the approximation $\vec{h}_n(\theta_x, \theta_y) \approx \vec{h}_n(0,0)$ for small angles, this analysis reduces once more to study the behavior of the sinc function. Considering the results of the former section, it is necessary to analyze the beam profile for energies different from the on-axis resonance energies $\hbar\omega_{no}$, *i.e.*, for $\beta \neq 0$. For this reason we continue using the parameterization of equation (2.9). The behavior of the sinc function for different values of β is shown in Figure 2.6 and in Figure 2.7. The graphs show that at the resonance frequency ($\beta = 0$) the beam profile has its maximum value of 1 at the center of the distribution ($\theta = 0$). For photon energies above the resonance energy ($\beta > 0$) the profile looks similar but the peak value is smaller than 1. A different behavior is observed at energies below the resonance ($\beta < 0$): the distribution shows a saddle like structure around the

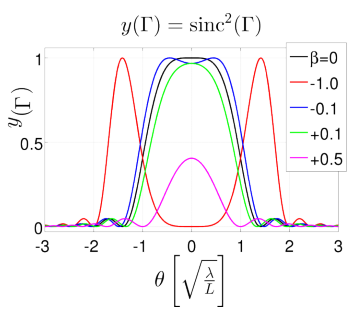


Figure 2.6. $\mathrm{sinc}^2\Gamma$ *for some values of* β*.*

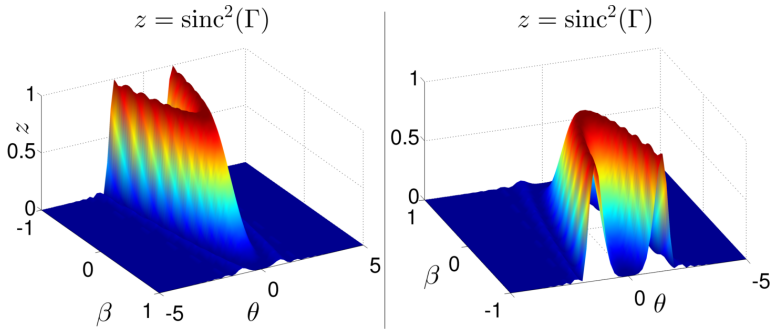


Figure 2.7. $\mathrm{sinc}^2\Gamma$ *in function of the angle of emission and photon energy (implicitly through* β*). The angle is in units of* $\sqrt{\lambda/L}$*. The two figures are the same function seen from two different point of view.*

center ($\theta = 0$). This behavior can be clarified by analyzing the properties of the $\mathrm{sinc}\,\Gamma$ function. Considering that it has a maximum value of 1 when $\Gamma = 0$, and using equation (2.7) we can show that

$$\theta_{\mathrm{peak}} = \sqrt{-\beta\frac{2\lambda}{L}}, \tag{2.12}$$

where θ_{peak} is the angle where the distribution has the maximum value of 1 (see Figure 2.8). Clearly for $\beta = 0$ we have that $\theta_{peak} = 0$ and the maximum intensity is on the axis. For $\beta > 0$ the equation (2.12) has no real roots, what means that the function is different from 1 and therefore has a smaller peak value. Finally, in the case of $\beta < 0$ the function has two roots (positive and negative), which results on the observed two peaks.

With this analysis it is also possible to obtain the aperture ϑ and the full width at half maximum (FWHM) Δ' (Figure 2.8). The first is used to know the minimum angular aperture that collects the whole photon beam. It is given by the value of θ for what the value of Γ is π, and then $\mathrm{sinc}\,\Gamma = 0$. This results in

$$\vartheta = \sqrt{(1-\beta)\frac{2\lambda}{L}}. \tag{2.13}$$

Finally, the FWHM value Δ' is a measure of the beam natural divergence, and is obtained by the value of θ when $\Gamma \approx 0.442\pi$:

$$\Delta' = 2\sqrt{(0.442-\beta)\frac{2\lambda}{L}}. \tag{2.14}$$

Spatial Distribution

The spatial profile of the emitted radiation is obtained by using again the former approximation of small angles, where we can approxi-

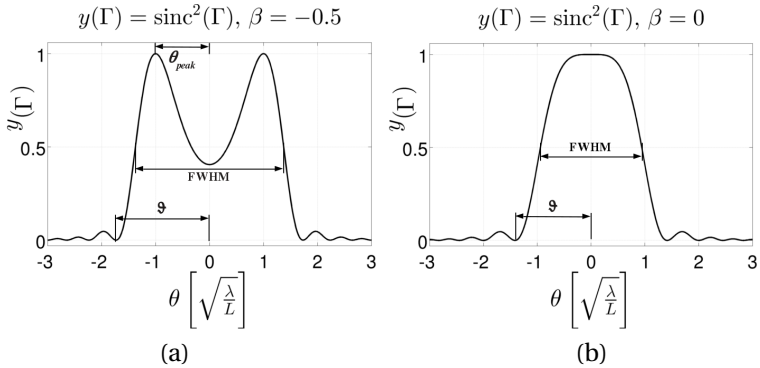


Figure 2.8. *(a) Definition of θ_{peak}, FWHM and ϑ for $\beta > 0$. (b) FWHM and ϑ for $\beta = 0$*

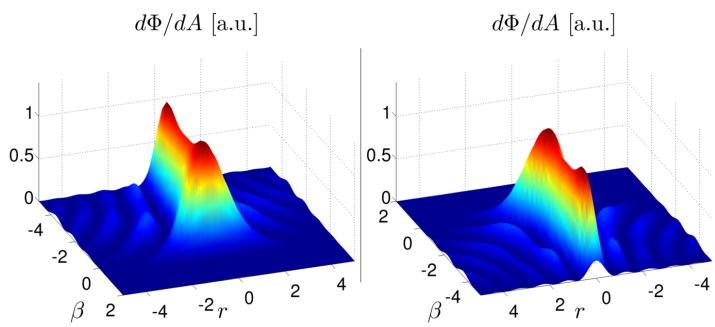


Figure 2.9. *Spatial profile of the radiation in function of the parameter β. The two figures are the same function seen from two different point of view.*

mate $\vec{h}_n(\theta_x, \theta_y) \approx \vec{h}_n(0,0)$. Using such approximation the spectral flux per unit surface $\mathscr{A}$ is given by [Onuki and Elleaume, 2003]:

$$\frac{d\Phi_n}{d\mathscr{A}}(r,\omega,\hat{u}) = \alpha \frac{I}{e} N^2 \left(\frac{\omega}{2\pi c}\right)^2 |\vec{h}_n(0,0)|^2 \times \times \left| \int_{-\infty}^{\infty} \mathrm{sinc}\left[\pi N\left(\frac{\omega}{\omega_R(\theta)} - n\right)\right] J_0\left(\frac{\omega\theta r}{c}\right) \theta\, d\theta \right|^2 \quad (2.15)$$

where J_0 is the Bessel function of zeroth order [Arfken et al., 2012] and r is a radial coordinate $r^2 = x^2 + y^2$. Due to the Bessel function, the analytical analysis is complicated and a numerical method is more informative. Figures (2.10) and (2.9) show the spatial distribution for different energies (through the variation of β, equation (2.9)) around the resonance energy given by β=0.

Although these distributions have a Gaussian-like shape, they are not actually that. To obtain the values of FWHM a Gaussian func-

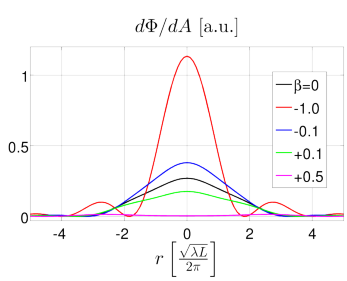


Figure 2.10. *Spatial distribution.*

	β	ϑ	Δ	Δ'	$\Delta\cdot\Delta'$
on-axis resonance	0	$\sqrt{\frac{2\lambda}{L}}$	$2.27\frac{\sqrt{2\lambda L}}{2\pi}$	$1.33\sqrt{\frac{2\lambda}{L}}$	$12.076\cdot\frac{\lambda}{4\pi}$
max. ang. int. flux	-1	$1.414\sqrt{\frac{2\lambda}{L}}$	$1.23\frac{\sqrt{2\lambda L}}{2\pi}$	$2.4\sqrt{\frac{2\lambda}{L}}$	$11.808\cdot\frac{\lambda}{4\pi}$

Table 2.2. *Numeric expressions for aperture, beam size and divergence.*

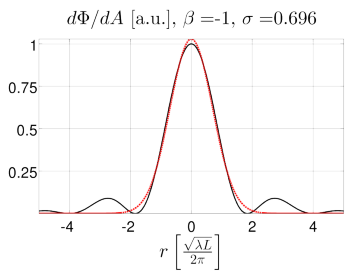


Figure 2.11. *Gaussian function fit to the spatial distribution for $\beta=-1$.*

tion is fitted to the spatial profile, as in Figure 2.11. Like in the angular analysis, the most interesting energy values are at the ones for on-axis resonance ω_{no} and at ω', for which the values of FWHM are listed on Table 2.2. We can note a clear decrease of the beam size when tuning the undulators to ω'.

Due to the similarity between the spatial distribution of the undulator radiation and a Gaussian distribution, undulators radiation is commonly approximated as a Gaussian beam[2] For instance, the emittance ε for Gaussian beams is defined as the product between the standard deviation values of spatial and angular distributions[3] (σ and σ'):

$$\varepsilon_{x,y}=\sigma_{x,y}\sigma'_{x,y}=\frac{\lambda}{4\pi}. \tag{2.16}$$

In order to compare undulator radiation and Gaussian beam we combine the values of Table 2.2 to obtain the undulators emittance. This results

$$\Delta\cdot\Delta'=12.076\cdot\frac{\lambda}{4\pi}\quad\text{for}\quad\beta=0 \tag{2.17}$$

$$\Delta\cdot\Delta'=11.808\cdot\frac{\lambda}{4\pi}\quad\text{for}\quad\beta=-1. \tag{2.18}$$

Since the difference between the two equations is small ($\approx 2\%$), we can assume that

$$\left(\Delta\cdot\Delta'\right)_{\beta=-1}.\approx\left(\Delta\cdot\Delta'\right)_{\beta=0} \tag{2.19}$$

Assuming now that the undulator radiation can be approximated by Gaussian distributions, where the standard deviation σ is equal to $\frac{\Delta}{2.35}$, results in a emittance given as [Onuki and Elleaume, 2003]

[2] It must be noted that Gaussian beam is not a simple beam with Gaussian distribution of rays. Gaussian beam is a solution for the wave equation under paraxial approximation[Saleh and Teich, 2007] and describes a completely coherent beam. See the definition and discussion about Gaussian beam at pg. 67, where the approximation for undulators radiation is also discussed.

[3] Remembering that for a Gaussian distributions the FWHM value Δ is equal to $2\sqrt{2\ln 2}\sigma\approx 2.35\sigma$.

$$\varepsilon_{x,y} = \sigma_{x,y}\sigma'_{x,y} \approx \frac{\lambda}{2\pi}. \qquad \begin{matrix}\text{emittance undulator}\\ \text{radiation}\end{matrix} \tag{2.20}$$

That is, two times larger than a Gaussian beam. This shows that some caution is needed when using Gaussian beams as approximation for undulator radiation.

As conclusion, the analysis of the sinc term shows an increase of flux by a factor 2 when properly detuning the undulator. The drawback is that the angular profile will be non-Gaussian. At the resonance, however, the angular profile is close to a Gaussian, as commonly assumed. Additionally, the FWHM value of the divergence and of the aperture is changed. For instance detuning the undulator to double the flux has the effect of increasing the FWHM of the angular distribution by *ca.* 1.8 times and increasing the aperture by *ca.* 1.4 (see Table 2.2).

Numerical Simulation and Discussion

The descriptions in the former sections are based in analytical models and assuming a filament electron beam. The next step for obtaining a more realistic description is to include the profile of the magnetic field, in order to calculate the vector function $\vec{h}_n(\theta_x, \theta_y)$, and to consider the spatial and angular distributions of the (non filament) electron beam.

In the present work, *SPECTRA* [Tanaka and Kitamura, 2001] software is used for calculating numerically the properties of synchrotron radiation. SPECTRA performs the calculation of the undulator radiation either by using experimental results of the magnetic field or by assuming a sinusoidal magnetic field. In the simple case of a planar undulator (electron oscillating on the xz plane, see Figure 2.1), the magnetic field $\vec{B}_{PU}(z)$ in the sinusoidal approximation is given by:

$$\vec{B}_{PU}(z) = B_y(z)\cdot\hat{y} = B_{yo}\sin(kz+\phi_y)\cdot\hat{y}. \tag{2.21}$$

where $k = 2\pi/\lambda_u$. B_{yo} is the peak value of the magnetic field and depends only on the physical gap between the magnets. Many references present the analytical equations for the planar undulator with a sinusoidal magnetic field [as in Clark, 2004; Onuki and Elleaume, 2003] and these results are extensively used by SPECTRA during the calculations.

To include the spatial and angular profile of the electron beam, it is necessary to convolute the results obtained for the filament beam with the geometric distribution of the electron beam in the storage ring. As a first approximation, both the geometrical distribution of the electron beam and of the radiation emission pattern are approximated by Gaussian distributions with standard deviations $\sigma_{x,y}$ and σ_R, respectively. For Gaussian distributions, quadratic summation

can be used to obtain the total size $\Sigma_{x,y}$ and total divergence $\Sigma'_{x,y}$ of the source as

$$\Sigma^2_{x,y} = \sigma^2_{x,y} + \sigma^2_R, \tag{2.22a}$$

$$\Sigma'^2_{x,y} = \sigma'^2_{x,y} + \sigma'^2_R. \tag{2.22b}$$

However, as discussed previously, maximizing the flux results in a non-Gaussian angular profile (Fig. 2.8), which means that equations 2.22 are only valid at some situations, namely when the Gaussian electron beam divergence is much bigger than the angular distribution of the radiation (*i.e.*, $\sigma'_{x,y} \gg \sigma'_R$). In the past, with high(er) emittance storage rings, this condition was very likely to be achieved. However, as the emittance has reduced over the years this condition is only achieved at very small wavelengths (~1Å) and equations 2.22 are not necessarily valid anymore.

To illustrate the results presented so far, a set of calculation was run with SPECTRA considering again the radiation at SPECIES beamline. The goal of these calculations is to maximize the flux through different sizes of aperture by tuning the undulator. Later, the angular distribution is evaluated at the maximum flux condition.

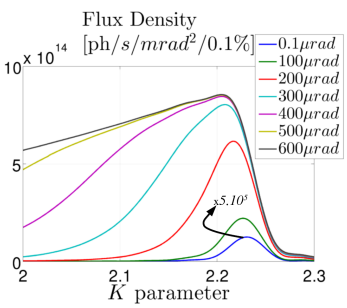


Figure 2.12. *Photon flux at 100e V when the K parameter is scanned at different angular apertures. The apertures have a squared shape with the values on the legend.*

Figure 2.12 shows how the undulator must be tuned towards smaller K values when the aperture size is increased. In this case, the 100eV photon energy on-axis is obtained with K =2.237. For the largest considered aperture of $600 \times 600 \mu m^2$ the K value must be reduced to K =2.2045, which is equivalent to an on axis resonant energy of 102.24eV, very similar to the value of 102.56eV obtained by the analysis performed earlier using equation (2.11) and summarized in Figure 2.5.

The angular distribution corresponding to different K values and to different emittances are shown in Figure 2.13. First in Figure 2.13.(a) the K parameter is tuned to the on-axis photon energy of $100eV$ and the angular profile is very close to a two dimensional Gaussian distribution. However, when tuned to $102.63eV$ to increase the integrated flux, the two peak structure is observed in the vertical direction(Figure 2.13.b). The deviation from Gaussian distribution is only observed in the vertical direction because the electron beam divergence is very small in this direction and thus the radiation profile dominates. On the other hand, in the horizontal direction the electron beam divergence dominates and thus we have a Gaussian distribution in this direction. And finally, Figure 2.13.(c) and Figure 2.13.(d) illustrate the effects of very small and very high emittance electron beams, respectively. At small emittance condition the two peaks structure is dominant in both direction and results in a circular structure. On the other hand, at high emittance condition the Gaussian profile of the electrons angular distribution dominates, and the resulting radiation angular profile presents a two dimensional Gaussian distribution.

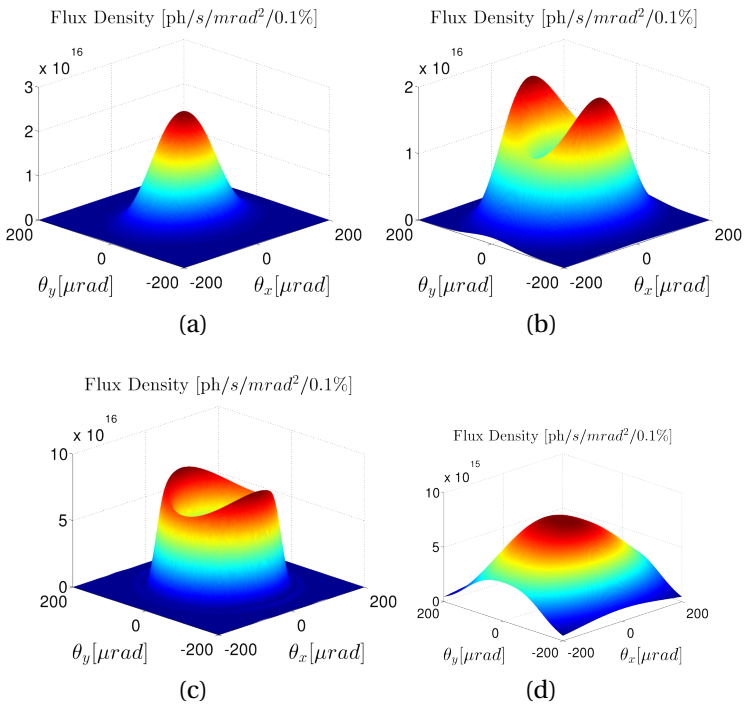


Figure 2.13. *Angular photon distribution at 100eV simulated with SPECTRA for the SPECIES beamline undulator. (a) and (b) MAX II parameters and K parameter tuned to 100eV and 102.63eV, respectively. (c) Low emittance example: K for 102.63eV and MAX IV 3GeV ring parameters. (d) High emittance example: K for 102.63eV and emittance value 10 times the values of MAX II.*

Concluding, the non-Gaussian profiles discussed above illustrate the need to consider the angular emission of radiation by an undulator in low emittance storage rings. The consequence is that for any point out of the focus on the beamline a non-Gaussian profile will be observed. For instance, this will result in a non Gaussian photon distribution on the optical elements and on the diagnostic instrumentation (beam profilers and cameras, for example). Specially, it will affect the beam spatial profile of any out-of-focus location, like in the astigmatic focus proposed on the Paper **I** and on Chapter 2.4, as we will see.

2.2 Petersen geometry

We now concentrate in the optical aspects of a beamline, in particular in the Petersen geometry. The central concept in the Petersen geometry is the focusing property of a plane grating [Murty, 1962]. For synchrotron radiation beamlines this was first applied by Petersen

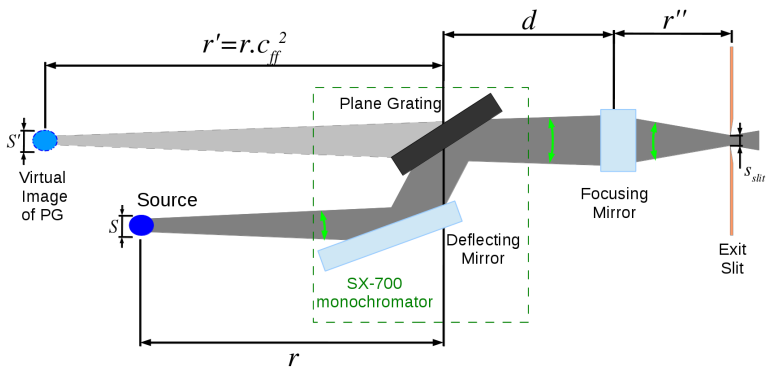


Figure 2.14. *PGM beamline at Petersen geometry. The light gray rays represent the virtual rays that form the virtual image S' at r'.*

[1982], who introduced the fixed focus concept and defined the standard configuration for a PGM beamline: a deflecting plane mirror, a plane grating and a focusing mirror (Figure 2.14).

The plane mirror–plane grating pair forms the so-called SX700 monochromator [Petersen and Baumgärtel, 1980; Riemer and Torge, 1983]. The plane mirror is used to allow that the grating can always be put in such orientation that the exiting beam has always the same direction (*e.g.*, along the horizont). In this geometry, the plane grating is placed at fixed distance from the source and it rotates in the vertical plane around an axis located at the center of the grating surface. The plane mirror adjusts the angle of the incoming beam at the grating. This could be realized by moving and rotating the mirror, but it can actually be done by a single rotation if only the rotation axis is chosen properly (for details see [Petersen and Baumgärtel, 1980], [Riemer and Torge, 1983] and [Petersen, 1985]) .

The focusing property of a plane grating can be studied by applying the Fermat's principle[4], as has been done for instance for optical aberration analysis, but adding a term due to the diffraction on the grating [Noda et al., 1974]. This has been done earlier for spherical gratings [Namioka, 1959; Peatman, 1997] and these results were then applied to plane grating[Murty, 1962; Peatman, 1997], since a plane can be assumed as a spherical surface with infinite radius. Most importantly, such approach results in the diffraction in the grating, described by the grating equation as

$$\boxed{mG\lambda = \sin\alpha + \sin\beta,} \qquad \begin{matrix}\text{grating}\\ \text{equation}\end{matrix} \qquad (2.23)$$

[4]Fermat's Principle states that optical rays traveling between two points in space must follow the trajectory where the optical path length between these two points is an extremum relative to neighboring paths [Born and Wolf, 1999; Saleh and Teich, 2007].

where λ is the wavelength of the diffracted light, m is the order of the diffraction and G the grating line density (in lines per meters). The angles α and β are defined in Figure 2.15.

Related to the focusing property, such analysis results that a source of the size S at a distance r from the plane grating will be focused to a distance r' given by

$$\frac{r'}{r} = -\frac{\cos^2\beta}{\cos^2\alpha}. \qquad \begin{matrix}\text{plane grating}\\ \text{focus}\end{matrix} \tag{2.24}$$

The magnification M_{PG} is defined as the ratio between the image size S' and the source (object) size S and can be shown to be

$$M_{PG} = \frac{S'}{S} = \frac{r'}{r}\frac{\cos\alpha}{\cos\beta}. \qquad \begin{matrix}\text{plane grating}\\ \text{magnification}\end{matrix} \tag{2.25}$$

The negative value of r' means that a virtual image is formed. Moreover, the magnification is very different from a focusing mirror or a focusing lens, as the magnification depends also on the incidence angles α and β.

An additional result from this analysis is that the aberration terms are non-zero[Petersen, 1982], where the most relevant term is the so-called coma aberration[Murty, 1962]. These aberrations, however, have not been found to be critical for the grating performance.

Equation 2.24 shows that the location of the virtual image changes if the ratio $\cos\beta/\cos\alpha$ changes. Moreover, the virtual image of the plane grating is the virtual source of the focusing mirror (see Figure 2.14), which means that the exit slit position also changes for changes of the ratio $\cos\beta/\cos\alpha$. Movable exit slit has been used in many geometries (*e.g.* SGM), but they have the drawback to require a bendable and/or movable mirrors on the refocusing system in the downstream optics, in order to refocus the radiation into the end station. Alternatively, the focusing mirror could be bendable or movable in order to keep the exit slit and upstream optics at fixed positions. However, optical elements at fixed position are easier to operate and thus a virtual source at fixed position would make possible a setup with a higher performance.

In this sense, to keep the virtual source fixed, the plane mirror can be used together with the grating in order to select the angles α and β such that the ratio $\cos\beta/\cos\alpha$ is fixed at a constant value for different wavelengths. Mathematically we have[Petersen, 1982]

$$\boxed{\frac{\cos\beta}{\cos\alpha} = c_{\text{ff}}.} \qquad \begin{matrix}\text{fixed focus}\\ \text{condition}\end{matrix} \tag{2.26}$$

The constant c_{ff} is the so-called fixed focus constant, as it literally fixes the virtual source location r' in equation (2.24) at

$$\boxed{r' = -r \cdot {c_{\text{ff}}}^2.} \qquad \begin{matrix}\text{focus location at}\\ \text{fixed focus condition}\end{matrix} \tag{2.27}$$

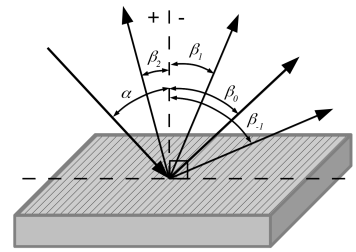


Figure 2.15. *Definition of the grating incidence angle α and diffracted angles β_m, where m is the order of the diffraction. $m = 0$ is due to specular reflection and contain all wavelengths of the incidence beam. The signs close to the grating normal define the convention for the signs of α and β.*

The magnification in equation (2.25) is then given simply by the value of c_{ff}:

$$M_{\mathrm{ff}} = -c_{\mathrm{ff}} = -\sqrt{\left|\frac{r'}{r}\right|}. \qquad \begin{matrix}\text{Magnification at}\\ \text{fixed focus condition}\end{matrix} \tag{2.28}$$

Therefore, by operating at the fixed focus condition it is possible to use a focusing mirror with fixed position and fixed radius of curvature that focuses the dispersed radiation to a fixed exit slit location. In addition, the downstream refocusing optics can also be fixed. With the additional constraint given by the fixed focus condition in equation (2.26) the angles in the grating equations can be determined. This results for angle α

$$\sin\alpha = -\frac{mG\lambda}{{c_{\mathrm{ff}}}^2 - 1} + \sqrt{1 + \left(\frac{mG\lambda \cdot c_{\mathrm{ff}}}{1 - {c_{\mathrm{ff}}}^2}\right)^2}, \qquad \begin{matrix}\text{angle } \alpha \text{ at fixed focus}\\ \text{condition}\end{matrix} \tag{2.29}$$

and for β

$$\sin\beta = \frac{mG\lambda \cdot {c_{\mathrm{ff}}}^2}{{c_{\mathrm{ff}}}^2 - 1} - \sqrt{1 + \left(\frac{mG\lambda \cdot c_{\mathrm{ff}}}{1 - {c_{\mathrm{ff}}}^2}\right)^2}. \qquad \begin{matrix}\text{angle } \beta \text{ at fixed focus}\\ \text{condition}\end{matrix} \tag{2.30}$$

As can be seen from these equations, the angles are defined by the line density of the grating, the order of the diffraction m, the desired wavelength and the chosen c_{ff} value.

Although equations 2.29 and 2.30 seem complicated, they can be easily evaluated and an example is shown in Figure 2.16. We can see that for higher energies and higher values of c_{ff} the incidence angles are higher, *i.e.*, the grating is operating at more grazing incidence angles. This requires long gratings in order to collect all radiation and hence the grating dimensions and geometric acceptance must be carefully evaluated for high c_{ff} values.

Energy Resolution

The resolution $\Delta\lambda/\lambda$ of a Petersen PGM monochromator is defined by many factors, but all of them are somehow related to a variation (or "error") $\Delta\alpha$ or $\Delta\beta$ in the grating angles that results in a variation of wavelength. Mathematically it is represented by contributions

$$\frac{\Delta\lambda}{\Delta\alpha} = \frac{\partial\lambda}{\partial\alpha} \qquad \text{and} \qquad \frac{\Delta\lambda}{\Delta\beta} = \frac{\partial\lambda}{\partial\beta}. \tag{2.31}$$

Applied to the grating equation, this results in the energy resolution terms

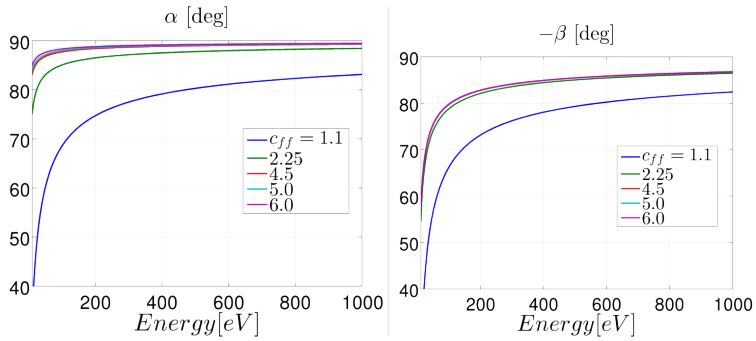


Figure 2.16. *Grating angles in function of photons energy at fixed focus condition for $N = 1221$ lines/mm and $m = 1$.*

$$\left(\frac{\Delta\lambda}{\lambda}\right)_{\alpha} = \frac{\cos\alpha}{mN\lambda}\Delta\alpha \qquad \text{and} \qquad \left(\frac{\Delta\lambda}{\lambda}\right)_{\beta} = \frac{\cos\beta}{mN\lambda}\Delta\beta. \tag{2.32}$$

There are many reasons for the variations $\Delta\alpha$ and $\Delta\beta$, namely the finite source size, the slope errors[5] of the mirrors, imaging aberrations like grating coma and the exit slit size. The effect of coma is small and will not be considered, but all the other terms must be studied in order to evaluate the resulting energy resolution of this type of monochromator. The effects due to exit slit opening and finite source size can be evaluated quite easily. For example, using the notation of Figure 2.14 the finite source size causes a variation of angle $\Delta\alpha$ as

$$\Delta\alpha = \frac{S}{r}, \tag{2.33}$$

and the resolution term corresponding to the finite source size becomes

$$\left(\frac{\Delta\lambda}{\lambda}\right)_{\text{src}} = \frac{\cos\alpha}{mN\lambda}\frac{S}{r} \qquad \begin{matrix}\text{contribution from}\\ \text{finite source size}\end{matrix} \tag{2.34}$$

To obtain the variation due to the exit slit opening s_{slit} it is easier to consider the variation on $\Delta\beta$ due to a virtual source size S' and then use the magnification of the focusing mirror to relate it to s_{slit}:

$$\Delta\beta = \frac{S'}{r'} = \frac{(r'+d)s_{\text{slit}}}{r'r''}, \tag{2.35}$$

[5] Slope errors are consequence of the manufacturing process of the mirrors and of the grating. It is a deviation of the optical elements surface from the ideal surface figure (plane, sphere, toroid, *etc.*). See [Peatman, 1997, Chap. 5.5] for more details.

and by assuming $r' \gg d$ we obtain

$$\Delta\beta = \frac{s_{\text{slit}}}{r''}, \tag{2.36}$$

resulting

$$\left(\frac{\Delta\lambda}{\lambda}\right)_{\text{slit}} = \frac{\cos\beta}{mN\lambda}\frac{s_{\text{slit}}}{r''}. \qquad \begin{matrix}\text{contribution from}\\ \text{exit slit size}\end{matrix} \tag{2.37}$$

The effects of slope errors in the energy resolution are similar to the contributions discussed earlier: the slope errors on the mirrors and on the grating cause a variation $\Delta\alpha$ or $\Delta\beta$ in the grating angles. In the case of the grating, the slope error can be directly related to the variations $\Delta\alpha$ and $\Delta\beta$. For the mirrors the slope errors cause an enlargement of the images, which in turn has an effect similar of a finite source size or of a slit opening. In this way, the derivations of the effects in the energy resolution are similar to the derivations discussed above and results in the following equations[see *e.g.* Follath, 1997]

$$\left(\frac{\Delta\lambda}{\lambda}\right)^{\text{se}}_{\text{po}} = \frac{\cos\alpha}{mN\lambda}\cdot 2\sigma_{\text{po}}, \tag{2.38a}$$

$$\left(\frac{\Delta\lambda}{\lambda}\right)^{\text{se}}_{\text{fo}} = \frac{\cos\beta}{mN\lambda}\cdot 2\sigma_{\text{fo}}, \qquad \begin{matrix}\text{contributions from}\\ \text{slope errors}\end{matrix} \tag{2.38b}$$

$$\left(\frac{\Delta\lambda}{\lambda}\right)^{\text{se}}_{\text{gr}} = \frac{\cos\alpha}{mN\lambda}\cdot 2\sigma_{\text{gr}}\cdot(c_{\text{ff}}+1), \tag{2.38c}$$

where the sub-indexes *po*, *fo*, and *gr* refer to the pre-optics, focusing optics and grating, respectively. The slope errors are assumed to obey Gaussian distribution with standard deviation value of σ and therefore we can use quadratic summing to obtain the total contribution of the slope errors as

$$\left(\frac{\Delta\lambda}{\lambda}\right)^2_{\text{se}} = \left[\left(\frac{\Delta\lambda}{\lambda}\right)^{\text{se}}_{\text{po}}\right]^2 + \left[\left(\frac{\Delta\lambda}{\lambda}\right)^{\text{se}}_{\text{fo}}\right]^2 + \left[\left(\frac{\Delta\lambda}{\lambda}\right)^{\text{se}}_{\text{gr}}\right]^2. \tag{2.39}$$

The values σ in equations (2.38) are the effective values of slope error at each optical element and have different values for the tangential and sagittal directions. For the tangential direction it is equal to the tangential slope error $\sigma = \sigma_t$. In the sagittal direction, however, it has an additional dependence with the grazing incidence angle of the mirrors θ (related to the surface) and is given by $\sigma = \theta\sigma_{sag}$. Due to the use of grazing incidence angles, the values of θ are typically small (*e.g.* 2 deg= 35mrad), and as a result the effective value of slope error is smaller in sagittal focusing[Cash, 1987; Petersen, 1982]. Sagittal focusing also shows smaller coma aberration and thus it has a smaller

effect on the energy resolution. For these reasons, the sagittal focusing is preferred over the tangential one, especially on the pre-optics and in the focusing optics where they are directly affecting the energy resolution.

By assuming that also the contributions of the source size and exit slit follow Gaussian distribution, quadratic summing can be used again to write the final resolution as

$$\left(\frac{\Delta\lambda}{\lambda}\right)_T^2 = \left(\frac{\Delta\lambda}{\lambda}\right)_{\text{src}}^2 + \left(\frac{\Delta\lambda}{\lambda}\right)_{\text{slit}}^2 + \left(\frac{\Delta\lambda}{\lambda}\right)_{\text{se}}^2. \tag{2.40}$$

We know, however, from the discussion in Chapter 2.1 that the spatial distribution of the source is not Gaussian. The exit slit cannot be described by a Gaussian distribution either (a rectangular function would be more suitable). This means that for more realistic analytical results it is necessary to convolute these different contributions, not simply to add them quadratically. Nevertheless, assuming that these quantities convolute like Gaussian distributions (eq. 2.40) we are able to provide a good estimation of the real values and also a good description of how the energy resolution depends on individual parameters (*e.g.* optical distances and slope error). Moreover, much better results can be achieved with ray tracing simulations (to be discussed later) and therefore the use of these analytical results is a reasonable starting point.

In many practical situations the energy resolution is presented in terms of the resolving power R, defined as the inverse of the energy resolution: $R = \frac{\lambda}{\Delta\lambda}$. Eq. 2.40 becomes then

$$\frac{1}{R_T^2} = \frac{1}{R_{\text{src}}^2} + \frac{1}{R_{\text{slit}}^2} + \frac{1}{R_{\text{se}}^2}. \qquad \text{Total Resolving Power} \tag{2.41}$$

The analytical description of the total resolving power R_T^2 can be further simplified by introducing some additional approximations when computing $\sin\alpha$ in equation (2.29). The first step is to estimate the maximum values of the quantities c_{ff}, grating line density G and the wavelength λ. A typical grating has a line density of the order of 1000 lines/mm that results in a maximum wavelength of the order of 100nm (12.4eV).

The choice of c_{ff} value is done in order to optimize the grating efficiency and the resolving power in a broad energy range. Petersen [1982] determined that, for his setup and energy range, operating at c_{ff}=2.25 provides a good performance. A more appropriate method would be to analyze the grating efficiency for different gratings, but this value is commonly accepted as a good condition and widely used in operation. For the current analysis the important point is to define an approximate value of c_{ff} and the value determined by Petersen suggest values of the order of unity.

Using these values results that in equation (2.29) the $G\lambda c_{\mathrm{ff}}$ term is of the order of $\approx 10^{-8}$ and therefore the term inside the brackets can be ignored. Assuming also that $c_{\mathrm{ff}} > 1$, we can write $\sin\alpha$ as

$$\sin\alpha \approx 1 - \frac{mG\lambda}{{c_{\mathrm{ff}}}^2 - 1}, \tag{2.42}$$

resulting $\cos\alpha$ as

$$\cos\alpha \approx \sqrt{\frac{2mG\lambda}{{c_{\mathrm{ff}}}^2 - 1}}. \tag{2.43}$$

Combining now these approximations with the fixed focus condition of equation (2.26) we obtain that the resolving power terms are given by

$$R_{\mathrm{src}} = \frac{r}{S} \cdot T(c_{\mathrm{ff}}), \tag{2.44a}$$

$$R_{\mathrm{slit}} = \frac{r''}{s_{\mathrm{slit}} \cdot c_{\mathrm{ff}}} \cdot T(c_{\mathrm{ff}}), \tag{2.44b}$$

$$R_{\mathrm{po}} = \frac{1}{2\sigma_{\mathrm{po}}} \cdot T(c_{\mathrm{ff}}), \tag{2.44c}$$

$$R_{\mathrm{fo}} = \frac{1}{2\sigma_{\mathrm{fo}} \cdot c_{\mathrm{ff}}} \cdot T(c_{\mathrm{ff}}), \tag{2.44d}$$

$$R_{\mathrm{gr}} = \frac{1}{2\sigma_{\mathrm{gr}} \cdot (c_{\mathrm{ff}} + 1)} \cdot T(c_{\mathrm{ff}}), \tag{2.44e}$$

where the function $T(c_{\mathrm{ff}})$ is defined as (following the notation of [Follath, 1997])

$$T(c_{\mathrm{ff}}) = \sqrt{\frac{mG\lambda}{2}({c_{\mathrm{ff}}}^2 - 1)}. \tag{2.45}$$

This behavior is plotted in Figure 2.17 for different c_{ff} values.

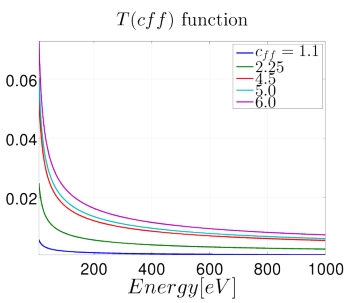


Figure 2.17. *Function $T(c_{ff})$ for a grating with line density of 1221 lines/mm for typical values of c_{ff}.*

To illustrate these results, we now apply this analytical treatment to a real case: beamline I511 at MAX II (main parameters listed in Table 2.3)[Denecke et al., 1999]. Figure 2.18.a shows the individual terms and the total resolving power calculated using equations (2.44), where we see that the behavior of the function $T(c_{\mathrm{ff}})$ (shown in Figure 2.17) dominates the general behavior of all the terms. In addition, Figure 2.18.b shows the total resolving power for different values of c_{ff}, where we observe that the resolving power has a strong dependence on c_{ff} for values in the range $1 < c_{\mathrm{ff}} < 3$. For values of $c_{\mathrm{ff}} > 3$ the resolving power is less sensitive to c_{ff}, which make possible the choose c_{ff} in order to optimize other parameters (*e.g.* flux).

Finally, from equation (2.41) is expected that eventualy one of the terms R_{src}, R_{slit} or R_{se} will be the limiting term of the final resolving power. The limiting term is actually chosen in the design phase of the beamline, when requirements of flux (connected to the opening of the exit slit) and resolution are evaluated considering constraints

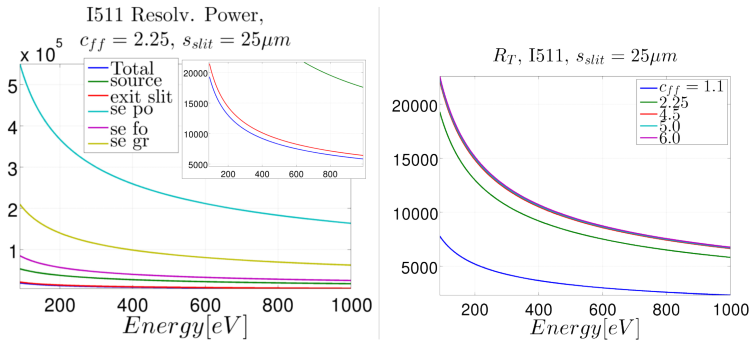


Figure 2.18. *Resolving power values for I511 beamline. Optical distances on Table 2.3. (a) The values of the different components of the resolving power: term due to source size, exit slit aperture, slope error due to pre-optics (se po), slope error due to focusing optics (se fo) and slope error due to the grating (se gr). In the inset a zoom on the total resolving power curve showing the small difference with the term due to the exit slit size. (b) Total resolving power calculated for different values of c_{ff}. The source size dependence of the photon energy, caused by the diffraction limit, is also included.*

like source size and quality limits for optical surfaces. Optical elements with good quality are commonly available nowadays and this is not a limiting factor for ordinary beamlines. The quality of the optical components is indeed specified based on the desired resolving power by using equations (2.44). This means that eventually flux and resolution requirements will define whether the resolving power is limited by the source size or exit slit opening.

In this respect, due to a combination of source size and ring current at MAX II, beamline I511 needs to operate at a relative big exit slit opening ($50\mu m$) to reach high photon flux. This results in a resolving power limited by the slit opening, as can be observed in Figure 2.18.a. We can see in this figure that for I511 beamline the term R_{slit} has the smaller value and thus it is the limiting factor. On the other hand, beamlines at MAX IV will be able to take advantage of the smaller source size and of the higher ring current to operate at very small slit opening (*e.g.* $5\mu m$) up to the limit where the resolving power is source limited.

2.3 cPGM

It was mentioned in the previous sections that the c_{ff} values can be optimized to different modes of operation of the beamline. This means that for a flexible PGM setup it is necessary to be able to adjust the value of c_{ff} during the operation. However, a change in c_{ff} changes the position of the virtual image created by the grating

	I511	SPECIES	HIPPIE
r [m]	15	12	24
r'' [m]	6.487	7	12
electron beam size (FWHM) [μm]	110.73	110.73	9.389
Undulator Length L [m]	2.6615	2.379	3.93

Table 2.3. *Beamlines optical parameters used for the calculation of R_T. The source vertical sizes for I511 and for SPECIES are the size at MAX II and the one for HIPPIE is the value of MAX IV 3GeV ring.*

and thus the entrance arm of the focusing mirror (in the same way as discussed on page 19). This results that at first sight a beamline with variable c_{ff} requires bendable mirrors, or movable exit slit, or both.

To overcome this problem different solutions have been developed. Beamline I311 at MAX-lab for instance uses an exit slit that can be displaced by *ca* 1 meter along the beam in order to adapt to the different entrance and exit arms of the focusing mirror[Nyholm et al., 2001; Petersen et al., 1995]. This makes possible to select c_{ff} in the range between 1.7 and 10 and an optimal refocusing performance was designed for c_{ff}=4.5 (*i.e.*, the optical distances and the focus was optimal only for this condition). To obtain an optimal focus at sample plane for different values of c_{ff} the properties of the focusing or the refocusing mirrors need to be changed. This can be done by moving them or by changing the radius of curvature, the second solution being easier to apply. Therefore, bendable refocusing mirrors were installed in order to follow the focus for any value of c_{ff}. This solution was then also used for beamline I511, although there are many drawbacks related to the operation of bendable mirrors[Follath, 1997], one of them being forced to use a tangential focusing, which is more sensitive to slope errors and coma aberration[Peatman, 1997].

As a major improvement to the Petersen type monochromator, Follath [1997] proposed illuminating the plane grating with collimated light. By doing so, the source of the focusing mirror will always be at infinity, making possible to have an exit slit at a fixed position using a non-bendable focusing mirror. This new setup is normally referred as cPGM and its design and performance are well reported in [Follath et al., 1998] and [Follath, 2001].

The collimation of the light is done with an additional mirror placed before the SX700 monochromator, as shown in Figure 2.19 (the collimating mirror, compare with Figure 2.14). Although this additional optical element results in some reflection losses, it has some advantages, namely the possibility for horizontal focusing and absorption of the high energy radiation, decreasing the thermal load at

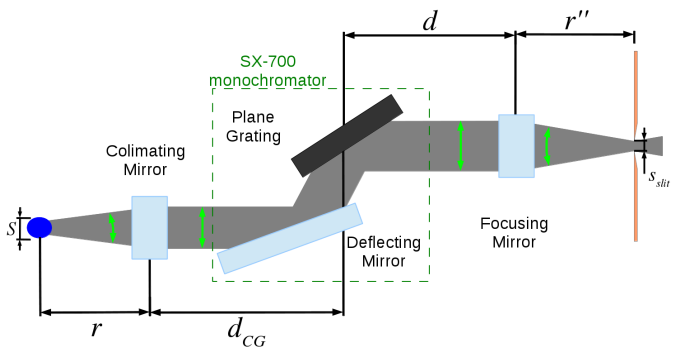


Figure 2.19. *cPGM layout. The notation for the optical distances are defined to make possible the use of the same equations of a Petersen geometry.*

the monochromator [Follath, 1997].

The performance of this geometry is also evaluated in terms of the energy resolution as done in Section 2.2 for the Pettersen geometry. Since there were no assumptions about the divergence of the radiation, the equations 2.27 and 2.28 are still valid. There is only one minor change in the calculation of $\Delta\alpha$, where the entrance arm must be changed from source-to-grating distance to source-to-collimating mirror distance. However, by having the optical distances defined as in Figure 2.19, all the equations for the Petersen geometry in the Section 2.2 are also valid to a cPGM monochromator. Therefore, we can conclude that the general behavior of the energy resolution presented for the Petersen setup will be similar for a cPGM monochromator.

To illustrate this, Figure 2.20 presents the total resolution calculated with the optical parameters of beamlines I511 (Petersen geometry), SPECIES (cPGM)[Schnadt et al., 2012] and HIPPIE (cPGM)[Sankari, 2012]. The beamline parameters are listed in Table 2.3. It is evident that, besides the difference of performance of these setups, the dependences with energy and c_{ff} are quite similar.

The cPGM has, however, the advantage of easily adjust the c_{ff} value during the operation, with no need to move the exit slit and/or adjust the mirrors properties to obtain a focus at a fixed position.

2.4 cPGM with astigmatic re-focusing

The flexibility of a cPGM can be further improved by combining this monochromator geometry with a refocusing system that uses astigmatism. In a normal (stigmatic) re-focusing the exit slit is imaged into the sample position. This results that the vertical spot size has a linear dependence with the exit slit opening, defined by the vertical

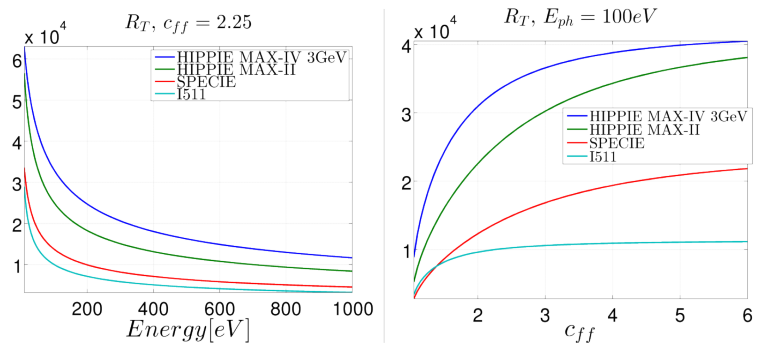


Figure 2.20. *Total Resolution R_T values due to source and exit slit size for SPECIES, I511 and HIPPIE beamlines at (a) c_{ff} =2.25 and (b) 100eV. Optical parameters in Table 2.3. Slit size of 25μm, grating lines density of 1221 lines/mm and $m = 1$. HIPPIE beamline is planned to MAX IV 3GeV ring, but its performance is also calculated considering the electron beam size of MAX II in order to show the improvement due to a low emittance source. Other parameters in Table 2.3.*

magnification of the refocusing system. This is, however, not true for an astigmatic spot where the vertical focus is not at the sample position.

It can be indeed shown that for this astigmatic spot the vertical size is mostly defined by the vertical divergence. We show in Paper **I** that as a consequence the vertical beam size at the sample position is simply determined by the distances between the optical elements, the vertical source divergence and by the c_{ff} value. In the next sections the astigmatism is described and the concept of using it as described in Paper **I** is verified through ray tracing simulation. In order to illustrate this here results for HIPPIE beamline, to be built at MAX IV 3GeV ring, are presented.

Astigmatism

Astigmatism is characterized by having a focus at tangential plane in a different position than in the sagittal plane [Hecht, 2002](see Figure 2.21). It occurs mainly with spherical surfaces where the object is out of the symmetry axis of the lens (as will be discussed soon). It is in general a problem for image formation and focusing of light, but the study of this effect makes possible to correct astigmatism and, as done on this work, to use it to achieve specific goals.

Toroidal mirrors are often used to obtain stigmatic focus(*i.e.*, both tangential and sagittal focus are at same position). The position of the two focuses as a function of the torus radii R and ρ can be derived either by geometric optics[as done by Born and Wolf, 1999, in section 4.6.2] or by using the Fermat's principle[as done by Peatman,

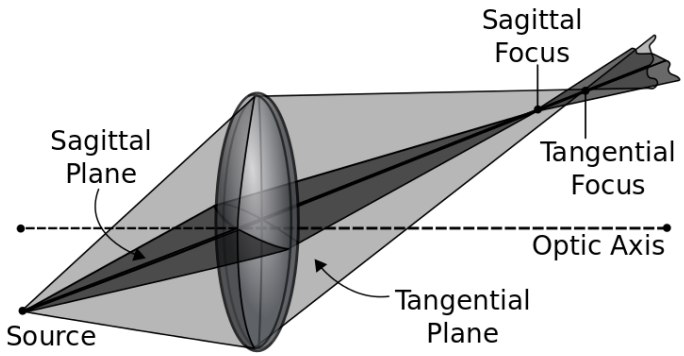


Figure 2.21. *Graphical illustration of the astigmatism effect on a spherical lens. For clarity's sake a lens is used as example instead of a mirror. The rays from an out of axis point source are focused on two different positions on the tangential plane (light-gray) and sagittal plane (dark-gray).*

Image edited from the original "Astigmatism" by I, Sebastian Kroch. Licensed under CC BY-SA 3.0 via Wikimedia Commons, `http://en.wikipedia.org/wiki/File:Astigmatism.svg`

1997]. In both methods we get that

$$\boxed{\frac{1}{p_t}+\frac{1}{q_t}=\frac{2}{R\cos\theta},} \quad \text{tangential focusing} \tag{2.46}$$

$$\boxed{\frac{1}{p_s}+\frac{1}{q_s}=\frac{2\cos\theta}{\rho},} \quad \text{sagittal focusing} \tag{2.47}$$

where p is the source to mirror distance and q the mirror to image distance, θ is the incidence angle, R the tangential radius and ρ the sagittal radius. In a spherical mirror we have $R=\rho$.

As an example, let us consider a spherical mirror operating at $\theta = 88$deg, with the source and image placed at 10m and 5m from the mirror, respectively. Using this condition the tangential focus results in a radius of curvature of $R = 191$m. A sagittal focus with the same radius of curvature $\rho = 191$m results in a image at $r' \approx -10$m, where the negative sign means that a virtual image is formed. This is the same result as obtained for a plane mirror, which means that in the sagittal direction the mirror has no focusing effect at all. To obtain coinciding focuses (stigmatic focus) we can combine equations 2.46 and 2.47 by imposing that $q_t = q_s$. The stigmatic focus is obtained for a radius of curvature $\rho = R\cos^2\theta$, which in this example results in a radius $\rho = 0.232$m, two orders of magnitude smaller than the tangential radius R. Equations 2.46 and 2.47 can also be used to deliberately obtain an astigmatic focus, as discussed next.

Use of astigmatism in a cPGM beamline

Paper **I** and Paper **II** discuss the use and application of astigmatism to improve the performance of a cPGM based beamline, in this case the SPECIES beamline at MAX II. It is studied whether a single toroidal mirror can be used as a refocusing system, rather than two bendable mirrors in a *KB* configuration [Kirkpatrick and Baez, 1948; Peatman, 1997]. However, a conventional re-focusing system based on a single mirror results in a very asymmetric spot, where the vertical dimension would be much smaller than the horizontal (see Paper **I** for a discussion about the advantages of a symmetric spot and about the advantages of a single toroidal mirror over a *KB* system). This asymmetry is due to some hard constraints like the source properties, available space in the experimental hall and required exit slit opening to reach the required resolving power, to name a few. Especially the relatively large horizontal source size of MAX II and the small opening of the exit slit have a big influence, since they are both imaged at the sample position in a stigmatic focus.

In order to match the horizontal dimension of the beam, it was evaluated how astigmatism increases the vertical size of the spot in this single mirror geometry. It turned out that the vertical beam size at the astigmatic focus has a linear dependence with the monochromator operation mode through c_{ff}. It means that in this geometry c_{ff} can be used to control the vertical beam size. This concept is derived and described in details in Paper **I**, and Paper **II** shows the performance of the astigmatic refocusing measured during the commissioning of the beamline. The resulting geometrical model for the vertical spot size and results from simulations for the SPECIES beamline are also discussed. This idea was first developed for SPECIES beamline at MAX II and the concept has now been used in the design of HIPPIE beamline to MAX IV [Sankari, 2012], having similar requirements for the spot size.

In the following, some additional results for HIPPIE beamline at MAX IV obtained from ray-tracing simulations are presented, supporting and further illustrating the concept of a cPGM with astigmatic focus.

Ray-Tracing Simulation

Ray-tracing simulation is a powerful computational method for optical design [Cerrina and del Rio, 2010; Peatman, 1997]. It consists of using Monte Carlo technique[Bevington and Robinson, 2003] to define a number of rays with specific spatial and angular distribution (the source), and then calculating the trajectory of each ray until it reaches an optical element (mirror, lens, gratings, absorber, *etc.*). The interaction with such an element will define a new spatial and angular distribution that is then propagated further to the next element. By placing the source and the optical elements properly in

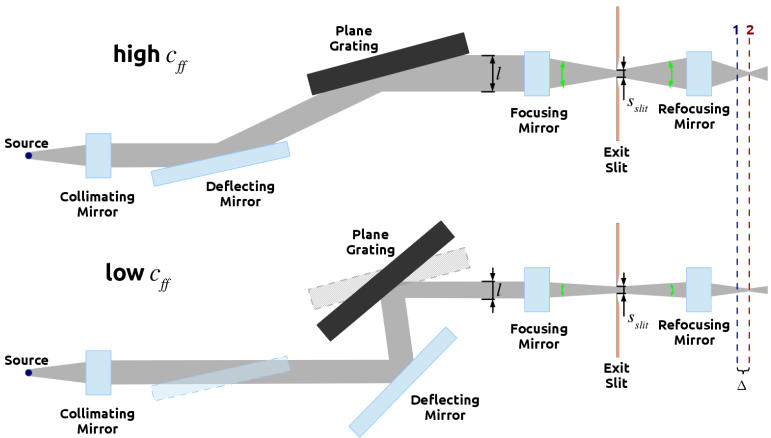


Figure 2.22. *Geometry of a cPGM beamline with astigmatic focus, as used at beamlines SPECIES and HIPPIE. The position 1 (blue) is the position of horizontal focus and of the sample. Position 2 (red) is the vertical focus position, displaced by* Δ *from the horizontal focus. The top and bottom figures illustrate the monochromator geometry for high and low* c_{ff} *values, respectively. We observe for instance that at an higher* c_{ff} *value a bigger length* l *is projected on the focusing mirror, which results in a bigger vertical divergence on the following optics and consequently in a bigger vertical size on the astigmatic focus.*

a reference frame, the trajectories of the rays are calculated through this optical system from one element to the next.

One of the key points of ray-tracing simulation is that it needs very few assumptions about the physical processes. For a typical beamline composed by mirrors and gratings, the only needed assumptions are: (a) a ray propagates in a straight line; (b) reflection law; (c) and grating equation. The source and the optical elements descriptions can be made more complex to have more realistic simulations, but there are no assumptions about geometric optics and therefore the simulation is not limited by the paraxial approximation[Born and Wolf, 1999]. Due to these advantages, ray tracing simulations have been extensively used over the last 40 years in designing synchrotron radiation beamlines [del Rio et al., 2011]. Eventually two computing programs have stood out in this area: SHADOW [del Rio et al., 2011] and RAY [Schäfers, 2008]. In this work SHADOW has been used through a computational library for the Python programing language [Canestrari et al., 2011].

HIPPIE beamline

HIPPIE beamline[6] is designed for MAX IV 3GeV ring [Sankari, 2012] for high pressure X-ray photoelectron spectroscopy (HP-XPS) and X-

[6] `http://www.maxlab.lu.se/hippie`

ray absorption spectroscopy (XAS). It will permit experiments from ultra high vacuum ($\approx 10^{-9}$mbar) up to near atmospheric pressure ($\approx$ 10mbar). The optical design allows the beamline to operate in the photon energy range from 263eV to 2000 eV with variable polarization provided by an APPLE-II undulator[Sasaki, 1994]. The layout of the beamline is the same as in Figure 2.22.

Some features of MAX IV 3GeV ring make possible to obtain very high performance when compared to MAX II: smaller vertical source size, higher ring current and possibility to build longer beamlines. Longer entrance and exit arms (r and r'' on Fig. 2.19) in turn increase the resolving power, as shown in equations 2.44.a and b. Higher electron beam current of MAX IV combined with the small source size makes possible to operate at small exit slit apertures, obtaining a high resolving power without compromising flux so much.

To be able to operate over the full energy range of the beamline, the astigmatic focus must be set at the minimum photon energy (maximum divergence due to the diffraction limit) and minimum c_{ff} value, since from this condition it will be easy to compensate the decrease of the divergence at higher energies by increasing c_{ff}. The change on the energy resolution due to the change of c_{ff} can be compensated by adjusting the exit slit aperture, since it has no effect on the spot size of the astigmatic spot [Paper **I**]. Therefore, the design condition for the astigmatic focus was to obtain a vertical spot size of 50μm (FWHM) at photon energy of 263eV with c_{ff}=2.25. The value of 50μm is determined by the spectrometer and this value maximizes the efficiency of detection [Paper **I**].

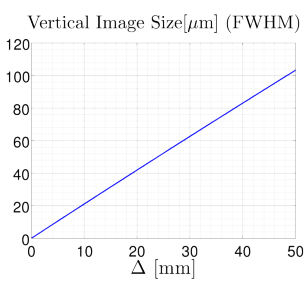


Figure 2.23. *Calculated values using the geometric model for the vertical image size in function of the distance Δ for HIPPIE beamline.*

Considering these requirements, the first step related to the astigmatic refocusing is to define the distance Δ between the vertical and the horizontal focus that results in a vertical spot size of 50μm at photon energy of 263eV and c_{ff}=2.25. Based on the geometric model presented in the Paper **I** we obtain that this condition is achieved for $\Delta = 25mm$ (see also Figure 2.23).

With all the optical distances defined, we can now use ray tracing simulation to compare with the results of the analytical model and to evaluate the beamline performance. The analysis is done in two steps: the cPGM monochromator performance is evaluated by the resulting resolving power; and the astigmatic refocusing is evaluated in relation to the dependence of the vertical spot size on the values of Δ, c_{ff} and the slit opening.

Resolving power HIPPIE beamline is expected to reach a very high resolving power, *e.g.* 40000 at 400eV and 20000 at 1000eV. This requires very small exit slit apertures ($\approx 5\mu m$). The simulations results described here are performed with a slit size of 50μm for practical reasons, namely to increase the number of simulated rays reaching the spot and to compare the results with beamlines I511 and SPECIES. This results in a relatively high resolution for HIPPIE beamline when compared with SPECIES, even though this is not present-

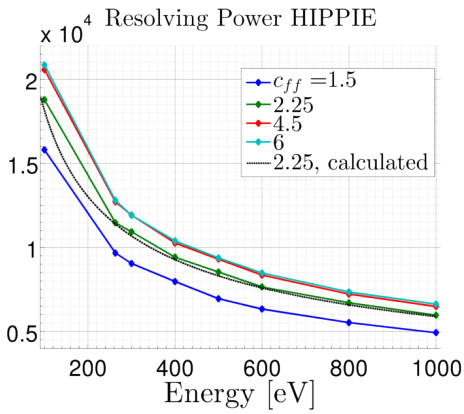


Figure 2.25. *Resolving power for HIPPIE beamline at several c_{ff} values and photon energy with exit slit aperture of 50μm. Each point represents the result from ray-tracing simulation of one run with 1 million rays. The dashed black line represents the calculated value for c_{ff}=2.25 based on the model of Section 2.3, showing a very good agreement between model and simulation.*

ing the top performance of this beamline.

The energy distribution with a slit opening of 50μm at the design condition is shown in Figure 2.24. It is evident that it does not follow a Gaussian distribution, which indicates that the dominant term in resolving power is the term due to the slit opening (eq. 2.44.b). Operating at smaller apertures will eventually result in a Gaussian profile, but since the beamline has high performance with this aperture (resolving power of 12000) this condition is likely to be used and thus this behavior needs to be considered on the data analysis of the experiments (as done for instance by Jiang et al. [2004]).

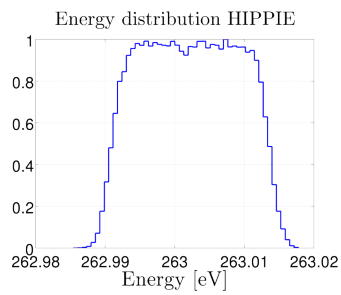


Figure 2.24. *Photon energy distribution at 263eV, c_{ff}= 2.25 and exit slit opening of 50μm. ΔE =22meV (FWHM). Simulation with 250.000 rays*

Finally ray tracing simulation is used to evaluate the beamline performance regarding energy resolution over the photon energy range for different values of c_{ff}. Figure 2.25 shows the results from the simulations, where we observe the features predicted by the analytical models discussed in Section 2.3. In addition, the curve based on the analytical models are also plotted for c_{ff}=2.25 and shows a very good agreement with the ray tracing results.

Astigmatic re-focusing In order to check the geometric result, several ray tracing simulations were performed with different values of Δ. In addition, different photon energies were simulated to examine the energy dependence and the validity of the model.

The results are shown in Figure 2.26 where we can see that the value of Δ=25mm obtained by using the geometric model fulfills the design criteria (*i.e.*, vertical size of 50μm at 263eV). Additionally, we observe that for all astigmatic conditions ($\Delta \neq 0$) there is a dependence of roughly $\propto 1/\sqrt{E_{ph}}$ with photon energy. This is a result of

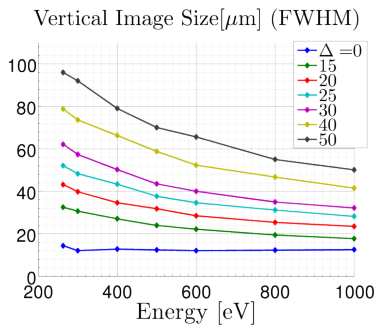


Figure 2.26. *Ray-tracing simulation results for HIPPIE beamline. (a) Vertical image size at several conditions of astigmatic focus in relation to the photon energy for c_{ff}=2.25.*

the linear dependence of the image size with the source divergence, which in turn has direct dependence with photon energy when the source is limited by diffraction.

Considering the layout with Δ=25mm, we evaluate now the possibility of changing the c_{ff} value in order to obtain a constant vertical spot size over the energy range. First, the geometric model is used to calculate the c_{ff} values that results in a specific vertical beam size. Experimental procedure is indeed more suitable to determine the precise value of c_{ff}, but the calculated values are suitable for this analyzes too. The results of the calculation are shown in Figure 2.27.a. Ray tracing simulations were then performed using the calculated c_{ff} values for a vertical size of 50μm (the green curve on fig. 2.27.a). The results of the ray tracing simulations are presented in Figure 2.27.b. It can be seen that the vertical spot size have values of 50μm ±10%. This variation is however due to statistical fluctuations, a problem intrinsic to any Monte Carlo simulation. We can nevertheless conclude the feasibility of obtaining constant vertical beam size over the energy range by adjusting the value of c_{ff} of the grating.

Finally, the independence of the image size from the slit size is checked. The exit slit aperture is an important point in the approximations used for modeling this geometry. The beam at the focus, which is an image of the exit slit, is assumed to be a point (*i.e.*, size equal to zero) which means that the exit slit aperture is also assumed to be zero. This approximation is based on the optical distances: the entrance and exit arms are of the order of meters; the value of Δ is of the order of tens of millimeter; and the image size at the vertical focus is of the order of $\approx 100\mu m$, that is, much smaller than the other values and thus it validates the approximation. On the other hand, it is expected that at some (large) exit slit opening this approximation will not be valid anymore. Moreover, the approximation is more sensitive to the value of Δ, in such way that bigger the value of Δ the

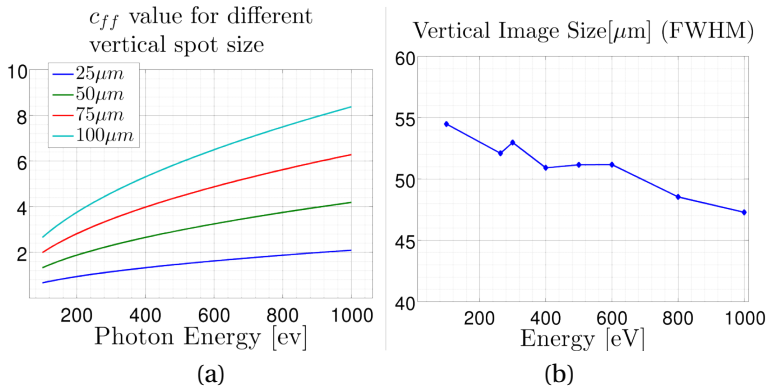


Figure 2.27. *(a) Calculated values of c_{ff} that produce a vertical astigmatic spot (Δ=25mm) of 25μm, 50μm, 75μm and 100μm. (b) Vertical spot size values from ray tracing simulation using the previous values of c_{ff} for 50μm.*

better the approximation.

To conclude this analysis, figures 2.28 (a) and (b) present the simulated values of vertical beam size as a function of slit apertures. By taking the aperture of 50μm as reference, we observe an increase of $\approx$ 20% when the slit aperture is increase five times from 50μm to 250μm. SPECIES beam line in turn presented the same relative increase in image size when the slit opening increases 10 times. This shows that HIPPIE beamline is more sensitive to the approximation and consequently to the slit aperture. This is due to the smaller value of Δ used by HIPPIE beamline: Δ = 25mm, whilst SPECIES has a value of Δ = 62mm. The size of this increase is, however, relatively small and we can approximate that the image size does not depend on the slit aperture. Moreover for HIPPIE beamline an opening of 50μm is relatively big and is unlikely to operate the beamline at bigger exit slit apertures, indicating that the insensitivity for the slit aperture is even better.

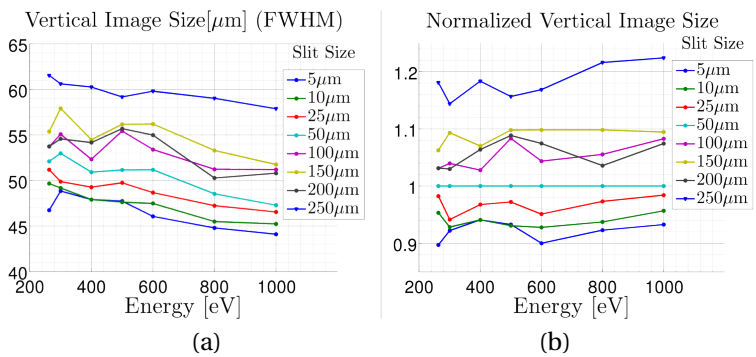


Figure 2.28. *(a)Vertical image size for different values of the exit slit aperture. c_{ff} values values from Figure 2.27.a for beam size of 50μm. (b) Same graph with the values normalized by the results of 50μm slit opening. We observe that for an exit slit apertures smaller than 200μm the resulting image size has a variation of less than 10% compared to the results at 50μm slit opening.*

CHAPTER 3

POLARIMETRY

Polarimetry is a technique for measuring and determining the state of polarization of the radiation[Goldstein, 2011]. Such information has many applications for synchrotron radiation science, namely to provide the actual state of polarization for the light meeting the sample in an experiment, to characterize the effects of the beamline optics on the polarization, and finally to be able to tune the undulator in order to obtain the desired polarization state at the sample, overcoming the changes in polarization caused by the beamline optics.

Due to the energy dependence of the optical properties of materials, different techniques are required to measure the polarization of the light. At the visible and ultraviolet regime the classical method uses a phase retarder[1] based on an anisotropic crystal (*e.g.* mica), and a polarizer[2] based on dichroic materials (*e.g.* Polaroid H-sheet), both elements in a transmission geometry [Saleh and Teich, 2007]. However, for higher energies like VUV radiation there are no transparent materials and practically all the radiation is absorbed in transmission geometry. Therefore mirrors with metal coating are used. On the soft X-rays regime (≈100-2400eV) in turn, mirrors have low reflectivity and designs using multilayers in reflection or transmission mode have been used [Schäfers et al., 1999].

This chapter discusses a reflection based VUV polarimeter manufactured at MAX IV Laboratory that utilizes gold coated mirrors as retarders and attenuators. It starts with a description of different polarizations of light emitted by undulators in storage rings. Then, concepts of electromagnetic optics are used to review the optical properties of conducting materials and to introduce the formalism of the Stokes parameters and the Mueller Matrices. All this is com-

[1] A retarder is an optical element that add a phase shift between the two orthogonal components of the incident light.

[2] Also named attenuator and more precisely diattenuattor. It is an optical element that attenuate the two orthogonal components of the incident light by different amounts.

bined to understand how the Stokes parameters and Mueller matrices are used in order to do polarimetry with the VUV polarimeter. In addition, Paper **III** presents a new method for polarimetry, which can be very advantageous for the VUV polarimeter due its particular mechanical design. Finally, the commissioning of a soft X-rays polarimeter is described in Paper **IV**. This polarimeter is based on multilayers but it shares many basic concepts of the VUV polarimeter; this is discussed in the end of the chapter.

3.1 Polarized light at synchrotron radiation sources

The ultimate goal of the polarimeter is to measure and define the polarization of light at the end station. To achieve certain polarization it will be necessary to tune the undulator to compensate for the changes of the polarization due to the beamline optics. For this reason, it is also essential to know how to change the undulator parameters in order to obtain the desired polarization state.

The polarization of the radiation emitted by a charged particle is defined by the particle's trajectory when it is accelerated [Jackson, 1997]. In the case of storage rings, the radiation emitted by the charged particles has a well defined polarization profile. In bending magnet based synchrotron radiation sources (second generation sources), this polarization profile was used to obtain specific polarization by selecting part of the beam with the desired polarization. This method, however, has no control of the polarization and actually it is possible to obtain only elliptically polarized light (but not circularly polarized) [Clark, 2004].

For this reason, several types of exotic insertion devices have been developed with the aim to force the electrons into specific trajectories to generate specific states of polarization [Clark, 2004; Onuki and Elleaume, 2003; Walker, 1998]. Among them, a variation of the planar undulators, the APPLE II undulator, has stood out and nowadays it is the standard device for generating polarized light in storage rings. For this reason, we start with a general discussion about undulators and then we concentrate in the APPLE II characteristics.

Accelerated Charged Particles and Generation of Polarized Light

Considering the light propagating in the vacuum in the $\hat{z}$ direction, the electric field $\vec{E}(z,t)$ is described as a plane wave of the form

$$\begin{aligned} \vec{E}(z,t) &= E_x(z,t)\hat{x} + E_y(z,t)\hat{y}, \\ E_x(z,t) &= E_{xo}\exp(kz-\omega t+\delta_x), \\ E_y(z,t) &= E_{yo}\exp\big(kz-\omega t+\delta_y\big), \end{aligned} \tag{3.1}$$

where ω and k are the frequency and wavenumber of the radiation, and $c = k/\omega$, the speed of light in vacuum. The state of polarization of the light is then determined by the amplitudes E_{xo} and E_{yo},

and by the relative phase $\delta = \delta_y - \delta_x$ of these two orthogonal components of the electric field [Born and Wolf, 1999]. Considering now the synchrotron radiation emitted by accelerated electrons, in a first order approximation the electric field of the emitted radiation has a dependence with the electrons acceleration $\vec{\dot{\beta}} = (\dot{\beta}_x, \dot{\beta}_y, \dot{\beta}_z)$ given by [Walker, 1998]

$$E_x \propto \dot{\beta}_x - \theta_x \dot{\beta}_z, \qquad E_y \propto \dot{\beta}_y - \theta_y \dot{\beta}_z, \tag{3.2}$$

where θ_{xy} are the angles relative to the $\hat{z}$ axis. Applying these equations to an electron traveling in a storage ring, these equations explain, for instance, the fact that there is no linearly vertical polarized light emitted on axis ($\theta_y = 0$) from a bending magnetic, since in that device we have $\dot{\beta}_y = 0$.

Equations (3.2) can be used in order to obtain the polarization of the light emitted in an undulator. To this end, the electron trajectory on the undulator must be derived from the magnetic field therein, as done below using a sinusoidal approximation for the magnetic field.

Planar Undulators - Sinusoidal Magnetic Field

Using sinusoidal approximation, the on axis (*i.e.*, for x and $y \approx 0$) magnetic field of a undulator $\vec{B}(z)$ is given by

$$\vec{B}(z) = \big(B_x(z), B_y(z), 0\big), \tag{3.3a}$$

$$B_x(z) = B_{xo} \sin(k_u z + \phi_x), \tag{3.3b}$$

$$B_y(z) = B_{yo} \sin(k_u z + \phi_y), \tag{3.3c}$$

where $k_u = 2\pi/\lambda_u$ and λ_u is the undulator period. Note that unlike earlier, the undulator field is not limited to one plane, which was the case in equation (2.21). Using this sinusoidal approximation, the trajectory of the electrons $\vec{r(t)} = (x(t), y(t), z(t))$ can be shown to be [Walker, 1998]

$$x(t) = \frac{K_y c}{\gamma\Omega} \sin(\Omega t + \varphi_x), \qquad y(t) = \frac{K_x c}{\gamma\Omega} \sin(\Omega t + \varphi_y), \tag{3.4}$$

where $\Omega = \bar{\beta}_z 2\pi c/\lambda_u$ and $\bar{\beta}_z$ is the mean speed in the $\hat{z}$ direction. The parameters $\varphi_{x,y}$ are temporal phases related to the magnetic fields. For the sinusoidal magnetic field the K parameters are given by

$$K_x = \frac{e B_{xo} \lambda_u}{2\pi m c} = 93.4 \cdot B_{xo}[T]\lambda_u[m] \qquad \text{and} \qquad K_y = \frac{e B_{yo} \lambda_u}{2\pi m c}. \tag{3.5}$$

The acceleration in the $\hat{z}$ direction $\dot{\beta}_z$ is much smaller than the acceleration $\dot{\beta}_x$ and will not be considered here [Walker, 1998]. This

allows assuming the electron position at the $\hat{z}$ axis to be $z(t) = \bar{\beta}_z\, t$ and this makes possible to convert the spatial dependence of the magnetic field to temporal dependence. With this assumption we obtain that the electron acceleration and the magnetic field are proportional as

$$\dot{\beta}_x(t) \propto B_y(t) \qquad \text{and} \qquad \dot{\beta}_y(t) \propto B_x(t). \tag{3.6}$$

Therefore, the electric field of the radiation, as given by equation (3.2), is directly proportional to the magnetic field of the undulator

$$E_x \propto \dot{\beta}_x \propto B_y \qquad \text{and} \qquad E_y \propto \dot{\beta}_y \propto B_x. \tag{3.7}$$

From this we can conclude that the amplitude of the electric field of the emitted radiation can be modulated by the amplitude of the magnetic field of the undulator. Furthermore, it can be derived that the difference of phase between $B_y(z)$ and $B_x(z)$ defines the phase between the electric fields $E_x\,\hat{x}$ and $E_y\,\hat{y}$ [Walker, 1998].

For a planar undulator, using the reference frame of Figure 2.1, there is no acceleration in $\hat{y}$ direction and thus the radiation is linearly polarized in the plane of oscillation, in this case the horizontal plane. Although the magnitude of $\dot{\beta}_z$ is small, vertical components of the electric field are still produced out of the axis, which causes an inclination in the plane of oscillation of the resulting electric field. In other words, off-axis the plane of polarization is tilted and has a dependence with the angle of observation[3]. However, this does not generate circularly or elliptically polarized light since there is no phase difference between the electric field components.

Some experiments, on the other hand, require circularly and elliptically polarized light and also the possibility to control the state of polarization. For this, many variations of the undulators have been developed with the purpose of controlling both the amplitudes and the phase between the magnetic fields $B_x(z)$ and $B_y(z)$ [Carr, 1993; Hwang and Yeh, 1999]. Among these, the device known as APPLE-II [Sasaki, 1994] has become popular due to the high peak fields for the circular mode and due to the possibility to generate all states of polarization in a relative simple way when compared to other devices, as is discussed next.

APPLE-II Undulators

APPLE-II type undulators are based in four rows of magnets arranged in two rows above and two rows below the electron orbit (Figure 3.1). The magnets in the rows are assembled in a Halbach configuration

[3] For a detailed description of the observation angle dependence of polarization in a planar undulator see Fig. 4.14 in [Clark, 2004] and [Kitamura, 1980].

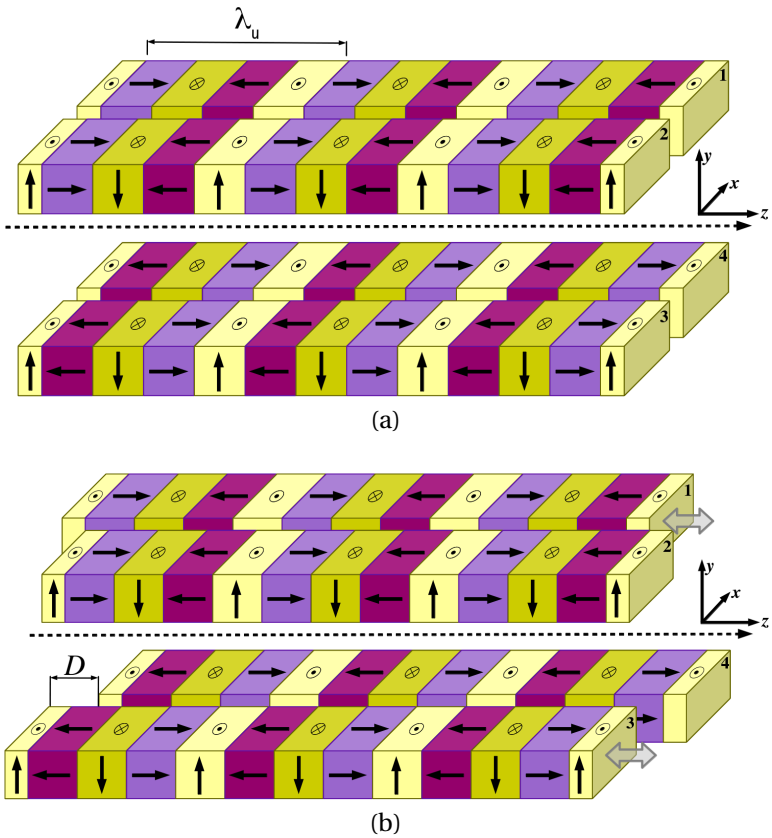


Figure 3.1. *APPLE-II Undulator type. (a) No phases, the Undulator emits horizontally linear polarized light. (b) Two row displaced by D on the $-\hat{z}$ direction (symmetric mode).*

(Figure 3.2) [Halbach, 1981], and actually can be considered as an ordinary planar Halbach undulator where the upper and lower rows are split into two.

Using the sinusoidal description of the magnetic field, the field in the $\hat{x}$ and $\hat{y}$ directions is described as [Hwang and Yeh, 1999; Schmidt and Zimoch, 2007]

$$B_y(z) = B_{yo}(g)\Big[\sin(k_u z+\phi_1)+\sin(k_u z+\phi_2)+ \sin(k_u z+\phi_3)+\sin(k_u z+\phi_4)\Big], \quad (3.8a)$$

$$B_x(z) = B_{xo}(g)\Big[-\sin(k_u z+\phi_1)+\sin(k_u z+\phi_2)+ -\sin(k_u z+\phi_3)+\sin(k_u z+\phi_4)\Big]. \quad (3.8b)$$

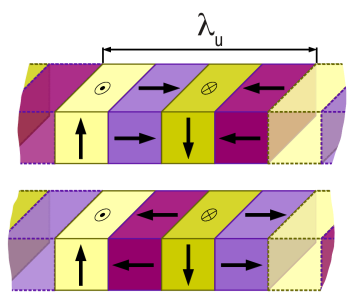


Figure 3.2. *Halbach undulator, where the magnetic period is based on 4 poles with the magnetic field pointing in different direction. Note that the up and bottom rows are assembled in different manners.*

The amplitudes $B_{xo}(g)$ and $B_{yo}(g)$ depend on the gap opening g (*i.e.* the distance between the upper and lower plane of magnets) and will be discussed in more detail later. The undulator phases $\phi_i = 2\pi D_i/\lambda_u$ are related to the displacement D_i in $-\hat{z}$ direction of each row. The number of the quadrant is used as the index i: index 1 for the upper back row, 2 for upper front, 3 for lower front and 4 for lower back (see Figure 3.1).

It can be seen from equations (3.8) that it is indeed possible to modulate the strength and phase between the two components of the magnetic field by moving the individual rows. Regarding equations (3.7), it means that it is possible to change the energy and polarization of the emitted photons by moving the rows individually.

Therefore, to generate linearly or circularly polarized light the rows must be placed properly in order to obtain the correct phase and amplitude ratios ρ. There are two simplified ways to operate an APPLE-II undulator in order to generate these different polarization states: the circular and linear polarization modes, as discussed next.

Circular Polarization Mode (Symmetrical movement)

Circularly polarized light (CPL) can be obtained in an APPLE-II undulator by moving a pair of opposite rows (1 and 3 or 2 and 4, Figure 3.1.b) in the same direction with the same displacement(symmetrical movement), while keeping the other pair fixed (the CP mode). Arbitrarily choosing the rows 1 and 3 to move while keeping rows 2 and 4 fixed, we have

$$\phi_1 = \phi_3 = \phi \quad \text{and} \quad \phi_2 = \phi_4 = 0 \qquad \text{(CP mode)} \tag{3.9}$$

and equations (3.8) reduce to

$$B_y(z) = 4B_{yo}(g)\cos\left(\frac{\phi}{2}\right)\sin\left(k_u z + \frac{\phi}{2}\right), \qquad \text{(CP mode)} \tag{3.10a}$$

$$B_x(z) = 4B_{xo}(g)\sin\left(\frac{\phi}{2}\right)\cos\left(k_u z + \frac{\phi}{2}\right). \tag{3.10b}$$

These equations can be rewritten as

$$B_y(z) = \tilde{B}_y \sin\left(k_u z + \frac{\phi}{2}\right), \qquad \text{(CP mode)} \tag{3.11a}$$

$$B_x(z) = \tilde{B}_x \sin\left(k_u z + \frac{\phi}{2} + \frac{\pi}{2}\right), \tag{3.11b}$$

where $\tilde{B}_{x,y}$ are the peak of the magnetic field when operating in the CP mode.

From these equations[4] we conclude that the phase between the fields $B_y(z)$ and $B_x(z)$ in the CP mode is $\pi/2$. By recalling the linear dependence of the emitted electric fields on the magnetic fields (equations 3.7), we have the result that the phase δ between the electric field components of the light is also $\pi/2$.

The remaining parameter to be determined for the polarization of the light is then the amplitude ratio $\rho = E_y/E_x$. Recalling again equations 3.7, we can relate the ratio ρ to the peak values of the magnetic field $\tilde{B}_{x,y}$ by

$$\rho = \frac{E_y}{E_x} = \frac{\tilde{B}_x}{\tilde{B}_y} = \frac{B_{xo}(g)}{B_{yo}(g)} \tan\left(\frac{\phi}{2}\right). \qquad \text{(CP mode)} \tag{3.12}$$

Hence the ratio ρ can have values from $-\infty$ to ∞, and the different values of ρ are related to different polarization states as follows

ρ	rows phase ϕ	Polarization
0	0	horizontal linear
$\pm\infty$	$\pm\pi$	vertical linear
$0 < \rho < \infty$	$-\pi < \phi < 0$	left-handed elliptical
$0 > \rho > \infty$	$0 < \phi < \pi$	right-handed elliptical

By solving equation (3.12) for ϕ we obtain the rows phase necessary to produce a specific state of polarization with a certain value of ρ

$$\phi = 2 \cdot \arctan\left(\rho \frac{B_{yo}(g)}{B_{xo}(g)}\right). \qquad \text{(CP mode)} \tag{3.13}$$

For circular polarized light it is necessary that $\rho = \pm 1$ (left or right hand CPL) and then

$$\phi_{CPL} = \pm 2 \cdot \arctan\left(\frac{B_{yo}(g)}{B_{xo}(g)}\right). \qquad \text{(CP mode)} \tag{3.14}$$

Finally, it must be noted that the main axis of the elliptically polarized light is limited to the directions $\hat{x}$ or $\hat{y}$, and then it is not possible to produce inclined (tilted) elliptically polarized states with symmetric displacement.

Defining the row phase and gap values By equation (3.12), we have that the polarization state is determined by both the row phase ϕ and the undulator gap g. However, the resonant energy of the undulator

[4] There are two equivalent ways to read these equations: first, the peaks of the magnetic field $\tilde{B}_x$ and $\tilde{B}_y$ are positives and the phase can assume values $+\pi/2$ or $-\pi/2$; in the second way, the phase is $+\pi/2$ and the peaks $\tilde{B}_x$ and $\tilde{B}_y$ can assume negative values. In this work the later way is used.

is also a function of the amplitudes of the magnetic field components $\tilde{B}_{x,y}$ (equations 2.3 and 3.5), and hence it is also a function of the gap and the rows phase[5] ϕ.

To make this relation clear, we start by writing the deflection parameter K in equation (3.5) as a function of the gap opening

$$K_o(g) = \frac{e\lambda_u}{2\pi mc} \cdot B_o(g).$$

With this definition we use the peak of the magnetic field when operating on the CPL mode in equation (3.10) to obtain the photon energy as

$$E_1[eV] = \frac{2 \cdot 1.24 \cdot 10^{-6} \cdot \gamma^2}{\lambda_u \left[1 + 8K_{yo}(g)^2 \cos^2\left(\frac{\phi}{2}\right) + 8K_{xo}(g)^2 \sin^2\left(\frac{\phi}{2}\right)\right]}. \quad \text{(CP mode)} \tag{3.15}$$

We have then that two quantities, photon energy E and the amplitude ratio ρ, depend on the same two parameters, gap opening g and rows phase ϕ. Mathematically we have $E(g,\phi)$ and $\rho(g,\phi)$. It means that to tune the undulator to a specific polarization state and photon energy is necessary to find the proper values of gap g and rows phase ϕ.

The proper way to determine the gap dependence of the magnetic field functions $B_{xo}(g)$ and $B_{yo}(g)$ is by experiment. However, Elleaume et al. [2000] presented an analytical function that can be used to estimate the peak magnetic field for several models of undulators as:

$$B_{peak} = a \exp\left[b\frac{g}{\lambda_u} + c\left(\frac{g}{\lambda_u}\right)^2\right]. \tag{3.16}$$

The parameters a, b and c depend on the direction of the magnetic field in the undulator. The results obtained using this model for an APPLE-II undulator are plotted in Figure 3.3. Figure 3.4 shows the photon energy as given by equation (3.15), using the model for the magnetic field $B_{xo}(g)$ and $B_{yo}(g)$ as given by equation (3.16).

In practice the interest lays in finding out the correct gap opening and rows phase for producing certain polarization state at certain energy, *i.e.* $\phi(E_{h\nu},\rho)$ and $g(E_{h\nu},\rho)$. Solving these parameters analytically is complicated but numerical methods can be used for that.

The graphs of Figure 3.5 show the values of gap and phase required to obtain different polarization states. The graphs are produced with numerical calculation by using the undulator parameters, the magnetic field of equation (3.16), the photon energy and

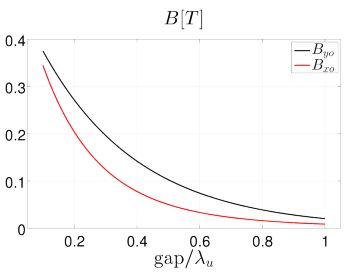


Figure 3.3. *Dependence of the magnetic fields $B_{xo}(g)$ and $B_{yo}(g)$ with the gap for an APPLE-II undulator.*

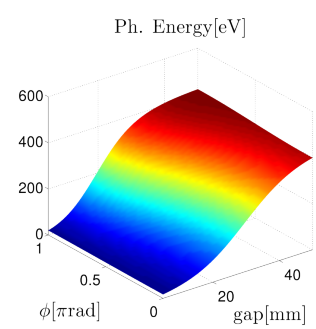


Figure 3.4. *(CP Mode) Photon Energy in function of the gap aperture and the phase ϕ. ($\lambda_u = 5cm$ and electron energy $1.5GeV$)*

[5] In fact it is possible to tune the resonant photon energy in a APPLE-II undulator by using the row movement instead of changing the gap opening, making use of the Adjustable Phase Undulator concept. The concept is described by Lidia and Carr [1994] and an application is reported for instance by Schmidt et al. [2007].

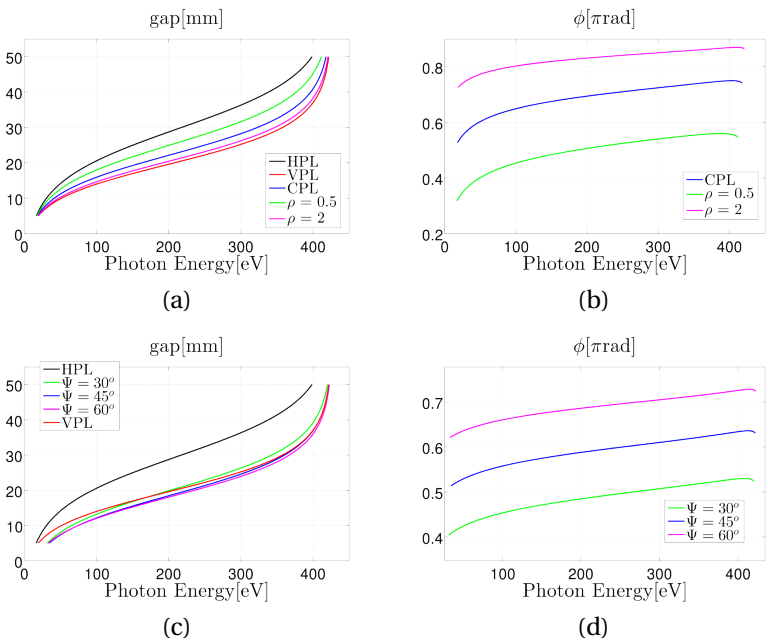


Figure 3.5. *Gap and row phase ϕ in function of the photon energy and different state of polarization: elliptically (defined by different values of ρ), horizontally (HPL), vertically (VPL) and circularly (CPL) polarized light. (a) and (b) CP mode. (c) and (d) LP mode.*

rows phase equations (equations 3.12 and 3.15). For instance, as an illustrative example, we can see that to generate circular polarized light at $200eV$ it is necessary to have a gap of 20mm and a phase of 0.67πrad. This shows the feasibility of using numerical calculation to obtain the values of gap and row phase for a specific set of photon energy and polarization.

Linear Polarization Mode(Antisymmetric movement)

In the CP mode discussed above it is possible to generate linearly polarized light only in the vertical or horizontal directions. As the row phase changes in the interval $0 < \phi < \pi$, the polarization of the light changes from horizontal linearly polarized to elliptically polarized (passing by circularly polarized) and finally ending up to vertical linearly polarized light.

However, by moving the rows in opposite directions (antisymmetric movement) it is possible to obtain inclined linearly polarized light. In this case, as will be discussed, the inclination of the linearly polarized light can also be adjusted.

Similarly to the CP mode, when operating in the linear polarization mode (LP mode) is necessary to move two opposite rows by the

same displacement, but now in opposite directions (*i.e.*, antisymmetric movement). Having again rows 1 and 3 to be moved, we have that

$$\phi_1 = -\phi_3 = \phi \quad \text{and} \quad \phi_2 = \phi_4 = 0. \qquad \text{(LP mode)} \tag{3.17}$$

These conditions in equations 3.8 yield

$$B_y(z) = 4B_{yo}(g)\cos^2\left(\frac{\phi}{2}\right)\sin(k_u z), \qquad \text{(LP mode)} \tag{3.18a}$$

$$B_x(z) = 4B_{xo}(g)\sin^2\left(\frac{\phi}{2}\right)\sin(k_u z), \tag{3.18b}$$

which becomes

$$B_y(z) = \tilde{B}_y \sin(k_u z), \qquad \text{(LP mode)} \tag{3.19a}$$

$$B_x(z) = \tilde{B}_x \sin(k_u z), \tag{3.19b}$$

and now the components of the magnetic field are in phase, which means that only linearly polarized light is produced. Again, the peak values of magnetic field components $\tilde{B}_x$ and $\tilde{B}_y$ depend on the phase ϕ, though the dependence is different compared with the CP mode.

Since the fields are in phase in the LP mode, the ratio $\rho = E_{yo}/E_{xo}$ of the amplitudes of the electric field defines the angle of inclination Ψ of the polarization plane

$$\rho = \frac{E_y}{E_x} = \tan\Psi.$$

Using again equation (3.2) we have

$$\rho = \frac{\tilde{B}_x}{\tilde{B}_y} = \frac{B_{xo}(g)}{B_{yo}(g)}\tan^2\left(\frac{\phi}{2}\right). \qquad \text{(LP mode)} \tag{3.20}$$

Solving this equation for ϕ, we obtain the row phase necessary to produce linearly polarized light with an inclination angle Ψ:

$$\phi_{LPL} = 2\cdot\arctan\left(\sqrt{\tan\Psi\cdot\frac{B_{yo}(g)}{B_{xo}(g)}}\right). \qquad \text{(LP mode)} \tag{3.21}$$

Similarly to the CP mode, the change of phase changes the amplitude of the magnetic field, which in turns changes the resulting photon energy. Carrying on the calculation for the on-axis photon energy for the LP mode gives

$$E_1[eV] = \frac{2 \cdot 1.24 \cdot 10^{-6} \cdot \gamma^2}{\lambda_u \left[1 + 8K_{yo}(g)^2 \cos^4\left(\frac{\phi}{2}\right) + 8K_{xo}(g)^2 \sin^4\left(\frac{\phi}{2}\right)\right]}. \quad \text{(LP mode)} \tag{3.22}$$

It can be seen that although the results are different from the CP mode (equation (3.15)), we face the same problems in analytically defining the proper gap opening and row phase. Again a numerical method is used to define these parameters. The results are shown in Figure 3.5. For the parameters considered here, we have for instance that to generate LPL of 200eV inclined at 45deg it is necessary to have a gap of 20mm and an asymmetric row phase of 0.58πrad.

The limitation in this case is that $\tan\Psi$ in equation (3.21) must be positive, limiting the angle of inclination to the interval $0 \leqslant \Psi \leqslant 90$ deg. That means that the plane of oscillation is restricted to the first and third quadrant. To obtain linearly polarized light inclined on the other quadrants it is necessary to move rows 2 and 4, instead of the 1 and 3 as assumed here. Thus, to be able to rotate the oscillation plane of the radiation in all direction, it is necessary that all the rows are movable [Schmidt et al., 2007].

It must be pointed out that some approximations are used in the development of these models. The main approximations are related to the magnetic fields of the undulator: a sinusoidal approximation was used, only the on-axis magnetic fields were considered (x and $y \approx 0$), the magnetic field on $\hat{z}$ direction was assumed to be zero, and it was assumed that the gap dependence of the magnetic field is as given by equation (3.16). In principle, all deviations from these approximations must be considered when applying the models described here to real undulators.

The better way to overcome the problems due to the use of these approximations is by determining experimentally the polarization of the light emitted by the undulator. This, together with the change of polarization of the light due to optical effects (as will be discussed in Section 3.2), motivated us to develop a polarimeter to be operated at the MAX IV laboratory, discussed in the next sections.

3.2 Optical properties of metals

The study of the optical properties of the materials starts by developing a model for the response of the material to an external electric field and then solving the Maxwell's equation for the corresponding electromagnetic wave [Born and Wolf, 1999].

In the case of metals, the common approach is to consider the medium with both dielectric and conductive properties with a linear dependence on the electric field [Saleh and Teich, 2007]. Such media are called linear media, and by assuming additionally that they are homogeneous and isotropic, one derives that the resulting electro-

magnetic wave can be described by a harmonic field with a complex wave number $\tilde{k}$.

Such result is equivalent to assuming a dielectric material (non-conducting) with a complex permittivity. Therefore, some results from dielectric materials can be applied to metals by simply replacing the real value of the wave number and permittivity by corresponding complex values. Regarding the optical phenomenon, this means that some well known descriptions for the dielectrics can also be used for metals just by adding a complex term to the refractive index $\tilde{n}$ as

$$\tilde{n} = (1-\delta) + i\beta, \tag{3.23}$$

where the parameters δ and β are energy dependent. The real part is written as $1-\delta$ for practical reasons since it is very close to unity in the X-ray range. The imaginary term β is the attenuation index and is related to the absorption effect. Figures 3.6 and 3.7 show the values of δ and β for gold obtained from tabulated data [Palik, 1985].

One of the results for dielectrics that can be extended to metals is the Snell's law (law of refraction). Considering an electromagnetic wave propagating from the vacuum into a metal with an incidence angle θ, the Snell's law becomes

$$\sin\theta = \tilde{n}\sin\tilde{\phi}. \qquad \text{Snell's law for metals} \tag{3.24}$$

Since the refractive index $\tilde{n}$ is complex, the quantity $\tilde{\phi}$ needs also to be complex, and cannot be interpreted as the angle of refraction anymore [Born and Wolf, 1999]. It must be recalled that also the wave number $\tilde{k}$ is complex, which in turn results in a complex propagation vector. Considering these complex quantities and carrying on the calculation for the refracted beam actually results in an inhomogeneous wave[6], which requires a special interpretation of the quantities $\tilde{n}$, $\tilde{\phi}$ and $\tilde{k}$ [Born and Wolf, 1999]. However, since we are interested in the reflected radiation only, the properties of the inhomogeneous wave inside the metal can be ignored and all the properties of the reflection can be studied based only on the parameters δ and β.

The next step is to derive the Fresnel's law for metal surfaces. The derivation of the Fresnel's law are omitted here but it is described in details for instance by Fowles [1975] and Born and Wolf [1999]. The Fresnel's law considering a complex refractive index is given by

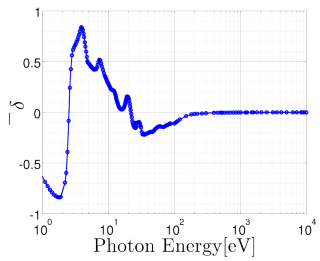


Figure 3.6. *Values of $-\delta$ for gold as a function of photon energy [Palik, 1985].*

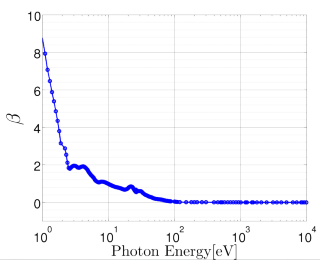


Figure 3.7. *Values of β for gold as a function of photon energy [Palik, 1985].*

[6] By inhomogeneous waves one means waves in which the surfaces of constant amplitudes do not coincide with the surfaces of constant phase. In particular on the metals the surfaces of constant amplitude are normal to the metal-dielectric interface whereas the surfaces of constant phase are not [Fowles, 1975].

$$\tilde{r}_s = \frac{\cos\theta - \sqrt{\tilde{n}^2 - \sin^2\theta}}{\cos\theta + \sqrt{\tilde{n}^2 - \sin^2\theta}},$$
$$\tilde{r}_p = \frac{-\tilde{n}^2\cos\theta + \sqrt{\tilde{n}^2 - \sin^2\theta}}{\tilde{n}^2\cos\theta + \sqrt{\tilde{n}^2 - \sin^2\theta}}. \quad \text{Fresnel's law} \quad (3.25)$$

Here $\tilde{r}_s$ is the complex coefficients of reflection of the incident waves with the electric field parallel to the optical surface (direction x_2 in Figure 3.8), named as transverse electric (TE) waves. The $\tilde{r}_p$ in turn is the complex coefficients of reflection for transverse magnetic (TM) waves, which have the magnetic vector of the field parallel to the surface. These results are the same for dielectric except by the fact that the refractive index is complex. Furthermore, for $\beta = 0$ these equations reduce to the results for dielectrics, as expected.

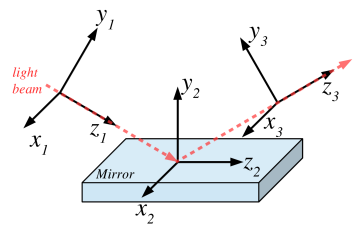


Figure 3.8. *Coordinates system for the reflection on mirror. TE waves have the electric field oscillating on x_1 direction, whilst TM waves the electric field oscillates on z_1 direction.*

It can be seen directly from equation (3.25) that due to the complex value of refractive index also the coefficients of reflection are complex, *i.e.*

$$\tilde{r}_s = r_s e^{i\delta_s} \qquad \text{and} \qquad \tilde{r}_p = r_p e^{i\delta_p}. \quad (3.26)$$

A direct result of these coefficients being complex is that the electromagnetic wave is subject to a change of phase when reflected. Reflection on dielectrics also causes a change in phase but, except in a total internal reflection, the phase shift is always equal to zero or π [Saleh and Teich, 2007]. The change of phase due to reflection on metals has a more complicated behavior and there is no simple analytical function for it. However, by using tabulated values and the Fresnel's equations for metals, a numeric computational software(*e.g.* MATLAB [2013]) can be used to calculate the phase shifts in reflection.

This method is now used to illustrate the energy and angular dependence of the reflectance $R_{s,p} = |\tilde{r}_{s,p}|^2 = r_{s,p}^2$ for gold. The energy dependence is shown in Figure 3.9 for $\theta = 80$deg. Figure 3.10 shows the dependence of angle of incidence for three different energies. We can see that in general R_s increases for larger values of θ (*i.e.* grazing incidence) and that R_p has a minimum value at some angle, called the principal angle of incidence, which corresponds to the Brewster angle of the dielectrics [Fowles, 1975].

From these figures it becomes clear that we can calculate the reflectivity coefficients and the phase shifts due to reflection just by using the Fresnel's law for metals and tabulated data for the optical constants of materials. This information is essential to study the performance of the polarimeter and also to optimize the instrument and its mode of operation. Moreover, analyzing the values of reflectance and phase shift it is clear that a beamline consisting on many reflective optical elements at grazing incidence (*e.g.* $\theta = 88$deg) changes the polarization state of the light. Indeed the same concepts used to

study the polarimeter can be used later on to study the beamlines influence in polarization, as done *e.g.* by Bahrdt et al. [2010].

3.3 Stokes Parameters and Mueller Matrices

There are several ways to represent and study the polarization state of the light, namely the polarization ellipse, Jones matrices, Stokes Parameters and Poincaré sphere [Born and Wolf, 1999; Goldstein, 2011]. For the present work the Stokes parameters and the Mueller matrices are well suitable since they represent the polarization state in terms of measurable quantities and can be applied also for unpolarized and partially polarized light.

For an electromagnetic wave described by a plane wave the Stokes parameters are defined by

$$S_0 = E_{x0}^2 + E_{y0}^2, \tag{3.27a}$$
$$S_1 = E_{x0}^2 - E_{y0}^2, \tag{3.27b}$$
$$S_2 = 2E_{x0}E_{y0}\cos(\delta_y - \delta_x), \tag{3.27c}$$
$$S_3 = 2E_{x0}E_{y0}\sin(\delta_y - \delta_x), \tag{3.27d}$$

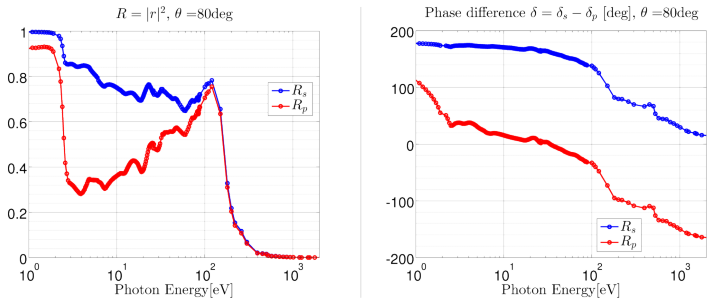
where E_{x0} and E_{y0} are the amplitudes of the two orthogonal components $E_x(t)$ and $E_y(t)$ of the electric field $\vec{E}(t)$ of the light, which in turn are given by equations 3.1. The coordinate system is the same as already shown in Figure 3.8. The derivation is omitted here, but the Stokes parameters are extensively discussed in the literature and a formal derivation can be found for instance in Goldstein [2011, Chapter 5].

Besides the polarization of the light itself, it is necessary to describe how an optical element affects the polarization. For the majority of the optical elements this interaction can be assumed to be linear (linear optics), *i.e.*, the resulting electric field that comes out of an optical system can be described as a linear combination of the electric field that comes into the element. This can be written as

$$E'_x = T_{11}E_x + T_{12}E_y,$$
$$E'_y = T_{21}E_x + T_{22}E_y,$$

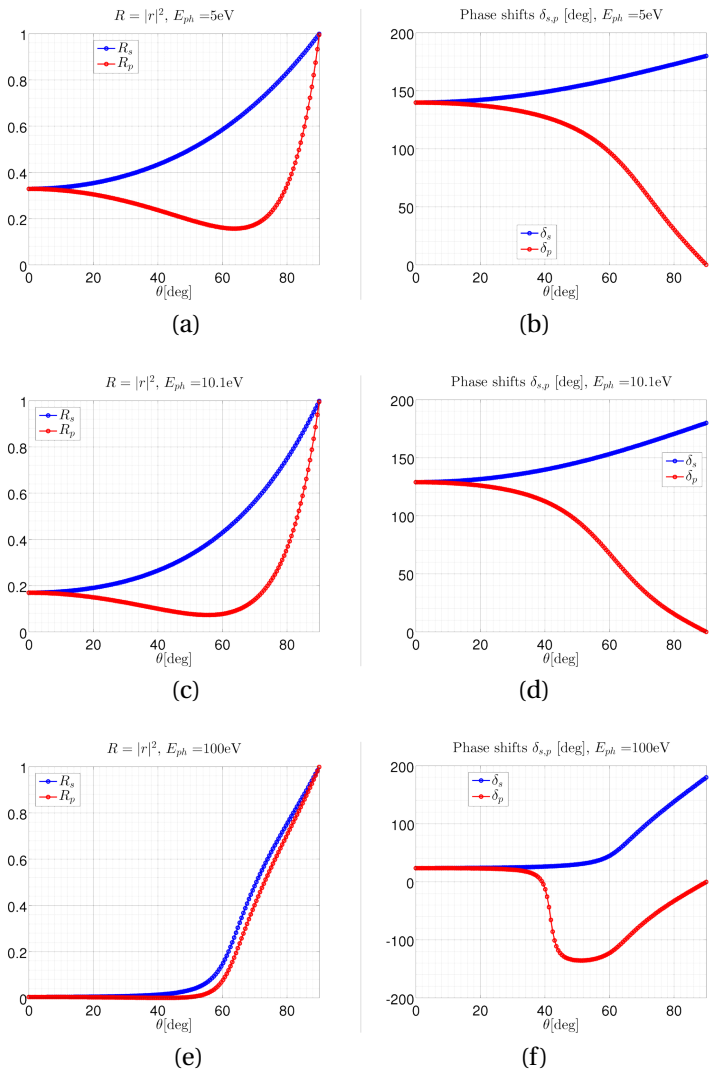
where $E_{x,y}$ and $E'_{x,y}$ are the two components x and y of the electric field of the incoming and outgoing waves respectively. T_{11}, T_{12}, T_{21} and T_{22} are (complex) constants. In the case of a mirror oriented such that the $\hat{x}$ direction is on the mirror surface (as in Figure 3.8), we would have that $T_{12} = T_{21} = 0$, $T_{11} = \tilde{r}_s$ and $T_{22} = \tilde{r}_p$, where $\tilde{r}_{s,p}$ are the complex coefficients of reflection (equations 3.25).

Linear equations as the ones shown above can be presented in simpler way by using matrices, and thus the equations above can be written as



(a) (b)

Figure 3.9. *(a)Reflectance R and (b) phase shifts due to reflection on gold in function of the photon energy. Incidence angle 80deg.*



(a) (b)

(c) (d)

(e) (f)

Figure 3.10. *Reflectance and phase shifts in function of the incidence angle for 5eV (a-b), 10eV (c-d) and 100eV (e-f).*

$$\begin{pmatrix} E'_x \\ E'_z \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \cdot \begin{pmatrix} E_x \\ E_z \end{pmatrix},$$
$$\mathbf{J}_1 = \mathbf{T} \cdot \mathbf{J}_2,$$

where in the context of optics the vectors $\mathbf{J}_1$ and $\mathbf{J}_2$ are called the Jones vectors and $\mathbf{T}$ the Jones matrix. In this case, the Jones matrix $\mathbf{R}$ for reflection is given by

$$\mathbf{R} = \begin{pmatrix} \tilde{r}_s & 0 \\ 0 & \tilde{r}_p \end{pmatrix}.$$

The same reasoning applies to the Stokes parameters: for an optical system that is linear regarding the electric fields $E_{x,y}$, the Stokes parameters are also subjected to linear transformations like [Goldstein, 2011]

$$\begin{pmatrix} S'_0 \\ S'_1 \\ S'_2 \\ S'_3 \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & m_{44} \end{pmatrix} \cdot \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix}$$
$$\mathbf{S}' = \mathbf{M} \cdot \mathbf{S}_0$$

where the vectors $\mathbf{S}'$ and $\mathbf{S}_0$ are called Stokes vectors and are related to the outgoing and incoming light, respectively. $\mathbf{M}$ is called the Mueller matrix and it describes the optical element.

The derivation of the Mueller matrices for the different optical elements is slightly more complicated than that of the Jones matrices and is done elsewhere (*e.g.* in Goldstein [2011]). Additionally, Azzam and Bashara [1987] discuss how to make the conversion from a Jones matrix to a Mueller matrix.

In the present case the main interest is in the Mueller matrix for a metal coated mirror, or in other words, a Mueller matrix for the Fresnel's law for metals, which is given by [Goldstein, 2011]

$$\mathbf{M}(r_s, r_p, \delta) = \frac{1}{2} \times \begin{pmatrix} r_s^2 + r_p^2 & r_s^2 - r_p^2 & 0 & 0 \\ r_s^2 - r_p^2 & r_s^2 + r_p^2 & 0 & 0 \\ 0 & 0 & 2r_s r_p \cos\delta & 2r_s r_p \sin\delta \\ 0 & 0 & -2r_s r_p \sin\delta & 2r_s r_p \cos\delta \end{pmatrix}, \tag{3.28}$$

where $\delta = \delta_s - \delta_p$. By comparing this matrix with the Mueller matrices for a polarizer and a retarder (as described in Chapters 6.2 and 6.3 of Goldstein [2011]), we conclude that a mirror has the same effect than a polarizer-retarder pair.

It is assumed in equation (3.28) that the reflective surface is in the orientation shown in Figure 3.8: the mirror is orientated so that

the x_2 axis of the mirror coincides with the x_1 axis of the incoming wave. To obtain other orientations the reference frame must be rotated. Considering the reference frame presented in Figure 3.11, we have that the reflection happens on the yz plane at each mirror. To obtain the orientation of the second mirror, the reference frame must be rotated around the z_3 axis, and in this particular example the reference frame must be rotated by 180deg. Regarding the Stokes parameters, a rotation is done by applying a Mueller matrix of rotation $\mathbf{M_R}(\eta)$ given by

$$\mathbf{M_R}(\eta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(2\eta) & \sin(2\eta) & 0 \\ 0 & -\sin(2\eta) & \cos(2\eta) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.29}$$

where η is the rotation angle. The particular way that vectors and matrices transform due to rotation must be recalled: a rotated Stokes vector $\mathbf{S}(\eta)$ is given by[7]

$$\mathbf{S}(\eta) = \mathbf{M_R}(\eta) \cdot \mathbf{S}$$

whilst a rotated matrix $\mathbf{M}(\eta)$ is

$$\mathbf{M}(\eta) = \mathbf{M_R}(-\eta) \cdot \mathbf{M} \cdot \mathbf{M_R}(\eta).$$

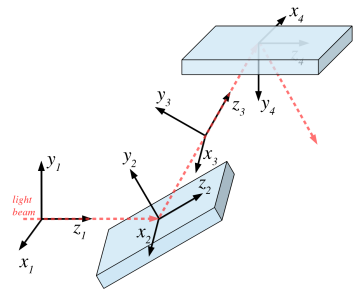


Figure 3.11. *Geometry of two mirrors of the retarder unit. To obtain this geometry the second mirror must be rotated by 180deg around the $\hat{y}$ axis (see text for details).*

3.4 Polarimeter

The Mueller matrices and Stokes parameters are now used to evaluate the performance of the polarimeter and to study how it can be used for polarimetry. The discussion is focused in a polarimeter based on reflectors, like the one presented in Figure 3.12. The optical design of the polarimeter is based on two main units like any other polarimeter: the retarder, consisting on three mirrors, and the analyzer unit, consisting on a single mirror and a light detector. The measurement of the polarization of the incoming light is done by rotating the units in proper manner and measuring the transmitted intensity with the detector.

The function of the first unit is to introduce different phase shifts to the different components of the electric field (s and p). Such function could be performed by a single mirror, but then the beam that comes out of this unit would be deflected in different directions during the rotation η_1. With this three mirrors assembly it is possible to place the mirrors in such a way that the outgoing beam is always at the same axis as the incoming light, and consequently the final design and operation of the polarimeter is made simpler. The analyzer

[7] For the present case geometric arguments can be used to prove these relations, as done *e.g.* by Goldstein [2011, Chapter 6.5]. A formal and general derivation can be found for instance in Arfken et al. [2012, Chapter 5.6 - Transformations of Operators].

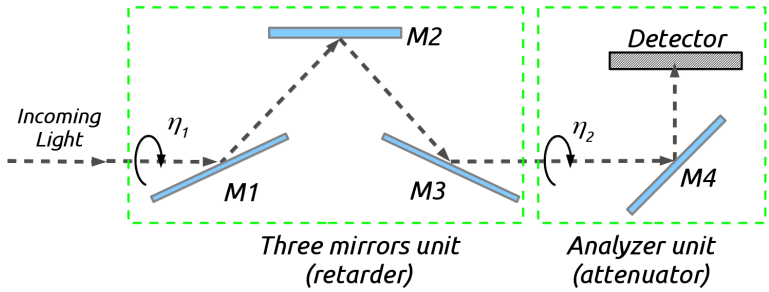


Figure 3.12. *Optical scheme of the polarimeter. The two units can be rotated independently by angles η_1 and η_2 around a main axis determined by the incoming light. In this configuration the angles η_1 and η_2 are defined as zero.*

unit has the function of introducing a (further) attenuation to the s and p components of the light. Finally, the intensity of the light is measured with a photon detector.

It must be noted that the retarder unit also attenuates the s and p components of the light and this actually decreases the final performance of the polarimeter. The ideal situation is not having any attenuation at all in this unit. The analyzer unit in turn also introduces phase shifts, but that has no effect on the polarimeter performance.

In order to describe the performance of the polarimeter, we can now make use of the Stokes and Mueller formalism discussed in the former section. By describing the incoming light by the Stokes vector $\mathbf{S}_0$ and the light that reaches the detector as $\mathbf{S}'$ we have

$$\mathbf{S}' = \mathbf{M}_{\text{pol}}(\eta_1, \eta_2) \cdot \mathbf{S}_0 \tag{3.30}$$

where $\mathbf{M}_{\text{pol}}$ is the Mueller matrix that describes the polarimeter and is a function of the optical constants and the angles η_1 and η_2. The detector measures the total intensity of the light reaching it, which correspond to the first Stokes parameters S_0' of the vector $\mathbf{S}'$.

To obtain the matrix $\mathbf{M}_{\text{pol}}(\eta_1, \eta_2)$ we need to consider the Mueller matrices of every mirror $\mathbf{M}_1$, $\mathbf{M}_2$, $\mathbf{M}_3$ and $\mathbf{M}_4$ as defined by equation (3.28). Additionally, it is necessary to consider the rotation angles between each element related to the optical axis, η_1 and η_2. This gives

$$\begin{aligned} \mathbf{M}_{\text{pol}}(\eta_1, \eta_2) = \quad & \mathbf{M_R}(-\eta_2) \cdot \mathbf{M_4} \cdot \mathbf{M_R}(\eta_2) \cdot \\ & \cdot \mathbf{M_R}(-\eta_1) \cdot \mathbf{M_3} \cdot \mathbf{M_R}(\eta_1) \cdot \\ \cdot \mathbf{M_R}(-180^o) \cdot & \mathbf{M_R}(-\eta_1) \cdot \mathbf{M_2} \cdot \mathbf{M_R}(\eta_1) \cdot \mathbf{M_R}(180) \cdot \\ & \cdot \mathbf{M_R}(-\eta_1) \cdot \mathbf{M_1} \cdot \mathbf{M_R}(\eta_1) \end{aligned} \tag{3.31}$$

where $\mathbf{M_R}$ is the Mueller matrix of rotation (eq. 3.29).

As shown in Appendix B, the 3 mirrors assembly of the retarder unity can be treated as a single element due to the fixed rotation angle of 180deg between the mirrors. This reduces equation (3.31) to

$$\mathbf{M}_{\text{pol}}(\eta_1,\eta_2)=\mathbf{M_R}(-\eta_2)\cdot\mathbf{M_4}\cdot\mathbf{M_R}(\eta_2)\cdot\mathbf{M_R}(-\eta_1)\cdot\mathbf{M_{ret}}\cdot\mathbf{M_R}(\eta_1). \quad (3.32)$$

where $\mathbf{M_{ret}}$ is the Mueller matrix of the retarder unity defined in Appendix B.

Before calculating this matrix, rearranging some variables will simplify the Mueller matrix of a mirror. Namely, defining the parameter ψ such that

$$\tan(\psi)=\frac{r_p}{r_s} \qquad \delta=\delta_s-\delta_p \quad (3.33a)$$

$$\cos(\psi)=\frac{r_s}{\sqrt{r_s^2+r_p^2}} \qquad \sin(\psi)=\frac{r_p}{\sqrt{r_s^2+r_p^2}} \quad (3.33b)$$

$$\cos(2\psi)=\frac{r_s^2-r_p^2}{r_s^2+r_p^2} \qquad \sin(2\psi)=\frac{2\cdot r_s\cdot r_p}{r_s^2+r_p^2} \quad (3.33c)$$

the matrix in equation (3.28) becomes

$$\mathbf{M}(r_s,r_p,\delta)=\frac{r_s^2+r_p^2}{2}\times \times\begin{pmatrix} 1 & -\cos(2\psi) & 0 & 0 \\ -\cos(2\psi) & 1 & 0 & 0 \\ 0 & 0 & \sin(2\psi)\cos\delta & \sin(2\psi)\sin\delta \\ 0 & 0 & -\sin(2\psi)\sin\delta & \sin(2\psi)\cos\delta \end{pmatrix}. \quad (3.34)$$

This reduces the number of parameters inside the brackets from three (r_s, r_p and δ) to two (ψ, δ).

Now we have all the tools to perform the matrix multiplication presented in equation (3.30). However, the result is complicated and only the relevant result will be discussed, which in this case is the result for the parameter S_0', because it is related to the intensity measured on the detector. Solving the matrix multiplication for the term S_0' yields

$$S_0'=\frac{1}{4}\cdot\left({r_{s3}}^2+{r_{p3}}^2\right)\left({r_{s4}}^2+{r_{p4}}^2\right)(c_0\cdot S_0+c_1\cdot S_1+c_2\cdot S_2+c_3\cdot S_3) \quad (3.35)$$

where the parameters c_0, c_1, c_2 and c_3 are given as

$$\begin{aligned} c_0 = & \\ & +1 \\ & +\cos\left(2\eta_2-2\eta_1\right) \quad \times\cos\left(2\psi_3\right)\cos\left(2\psi_4\right), \end{aligned}$$

$$
\begin{aligned}
c_1 = & \\
& -\cos(2\eta_1) && \times \cos(2\psi_3) \\
& -\frac{1}{2}\sin(4\eta_1)\sin(2\eta_2) && \times \cos(2\psi_4)\left[1-\sin(2\psi_3)\cos(\delta_3)\right] \\
& -\cos(2\eta_1)^2\cos(2\eta_2) && \times \cos(2\psi_4) \\
& -\sin(2\eta_1)^2\cos(2\eta_2) && \times \cos(2\psi_3)\sin(2\psi_4)\cos(\delta_3),
\end{aligned}
$$

$$
\begin{aligned}
c_2 = & \\
& -\sin(2\eta_1) && \times \cos(2\psi_3) \\
& -\frac{1}{2}\sin(4\eta_1)\cos(2\eta_2) && \times \cos(2\psi_4)\left[1-\sin(2\psi_3)\cos(\delta_3)\right] \\
& -\sin(2\eta_1)^2\sin(2\eta_2) && \times \cos(2\psi_4), \\
& -\cos(2\eta_1)^2\sin(2\eta_2) && \times \cos(2\psi_3)\sin(2\psi_4)\cos(\delta_3)
\end{aligned}
$$

$$
\begin{aligned}
c_3 = & \\
& -\sin(2\eta_2-2\eta_1) && \times \sin(2\psi_3)\cos(2\psi_4)\sin(\delta_3).
\end{aligned}
$$

For the sake of clarity, the equations have the terms which depend on the rotation angle written on the left-hand side of the multiplication symbol "×". The terms that depend only on the optical parameters are written on the right-hand side.

Finally, normalizing the equation by the incoming intensity S_0 results in the polarimeter equation (following the notation of Appendix B)

$$
\bar{S}_0' = \frac{1}{4}\cdot\left({r_{s3}}^2+{r_{p3}}^2\right)\left({r_{s4}}^2+{r_{p4}}^2\right)\left(c_0+c_1\cdot\bar{S}_1+c_2\cdot\bar{S}_2+c_3\cdot\bar{S}_3\right),
\qquad \begin{matrix}\text{polarimeter}\\ \text{equation}\end{matrix} \tag{3.36}
$$

where the terms with bar are normalized by S_0 as $\bar{S}_i = S_i/S_0$.

Performance of the VUV Polarimeter

Considering the performance of a polarimeter based on retarder and attenuator, an ideal retarder would present a phase shift δ such that $\sin(\delta) = \pm 1$ and no attenuation resulting that $\sin(2\psi) = 1$. An ideal attenuator in turn would totally absorb one component of the radiation and be transparent to the other, resulting in $\cos(2\psi) = \pm 1$.

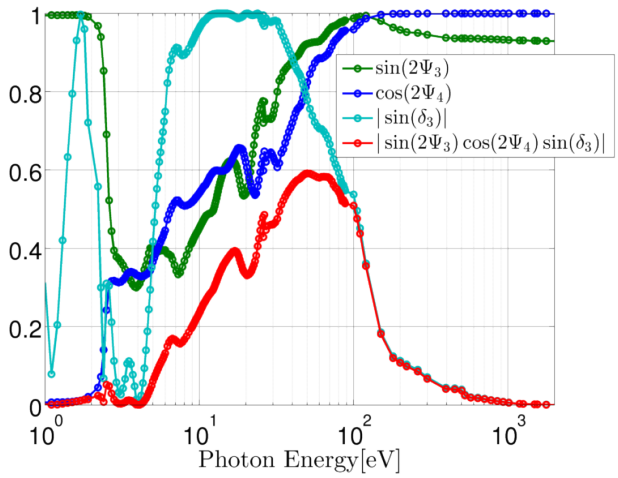


Figure 3.13. *Optical constants of the gold mirrors calculated with Fresnel's equations and data from Palik [1985]. Incidence angles for mirrors M1 to M4 are: 80º , 70º , 80º and 45º , respectively.*

Using the previous notation for the current case, an ideal polarimeter would have optical parameters such that

$$\begin{aligned} \sin(2\psi_3) &= 1, \\ \sin(\delta_3) &= \pm 1, \\ \cos(2\psi_4) &= \pm 1. \end{aligned} \qquad \begin{array}{c} \text{optical parameters} \\ \text{of an ideal} \\ \text{polarimeter} \end{array} \tag{3.37}$$

The value of δ_4 is not relevant since it does not appear in equation (3.36). Further analysis shows that these conditions would greatly simplify equation (3.36) and would maximize the sensitivity of the instrument to circularly polarized light by maximizing the amplitude of the c_3 parameter. With this we assume that the figure of merit for the polarimeter performance is a high value of the product $\left|\sin(2\psi_3)\cos(2\psi_4)\sin(\delta_3)\right|$, which has a maximum possible value of 1.

Figure 3.13 shows the values of these optical parameters for gold for the incidence angles of the polarimeter that has been manufactured. We observe that the best performance occurs around $50eV$ and a good performance is expected in the range of 10-$100eV$.

In addition, it is also possible to use equation (3.36) to study the behavior of the polarimeter for different states of polarization. Noting that the current I measured by the detector is linearly proportional to the photon intensity, we have $I(\eta_1,\eta_2) = E_{\text{fficiency}} \cdot S_0'$, where $E_{\text{fficiency}}$ is the detector efficiency. Therefore it is possible to use tabulated optical constants to estimate the polarimeter response to different states of polarization. The dependence of $I(\eta_1,\eta_2)$ as a function of the rotation angles η_1 and η_2 is shown in Figure 3.14. It can be

observed for instance that linearly polarized light (from fig. 3.14.b to f) results in a multi-peak surface. A completely different behavior is obtained for circular polarized light (fig. 3.14.g and h), which in turn is similar to unpolarized light (fig. 3.14.a).

3.5 Determination of the polarization state of the light

From the analysis shown in the previous sections we know that the intensity measured by the polarimeter has a linear dependence of the Stokes parameters of the incoming light, though it has a complicated dependence with the rotation angles η_1 and η_2 and with the optical constants.

Although the equations have been derived to a polarimeter based on mirrors, they are still quite general and very similar to the corresponding equation for different devices built by others, *e.g.* Gaupp and Mast [1989] using mirrors and Schäfers et al. [1999] and MacDonald et al. [2009] using multilayers.

This theory can now be used to determine the polarization of the light, *i.e.* by measuring $I(\eta_1, \eta_2)$ we can determine the Stokes parameters S_1, S_3 and S_3. Over the years two methods for measuring the polarization of the light have been used. The first combines measurements at specific pairs of angles η_1 and η_2 where the equation can be simplified, as done for instance by Koide et al. [1991], Schledermann and Skibowski [1971] and Hamm et al. [1965]. This method is also shown in Appendix C.3 where it is applied to the present polarimeter with the notations used in this work. The second method is in turn based on measuring the intensity while scanning the angles η_1 and η_2 and performing the least square fitting procedure to the data using equation (3.36) (or other similar, depending on the device) to obtain the Stokes parameters and the optical constants of the polarimeter. The later method has been used for instance by Schäfers et al. [1999] and by Nahon and Alcaraz [2004].

The first method has the advantage of being fast, since it requires fewer measurements, but it is more sensitive to experimental errors, like misalignment. It can however be very useful for preliminary analysis and determination of initial values for the fitting method. The second methods is less sensitive to errors but due to the fitting process it is difficult to estimate the error in determination of the Stokes parameters, since this is generally done by statistical estimations [Bevington and Robinson, 2003; Nahon and Alcaraz, 2004]. Furthermore, due to the complicated function, it is difficult to reliably estimate the fitting convergence, existence of local minimums when minimizing χ^2 and existence of a unique solution.

In respect to that, Paper **III** describes a new method, where the angles η_1 and η_2 are scanned under certain conditions. These con-

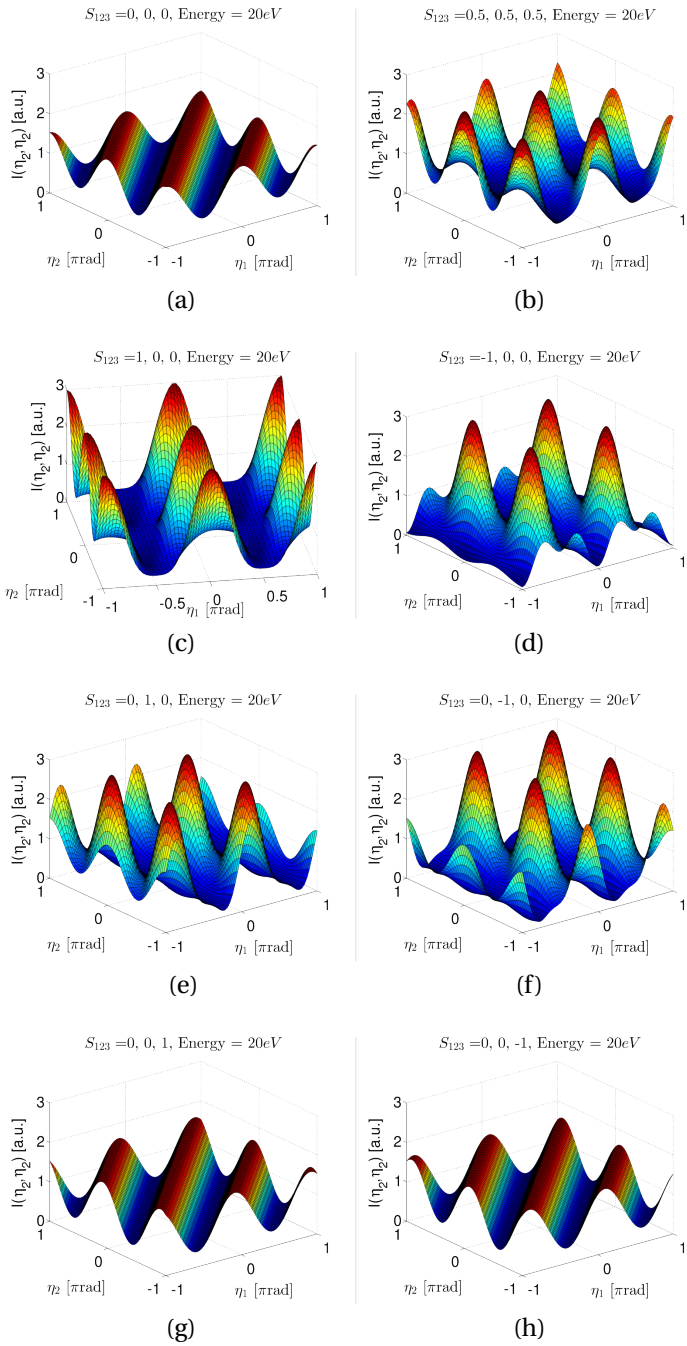


Figure 3.14. *Plot of equation (3.36) for different polarization states (Stokes parameters on the titles of each figure): (a) Unpolarized light. (b) Elliptically polarized light. (c-f) Linearly polarized light. (g-h) Circular polarized light. Photon energy 20e V . Incidence angles for mirrors M1 to M4 are: 80º , 70º , 80º and 45º , respectively. Gold data from Palik [1985].*

ditions are discussed in the paper and derived in more details in the Appendix C.2.

The initial motivation to develop this method was to simplify the data analysis by seeking rotation angles that simplify equation (3.36). It turned out that fixing the angular difference $\eta_2 - \eta_1$ to $\frac{m\pi}{4}$, with $m = -1, 0, 1, 2$ results in four special cases where the measured intensity $I(\eta_1, \eta_2) = I(\eta, \eta + \frac{m\pi}{4})$ can be always reduced to the form

$$I(\eta) = \bar{I}\left[1 + V \cdot \sin(2\eta + \Psi)\right],$$

where $\bar{I}$, V and Ψ are the experimental quantities mean value of $I(\eta)$, visibility $V = (I_{max} - I_{min})/(I_{max} + I_{min})$ and phase Ψ, respectively. Therefore, these special cases greatly simplify the intensity equation and link the analysis to the experimental quantities $\bar{I}$, V and Ψ that can be easily measured.

Moreover, it can be shown [Paper **III**] that these three experimental quantities of the four special cases can be used to fully determine the Strokes parameters of the incoming light. These direct relations between the Stokes parameters and the experimental quantities $\bar{I}$, V and Ψ have the further advantage of avoiding any fitting procedure and allowing to study the effects of the experimental errors to the errors for the Stokes parameters.

Multilayer based polarimeter

Figures 3.9 and 3.13 show that the gold mirrors have poor performance for energies above 100eV. This is the case for most of the materials since the refractive index in general behaves as shown in Figure 3.6. To overcome this limitation, alternative optical devices are required for doing polarimetry, and the most common solution for the energy range above 100eV and up to *ca* 1500eV is to use reflection and transmission multilayers [Gaupp et al., 2010; Schäfers et al., 1999].

For this energy range a second polarimeter based on multilayers has been commissioned at MAX-lab and it is described in Paper **IV**. The motivation for using multilayers is discussed elsewhere, for instance by Schäfers et al. [1999]. The discussion here is based on the fact that the polarimeter equation (eq. 3.37) is also valid for the multilayer-based polarimeter, where only minor changes are necessary, the most important being to interpret some optical parameters as transmission coefficients when working with transmission multilayers. Regarding the new method on Paper **III**, some caution is necessary when applying it for the multilayer based polarimeter. The reason for this are assumptions made, for instance that $\cos\psi > 0$, which very likely is not true for the working conditions of the transmission multilayer. Nevertheless, the same reasoning of Paper **III** can be applied to a multilayer-based polarimeter and a new set of equations can be derived, and only minor changes are expected.

CHAPTER 4

WAVE OPTICS APPLIED TO SR

The Diffraction Limited Storage Rings (DLSR) put new challenges in beamline modeling as the light has a high degree of coherence. The models must provide information about how this kind of light behaves when propagating trough the beamline and eventually the degree of coherence of the beam at the experimental station. There are various strategies for modeling beamlines at DLSR, but in general they all have two main steps in common: first, the light emitted by a single electron is propagated trough the optical elements of a desired optical system; secondly, the light from many electrons are summed up. This latter step can be done either by properly convoluting the electron beam size with the light beam dimensions, or by propagating the light emitted by each electron in the bunch and then summing up the light in the end (brute force method). The propagation of the light emitted by a single electron is completely based on wave optics and is referred as *wavefront propagation*. The rules for summing up the resulting light is dictated by statistical optics, which also provides information about the degree of coherence.

This chapter will start with a brief discussion about diffraction limited sources and the motivation to use wavefront propagation. Then the main concepts of wave optics are introduced in order to understand the main tools in modeling optical system, namely Fourier optics and the stationary phase approximation method. Finally, statistics optics will be introduced in order to understand the modeling of partially coherent beam.

4.1 Diffraction Limited Storage Rings

Synchrotron radiation from storage rings is an outstanding light source due to the high brilliance of the produced photon beam. Brilliance is defined as the photon flux divided by the photon beam

area and divergences[1]. Using the statistical root-mean-square (rms) values for the photon beam size σ_{xy} and divergence σ'_{xy}, we have that the brilliance $\mathscr{B}(\omega)$ can be written as (following notation of Ch. 2)

$$\mathscr{B}(\omega) = \frac{\Phi(\omega, \hat{u})}{\sigma_x \sigma'_x \sigma_y \sigma'_y}, \qquad \text{Brilliance}$$

where the sub indexes x and y refer to the horizontal and vertical directions respectively. The product of the rms values of size and divergence $\sigma\sigma'$ can be interpreted as the area in the phase space, which is referred as *emittance*, ε,

$$\varepsilon = \sigma\sigma', \qquad \text{Emittance} \tag{4.1}$$

and thus the expression for brilliance becomes

$$\mathscr{B}(\omega) = \frac{\Phi(\omega, \hat{u})}{\varepsilon_x \varepsilon_y}. \qquad \text{Brilliance} \tag{4.2}$$

Since the brilliance is the main quality of light generated at storage rings, this means that improvements are either done by increasing flux or by decreasing the beam emittance. Third generation storage rings provided an increase in brilliance by using undulators to increase the flux. Since the flux from the undulators scales with the square of the number of periods ($\Phi(\omega, \hat{u}) \propto N^2$, see Section 2.1) and $N \approx 10-30$ periods, undulators provided a increase of brilliance of the order of $\approx 100-1000$ times compare to second generation storage rings.

Another obvious way to increase flux, and thus brilliance, would be by simply increasing the number of electrons circulating in the storage ring (ring current). However this has some limitations, since operating at high currents will lead to beam instabilities, which eventually increases emittance and decreases the life time of the electron beam [Skripka et al., 2014]. Therefore, since the flux increases only linearly with the number of electrons, this option is not very attractive. For instance, MAX IV rings will operate with 500mA, only two times higher than the 250mA previously used at MAX II (updated values for other facilities are listed by Hettel [2014]).

More recently the strategy has been to decrease the beam emittance, since it is more effective than increasing the flux, as the brilliance scales as $1/\varepsilon^2$. For a storage ring with electron energy E, the emittance scales as E^2/N_d^3, N_d being the number of dipoles in the ring. Therefore an increase in the number of dipoles N_d by a factor two contributes alone to a decrease in the emittance almost by a factor of ten, and thus an increase in brilliance by a factor of *circa* 100.

[1] In the literature there are some confusion with the terms *brilliance* and *brightness*, and often they are used to refer to the same quantity as discussed in Mills et al. [2005], Peatman [1997] and Clark [2004].

This is the main idea in the so-called multi-bend achromat concept [Einfeld et al., 2014] that is used for the first time at MAX IV, where the 3GeV ring uses 140 dipoles compared with only 20 used at MAX II [Tavares et al., 2014].

There is, however, a limit for decreasing the emittance of the photon beam. This is called *diffraction limit*, which has origin in the uncertainty principle and is derived in the context of quantum optics [Mandel and Wolf, 1995; Saleh and Teich, 2007]. This limitation gives the minimum value of the emittance for a photon beam ε_r, and it is given by[2]

$$\varepsilon_r(\lambda) = \sigma\sigma' \geq \frac{\lambda}{4\pi}. \qquad \text{Diffraction Limit} \tag{4.3}$$

where λ is the wavelength of the radiation. This means that even a source of negligible dimensions (like a single electron orbiting a storage ring) produces radiation with dimensions dictated by the diffraction limit in Eq. 4.3.

It must be noted that this is a lower limit that is reached for instance for a Gaussian beam [see discussion in Saleh and Teich, 2007, Section 12.3]. However, since the value of emittance is defined in terms of rms values of size and divergence, the emittance of the radiation will actually depend on the spatial and angular distribution of the radiation in question, which in turn depends on the emission process. This can be seen explicitly for the case of undulators in Eq. 2.17, where the resulting undulator emittance is bigger than the minimum value provided by Eq. 4.3. In addition, we can see in Eq. 2.18 that out of the resonance frequency the emittance can be slightly different [3].

Due to the finite emittance of the electron beam, in order to obtain the real dimensions and divergences of the source it is necessary to take into account both the emittance of electron beam and the diffraction limited emittance of the photon beam. This can be done by convolution of the two emittances. However, the diffraction limited emittance is energy dependent, and at some low photon energies that will be much larger than the electron beam emittance, that is $\varepsilon_r(\lambda) \gg \varepsilon_e$. At the photon energies where $\varepsilon_r(\lambda)$ is dominant, the source is said to be diffraction limited.

For instance MAX IV 3GeV and NSLS-II storage rings have both a design goal to be diffraction limited in the vertical direction up

[2] The term *diffraction limited system* is also used to described aberration free optical systems, where the only limit for spatial resolution is given by the response function of this system, which in turn depends on the diffraction of light by the apertures of such system. An illustrative example is a circular lens imaging a point source. The image size will be defined by the diffraction on the lens aperture and it will result in an Airy pattern. Such (idealized) image is said to be diffraction limited [Goodman, 2005; Saleh and Teich, 2007].

[3] Strictly speaking, as argued by Bazarov [2012], this mean that undulators radiation is not diffraction limited. In this sense, concerning storage rings, diffraction limited actually means limited by the properties of undulators radiation.

to $\lambda = 1$Å [MAX IV Facility, 2010] [Brookhaven National Laboratory, 2006], and therefore they are said to be diffraction limited storage rings[4]. This also means that not all the beamlines in the same storage ring operate in the diffraction limited regime, since not all of them operate at wavelengths where the source is diffraction limited.

Coherence at DLSR

Statistical optics provides the proper tools for discussing coherence, and such approach is discussed in details later (Section 4.4). A statistical approach is necessary since the source is an ensemble of individual point sources, that is, the *ca.* 10^{18} electrons at each bunch circulating in the ring.

The degree of coherence can be interpreted as the correlation between the radiation in two different points in space or in time (or both). Mathematically, degree of coherence is the normalized value of the correlation function of the EM field in those two points (see eq. 4.29). The degree of coherence assumes the form of a distribution and the rms value corresponding to this distribution is called the coherence length. Within this length the degree of coherence is high and close to the maximum value of 1 (which means completely coherent radiation).

This degree of coherence depends on the correlation between the different points of the source. If these source points are completely uncorrelated and emit radiation randomly, then this source is said to be completely incoherent. On the other hand if the point sources are emitting radiation in an ordered manner, then the source is coherent (or partially coherent) with some degree of coherence. In a storage ring there are roughly $\approx 10^{18}$ electrons in a volume of about $100\mu m \times 10\mu m \times 10mm$ (approximated rms values of a electron bunch), and thus they are assumed to behave as incoherent source. However, one important result from statistical optics is that the light from a completely incoherent source can gain coherence trough propagation in free space. This is a consequence of the van Cittert-Zernike theorem [Born and Wolf, 1999; Mandel and Wolf, 1995; Wolf, 2007], which is discussed later in pg. 82. In this case the coherence length l_c is proportional to

$$l_c \propto \frac{\lambda d_z}{\sigma} \tag{4.4}$$

where λ is the wavelength of the radiation, σ the source size and d_z the propagating distance between the source and the observation

[4]The definition of diffraction limited storage ring is a bit arbitrary. First of all, it is commonly used the diffraction limit relation for Gaussian sources, instead of real undulators distribution. Secondly, there is no definition of a reference wavelength. And finally, often the horizontal and vertical electron emittance are quite different ($\varepsilon_{ey}/\varepsilon_{ex} \lesssim 100$) and thus for many cases the source is diffraction limited in only one direction (in general in the vertical). Hence, this definition may be used only as a guiding parameter.

plane. Therefore, the larger the propagation distance and the smaller the source size, the bigger the coherence length.

This means that storage rings produce light with a certain coherence length, even though the electrons are randomly distributed in the electron bunch and that they form an incoherent source. Further on, a diffraction limited beam that has propagated a distance d_z will produce a beam with size Σ given by

$$\Sigma = d_z \cdot \sigma' = \frac{\lambda d_z}{\sigma 4\pi} \tag{4.5}$$

and thus of the order of the coherence length presented in Eq. 4.4. In other words, for a diffraction limited source the coherence length l_c is of the order of the (propagated) beam size Σ and therefore the beam has a high degree of coherence (See Figures 4.1 and 4.2).

This means that the beam generated at a DLSR has a high degree of coherence for the wavelengths where the source is diffraction limited. Based on equation (4.3), one sees that this happens at long wavelength (low photon energies). In case of short wavelengths (high photon energies) the source produces radiation that is incoherent (*i.e.*, it has low degree of coherence), and for wavelengths between those extremes the radiation is said to be partially coherent.

Coherent X-rays brings new opportunities in science mostly due to the previous lack of sources for coherent X-rays [Eriksson et al., 2014]. This also put new challenges in beamline modeling, where the most important challenge is the propagation of coherent beam trough grazing incidence optics, a particular feature of X-ray optics. This means that most of the theory developed for visible light optics needs to be revised. For instance, propagation of coherent light in visible optics relies heavily in the concepts of *Fourier optics*, which extensively makes use of paraxial approximations and also assumes that optical elements are thin. These two approximations alone put serious restrictions for the use of Fourier optics for X-ray beamlines.

4.2 Wave Optics

Although geometric optics is very powerful in explaining many phenomena in optics, it cannot explain for example diffraction and interference of light. These two phenomena are a consequence of the wave nature of light, the light being constituted by electromagnetic waves [Born and Wolf, 1999]. Electromagnetic waves are described by the Maxwell's equations, but many properties of light are resulting of the wave properties alone and can be studied even if the electromagnetic properties are ignored (or put aside). Wave optics assumes this and light is then described by a complex scalar wave function, subject to the wave equation[5].

[5]The scalar approximation of the vectorial waves is discussed in [Born and Wolf, 1999, Ch. 8.4], [Goodman, 2005, Ch. 3.2] and [Jackson, 1997, Ch. 10.7]

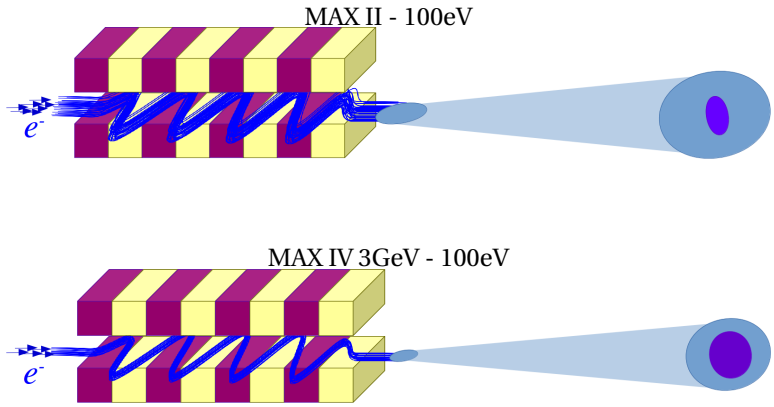


Figure 4.1. *Comparison of photon beam size (light blue) and coherence length (dark blue) after propagation in free space, where actual values for MAX II and MAX IV 3GeV are used to compare a typical 3rd generation source and a DLSR (values given in Figure 4.13). At low energy (≈100eV) the propagated beam size is mostly defined by the divergence of the photon beam, and it results in an almost round beam. (a) MAX II: The coherence length in vertical is about 50% of the vertical size but only about 10% of the horizontal dimension. (b) MAX IV 3GeV: there is a small decrease in the beam size due to the smaller emittance of the electron beam. In addition, the coherence length increases to 63% and 57% in the vertical and horizontal direction, respectively. The values are calculated 10m from the undulator and they are obtained by using van Cittert-Zernike theorem.*

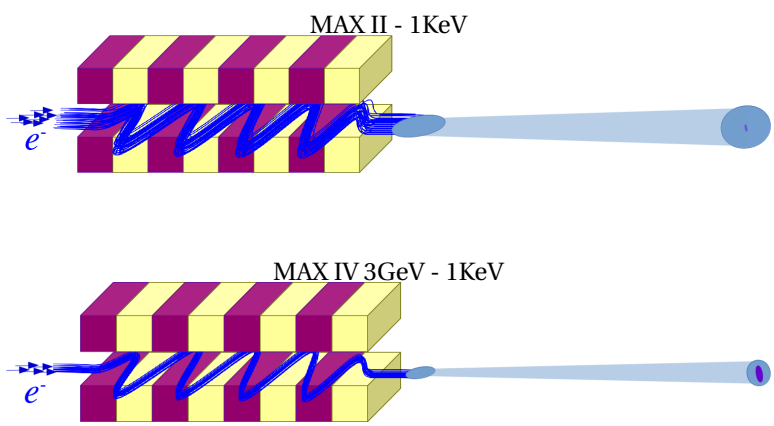


Figure 4.2. *Same as above, but for photon energy of 1keV. (a) MAX II: beam size is mostly defined by emittance of the electron beam and the fraction covered by the coherence length is ca 20% in the vertical and only about 2% in the horizontal direction. (b) MAX IV 3GeV: in vertical direction the fraction is still around 62% whereas in the horizontal it is 32%, a decrease compared to low energy cases but a big improved compared to MAX II.*

As mentioned earlier, wave effects manifest themselves when the wavelength of the radiation is of the order of the geometrical distances involved or when the degree of coherence is high. This makes wave optics essential for studying some beamlines at DLSR, since the low emittance makes possible to obtain high coherent flux in a wide energy range. In addition, only a combination of wave and statistical optics can estimate the degree of coherence of the beam at the end station, which is crucial in some projects.

It is not the goal of this work to make an extensive review of wave optics, that is extensively done in the literature [*e.g.* Born and Wolf, 1999; Goodman, 2005; Saleh and Teich, 2007]. Instead, the following sections will discuss some concepts of wave optics and how they apply in modeling a X-ray beamlines. Special attention is given to how some concepts of visible/paraxial optics have limited applicability for X-ray optics.

Wave Equation

The main postulate of wave optics is that light waves must satisfy the wave equation:

$$\nabla U(\vec{r},t) - \frac{1}{c}\frac{\partial^2 U(\vec{r},t)}{\partial t^2} = 0, \qquad \begin{matrix}\text{Wave}\\ \text{Equation}\end{matrix} \tag{4.6}$$

where U is a scalar complex function, $\vec{r}$ is a spatial coordinate, t is time and c is the speed of light.

Different symmetries, approximations and boundary conditions will produce different results for $U(\vec{r},t)$, the most elementary ones being a plane wave and a spherical wave. Wave optics studies the effects in propagating waves caused by the optical elements like mirrors, lenses and prisms (for a list of examples see [Saleh and Teich, 2007, Chap. 2]). For visible optics one can take advantage of paraxial approximations and the usual circular symmetry of lenses. In addition, the optical elements are usually in the same optical axis, which simplifies any change of coordinates.

Wave optics with these approximations for visible light is extensively studied in the literature. However these approximations are not valid for X-ray optics, which is mainly based on using curved grazing incidence mirrors. There one can not use paraxial approximation, change of coordinates is not simple and the curved surfaces have a complicated mathematical description in a generic reference system. Lack of coherent X-ray sources also contributed to a lack of interest in finding analytical solutions, and for this reason there are not many examples in the literature.

Gaussian beam

Most of the results in wave optics can be derived by using planar and spherical waves. However, this kind of waves are not appropriate for

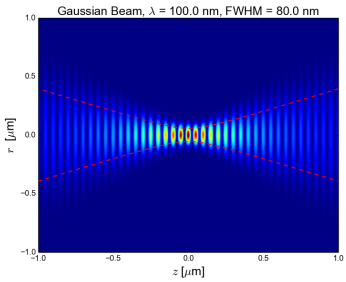


Figure 4.3. *Representation of Gaussian beam of $\lambda = 100$nm and beam waist FWHM equals to 80nm propagating in z direction. r is radial direction. Red dashed lines show FWHM values as the beam propagates*

describing a beam of light, that is, a spatially confined light.

There are alternatives for describing a light beam[6], and the most common one is the so-called Gaussian beam. Gaussian beam is a solution of the paraxial wave equation (wave equation under paraxial approximation) and its main properties are that the beam power is concentrated around the beam axis (see Fig. 4.3) and that when it is propagated trough an optical system it remains a Gaussian beam. In addition, Gaussian functions work in a well known manner under Fourier transform, making some studies in Fourier Optics straightforward for the Gaussian beams[7].

Gaussian beams is very often used as approximation for undulator radiation in the diffraction limited regime, as done by Kim [1986], Onuki and Elleaume [2003] and Bazarov [2012], to cite a few. However we can note from Section 2.1 that this is not always true and many of the above mentioned authors have actually mentioned this fact [*e.g.* Bazarov, 2012; Kim, 1986]. Gaussian beam has smaller emittance than undulator beam, which means that an undulator beam and a Gaussian beam of same size will have different divergences. Nevertheless Gaussian beams provide, in general, a useful approximation but that requires caution when studying undulator based beamlines at low emittance storage rings (see also the discussion in Onuki and Elleaume [2003, Chapter 3 Section 2.6]).

4.3 Diffraction Theory

Diffraction is any deviation of light that cannot be interpreted as reflection or refraction (following the definition by Sommerfeld [1949]), and it is caused by a transverse confinement of a wave [Goodman, 2005]. From a physical point of view there is no difference between diffraction and interference: diffraction is seen simply as a superposition of a large number of waves [Hecht, 2002].

Calculating rigorous solutions for a diffracted wave is one of the most difficult problems in optics [Born and Wolf, 1999; Hecht, 2002] and general analytical solutions do not exist. Therefore, approximate models limited to certain conditions are necessary. Of such models, the most commonly used are [Born and Wolf, 1999; Goodman, 2005]: the *Huygens-Fresnel theory*, which was first derived based on intuitive assumptions; the *Kirchhoff's diffraction theory*, which has some internal theoretical inconsistencies; and the *Rayleigh-Sommerfeld diffraction theory*, which is restricted to the diffraction by a aperture in a planar screen.

[6] For instance Hermite-Gaussian and Laguerre-Gaussian beams [Saleh and Teich, 2007], where the Gaussian beam is a particular case of both kinds of beam.

[7] There are additional advantages in laser optics that make Gaussian beams even more useful. For instance it is the fundamental mode of light in a spherical mirror resonators, which in turn is present in many kinds of lasers and for this reason Gaussian beam is a very good description of this kind of laser beam.

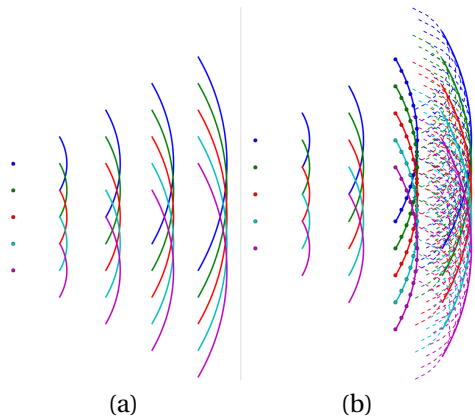


Figure 4.5. *Representation of Huygens principle. In this example each point source emits a spherical wave in phase with each other (a) Point sources and their related wave fronts (associated by the color). The resulting wave is the envelop of these wavefronts. (b) In addition, Huygens principle states that each point in the wavefront can be treat as a new secondary point source that generates secondary spherical waves, the wavelets (represented by dashed lines). Similarly, the final wave is the envelop of the wavelets.*

The Huygens principle is illustrated in Figure 4.4 and Figure 4.5 and it states that each point of a wavefront acts as a secondary point source for secondary spherical waves (*wavelets*, see Figure 4.5.b). As these wavelets propagate further, the total wavefront is the envelope of all wavelets. Although it was first derived based on some incorrect concepts, as that light propagates in the ether [Huygens, 1690], it can actually be derived from the Kirchhoff's and Rayleigh-Sommerfeld diffraction theories [Goodman, 2005]. Eventually, it can be shown that Huygens principle is a consequence of the wave equation [Hecht, 2002; Sommerfeld, 1949] and therefore a consequence of the wave nature of light (see for instance [Goodman, 2005, ch. 3.7]).

Diffraction can be described with reasonable accuracy by using Huygens principle and assuming interference of the waves, which is called *Huygens-Fresnel* principle. A mathematical description of diffraction based on the Huygens-Fresnel principle is discussed by[8] Born and Wolf [1999]. Besides the accuracy, Huygens-Fresnel principle is an important model as it offers a physical interpretation of otherwise complex mathematical results.

More formal mathematical descriptions of diffraction are provided by the Kirchhoff's and Rayleigh-Sommerfeld diffraction theo-

Figure 4.4. *Huygens principle states that each point of a light source, like A, B or C, can be assumed as an individual point sources that radiate a spherical wave. Figure from the original by Huygens [1690].*

[8] Once more there is no agreement about conventions in the literature. Goodman [2005] uses the term Huygens-Fresnel principle to refer to the first solution of Rayleigh-Sommerfeld diffraction theory, while Born and Wolf [1999] treat them as different descriptions despite the similar results. Here the results of Rayleigh-Sommerfeld diffraction theory are regarded separately of Huygens-Fresnel principle.

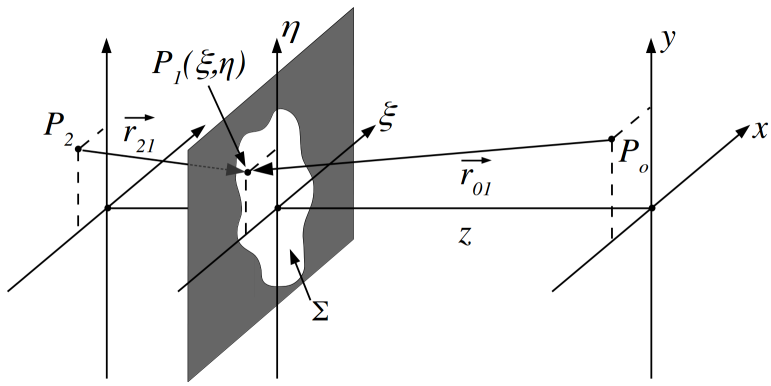


Figure 4.6. *Geometry for the Fresnel-Kirchoff and Rayleigh-Sommerfeld theory of diffraction. A point source in P_2 will generate a field $U(P_o)$ at P_o. The radiation is diffracted when passing trough the aperture Σ in the plane screen. This planar geometry of the screen is a requirement for the Rayleigh-Sommerfeld solution, but not for the Fresnel-Kirchoff description.*

ries. Both of them are based on Green theorem and wave equation [Goodman, 2005]. In fact, the main difference between them is the choice of the auxiliary Green's function: while the Kirchhoff's theory uses a spherical wave, Rayleigh-Sommerfeld description is based on a combination of two spherical waves. The reason for the difference is to fulfill different boundary conditions. The choice of the Green function in the Kirchhoff theory contradicts the boundary conditions used to derive the result [Goodman, 2005, Ch. 3.5]. This is solved in the Rayleigh-Sommerfeld model by using different Green's function and different boundary conditions. However, these later calculations are valid only for diffraction of light by apertures in plane surfaces, as in Figure 4.6 [Goodman, 2005, Ch. 3.6] [Born and Wolf, 1999, Ch. 8.11.2]. Mathematically both theories shows that a point source at P_2 in Figure 4.6 will generate a field $U_o(P_0)$ given by [Goodman, 2005]

$$U_o(P_o) = \frac{A}{i\lambda}\iint_{\Sigma}\frac{\exp\left[jk(r_{21}+r_{01})\right]}{r_{21}\,r_{01}}\Psi\,ds \tag{4.7}$$

where A, λ and k are the amplitude, wavelength and wavenumber of the radiation from the point source, respectively. The other geometric factors are defined in Figure 4.6. The parameter Ψ is known as obliquity factor and, depending on the model, is given by

$$\Psi = \begin{cases} \frac{1}{2}\left[\hat{n}\cdot\hat{r}_{21} - \hat{n}\cdot\hat{r}_{01}\right] & \text{Kirchhoff,} \\ \hat{n}\cdot\hat{r}_{01} & \text{Rayleigh-Sommerfeld I,} \\ -\hat{n}\cdot\hat{r}_{21} & \text{Rayleigh-Sommerfeld II.} \end{cases} \tag{4.8}$$

The dot $(\cdot)$ in equation (4.8) is used to note the scalar product of vectors. The first solution is known as *Fresnel-Kirchoff integral*.

Rayleigh-Sommerfeld theory derives two solutions based on two different Green's functions[9].

If we consider now a source that is not a point but rather a distribution given by $U(P_2)$, then the field $U(P_o)$ is a sum of the field generated by each point in the source plane containing P_2. Mathematically this results in

$$U_o(P_o) = \frac{1}{i\lambda}\int\!\!\int_2 \left\{ \int\!\!\int_\Sigma \frac{\exp\left[jk(r_{21}+r_{01})\right]}{r_{21}\,r_{01}} \Psi\, ds \right\} U(P_2)\, dx_2\, dy_2, \quad (4.9)$$

where $\int\!\!\int_2 dx_2\, dy_2$ represents integration over the source plane.

The discussion of the validity of these models is extensive and done elsewhere, for instance in the papers by Wolf and Marchand [1964] and by Lucke [2006], and in the books by Goodman [2005, Ch. 3.6] and Born and Wolf [1999, pg. 516]. There is however a consensus that all of them produce accurate results far away from the diffraction screen (*i.e.*, a distance of many wavelengths far from the plane Σ).

Rayleigh-Sommerfeld is favored for solving problems involving diffraction in paraxial optics since it is simpler than Fresnel-Kirchoff and the limitation of planar diffracting screen is not serious in paraxial optics, as it is actually almost always the case. Additional paraxial approximations simplify the equations even further and results in the so-called Fourier optics, to be discussed soon.

On the other hand the plane screen limitation of Rayleigh-Sommerfeld is a serious constraint for grazing incidence optics and for this reason Fresnel-Kirchoff is more appropriate for X-ray optics. However the calculation of the field by using Fresnel-Kirchoff is still very complicated to be done analytically and also very demanding to be done numerically, and an additional approximation is used, namely the Stationary Phase Approximation.

Fourier Optics

An alternative form for Rayleigh-Sommerfeld solution is given by

$$U_o(P_o) = \frac{z}{i\lambda}\int\!\!\int_\Sigma U_1(P_1)\frac{\exp(k\,r_{01})}{r_{01}}\, ds \qquad \begin{matrix}\text{Rayleigh-}\\ \text{Sommerfeld}\end{matrix} \qquad (4.10)$$

where $U_1(P_1)$ is the field in the aperture Σ. Note that substituting $U_1(P_1)$ by a spherical wave one gets equation (4.7). In a rectangular coordinate system the distance r_{01} is then given by (see Figure 4.6)

$$r_{01} = \sqrt{z^2 + (x-\xi)^2 + (y-\eta)^2}.$$

[9]Hereafter Rayleigh-Sommerfeld is used to refer only to the first Rayleigh-Sommerfeld solution.

This allows rewriting equation (4.10) as

$$U_o(x,y) = \iint_{\Sigma} U(\xi,\eta) h(x-\xi, y-\eta)\, d\xi\, d\eta \quad \begin{matrix}\text{Rayleigh-}\\ \text{Sommerfeld}\end{matrix} \tag{4.11}$$

where

$$h(x,y) = \frac{z}{i\lambda}\frac{\exp(kr)}{r}, \qquad r = \sqrt{z^2+x^2+y^2}. \quad \begin{matrix}\text{Response function}\\ \text{Rayleigh-Sommerfeld}\end{matrix} \tag{4.12}$$

The previous manipulation allows interpreting equation (4.11) as a convolution integral where $h(x,y)$ is the response function (also called point spread function). By applying the Fourier convolution theorem one obtains [Voelz, 2011]

$$U(x,y) = \mathscr{F}^{-1}\left\{\mathscr{F}\left\{U(x,y)\right\}\mathscr{F}\left\{h(x,y)\right\}\right\}, \tag{4.13}$$

where $\mathscr{F}$ and $\mathscr{F}^{-1}$ are the Fourier and inverse Fourier transforms, respectively. For a generic function $g(x,y)$ the Fourier Transform is defined as

$$G(f_X, f_Y) = \mathscr{F}\left\{g(x,y)\right\} = \int_{-\infty}^{\infty}\int g(x,y)\exp\left[-j2\pi(f_X x + f_Y y)\right] dx\, dy. \tag{4.14}$$

This is the basis of the Fourier optics and in fact this applies not only for problems in diffraction but also to any wave propagation problem with a geometry similar to the one presented in Figure 4.6. There are many advantages in this approach: diffraction and wave propagation can be understood as a convolution problem; the analytical Fourier transform of many functions are well known, which make easier to obtain analytical solution for many problems in optics; the problem can be studied in the frequency domain; and finally, when performing numerical integration one can take advantage of the Fast Fourier Transform (FFT) algorithm[Brigham, 1988], which is much more efficient[10] than simple integration (more about numerical integration at the discussion about practical issues on 78).

Concerning paraxial optics, Rayleigh-Sommerfeld theory is the most accurate model that is usually used. Two additional approximations simplify the problem for analytical analysis and will be discussed next, but it must be noted that for computational purposes

[10] The efficiency in the calculation (also known as *time complexity*) of a computational problem is how the computational running time grows related to the number of entries. For instance, the computational running time of a problem with n entries can be proportional to n or n^2 . There are formal definitions of how to express the complexity of a computational problem based on asymptotic notation, where the so called (big) O-notation is used to express the worst case running time of an algorithm[Cormen, 2009]. For instance the Fourier Transform of a vector with n entries runs in $O(n\log n)$ time when calculated by the Fast Fourier Transform algorithm, opposite to $O(n^2)$ in a simple (naive) numerical integration [Brigham, 1988].

the above description is enough, since further approximations will not improve the efficiency of the computational calculation (the running time will still be $O(n \log n)$, see below[10]), but only simplify the analytical results.

Fresnel and Fraunhofer Approximations

Fresnel approximation often appears when applying paraxial approximations to problems in optics, for instance when using parabolic waves as approximations to spherical waves [Saleh and Teich, 2007]. Following Figure 4.6, Fresnel approximation assumes that the wave propagates at distances close to the main axis $\hat{z}$ (the paraxial approximation *per se*) but for a distance z long enough so that $\sqrt{x^2+y^2} \ll z$. The paraxial approximation is then applied to the terms $1/r_{01}^2$ and $k\,r_{01}$ in equation (4.10). The term $1/r_{01}^2$ becomes $1/z^2$. Then the binomial expansion

$$\sqrt{1+b} = 1 + \frac{1}{2}b - \frac{1}{8}b^2 + \cdots$$

can be used and when applied (up to the first order in b) one gets:

$$k\,r_{01} \approx k\,z\left[1 + \frac{1}{2}\left(\frac{x-\xi}{z}\right)^2 + \frac{1}{2}\left(\frac{y-\eta}{z}\right)^2\right].$$

The above approximation (up to the first order in b) is the so called *Fresnel approximation* and for the values of z where this approximation is valid *Fresnel diffraction* applies. Rayleigh-Sommerfeld solution under this approximation can be reduced to the convolution form of equation (4.11), where the response function is given by

$$h(x,y) = \frac{e^{ikz}}{i\lambda z}\exp\left[\frac{ik}{2z}(x^2+y^2)\right] \qquad \begin{matrix}\text{Response Function}\\ \text{Fresnel Diffraction}\end{matrix} \qquad (4.15)$$

and equation (4.13) also holds.

In the region where we can also ignore the linear term of the expansion (linear term in b), *Fraunhofer diffraction* applies. This simplification reduces equation (4.10) to

$$U_o(P_o) = \frac{e^{ikz}}{i\lambda z}\exp\left[\frac{ik}{2z}(x^2+y^2)\right]\mathscr{F}\left\{U_1(\eta,\xi)\right\}. \qquad \begin{matrix}\text{Fraunhofer}\\ \text{Diffraction}\end{matrix} \qquad (4.16)$$

That is, besides phase factors, the Fraunhofer diffraction results in the Fourier transform of the source field.

In order to these approximations to hold, it is necessary that the distance z respect the relations given by

$$z^3_{\text{Fresnel}} \gg \frac{\pi}{4\lambda}\left[(x-\eta)^2 + (y-\xi)^2\right]^2_{\max} \qquad \begin{matrix}\text{Fresnel}\\ \text{Diffraction}\end{matrix} \qquad (4.17a)$$

and

$$z_{\text{Fraunhofer}} \gg \frac{\pi}{\lambda}\left(\eta^2+\xi^2\right)_{\max}, \qquad \begin{matrix}\text{Fraunhofer}\\ \text{Diffraction}\end{matrix} \tag{4.17b}$$

where the subscript $_{\max}$ means the maximum value assumed by the therm inside the brackets. We note that $z_{\text{Fraunhofer}} \gg z_{\text{Fresnel}}$, which means that for short distances (*near field*) we have Fresnel diffraction[11] and for longer distances (*far field*) we have Fraunhofer diffraction (See Figure 4.7).

An alternative definition in commonly used in terms of the Fresnel number, which is give by

$$N_F = \frac{a^2}{\lambda z} \qquad \text{Fresnel Number} \tag{4.18}$$

where a is the aperture of the diffracting screen. Fresnel number is defined in the context of the Huygens-Fresnel theory [Born and Wolf, 1999] but it is commonly used in other areas of optics. In terms of the Fresnel number, equation (4.17).b becomes

$$N_F \ll 1. \qquad \text{Fraunhofer Diffraction} \tag{4.19}$$

Figure 4.7 shows how the diffraction profile changes as $N_F \to 0$.

By comparing the response function in the Rayleigh-Sommerfeld and Fresnel diffraction one sees that the first one is a spherical wave while the second one is a parabolic wave. The convolution of the field with the response function can then be interpreted as a mathematical description of the Huygens principle, where in the Rayleigh-Sommerfeld case we have spherical waves and in the Fresnel diffraction we have parabolic waves. An additional advantage of Fraunhofer diffraction is that the same result is obtained for some geometries using lens, and thus it can be used when thin lens approximation is in place, like for compound refractive lenses, zone plates and also when some computations use thin lens approximation for representing grazing incidence mirrors.

Stationary Phase Approximation

The method of stationary phase approximation (SPA) uses asymptotic solution to the Fresnel-Kirchhoff integral, with the goal to solve problems of diffraction and wave propagation trough optical elements [Bahrdt, 2007a][Paper **V**]. The method is based in a asymptotic approximation in solving the following type of integrals [Born and Wolf, 1999; Mandel and Wolf, 1995]:

[11] It can be shown that actually Fresnel diffraction works well for values smaller than z_{Fresnel} [Goodman, 2005; Southwell, 1981], but the expression presented here is used as reference.

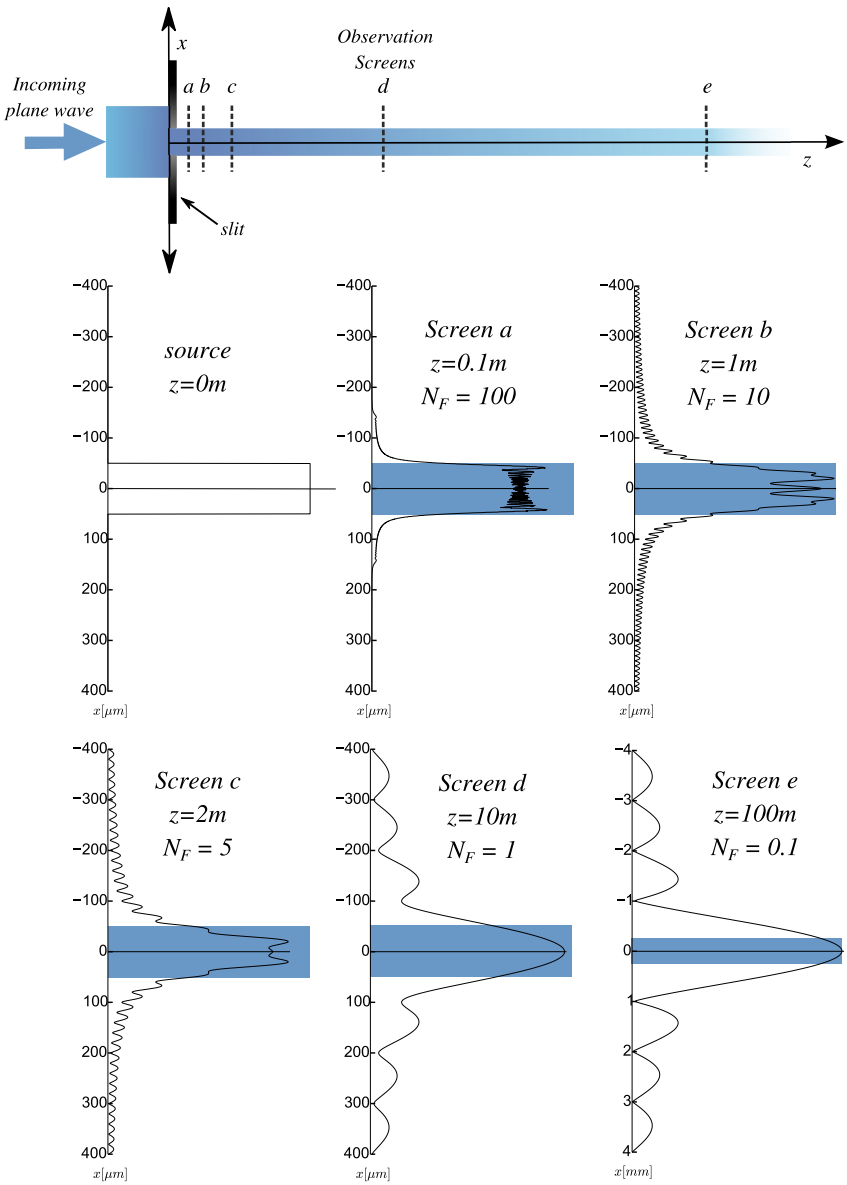
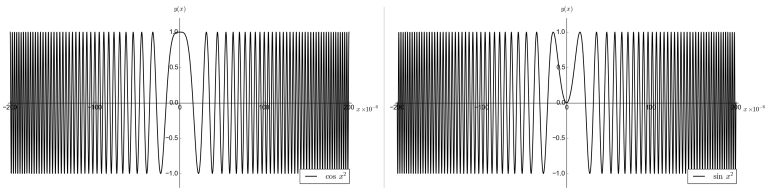


Figure 4.7. *Representation of near and far field diffraction for a plane wave diffracting in a rectangular aperture. (Top) The blue area represents the result predicted by ray optics, which is only the shadow of the aperture. (Middle and bottom rows) Beam intensity profile at the different observations screens, illustrating how the diffraction pattern changes as light propagates from the near to the far field. Values of Fresnel's number, N_F, are also noted. The profile at the screen e approaches th condition of Fraunhofer's diffraction and the profile resembles a Fourier transform of the aperture. Profiles are obtained by numerical calculation assuming aperture of 50μm and wavelength of λ=1nm.*



(a) (b)

Figure 4.8. *Plots of the real and imaginary part of e^{ikx^2} for $k = 2\pi/10^{-9}$. (a) $Re(e^{ikx^2}) = \cos kx^2$. (a) $Im(e^{ikx^2}) = \sin kx^2$.*

$$F(k) = \int_a^b f(x) e^{ikg(x)} dx, \tag{4.20}$$

where k, x, a, b, $f(x)$ and $g(x)$ are real.

The basic idea is that for large values of k the integrand oscillates so fast between positive and negative values that there is only minor contribution to the value of the integral. Figure 4.8 shows a plot of the real and imaginary part of the integrand for $f(x) = 1$ and $g(x) = x^2$, showing that after few oscillations the function oscillates very fast (related to the x coordinate) between -1 and 1. Figure 4.9 shows the numerical integration and the asymptotic value obtained by the SPA, and again, after few ($\approx$10) oscillations the value of the integral become very close to the asymptotic value.

The asymptotic solutions depend on the values of the functions $f(x)$ and $g(x)$ at the so called *critical points* of first and second kind. Critical points of the first kind, x_i^1, are the values of x where the first derivative of $g(x)$ is zero, that is $g'(x_i^1) = 0$. Critical points of the second kind are the limits of integration, a and b.

The condition of validity is that $f(x)$ is continuous and $g(x)$ is twice continuously differentiable in the interval $a < x < b$. In addition, the second derivative of $g(x)$ must be different to zero in the critical points of first kind, that is $g''(x_i^1) \neq 0$.

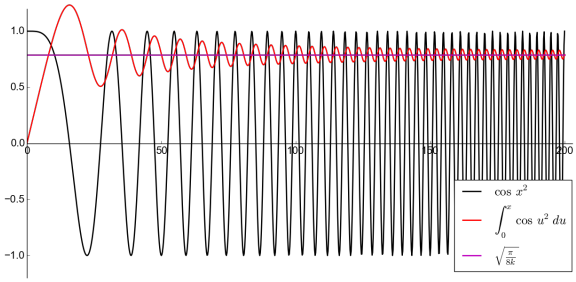
The contribution to the integral by the critical points of the first kind is given by [Mandel and Wolf, 1995]

$$F^{(1)}(k) \approx \sqrt{\frac{2\pi}{k}} \sum_1^n \frac{e^{\pm j\pi/4}}{\sqrt{|g''(x_i)|}} f(x_i) e^{ikg(x_i)}, \tag{4.21}$$

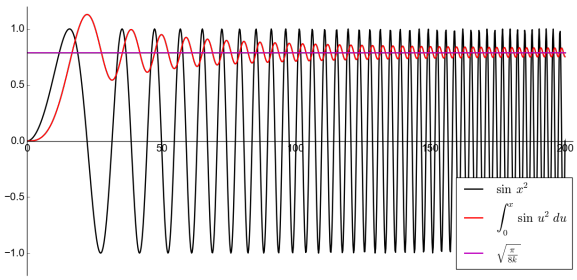
where the signal $\pm$ follows $g''(x_i) \gtrless 0$. The critical points of the second kind, in turn, have a contribution of the form:

$$F^{(2)}(k) \approx \frac{1}{ik} \left[\frac{f(b)}{g'(b)} e^{ikg(b)} - \frac{f(a)}{g'(a)} e^{ikg(a)} \right], \tag{4.22}$$

and the asymptotic solution of equation (4.20) is then given by



(a)



(b)

Figure 4.9. *(a) $Re(e^{ikx^2}) = \cos kx^2$. (a) $Im(e^{ikx^2}) = \sin kx^2$. The red curves are the numerical integration from 0 to x. Pink curves are the value of asymptotic solution obtained by the SPA.*

$$F(k) \approx F^{(1)}(k) + F^{(2)}(k). \tag{4.23}$$

It is important to note that $F^{(1)}(k)$ scales with $1/\sqrt{k}$ while $F^{(2)}(k) \propto 1/k$, and for this reason the contribution of $F^{(2)}(k)$ can be ignored for large values of k.

The application of SPA in the Kirchhoff integral requires a two dimensional treatment, which uses asymptotic solutions to solve the double integral given by

$$F(k) = \iint_D f(x,y) e^{ikg(x,y)} dx\, dy. \tag{4.24}$$

For the sake of simplicity the discussion about SPA is limited here to the one dimensional case. The properties of the two dimensional functions $f(x,y)$ and $g(x,y)$ are the same as for the functions $f(x)$ and $g(x)$ discussed above[12].

[12] Two dimensional case is discussed by Born and Wolf [1999, Appendix III.3] and Mandel and Wolf [1995, Ch. 3.3.3], where the respective asymptotic values are presented.

The SPA can then be applied to the Fresnel-Kirchoff integral of the form shown in equation (4.9). More precisely, SPA is applied to the integration over Σ (term inside curly brackets), where

$$f(\xi,\eta)=\frac{\Psi}{r_{21}\,r_{01}} \qquad \text{and} \qquad g(\xi,\eta)=r_{21}+r_{01}=PL(\xi,\eta), \tag{4.25}$$

and in the above equation PL stands for the path length.

The fact that $g(\xi,\eta)$ is the path length $PL(\xi,\eta)$ has an interesting physical interpretation: following the discussing above, we know from the SPA that the main points which contribute to the integral are the critical points of first kind ξ_i^1,η_i^1, where the first derivative of $g(\xi,\eta)$ is equal to zero. Substituting $g(\xi,\eta)$ by $PL(\xi,\eta)$ results that the main contribution to the Fresnel-Kirchoff integral are the points where

$$\left.\frac{\partial\, PL(\xi,\eta)}{\partial\,\xi}\right|_{\xi=\xi_i^1}=0 \qquad \text{and} \qquad \left.\frac{\partial\, PL(\xi,\eta)}{\partial\,\eta}\right|_{\eta=\eta_i^1}=0 \tag{4.26}$$

which is nothing else than the Fermat principle [Born and Wolf, 1999] (see also footnote on page 18). In other words, for a given point in the source plane and one point in the image plane, the main contribution to the field in the image plane is (only) due to the points in Σ that fulfill the Fermat principle. This provides a physical interpretation to the stationary phase approximation that is very useful when choosing the propagation planes and to understand the method for considering figure errors developed in Paper **V**.

Wave optics modeling - practical issues

When modeling X-ray beamlines by wave optics, two programs have been preferred to run computational simulations [Bowler, Bahrdt, and Chubar, 2008]: *Synchrotron Radiation Workshop* (SRW) [Chubar et al., 2011] and PHASE [Bahrdt, 2007b][Paper **V**]. It is not the goal here to compare the programs, but rather to discuss the methods and their accuracy. SRW makes use of thin lens approximations and of the first Rayleigh-Sommerfeld solution, which is restricted to planar geometry. PHASE is based on the SPA, which means that it can treat grazing incidence optics, and it uses a 2nd order expansion of the path length.

As we can see (and mentioned earlier) there is no existing method for rigorous calculation of diffraction and wave propagation, and any present method makes use of further approximations, in addition to the approximations already used to obtain Fresnel-Kirchoff and Rayleigh-Sommerfeld solutions.

For evaluating the calculation performance of the programs we follow the notation of Figure 4.6 and define a source plane as a mesh of $n_{x1}\times n_{y1}$ points, an image plane with $n_{x2}\times n_{y2}$ points, and an

optical device with $n_\xi \times n_\eta$ points. SRW propagates the field in two steps: first from the source to the optics, then from the optics to the image plane. SRW then profits from the performance of the FFT algorithm and therefore its calculation time is, following the big-O notation, equal to $O\left(n_{x1} \times n_{y1} \log(n_{x1} \times n_{y1})\right)$ (considering that there is no rescaling in the two steps propagations). PHASE on the other hand scans the plane Σ for each point in the image plane, which results $O(n_{x2} \times n_{y2} \times n_\xi \times n_\eta)$.

However in SRW the source and image plane have the same size and spatial resolution, and some re-scaling techniques are necessary. In the cases of high demagnification (when the beam size changes too much), this requires high spatial resolution and thus a large number of points. At the same time, PHASE can have a faster computational time by using a smaller number of points in the image plane, without loss of spatial resolution.

In case of PHASE, 2nd order expansion of the path length is a good approximation if the critical points are well separated. A proper choice of image and source plane ensure that the critical points are well spaced (see Paper **V**), but on the other hand this requires a two steps calculation (for instance, by first propagating from the source to an intermediary plane, and then from the intermediary to the image plane).

4.4 Statistical Optics and Coherence

Statistical optics deals with fluctuations in the optical (electromagnetic) field and their propagation trough an optical system [Wolf, 2007]. The fluctuations of the field may be completely random or, alternatively, described by some probabilistic distribution. They can be either intrinsic to the process of emission of radiation or due to instabilities of the source. For instance in case of thermal light (like light from lamps or from stars) the fluctuations are a consequence of the spontaneous emission of radiation by the atoms in the source. On the other hand, in the stimulated emission in lasers the atoms or molecules produce light in a ordered manner, with small(er) random fluctuations [Goodman, 2015]. Even when the fluctuations are random, the EM field still follows the wave equation and Maxwell's equations [Saleh and Teich, 2007], which makes possible to derive some properties of the light as a function of statistical properties of the source.

In general the fluctuations of the EM field are too fast to be perceived and directly detected. Real detectors make measurements in a finite area and in a finite time interval and in practice the final measurements are average measurements which are in fact also affected by the fluctuations of the field. Therefore one can say that the fluctuations are observable, but only under a proper statistical approach.

The statistical properties of the field also determine the response

of the optical systems, and eventually these properties can be exploited to achieve varying goals, for instance in coherent and in incoherent imaging techniques [Goodman, 2005].

Second Order Coherence

Coherence is a second order property of the EM field in the sense that it depends on the properties of two points in time or space[13]. The basic geometry for this kind of problem is shown in Figure 4.10. In the case where the individual points of the source emit radiation with random fluctuation we have that the radiation at the points P_1 and P_2 are also subject to random fluctuations. However, if the sources of the radiation at P_1 and P_2 are is same degree correlated (for instances if they emit in phase) then the fluctuations in the two points are in some way correlated. If the fields are completely correlated the light is said to be *coherent*, and the *degree of coherence* is a value that quantifies the correlation between the fields at these two points. In the case where the radiation at P_1 and P_2 is completely uncorrelated the light is said to be *incoherent*.

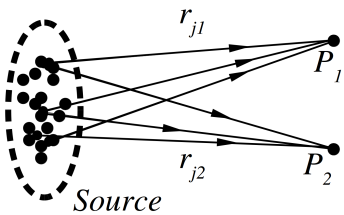


Figure 4.10. *Second order coherence studies the relation between two points P_1 and P_2. In this case the fields are generated from a common source.*

Therefore the study of coherence relies strongly in the correlation of the fields, and one of the basic quantities is the correlation of the field at two points in space, P_1 and P_2, and two points in time, t and $t+\tau$. Mathematically[14] we have [Wolf, 2007]

$$\Gamma(P_1, P_2, \tau) = \Gamma_{12}(\tau) = \langle V(P_1, t) V^*(P_2, t+\tau) \rangle , \tag{4.27}$$

where $\langle\ \rangle$ means time average defined as

$$\big\langle f(t) \big\rangle = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} f(t) dt . \tag{4.28}$$

The function $\Gamma_{12}(\tau)$ is the so-called *mutual coherence function*, a key concept in the study of second order coherence. We identify $\Gamma_{11}(0) = \sum_{m=1} V_{m1}^2 = I_1$, that is, $\Gamma_{11}(0)$ is equal to the sum of the fields V_{m1} at the point P_1, which results in the total intensity I_1. Similarly we have $\Gamma_{22}(0) = I_2$.

When we have that $P_1 = P_2$ and $\tau = 0$, the fields are the same and completely correlated, and the correlation $\Gamma_{12}(\tau)$ reaches its maximum value. As the points move apart from each other, a decrease in the value of $\Gamma_{12}(\tau)$ is expected due to the fluctuations of the field and eventually that $\Gamma_{12}(\tau)$ vanishes. For the distance $l_c = |P_1 - P_2|$ where $\Gamma_{12}(0)$ is equal to zero[15] we define as *coherence length*. The same applies for time: considering $\Gamma_{11}(\tau) = \Gamma(P_1, P_1, \tau)$ we have that $\Gamma_{11}(\tau_c) = 0$ for the *coherence time* τ_c.

[13] Previous discussions were part of what is labeled first order properties, that is, properties of a single point in space and time. Similarly, higher-order properties are related to properties at three or more points in time or space [Goodman, 2015; Mandel and Wolf, 1995; Wolf, 2007].

[14] The notation here closely follows [Wolf, 2007] and [Mandel and Wolf, 1995].

[15] Alternative definitions based on FWHM or RMS values are also used.

Another useful concept is the normalized value of $\Gamma_{12}(\tau)$ given by

$$\gamma_{12}(\tau) = \gamma(P_1, P_2, \tau) = \frac{\Gamma_{12}(\tau)}{\sqrt{\Gamma_{11}(0)}\sqrt{\Gamma_{22}(0)}}. \tag{4.29}$$

It can be shown that $0 \leqslant |\gamma_{12}(\tau)| \leqslant 1$ and, since it scales with $\Gamma_{12}(\tau)$, then $|\gamma_{12}(\tau)|$ is a quantitative value of the correlation of the fields. For this reason $\gamma_{12}(\tau)$ is referred as the *complex degree of coherence* and the modulus value $|\gamma_{12}(\tau)|$ is used as a quantitative description of the degree of coherence.

Coherence in the frequency Domain

For optical simulations it is useful to have these quantities in the frequency domain. By using the Fourier Transform of the mutual coherence function one obtains the *cross-spectral density function* [Wolf, 2007]:

$$W_{12}(\omega) = W(P_1, P_2, \omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \Gamma(P_1, P_2, \tau) e^{i\omega\tau} d\tau. \tag{4.30}$$

Similarly to equation (4.29) we have the normalized value of $W(P_1, P_2, \omega)$ as

$$\mu_{12}(\omega) = \mu(P_1, P_2, \omega) = \frac{W_{12}(\omega)}{\sqrt{I_1(\omega)}\sqrt{I_2(\omega)}}. \tag{4.31}$$

$I_j(\omega)$ is equal to $W_{jj}(\omega)$ and is the spectral intensity. The normalized value $\mu_{12}(\omega)$ is called *spectral degree of coherence* and it holds the same properties than $\gamma_{12}(\tau)$, which have been discussed previously. Therefore the modulus $|\mu_{12}(\omega)|$ also describes the degree of coherence [Wolf, 2007], and the use of either $\gamma_{12}(\tau)$ or $\mu_{12}(\omega)$ depends on which domain such properties are being discussed.

One important property of the mutual coherence function and the cross-spectral density function is that in free space they obey the wave equation [Wolf, 1955, 2007]:

$$\nabla_1^2 \Gamma_{12}(\tau) = \frac{1}{c^2}\frac{\partial^2 \Gamma_{12}(\tau)}{\partial \tau^2}, \qquad \nabla_2^2 \Gamma_{12}(\tau) = \frac{1}{c^2}\frac{\partial^2 \Gamma_{12}(\tau)}{\partial \tau^2}, \tag{4.32}$$

$$\nabla_1^2 W_{12}(\omega) = -\frac{\omega^2}{c^2} W_{12}(\omega) \quad \text{and} \quad \nabla_2^2 W_{12}(\omega) = -\frac{\omega^2}{c^2} W_{12}(\omega). \tag{4.33}$$

where the operators ∇_1^2 and ∇_2^2 are the Laplacian with respect to the points P_1 and P_2.

This means that the mutual coherence propagates similarly to a plane wave, reducing a problem in statistical optics into a problem of partial differential equations.

This also offers some insights about how $\Gamma_{12}(\tau)$ or $W_{12}(\omega)$ propagate in space. For instance, we know that in a divergent light beam

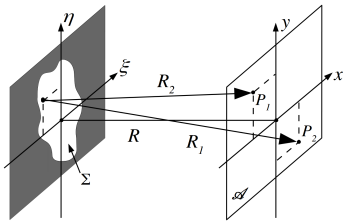


Figure 4.11. *Geometry for the VCZ theorem.*

two points in the wavefront move apart as the beam propagates. By analogies we can expect that with certain conditions the coherence length will also increase as the beam propagates. This will be described under the Gauss Schell approximation discussed later, and it is the base for deriving the van Cittert–Zernike theorem discussed next.

van Cittert–Zernike theorem

Equations 4.32 and 4.33 can be further developed to describe the propagation of the correlations in free space (as done for instance in [Mandel and Wolf, 1995, Ch. 4.4]). Such results are used to obtain the propagated function $\Gamma_{12}(0)$ of a completely incoherent source. For this case the source is assumed to have mutual coherence function of the form:

$$\Gamma'_{12}(0) = I(P_1)\delta(P_2 - P_1), \tag{4.34}$$

where the prime symbol means that the quantities are at the source plane Σ (see Figure 4.11). It can actually be shown that for a propagating wave the coherence length need to be at least one wavelength long in any direction [Goodman, 2015, Ch. 5.5.2]. This means that completely incoherent beam is actually a limiting case and that propagating light has some degree of coherence at least in a very short range.

For an incoherent source as described in equation (4.34), the propagated mutual coherence function $\Gamma_{12}(0)$ results in [Mandel and Wolf, 1995, Eq. 4.4.34] [Mandel and Wolf, 1995, Eq. 3.2.11]

$$\gamma_{12}(0) = \frac{1}{\sqrt{I(P_1)}\sqrt{I(P_2)}} \int_\sigma I(S) \frac{e^{ik(R_{2S} - R_{1S})}}{R_{2S} R_{1S}} dS \tag{4.35}$$

and

$$I(P_j) = \Gamma'_{12}(0) = \int_\sigma \frac{I(S)}{R_{jS}} dS, \tag{4.36}$$

where R_{1S} and R_{2S} are the distances from points at the source to the points P_1 and P_2, respectively. k is the wavenumber equal to $2\pi/\lambda$, λ being the wavelength.

Equation 4.35 is the mathematical form of the *van Cittert–Zernike theorem* (VCZ) [Mandel and Wolf, 1995; Wolf, 2007]. Applying far field approximations, the VCZ theorem becomes [Wolf, 2007, Eq. 3.2.20]

$$|\gamma_{12}(0)| = \left| \frac{\int\int_\sigma I(\xi,\eta) e^{-ik(p\xi + q\eta)} d\xi\, d\eta}{\int\int_\sigma I(\xi,\eta) d\xi\, d\eta} \right|. \tag{4.37}$$

where

$$p = \frac{X_2 - X_1}{R}, \quad \text{and} \quad q = \frac{Y_2 - Y_1}{R}. \tag{4.38}$$

Equation 4.37 is the far field form of VCZ[16]. Note that the right hand part can be identified as a (normalized) Fourier transform, and thus the intensity distribution in the source $I(\xi,\eta)$ and the mutual coherence function are related though a Fourier transform. From the properties of the Fourier transform we know that two intervals $\Delta\xi$ and Δp are reciprocal and therefore inversely proportional as [Goodman, 2015, Ch. 5.7.2]

$$k\Delta p \propto \frac{1}{\Delta\xi}. \tag{4.39}$$

If we identify now $\Delta\xi$ as source size σ and Δp as l_c/R, where l_c is the coherence length, we obtain that

$$l_c \propto \frac{\lambda R}{\sigma}, \tag{4.40}$$

which is the result shown without derivation as equation (4.4) in the beginning of this chapter.

To illustrate the VCZ theorem we consider the radiation generated by a circular and completely incoherent source of radius a at a distance R (see Figure 4.12). The degree of coherence is then given by [Wolf, 2007]

$$|\gamma_{12}(0)| = \left| \frac{2J_1(\frac{kad}{R})}{\frac{kad}{R}} \right|, \tag{4.41}$$

where J_1 is the Bessel function of the first kind and of the first order [Arfken et al., 2012], and $d = \sqrt{(X_2-X_1)^2+(Y_2-Y_1)^2}$ is the distance between the two points in the observation plane $\mathscr{A}$. Figure 4.12 shows the behavior of $|\gamma_{12}(0)|$ for P_1 at the optical axis.

If we define now the coherence length l_c as the distance d when $|\gamma_{12}(0)|$ is zero, we obtain that for the current case the coherence length is

$$l_c = \frac{0.61R\lambda}{a}, \tag{4.42}$$

which resembles equations 4.4 and 4.40, but now it has been derived in a more formal manner.

Equation 4.42 can be used to estimate the coherence length of a propagated beam originating from an undulator. Figure 4.13 shows this result for MAX II and MAX IV storage rings. This is of course an estimation since undulators radiation is very different of the radiation assumed here (circular aperture illuminated by a planar wave). In Figure 4.13.b it is shown how the coherence lengths compare to the beam size by plotting the ratios q_x and q_y defined as

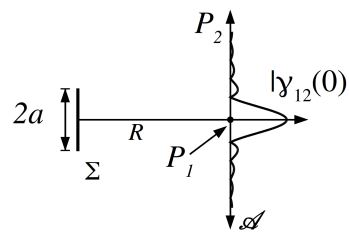
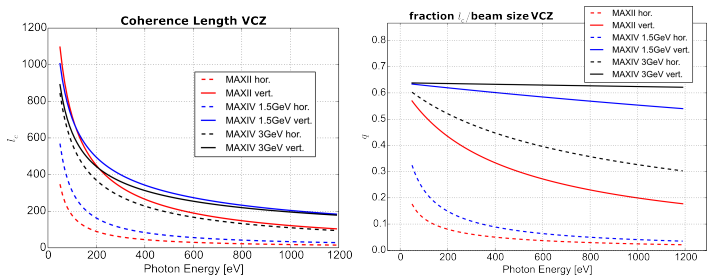


Figure 4.12. *Example of application of VCZ theorem, where in the observation plan $\mathscr{A}$ is ploted the value of $|\gamma_{12}(0)|$ resulting from a circular incoherent source of radius a.*

[16] Besides the name far field, the far field approximation here is less restrictive than the Fraunhofer approximation and it is valid also in the Fresnel regime [Goodman, 2015, Ch. 5.7.2].



(a) (b)

Figure 4.13. *Values of the ratio q obtained by the VCZ theory. Results for MAX II, MAX IV 1.5GeV and MAX IV 3GeV rings. Undulator lengths are assumed to be the maximum length of the straight section.*

$$q_x = \frac{l_{cx}}{\sigma_x} \quad \text{and} \quad q_y = \frac{l_{cy}}{\sigma_y}. \tag{4.43}$$

These results shows that at low photon energies MAX II is fairly coherent in the vertical direction, but the small horizontal coherence length results that only a small fraction of the beam is inside the coherence area of $A_c = l_{cx} \times l_{cy}$ (see also Figures 4.1 and 4.2). When comparing the curves we see that MAX IV 3GeV ring has much larger fraction of the beam inside the coherence area, and that there is a big improvement in the horizontal coherence length compared to MAX II ring.

It must be noted that in fact a generalized van Cittert–Zernike theorem can be derived for partially coherent light by using a mutual coherence functions different from equation (4.34) [Goodman, 2015, Ch. 5.8]. Goodman [2015, eq. 5.8.10] for instance derives the analytical solution for a Schell-model beam (to be discussed next). However other kind of analytical solutions are complex and when applied to undulators it results in complicated solutions. Geloni et al. [2005] makes a formal derivation of the coherent properties of undulators and it discusses that van Cittert–Zernike produces satisfactory results in the far field, despite the fact that undulator radiation do not satisfy the assumptions of the theorem. However, deviations are expected between the experimental values and the ones predicted by van Cittert-Zernike theorem, and for realistic comparison results from Geloni et al. [2005] must be considered.

Gaussian Schell-model Field

Schell-model fields are characterized by a spectral degree of coherence $\mu(P_1, P_2, \omega)$ that depends only on the distance between the points P_1 and P_2, and therefore $\mu(P_1, P_2, \omega) = \mu(P_2 - P_1, \omega)$ [Mandel

and Wolf, 1995; Wolf, 2007]. Using equation (4.31) we obtain that the cross-spectral density of a Schell-model field is given by

$$W_{12} = \sqrt{I(P_1)}\sqrt{I(P_2)}\mu(P_2 - P_1), \tag{4.44}$$

where the dependence with ω is omitted.

A Gaussian Schell-model (GSM) beam is obtained by also assuming that the spatial profile of both $I(P)$ and $\mu(P_2 - P_1)$ are described by two dimensional Gaussian functions:

$$I(P) = \exp\left(-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}\right), \tag{4.45}$$

and

$$\mu(P_2 - P_1) = \exp\left(-\frac{(x_2 - x_1)^2}{2\sigma_{cx}^2} - \frac{(x_2 - x_1)^2}{2\sigma_{cy}^2}\right), \tag{4.46}$$

where σ_x and σ_y are the beam sizes in x and y directions. The quantities σ_{cx} and σ_{cy} are the rms (root mean square) values of the spectral degree of coherence $\mu(P_2 - P_1)$, and thus in the GSM context they are defined as the coherence lengths.

The use of GSM beam simplifies the calculation of the propagated beam and makes possible to correlate the coherence length and beam size of the source and of the propagated beam. The mathematical result is omitted here (see for instance [Mandel and Wolf, 1995, eq. 5.6-95]), but one important result of the GSM model is that the ratios q_x and q_y are constant as the beam propagates. If we use these ratios as measurements of the degree of coherence of the beam, then we have that the degree of coherence is invariant in propagation [Mandel and Wolf, 1995].

We can use the results of the GSM theory in order to estimate the coherence properties of the beam in a storage ring (as done for instance by [Vartanyants and Singer, 2010]). The ratio q_{xy} can be written in function of the electron beam emittance ε and it is given by

$$q_{xy} = \frac{2}{\sqrt{4k^2\varepsilon_{xy}^2 - 1}}, \tag{4.47}$$

that is, in the GSM model the ratio q_{xy} depends only on the electron beam emittance and wavelength of the radiation. Moreover, it reaches its maximum value $q_{xy} \to \infty$ when $4k^2\varepsilon_{xy}^2 \to 1$, which is exactly the diffraction limit condition of equation (4.3).

Similarly to van Cittert–Zernike, Gaussian Schell-model beams have limited application for undulator radiation, the most obvious problem being due to the different expressions for diffraction limit of Gaussian beam and undulator radiation. Hence, once more this results are used to study general properties and to provide some insights about the physics of partially coherent beam. Nevertheless,

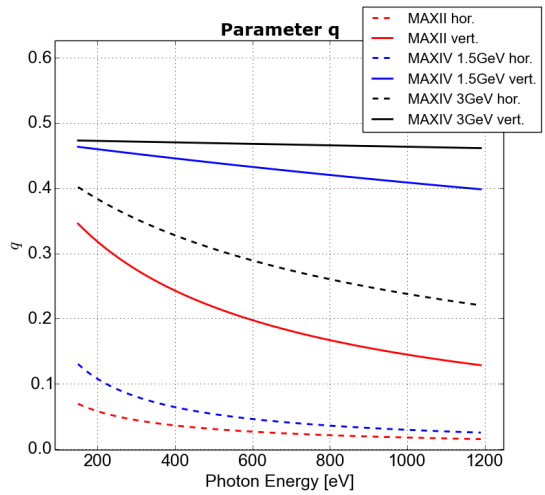


Figure 4.14. *Values of the ratio q obtained when using Gaussian Schell-model to describe undulator radiation. Results for MAX II, MAX IV 1.5GeV and MAX IV 3GeV rings. Undulator lengths are assumed to be the maximum length of the straight section.*

under certain conditions it can produce satisfactory results that are in agreement with results obtained with SRW [Vartanyants and Singer, 2010] and with analytical results obtained for undulator, as the ones derived by Geloni et al. [2005].

Discussion

This final discussion focuses now in how the codes for modeling beamline optics make use of wave and statistical optics.

As discussed before, PHASE uses the stationary phase approximation to calculate the Fresnel-Kirchoff integral and for this reason it can treat grazing incidence optics. The code was limited to a second order expansion of the path length but this expansion, combined with the method for considering surface errors, can produce accurate results. This is discussed in Paper **V** and in special it is quantified by equation (25) in the same paper.

SRW on the other hand makes use of thin lens approximation, which can be used only as rough approximation for grazing incidence optics. This restricts the use of the code for many situations but it can produce satisfactory results for some geometries, for instance for (over illuminated) zone plates. In addition, SRW has integrated in the same program simulations of real sources, *i.e.*, undulators. It can also perform simulations for partially coherent light, where the radiation generated by many electrons is propagated individually as coherent light and then the radiation is summed up inco-

herently (that is, sum of intensities rather than the electric field). At present this is the only method available to estimate the coherence length in the experimental station (apart from analytical approximative studies). In principle PHASE is also suitable for this kind of calculations, but as today, it does not have this integrated into the code in a easy way.

With this, SRW was the program of choice when performing initial studies for SoftiMAX beamline, as discussed in Paper **VI**. By using SRW it was possible to evaluate the degree of coherence at the coherent imaging end station. The degree of coherence can be quantified in several ways, one of them being the spectral degree of coherence, described in equation (4.31). SRW can calculate this, as done for instance in [Laundy et al., 2013] and [Laundy et al., 2014]. However the results are difficult to interpret and, more importantly, difficult to compare with experimental results. For this reason, it is quite common to evaluate the degree of coherence by simulating a Young's double slit experiment, which can actually be used experimentally to quantify the degree of coherence. This method has been used by many authors[Chubar et al., 2010, 2011; Fluerasu et al., 2011; Shapiro et al., 2011], which means that the simulation of double slit experiment provides a good way to compare performance of different beamlines.

The same approach was used for the coherent X-ray imaging branch of the beamline SoftiMAX. The results of such simulations are shown in Figure 4.15 and they are discussed in more details in Paper **VI**.

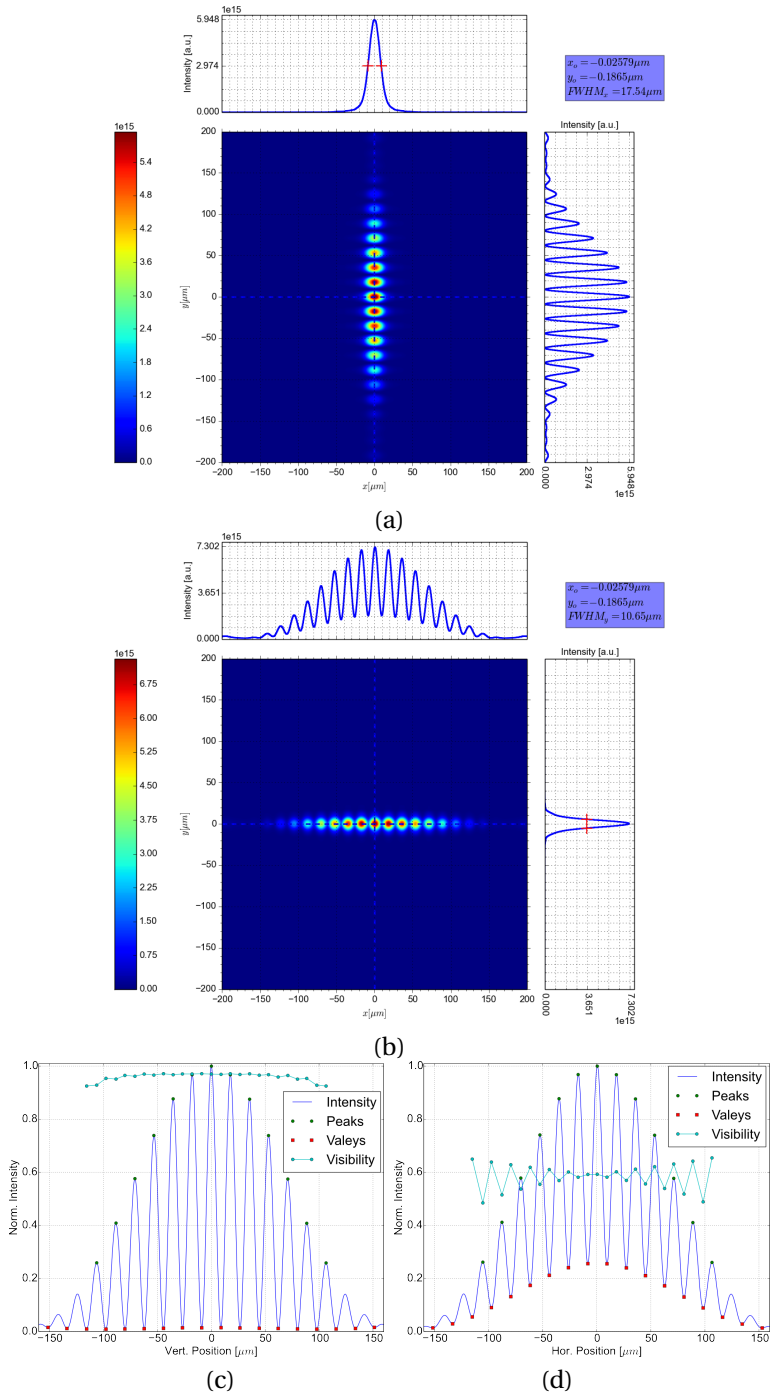


Figure 4.15. *Results of simulations of the double slit experiment for the CXI branch of SoftiMAX beamline [Paper **VI**]. (a) and (c) Intensities measured on the observation screen. (b) and (d) Center profile showing the visibility measured at each peak. Double slit consists of two 1μm rectangular slits spaced by 10μm (measured from the center of the apertures). The observation screen is placed 100 mm from the slits. Photon energy is 700e V .*

CHAPTER 5

CONCLUSION

In this work the concepts of X-rays optics and synchrotron radiation emission are used to develop and study optical devices for synchrotron radiation beamlines. The projects described here and in the papers include many areas of optics, namely geometric optics, X-ray optics, ray tracing simulations, polarization of the light, Stokes parameters, Mueller matrices, optical properties of metals, wave optics, diffraction and statistical optics. Though not being an exhaustive description, this work aims to give a general view of optics applied to synchrotron radiation and how it distinguishes from other results in optics, mainly those for visible light and paraxial optics. The contributions of this work can be summarized as follows: a new concept for refocusing system of cPGM beamlines has been developed, which is already in use at SPECIES beamline and it will be used at forthcoming HIPPIE beamline at MAX IV; a soft X-ray polarimeters have been commissioned, the commissioning of the VUV polarimeter is in progress and a novel method for data analysis for X-ray polarimetry has been proposed; a method for considering surface errors of optics elements has been proposed and included in the PHASE code for wave optics simulations.

More specifically, Chapter 2 together with Paper **I** give a detailed description of a cPGM beamline with an astigmatic focus, which adds two additional flexibilities to cPGM beamlines: possibility to control the vertical beam size through c_{ff} and independence of the image size from the slit opening size. This makes possible to adjust the beam size to optimize geometric requirements (*e.g.* geometric detector efficiency) and freely adjust the exit slit aperture considering only energy resolution requirements. It must be noted that the astigmatic focus can be implemented on existing beamlines by displacing and/or changing the angle of incidence of the refocusing mirrors. It can easily be done in refocusing system based on cylindrical or toroidal mirrors. The cPGM with astigmatic focus concept is already in use at SPECIES beamline at MAX II (later on the beam-

line will be transferred to MAX IV) and a description of the beamline performance is done in Paper **II**. This concept will also be used for HIPPIE beamline at MAX IV and it is expected that other beamlines take advantage of this concept in the near future.

Chapter 3 described polarimetry for the energy range of vacuum ultraviolet. The particular case of a polarimeter based on metal coated mirrors, as the one being commissioned at MAX IV laboratory, is discussed. The theoretical background covered the topics of reflection on metal surfaces, the Stokes parameters and the Muller Matrices. This basic information made possible to derive the polarimeter equation (eq. 3.36 in Ch. 3.4) for the MAX IV polarimeter and to estimate the performance for different polarization states (Fig. 3.14). Similar calculations were then developed for a multilayer based polarimeter, providing information about performance and eventually being part of the data analysis, as described in Paper **IV**. In addition, a new method for data analysis is developed in Paper **III**, which directly links the Stokes parameters to the experimentally observed quantities and as a consequence it does not rely in any fitting procedure. This makes possible to determine the experimental error of the Stokes parameters in the measurements in a straight forward manner.

Chapter 4 review the concepts of wave and statistical optics and discuss their application for synchrotron radiation beamlines. Special attention was dedicated to diffraction and to coherence in diffraction limited storage rings. Distinction was made between the methods used by SRW and PHASE, the most widely used computational programs for wave optics simulations of beamlines. The understanding of the physics behind PHASE made possible to propose and implement a method for considering the surface errors of optical elements, as described in Paper **V**. This was previously missing in PHASE and in the paper it is discussed how the method provides accurate results. Finally, SRW was used to study the degree of coherence at the experimental station of SoftiMAX beamline; that is included in Paper **VI**.

APPENDIX A

Stokes parameters and partially polarized light

This appendix gathers some useful relations of the Stokes parameters that are used in this work, mainly in Chapter 3. The goal is to use Stokes parameters to fully describe the state of polarization of the light as well as the fraction of linear, circular and unpolarized light.

Considering a polarized electromagnetic wave propagating in $\hat{z}$ direction (coordinate system of Figure 3.8), the two orthogonal components $E_x(t)$ and $E_y(t)$ of the electric field $\vec{E}(t)$ are given by

$$\begin{aligned} E_x(t) &= E_{x0}\exp[i(kz-\omega t+\delta_x)],\\ E_y(t) &= E_{z0}\exp[i(kz-\omega t+\delta_y)],\\ \vec{E}(t) &= E_x(t)\hat{x}+E_y(t)\hat{y}. \end{aligned}$$

E_{x0} and E_{y0} are the corresponding amplitudes of the fields, k is the wave number, ω is the angular frequency whereas δ_x and δ_y are the phases. The Stokes parameters are given by

$$\begin{aligned} S_0 &= E_{x0}^2+E_{y0}^2,\\ S_1 &= E_{x0}^2-E_{y0}^2,\\ S_2 &= 2E_{x0}E_{y0}\cos(\delta_z-\delta_y), \qquad \text{Polarized light}\\ S_3 &= 2E_{x0}E_{y0}\sin(\delta_z-\delta_y),\\ S_0^2 &= S_1^2+S_2^2+S_3^2. \end{aligned} \tag{A.1}$$

These definitions apply for a completely polarized light. It has been shown experimentally that the Stokes parameters for completely unpolarized light are[1]

$$\begin{aligned} S_0 &> 0,\\ S_1 = S_2 = S_3 &= 0. \qquad \text{Unpolarized light} \end{aligned} \tag{A.2}$$

It can also be shown that the Stokes parameters of two completely independent beams can be summed to yield the Stokes parameters of

[1] The experimental motivation is discussed in Goldstein [2011, Chapter 5.5].

the combined beam [Goldstein, 2011]. Therefore, the Stokes parameters of partially polarized light can be described as a combination of the Stokes parameters of a polarized and an unpolarized beam.

By using the Stokes vector notation (Ch. 3.3), the vectors for completely polarized $\mathbf{S}^{(p)}$ and unpolarized light $\mathbf{S}^{(u)}$ are given by

$$\mathbf{S}^{(p)} = \begin{pmatrix} P \cdot S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} \quad \text{and} \quad \mathbf{S}^{(u)} = \begin{pmatrix} P_u \cdot S_0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{A.3}$$

where P and $P_u = (1-P)$ are the fractions of polarized and unpolarized light of the beam. The sum is simply

$$\mathbf{S} = \mathbf{S}^{(p)} + \mathbf{S}^{(u)} = \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix}. \quad \begin{matrix}\text{Partially} \\ \text{polarized} \\ \text{light}\end{matrix} \tag{A.4}$$

It must be noted that the parameter S_0 is now the total intensity, *i.e.* the sum of the polarized and unpolarized light, and hence is no longer defined by equation (A.1). For the partially polarized light this parameter must be redefined, and the four Stokes parameters become

$$\boxed{\begin{aligned} P \cdot S_0 &= E_{x0}^2 + E_{z0}^2, \\ S_1 &= E_{x0}^2 - E_{z0}^2, \\ S_2 &= 2E_{x0}E_{z0}\cos(\delta_z - \delta_x), \\ S_3 &= 2E_{x0}E_{z0}\sin(\delta_z - \delta_x), \\ S_0^2 &\geqslant S_1^2 + S_2^2 + S_3^2, \\ P^2 \cdot S_0^2 &= S_1^2 + S_2^2 + S_3^2. \end{aligned}} \quad \begin{matrix}\text{Partially} \\ \text{polarized} \\ \text{light}\end{matrix} \tag{A.5}$$

Some further relations are also useful: the parameters S_1 and S_2 are related to the amount of linearly polarized light while the parameter S_3 is related to the amount of circularly polarized radiation. With this respect, the degreed of linearly (P_L) and circularly (P_C) polarized light are defined as

$$P_L = \sqrt{S_1^2 + S_2^2} \quad \text{and} \quad P_C = S_3,$$

which, using equation (A.5), results in

$$P^2 \cdot S_0^2 = P_L^2 + P_C^2. \tag{A.6}$$

Normalizing this by the total intensity S_0 yields

$$
\begin{aligned}
P^2 &= \bar{P}_L^2 + \bar{P}_C^2, \\
\bar{P}_L &= \frac{\sqrt{S_1^2 + S_2^2}}{S_0} = \sqrt{\bar{S}_1^2 + \bar{S}_2^2} \\
\bar{P}_C &= \frac{S_3}{S_0} = \bar{S}_3
\end{aligned}
$$

where the parameters with the bar are normalized by S_0. By using the normalized Stokes parameters it is not necessary to determine the parameter S_0, requiring an absolute measurement (*i.e.*, requiring the use of a detector with known efficiency).

Finally, to be able to describe the polarization ellipse [Goldstein, 2011] of the polarized fraction of the light, it is necessary to normalize the Stokes parameters by P, resulting in

$$
\begin{aligned}
1 &= \frac{S_1^2 + S_2^2 + S_3^2}{P^2 \cdot S_0^2}, \\
1 &= \bar{s}_1^2 + \bar{s}_2^2 + \bar{s}_3^2, \\
1 &= \bar{p}_L^2 + \bar{p}_C^2,
\end{aligned}
$$

where the parameters normalized by P are written in lowercase.

APPENDIX B

Mueller Matrix applied for the three mirror unit

This appendix shows that the three mirrors of the retarder unit of the polarimeter can be treated as a single mirror, with effective optical parameters dependent on the three original mirrors.

Defining the Mueller matrix of the mirror i as $\mathbf{M_i}$, the Mueller matrix of the retarder unit $\mathbf{M_{ret}}$ is given by:

$$\begin{aligned}\mathbf{M_{ret}}(\eta_1) = \quad & \mathbf{M_R}(-\eta_1)\cdot\mathbf{M_3}\cdot\mathbf{M_R}(\eta_1)\cdot \\ \cdot\mathbf{M_R}(-180^o)\cdot & \mathbf{M_R}(-\eta_1)\cdot\mathbf{M_2}\cdot\mathbf{M_R}(\eta_1)\cdot\mathbf{M_R}(180)\cdot \\ & \cdot\mathbf{M_R}(-\eta_1)\cdot\mathbf{M_1}\cdot\mathbf{M_R}(\eta_1)\end{aligned}$$

where $\mathbf{M_R}$ is the Mueller matrix of rotation given by equation (3.29). It is also assumed that the mirror 2 is rotated by 180deg related to the first mirror, as discussed in Chapter 3.3 (see Figure 3.11).

Calculating the rotation matrices for the angles 180deg and -180deg results $\mathbf{M_R}(180) = \mathbf{M_R}(-180^o) = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix. Additionally, we have that $\mathbf{M_R}(-\eta)\cdot\mathbf{M_R}(\eta) = \mathbf{I}$. These two results simplify the former equation to

$$\begin{aligned}\mathbf{M_{ret}}(\eta_1) &= \mathbf{M_R}(-\eta_1)\cdot\mathbf{M_3}\cdot\mathbf{M_2}\cdot\mathbf{M_1}\cdot\mathbf{M_R}(\eta_1), \\ &= \mathbf{M_R}(-\eta_1)\cdot\mathbf{M_{ret}}\cdot\mathbf{M_R}(\eta_1),\end{aligned}$$

where $\mathbf{M_{ret}} = \mathbf{M_3}\cdot\mathbf{M_2}\cdot\mathbf{M_1}$.

We must recall that all the mirrors are of the same material (gold) and that the light has the same incidence angle θ on mirrors 1 and 3. That means that mirrors 1 and 3 have the same values of reflectivity and phase shift, and as a result $\mathbf{M_1} = \mathbf{M_3}$. Therefore we can write

$$\mathbf{M_{ret}} = \mathbf{M_1}(r_{s1}, r_{p1}, \delta_1)\cdot\mathbf{M_2}(r_{s2}, r_{p2}, \delta_2)\cdot\mathbf{M_1}(r_{s1}, r_{p1}, \delta_1),$$

resulting

$$\mathbf{M_{ret}} = \begin{pmatrix} m_{11} & m_{12} & 0 & 0 \\ m_{12} & m_{11} & 0 & 0 \\ 0 & 0 & m_{33} & m_{34} \\ 0 & 0 & -m_{34} & m_{33} \end{pmatrix}.$$

The non-zero matrix elements above are given by

$$\begin{aligned} m_{11} &= \frac{{r_{s1}}^4 {r_{s2}}^2}{2} + \frac{{r_{p1}}^4 {r_{p2}}^2}{2}, \\ m_{12} &= \frac{{r_{s1}}^4 {r_{s2}}^2}{2} - \frac{{r_{p1}}^4 {r_{p2}}^2}{2}, \\ m_{33} &= {r_{s1}}^2 r_{s2} {r_{p1}}^2 r_{p2} \cos(2\delta_1 + \delta_2) \quad \text{and} \\ m_{34} &= {r_{s1}}^2 r_{s2} {r_{p1}}^2 r_{p2} \sin(2\delta_1 + \delta_2). \end{aligned}$$

Then, by defining

$$\begin{aligned} r_{s3} &= r_{s1}^2 \cdot r_{s2}, \\ r_{p3} &= r_{p1}^2 \cdot r_{p2} \quad \text{and} \\ \delta_3 &= 2\delta_1 + \delta_2 \end{aligned}$$

we can simplify the former matrix to

$$\mathbf{M_{ret}} = \frac{1}{2} \times \begin{pmatrix} {r_{s3}}^2 + {r_{p3}}^2 & {r_{s3}}^2 - {r_{p3}}^2 & 0 & 0 \\ {r_{s3}}^2 - {r_{p3}}^2 & {r_{s3}}^2 + {r_{p3}}^2 & 0 & 0 \\ 0 & 0 & 2 r_{s3} r_{p3} \cos(\delta_3) & 2 r_{s3} r_{p3} \sin(\delta_3) \\ 0 & 0 & -2 r_{s3} r_{p3} \sin(\delta_3) & 2 r_{s3} r_{p3} \cos(\delta_3) \end{pmatrix}.$$

This is exactly the same form as the matrix for a single mirror (see eq. 3.28). It means that the three mirrors of the retarder can be considered as a single mirror with effective reflectivity parameters and phase shift r_{s3}, r_{p3} and δ_3 as given above.

APPENDIX C

Special Cases for Polarimetry

The so-called special cases are discussed in Chapter 3 and Paper **III**. They are special conditions for measuring the Stokes parameters with the polarimeter. To derive the functions used in these special cases it is necessary to perform some repetitive algebraic calculation and for the sake of clarity these calculation are done in this Appendix.

C.1 Useful identity

We will soon see that with certain conditions the polarimeter equation 3.36 reduces to

$$I(\eta) = \alpha\big(B + a \cdot \cos(2\eta) + b \cdot \sin(2\eta)\big). \tag{C.1}$$

A particular trigonometric identity is extremely useful in simplifying this equation. We start by noting that the function $A \cdot \sin(2\eta + \Psi)$ can be expanded as

$$A \cdot \sin(2\eta + \Psi) = A\sin(\Psi)\cos(2\eta) + A\cos(\Psi)\sin(2\eta). \tag{C.2}$$

Then, by defining

$$A\sin(\Psi) = a, \qquad A\cos(\Psi) = b, \tag{C.3a}$$

$$A^2 = a^2 + b^2, \qquad \text{and} \qquad \tan(\Psi) = \frac{a}{b} \tag{C.3b}$$

we can rewrite equation (C.2) as

$$A \cdot \sin(2\eta + \Psi) = a \cdot \cos(2\eta) + b \cdot \sin(2\eta). \tag{C.4}$$

Using this identity we can write $I(\eta)$ in equation (C.1) as

$$I(\eta) = \alpha\big(B + A \cdot \sin(2\eta + \Psi)\big). \tag{C.5}$$

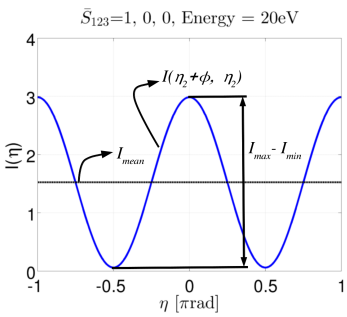


Figure C.1. *Example of a (simulated) experimental curve $I(\eta)$ where are shown the experimental parameters used on the special cases.*

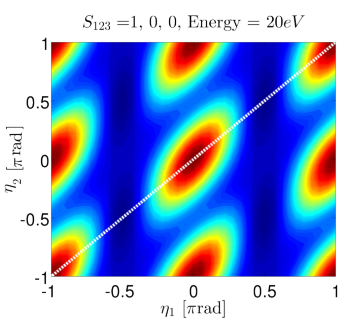


Figure C.2. *Simulated values of $I(\eta_1,\eta_2)$ for horizontally polarized light, where the white line represents the geometric interpretation of special case 1: on this special case the intensity is measured over the line $\eta_1=\eta_2$.*

Considering the application of this identity to the polarimeter, we can show that by fixing the angular difference $\eta_2-\eta_1$ to some specific values the polarimeter equation (eq. 3.36) can be reduced to the form of equation (C.1), and thus to the form of equation (C.5). The advantage of equation (C.5) is that it can be easily written in terms of experimental, measurable values. By writing

$$\bar{I}=\alpha B \qquad \text{and} \qquad V=A/B, \tag{C.6}$$

we obtain

$$I(\eta)=\bar{I}\big(1+V\sin(2\eta+\Psi)\big). \tag{C.7}$$

From an experimental point of view, the function $I(\eta)$ in equation (C.7) can be interpreted either as light intensity, or electric current at the detector, depending on the definition of the parameter α. The parameter $\bar{I}$ is the mean intensity (or current) and is defined as (see Figure C.1)

$$\bar{I}=\frac{1}{2\pi}\int_0^{2\pi} I(\eta)\cdot d\eta=\frac{1}{N}\sum_1^{i=N} I(\eta_i), \tag{C.8}$$

where the last identity corresponds to the case where N discrete measurements $I(\eta)$ are performed at angles $\eta=\eta_i$, $i=(1,2,..,N)$ and $0\le\eta_i\le 2\pi$. The parameter V is called visibility[2] of the resulting data and can be written as a function of the experimental values for the maximum and minimum values of $I(\eta)$:

$$V=\frac{(I_{max}-I_{min})}{2\bar{I}}. \tag{C.9}$$

Next, we will demonstrate how the Stokes parameters can be determined from the mean current $\bar{I}$, the visibility V and phase Ψ measured at different conditions.

C.2 Special cases

As explained above, choosing the angles η_1 and η_2 properly allows to simplify the polarimeter equation (eq. 3.36) to the form of equation (C.5). Consequently, both the data analysis and determination of the polarization become simpler. Common to all these choices is that the angle between the retarder and analyzer are kept constant, although the absolute values change from case to case. These conditions reduce $I(\eta_1,\eta_2)$ from a two variables function to one variable function $I(\eta)$. The geometric interpretation is shown in Figure C.2,

[2] Named visibility due to the similarity with the equation for fringe visibility used on the study of wave interference, as given for instance in Eq. 3.19 of [Fowles, 1975].

where instead of mapping the function $I(\eta_1, \eta_2)$ by measuring at all the possible angle pairs (η_1, η_2), the measurement is only done for example along the line $\eta_2 = \eta_1$.

To completely determine the polarization of the light four special cases are necessary:

$$\begin{aligned} &\text{case 1:} \quad \eta_2 = \eta_1, \\ &\text{case 2:} \quad \eta_2 = \eta_1 + \pi/2, \\ &\text{case 3:} \quad \eta_2 = \eta_1 + \pi/4, \\ &\text{case 4:} \quad \eta_2 = \eta_1 - \pi/4. \end{aligned}$$

In general we have that the special cases have an angular difference $\eta_2 - \eta_1 = m\pi/4$ for $m = -1, 0, 1, 2$. The geometric interpretation of the four cases are shown in Figure C.3 for the four values of m at different polarizations: linearly (Figure C.3.a) and circularly polarized light (Figure C.3.b). Figures C.4 show the expected curves $I(\eta)$ for the different special cases shown in Figure C.3.

We can see from the intensity curves in Figure C.4 that the intensity oscillates with a period π when measuring linearly polarized light in these special cases. For circularly polarized light the measurements will result in constant intensities with different values. Based on Figure 3.14.a, we note that completely unpolarized light will also present the same feature. The derivation of the analytical functions and the method to determine the Stokes parameters from these four cases is done next, case by case.

Case 1: $\eta = \eta_1 = \eta_2$

First of all, we must recall that the intensity registered by the detector is proportional to the intensity of the light, *i.e.*

$$I(\eta_1, \eta_2) \propto \bar{S}_0'$$

where $\bar{S}_0'$ is the Stokes parameter of the light that hits the detector, following the notation in Chapter 3. If we assume that $I(\eta_1, \eta_2)$ is electric current measured from a photodiode, there is an efficiency factor $E_{\text{fficiency}}$ connecting electric current and light intensity (Ampere per Watt) so that

$$I(\eta_1, \eta_2) = E_{\text{fficiency}} \cdot \bar{S}_0'.$$

From now on we use a constant

$$\mathscr{E}_f = \frac{E_{\text{fficiency}}}{4} \cdot \left(r_{s3}{}^2 + r_{p3}{}^2\right)\left(r_{s4}{}^2 + r_{p4}{}^2\right),$$

which combined with the polarimeter equation (3.36) allows to write

$$I(\eta_1, \eta_2) = \mathscr{E}_f S_0 \cdot \left(c_0 + c_1 \cdot \bar{S}_1 + c_2 \cdot \bar{S}_2 + c_3 \cdot \bar{S}_3\right). \tag{C.10}$$

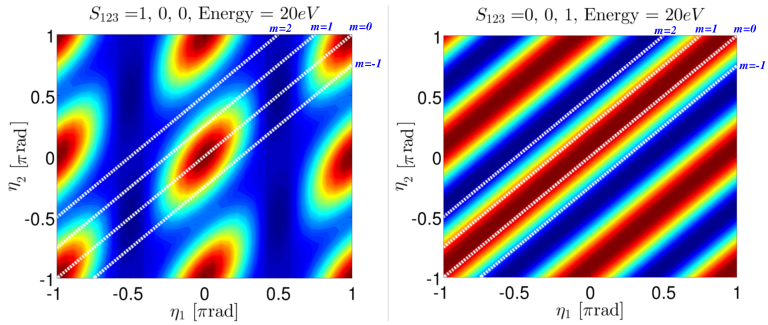


Figure C.3. *Intensity map expected for (a) horizontal linearly and (b) circularly polarized light. The white lines represent the values of η_1 and η_2 that are measured for each special case. They are labeled by the respective value of m. Cases 1, 2, 3 and 4 have values $m = 0, 2, 1, -1$, respectively.*

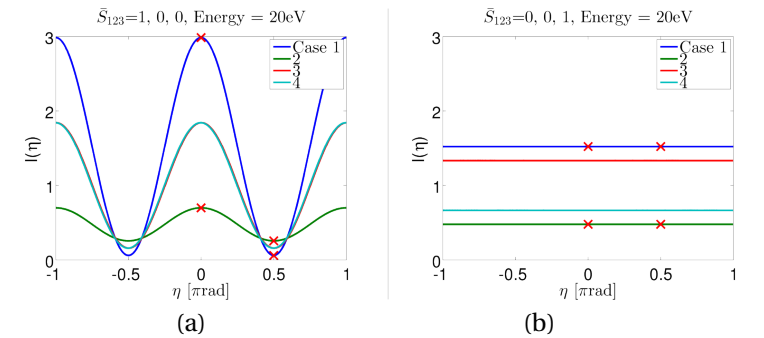


Figure C.4. *Expected intensity curves as measured along the lines in Figure C.3. a) horizontal linearly and (b) circularly polarized light. The red crosses are the points A, C, F and E discussed in section C.3*

The constant $\mathscr{E}_f$ can be eliminated by simply applying normalizations, as will be shown soon, and thus it will not be discussed any further. We can now apply the condition $\eta = \eta_1 = \eta_2$ of special case 1 into equation (C.10), resulting

$$I_1(\eta) = I(\eta,\eta) = \mathscr{E}_f S_0\big\{1+\cos(2\psi_3)\cos(2\psi_4) - \big[\cos(2\psi_3)+\cos(2\psi_4)\big]\big[\bar{S}_1\cos(2\eta)+\bar{S}_2\sin(2\eta)\big]\big\}. \tag{C.11}$$

Here we can use the identity of equation (C.2) to write

$$I_1(\eta) = \bar{I}_1\big(1+V_1\sin(2\eta+\Psi_1)\big). \tag{C.12}$$

Using further equations C.3 and C.6 we obtain that

$$\bar{I}_1 = \mathscr{E}_f S_0\big(1+\cos(2\psi_3)\cos(2\psi_4)\big), \tag{C.13a}$$

$$V_1 = \frac{\big(\cos(2\psi_3)+\cos(2\psi_4)\big)\bar{P}_L}{1+\cos(2\psi_3)\cos(2\psi_4)}, \tag{C.13b}$$

$$\tan\Psi_1 = \frac{\bar{S}_1}{\bar{S}_2}. \tag{C.13c}$$

where the term $\bar{P}_L$ is the degree of linear polarization, $\bar{I}$ the mean current and all the terms with bar (besides $\bar{I}$) are normalized by the first Stokes parameter, S_0(as in Appendix A).

It is actually simpler to use the product $\bar{I}\cdot V$ instead of only the visibility V, which results in

$$\bar{I}_1 \cdot V_1 = \mathscr{E}_f S_0\big(\cos(2\psi_3)+\cos(2\psi_4)\big)\bar{P}_L. \tag{C.13d}$$

We can see therefore that equations C.13 provide a link between the experimental quantities $\bar{I}_1$, V_1 and Ψ_1 with the Stokes parameters $\bar{S}_1$ and $\bar{S}_2$ and the optical constants (through $\cos(2\psi_3)$ and $\cos(2\psi_4)$). It is still necessary to study the cases 2, 3 and 4 which, together with equations C.13, make possible to determine the three Stokes parameters and the (relevant) optical constants directly.

Case 2: $\eta = \eta_1 = \eta_2 + \pi/2$

The derivation of the cases 2, 3 and 4 is very similar to case 1, and thus they are presented here in a streamlined manner.

Let us start by applying the condition $\eta = \eta_1 = \eta_2 + \pi/2$ to equation (C.10), yielding

$$\begin{aligned} I_2(\eta) &= I(\eta, \eta+\pi/2) \\ &= \mathscr{E}_f S_0 \cdot \Big\{(1-\cos(2\psi_3)\cos(2\psi_4)) \\ &\quad -\big[\cos(2\psi_3)-\cos(2\psi_4)\big]\big[\bar{S}_1\cos(2\eta)+\bar{S}_2\sin(2\eta)\big]\Big\}. \end{aligned} \tag{C.14}$$

This results in

$$\bar{I}_2 = \mathscr{E}_f S_0\big(1-\cos(2\psi_3)\cos(2\psi_4)\big), \tag{C.15a}$$

$$\bar{I}_2 \cdot V_2 = \mathscr{E}_f S_0\big(\pm\cos(2\psi_3)\mp\cos(2\psi_4)\big)\bar{P}_L, \tag{C.15b}$$

$$\tan\Psi_2 = \frac{\bar{S}_1}{\bar{S}_2}, \tag{C.15c}$$

where the symbols $\pm,\mp$ are valid either for the case $\cos(2\psi_3) > \cos(2\psi_4)$ or for the case $\cos(2\psi_3) < \cos(2\psi_4)$, respectively. The determination of which of the two term, $\cos(2\psi_3)$ or $\cos(2\psi_4)$, is bigger is for clarity discussed in the end of the section C.3, as derivation requires some arithmetic.

Using Case 1 and Case 2 to determine $\bar{S}_1$, $\bar{S}_2$, $\cos(2\psi_3)$ and $\cos(2\psi_4)$

Combining now equations C.13 and C.15 we get

$$\bar{P}_L = \sqrt{\frac{(V_1^2\bar{I}_1^2 - V_2^2\bar{I}_2^2)}{(\bar{I}_1^2 - \bar{I}_2^2)}} \tag{C.16a}$$

$$\cos(2\psi_3) = \frac{1}{\bar{P}_L}\cdot\frac{V_1\bar{I}_1 \pm V_2\bar{I}_2}{\bar{I}_1 + \bar{I}_2} \tag{C.16b}$$

$$\cos(2\psi_4) = \frac{1}{\bar{P}_L}\cdot\frac{V_1\bar{I}_1 \mp V_2\bar{I}_2}{\bar{I}_1 + \bar{I}_2}. \tag{C.16c}$$

$$\bar{S}_1 = \bar{P}_L\sin(\Psi_{1,2}), \tag{C.16d}$$

$$\bar{S}_2 = \bar{P}_L\cos(\Psi_{1,2}). \tag{C.16e}$$

where $\Psi_{1,2}$ means that either of the phases Ψ_1 and Ψ_2 can be used. Therefore we are able to directly determine the Stokes parameters $\bar{S}_1$ and $\bar{S}_2$ of the incoming light and the optical constants of the polarimeter $\cos(2\psi_3)$ and $\cos(2\psi_4)$ by measuring the intensities $I(\eta_1,\eta_2)$ in the conditions of the cases 1 and 2 and using the quantities mean currents $\bar{I}_1$ and $\bar{I}_2$, visibilities V_1 and V_2, and phases Ψ_1 and Ψ_2. Furthermore, it is not necessary to use any tabulated values or to use any fitting procedure, and it is possible to evaluate the experimental error of the Stokes parameters directly from the experimental errors. The discussion about the advantages of these features is presented in more details in Paper **III**. We observe that so far it is not possible to determine the Stoke parameter $\bar{S}_3$ related to circularly polarized light. For this we need to study the cases 3 and 4.

Case 3: $\eta = \eta_1 = \eta_2 - \pi/4$

Condition $\eta = \eta_1 = \eta_2 - \pi/4$ in equation (C.10) produces

$$\begin{aligned} I_3(\eta) &= I(\eta, \eta+\pi/4), \\ &= \mathscr{E}_f S_0\Big\{1 - \bar{S}_3\cos(2\psi_4)\sin(2\psi_3)\sin\delta_3 \\ &\quad - \cos(2\eta)\big[\bar{S}_1\cos(2\psi_3) + \bar{S}_2\cos(2\psi_4)\sin(2\psi_3)\cos\delta_3\big] \\ &\quad - \sin(2\eta)\big[-\bar{S}_1\cos(2\psi_4)\sin(2\psi_3)\cos\delta_3 + \bar{S}_2\cos(2\psi_3)\big]\Big\}, \end{aligned}$$

and again using the identity in equation (C.2) we obtain

$$\bar{I}_3 = \mathscr{E}_f S_0\Big[1 - \bar{S}_3\cos(2\psi_4)\sin(2\psi_3)\sin\delta_3\Big], \tag{C.17a}$$

$$\bar{I}_3\cdot V_3 = \mathscr{E}_f P_L\sqrt{\cos^2(\delta_3)\sin^2(2\psi_3)\cos^2(2\psi_4) + \cos^2(2\psi_3)}, \tag{C.17b}$$

$$\tan\Psi_3 = \frac{\bar{S}_1\cos(2\psi_3) + \bar{S}_2\cos(2\psi_4)\sin(2\psi_3)\cos\delta_3}{-\bar{S}_1\cos(2\psi_4)\sin(2\psi_3)\cos\delta_3 + \bar{S}_2\cos(2\psi_3)}. \tag{C.17c}$$

Case 4: $\eta = \eta_1 = \eta_2 + \pi/4$

Finally we apply $\eta = \eta_1 = \eta_2 + \pi/4$ in equation (C.10):

$$
\begin{aligned}
&I_4(\eta) = I(\eta, \eta - \pi/4),\\
&I_4(\eta) = \mathscr{E}_f S_0 \Big\{ 1 + \bar{S}_3 \cos(2\psi_4)\sin(2\psi_3)\sin\delta_3 \\
&\quad - \cos(2\eta)(\bar{S}_1 \cos(2\psi_3) - \bar{S}_2 \cos(2\psi_4)\sin(2\psi_3)\cos\delta_3) \\
&\quad - \sin(2\eta)(\bar{S}_1 \cos(2\psi_4)\sin(2\psi_3)\cos\delta_3 + \bar{S}_2 \cos(2\psi_3)) \Big\},
\end{aligned}
$$

and obtain

$$\bar{I}_4 = \mathscr{E}_f S_0 \Big[1 + \bar{S}_3 \cos(2\psi_4)\sin(2\psi_3)\sin\delta_3 \Big], \tag{C.18a}$$

$$\bar{I}_4 \cdot V_4 = \mathscr{E}_f P_L \sqrt{\cos^2(\delta_3)\sin^2(2\psi_3)\cos^2(2\psi_4) + \cos^2(2\psi_3)}, \tag{C.18b}$$

$$\tan\Psi_4 = \frac{\bar{S}_1 \cos(2\psi_3) - \bar{S}_2 \cos(2\psi_4)\sin(2\psi_3)\cos\delta_3}{\bar{S}_1 \cos(2\psi_4)\sin(2\psi_3)\cos\delta_3 + \bar{S}_2 \cos(2\psi_3)}. \tag{C.18c}$$

We observe that in cases 3 and 4 we have an additional dependence with the parameters $\bar{S}_3$ and δ_3. This will make possible to determine their values, as is shown next.

Using Case 3 and Case 4 to determine $\bar{S}_3$ and $\cos\delta_3$

Although the Equations C.17.c and C.18.c seem complicated, all parameters except δ_3 can be determined by the analysis of cases 1 and 2. Therefore, by measuring Ψ_3 or Ψ_4 it is possible to determine δ_3 by

$$\cos\delta_3 = \frac{\cos(2\psi_3)}{\cos(2\psi_4)\sin(2\psi_3)} \cdot \frac{-\bar{S}_1 + \bar{S}_2 \tan\Psi_3}{\bar{S}_2 + \bar{S}_1 \tan\Psi_3} \tag{C.19a}$$

or

$$\cos\delta_3 = \frac{\cos(2\psi_3)}{\cos(2\psi_4)\sin(2\psi_3)} \cdot \frac{\bar{S}_1 - \bar{S}_2 \tan\Psi_4}{\bar{S}_2 + \bar{S}_1 \tan\Psi_4}. \tag{C.19b}$$

Finally, from equations C.17.a and C.18.a we have that

$$\frac{\bar{I}_4 - \bar{I}_3}{\bar{I}_4 + \bar{I}_3} = \bar{S}_3 \cos(2\psi_4)\sin(2\psi_3)\sin\delta_3,$$

allowing to obtain Stokes parameter $\bar{S}_3$ by

$$\bar{S}_3 = \frac{1}{\cos(2\psi_4)\sin(2\psi_3)\sin\delta_3} \cdot \frac{\bar{I}_4 - \bar{I}_3}{\bar{I}_4 + \bar{I}_3}. \tag{C.20}$$

The values for $\bar{P}_L$, $\bar{S}_1$ and $\bar{S}_2$ can also be determined based on equations C.17 and C.18, resulting in

$$\bar{P}_L = \frac{\bar{I}_4 \cdot V_4}{\bar{I}_3 + \bar{I}_4} \frac{2}{\sqrt{\cos^2(\delta_3)\sin^2(2\psi_3)\cos^2(2\psi_4) + \cos^2(2\psi_3)}} \tag{C.21a}$$

$$= \frac{\bar{I}_3 \cdot V_3}{\bar{I}_3 + \bar{I}_4} \frac{2}{\sqrt{\cos^2(\delta_3)\sin^2(2\psi_3)\cos^2(2\psi_4) + \cos^2(2\psi 3)}}, \tag{C.21b}$$

$$\frac{\bar{S}_1}{\bar{S}_2} = \frac{\cos\Psi_3 + \cos\Psi_4}{\sin\Psi_3 + \sin\Psi_4}. \tag{C.21c}$$

In summary, by measuring the special cases 1, 2, 3 and 4 it is possible to determine the Stokes parameters and the optical properties with the use of equations C.16, C.19, C.20 and C.21. The use and additional discussion about this method is presented in more details in Paper **III**.

C.3 Method of measuring at specific pairs η_1 and η_2

The strategy for determining the polarization of the light with this polarimeter is by doing the measurements $I(\eta_1, \eta_2)$ at specific pairs of angles η_1 and η_2. Variations of this method were used with different purposes for instance by Koide et al. [1991], Schledermann and Skibowski [1971], Hamm et al. [1965] and Cubric et al. [1999]. These relations are derived here to adapt the results of these authors to the current case and notation.

The initial motivation is to choose proper angles pairs η_1, η_2 that simplify $I(\eta_1, \eta_2)$ in equation (C.10). It is indeed the same goal as for the method discussed previously in section C.2, but instead of measuring along the lines shown in Figure C.3, the current method is based on measuring at specific points of the intensity map $I(\eta_1, \eta_2)$.

Due to the presence of many terms including $\cos(2\eta_{12})$ and $\sin(2\eta_{12})$ in equation (3.36), the function $I(\eta_1, \eta_2)$ can be made simpler at angular values η_1, η_2 equal to 0, $\pm\pi/4$, $\pi/2$ or $\pm 3\pi/4$. In addition, it can be made even simpler by using angles which make the term c_3 of equation (3.36) equal to 0 or to 1. Using this, by combining the intensities $I(\eta_1, \eta_2)$ at different values of η_1, η_2 we obtain that the desired Stokes parameters and optical constant are obtained from the following equations

$$\cos(2\psi_3)\cos(2\psi_4) = \frac{A+C-F-E}{A+C+F+E},$$
$$\frac{\cos(2\psi_3)}{\cos(2\psi_4)} = \frac{-A+C+F-E}{-A+C-F+E},$$
$$\bar{S}_1 = \frac{1+\cos(2\psi_3)\cos(2\psi_4)}{\cos(2\psi_3)+\cos(2\psi_4)} \cdot \frac{C-A}{C+A},$$
$$\bar{S}_2 = \frac{1+\cos(2\psi_3)\cos(2\psi_4)}{\cos(2\psi_3)+\cos(2\psi_4)} \cdot \frac{D-B}{D+B},$$
$$\bar{S}_3 = \frac{1}{\sin(2\psi_3)\cos(2\psi_4)\sin(\delta_3)} \cdot \frac{M+Q-J-N}{M+Q+J+N},$$
$$\cos(\delta_3) = \frac{1}{\bar{S}_1 \sin(2\psi_3)\cos(2\psi_4)} \cdot \frac{O+L-K-P}{O+L+K+P},$$

where the points A to Q are the intensity values $I(\eta_1, \eta_2)$ measured at specific values of η_1, η_2 as following

$$A = I(0,0), \qquad J = I(0,\pi/4),$$
$$B = I(\pi/4,\pi/4), \qquad K = I(\pi/4,0),$$
$$C = I(\pi/2,\pi/2), \qquad L = I(\pi/4,\pi/2),$$
$$D = I(3\pi/4,3\pi/4), \qquad M = I(\pi/2,\pi/4),$$
$$E = I(0,\pi/2), \qquad N = I(\pi/2,3\pi/4),$$
$$F = I(\pi/2,0), \qquad O = I(3\pi/4,\pi/2),$$
$$P = I(3\pi/4,\pi),$$
$$Q = I(\pi,3\pi/4).$$

It is indeed possible to make many other combinations and to use other values of η_1 and η_2, but the discussion here is limited to the cases above. In figures C.4.(a) and (b) we can see the points A, C, F, E for these measurements as red crosses.

This method shows that is possible to determine all the desired parameters with fourteen measurements of intensity $I(\eta_1, \eta_2)$. It has a smaller number of experimental points and it is more sensitive to misalignment and experimental errors compared to the method shown in Section C.2 or the full mapping of the intensity $I(\eta_1, \eta_2)$. On the other hand it can be an easy and fast way to estimate all the experimental quantities of interest. In particular, it can be used to determine whether $\cos(2\psi_3) > \cos(2\psi_4)$ or $\cos(2\psi_3) < \cos(2\psi_4)$, and thus the correct sign in equation (C.16).

REFERENCES

Als-Nielsen, J. and D. McMorrow (2001). *Elements of Modern X-Ray Physics*. John Wiley and Sons.

Arfken, G. B., H. J. Weber, and F. E. Harris (2012). *Mathematical Methods for Physicists* (7 ed.). Elsevier.

Azzam, R. M. A. and N. Bashara (1987). *Ellipsometry and polarized light*. North-Holland personal library. North-Holland.

Bahrdt, J. (2007a, Jun). Wavefront tracking within the stationary phase approximation. *Phys. Rev. ST Accel. Beams 10*, 060701.

Bahrdt, J. (2007b, Jun). Wavefront tracking within the stationary phase approximation. *Phys. Rev. ST Accel. Beams 10*, 060701.

Bahrdt, J., R. Follath, and W. Frentrup (2010). Compensation of Beam Line Polarizing Effects at UE112 of BESSY II. *AIP Conference Proceedings 1234*(1), 335–338.

Bazarov, I. V. (2012, May). Synchrotron radiation representation in phase space. *Phys. Rev. ST Accel. Beams 15*, 050703.

Bevington, P. and D. Robinson (2003). *Data reduction and error analysis for the physical sciences* (3rd ed.). McGraw-Hill Higher Education. McGraw-Hill.

Born, M. and E. Wolf (1999). *Principles of Optics* (7th ed.). Cambridge University Press.

Bowler, M., J. Bahrdt, and O. Chubar (2008). Wavefront Propagation. In A. Erko, M. Idir, T. Krist, and A. G. Michette (Eds.), *Modern Developments in X-Ray and Neutron Optics*, Volume 137 of *Springer Series in optical science*, Chapter 2. Berlin, Heidelberg: Springer Berlin Heidelberg.

Brigham, E. (1988). *The Fast Fourier Transform and Its Applications*. Prentice-Hall Signal Processing Series: Advanced monographs. Prentice Hall.

Brookhaven National Laboratory (2006). NSLS-II Conceptual Design Report.

Canestrari, N., D. Karkoulis, and M. S. del Rio (2011). SHADOW3 API: The Application Programming Interface for the ray tracing code SHADOW. In DelRio, MS and Chubar, O (Ed.), *Proceedings of SPIE Conference 2011*, Volume 8141. SPIE.

Carr, R. (1993, May). Planar helical undulator sources of circularly polarized X-rays. In *Particle Accelerator Conference Proceedings, 1993*, pp. 1596–1598 vol.2.

Cash, W. (1987, Jul). X-ray optics: a technique for high resolution imaging. *Appl. Opt. 26*(14), 2915–2920.

Cerrina, F. and M. S. del Rio (2010). Ray-tracing of X-ray Optical Systems. In M. Bass (Ed.), *Handbook of Optics (Volume V), Atmospheric Optics, Modulators, Fiber Optics, X-Ray and Neutron Optics*, Chapter 35. Mc Graw Hill.

Chubar, O., L. Berman, Y. S. Chu, A. Fluerasu, S. Hulbert, M. Idir, K. Kaznatcheev, D. Shapiro, Q. Shen, and J. Baltser (2011). Development of partially-coherent wavefront propagation simulation methods for 3rd and 4th generation synchrotron radiation sources. Volume 8141, pp. 814107–814107–10. Proc. SPIE.

Chubar, O., Y. S. Chu, K. Kaznatcheev, and H. Yan (2010). Performance Optimization for Hard X-ray Microscopy Beamlines Guided by Partially Coherent Wavefront Propagation Calculations. *AIP Conference Proceedings 1234*(1), 75–78.

Chubar, O., Y. S. Chu, K. Kaznatcheev, and H. Yan (2011). Application of partially coherent wavefront propagation calculations for design of coherence-preserving synchrotron radiation beamlines. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 649*(1), 118–122. National Synchrotron Radiation Instrumentation conference in 2010.

Clark, J. A. (2004). *The Science and Technology of Undulators and Wigglers*. Oxford University Press.

Cormen, T. (2009). *Introduction to Algorithms* (3rd ed.). MIT Press.

Cubric, D., D. R. Cooper, M. C. A. Lopes, P. Bolognesi, and G. C. King (1999). Polarization measurements in the VUV region of synchrotron radiation. *Measurement Science and Technology 10*(6), 554.

del Rio, M. S., N. Canestrari, F. Jiang, and F. Cerrina (2011, SEP). SHADOW3: a new version of the synchrotron X-ray optics modelling package. *JOURNAL OF SYNCHROTRON RADIATION 18*(Part 5), 708–716.

Denecke, R., P. Väterlein, M. Bässler, N. Wassdahl, S. Butorin, A. Nilsson, J.-E. Rubensson, J. Nordgren, N. Mårtensson, and R. Nyholm (1999, June). Beamline I511 at MAX II, capabilities and performance. *Journal of Electron Spectroscopy and Related Phenomena 101-103*, 971–977.

Einfeld, D., M. Plesko, and J. Schaper (2014, Sep). First multi-bend achromat lattice consideration. *Journal of Synchrotron Radiation 21*(5), 856–861.

Elleaume, P., J. Chavanne, and B. Faatz (2000). Design considerations for a 1Å SASE undulator. *NIM A 455*(3), 503–523.

Eriksson, M., J. F. van der Veen, and C. Quitmann (2014, Sep). Diffraction-limited storage rings – a window to the science of tomorrow. *Journal of Synchrotron Radiation 21*(5), 837–842.

Fluerasu, A., O. Chubar, K. Kaznatcheev, J. Baltser, L. Wiegart, K. Evans-Lutterodt, M. Carlucci-Dayton, and L. Berman (2011). Analysis of the optical design of the NSLS-II coherent hard X-ray beamline. Volume 8141, pp. 81410J–81410J–7.

Follath, R. (1997, May). New plane-grating monochromators for third generation synchrotron radiation light sources. *NIM A 390*(3), 388–394.

Follath, R. (2001). The versatility of collimated plane grating monochromators. *NIM A 467–468, Part 1*(0), 418–425. 7th Int.Conf. on Synchrotron Radiation Instrumentation.

Follath, R., F. Senf, and W. Gudat (1998, May). Plane-grating monochromator at BESSY II using collimated light. *J. Synchrotron Rad. 5*(Pt 3), 769–71.

Fowles, G. R. (1975). *Introduction to Modern Optics* (2 ed.). Holt, Rinehart, and Winston.

Gaupp, A., M. MacDonald, F. Schaefers, R. Garrett, I. Gentle, K. Nugent, and S. Wilkins (2010). A W/B_4C Transmission Multilayer as an Achromatic Phase Shifter in the XUV: Some Experimental Aspects. In *SRI2010*, pp. 665–668.

Gaupp, A. and M. Mast (1989). First experimental experience with a VUV polarimeter at BESSY. *Review of Scientific Instruments 60*(7), 2213–2215.

Geloni, G., E. Saldin, E. Schneidmiller, and M. Yurkov (2005, June). Understanding transverse coherence properties of X-ray beams in third generation Synchrotron Radiation sources. *ArXiv Physics e-prints*.

Goldstein, D. (2011). *Polarized Light* (3rd ed.). CRC Press.

Goodman, J. (2005). *Introduction to Fourier Optics* (2nd ed.). Roberts & Company.

Goodman, J. (2015). *Statistical Optics* (2nd ed.). Wiley.

Halbach, K. (1981). Physical and optical properties of rare earth cobalt magnets. *NIM 187*(1), 109–117.

Hamm, R. N., R. A. Macrae, and E. T. Arakawa (1965, Nov). Polarization Studies in the Vacuum Ultraviolet. *J. Opt. Soc. Am. 55*(11), 1460–1463.

Hecht, E. (2002). *Optics* (4 ed.). Addison Wesley.

Hettel, R. (2014, Sep). DLSR design and plans: an international overview. *Journal of Synchrotron Radiation 21*(5), 843–855.

Huygens, C. (1690). *Traite de la Lumiere*. English translation: Treatise on Light by S. P.Thompson (London, Macmillan & Co., 1912).

Hwang, C. and S. Yeh (1999). Various polarization features of a variably polarized undulator with different phasing modes. *NIM A 420*(1–2), 29–38.

Jackson, J. D. (1997). *Classical Electrodynamics* (3 ed.). John Wiley and Sons.

Jiang, Y. H., R. Püttner, M. Martins, R. Follath, J. M. Rost, and G. Kaindl (2004, May). Isotope shifts of double-excitation resonances in helium. *Phys. Rev. A 69*, 052703.

Kim, K.-J. (1986). Brightness, coherence and propagation characteristics of synchrotron radiation. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 246*(1–3), 71–76.

Kirkpatrick, P. and A. V. Baez (1948, Sep). Formation of Optical Images by X-Rays. *J. Opt. Soc. Am. 38*(9), 766–773.

Kitamura, H. (1980). Polarization of Undulator Radiation. *Japanese Journal of Applied Physics 19*(4), L185–L188.

Koide, T., T. Shidara, M. Yuri, N. Kandaka, K. Yamaguchi, and H. Fukutani (1991, October). Elliptical-polarization analyses of synchrotron radiation in the 5–80-eV region with a reflection polarimeter. *NIM A 308*(3), 635–644.

Laundy, D., S. G. Alcock, L. Alianelli, J. P. Sutter, K. J. S. Sawhney, and O. Chubar (2014). Partial coherence and imperfect optics at a synchrotron radiation source modeled by wavefront propagation. Volume 9209, pp. 920903–920903–7.

Laundy, D., J. P. Sutter, U. H. Wagner, C. Rau, C. A. Thomas, K. J. S. Sawhney, and O. Chubar (2013). Parallel simulations of partially coherent wavefront propagation from a finite emittance electron beam. *Journal of Physics, Conference Series 425*(16), 162002.

Lidia, S. and R. Carr (1994). An elliptically-polarizing undulator with phase adjustable energy and polarization. *NIM A 347*(1–3), 77–82.

Loewen, E. and E. Popov (1997). *Diffraction Gratings and Applications*. Optical engineering. Taylor & Francis.

Lucke, R. L. (2006). Rayleigh–Sommerfeld diffraction and Poisson's spot. *European Journal of Physics 27*(2), 193.

MacDonald, M. A., F. Schäfers, and A. Gaupp (2009, Dec). A single W/B_4C transmission multilayer for polarization analysis of soft X-rays up to 1keV. *Opt. Express 17*(25), 23290–23298.

Mandel, L. and E. Wolf (1995). *Optical Coherence and Quantum Optics*. Cambridge University Press.

MATLAB (2013). version 8.0.0.783 (R2012b).

MAX IV Facility (2010). Detailed Design Report on the MAX IV Facility. (in progress).

Mills, D. M., J. R. Helliwell, Å. Kvick, T. Ohta, I. A. Robinson, and A. Authier (2005, May). Report of the Working Group on Synchrotron Radiation Nomenclature – brightness, spectral brightness or brilliance? *Journal of Synchrotron Radiation 12*(3), 385.

Murty, M. V. R. K. (1962, Jul). Use of Convergent and Divergent Illumination with Plane Gratings. *J. Opt. Soc. Am. 52*(7), 768–773.

Nahon, L. and C. Alcaraz (2004, February). SU5: a calibrated variable-polarization synchrotron radiation beam line in the vacuum-ultraviolet range. *Applied optics 43*(5), 1024–37.

Namioka, T. (1959, May). Theory of the Concave Grating. *J. Opt. Soc. Am. 49*(5), 446–460.

Noda, H., T. Namioka, and M. Seya (1974, Aug). Geometric theory of the grating. *J. Opt. Soc. Am. 64*(8), 1031–1036.

Nyholm, R., J. Andersen, U. Johansson, B. Jensen, and I. Lindau (2001, July). Beamline I311 at MAX-LAB: a VUV/soft X-ray undulator beamline for high resolution electron spectroscopy. *NIM A 467-468*, 520–524.

Onuki, H. and P. Elleaume (Eds.) (2003). *Undulators, Wigglers and Their Applications*. Taylor & Francis.

Padmore, H. A. (1989). Optimization of soft X-ray monochromators (invited). *Review of Scientific Instruments 60*(7), 1608–1615.

Palik, E. (1985). *Handbook of Optical Constants of Solids, Volumes I, II, and III: Subject Index and Contributor Index*. Academic Press Handbook Series. Elsevier Science & Tech.

Peatman, W. B. (1997). *Gratings, mirrors, and slits: beamline design for soft X-ray synchrotron radiation sources*. Gordon & Breach.

Petersen, H. (1982). The plane grating and elliptical mirror: a new optical configuration for monochromators. *Opt. Commun. 40*(6), 402–6.

Petersen, H. (1985, nov). Plane-grating monochromator. US Patent US 4553253.

Petersen, H. and H. Baumgärtel (1980). Bessy SX/700: A monochromator system covering the spectral range 3 eV$\lesssim \hbar\omega \lesssim$700eV. *NIM 172*(1–2), 191–3.

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

Riemer, F. and R. Torge (1983). Bessy SX 700 UHV monochromator: Design features and kinematic concept. *NIM 208*(1–3), 313–314.

Saleh, B. and M. Teich (2007). *Fundamentals of Photonics*. Wiley Series in Pure and Applied Optics. John Wiley & Sons.

Sankari, R. (2012, May). HIPPIE beamline at MAX IV: Detailed optical design report . Internal Report.

Sasaki, S. (1994). Analyses for a planar variably-polarizing undulator. *NIM A 347*(1–3), 83–86.

Schäfers, F. (2008). The BESSY Raytrace Program RAY. In A. Erko, M. Idir, T. Krist, and A. G. Michette (Eds.), *Modern Developments in X-Ray and Neutron Optics*, Volume 137 of *Springer Series in optical science*, Chapter 2. Berlin, Heidelberg: Springer Berlin Heidelberg.

Schäfers, F., H.-C. Mertins, A. Gaupp, W. Gudat, M. Mertin, I. Packe, F. Schmolla, S. D. Fonzo, G. Soullié, W. Jark, R. Walker, X. L. Cann, R. Nyholm, and M. Eriksson (1999, Jul). Soft-X-Ray Polarimeter with Multilayer Optics: Complete Analysis of the Polarization State of Light. *Appl. Opt. 38*(19), 4074–4088.

Schledermann, M. and M. Skibowski (1971, February). Determination of the Ellipticity of Light and of Optical Constants by Use of Two Reflection Polarizers. *Applied Optics 10*(2), 321.

Schmidt, T., A. Imhof, G. Ingold, B. Jakob, and C. Vollenweider (2007). A Fixed Gap APPLE II Undulator for SLS. *AIP Conference Proceedings 879*(1), 400–403.

Schmidt, T. and D. Zimoch (2007). About APPLE II Operation. *AIP Conference Proceedings 879*(1), 404–407.

Schnadt, J., J. Knudsen, J. N. Andersen, H. Siegbahn, A. Pietzsch, F. Hennies, N. Johansson, N. Mårtensson, G. Öhrwall, S. Bahr, S. Mähl, and O. Schaff (2012, Sep). The new ambient-pressure X-ray photoelectron spectroscopy instrument at MAX-lab. *Journal of Synchrotron Radiation 19*(5), 701–704.

Shapiro, D. A., O. Chubar, K. Kaznatcheev, R. Reininger, C. Sanchez-Hanke, and S. Wang (2011). Optimization of a coherent soft X-ray beamline for coherent scattering experiments at NSLS-II. Volume 8141, pp. 81410I–81410I–6.

Skripka, G., P. F. Tavares, M. Klein, and R. Nagaoka (2014). Transverse Instabilities In The Maxiv 3GeV Ring. In *Proceedings of IPAC2014, Dresde, Germany*, pp. 1689–16891.

Sommerfeld, A. (1949). *Lectures on Theoretical Physics: Optics*. Lectures on Theoretical Physics. Academic Press.

Southwell, W. H. (1981, Jan). Validity of the Fresnel approximation in the near field. *J. Opt. Soc. Am. 71*(1), 7–14.

Tanaka, T. and H. Kitamura (2001, November). SPECTRA: a synchrotron radiation calculation code. *Journal of Synchrotron Radiation 8*(6), 1221–1228.

Tavares, P. F., S. C. Leemann, M. Sjöström, and Å. Andersson (2014, Sep). The MAXIV storage ring project. *Journal of Synchrotron Radiation 21*(5), 862–877.

Vartanyants, I. A. and A. Singer (2010). Coherence properties of hard X-ray synchrotron sources and X-ray free-electron lasers. *New Journal of Physics 12*(3), 035004.

Voelz, D. (2011). *Computational Fourier Optics: A MATLAB Tutorial*. Tutorial Texts in Optical Engineering . SPIE - Society of Photo-Optical Instrumentation Engineers.

Walker, R. P. (1998). Insertion devices: undulators and wigglers. pp. 129–190. In Proceedings of the CERN Accelerator School, CAS-98-04, Edited by S. Tuner.

Wiedemann, H. (2003). *Synchrotron Radiation*. Springer-Verlag Berlin Heidelberg.

Willmott, P. (2011, Sep). *An Introduction to Synchrotron Radiation: Techniques and Applications* (1 ed.). Wiley.

Wolf, E. (1955). A Macroscopic Theory of Interference and Diffraction of Light from Finite Sources. II. Fields with a Spectral Range of Arbitrary Width. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences 230*(1181), pp. 246–265.

Wolf, E. (2007). *Introduction to the Theory of Coherence and Polarization of Light*. Cambridge University Press.

Wolf, E. and W. Marchand (1964, May). Comparison of the Kirchhoff and the Rayleigh—Sommerfeld Theories of Diffraction at an Aperture. *J. Opt. Soc. Am. 54*(5), 587–594.

COMMENTS ON THE PAPERS

I **Use of astigmatic re-focusing at HP-XPS end-station**
In this paper I present the astigmatic refocusing optics for the new high pressure photoelectrons spectroscopy (HP-XPS) branch line of SPCEIES beamline at MAX-lab. In the text I derived a geometric model for the vertical beam size and present some results from ray-tracing simulation. Moreover, the motivation for the project and the discussion about the advantages of the astigmatic focus are discussed. I performed all the simulations and wrote most of the article.

II **The SPECIES beamline at MAX-lab: a facility for soft-X-ray RIXS and HP-XPS experiments**
This paper reports the structure and performance of the SPECIES beamline, including the performance of the astigmatic re-focusing, which I studied theoretically in Paper I. Main contribution is in the comparison of experimental results to the theoretical predictions; I am also working in relevant parts of the manuscripts regarding the astigmatic refocusing.

III **Simplification of data analysis and experimental error propagation for a VUV polarimeter**
This manuscript is in preparation. It presents a novel data analysis of the polarimeter data and a straightforward way for error estimation. I developed the method and I worked in getting experimental confirmation of these ideas. This is also part of the commissioning of the VUV polarimeter, which I have been in charge. I am the main author of this manuscript and I have wrote most of it.

IV **Multilayer based EUV polarimeter at MAX IV laboratory**
This paper presents the EUV polarimeter, its performance, and possibilities for defining polarization with high precision. I actively participated in the commission phase of the instrument and experiments at beamlines I411 and I1011 at MAX II. I was also in charge of analyzing the results. I have contributed to the whole article and my main contribution is in the experimental part and discussion of the results.

V **Propagation of coherent light pulses with PHASE**
The paper describes the development of the PHASE code which is one of the available codes for wave front propagation dedicated for synchrotron light sources. My main contribution here was about developing a method for inclusion of the figure errors of optical elements in the simulations. I describe this method in section 4 of the article and I have wrote the most of this section.

VI **Design for a coherent imaging beamline SoftiMAX at MAX IV Laboratory**
This paper describes the proposed design of the first MAX IV soft-X-ray beamline utilizing the high coherence of the low emittance storage ring. Here the coherence is the key-factor and I studied its effects on image forming by using wavefront propagation simulations. I am also contributing in the preparation of the manuscript.

Papers