Investigations on the wavefront of the seed laser at the XUV seeding experiment at FLASH

Diploma Thesis
by
Ulrich Hipp

submitted to the
Department of Physics of the University of Hamburg

prepared at
Deutsches Elektronen-Synchrotron (DESY), Hamburg

Hamburg, January 2011
Erklärung

Hiermit bestätige ich, dass die vorliegende Arbeit von mir selbstständig verfasst wurde und ich keine anderen als die angegebenen Hilfsmittel benutzt habe. Ich bin damit einverstanden, dass die Diplomarbeit veröffentlicht wird.

Hamburg, Januar 2011

(Ulrich Hipp)

Gutachter der Diplomarbeit: Prof. Dr. Jörg Roßbach (Universität Hamburg)
Dr. Velizar Miltchev (Universität Hamburg)
Abstract

The Free-Electron Laser in Hamburg (FLASH) has been operated so far in the so-called Self-Amplified-Spontaneous-Emission (SASE) mode. In this operation mode the FEL process starts from noise and as a result the output radiation has poor longitudinal coherence. Moreover, the electron acceleration process introduces arrival-time fluctuations of the electron bunch at the undulator entrance, which lead to a temporal jitter of the XUV pulses and limit the resolution of XUV/optical pump-probe experiments. One option to improve coherence and synchronization is to start the FEL process not from noise, but to stimulate the FEL radiation with an external laser. This concept has been implemented in the framework of the seeding project at FLASH, called "sFLASH". The aim is to study the technical feasibility of seeding in the XUV spectral range. Along a variable-gap undulator the electron beam is overlapped with the external so-called "seed" laser beam of a corresponding XUV FEL resonance wavelength. This XUV seed is produced by high-harmonic generation (HHG) of powerful near-infrared (NIR) femtosecond laser pulses interacting with rare gas atoms.

In this thesis, the wavefront of the HHG seed at 38 nm is investigated. The influence of aberrant wavefronts on the experiment is studied in numerical simulations as well as experimentally. For this purpose, a Hartmann wavefront sensing system applicable for the XUV spectral range has been constructed. The laser beam is sampled by a rectangular grid and generates a spot pattern on a CCD chip. Compared to a known reference, this yields the wavefront's local slope at a large number of points and the wavefront's shape can be reconstructed using suitable algorithms with a resolution down to 0.2 \( \lambda \) (\( \lambda = 38 \) nm). The measurement process and the wavefront reconstruction algorithm are discussed in detail. Moreover, different wavefront measurement techniques are presented.

It is shown that the XUV wavefront shape has a big impact on the focus size in the undulator area. The need for a good coupling between XUV seed and the electron beam determines the required properties of the wavefront.
Zusammenfassung


Es wird gezeigt, dass die XUV-Wellenfront einen großen Einfluss auf die Fokussgrößen im Bereich des Undulators hat. Die Notwendigkeit einer präzisen Überlagerung zwischen XUV-Seed und Elektronenstrahl bestimmt die erforderlichen Eigenschaften der Wellenfront.
## Contents

1 Introduction
   1.1 Synchrotron light sources ...................................... 1
   1.2 FLASH ............................................................. 3
   1.3 sFLASH ............................................................ 5
   1.4 Relevance of wavefronts .......................................... 7

2 Theoretical concepts .................................................. 9
   2.1 The fundamental Gaussian mode .................................. 9
   2.2 Interpretation of wavefronts .................................... 12
   2.3 Definition of the beam quality factor $M^2$ .................... 14
   2.4 Propagation of laser beams ..................................... 16
      2.4.1 The paraxial wave equation ............................... 16
      2.4.2 Huygens' integral ........................................... 17
   2.5 Mathematical description of wavefronts ....................... 19
      2.5.1 Zonal approach ............................................ 20
      2.5.2 Modal approach ............................................ 20
   2.6 High-harmonic generation ...................................... 22
   2.7 Basic theory on free-electron lasers ......................... 25
      2.7.1 Comparison of FELs to synchrotrons ..................... 25
      2.7.2 Interaction between electron beam and light wave .......... 27
      2.7.3 One-dimensional description of FELs ..................... 29
      2.7.4 One-dimensional solution for seeding ..................... 32

3 Experimental setup of sFLASH ....................................... 35
   3.1 sFLASH HHG source ............................................ 35
   3.2 Injection beam line ............................................. 36
   3.3 Diagnostic beam line ........................................... 38

4 Numerical simulations ................................................. 39
   4.1 The "Physical Optics Propagation" mode in ZEMAX ................ 40
   4.2 Creating the wavefront model .................................. 41
   4.3 Propagation through the sFLASH injection beam line .......... 44
# List of Figures

1.1 Peak brilliance of various synchrotron radiation sources and FELs . . 3
1.2 Layout of FLASH . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 4
1.3 Microbunching and intensity growth in FELs . . . . . . . . . . . . . . 5
1.4 Experimental layout of sFLASH . . . . . . . . . . . . . . . . . . . . . 6
1.5 Exemplary wavefronts . . . . . . . . . . . . . . . . . . . . . . . . . . . 7

2.1 Properties of the fundamental Gaussian mode . . . . . . . . . . . . . 11
2.2 Wavefront as a surface of constant optical path length . . . . . . . . . 13
2.3 Wavefront as a surface of constant phase . . . . . . . . . . . . . . . . . 14
2.4 $M^2$ of a laser beam . . . . . . . . . . . . . . . . . . . . . . . . . . . 16
2.5 Patch of a wavefront . . . . . . . . . . . . . . . . . . . . . . . . . . . . 19
2.6 Zonal representation of a wavefront . . . . . . . . . . . . . . . . . . . 20
2.7 Superposition of spatial modes . . . . . . . . . . . . . . . . . . . . . . 20
2.8 Examples of low-order Zernike polynomials . . . . . . . . . . . . . . . 22
2.9 Three step model of HHG . . . . . . . . . . . . . . . . . . . . . . . . . 23
2.10 Typical HHG spectrum . . . . . . . . . . . . . . . . . . . . . . . . . . 24
2.11 Simulated HHG time profile . . . . . . . . . . . . . . . . . . . . . . . 25
2.12 Synchrotron radiation emitted by a circulating relativistic electron . . 25
2.13 Schematic view of an undulator . . . . . . . . . . . . . . . . . . . . . . 26
2.14 Condition for sustained energy transfer . . . . . . . . . . . . . . . . . 28
2.15 Power rise for seeding in the one-dimensional theory . . . . . . . . . . 33

3.1 Beam line of the 800 nm drive laser . . . . . . . . . . . . . . . . . . . . 35
3.2 Experimental hutch with HHG source besides the FLASH tunnel . . . . 36
3.3 Pit with HHG source inside the experimental hutch . . . . . . . . . . . 37
3.4 6-fold focal mirror wheel . . . . . . . . . . . . . . . . . . . . . . . . . 37
3.5 sFLASH injection beam line . . . . . . . . . . . . . . . . . . . . . . . . 37
3.6 Scheme of the HHG source and the diagnostic section . . . . . . . . . . 38

4.1 Scheme of the simulation workflow . . . . . . . . . . . . . . . . . . . . 39
4.2 Wavefront of the drive laser ($\lambda=800$ nm) . . . . . . . . . . . . . . . 42
4.3 Initial Gaussian intensity distribution . . . . . . . . . . . . . . . . . . . 43
4.4 Simulated initial wavefront aberrations . . . . . . . . . . . . . . . . . . 43
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.5</td>
<td>Beam sizes near undulator</td>
<td>44</td>
</tr>
<tr>
<td>4.6</td>
<td>Computed intensity profile and wavefront at undulator begin</td>
<td>46</td>
</tr>
<tr>
<td>4.7</td>
<td>Power along $z$ (initial beam power 50 kW)</td>
<td>47</td>
</tr>
<tr>
<td>4.8</td>
<td>Cross section along $z$ (initial beam power 50 kW)</td>
<td>48</td>
</tr>
<tr>
<td>4.9</td>
<td>Power density along $z$ (initial beam power 50 kW)</td>
<td>48</td>
</tr>
<tr>
<td>4.10</td>
<td>Power along $z$ near beam waist</td>
<td>49</td>
</tr>
<tr>
<td>4.11</td>
<td>Electron energy along $z$ near beam waist</td>
<td>50</td>
</tr>
<tr>
<td>4.12</td>
<td>Power along $z$ (initial beam power density 2 W/μm$^2$)</td>
<td>51</td>
</tr>
<tr>
<td>4.13</td>
<td>Cross section along $z$ (initial beam power density 2 W/μm$^2$)</td>
<td>52</td>
</tr>
<tr>
<td>4.14</td>
<td>Power density along $z$ (initial beam power density 2 W/μm$^2$)</td>
<td>52</td>
</tr>
<tr>
<td>5.1</td>
<td>Scheme of the Hartmann wavefront sensing technique</td>
<td>56</td>
</tr>
<tr>
<td>5.2</td>
<td>Dynamic range of a Hartmann sensor</td>
<td>56</td>
</tr>
<tr>
<td>5.3</td>
<td>Principle of the linear reconstruction</td>
<td>58</td>
</tr>
<tr>
<td>5.4</td>
<td>Description of a Hartmann measurement by a vector field</td>
<td>59</td>
</tr>
<tr>
<td>5.5</td>
<td>Curvature sensor</td>
<td>64</td>
</tr>
<tr>
<td>5.6</td>
<td>Radial shearing interferometer</td>
<td>65</td>
</tr>
<tr>
<td>5.7</td>
<td>Point diffraction interferometer</td>
<td>66</td>
</tr>
<tr>
<td>5.8</td>
<td>Knife-edge test</td>
<td>67</td>
</tr>
<tr>
<td>5.9</td>
<td>Ronchi test</td>
<td>68</td>
</tr>
<tr>
<td>5.10</td>
<td>Shack-Hartmann sensor</td>
<td>69</td>
</tr>
<tr>
<td>6.1</td>
<td>Adapter for the Hartmann masks</td>
<td>71</td>
</tr>
<tr>
<td>6.2</td>
<td>Photo of the experimental setup</td>
<td>72</td>
</tr>
<tr>
<td>6.3</td>
<td>Sketch of the experimental setup</td>
<td>72</td>
</tr>
<tr>
<td>6.4</td>
<td>Microscopy images of the fabricated Hartmann masks</td>
<td>74</td>
</tr>
<tr>
<td>6.5</td>
<td>Photo of the experimental setup for the reference measurement</td>
<td>75</td>
</tr>
<tr>
<td>6.6</td>
<td>Sketch of the experimental setup for the reference measurement</td>
<td>75</td>
</tr>
<tr>
<td>6.7</td>
<td>Computation of aperture positions</td>
<td>77</td>
</tr>
<tr>
<td>6.8</td>
<td>&quot;Virtual&quot; reference produced by a perfect Hartmann mask</td>
<td>78</td>
</tr>
<tr>
<td>6.9</td>
<td>Vector field of a tilted plane wavefront and reconstructed surface</td>
<td>79</td>
</tr>
<tr>
<td>6.10</td>
<td>Vector field of a spherical wavefront and reconstructed surface</td>
<td>79</td>
</tr>
<tr>
<td>6.11</td>
<td>Vector field of an arbitrary wavefront and reconstructed surface</td>
<td>80</td>
</tr>
<tr>
<td>6.12</td>
<td>Comparison of an arbitrary wavefront and its reconstructed surface</td>
<td>80</td>
</tr>
<tr>
<td>7.1</td>
<td>Measured spot pattern</td>
<td>81</td>
</tr>
<tr>
<td>7.2</td>
<td>Derived gradient vector field</td>
<td>82</td>
</tr>
<tr>
<td>7.3</td>
<td>Reconstructed wavefront compared to a plane wave</td>
<td>82</td>
</tr>
<tr>
<td>7.4</td>
<td>Reconstructed wavefront - surface plot</td>
<td>83</td>
</tr>
<tr>
<td>7.5</td>
<td>Reconstructed wavefront - color plot</td>
<td>83</td>
</tr>
</tbody>
</table>
7.6 Derived Zernike coefficients ........................................... 84
7.7 Wavefront reconstructed using the measured reference - surface plot . 85
7.8 Wavefront reconstructed using the measured reference - color plot ... 85
7.9 Error of the reconstructed wavefront by using the virtual reference ... 87
7.10 Error of the reconstructed wavefront by using the measured reference 88

8.1 Phase-space distribution and evaluated $M^2$-values .................... 92
8.2 Phase-space distribution without outliers and evaluated $M^2$-values . . 93
8.3 Embedded Gaussian beam .................................................. 94
8.4 Results of propagation to undulators with ZEMAX ....................... 95
8.5 Intensity distribution at undulator begin .................................. 96
8.6 Wavefront at undulator begin ............................................. 97
8.7 Results of back-propagation with ZEMAX ............................... 97
8.8 Power of the real seed (initial power 50 kW) ............................ 98
8.9 Power of the real seed (initial power 2 kW) ............................... 99
List of Tables

4.1 Rms beam sizes at undulator begin with a focal length of 7 m . . . . . . 45
4.2 Start conditions with constant initial seed power . . . . . . . . . . . . 47
4.3 Start conditions with constant initial seed power density . . . . . . . 51

8.1 Parameters of the fitted Gaussian mode . . . . . . . . . . . . . . . . . 89
8.2 $M^2$ of the seed laser . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 90
8.3 Recalculated beam waist sizes at the gas target . . . . . . . . . . . . 91
8.4 Beam sizes at undulator begin . . . . . . . . . . . . . . . . . . . . . . . . . 96
1 Introduction

1.1 Synchrotron light sources

Experiments using electromagnetic waves, in a range of wavelengths going well beyond the range of visible light, have played an essential role in the development of modern science. Scientists have been eager to obtain the brightest light sources to continuously improve the understanding of the atomic and electronic structure of matter. With the advent of synchrotron radiation in 1947 [1] research has been revolutionized in various areas of fundamental and applied science like solid state physics, material science, biology, and medicine. Synchrotron light sources are classified by their brilliance which is the crucial figure of merit for most experiments.

The brilliance $B$ is defined as the number of photons $N_{\text{photons}}$ emitted by the source per unit of time $\Delta t$ into a unit of solid angle $\Delta \Omega$, per unit of surface $\Delta A$ of the source and into a unit bandwidth around a given frequency $\Delta \omega/\omega$. Mostly $B$ is given in units [photons/(sec·mrad²-mm²·0.1% bandwidth)].

$$B = \frac{N_{\text{photons}}}{\Delta t \ \Delta \Omega \ \Delta A \ (\Delta \omega/\omega)}$$

(1.1)

Since the first attempts for a systematic exploitation of storage rings as light sources for scientific experiments in the 1960's [2], the brilliance of synchrotron light sources has been growing rapidly. The development of the synchrotron radiation sources can be divided in several periods which are referred as "generations". The first synchrotron facilities, built for high-energy physics research, have been parasitically used for some fraction of their beam time as sources of synchrotron radiation. The results obtained with these so-called "first generation" light sources were so interesting and promising that dedicated storage rings for production of synchrotron radiation have been constructed. Examples for such "second generation" light sources are the BESSY I ring in Berlin and the two National Synchrotron Light Sources in Brookhaven. In the 1990's the "third generation" machines started operation, which are characterized by reduced emittance of the circulating particle beam and by the extensive use of undulators as radiation sources [2]. Thus, the brilliance has been increased by several orders of magnitude. BESSY II, the European Synchrotron Radiation Facility (ESRF) in Grenoble, and the Advanced Light Source...
1 Introduction

(ALS) in Berkeley are examples for such third generation rings. However, there is a fundamental lower limit of the reachable emittance in the storage rings determined by the equilibrium between quantum excitations that cause the particles to oscillate transversely and the dumping provided by the RF accelerating system. One possibility to reduce the emittance is to abandon the ring geometry and to consider linear accelerators (linac) which feed the accelerated electrons into a long undulator. Thus, very low emittances and short electron bunches ($\sim 100$ fs as opposed to the typical $\sim 50$ ps in the storage rings) can be reached [2]. The fourth generation machines, called "Free-Electron Lasers" (FELs), are able to produce laser-like fully coherent radiation in sub-ps pulses of extremely high peak brilliance exceeding third generation sources by more than eight orders of magnitude. The Free-Electron Laser in Hamburg (FLASH) at DESY is an example for a fourth generation light source and has been playing an essential role in the development of X-ray FELs. Technological studies at FLASH have contributed to the development of the European X-ray FEL (XFEL) which is currently under construction and is supposed to start user operation in 2015 [3]. With a 3 km long accelerator being one of the world's largest FEL it will produce leading-edge research. Another example of important FEL facilities is the Linac Coherent Light Source (LCLS) in Stanford, which was commissioned in 2009 [4]. An even further increase in brilliance of FELs may be provided by the so-called "seeding" - stimulating the emission of radiation by an external laser (the so-called "seed") resulting in an additional improvement of the longitudinal coherence. With the currently ongoing project sFLASH, this idea is being tested at FLASH in the XUV spectral range. In Fig.1.1 peak brililances of different types and generations of X-ray radiation sources are shown with sFLASH being one of the most brilliant sources [5].
1.2 FLASH

The Free-Electron Laser FLASH at DESY, Hamburg, is a unique source of extremely bright, coherent, and ultrashort pulses of extreme ultraviolet (XUV) and soft X-ray radiation. With a time duration down to 10-30 fs and a power of the pulses in the Gigawatt range, the peak brilliance of FLASH is about eight orders of magnitude higher than the most existing X-ray sources [6]. In combination with the coherence of the pulses, this opens new frontiers for science and technology and allows experiments ranging from solid-state physics, chemistry, biology, material sciences, geophysics, and medical diagnostics. A recently completed energy upgrade of FLASH now allows lasing at 4.12 nm [6]. This has opened the view through the so-called water window which is the wavelength region between 2.3 and 4.4 nm where water is transparent (for the FEL radiation). This feature enables investigations on...
living biological specimens. Furthermore, it will be possible to make time resolved observations of chemical reactions on the femtosecond scale. This is coming close to the dream of making ultrafast movies of the nanoscale dynamics.

FLASH, schematically shown in Fig. 1.2, is an about 300 m long, linac based, single-pass, high-gain FEL which operates at the XUV and the soft X-ray wavelength range.

![Figure 1.2: Schematic layout of FLASH. (not to scale, taken from [6])](image)

Its electron source is a laser-driven photocathode mounted in a 1.3 GHz, 1.6-cell RF cavity which is able to produce electron bunches of below 2 mm·mrad normalized transverse emittance. The emittance is a measure of the spread of the electrons in the transverse phase-space and is of utmost importance for the FEL performance. The photo injector is followed by seven TESLA-type superconducting accelerator modules, each of them having eight niobium cavities. The maximum possible energy of the accelerated electrons is about 1.2 GeV [5], which corresponds to the shortest achievable wavelength below 5 nm. The FEL process requires high peak current of few kA which cannot be produced directly in the gun. Therefore, the electron bunches are longitudinally compressed in two 4-dipole magnetic chicanes - the so called "bunch compressors" - located after the first and the third accelerating module. The 3rd-harmonic RF module downstream the first accelerating module (see Fig.1.2) is used to linearize the electron bunch longitudinal phase-space, what makes the compression more effective and enables the generation of longer (~100 fs) bunches of high peak current. This in turn makes a stable seeded FEL operation possible.

After the acceleration section, the electron bunches pass through the undulator magnets. The 30 m long magnets of FLASH consist of permanent NdFeB with a fixed gap of 12 mm, a period length $\lambda_u$ of 27.3 mm, and peak magnetic field $B$ of 0.47 T. Undergoing a wavelike trajectory, the electrons emit radiation in a cone in forward direction. As a result of the interaction between the emitted radiation and the electron beam, a longitudinal charge density modulation develops. The electrons are concentrated in longitudinal slices (so-called "microbunches") shorter than the radiation wavelength, which radiate coherently as a single high-charge macro particle. This effect leads to exponential growth of the FEL power along the
undulator length $z$ as depicted in Fig.1.3. Herein also the process of microbunching is schematically shown.

![Figure 1.3: Exponential growth of radiation power as a function of undulator length $z$. The evolution of the microbunching is indicated schematically. Along $z$, an electron bunch (red cloud) is being structured into longitudinal slices (microbunches) shorter than the radiation wavelength. (taken from [7])](image)

Finally, a dipole magnet deflects the electron beam into a dump, while the FEL radiation propagates to the experimental hall. There, the radiation is available for user experiments. The sFLASH project, also shown in Fig.1.2, will be discussed in the following.

### 1.3 sFLASH

Currently FLASH operates in the Self-Amplified-Spontaneous-Emission (SASE) mode. As a consequence of the start-up from noise, the SASE-radiation consists of a number of uncorrelated modes resulting in reduced longitudinal coherence and shot-to-shot fluctuations of the output pulse energy [8]. Moreover, the temporal jitter of the XUV pulses relative to a synchronization optical laser limits the temporal resolution of time resolved pump-probe experiments.

An alternative possibility is to operate FLASH as an amplifier for an injected seed laser at the FEL resonance wavelength. The seed pulses are produced using high-harmonic generation (HHG) of a near-infrared laser focused in a noble-gas jet. Using this seeding mode, the longitudinal coherence can be greatly improved and
1 Introduction

a higher shot-to-shot stability at GW-power level with a pulse duration given by
the seed pulse of the order of 20 fs FWHM can be achieved [9]. This will increase
the brilliance by some orders of magnitude. In this case, the FEL output would
be naturally synchronized with the external seed laser and jitter-free XUV/optical
pump-probe experiments may be performed.

For sFLASH 10 m of gap-tunable undulators were installed in a distance of about
40 m upstream the existing fixed-gap SASE undulators (see Fig.1.2). In order to
achieve the transverse overlap between seed and electron beam, a dedicated beam
line was designed and installed, giving the possibility to steer and focus the laser
beam within the undulator [8]. As the spatial and temporal overlap of the seed
and the electron beam have to be very precise, the efficient transport of a short
wavelength laser ensuring accurate spatial overlap is one of the most challenging
tasks. After amplification in the sFLASH undulators, the seeded FEL-radiation
is extracted using a mirror chamber and transported to the photon diagnostics
equipment. Fig.1.4 shows a detailed overview of the sFLASH layout.

Figure 1.4: The FLASH facility (top) comprises a 260 m long tunnel housing the linac
and undulators of a SASE FEL. A 40 m long section (bottom) preceding
the SASE undulators accommodates four additional undulators for sFLASH.
Seed pulses from high-harmonic generation (HHG) in a building adjacent to
the FLASH tunnel are aligned to the electron beam (left). At the undulator
exit, the electron beam is displaced while FEL radiation is sent by mirrors to
an experimental hutch. (taken from [10])
1.4 Relevance of wavefronts

A French-Japanese collaboration at SPring-8 Compact SASE source has already successfully demonstrated HHG seeding at 61.2 nm [11]. At sFLASH the goal is to study the technical feasibility of seeding at wavelength as short as 38 nm.

1.4 Relevance of wavefronts

The most frequently used definition for a wavefront is a surface connecting all points of a wave having the same phase. See Fig.1.5 showing two exemplary wavefronts.

Fig. 1.5: Exemplary wavefronts

(a) Spherical wavefronts of a point source
(b) Wavefronts of a plane wave

The wavefronts of a point source (left) are spheres all having the same central point whereas the wavefronts of a plane wave (right) are planes. One of the widest spread wavefront sensing methods is the Hartmann sensor. This principle is applicable in many diverse areas from medicine to astronomy. For example, this method is routinely used to measure aberrations of the human eye or to control adaptive mirrors of telescopes to compensate for atmospheric turbulences. In these examples one can get an impression of the importance of wavefronts. They represent the phase information of a beam, which is, in addition to the intensity profile, needed for simulating beam propagation. Therefore, wavefront measurements have a direct impact on the optimization of optical devices.

Going back to sFLASH, also here an influence is expected. The energy gain of the seeded FEL depends on the relative phase between the electron transverse oscillatory motion and the co-propagating wave (e.g. the seed laser or photons from spontaneous emission). Besides the FEL resonance condition, a certain phase relation must be hold in order the light wave to gain and the electrons to lose energy. The latter condition may vary in the plane perpendicular to the propagation direction, depending on the wavefront of the electromagnetic radiation. This transverse variation of the energy gain has an impact not only on the FEL output power, but also on the FEL radiation properties (e.g. divergence) which are determined by the gain-guiding effects.
2 Theoretical concepts

2.1 The fundamental Gaussian mode

The basis for describing electromagnetic wave propagation is given by the Helmholtz electromagnetic wave equation derived from the Maxwell equations. Solving the Helmholtz equation within the paraxial approximation leads to the laser beam modes as shown for example in [12]. The laser modes can also be derived using a different technique which is presented in [13]. Neglecting the vector character of the electromagnetic field (scalar field approximation), any electric field amplitude distribution can be represented as a superposition of plane waves by

\[ E \propto \int \int \frac{d\kappa_x d\kappa_y}{(2\pi)^2} A(k_x, k_y) \exp(ik_x x + ik_y y + i\sqrt{k^2 - k_x^2 - k_y^2}). \] (2.1)

\( A(k_x, k_y) \) is the amplitude of the plane wave contribution with the wave vector \( \vec{k} = (k_x, k_y, \sqrt{k^2 - k_x^2 - k_y^2}) \) whose modulus is given by \( k = |\vec{k}| = 2\pi n/\lambda \). Thereby, \( n \) is the refractive index of the medium and \( \lambda \) the wavelength of light in vacuum. As derived in [13], the divergence can be defined as

\[ \text{divergence} \propto \int \int \frac{dk_x}{2\pi} \int \frac{dk_y}{2\pi} (k_x^2 + k_y^2)|A|^2 \] (2.2)

and the transversal spatial extension (spread of the field intensity perpendicular to the propagation direction) is given by

\[ \text{transversal extension} \propto \int \int \frac{dk_x}{2\pi} \int \frac{dk_y}{2\pi} \left( \left| \frac{\partial A}{\partial x} \right|^2 + \left| \frac{\partial A}{\partial y} \right|^2 \right). \] (2.3)

As by definition the fundamental Gaussian mode has both minimal divergence and minimal transversal extension, this mode can be derived by looking for an electromagnetic field that minimizes the product of divergence and transversal extension. Due to symmetry reasons, this simplifies to an amplitude function \( A(k_x, k_y) \) minimizing the product

\[ \left( \int \int \frac{dk_x}{2\pi} |A|^2 \right) \left( \int \int \frac{dk_x}{2\pi} k_x^2 \left| \frac{\partial A}{\partial k_x} \right|^2 \right) \] (2.4)
and the similar product for the $y$-direction. Using the Schwartz inequality, this product is found to be minimal if $k_x A$ and $\partial A/\partial k_x$ differ solely by a constant factor. The same applies for the $y$-direction, which leads to the explicit amplitude expression

$$A(k_x, k_y) \propto \exp \left( -\frac{w_0^2}{4} (k_x^2 + k_y^2) \right).$$  

(2.5)

Here an arbitrary proportionality factor $-w_0^2/2$ was introduced. The fundamental laser mode is now written as

$$E_0 \propto \int_{-\infty}^{\infty} \frac{dk_x}{2\pi} \int_{-\infty}^{\infty} \frac{dk_y}{2\pi} \exp \left( ik_x x + ik_y y + ik_z z - i \frac{k_x^2 + k_y^2}{2k} z - \frac{w_0^2}{4} (k_x^2 + k_y^2) \right),$$  

(2.6)

where carrying out the integration leads to

$$E_0 = \frac{1}{w_0^3 + 2iz/k} \exp \left[ ikz - \frac{x^2 + y^2}{w_0^2 + 2iz/k} \right].$$  

(2.7)

Usually, the longitudinal distance is expressed by the reduced coordinate $\zeta$

$$\zeta = \frac{z}{z_R}$$  

(2.8)

introducing the so-called Rayleigh length $z_R$. This is the distance from the beam focus at which the beam area doubles.

$$z_R = \frac{k w_0^2}{2} = \frac{\pi w_0^2}{\lambda}$$  

(2.9)

With the abbreviation $\rho^2 = x^2 + y^2$ and due to the fact that

$$\frac{1}{1 + i\zeta} = \frac{1 - i\zeta}{1 + \zeta^2} = \frac{\exp(-i \arctan(\zeta))}{\sqrt{1 + \zeta^2}},$$  

(2.10)

the fundamental Gaussian mode can be rewritten in the more common form

$$E_0 = \frac{1}{w_0 \sqrt{1 + \zeta^2}} \exp \left[ ikz - \frac{1 - i\zeta}{w_0^2(1 + \zeta^2)} \rho^2 - i \arctan(\zeta) \right].$$  

(2.11)

The corresponding transversal intensity profile is Gaussian

$$|E_0|^2 = \frac{1}{w_0^2(1 + \zeta^2)} \exp \left[ -\frac{2\rho^2}{w_0^2(1 + \zeta^2)} \right],$$  

(2.12)

which is the reason why the fundamental mode is called Gaussian. The beam size of the Gaussian mode is denoted as $w(z)$ and defines the radial distance from the optical axis where the intensity has fallen to $1/e^2$ times the peak intensity at $\rho = 0$. $w(z)$ is just twice the variance of the corresponding Gaussian intensity profile and can be calculated as a function of $z$ by

$$w(z) = w_0 \sqrt{1 + \zeta^2}.$$  

(2.13)
2.1 The fundamental Gaussian mode

According to this, \( w_0 \) is the size of the beam waist at focus position \( (\zeta = 0) \). This is the \( z \)-position where the beam has the lowest transversal extension. The \( \arctan(\zeta) \)-term in equation 2.11 is called Gouy’s phase \( \psi_0 \)

\[
\psi_0 = \arctan(\zeta).
\] (2.14)

Gouy’s phase describes the rapid phase change of the electric field when traversing the point of minimal beam diameter at the waist position \( (z = 0) \).

To determine the wavefront of Gaussian beams, one has to seek for the surface of constant phase in equation 2.11.

\[
\phi = kz + \frac{\zeta \rho^2}{w(\zeta)^2} - \psi_0 \approx kz + \frac{\zeta \rho^2}{w(\zeta)^2} = \text{const}
\] (2.15)

Neglecting Gouy’s phase in this derivation is valid as \( |\zeta| \) has large values and \( \arctan(\zeta) \) only changes slowly with changing \( \zeta \) \[13\]. Solving for \( z \) defines parabolas

\[
z = -\frac{\zeta r^2}{k w(\zeta)^2}
\] (2.16)

with apex curvature radius

\[
R(z) = \frac{k w_0^2}{2} \frac{1 + \zeta^2}{\zeta^2} = z(r) + \frac{z_R^2}{z(r)}.
\] (2.17)

At waist position the wavefront is plane \( (R(0) = \infty) \) whereas at position \( z = z_R \) the wavefront reaches its maximum curvature \( (R(z_R) = 2z_R) \) and becomes again plane at very long distances.

\[
R(z) = z_R + \frac{z^2}{z_R^2} \approx \begin{cases} 
\infty & \text{for } z << z_R \\
2z_R & \text{for } z = r_R \\
\infty & \text{for } z >> z_R
\end{cases}
\] (2.18)

In Fig.2.1 the basic properties of the fundamental Gaussian mode are depicted.

![Figure 2.1: Properties of the fundamental Gaussian mode](image)
2 Theoretical concepts

From equation 2.13 one can see that at long distances \((z \gg z_R)\) the beam radius increases approximately linear with \(z\).

\[
w(z) \approx w_0 \frac{z}{z_R} \quad \text{for} \quad z \gg z_R \tag{2.19}
\]

Due to this linearity, a divergence angle can be defined which is derived as follows.

\[
\text{divergence} = \lim_{z \gg z_R} \frac{w(z)}{z} = \frac{w_0}{z_R} = \frac{\lambda}{\pi w_0}
\tag{2.20}
\]

The Gaussian mode is the simplest among a whole family of modes, which contains an infinite number of modes [14]. Hermite-Gaussian and Laguerre-Gaussian modes are the most frequently used mode families with the Gaussian mode being the fundamental one \((TEM_{00})\). All other modes are called higher-order modes and have in general more complicated intensity profiles and wavefront shapes. Only the Gaussian mode reaches the diffraction limitation and is therefore the best collimated beam one can think of.

2.2 Interpretation of wavefronts

To illustrate the wavefront concept, consider a point source of light. It emits light rays in all directions in space, or into the solid angle \(4\pi\) steradians. Next consider the ensemble of rays having a certain optical path length \((OPL)\). The \(OPL\) is the product of the geometric length of the path light follows through a system and the index of refraction of the medium through which it propagates. In a medium of constant refractive index \(n\), the \(OPL\) for a path of physical length \(d\) is just

\[
OPL = n \cdot d.
\tag{2.21}
\]

Assuming a homogeneous medium, the tips of this ensemble of rays lie on a surface given by a sphere which is centered at the point source. This surface is a basic example of what is meant by a wavefront. Going further, one can imagine an anisotropic refractive index \(n(\vec{r})\). For a non-constant refractive index \(n(\vec{r})\), the \(OPL\) is defined by

\[
OPL = \int_p n(s) \, ds
\tag{2.22}
\]

where \(n(s)\) is the local refractive index as a function of distance \(s\) along the path \(p\). Now, the tips of the ensemble of rays having a constant \(OPL\) will no longer form a sphere but a different surface which is determined by the refractive index variation. Still, this surface represents the wavefront which therefore may be defined as a surface of constant \(OPL\). An essential feature is that the wavefronts and the light
2.2 Interpretation of wavefronts

rays are always orthogonal to each other. In Fig. 2.2 the interpretation of wavefronts as surfaces of constant OPL is shown schematically.

Figure 2.2: Wavefront of a point source in an anisotropic medium. The wavefront can be defined by the surface of constant optical path length OPL.

Often the wavefront is called "phasefront" because it is not only a surface of constant OPL but is more generally defined by a surface of constant phase as already stated in chapter 1.4. Both definitions are equivalent and there is a direct connection between them, which will be explained in the following. Assuming a light ray propagating in z-direction, the electric field is described by

\[ E = E_0 e^{i(kz - \omega t)} = E_0 e^{i\phi}, \]  

(2.23)

with the light’s angular frequency \( \omega \), the wave number \( k = 2\pi n/\lambda \), and the phase \( \phi = (kz - \omega t) \). \( E \) is cyclic with a phase period of \( 2\pi \) where \( \phi \) refers to some point in this cycle. For example, \( E = E_0 \) when \( \phi = \pi/2 \) and \( E = 0 \) when \( \phi = \pi \). The temporal variation of the electric field is often ignored as only phase differences are of interest which are constant in time. Assuming \( t = 0 \), the phase is given by \( \phi = kz = (2\pi n/\lambda)z \). Using the definition of the OPL, \( \phi \) can be also expressed by

\[ \phi = \frac{2\pi n}{\lambda} z = \frac{2\pi}{\lambda} \cdot OPL. \]  

(2.24)

The wavelength in vacuum \( \lambda \) is constant and therefore the phase \( \phi \) and the optical path length OPL only differ by a constant factor \( 2\pi/\lambda \). So a constant OPL refers directly to a constant phase and vice versa, both defining the wavefront. If one considers two rays leaving a point source, they are observed to have equal OPL and simultaneously equal phase \( \phi_0 \) at the wavefront. Note that it is the nature of a point
source that all rays leave the source having the same phase. See Fig. 2.3 showing this schematically.

![Wavefront of a point source in an anisotropic medium.](image)

**Figure 2.3:** Wavefront of a point source in an anisotropic medium. The wavefront can be defined by the surface of both, constant phase and constant optical path length $OPL$.

### 2.3 Definition of the beam quality factor $M^2$

As already mentioned, the Gaussian fundamental mode represents the best collimated beam and has therefore the highest beam quality. But this case is usually not reached in the real applications. The higher-order modes might induce more complicated intensity profiles and cause larger divergence angles and waist sizes. The deviation from the perfect case can be described by the "$M^2$-factor", also called beam quality factor or beam propagation factor. It is defined as the beam parameter product divided by $\lambda/\pi$. Thereby, the beam parameter product is the product of the beam’s divergence $\theta$ and waist size $w_0$.

$$M^2 = \theta \cdot w_0 \cdot \frac{\pi}{\lambda}$$  \hspace{1cm} (2.25)

As for the diffraction-limited fundamental Gaussian mode the beam parameter product is $\lambda/\pi$, the $M^2$ equals 1 in this case. All other beam modes end up in an $M^2 > 1$. The $M^2$ quantifies the quality of a laser beam and how well it can be focused to a small spot. Furthermore, the $M^2$ can be used to describe beam propagation of higher order modes just by modifying slightly the equations for the fundamental Gaussian mode, which are comparatively simple. The intensity profiles of higher-order modes are no longer necessarily Gaussian and the beam’s transversal extension...
2.3 Definition of the beam quality factor $M^2$

might depend on the transversal direction $x$ or $y$. By convention, the transversal beam sizes $W_x(z)$ and $W_y(z)$ of higher-order modes are given by two times the second moment of the beam’s intensity profile $I(x,y)$ across the rectangular coordinate $x$ and $y$ respectively [15]. Upper-case letters indicate beam sizes of non-Gaussian modes.

$$W_x = 2\sigma_x = 2\sqrt{\frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x,y)x^2 dxdy}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x,y) dxdy}}$$

$$W_y = 2\sigma_y = 2\sqrt{\frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x,y)y^2 dxdy}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x,y) dxdy}}$$

This would match the definition of the fundamental mode’s beamsize $w(z)$ as given before if $I(x,y)$ was a Gaussian intensity profile. The evolution of the beam sizes for higher-order modes reads as follows [15].

$$W_x(z) = \sqrt{W_{0x}^2 + M_x^2 \left( \frac{\lambda}{\pi W_{0x}} \right)^2 z^2}$$

$$W_y(z) = \sqrt{W_{0y}^2 + M_y^4 \left( \frac{\lambda}{\pi W_{0y}} \right)^2 z^2},$$

where $M_x^2$ and $M_y^2$ are the beam quality factors for the transversal $x$- and $y$-direction and $W_{0x}$ and $W_{0y}$ are the respective waist sizes. A very useful feature of the $M^2$ is that one can envision an "embedded Gaussian beam" to any higher-order mode in either transversal direction $x$ and $y$. This embedded Gaussian beam does not necessarily have to represent the lowest-order mode component of the real beam [15]. With a given $M^2$, the beam size $w_{em}(z)$ and waist size $w_{0m}^m$ of the embedded Gaussian beam yield

$$w_{0m}^m = \frac{W_0}{M^2} \quad \text{and} \quad w_{em}(z) = \frac{W(z)}{M^2} \quad \text{at any } z.$$  \hspace{1cm} (2.28)

When calculating beam propagation through optical elements, lenses or mirrors for example, one makes use of this hypothetical embedded Gaussian beam. By finding properties as focal points and beam sizes for the embedded Gaussian beam, one can assume that the real beam propagates through the system in exactly the same way. Only the beam sizes $W_x(z)$ and $W_y(z)$ of the real beam will be $M_x$-times, respectively $M_y$-times larger than the beam size $w_{em}(z)$ found for the embedded Gaussian beam. In Fig.2.4 this concept is depicted schematically.
2 Theoretical concepts

2.4 Propagation of laser beams

In this section, the theory of propagating arbitrary beams in free space is briefly presented. As laser beams are almost always well enough collimated, their diffraction properties may be described using a scalar wave theory [16]. Furthermore, the paraxial wave approximation is applicable.

2.4.1 The paraxial wave equation

One fundamental way of analyzing free-space wave propagation is deduced by using the paraxial wave equation. It is derived as follows. Any electromagnetic field \( E(x, y, z) \) in free space is governed by the scalar wave equation 2.29 which follows directly from the fundamental Maxwell equations.

\[
(\nabla^2 + k^2)E(x, y, z) = 0 \quad (2.29)
\]

Consider a beam which is primarily propagating in \( z \)-direction. Then the primary spatial dependence of the electric phasor amplitude \( E(x, y, z) \) is \( \exp(-ikz) \) with a spatial period of one wavelength \( \lambda \) in \( z \)-direction. Extracting this factor out of \( E(x, y, z) \) leads to the form

\[
E(x, y, z) = u(x, y, z)e^{-ikz}. \quad (2.30)
\]

\( u(x, y, z) \) is a complex scalar wave amplitude describing the transverse profile of the beam. Substituting this in the scalar wave equation 2.29 leads to the reduced equation

\[
\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} - 2ik\frac{\partial u}{\partial z} = 0. \quad (2.31)
\]

Due to diffraction and propagation effects, the transverse amplitude and phase profile \( u(x, y, z) \) will change slowly with distance \( z \). Compared to the plane-wave variation \( \exp(-ikz) \) in \( z \)-direction, both the transverse variations across any plane \( z \) and especially the variation in beam profile with distance along \( z \) will be slow for
2.4 Propagation of laser beams

reasonably well-collimated beams. Mathematically, this slowly varying dependence of $u(x, y, z)$ on $z$ is expressed by the paraxial approximation [16].

$$\left| \frac{\partial^2 u}{\partial z^2} \right| << \left| 2k \frac{\partial u}{\partial z} \right| \quad \text{or} \quad \left| \frac{\partial^2 u}{\partial x^2} \right| \quad \text{or} \quad \left| \frac{\partial^2 u}{\partial y^2} \right|$$

(2.32)

Thus, the second partial derivative in $z$ can be dropped and the exact wave equation 2.29 is reduced to the paraxial wave equation.

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} - 2ik \frac{\partial u}{\partial z} = 0$$

(2.33)

By introducing the operator $\nabla^2_t u(x, y, z)$, denoting the laplacian operator operating only on the coordinates in the transverse plane, and reshaping this equation, the more common form of the paraxial wave equation is given by

$$\frac{\partial u(x, y, z)}{\partial z} = -\frac{i}{2k} \nabla^2_t u(x, y, z).$$

(2.34)

Forward propagation and diffraction spreading of an arbitrary paraxial beam can be computed by integrating this equation forward in $z$-direction [16]. With a given transverse derivative $\nabla^2_t u(x, y, z)$ at a given plane $z$, one can compute the forward propagation to a new plane $z + \Delta z$. Thus, numerical forward propagation of an arbitrary wavefront is accomplished.

2.4.2 Huygens’ integral

Paraxial-spherical waves

Another effective way of analyzing paraxial wave propagation is to employ "Huygens’ principle, expressed in the Fresnel approximation" [16]. To study this concept, note that one very general solution of the exact wave equation corresponds physically to a uniform spherical wave diverging from a point source $\vec{r}_0$. This may be written in the form

$$E(\vec{r}, \vec{r}_0) = \frac{\exp[-ik\rho(\vec{r}, \vec{r}_0)]}{\rho(\vec{r}, \vec{r}_0)}.$$ 

(2.35)

$E(\vec{r}, \vec{r}_0)$ denotes the field at point $\vec{r}$ diverging from a source at point $\vec{r}_0$. $\rho(\vec{r}, \vec{r}_0)$ means the distance from the source $\vec{r}_0$ to the observation point $\vec{r}$.

$$\rho(\vec{r}, \vec{r}_0) = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}$$

(2.36)

Next, consider a point source that is not too far off the $z$-axis. Furthermore, suppose the field distribution $E(x, y, z)$ of this spherical wave being on some transverse plane $(x, y)$ that is farther along the $z$-axis for values $x$ and $y$ that are also not too far off
the $z$-axis. Then the distance $\rho(\vec{r}, \vec{r}_0)$ can be expanded in a power series in the form

$$\rho(\vec{r}, \vec{r}_0) = z - z_0 + \frac{(x - x_0)^2 + (y - y_0)^2}{2(z - z_0)} + \ldots.$$  \hspace{1cm} (2.37)

In this expression, all terms higher than quadratic can be dropped. In the $1/\rho(\vec{r}, \vec{r}_0)$ denominator in equation 2.35 even the quadratic terms will be dropped, replacing $\rho(\vec{r}, \vec{r}_0)$ simply by $z - z_0$. These approximations are called "Fresnel approximations" [16] which allow to convert equation 2.35 in

$$E(x, y, z) \approx \frac{1}{z - z_0} \exp[-ik(z - z_0) - ik\frac{(x - x_0)^2 + (y - y_0)^2}{2(z - z_0)}].$$  \hspace{1cm} (2.38)

This expression is known as the "paraxial-spherical wave" and is an exact analytical solution to the paraxial wave equation.

**Derivation of Huygens’ integral**

Huygens’ integral is a basic principle which says that if an incident field distribution $E_0(x_0, y_0, z_0)$ is given over some closed surface $S_0$, each point of that surface may by regarded as the source for a uniform spherical wave or "Huygens' wavelet" [16]. The total field at any other point $(x, y, z)$ beyond the surface $S_0$ can then be calculated by summing the fields of all Huygens’ wavelets coming from all the points on the surface $S_0$.

These ideas were put into more formal mathematical form by Fresnel and Kirchhoff, later on also by Rayleigh and Sommerfeld. In principle, each of the Huygens’ wavelets should be viewed as a spherical wave, which leads to Huygens’s integral [16]

$$E(x, y, z) = \frac{i}{\lambda} \int \int_{S_0} E_0(x_0, y_0, z_0) \frac{\exp[-ik\rho(\vec{r}, \vec{r}_0)]}{\rho(\vec{r}, \vec{r}_0)} \cos \theta(\vec{r}, \vec{r}_0) dS_0.$$  \hspace{1cm} (2.39)

d$S_0$ is an incremental element of surface area at point $(x_0, y_0, z_0)$ on the surface $S_0$. The factor $\cos \theta(\vec{r}, \vec{r}_0)$ is an "obliquity factor" and depends on the angle $\theta(\vec{r}, \vec{r}_0)$ between the line element $\rho(\vec{r}, \vec{r}_0)$ and the normal to the surface element $dS_0$.

**The Fresnel approximation to Huygens’ integral**

If now the beam profile of a paraxial optical beam across an input transverse plane is known, the output beam profile across another plane at distance $z$ can be calculated. Thereby, the Huygens’s principle is used, except that the exact spherical wave is replaced with a paraxial-spherical wave. Thus, the term $\exp(-ik\rho)/\rho$ in equation 2.39 is replaced by the paraxial-spherical wave derived in equation 2.38. Thereafter,
2.5 Mathematical description of wavefronts

Huygens’s integral in the Fresnel approximation reads [16]

\[ E(x, y, z) \approx i e^{-ik(z-z_0)} \lambda \int \int E_0(x_0, y_0, z_0) \exp \left[ -ik \left( \frac{(x-x_0)^2 + (y-y_0)^2}{2(z-z_0)} \right) \right] dx_0 dy_0. \]  

\[ (2.40) \]

2.5 Mathematical description of wavefronts

Generally, one is only interested in that part of a wavefront that interacts with certain components and therefore influences the result of an experiment or the performance of an optical device. Hence, only this part is referred to as the wavefront as depicted in Fig.2.5.

![Wavefront](image)

**Figure 2.5:** Patch of a wavefront selected by a square aperture

Considering the wavefronts of laser beams, this area is mostly defined by the transversal intensity distribution namely the beam sizes \( W_x \) and \( W_y \) as defined in equation 2.27.

The wavefront can be described mathematically by a surface \( W \) defined by the phase values as a function of \( x \) and \( y \) at a certain \( z \)-position assuming that the propagation direction is along the \( z \)-axis.

\[ W = F(x, y) \]  

\[ (2.41) \]

\( F(x, y) \) is a function defined over the aperture \( \Omega \). As mentioned above, \( \Omega \) might be the beam’s cross section itself or the aperture of an observed component.

There are two ways to describe mathematically the surface \( W \), either discrete or continuous [17]. Discrete description refers to the zonal approach. Herein, the wavefront \( W \) is calculated independently at each sampling point, which is therefore also called local method. On the other hand, continuous description refers to the modal approach that uses a full set of basis functions and reconstructs the wavefront using the whole set of sampling points at once. In style of the above mentioned local method, this is also known as a global reconstruction method.
2 Theoretical concepts

2.5.1 Zonal approach

Zonal description of a wavefront uses a discrete set of points in space (vertexes). On these vertexes which are distributed over the aperture $\Omega$, the wavefront is defined as height values. The wavefront is then approximated by connecting these height values. As in Fig.2.6 depicted, the wavefront becomes smoother with increasing number of vertexes.

![Figure 2.6: Zonal approach: Representation of a wavefront with different numbers of vertexes.](image)

The sampling theorem implies that in order to describe any spatial frequency, the vertex spacing has to have at least twice this frequency. Thus, the accuracy of wavefront representation depends on the number of vertexes sampling it. That means, the more vertexes (sampling points) the higher accuracy can be reached.

2.5.2 Modal approach

The modal approach uses two-dimensional continuous functions, so-called "spatial modes". Weighted superposition of these spatial modes yields the wavefront’s shape. Fig.2.7 illustrates the idea of this concept by superposition of 1-dimensional spatial modes.

![Figure 2.7: Superposition of 1-dimensional spatial modes resulting in the desired function](image)
2.5 Mathematical description of wavefronts

Mathematically, the wavefront can be expressed as in equation 2.42. $f_i(x, y)$ is a set of basis functions which are defined over the aperture $\Omega$. The coefficient $c_i$ defines the weight of the basis function $f_i(x, y)$ of order $i$.

$$W = \sum_{i=0}^{n-1} c_i f_i(x, y), \quad \forall (x, y) \supset \Omega$$

(2.42)

Possible basis functions are e.g. Zernike polynomials or sine and cosine functions resulting in Fourier reconstruction as proposed by Poyneer [18]. The Zernike polynomials which are extensively used to describe the measured wavefronts in this thesis, are discussed below.

Zernike polynomials

Zernike polynomials are often used to represent wavefronts in polynomial form and have several important properties. They are a complete set of basis functions of two real variables $\rho$ and $\theta$ that are orthogonal on the interior of a unit circle. Here $\theta$ is the azimuthal angle and $\rho$ is the radial distance ($0 < \rho < 1$). Since the polynomials are made up of terms that are of the "same form as the types of aberrations often observed in optical tests" [19], their interpretation is easy. Most aberrations occurring in optical tests can be described by low-order Zernike polynomials (primary aberrations). These low order terms correspond to classical aberrations like astigmatism or coma for example. An arbitrary wavefront can be expressed by

$$W = \sum_{i=0}^{\infty} c_i Z_i(\rho, \theta)$$

(2.43)

where $Z_i(\rho, \theta)$ are the Zernike polynomials which can be divided in a radial and an azimuthal part.

$$Z_i(\rho, \theta) \equiv Z_n^m(\rho, \theta) = \begin{cases} R_n^m(\rho) \cos(m\theta), & m > 0 \\ R_n^m(\rho) \sin(m\theta), & m < 0 \\ R_n^m(\rho), & m = 0 \end{cases}$$

(2.44)

The polynomials with negative $m$ are called odd, the polynomials with positive $m$ are called even polynomials respectively. For simplicity, $n$ and $m$ are often combined to a single index $i$ called mode ordering number. Thereby, following rules have to be complied:

$$n \in \mathbb{N}_0, \quad m \in \mathbb{Z}$$

for a given $n$, $m$ ranges from $-n$ to $n$ with stepsize $2$

The index $i$ derived from $n$ and $m$ numbers consecutively the modes of the Zernike polynomials. There is one mode for $n = 0$, two for $n = 1$, three for $n = 2$, four for
\[ i_n = \frac{(n+1)(n+2)}{2} \tag{2.45} \]

The radial term \( R_n^m \) is given by
\[
R_n^m (\rho) = \begin{cases} 
\sum_{k=0}^{(n-m)/2} \frac{(-1)^k (n-k)!}{k! [n+m-k]! [n-m-k]!} \rho^{n-2k} & \text{for } n - m \text{ even} \\
0 & \text{for } n - m \text{ odd}
\end{cases}
\tag{2.46}
\]

Orthogonality of Zernike polynomials means that the inner product of any two Zernike polynomials \( Z_n^m (\rho, \theta) \) is zero when the inner product is performed over the unit circle. The inner product is the integral of the product of one polynomial times the complex conjugate of the other over the specified limits of orthogonality.
\[
\int_0^1 \int_0^{2\pi} Z_n^m (\rho, \theta) Z_{n'}^{m'} (\rho, \theta) \rho \, d\rho \, d\theta = \frac{\pi (1 + \delta_{m,0})}{2n+1} \delta_{n,n'} \delta_{m,m'} \tag{2.47}
\]

To give an example, Fig.2.8 shows some Zernike polynomials with the corresponding coefficients and mode ordering number.

\textbf{Figure 2.8:} Examples of low-order Zernike polynomials

\subsection*{2.6 High-harmonic generation}

High-harmonic generation (HHG) is a phenomenon which occurs during the interaction between intense laser pulses with rare gas atoms. It results in the generation of higher odd harmonics of the driving laser frequency. In this process, photons of high
2.6 High-harmonic generation

energies (high frequencies) can be generated. HHG in gases was first observed in 1988 [20]. It is a highly non-linear effect which requires high laser intensities in the order of $10^{14}$ W/cm$^2$ or higher. Mostly femtosecond lasers are used which produce very short and intense pulses. For example, wavelength down to the XUV spectral range can be reached with a fundamental laser in the visible.

The physics of atoms in strong laser fields has been a subject of intensive studies in the last decades. A semiclassical model has been developed to describe the fundamental mechanism of HHG. During the laser-atom interaction, ionization can occur close to the peak of the laser electric field. In this case, an electron tunnels through the potential barrier formed by the combined Coulomb and laser field. Further, the electron is accelerated in the electric field of the laser. The electron oscillates in the laser field, gaining kinetic energy. Eventually, the electron again hits the parent ion. Doing so, different processes can occur. One possibility is that the electron rescatters elastically, which leads to highly energetic electrons. This is called above threshold ionization (ATI). Another possibility is that the electron rescatters inelastically and frees another electron. This phenomenon is called non-sequential double ionization (NSDI). The third possible phenomenon is high-harmonic generation. Hereby, the electron converts its gained energy into a high-energy photon by recombination with the parent ion. The model described above is often called three-step-model because it consists of three different stages, ionization, acceleration, and recombination. In Fig.2.9 the process of HHG within the three-step-model is schematically shown.

![Three step model of HHG](image)

**Figure 2.9:** Three step model of HHG

As this process occurs twice the laser oscillation period, only odd higher harmonics of wavelength $\lambda_n$ are generated.

$$\lambda_n = \frac{\lambda_0}{n} \quad \text{with} \quad n = 3, 5, 7, \ldots$$  \hspace{1cm} (2.48)
Hereby, $\lambda_0$ is the wavelength of the fundamental drive laser. High-order harmonics up to $n > 100$ can be produced. A typical HHG frequency spectrum like in Fig. 2.10 shows a characteristic comb-form.

![Figure 2.10: Typical HHG spectrum for neon gas. (taken from [21])](image)

The intensity of the harmonics decreases with the harmonics number in the low order region which is followed by the so-called "plateau" where the harmonics intensity remains approximately constant. The plateau region spans over many harmonics down to the X-ray regime [22]. Finally, the plateau is sharply terminated by the "cutoff"-frequency which depends on the laser intensity and the ionization potential of the atom.

Also in the time domain, HHG radiation forms a comb resulting in an overall pulse shape. This pulse is composed of a number of peaks separated by half a driving laser period. See Fig. 2.11 for a simulated time profile of an HHG pulse.
2.7 Basic theory on free-electron lasers

2.7.1 Comparison of FELs to synchrotrons

In bending magnets synchrotron radiation is emitted when charged particles with relativistic energies $W$ are transversally accelerated. Mostly electrons are used for that purpose because they have a small rest mass $m_e$. In a synchrotron or storage ring the electrons are accelerated toward the center of the ring. For a Lorentz factor $\gamma = W/(m_e c^2) \gg 1$ the electrons emit synchrotron radiation tangentially to their orbit [23]. See Fig. 2.12 for a schematic sketch.

**Figure 2.11:** Simulated HHG time profile of Fig. 2.10. (taken from [21])

**Figure 2.12:** Synchrotron radiation emitted by a circulating relativistic electron
Theoretical concepts

The typical opening angle of the cone of emitted radiation can be given by

\[ \theta = \frac{1}{\gamma}. \] (2.49)

The radiation is incoherent because the electrons in a bunch radiate independently from each other. The phases do not have any correlation to each other and therefore some of the emitted electric field will be canceled out. The frequency spectrum extends beyond the critical frequency \( \omega_c = \frac{3e^2}{2\pi m c} \) which is a characteristic quantity of the emitted radiation separating the frequency spectrum into two equal parts of emitted energy. Herein, \( R \) is the radius of curvature in the bending magnets. The integrated radiation power is

\[ P = \frac{e^2 c \gamma^4}{6\pi \varepsilon_0 R^2}, \] (2.50)

with the vacuum permittivity \( \varepsilon_0 \).

In contrast to bending magnets, the spectrum of the undulator radiation consists of narrow spectral lines and is emitted in a narrow cone with a typical opening angle \( K/\gamma \) [7]. The dimensionless quantity \( K = \frac{eB_0\lambda_u}{2\pi m c} \) is called undulator parameter which only depends on the peak magnetic field \( B_0 \), the undulator period \( \lambda_u \), and the type of used particles. A schematic sketch of an undulator is given in Fig. 2.13.

![Schematic view of an undulator magnet. Radiation is emitted in the forward direction.](image)

**Figure 2.13:** Schematic view of an undulator magnet. Radiation is emitted in the forward direction.

The fundamental wavelength can be estimated to

\[ \lambda_f = \frac{\lambda_u}{2\gamma^2} \left( 1 + \frac{K^2}{2} \right). \] (2.51)

The on-axis undulator radiation only contains odd higher harmonics with wavelengths

\[ \lambda_m = \frac{\lambda_f}{m}, \quad m = 1, 3, 5, 7, \ldots. \] (2.52)
The directional variation of the electron orbit can be calculated to
\[ \theta_{\text{max}} = \frac{K}{\beta \gamma} \approx \frac{K}{\gamma}. \] (2.53)

In the case when the $K$-parameter is smaller than 1, the opening angle of emitted radiation is smaller than the natural opening angle of the synchrotron radiation. Therefore, all of the radiation cones overlap and interfere with each other, what results in sharp coherent spectral lines.

### 2.7.2 Interaction between electron beam and light wave

To describe the basic principles how an FEL works, we have to take a closer look on the processes in the undulator. As already mentioned, relativistic electrons undergo a wavelike trajectory due to the alternating magnetic field. This was already depicted schematically in Fig.2.13. Note that in this sketch, for simplicity, the deflections are indicated in the vertical plane, but in reality the electrons oscillate perpendicular to the magnetic field. The motion of the electrons in the undulator can be described as follows [7].

\[ x(z) = \frac{K}{\gamma k_u} \sin(\omega_u t), \quad z(t) = \bar{v}_z t - \frac{K^2}{8 \gamma^2 k_u} \sin(2 \omega_u t), \] (2.54)

with the average longitudinal speed $\bar{v}_z$

\[ \bar{v}_z = \left( 1 - \frac{1}{2 \gamma^2} \left( 1 + \frac{K^2}{2} \right) \right) c \equiv \beta c, \] (2.55)

the wave-number $k_u$ and the angular frequency $\omega_u$.

\[ k_u = \frac{2 \pi}{\lambda_u} \quad \omega_u = \beta c k_u \] (2.56)

Now assume that a light wave is co-propagating with the relativistic electron beam, which can be described in the simplest case by a plane electromagnetic wave with a fixed amplitude. Hereby the initial light wave either is produced by noise (SASE) or it comes from an external laser (seeding).

\[ E_x(z, t) = E_0 \cos(k_l z - \omega_l t + \psi_0), \] (2.57)

with

\[ k_l = \omega_l / c = 2 \pi / \lambda_l. \] (2.58)

The time derivative of the electron energy $W_e$ is

\[ \frac{\partial W_e}{\partial t} = \bar{v} \cdot \bar{F} = -e v_x(t) E_x(t). \] (2.59)
In the formula above one can see that an electron loses energy when \( \frac{dW}{dt} < 0 \). The law of energy conservation implies that this energy is not lost but transferred to the light wave. Thus, if \( \frac{dW}{dt} \) stays negative, a constant energy transfer is guaranteed. Thereby, \( v_x \) and \( E_x \) have to point in the same direction at any time. As \( v_x \) changes its sign according to the sinusoidal electron trajectory, the electric field vector has to change its direction with the same temporal period. This is achieved if the light wave advances by \( \lambda/2 \) per half a period of electron trajectory. In Fig.2.14 one can proof this schematically. The electrons have a speed almost equal to \( c \), but the overall longitudinal speed is smaller than \( c \) due to the transversal deflections. Accordingly, it is possible for the light to pass the electrons by the right amount to ensure constant energy transfer.

![Figure 2.14: Condition for sustained energy transfer](image)

This condition is only fulfilled with certain radiation wavelengths \( \lambda_m \) from which, as stated in equation 2.51, the fundamental is

\[
\lambda_l = \frac{\lambda_m}{2\gamma^2} \left( 1 + \frac{K^2}{2} \right).
\] (2.60)

**Microbunching**

Microbunching is a periodic charge density modulation of an electron bunch. The microbunching is a result of the interaction between the electromagnetic radiation and the electrons in high-gain FELs. Due to this interaction, in dependence of their ponderomotive phase, some of the electrons lose energy and some of them gain energy. The energy change in turn modifies the particles’ longitudinal velocity, what combined with the sinusoidal transverse motion ends up in a charge density
modulation of the electron bunch. The bunch is split in slices which are shorter than
the emitted wavelength $\lambda_l$. All $N$ electrons in a slice can now radiate coherently.
One can interpret a slice as a high-charged particle with charge $Q = N \cdot e^-$, whose
emitted intensity grows quadratically with the number of coherently acting electrons
$N$: $I_N = N^2 \cdot I_1$. In contrast to that, the intensity of non-coherently acting particles
grows linearly with $N$: $I_N = N \cdot I_1$. Microbunching is a fundamental feature of the
high-gain FEL and the reason why FELs can reach such high brilliances.

2.7.3 One-dimensional description of FELs

The one-dimensional third-order equation

In the one-dimensional theory the transverse variation of electron bunch charge
density and electromagnetic fields is ignored. Furthermore, the electron bunches
are treated as very long, so the effects at the head and tail of a bunch can be
neglected. For the explanation of the basic principle of an FEL in section 2.7.2,
the co-propagating wave was described by a plane wave with constant amplitude
(equation 2.57). However, this model is not suitable to describe the operation of the
high-gain FELs, where saturation can be achieved in a single pass of the electrons
through the undulator. One needs a model which takes into account the variation
of the co-propagating wave's amplitude with $z$ as well as the microbunching effect.

\[
\tilde{E}_x(z,t) = \tilde{E}_x(z)\exp[i(k_l z - \omega_l t)]
\]

(2.61)

Hereby, complex quantities are indicated with a tilde. In the following the analytical
solution in the one-dimensional theory is presented and discussed. The complete
mathematical derivation can be looked up elsewhere in many theoretical books.
Worth mentioning is the book "The physics of free electron lasers - An introduction"
of E.L. Saldin, E.A. Schneidmiller and M.V. Yurkov [24], but also a good
introduction to free-electron laser physics is given in the book from P. Schm"user, M.
Dohhus and J. Rossbach "Ultraviolet and Soft-X-Ray Free-Electron Lasers" [7].

Starting with the above mentioned assumptions, the basic analytic description of
an FEL is obtained. Thereby, also microbunching was taken into account. As a
result, one gets the so-called "one-dimensional third-order equation" which reads as follows.

\[
\frac{\tilde{E}_x'''(z)}{\Gamma^3} + 2i \frac{\eta}{\rho_{FEL}} \frac{\tilde{E}_x''(z)}{\Gamma^2} + \left( \frac{k_p^2}{\Gamma^2} - \left( \frac{\eta}{\rho_{FEL}} \right)^2 \right) \frac{\tilde{E}_x'(z)}{\Gamma} - i\tilde{E}_x = 0
\]

(2.62)

$\eta$ is the relative electron energy deviation from the resonance energy $W_r$.

\[
\eta = \frac{W - W_r}{W_r} = \frac{\gamma - \gamma_r}{\gamma_r}, \quad W_r = \gamma_r m_e c^2
\]

(2.63)
2 Theoretical concepts

The resonance energy is defined by the wavelength of the co-propagating wave according to the following formula.

\[ \lambda_l = \frac{\lambda_u}{2\gamma_r^2} \left( 1 + \frac{K^2}{2} \right) \]

\[ \Rightarrow \gamma_r = \sqrt{\frac{\lambda_u}{2\lambda_l} \left( 1 + \frac{K^2}{2} \right)} \quad (2.64) \]

\( \Gamma \) is the so-called "gain parameter"

\[ \Gamma = \left[ \frac{\mu_0 \bar{K}^2 e^2 k_u n_e}{4\gamma_r^3 m_e} \right]^{1/3} \quad (2.65) \]

with modified undulator parameter \( \bar{K} \)

\[ \bar{K} = K \cdot \left[ J_0 \left( \frac{K^2}{4 + 2K^2} \right) - J_1 \left( \frac{K^2}{4 + 2K^2} \right) \right]. \quad (2.66) \]

Here, \( J_0(x) \) and \( J_1(x) \) are components of the Fourier-Bessel series. \( k_p \) is the space charge parameter

\[ k_p = \sqrt{\frac{2k_u \gamma_r m_e \omega_l}{\gamma_r m_e \omega_l}} = \sqrt{\frac{2\lambda_l}{\lambda_u} \cdot \frac{\omega_p^*}{c}}. \quad (2.67) \]

with the plasma frequency

\[ \omega_p^* = \sqrt{\frac{n_e e^2}{\gamma_r \varepsilon_0 m_e}}. \quad (2.68) \]

Both coefficients \( \Gamma \) and \( k_p \) have the dimension of an inverse length and depend on the layout of the undulator and the beam properties. \( \rho_{FEL} \) is called "FEL parameter" or "Pierce parameter".

\[ \rho_{FEL} = \frac{\Gamma}{2k_u} = \frac{1}{4\pi \sqrt{3} \cdot \lambda_u} \cdot \frac{\lambda_u}{L_g} \quad (2.69) \]

Here, the power gain length \( L_g \) was introduced which is a very important figure of merit for high-gain FELs. It describes the length along \( z \), in which the radiation power grows by a factor of \( e \).

**General solution of the one-dimensional third-order equation**

The third-order equation 2.62 can be solved by using the trial function \( \tilde{E}_x(z) = Ae^{\alpha z} \). This results in an cubic equation for the exponent \( \alpha \) with three different solutions \( \alpha_1, \alpha_2, \) and \( \alpha_3 \). The general solution can be written as a linear combination of the three eigenfunctions \( V_j(z) = \exp(\alpha_j z) \).

\[ \tilde{E}_x(z) = c_1 V_1(z) + c_2 V_2(z) + c_3 V_3(z), \quad V_j(z) = \exp(\alpha_j z) \quad (2.70) \]

The first and second derivatives read as follows:

\[ \tilde{E}_x'(z) = c_1 \alpha_1 V_1(z) + c_2 \alpha_2 V_2(z) + c_3 \alpha_3 V_3(z) \quad (2.71) \]

\[ \tilde{E}_x''(z) = c_1 \alpha_1^2 V_1(z) + c_2 \alpha_2^2 V_2(z) + c_3 \alpha_3^2 V_3(z) \quad (2.72) \]
The coefficients $c_j$ can be computed by specifying the initial conditions for $\tilde{E}_x(z)$, $\tilde{E}_x'(z)$, and $\tilde{E}_x''(z)$ at the beginning of the undulator at $z = 0$. This is expressed in matrix form like this:

$$\begin{pmatrix} \tilde{E}_x(0) \\ \tilde{E}_x'(0) \\ \tilde{E}_x''(0) \end{pmatrix} = \mathbf{A} \cdot \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \quad \text{with} \quad \mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ \alpha_1 & \alpha_2 & \alpha_3 \\ \alpha_1^2 & \alpha_2^2 & \alpha_3^2 \end{pmatrix} \quad (2.73)$$

Following, the coefficient-vector can be computed by taking the inverse of matrix $\mathbf{A}$.

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \mathbf{A}^{-1} \cdot \begin{pmatrix} \tilde{E}_x(0) \\ \tilde{E}_x'(0) \\ \tilde{E}_x''(0) \end{pmatrix} \quad (2.74)$$

In the simplest case, we assume that the energy deviation of the electron beam is 0. Moreover, one can assume $k_p = 0$ because the plasma frequency gets low for high electron energies and low electron density. With this in mind, the third-order equation simplifies to

$$\frac{\tilde{E}_x'''}{\Gamma^3} - i\tilde{E}_x = 0. \quad (2.75)$$

Using the trial function $\tilde{E}_x(z) = Ae^{\alpha z}$, one obtains

$$\alpha^2 = i\Gamma^3. \quad (2.76)$$

This has three solutions, namely

$$\alpha_1 = (i + \sqrt{3}) \cdot \frac{\Gamma}{2}, \quad \alpha_2 = (i - \sqrt{3}) \cdot \frac{\Gamma}{2}, \quad \alpha_3 = -i\Gamma. \quad (2.77)$$

Only the first solution has a positive real part which is leading to an exponential growth of the field. The imaginary part causes oscillations which stay constant in amplitude. The second solution corresponds to damping because of the negative real part. The third solution only contains an oscillatory part. Now, if $z$ is sufficiently large, the power of the light grows as

$$P(z) \propto \exp(2\Re\{\alpha_1\}) = \exp(\sqrt{3}\Gamma z) \equiv \exp(z/L_{g0}) \quad (2.78)$$

with the gain length $L_{g0}$ indicating that this result holds only for the one-dimensional theory. Here a connection between the gain length parameter and the gain parameter can be made.

$$L_{g0} = \frac{1}{\sqrt{3}\Gamma} \quad (2.79)$$
2 Theoretical concepts

2.7.4 One-dimensional solution for seeding

Now we can have a closer look at the solution if the FEL is seeded. For that, we take the obtained results of a beam with $\eta = 0$ and $k_p = 0$. Then the matrix $A$ is given by

$$A = \begin{pmatrix}
1 & 1 & 1 \\
(i + \sqrt{3})\Gamma/2 & (i - \sqrt{3})\Gamma/2 & -i\Gamma \\
(i + \sqrt{3})^2\Gamma^2/4 & (i - \sqrt{3})^2\Gamma^2/4 & \Gamma^2
\end{pmatrix} \quad (2.80)$$

and the inverse matrix $A^{-1}$ is given by

$$A^{-1} = \frac{1}{3} \begin{pmatrix}
1 & (\sqrt{3} - i)/(2\Gamma) & (-i\sqrt{3} + 1)/(2\Gamma^2) \\
1 & (\sqrt{3} - i)/(2\Gamma) & (-i\sqrt{3} + 1)/(2\Gamma^2) \\
1 & i/\Gamma & -1/\Gamma^2
\end{pmatrix} \quad (2.81)$$

As we assume the FEL is seeded, there is an initial plane wave generated by an external laser. This incident light wave with wavelength $\lambda_i$ is described by

$$E_x(z,t) = E_{seed} \cos(k_l z - \omega_l t), \quad \text{with} \quad k_l = \frac{\omega_l}{c} = \frac{2\pi}{\lambda_i}. \quad (2.82)$$

So $\tilde{E}_x(0) = E_{seed}$ while $\tilde{E}'_x(0) = 0$ because the seed has no density modulation. Also $\tilde{E}''_x(0)$ vanishes at the beginning of the undulator.

$$\begin{pmatrix}
\tilde{E}_x(0) \\
\tilde{E}'_x(0) \\
\tilde{E}''_x(0)
\end{pmatrix} = \begin{pmatrix}
E_{seed} \\
0 \\
0
\end{pmatrix} \quad (2.83)$$

Applying equation 2.74 gives the coefficients which in this case are all equal.

$$c_j = E_{in}/3 \quad (2.84)$$

Inserting the coefficients in equation 2.70 yields the electric field of the FEL along the undulator.

$$\tilde{E}_x(z) = \frac{E_{seed}}{3} \left[ \exp((i + \sqrt{3})\Gamma z/2) + \exp((i - \sqrt{3})\Gamma z/2) + \exp(-i\Gamma z) \right] \quad (2.85)$$

In Fig.2.15 the power growth is depicted as a function of $z$. One can see that during the first gain lengths $L_{g0}$ the power stays almost constant. This is called "lethargy regime" and comes from the damping and oscillatory term in equation 2.85. But after some gain lengths $L_{g0}$, the power grows as

$$P(z) \approx \frac{P_{seed}}{9} \exp\left(\frac{z}{L_{g0}}\right). \quad (2.86)$$

Hereby, $P_{seed}$ is the power of the seed laser. For FELs it is a quite typical behavior that the starting value of the exponential function is one ninth the incident power [7].
2.7 Basic theory on free-electron lasers

In the one-dimensional theory, the exponential growth continues infinitely. This is surely not true for the real case. Different effects lead to an upper limit for the reachable radiation power. For example, the energy of the electron beam decreases because it loses energy to the radiation. The average velocity of the electrons decreases, which leads to a mismatching of the resonance condition discussed in section 2.7.2. Therefore, after a certain length, the saturation regime is reached. Here, the radiation power oscillates around a certain value and stays almost constant. Note that thereby the saturation power does not depend on the initial power of the seed.

Figure 2.15: Computed power rise in the one-dimensional FEL theory for a seeded FEL
3 Experimental setup of sFLASH

3.1 sFLASH HHG source

For HHG, a high-power laser is needed. At sFLASH a pulsed laser system at a repetition rate of 10 Hz with a near-infrared wavelength of 800 nm, pulse-lengths of about 35 fs, and an energy of about 20 mJ is used. The source setup is shown in Fig. 3.1. The 800 nm laser is focused with a mirror (\(f = 3\) m) into a gas target within an ultra-high vacuum chamber. As the produced high harmonics are in the XUV spectral range, they can only propagate in vacuum. The 800 nm pulses have to enter the vacuum of about \(10^{-5}\) mbar through a vacuum window.

![Figure 3.1: Beam line of the 800 nm drive laser](image)

The HHG source itself is a noble-gas jet (argon) which is pulsed with the same repetition rate as the drive laser. For seeding only the 21. harmonic at 38 nm is used. Therefore, multilayer mirrors sort out other wavelength components and the fundamental IR-beam is blocked with a thin (200 nm) Al-filter.
3 Experimental setup of sFLASH

3.2 Injection beam line

The sFLASH beamline for the XUV seed starts at the HHG source. The HHG source together with the IR-laser system is located in an experimental hutch next to the FLASH tunnel. Fig. 3.2 shows the hutch and the FLASH tunnel schematically.

![Figure 3.2: Experimental hutch with HHG source besides the FLASH tunnel. (taken from [10])](image)

The HHG source itself is located in a pit from which a horizontal connection goes to the tunnel. A picture of this pit containing the HHG source is shown in Fig. 3.3. After the gas jet, the seed goes into the FLASH tunnel where a plane UHV mirror located at a distance of about 8.6 m from the source deflects the beam upwards to the electron-beam level. Thereafter, a mirror chamber is reached. In order to focus the seed into the undulator, multilayer normal incidence spherical mirrors are used. Three different focal lengths can be inserted. For this purpose, the mirrors with different radii are located on a wheel which can be remotely controlled. Fig. 3.4 shows a picture of that wheel with in total six mirrors. There are two mirrors for each focal length, but with different coatings. This gives options to use two different wavelengths (Si/Sc and Mo/Si multilayer coating) and three different focal lengths ($f_1 = 6.25$ m; $f_2 = 7$ m; $f_3 = 8.5$ m). But currently only the Si/Sc multilayer for the 38 nm seed is in use. For seeding the electron bunch, the seed has to overlap with the electrons and to propagate parallel to them through the undulators. Therefore, the seed is deflected on the electron beam axis by a triplet mirror. Triplet mirrors have to be used because only at grazing incidence enough power can be reflected. With three mirrors an overall reflection of 90° is reached in order to steer the seed parallel to the electron beam. A sketch of the sFLASH beam line from HHG source to undulators is given in Fig. 3.5.
3.2 Injection beam line

Figure 3.3: Pit with HHG source inside the experimental hutch

Figure 3.4: 6-fold focal mirror wheel. (taken from [25])

Figure 3.5: sFLASH injection beam line
3 Experimental setup of sFLASH

3.3 Diagnostic beam line

The wavefront sensor cannot be placed directly into the sFLASH injection beam line because of space limitations. This problem was circumvented by placing the sensor in the diagnostic section inside the same experimental hutch where the source is located. Fig.3.6 shows schematically the HHG source and the diagnostic section. Close to the point where the seed enters the FLASH tunnel, there is a movable remote controlled triplet mirror to reflect the photons to the diagnostic section. Grazing incidence has to be used again due to the higher reflectivity. Therefore, a minimum of three mirrors is needed to deflect the HHG beam by 90°. The first triplet deflects the HHG vertical by 90° to reach the table level. A second triplet again reflects the beam by 90° to horizontal propagation direction parallel to the experimental table in a height of about 30 cm above table level. The beam size has to be sufficiently large on the wavefront sensor because only if enough holes on the Hartmann plate are illuminated, reasonable results can be obtained. Therefore, before reaching the sensor, the beam is folded by two multilayer mirrors.

![Diagram](image.png)

Figure 3.6: Scheme of the HHG source and diagnostic section in the experimental hutch

Another important effect of these mirrors is that they reflect only the 21st harmonic which is used for seeding. Therefore, only the wavefront of one harmonic is investigated by the sensor. All in all, the wavefront sensor is placed about 7.2 m after the HHG source.
4 Numerical simulations

In this section the simulation process and its results are presented. In order to investigate the influence of the wavefront on the FEL performance, virtual beams containing certain assumptions for the wavefront shape were created. For this purpose, Matlab scripts were written which produce a file containing a full characterization of the seed in terms of electric field phase and amplitude at a certain z-position. Following, the virtual seed is imported in ZEMAX and propagated through the sFLASH beam line of about 24m in length down to the undulator. The one-dimensional theory discussed in section 2.7.3 is insufficient for the purpose of this work as it neglects completely the transverse variations of electromagnetic waves and electron bunch densities as well as the phase variations of the seed due to non-plane wavefronts. Since analytical solutions are not available, numerical simulation codes are needed. In order to simulate the FEL process inside the undulators, the FEL-code GENESIS has been used. GENESIS is a time-dependent three-dimensional code. It focuses on the simulation of single-pass free-electron lasers. A more precise description of GENESIS can be found in the official GENESIS manual [26]. In the last simulation step, the relevant figures of merit like power along the undulator or radiation beam size can be extracted from the GENESIS output. Fig.4.1 gives an overview of the needed programs and their combination to the whole simulation scheme of an initial wavefront model.

Figure 4.1: Scheme of the simulation workflow
4 Numerical simulations

4.1 The "Physical Optics Propagation" mode in ZEMAX

ZEMAX is an optical design program produced by ZEMAX Development Corporation [27]. Mainly it is used for the design and analysis of optical systems such as camera lenses and analysis of illumination systems. It performs standard sequential and non-sequential ray tracing for analysis of stray light. But in addition to that, a "Physical Optics Propagation" feature is implemented which can be used for problems where diffraction is important, including the propagation of laser beams. As "Physical optics is the modeling of optical systems by propagating wavefronts" [28], this is exactly what is needed to simulate the seed's beam propagation through the sFLASH beam line. The "Physical Optics Propagation" mode works in the paraxial approximation where the basic principles of beam propagation were already discussed in section 2.4.

The beam is represented by a discretely sampled complex electric field \( \tilde{E}(x, y, z) \) on a cartesian grid. The phase \( \varphi \) and amplitude \( E \) at each sampling point correspond to the local absolute value and angle of \( \tilde{E}(x, y, z) \).

\[ \varphi(x, y, z) = \text{angle} \left[ \tilde{E}(x, y, z) \right] \]  \hspace{1cm} (4.1)

\[ E(x, y, z) = \text{abs} \left[ \tilde{E}(x, y, z) \right] \]  \hspace{1cm} (4.2)

The entire grid is then propagated through the free space between optical surfaces. At each optical surface a transfer function is computed which transfers the beam from one side of the optical surface to the other. Thereby, ZEMAX uses diffraction propagation whose methods are described in detail in [29] and [30]. Basically, this corresponds to propagation by Fraunhofer or Fresnel diffraction depending on the Fresnel number. Conceptually speaking, the Fresnel number is the number of annular "Fresnel zones" from the center of the beam to the edge. Fresnel zones are the radial zones where the phase as seen from the observation point changes by \( \pi \). Thereby, the Fresnel number \( F \) is defined for a rotationally symmetric beam as

\[ F = \frac{2}{\lambda} \left( \sqrt{\Delta z^2 + a^2} - \Delta z \right), \]  \hspace{1cm} (4.3)

where \( \Delta z \) is the distance from the beam to the observation point and \( a \) the radial beam size.

If the Fresnel number is smaller than 1, the beam at the observation point is said to be in the "far field" relative to the current beam and Fresnel diffraction is used which is defined by

\[ \tilde{E}(x_2, y_2, z_2) = \left[ \frac{e^{ikz}}{i\lambda\Delta z} \right] q(r_2, \Delta z) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{E}(x_1, y_1, z_1)q(r_1, \Delta z)e^{-\frac{i\pi}{\lambda} \left( x_1 x_2 + y_1 y_2 \right)} dx dy, \]  \hspace{1cm} (4.4)
4.2 Creating the wavefront model

with

$$\Delta z = z_2 - z_1, \quad r^2 = x^2 + y^2, \quad q(r, \Delta z) = \exp\left(\frac{i\pi r^2}{\lambda \Delta z}\right).$$  \hspace{1cm} (4.5)

If the Fresnel number is larger than 1, the beam at the observation point is said to be in the "near field" relative to the current beam and Fraunhofer diffraction is used.

$$\tilde{E}(x_2, y_2, z_2) = \left[\frac{e^{ikz}}{i\lambda \Delta z}\right] \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{E}(x_1, y_1, z_1)e^{-\frac{i2\pi}{\lambda \Delta z}(x_1x_2+y_1y_2)} dx dy$$ \hspace{1cm} (4.6)

ZEMAX is able to read self-defined beam files (so-called "ZBF-files") with the complex electric field $\tilde{E}(x, y, z)$. Therefore, any initial start condition regarding wavefront and intensity profile can be externally defined and afterwards propagated. After propagation, ZEMAX generates again a ZBF-file in the same format as the input file, which can be analyzed not only with ZEMAX itself but also by other programs, e.g. Matlab.

4.2 Creating the wavefront model

First of all, it had to be decided where to start the simulations along the sFLASH beam line. A $z$-position had to be found where the initial beam is inserted and simulation of beam propagation is started. Keeping in mind that measurements of the seed’s real wavefront were to come and that the FEL performance should also be simulated using the measured wavefront, the most convenient way was to make simulations adaptable to the real measurements. According to this, it was decided to start the simulation at a distance from the seed source equal to the distance from wavefront sensor to source ($\approx 7.2$ m after the gas jet). Later on, the virtual wavefront data just have to be exchanged by the measured wavefront data. This way, propagation of the real seed and following simulation of FEL performance can be done easily without extra adjustment of measurements to the simulation framework.

In order to get a reasonable model for the seed’s phase distribution, the wavefront of the IR drive-laser was measured. For this purpose, an off-the-shelf wavefront sensor was used. This sensor uses the Shack-Hartmann principle and is able to reconstruct wavefronts in the visible spectral range. See Fig.4.2 for the measured wavefront.
One can clearly see an astigmatism in this wavefront. Assuming the wavefront shape is transferred to the high harmonics, they would resemble an astigmatic wavefront. This assumption is supported by the results reported in [31]. A similar system to the HHG system for sFLASH was investigated regarding wavefronts. Here, the drive laser was affected by astigmatism too. But more important, also the measurement of the higher harmonics showed the same behavior. It seems that as a matter of fact, during the process of HHG the wavefront shape is transferred from the fundamental beam to the high harmonics.

It was decided to create the wavefront model with a phase distribution following an astigmatism which is expressed by a saddle function. Two different strengths of distortion were created in order to investigate their influence. The strength of distortion was assumed to be in the order of 0.2 $\lambda$ rms. So the simulated wavefronts have an rms distortion of 0.2 $\lambda$ and 0.4 $\lambda$. Moreover, a wavefront with a perfect Gaussian wavefront, respectively an rms distortion of 0 $\lambda$, was created serving as a reference. Note that the wavefront distortion is always added to the perfect Gaussian wavefront which is presented for $z > z_R$ by the Zernike polynomial $Z^0_2$ (defocus). This means, this component of the wavefront is not taken into account as one is only interested in the deviations from the perfect case. The intensity distribution was assumed to be Gaussian in all cases. This resembles the measurements already done on the seed. The beam size of the seed was taken from measurements at sensor position as this is the position where the simulation begins. There, the beam size was 3 mm. Fig.4.3 shows the intensity profile used for each beam. Fig.4.4 shows the modeled wavefront aberrations.
4.2 Creating the wavefront model

Figure 4.3: Initial Gaussian intensity distribution

(a) Reference (no aberration)

(b) 0.2 \( \lambda \) rms wavefront deviation

(c) 0.4 \( \lambda \) rms wavefront deviation

Figure 4.4: Simulated initial wavefront aberrations
4 Numerical simulations

4.3 Propagation through the sFLASH injection beam line

As mentioned in section 3.2, three different focal lengths of the focusing mirror can be used. To investigate the impact of the focal length, first of all only the beam sizes near the position of the first undulator were analyzed. The beam sizes were also analyzed inside the undulator area. Note that thereby the interaction between electrons and electromagnetic wave inside the undulator was not taken into account. This simulation is meant to clarify what focal length is optimal to focus the beam. The focal length of 7m showed definitely the best results which are depicted in Fig.4.5.

![Figure 4.5: Beam sizes near undulator. The focal length of the focusing mirror is 7m.](image)

The focus lies inside the undulator area. Small beam sizes at undulator begin are the crucial factor for seeding. Otherwise the seed’s power density was too low to successfully seed the FEL. All following simulations were done with a focal length of 7m. Moreover, the effects due to wavefront distortions can be directly seen in this Fig.4.5. The focus positions in \(x\) and \(y\)-direction are drifting apart the bigger the initial rms wavefront distortion is. This is a direct effect of the induced astigmatism. For a wavefront distortion of 0.4 \(\lambda\) rms, the focuses are almost 1.5 m apart from each other. But still both are lying inside the area of the first undulator. Table 4.1 shows
the calculated rms beam sizes $\sigma_x$ and $\sigma_y$ and the rms cross section $\epsilon$ at undulator begin for each wavefront model. Thereby, $\sigma_x$, $\sigma_y$, and $\epsilon$ were calculated by

$$
\sigma_x = \sqrt{\frac{\sum_{i=1}^{n} I_i (x_i - x_0)^2}{\sum_{i=1}^{n} I_i}}, \quad \sigma_y = \sqrt{\frac{\sum_{i=1}^{n} I_i (y_i - y_0)^2}{\sum_{i=1}^{n} I_i}},
$$

(4.7)

$$
\epsilon = \sqrt{\sigma_x^2 \sigma_y^2 - \left(\frac{\sum_{i=1}^{n} I_i (x_i - x_0)(y_i - y_0)}{\sum_{i=1}^{n} I_i}\right)^2}.
$$

(4.8)

$n$ is the total number of sampling points. $I_i$ denotes the intensity at sampling point $i$ and $x_0$ and $y_0$ are the mean values of the spatial distribution.

$$
x_0 = \frac{\sum_{i=1}^{n} I_i x_i}{\sum_{i=1}^{n} I_i}, \quad y_0 = \frac{\sum_{i=1}^{n} I_i y_i}{\sum_{i=1}^{n} I_i},
$$

(4.9)

<table>
<thead>
<tr>
<th></th>
<th>0.0 $\lambda$ rms</th>
<th>0.2 $\lambda$ rms</th>
<th>0.4 $\lambda$ rms</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_x$ ($\mu m$)</td>
<td>136</td>
<td>189</td>
<td>242</td>
</tr>
<tr>
<td>$\sigma_y$ ($\mu m$)</td>
<td>136</td>
<td>85</td>
<td>34</td>
</tr>
<tr>
<td>$\epsilon$ ($\mu m^2$)</td>
<td>$186 \cdot 10^2$</td>
<td>$157 \cdot 10^2$</td>
<td>$78 \cdot 10^2$</td>
</tr>
</tbody>
</table>

Table 4.1: Rms beam sizes at undulator begin with a focal length of 7 m

The intensity profiles and wavefront shapes at undulator begin, derived by ZE-MAX, are depicted in Fig.4.6.
4 Numerical simulations

![Image](image.png)

(a) No aberration

(b) $0.2\lambda$ rms wavefront deviation

(c) $0.4\lambda$ rms wavefront deviation

Figure 4.6: Computed intensity profile and wavefront at undulator begin.

Again, one can clearly see the deformation of the originally round intensity profile to an elliptic beam. Moreover, the wavefront shape maintains its initial shape. The saddle function is still present but with different strength due to the initial rms distortion.

4.4 Simulation of the FEL process

In this section, the final results of the simulation work are presented. With transferring the generated beam file for the use of GENESIS, the beam power had to be defined. During simulation with ZEMAX, the absolute power of the beam had no relevance. Now, the beam power has an impact on the FEL process. For the following simulations, the total beam power was assumed to be 50 kW. This is expected for the real case. An overview of start conditions for the GENESIS simulation is presented in table 4.2.
4.4 Simulation of the FEL process

<table>
<thead>
<tr>
<th>initial power (kW)</th>
<th>0.0 λ rms</th>
<th>0.2 λ rms</th>
<th>0.4 λ rms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>initial power density (W/μm²)</td>
<td>2.69</td>
<td>3.18</td>
<td>6.42</td>
</tr>
<tr>
<td>initial cross section ϵ (μm²)</td>
<td>186·10²</td>
<td>157·10²</td>
<td>78·10²</td>
</tr>
</tbody>
</table>

Table 4.2: Start conditions with constant initial seed power

Results with constant initial seed power

Based on the above given assumptions, Fig. 4.7 shows the computed power along the undulator length z. Note that the sFLASH undulator consists of three separated sections of 2 m length each. The steps in the figures correspond to this layout as between the undulators no periodic magnetic field is present.

![Power graph](image)

Figure 4.7: Power along z (initial beam power 50 kW)

The power rise decreases due to the initial wavefront distortion. If we have a look at the seed's cross section (see Fig. 4.8), this behavior can be explained by the beam’s focusing. The beam without distortion has a smaller cross section in the first part of the undulator leading to a much higher power density. The power density is depicted in Fig. 4.9. This explains the higher power rise itself as for seeding high power densities at the beginning are crucial.
4 Numerical simulations

Figure 4.8: Cross section along \( z \) (initial beam power 50 kW)

Figure 4.9: Power density along \( z \) (initial beam power 50 kW)
4.4 Simulation of the FEL process

Actually, the distorted wavefronts have two minima of cross section and two maxima in power density respectively. Once more, this is a reason of the astigmatic initial wavefront. Firstly, the focus in $x$-direction and following in $y$-direction appears, shown as two peaks in the figures. But the overall power density does not reach the maximal value as with the perfect Gaussian wavefront. Only if in both transversal directions a focal point is present simultaneously, the power density increases to the maximum. Exactly this happens with the perfect wavefront and therefore seeding works better.

Influence of Gouy’s phase

The evolution of beam power shows an interesting behavior in the very first undulator section. In contrast to the expectations, the non-distorted beam causes a drop of power. Instead, the beam of $0.2 \lambda$ rms deviation shows only a small drop of power and the most distorted wavefront keeps the highest power (see Fig.4.10).

![Figure 4.10: Power along $z$ near beam waist](image)

Eye-catching is the fact that this drop of power occurs around the positions of the beam waists. This behavior can be explained be Gouy’s phase which was presented in section 2.1. The rapid phase change of $\pi$ around the Gaussian mode’s beam waist disrupts the condition for constant energy transfer. The relative phase between the electron transverse oscillatory motion and the co-propagating wave is abruptly
changed by $\pi$. Therefore, the energy flow between the electromagnetic wave and the electron beam, as discussed in section 2.7.2, is inverted. This leads to the loss of power. The beams with distorted wavefront are less affected as Gouy’s phase change distributes itself over a larger region. A consequence of the power loss is that the electrons gain energy. This way, a proof of the proposition given above is derived. The energy gain of the electrons shows that the FEL condition for constant energy is turned, which is only accomplished by a phase change by $\pi$. Instead of the electromagnetic wave, the electrons gain energy. Fig.4.11 shows exactly this reaction.

![Energy of electron beam](image)

**Figure 4.11:** Electron energy along $z$ near beam waist

**Results with constant initial seed power density**

The above discussed behavior only regards the focusing. The influence on the FEL process due to a distorted wavefront itself cannot be extracted. For further examinations, the GENESIS simulations were repeated not with constant beam power but constant initial power density. Hereby, the initial power density was assumed to be 2 W/µm². Table 4.3 shows the start conditions for this case.
4.4 Simulation of the FEL process

<table>
<thead>
<tr>
<th>initial power (kW)</th>
<th>0.0 ( \lambda ) rms</th>
<th>0.2 ( \lambda ) rms</th>
<th>0.4 ( \lambda ) rms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>37</td>
<td>31</td>
<td>16</td>
</tr>
<tr>
<td>initial power density (W/( \mu \text{m}^2 ))</td>
<td>2.00</td>
<td>2.00</td>
<td>2.00</td>
</tr>
<tr>
<td>cross section ( \epsilon (\mu \text{m}^2) )</td>
<td>186 ( \cdot 10^2 )</td>
<td>157 ( \cdot 10^2 )</td>
<td>78 ( \cdot 10^2 )</td>
</tr>
</tbody>
</table>

**Table 4.3:** Start conditions with constant initial seed power density

The simulated development of power with these start conditions is depicted in Fig.4.12. The differences in power are now even more distinctive. In Fig.4.14 one can see that in the undulator’s very first part the power density increases more rapidly with increasing wavefront distortion. This again follows from the smaller spot sizes. The cross section decreases more rapidly and therefore the power density increases faster. But after passing the first focus, the cross section increases again and the non-distorted wavefront acquires the smallest cross section.

![Graph showing power development along z](image)

**Figure 4.12:** Power along \( z \) (initial beam power density 2W/\( \mu \text{m}^2 \))
4 Numerical simulations

Figure 4.13: Cross section along $z$ (initial beam power density $2\,\text{W}/\mu\text{m}^2$)

Figure 4.14: Power density along $z$ (initial beam power density $2\,\text{W}/\mu\text{m}^2$)
4.5 Conclusion and Summary

According to the simulation results, the wavefront’s shape affects the focusing of the seed and thus the coupling to the electron beam in the undulators area. The beam’s cross section has to be sufficiently small in the undulator. But when the wavefront has initial distortions, the beam cannot be focused on a small spot with the present layout of the XUV injection beam line. The effect of the reduced coupling appears to be much stronger than the direct effects of the phase distribution.
5 Wavefront measurement methods

5.1 Hartmann method

The Hartmann wavefront sensor was invented over a century ago to perform optical metrology [32]. The success of this method is based on its simplicity and its adaption to a wide variety of applications. Adaptive optics, ophthalmology, and laser wavefront characterization are some examples. The measurement is an indirect wavefront sensing technique and is based on geometrical effects namely transverse ray aberration in an image plane [33]. In principle, this sensor only consists of a charge coupled device (CCD) and a mask. The mask is an array of apertures mounted a certain distance $D$ from the CCD. The obtained data can be interpreted as the gradients of the wavefront. A schematic of the sensing principle is shown in Fig.5.1.

Light hits the array of apertures and is sampled into as many spots as the number of apertures. As already brought up in section 2.2, the propagation direction of a light ray is orthogonal to the local wavefront. Hence, the spot position on the CCD is proportional to the respective average local phase tilt of the wavefront at aperture position [34]. To acquire the local tilt data, the spot positions have to be compared with those produced by a known reference wavefront. Considering the reference as a plane wavefront incident normal to the aperture array, the produced spot positions are lying exactly under the respective apertures as indicated in Fig.5.1. This way, a reference spot pattern is created and the spot positions produced by an unknown wavefront can be compared to these of the reference wavefront. With knowledge of the distance $D$ between mask and CCD, at each aperture position the wavefront’s local tilt values respective to a reference wavefront are computed by

$$\text{Tilt}_x = \frac{\Delta x}{D} \quad \text{and} \quad \text{Tilt}_y = \frac{\Delta y}{D}. \quad (5.1)$$

Basically, this is the acquisition of the wavefront’s gradient field. The measurement of the tilt data then is converted into a replica of the wavefront using certain methods. This is called wavefront reconstruction.

For an analysis of the spot pattern conducting wavefront reconstruction, it is important to reference each spot to the aperture where it was sampled. This leads
to the constraint that two neighboring spots must not overlap. Otherwise it was not obvious which spot was sampled by which aperture. According to this, a spot has to be inside a certain area on the CCD. For each aperture this area defines the region of potential spot position. For that reason it is also called "spatial range". The size of the spatial range is usually given by the aperture separation. Fig. 5.2 shows this principle schematically.

According to the spatial range, the maximal detectable tilt is given by

$$Tilt_{\text{max}} = \frac{\text{dynamic range}}{2D}. \quad (5.2)$$

In this formula it can be directly seen that the maximal detectable tilt can be adjusted by changing the distance $D$. To some extent the problem with overlapping spots can be overcome by sophisticated spot-finding algorithms. On the other hand, the minimal detectable tilt depends on the position resolution $\Delta_{\text{pos}}$ of the CCD.
and is defined by sensor noise, the spot detection mechanism, and the centroiding
algorithm. The minimal detectable tilt is given by

\[ \text{Tilt}_{\text{min}} = \frac{\Delta_{\text{pos}}}{D}. \]

(5.3)

So both minimal and maximal detectable tilt are simultaneously adjusted by chang-
ing the distance \( D \). With increasing \( D \), the sensitivity of the detector becomes
higher, whereas with decreasing \( D \) larger tilts can be detected.

5.2 Wavefront reconstruction from gradients

In this section wavefront reconstruction techniques are discussed. These are differ-
ent methods to recover a 3D surface shape from a discrete set of data. A measure-
ment with the Hartmann sensor gives a discrete set of slope data on a 2-dimensional grid
as described in section 5.1. In the following, each grid point is defined by two indices
\( i \) and \( j \) corresponding to the transversal directions \( x \) and \( y \). Thus defining a gradient
vector field which has to be converted in a replica of the wavefront. Again, there
are different reconstruction types which can be used for the Hartmann sensor. They
can be divided into zonal and modal reconstruction.

5.2.1 Zonal reconstruction

Zonal reconstruction is a type of numerical integration. The wavefront height is
calculated at each point independently. Only direct neighbors of a point are taken
into account. It is also known as local reconstruction method expressing the in-
dependence from remote sampling points. Mainly, there exist two zonal methods.
Linear integration is the most simple reconstruction type. Whereas the Southwell
reconstructor is an extension of the linear reconstruction and gives more accurate
data.

Linear integration

The computation of the surface shape begins at one edge of the grid of the measured
slope data. At the beginning, the wavefront height is defined as zero at all sampling
points. The height of the wavefront at the next adjacent point in \( x \)-direction is
calculated by taking the slope of the first point times the aperture separation. Fig.5.3
shows this principle schematically.
Now the next point in $x$-direction can be calculated the same way, but in addition the surface height of the previous point is added. The principle follows the simple mathematical concept that the height along a path increases by the length times the slope. This is done for every $x$-row starting at the first point of a row with zero height. Next step is to do exactly the same but with all $y$-columns. Mathematically, this is described by

$$W_{x_{i,j}} = W_{x_{i-1,j}} + \frac{\partial W_{i-1,j}}{\partial x} \Delta x \quad (5.4)$$

$$W_{y_{i,j}} = W_{y_{i,j-1}} + \frac{\partial W_{i,j-1}}{\partial y} \Delta y$$

$\Delta x$ and $\Delta y$ denote the aperture separation along the transversal directions. The overall wavefront surface is reconstructed by adding both height values from the individual calculations in $x$- and $y$-direction.

$$W_{i,j} = W_{x_{i,j}} + W_{y_{i,j}} \quad (5.5)$$

**Southwell Reconstructor**

William Southwell introduced this algorithm that is very popular [36]. The basic idea is to minimize errors of the linear reconstruction. Therefore, each wavefront height is calculated as the average height predicted by the neighbored slope data.
This is expressed mathematically by the following equation.

$$W_{i,j} = \sum_{m=-1}^{1} \sum_{n=-1}^{1} I_{i+n,j+m} \cdot [W_{i+n,j+m} + \left( \frac{\partial W}{\partial x_{i,j}} + \frac{\partial W}{\partial x_{i,n,j+m}} \right) \cdot \Delta],$$  \hspace{1cm} (5.6)$$

where \( \phi_{i,j} \) is the wavefront height at grid location of index \((i, j)\). \( \Delta \) is the separation between grid point \((i, j)\) and \((i + n, j + m)\). \( I_{i,j} \) is the intensity of the measured spot. The weighting with the measured spot intensity gives additional accuracy. The key is that high-intensity spots reproduce a more accurate slope measurement. Therefore, they are taken into account more than spots with low intensity.

### 5.2.2 Modal Zernike reconstruction

The big advantage of Zernike polynomials is their direct connection to wavefront aberrations. In optical meteorology Zernike polynomials are the standard set used for analyzing optical elements. As already mentioned, wavefront aberrations mostly have a form which can be described by Zernike polynomials. As these aberrations are well known, interpretation is easy.

In the following, the measured tilt in \( x \) is called \( a(x, y) \) and the tilt in \( y \) is called \( b(x, y) \). This can also be expressed by a vector-field with the vectors reproducing the wavefront’s tilts as depicted in Fig.5.4.

![Figure 5.4: Description of a Hartmann measurement by a vector field](image-url)
Most important is the fact that an analytical derivative of the Zernike polynomials can be obtained, which may be fitted to the slope data. As the basis functions are continuous and the first derivative is available, the wavefront gradients can be written as follows.

\[ W_x = \frac{\partial W(x, y)}{\partial x} = \sum_{i=0}^{n-1} c_i \frac{\partial f_i(x, y)}{\partial y} \]

\[ W_y = \frac{\partial W(x, y)}{\partial y} = \sum_{i=0}^{n-1} c_i \frac{\partial f_i(x, y)}{\partial x} \]  

This gives an analytic description of the wavefront’s gradients at each point \((x, y)\) and can be compared to the measured slope data. The measured slope data are related to the analytical derivatives by

\[ W_x = a(x, y), \quad \text{and} \quad W_y = b(x, y). \]  

This way, the coefficients \(c_i\) can be extracted. Inserted in formula 2.42, this gives a solution of the desired wavefront.

**Fitting methods**

It is now the task to find the best set of coefficients \(c_i\) in the wavefront representation (equation 2.43) so that \(W_x\) and \(W_y\) is as closely as possible at the measured gradient data \(a\) and \(b\). Basically, this means altering an arbitrary surface \(W(x, y)\) until its gradients coincide with the measured gradient data \(a(x, y)\) and \(b(x, y)\) as closely as possible. Mathematically, an error function has to be minimized.

Error function:  
\[ C = \iint_\Omega \left[ |W_x - a|^2 + |W_y - b|^2 \right] dxdy \]  

Different algorithms are applicable to find the minimum of that error function. The most simple way is least square fitting. Many algorithms use the calculus of variation which leads to an Euler-Lagrange equation that has to be solved. From the Euler-Lagrange equation a Poisson equation (5.10) is derived as shown in [37].

\[ \nabla^2 W = \frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} \]  

The reconstruction algorithm used in this thesis follows another concept. There is no error function is used but the coefficients \(c_i\) of the wavefront representation are calculated directly. This is possible by processing matrices with large numbers of indices. In the following, the principle of the reconstruction algorithm is outlined.

Again, the wavefront should be able to be represented by Zernike polynomials. Furthermore, the wavefront is assumed to have aberrations only up to a certain
5.2 Wavefront reconstruction from gradients

order. Mostly, aberrant wavefronts can be described by low order Zernikes. In the used reconstruction code, the wavefront is represented by Zernikes up to the 54th order \( (k = 55) \). This is a quite high number and should be much more than really needed to describe the wavefront’s aberrations. First of all, equation 5.8 is used in combination with the derivatives of the Zernike polynomials.

\[
a(x, y) = \sum_{i=0}^{k-1} c_i \frac{\partial Z_i(x, y)}{\partial x}, \quad b(x, y) = \sum_{i=0}^{k-1} c_i \frac{\partial Z_i(x, y)}{\partial y}
\]

(5.11)

For a simpler notation, the derivatives of the Zernike polynomials are called \( g_i(x, y) \) and \( q_i(x, y) \).

\[
\frac{\partial Z_i(x, y)}{\partial x} = g_i(x, y), \quad \frac{\partial Z_i(x, y)}{\partial y} = q_i(x, y)
\]

(5.12)

The set of measured data is used to create a vector \( \vec{t} \). The vector has a length of 2 times the number of sampling points \( N \). Let’s assume we have a rectangular sampling grid with \( n \) sampling points in a row and \( m \) sampling points in a column. Then \( N \) is just \( n \cdot m \). To each sampling point a measured wavefront tilt is connected. From index 1 to \( N \) the vector is defined as the tilts in \( x \), from \( N + 1 \) to \( 2N \) as the tilts in \( y \). This vector comprises all information gained by the Hartmann measurement.

\[
\vec{t} = \begin{pmatrix}
  a(x_1, y_1) \\
  a(x_1, y_2) \\
  \vdots \\
  a(x_n, y_1) \\
  a(x_n, y_2) \\
  \vdots \\
  a(x_n, y_m) \\
  b(x_1, y_1) \\
  b(x_1, y_2) \\
  \vdots \\
  b(x_n, y_1) \\
  b(x_n, y_2) \\
  \vdots \\
  b(x_n, y_m)
\end{pmatrix}
\]

(5.13)

The next step is to compute analytical derivatives of each Zernike polynomial \( g_i \) and \( q_i \) at the sampling points \((x_i, y_j)\). The analytical derivatives are stored in a matrix
5 Wavefront measurement methods

\[ \hat{M} \text{ with size } (2N \times k). \]

\[
\hat{M} = \begin{bmatrix}
g_1(x_1, y_1) & g_2(x_1, y_1) & \cdots & g_k(x_1, y_1) \\
g_1(x_1, y_2) & g_2(x_1, y_2) & \cdots & g_k(x_1, y_2) \\
\vdots & \vdots & \ddots & \vdots \\
g_1(x_n, y_1) & g_2(x_n, y_1) & \cdots & g_k(x_n, y_1) \\
g_1(x_n, y_2) & g_2(x_n, y_2) & \cdots & g_k(x_n, y_2) \\
\vdots & \vdots & \ddots & \vdots \\
g_1(x_n, y_m) & g_2(x_n, y_m) & \cdots & g_k(x_n, y_m)
\end{bmatrix}
\]

(5.14)

The analytical derivative at one sampling point can be rewritten in form of multiplication of vectors. For example, for the first sampling point \((x_1, y_1)\) this leads to

\[
a(x_1, y_1) = \sum_{i=0}^{k-1} c_i \cdot Z_i(x_1, y_1) = \vec{c} \cdot \vec{Z}(x_1, y_1).
\]

(5.15)

This can also be done for each sampling point at once. One gets a linear system of equations shown in equation 5.16.

\[
\begin{pmatrix}
a(x_1, y_1) \\
a(x_1, y_2) \\
\vdots \\
a(x_n, y_1) \\
a(x_n, y_2) \\
\vdots \\
a(x_n, y_m) \\
b(x_1, y_1) \\
b(x_1, y_2) \\
\vdots \\
b(x_n, y_1) \\
b(x_n, y_2) \\
\vdots \\
b(x_n, y_m)
\end{pmatrix}
= \begin{bmatrix}
g_1(x_1, y_1) & g_2(x_1, y_1) & \cdots & g_k(x_1, y_1) \\
g_1(x_1, y_2) & g_2(x_1, y_2) & \cdots & g_k(x_1, y_2) \\
\vdots & \vdots & \ddots & \vdots \\
g_1(x_n, y_1) & g_2(x_n, y_1) & \cdots & g_k(x_n, y_1) \\
g_1(x_n, y_2) & g_2(x_n, y_2) & \cdots & g_k(x_n, y_2) \\
\vdots & \vdots & \ddots & \vdots \\
g_1(x_n, y_m) & g_2(x_n, y_m) & \cdots & g_k(x_n, y_m)
\end{bmatrix}
\begin{pmatrix}
c_1 \\
c_2 \\
\vdots \\
c_k
\end{pmatrix}
\]

(5.16)
Equation 5.16 can also be written in the easy form as in equation 5.17.

\[ \vec{t} = \hat{M} \cdot \vec{c} \]  

(5.17)

In the equation above, each component is well known except for the vector \( \vec{c} \). The vector \( \vec{c} \) can be easily computed by taking the inverse of the matrix \( \hat{M} \) and multiplying it with \( \vec{t} \).

\[ \vec{c} = \hat{M}^{-1} \cdot \vec{t} \]  

(5.18)

Now, the vector \( \vec{c} \) or the coefficients of the Zernike polynomials are received. Applying this to

\[ W = \sum_{i=0}^{k-1} c_i \cdot Z_i(x, y), \]  

the 3D surface of the wavefront is reconstructed.

5.3 Other wavefront sensor techniques

5.3.1 Irradiance methods

As the name indicates, irradiance methods use radiant-flux densities to measure the wavefront. The key is to measure the intensity profile in different planes. In principle, there exists a relationship between the irradiances at these planes depending on the wavefront. This way, information about the wavefront shape can be deduced. There are two different methods, phase retrieval and curvature sensing [34].

Phase retrieval

Southwell suggested in 1977 to determine the wavefront from irradiance measurements in the near field and far field [38]. The method works as follows. An initial assumption of the wavefront in the first plane, the near field, is made. This assumption of phase distribution together with the known intensity pattern is used to propagate the beam. Sophisticated wave propagation codes have to be used. As a result, one gets the computer-generated far field. The intensity distribution is then compared to the real far field which might show huge differences to the computed one. Now the next iteration can be conducted by changing slightly the wavefront’s initial shape in order to get a better fitting far field pattern. After an extensive number of iterations, one gets an optimal trial input phase which should be equivalent to the real wavefront. The measurement setup is quite simple, but the computation process is very demanding. It is mostly a mathematical issue. Furthermore, there is no proof that a particular far-field pattern can be reached by only one special wavefront shape.
It is worth mentioning that phase retrieval was used to recover the Hubble Space Telescope [39]. There was a problem with the primary mirror and therefore the images were severely degraded. To analyze the specific aberrations causing this problem, phase retrieval methods played a major role. In 1993 a correction mirror could be installed and from then on great images could be achieved.

**Curvature sensor**

The curvature sensor was developed in the late 1980's. As in Fig. 5.5 depicted, it works by measuring simultaneously the irradiances at two closely spaced planes. The planes are located at the same distance but at opposite sides from a reference plane where a pupil is present. A lens is positioned in the beam line so that an incident plane wave has its focal spot at the reference plane. In this case, the irradiance profiles in both planes will be equal. If now a curved wavefront is present, the focal spot position changes and the irradiance profiles will not be longer equal. There is a specific relation between the cross sections of the irradiance profiles and the wavefront’s curvature. Thus, the phase curvature can be reconstructed. Almost the same computation effort applies as with phase retrieval.

![Principle of curvature sensor](image)

**Figure 5.5:** Principle of curvature sensor

### 5.3.2 Interferometric methods

Interferometric methods are direct sensing techniques because they directly measure the difference in the optical path length (OPL). There are two main methods namely radial shearing interferometry and point diffraction interferometry [40].
5.3 Other wavefront sensor techniques

**Shearing interferometry**

In radial shearing interferometry the beam is split in two parts from which one part is expanded by a telescope setup as depicted in Fig.5.6. After expansion of one part, the two beams again are overlapped. There the expanded beam acts as a reference and interference occurs in the overlap region. On a screen placed in the beam line, the interference pattern of the enlarged and the original beam is observed. Analyzing the fringe pattern, the wavefront shape can be reconstructed using certain reconstruction algorithms which are quite complex. The expanded beam’s wavefront has to be sufficiently flat to serve as a reference. Therefore, the expansion has to be large enough to flatten the wavefront in the overlap region. The quality of the resulting wavefront measurement strongly depends on the absence of aberrations in the reference. Any aberration will degrade the results.

![Principle of radial shearing interferometer](image)

**Figure 5.6:** Principle of radial shearing interferometer

**Point diffraction interferometry**

The point diffraction interferometer also produces an interference pattern which can be interpreted to the desired wavefront shape. In contrast to shearing interferometry, the reference is produced in another way. This scheme is shown in Fig.5.7. The beam hits a semitransparent filter with a pinhole in its center. This pinhole spatially filters the beam and acts as a diffraction aperture producing a spherical wave which represents the reference. As the filter is semitransparent, a part of the incident wave passes through the filter and interferes with the reference. In contrast to radial shearing, the reference is given by a spherical wavefront with well known curvature. Again, the observed fringe pattern can be used to reconstruct the wavefront shape [40].
Unfortunately, mechanical vibrations make obtaining an image with an interferometer difficult. In contrast to that, a Hartmann sensor is effectively insensitive to vibrations by averaging several frames together removing temporal noise. Another drawback to interferometric methods especially in the XUV spectral range comes from the number of optical components needed in the setup. Alignment is very difficult in the X-ray regime and fabricating optical components is very demanding. Besides difficult alignment of the components, they are additional sources of aberrations and influence the results if not perfectly fabricated.

5.3.3 Geometrical methods

These techniques have the advantage that they are relatively simple to use. Only with optical ray propagation the results can be interpreted. Mainly, there are three different types of geometrical methods with the Hartmann sensor being one of them. The other two methods are called Knife-edge test and Ronchi test. The Hartmann-Shack sensor is an extension of the Hartmann setup.

Knife-edge test

The knife-edge technique was invented by Foucault in 1856 and was originally used to test the surface quality of lens elements and mirrors in the production of telescopes [41]. Although it is the most primitive testing method, it was used for over a century for quality testing in the building of astronomical instruments.

To investigate the wavefront with the knife-edge test, the corresponding beam has to be focused. In the beam’s focal point a knife edge is placed. When moving the knife edge perpendicular to the optical axis, the lens aperture is observed directly from behind the knife edge. The shadow, observed by a screen, gives information about the aberration content. If the wavefront has a non-aberrant shape, the aperture darkens almost instantaneously when the knife edge passes through the focus.
Then all rays are covered by the knife edge. If the same is done with the knife edge placed in front of or after the focal point, the aperture darkens evenly starting at the top or at the bottom of the aperture respectively (see Fig.5.8). But if the wavefront has aberrations, some rays will cross before or after the focus. As a result of this, the observation pattern acquires special shapes. With the knife edge placed in the focal point, now the lens aperture will not darken evenly over the entire surface but uneven illumination will be observed [42]. These shadow patterns are the basis of the geometrical analysis. The observed pattern depends on the position of the knife edge along the optical axis.

![Figure 5.8: Principle of knife-edge test](image)

**Ronchi test**

The Ronchi test was invented in 1923 by Vasco Ronchi. Although it can be fully explained by diffraction and interference theory, it belongs to the geometrical methods as it was originally devised as such.
A perfect lens brings the incident beam to a focus in the vicinity of which a
diffraction grating is placed perpendicular to the optical axis. The grating is known
as the Ronchi ruling and consists of fine parallel lines \[33\]. The Ronchi ruling
produces multiple diffractive orders of the incident beam. After reaching a lens
called "pupil relay", the diffracted orders from the grating are collected on a screen
or CCD as depicted in Fig.5.9.

![Focusing lens and Ronchi grating](image)

**Figure 5.9:** Principle of Ronchi test

There, the produced interference fringes characterize the aberrations of the wave-
front. A plane wavefront will produce regular parallel lines whereas a distorted
wavefront yields curved lines. Analyzing these fringes determines the type of abber-
ration. See Fig.5.9 for an example of observed fringe patterns by a plane wave and
a wave with spherical aberration.

### Shack–Hartmann method

The Shack-Hartmann method is a modification of the Hartmann method. The array
of hard apertures is replaced by a discrete array of micro-lenses \[43\] (see Fig.5.10).
The main advantage in comparison to the Hartmann sensor is the better photon
efficiency. Almost the entire wavefront is examined, not just small samples as with
the Hartmann sensor. Another point is that the diffraction pattern at the focus of
the lenses is exactly the Fourier transform of the electric field pattern at the array of
micro-lenses. Therefore, the focal spot position is only proportional to the average
5.3 Other wavefront sensor techniques

wavefront slope over the aperture and is not dependent of higher-order aberrations and intensity profile variations.

Figure 5.10: Principle of the Shack-Hartmann sensor. In contrast to a Hartmann sensor using a hole array, the incoming beam is sampled by a lenslet array.

Because there are no conventional lenses suitable for the XUV spectral range, the Shack-Hartmann sensor is inappropriate for wavefront measurement of the seed laser.
6 Wavefront sensor - setup and usage

6.1 Experimental setup

The wavefront sensor was constructed according to the Hartmann principle. The experimental setup basically consists of the Hartmann mask and a CCD camera placed along the optical axis. Due to XUV absorption in air, the whole experimental setup is constructed inside vacuum chambers. The CCD camera is particularly suitable for X-rays. It can be connected directly to the diagnostic section. The mask is inserted in the optical axis by a manipulator. The manipulator is plugged to the side of the vacuum chamber ahead of the camera. This way, the mask can be moved transversally in order to find the optimal position for sampling the entire beam. The mask itself is attached to an adapter especially designed for this experiment. As actually two masks with different properties were fabricated, the adapter covers both Hartmann masks at once as depicted in Fig.6.1.

![Figure 6.1: Adapter for the Hartmann masks](image)

(a) Side view

(b) Top view

mask\((\text{h} \text{o} \text{l} e s: \text{200}\mu\text{m})\) mask\((\text{h} \text{o} \text{l} e s: \text{100}\mu\text{m})\)
6 Wavefront sensor - setup and usage

The adapter is fixed on the manipulator. By moving it in the transversal direction, one can easily switch the mask currently used for sampling the beam. This was done for practical reasons because ventilating and evacuating the vacuum chamber is a long lasting procedure. A photo of the experimental setup is shown in Fig.6.2. Fig.6.3 shows a sketch of the setup with its main components.

Figure 6.2: Experimental setup of the wavefront sensor. The Hartmann mask is mounted on an adapter attached to a manipulator. The CCD camera is placed at a distance of 218 mm from the Hartmann mask along the optical axis.

Figure 6.3: Sketch of the experimental setup

Unfortunately, the distance $D$ between Hartmann mask and CCD camera cannot be adjusted continuously. As discussed in section 5.1, the maximal and minimal
6.1 Experimental setup

detectable wavefront tilt ($\text{Tilt}_{\text{max}}, \text{Tilt}_{\text{min}}$) is adjusted by changing the distance $D$. The manipulator gives the option to change the distance $D$ by only a few millimeters. Larger changes of $D$ can only be achieved by adding or removing vacuum tube parts between the manipulator and the CCD camera. Therefore, $D$ can be changed only with step sizes given by the smallest available tube length. Furthermore, this again needs ventilating and subsequently evacuating the vacuum chamber. The distance $D$ for all following measurements was fixed to 218 mm. Based on the assumptions for the seed’s wavefront shape, this promised to yield reasonable limits for the detectable wavefront tilt.

6.1.1 CCD camera

The camera used for this experiment is a Princeton Instruments "Acton PIXIS-XO" camera. It is a high-sensitivity, thermoelectrically cooled X-ray camera [44]. The CCD chip has no AR coating and is therefore particularly suitable for low energy X-ray detection. With $2048 \times 512$ pixels of size $13.5 \times 13.5 \, \mu m^2$ and 100% fillfactor, the overall chip size is $27.6 \, mm$ in $x$ and $6.9 \, mm$ in $y$-direction.

6.1.2 Hartmann masks

Two Hartmann masks are available for this experiment. They were produced by the university workshop at DESY. Both masks have a size of $10 \times 10 \, mm^2$ and are made of aluminium plates of thickness 0.2 mm. The first mask has $21 \times 21$ (=441) holes with a hole diameter of 0.2 mm and a center to center distance of 0.4 mm. The second mask only differs from the first by its 0.1 mm hole diameter. The holes were made by drilling. For the holes of 0.1 mm diameter the smallest drillhead available on the market was used. Unfortunately, the mask with 0.1 mm hole diameter has a very low transmission. The spots could not be separated clearly from noise and therefore image processing was not possible. Therefore, all following measurements were acquired with the mask of hole diameter 0.2 mm. With this one, the produced spot pattern was sufficiently rich in contrast to be analyzed by image software. According to an aperture separation of 0.4 mm and the above mentioned properties of the experimental setup, $\text{Tilt}_{\text{max}}$ and $\text{Tilt}_{\text{min}}$ yield

$$\text{Tilt}_{\text{max}} = 0.92 \, mrad \quad \text{and} \quad \text{Tilt}_{\text{min}} = 0.06 \, mrad.$$ 

In Fig.6.4 images of both masks taken with a microscope are shown.
6 Wavefront sensor - setup and usage

(a) mask with 0.1 mm hole diameter
(b) mask with 0.2 mm hole diameter
(c) mask with 0.2 mm hole diameter (zoom)

Figure 6.4: Microscopy images of the fabricated Hartmann masks

6.1.3 Reference measurement

In order to create a well-known wavefront as a reference, a pinhole with 10 \( \mu \text{m} \) diameter was used. Illuminated from behind, the pinhole acts as a point source producing a spherical wavefront. It is placed in the optical axis in a distance \( d \) of 236 mm ahead of the Hartmann mask. To reach sufficiently rich contrast of the spot pattern, an external light source had to be used. The XUV seed could not illuminate the pinhole with sufficient intensity. As an external light source, a high-power LED is placed outside the vacuum tube. The LED has a power consumption of 3 W at a wavelength of 465 nm [45]. A setup of apertures and a focusing lens (\( f = 100 \) mm) was used in order to focus the light on the pinhole. The pinhole itself is placed inside the vacuum chamber on the optical axis. A vacuum window allows to couple the LED light into the vacuum chamber. A 45°-mirror deflects the beam to the pinhole. The acquired spot pattern of the produced spherical wavefront serves as a reference. Thereby, the wavefront’s radius of curvature is given by the distance from the pinhole to the Hartmann mask. Fig.6.5 and Fig.6.6 shows a photo and a sketch of the experimental setup. The pinhole and the 45°-mirror are inside the vacuum chamber and not visible on this photo, but they are shown schematically in Fig.6.6.
6.2 Software

6.2.1 Image processing

As part of this thesis, a complete wavefront reconstruction tool was created. The software was written in Matlab. First of all, the images taken by the camera had
to be analyzed. A spot-finding algorithm searches for each centroid coordinates.
Sub-pixel resolution can be achieved by determining the "center of gravity" which is
the first moment of each sub-aperture image. A coarse grid is applied to the images
to define the region of interest for each sub-aperture. Thereby, the respective region
of interest is given by the spatial dynamic range. Following, the center of gravity is
calculated successively for each sub-area. The center of gravity is defined by

\[ cx = \frac{\int \int I(x,y)xdxdy}{\int \int I(x,y)dxdy} \quad \text{and} \quad cy = \frac{\int \int I(x,y)ydydx}{\int \int I(x,y)dxdy}, \]  

with \( cx, cy \) the spot position in \( x- \) and \( y- \) direction respectively and the intensity
distribution \( I(x,y) \). The spot-finding algorithm is applied to both, the spot pattern
produced by the seed’s wavefront and the reference wavefront.

### 6.2.2 Gradient estimation

As input, the reconstruction algorithm uses the wavefront’s gradients. Thereby, the
gradients must be given in comparison to a plane wavefront. The reference wavefront
is normally given by a plane wavefront. It can be generated by a point source in
a reasonable large distance from the Hartmann mask. Unfortunately, this was not
possible in the present experimental setup. The only option was to position the
point source only shortly ahead of the mask. Thus creating a spherical wavefront
with diverging light rays. Therefore, the imaged spot pattern has to be processed.
By applying the rules of optical imaging, the spot positions created by a plane wave
can be computed with

\[ c^p_x = c^{sp}_x \cdot \frac{D}{d + D}, \]

\( c^p_x \) and \( c^{sp}_y \) are the measured spot positions generated by the spherical reference
wavefront. \( c^p_x \) and \( c^{sp}_y \) stand for the spot positions as they would be measured by a
plane reference wavefront. Alternatively, this can be interpreted as calculating the
exact positions of the holes in the mask. As a plane wavefront’s gradients are all
equal, the corresponding spot pattern is a "one-to-one" image of the mask. Fig.6.7
shows a 3D-sketch of the reference measurement and its underlying principle.
6.2 Software

Figure 6.7: Computation of aperture positions

Light rays starting at the point source go straight through the holes and produce a spot pattern on the CCD camera. From this, the hole positions are recalculated, which is already included in the sketch. It already shows the result based on the measured spot positions and real distances $d$ and $D$. Following, the local gradients of the seed’s wavefront can be easily computed. The spot positions acquired with the seed are called $(c_x^{seed}, c_y^{seed})_i$. With $i$ denoting the corresponding hole, respectively the spatial range. The difference between $(c_x^p, c_y^p)_i$ and $(c_x^{seed}, c_y^{seed})_i$ yields the local gradient $(\theta_x, \theta_y)_i$ at the hole-position $(c_x^p, c_y^p)_i$.

$$\theta_x(c_x^p, c_y^p)_i = \frac{\Delta x_i}{D}, \quad \text{with} \quad \Delta x_i = (c_x^p - c_x^{seed})_i$$

$$\theta_y(c_x^p, c_y^p)_i = \frac{\Delta y_i}{D}, \quad \text{with} \quad \Delta y_i = (c_y^p - c_y^{seed})_i$$

In Fig.6.7 one can see that the spot-to-spot distance on the CCD camera is enlarged by a factor of about 2 compared to the hole separation. This leads to an increase of the overall image area of about 4 compared to the mask area. In the present experimental setup, a big problem arises from this enlargement. The CCD chip is relatively small in the $y$-direction. Thus, only a cutout of the reference spot pattern is detected with the limits given by the CCD’s chip size. A considerable amount of information for wavefront reconstruction is lost. This definitely needs to be improved. But, for the time being, no other generation of the reference wavefront
could by provided. The problem could by eliminated by placing the pinhole reason-
able far away from the mask. This way, the magnification factor of the spot pattern
would be smaller and the chip size would be large enough to sample the hole image.
Unfortunately, the experimental setup in the diagnostic beam line allowed no other
position for inserting the pinhole.

A detour to solve this problem was made by generating a "virtual reference". Hereby, it was assumed that the apertures on the mask are perfectly spaced. According to this, the spots produced by a plane wavefront are equally spaced with the spacing given by the designed hole separation. This spot pattern, as shown in Fig.6.8, was created by software.

![Figure 6.8: "Virtual" reference produced by a perfect Hartmann mask](image)

The disadvantage of this method is that one assumes the mask to be perfect.
For each pair of holes, their distance will differ slightly from the design value. This
introduces errors to the reconstructed wavefront which have to be taken into account additionally. The error calculation is presented within the next chapter.

6.2.3 Reconstruction

The reconstruction algorithm processes the gradients \((\theta_x, \theta_y)\). The wavefront is
fitted to a set of Zernike Polynomials as described in section 5.2.2. Two options
regarding the reference wavefront can be chosen. One option is to use the measured reference. For the reasons discussed before, thereby only wavefronts with a diameter of maximal 4 mm can be reconstructed. The second option is to use the virtual reference. With this option, limits to the wavefront’s maximum detectable cross section are given only by the mask size itself.

For testing the reconstruction algorithm, different well-known gradient fields were
simulated which correspond to certain surface shapes. A very simple surface is
one that shows only a global tilt. The corresponding vector field is constant with e.g. only a $y$-component. As expected, the obtained surface is plane but tilted in $y$-direction. See Fig.6.9 for the vector field and the corresponding reconstructed surface.

\begin{figure}[h]
\centering
\includegraphics[width=0.4\textwidth]{figure6.9.png}
\caption{Constant vector field of a tilted plane wavefront and its reconstructed surface}
\end{figure}

As a further example, a spherical surface is simulated. The corresponding vector field is a gradient field with negative divergence. Fig.6.10 shows the simulated vector field and the reconstructed surface.

\begin{figure}[h]
\centering
\includegraphics[width=0.4\textwidth]{figure6.10.png}
\caption{Vector field of a spherical wavefront and its reconstructed surface}
\end{figure}

As one can see, the reconstruction algorithm correctly computes the wavefront. The same test was done for an arbitrary wavefront. Thereby, the test wavefront was simulated by a random set of Zernike polynomials. In Fig.6.11, the wavefront’s gradient field and its corresponding reconstructed surface are shown.

\begin{figure}[h]
\centering
\includegraphics[width=0.4\textwidth]{figure6.11.png}
\caption{Gradient field of an arbitrary wavefront and its reconstructed surface}
\end{figure}
6 Wavefront sensor - setup and usage

Figure 6.11: Vector field of an arbitrary wavefront and its reconstructed surface

For a comparison of simulated and reconstructed wavefront, a line plot is given in Fig. 6.12.

Figure 6.12: Comparison of the simulated and reconstructed surface of an arbitrary wavefront

Height values along the mean y-row are shown for both wavefronts. This comparison shows a very accurately reconstructed surface.
7 Measurements for sFLASH

7.1 Wavefront of the seed laser

During the measurements of the XUV seed, the drive laser had an average energy of 31.4 mJ with an rms stability of 4.94%. The CCD camera worked with full gain (gain=1) and a frequency of 2 MHz. For a sufficient signal to noise ratio, the chip was binned in both directions with a binning factor of 2. So the effective pixel size was 2 times 13.5 $\mu m$ ($=27 \mu m$). The chip’s exposure time was 1s for the measurements. This means, the acquired images are integrated over 10 pulses. Single-shot measurements were also performed but unfortunately showed too less contrast to be processed by the image software.

7.1.1 Using the virtual reference

Firstly, the wavefront was reconstructed by using the virtual reference. Hereby, the wavefront could be evaluated over the beam’s whole cross section. In Fig.7.1 the spot pattern produced by the seed is shown (white). Additionally, the virtual reference (red) is overlapped to show deviations in spot position from the reference.

![Figure 7.1: Virtual reference (red) and spot pattern produced by the XUV-seed (white)](image)
Derived from this spot pattern, the gradient vector field is shown in Fig. 7.2.

One can see that the gradient field resembles a negatively divergent vector field. According to this, the wavefront mainly consists of a spherical surface with minor deviations. This spherical part is explained by the fundamental Gaussian mode and therefore corresponds to the perfect case. The beam size at sensor position was 3 mm. Therefore, the wavefront was reconstructed on a circle $\Omega$ with radius 3 mm. Fig. 7.3 shows the wavefront of the seed.

**Figure 7.2:** Derived gradient vector field

**Figure 7.3:** Reconstructed wavefront compared to a plane wave
Mainly, one is interested in the deviations from the perfect case which is the fundamental Gaussian mode. According to this, all in the following displayed surfaces present the deviations from the fundamental Gaussian mode. The wavefront of the fundamental is given by the Zernike polynomial \( Z_{2}^{0} \) (ordering number \( i = 4 \)). This part is cut out and does not contribute any more to the presented surfaces. Doing so for the above presented wavefront, the deviations are in the order of 2.5 \( \lambda \) (see Fig.7.4 and Fig.7.5). Note that in the following, the term "wavefront" is referred to as the deviation from the fundamental Gaussian mode.

**Figure 7.4: Reconstructed wavefront**

**Figure 7.5: Reconstructed wavefront**
7 Measurements for sFLASH

Zernike polynomials up to ordering number \( i = 54 \) were used for reconstruction. Fig. 7.6 shows the derived set of Zernike coefficients.

![Figure 7.6: Derived Zernike coefficients](image)

As expected, the Zernike coefficient \( c_4 \), describing the wavefront shape of the fundamental mode, is dominating. Low-order polynomials contribute to the wavefront more than high-order polynomials.

7.1.2 Using the measured reference

For the reasons discussed in section 6.2.2, using the measured reference allowed to evaluate only a fraction of the wavefront’s area. The wavefront could be reconstructed on a circle \( \Omega \) with a radius of 1.6mm. This is less than a third of the beam’s cross section. Fig. 7.7 and Fig. 7.8 show the wavefront evaluated by the measured reference.
7.2 Measurement error

7.2.1 Possible error sources

There are several crucial error sources in the experimental setup. Each of them results in an erroneous wavefront estimation. It is very important to minimize and identify these errors in order to achieve reliable results.
Measurements for sFLASH

Hartmann mask
The Hartmann mask is the crucial part of the sensor. The wavefront gradients are sampled at the positions of the holes. These positions have to be defined as reconstruction is based on the gradients at these positions. A tool to find these positions is given by the reference measurement. Using the virtual reference, regular positions with constant hole-to-hole distance were assumed. That means, the mask is considered to be perfectly fabricated. In this case, errors of the hole positions have to be considered additionally.

CCD chip
The CCD chip's pixel size gives a limit for the tiniest detectable gradient. Therefore, the pixel size is an indicator for the accuracy. Spot finding algorithms are able to find the spot positions with a resolution limited by pixel size. The algorithm used in the thesis fits a Gaussian beam profile to each spot. Thus, the center of gravity is found with sub-pixel resolution.

Longitudinal distances in the setup
The distances in the experimental setup have to be known very exactly. Only small deviations result in a faulty wavefront estimation. Most crucial hereby is the distance Hartmann mask to CCD chip. The wavefront's gradients depend strongly on that distance and therefore influence the whole reconstruction dramatically. Knowledge of the exact distance and its possible error is essential.

Moreover, in the case of the measured reference, the distance from pinhole to Hartmann mask is important as well. A wrong length leads to a wrong estimation of the spot positions. Thereby, the relative hole-to-hole distances are not influenced. But the average hole-to-hole distance depends strongly on the distance pinhole - mask.

7.2.2 Error calculation
There is no general rule about how to analyze the errors introduced by the reconstruction algorithm. The error is usually maximal at the edges of the reconstruction area because there the sampled spots have lowest intensity. This makes finding the center of gravity more difficult and introduces bigger errors. Moreover, the absolute wavefront's height values at the edges change more rapidly with the Zernike coefficients as at the inner part of the wavefront. As reconstruction produces a set of Zernike coefficients, errors in the coefficients will express themselves mainly on the wavefront's edges.
7.2 Measurement error

To evaluate the error on the measured wavefront, instead of an analytical procedure a numerical simulation was performed. The wavefront was reconstructed 10000 times with all above discussed errors randomly distributed. The resulting wavefront heights at each point of the wavefront surface were then analyzed by calculating their standard deviation. As expected, the accuracy of the reconstruction algorithm is more precise in the inner part of the wavefront. The computation was done independently for the wavefront reconstructed by the virtual and by the measured reference. Fig.7.9 shows the error in the center and at the edge of the wavefront reconstructed by the virtual reference. Thereby, the error is maximal $0.2\lambda$ ($\lambda = 38\text{ nm}$).

![Error in the center](https://via.placeholder.com/150)

![Error at the edges](https://via.placeholder.com/150)

**Figure 7.9:** Error of the reconstructed wavefront by using the virtual reference

Fig.7.10 shows the errors calculated for the measured reference. One would expect a better accuracy in this case, but in fact the error shows now bigger values. The error is maximal $0.5\lambda$ ($\lambda = 38\text{ nm}$). This can be explained by the experimental setup for the reference measurement. As no plane wave could be generated as a reference wavefront, the spherical wavefront had to be used. In doing so, the distance pinhole-mask gives an additional error source and has to be taken into account. The spot positions on the CCD generated by the reference wavefront depend strongly on these errors. Therefore, the total error introduced by using the measured reference is bigger than the error introduced by the virtual reference. The conclusion is that the virtual reference is the better choice.
Measurements for sFLASH

Figure 7.10: Error of the reconstructed wavefront by using the measured reference
8 Analysis and interpretation

All following results are based on the wavefront which was reconstructed using the virtual reference. It was shown that errors introduced by assuming a perfect Hartmann mask are still small. With the wavefront known on the beam’s full cross section, further analysis of beam parameters is possible.

8.1 Fitting to the Gaussian mode

First of all, the measured wavefront was fitted to that of a Gaussian fundamental mode. This was done by developing an error function which takes into account the wavefront shape (phase $\phi(x, y)$) and also the beam sizes $W_x$ and $W_y$ at sensor position. By minimizing this error function, the optimal set of parameters of a Gaussian mode ($\zeta, w_0, z_R$) fitting the measured phase and intensity distribution is found.

\[
\text{Error function: } \frac{1}{\pi^2} \iint \Delta \phi(x, y, \zeta)^2 \, dx \, dy + \frac{\Delta W_x(\zeta)^2 + \Delta W_y(\zeta)^2}{(W_x + W_y)^2} \tag{8.1}
\]

$\Delta \phi(x, y, \zeta)$ denotes the deviation of the measured phase $\phi^{\text{meas}}(x, y)$ to the phase of the Gaussian mode $\phi^{\text{Gauss}}(x, y, \zeta)$.

$$\Delta \phi(x, y, \zeta) = \phi^{\text{Gauss}}(x, y, \zeta) - \phi^{\text{meas}}(x, y) \tag{8.2}$$

$\Delta W_x(\zeta)$ and $\Delta W_y(\zeta)$ are the deviations in beam size in the $x$- and $y$-coordinate respectively.

$$\Delta W_x(\zeta) = w^{\text{Gauss}}(\zeta) - W_x^{\text{meas}}$$
$$\Delta W_y(\zeta) = w^{\text{Gauss}}(\zeta) - W_y^{\text{meas}} \tag{8.3}$$

Thereby, the beam size $w^{\text{Gauss}}$ and the wavefront of the Gaussian mode $\phi^{\text{Gauss}}$ depend on the waist size $w_0$ and the reduced coordinate $\zeta = z/z_R$ (see equation 2.13 and 2.16 in section 2.1). The results of this fit are shown in table 8.1.

| $w_0$ | 23.3 $\mu$m |
| $z_R$ | 4.5 cm |
| $z$ | 6.545 m |

Table 8.1: Parameters of the fitted Gaussian mode
As the distance from gas-target to sensor position is 7.173 m, the focus of the fitted Gaussian beam is more than a half meter behind the HHG gas-jet.

### 8.2 Calculation of the beam quality factor $M^2$

The $M^2$ can be directly extracted from the beam’s wavefront shape and intensity distribution in only one transversal plane $z$. This is possible by calculating the second spatial moments $< x^2 >$ and $< y^2 >$ of the measured intensity profile and the second angular moments $< \theta_x^2 >$ and $< \theta_y^2 >$ of the measured gradient field.

\[
< x^2 > = \frac{\sum_{i,j} I_{i,j} (x_{i,j} - < x >)^2}{\sum_{i,j} I_{i,j}}, \quad < y^2 > = \frac{\sum_{i,j} I_{i,j} (y_{i,j} - < y >)^2}{\sum_{i,j} I_{i,j}} \quad (8.4)
\]

\[
< \theta_x^2 > = \frac{\sum_{i,j} I_{i,j} (\theta_{x,i,j} - < \theta_x >)^2}{\sum_{i,j} I_{i,j}} + \frac{1}{k^2} \frac{\sum_{i,j} (\partial I_{i,j}/\partial x)^2}{\sum_{i,j} I_{i,j}} \quad (8.5)
\]

\[
< \theta_y^2 > = \frac{\sum_{i,j} I_{i,j} (\theta_{y,i,j} - < \theta_y >)^2}{\sum_{i,j} I_{i,j}} + \frac{1}{k^2} \frac{\sum_{i,j} (\partial I_{i,j}/\partial y)^2}{\sum_{i,j} I_{i,j}} \quad (8.5)
\]

As shown in [46], the beam sizes and divergence angels are calculated from the second moments by

\[
w_x = 2\sqrt{< x^2 >}, \quad w_y = 2\sqrt{< y^2 >},
\]

\[
\text{divergence}_x = 2\sqrt{< \theta_x^2 >}, \quad \text{divergence}_y = 2\sqrt{< \theta_y^2 >}.
\]

The $M^2$-value for each transversal direction is calculated following by

\[
M_x^2 = \frac{4\pi}{\lambda} \sqrt{< x^2 > < \theta_x^2 > - < x\theta_x >^2},
\]

\[
M_y^2 = \frac{4\pi}{\lambda} \sqrt{< y^2 > < \theta_y^2 > - < y\theta_y >^2}.
\]

Thereby, $< x\theta_x >$ and $< y\theta_y >$ respectively is given by

\[
< x\theta_x > = \frac{\int_{-\infty}^{\infty} I(x,y)x\theta_x dxdy}{\int_{-\infty}^{\infty} I(x,y)dxdy} \quad \text{and} \quad < y\theta_y > = \frac{\int_{-\infty}^{\infty} I(x,y)y\theta_y dxdy}{\int_{-\infty}^{\infty} I(x,y)dxdy}.
\]

The results of this calculation are shown in table 8.2.

<table>
<thead>
<tr>
<th>$M_x^2$</th>
<th>$M_y^2$</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>30.4</td>
<td>25.1</td>
<td>1.6</td>
</tr>
</tbody>
</table>

**Table 8.2:** $M^2$ of the seed laser
8.2 Calculation of the beam quality factor $M^2$

The error was evaluated by conducting a Monte-Carlo simulation. The accuracy of the intensity distribution $I_{i,j}$, measured by the CCD chip, was assumed to be 10%. Furthermore, a systematic error due to the inaccuracy of the measured distance $D$ between Hartmann mask and CCD chip has to be included. $D$ was measured with an accuracy of 2 mm. Together with the accuracy of the analyzed spot positions of 10 $\mu m$, the wavefront’s gradients $\theta_{ix}$ and $\theta_{iy}$ are given with an error of 0.1 mrad. Finally, the positions of the holes in the Hartmann mask, $x_{i,j}$ and $y_{i,j}$ respectively, were assumed to be given with an accuracy of 10%. The Monte-Carlo simulation done with all these error sources randomly distributed lead to an accuracy of the analyzed $M^2$-values of 1.6.

According to these $M^2$-values and equation 2.27 on page 15, the waist sizes at the gas target are shown in table 8.3. The error of 80 $\mu m$ was analyzed in the same fashion as presented before.

<table>
<thead>
<tr>
<th>$W_{0x}$</th>
<th>$W_{0y}$</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>849 $\mu m$</td>
<td>660 $\mu m$</td>
<td>80 $\mu m$</td>
</tr>
</tbody>
</table>

Table 8.3: Recalculated beam waist sizes at the gas target

To cross-check the above derived $M^2$-values, they are additionally analyzed in a different way. This approach is directly connected to the emittance $\epsilon$ as used in accelerator physics. The emittance, which defines the quality of a particle beam, corresponds directly to the $M^2$ in laser physics. Both are constant in the linear regime and have in principle the same meaning. Naturally, the positions and gradients of a particle beam in phase-space ($x,x'$) and ($y,y'$) distribute themselves on the surface area of an ellipse. The cross section of this ellipse gives the emittance $\epsilon$ according to

$$\epsilon_x = \sqrt{( <x^2> - <xx'>^2 )}, \quad (8.10)$$

and in the same fashion for the $y$-coordinate. See for example [23] for a description of the emittance in particle physics. The same concept holds for the $M^2$. The Hartmann measurement also derives the positions and gradients of the generated spots. Following, this can be used to analyze the cross sections $A_x$ and $A_y$ of the phase-space distribution in $x$ and $y$ in the same fashion as for the emittance $\epsilon$.

$$A_x = \sqrt{( <x^2> - <x\theta_x'^2> )}, \quad A_y = \sqrt{( <y^2> - <y\theta_y'^2> )}$$

(8.11)

According to this, the $M^2$ is calculated by

$$M^2_x = \frac{A_x}{(4\pi/\lambda)} \quad \text{and} \quad M^2_y = \frac{A_y}{(4\pi/\lambda)}.$$  

(8.12)
In Fig. 8.1 the Hartmann measurement is presented in phase space \((x, \theta_x)\) and \((y, \theta_y)\) respectively. The plotted points are weighted with the intensity of the corresponding spot. Red stands for high intensity, black for low intensity. Additionally, the evaluated phase-space ellipse is shown together with the acquired results for \(A_x\), \(A_y\) and \(M_x^2\), \(M_y^2\) respectively. This results in the same \(M^2\)-values as already derived by using the equations from [46].

\[ A_x = 9.2 \cdot 10^{-2} \, \text{mm-mrad} \]
\[ M_x^2 = \frac{A_x}{4\pi/\lambda} = 30.4 \]

\[ A_y = 7.6 \cdot 10^{-2} \, \text{mm-mrad} \]
\[ M_y^2 = \frac{A_y}{4\pi/\lambda} = 25.1 \]

**Figure 8.1:** Phase-space distribution and evaluated \(M^2\)-values

As one can see in Fig. 8.1, there are a few low intensity outliers. Therefore, the
data analysis have been repeated without taking these points into account. In Fig. 8.2, the newly evaluated ellipses are shown again in the phase-space together with the results of the cross sections $A_x$, $A_y$ and $M^2_x$, $M^2_y$ respectively.

\[ A_x = 3.6 \cdot 10^{-2} \text{ mm-mrad} \]
\[ \Rightarrow M^2_x = \frac{A_x}{4\pi/\lambda} = 11.8 \]

\[ A_y = 3.5 \cdot 10^{-2} \text{ mm-mrad} \]
\[ \Rightarrow M^2_y = \frac{A_y}{4\pi/\lambda} = 11.6 \]

**Figure 8.2:** Phase-space distribution without outliers and evaluated $M^2$-values

As one can see, the exclusion of only a few outliers results in a reduction of the $M^2$. Points far off the mean value are very sensitive to the second moments $<\theta^2_x>$ and $<\theta^2_y>$ respectively, which leads to a bigger $M^2$.

In the above given phase-space plots and their evaluation regarding the $M^2$, all
positions \((x, y)_{i,j}\) and gradients \((\theta_x, \theta_y)_{i,j}\) measured by the Hartmann sensor are analyzed at once. Thinking of the Hartmann mask, there is a regular grid of holes which represent the positions on which the wavefront’s gradients are given. Consider for example the phase-space in \(x\). For each \(x\)-position there are several gradient values, each originating from a different \(y\)-position. In addition to the analysis presented above, the correlation between the \(x\)- and \(y\)-plane has been investigated. If, for example, the \((x, \theta_x)\) phase-space is calculated for each \(y\)-row separately \((i = 1, 2, 3, \ldots)\), the \(M^2\) evaluated for each row individually might show smaller values. However, if there is a correlation between \(x\) and \(y\), the shapes of these individual phase-space distributions are twisted against each other. Therefore, the projected phase-space area is bigger. Accordingly, the \(M^2\) will be larger. Based on these considerations, the phase-space was separately analyzed for each \(y\)-row and for each \(x\)-column respectively. Doing so, it was found that the shapes of the phase-space distributions are only very slightly twisted against each other. Hence, there is actually no correlation between \(x\) and \(y\). The individual phase-space distributions yielded an \(M^2\) of 8 to 12.

The obtained \(M^2\)-values as shown in table 8.2 are used to define the embedded Gaussian beam as described in section 2.3 on page 14. The results are depicted in Fig. 8.3.

![Figure 8.3: Embedded Gaussian beam](image)

In the context of calculating the \(M^2\), note that the \(M^2\) alone does not predict the wavefront shape, e.g. its smoothness and aberration content. Instead, as for example shown in [47], a beam with \(M^2 > 1\) may have perfectly shaped wavefronts. In this case, the intensity distribution would show an aberrated, non-Gaussian profile which leads to larger \(M^2\)-values. On the other hand, when measuring only the
intensity in a transversal plane, its profile might show a perfect Gaussian shape. Again, this information alone does not have any significance to the beam’s $M^2$. Despite the smooth Gaussian profile, the $M^2$ might show large values because of deformed wavefronts. Both, a smooth Gaussian intensity profile and a perfectly shaped wavefront are only given by a laser beam of $M^2 = 1$.

8.3 Propagation with ZEMAX

8.3.1 Propagation to the undulators

Considering the interaction between the seed and electron beam, the start conditions regarding wavefront shape and intensity distribution at undulator begin have to be found. As already mentioned, the knowledge of the $M^2$ alone is insufficient to predict the aberration content of the wavefront. Both, wavefront and intensity distribution of the real seed have to be known for further simulations of the FEL process. Therefore, beam propagation of the real seed is done along the sFLASH injection beamline in the same fashion as shown in chapter 4. The numerical propagation mode in ZEMAX allows to propagate the beam to the undulator. The start conditions for that computation are given by the measured wavefront and the corresponding intensity distribution. Thus analyzing the intensity distribution and wavefront in the area of the undulator begin. The propagation was done with a focal length of 7 m of the focusing mirror. The propagation was also done for the fitted fundamental Gaussian mode serving as a reference. In Fig.8.4 the beam sizes around the undulator begin are depicted.

![Diagram of beam sizes around the undulator begin](image)

**Figure 8.4:** Results of propagation to undulators with ZEMAX
8 Analysis and interpretation

The beam is still converging with the waist around 1 m inside the first undulator. The beam sizes exactly at undulator begin are shown in table 8.4.

<table>
<thead>
<tr>
<th>beam size x</th>
<th>213 µm</th>
</tr>
</thead>
<tbody>
<tr>
<td>beam size y</td>
<td>115 µm</td>
</tr>
</tbody>
</table>

Table 8.4: Beam sizes at undulator begin

The intensity distribution and the wavefront at undulator begin are shown in Fig. 8.5 and 8.6.

Figure 8.5: Intensity distribution at undulator begin
8.3.2 Back-propagation to the HHG source

Analogously, the beam propagation of the real seed is done to the other direction. Thus analyzing the intensity distribution in the area of the HHG source. Again, the propagation was also done for the fitted Gaussian mode wavefront to serve as a reference. The beam sizes were analyzed in an area of 2m, with the gas target being in the center. Fig.8.7 shows the results.
8 Analysis and interpretation

8.4 GENESIS simulation with the measured wavefront

The above shown results of the propagation along the sFLASH injection beam line were used to make a Genesis simulation of the real seed. Besides the real seed, also the simulated seeds as presented in chapter 4 are shown. This gives a direct comparison of the real case to the theoretical. Firstly, the overall power of the seed was assumed to be 50 kW. The power development of the seed is shown in Fig.8.8. It is shown that the real seed has the lowest power gain.

The same simulations were also done with an initial seed power of 2 kW. This corresponds to the current available seed power at sFLASH. The result is shown in figure 8.9. As expected, the power gain is reduced by a significant amount.

![Figure 8.8: Power of the real seed (initial power 50 kW)]
8.4 GENESIS simulation with the measured wavefront

Figure 8.9: Power of the real seed (initial power 2 kW)
9 Summary

Investigations on wavefronts of the seed at sFLASH were carried out. Firstly, simulations done with modeled wavefronts showed a huge impact on the outcome of the whole seeding experiment. Propagation of deformed wavefront shapes lead to poor coupling between electron beam and laser seed in the area of the undulator. If the overlap between seed and electron beam is insufficient, only a fraction of the seed will be able to initialize the emission of FEL radiation, which results in a low FEL gain. Therefore, the shape of the wavefront is crucial for successful seeding.

As a second part of this thesis, a Hartmann wavefront sensor for the XUV spectral range was designed and commissioned in the diagnostic section. The Hartmann mask, being the main part of the sensor, was fabricated with high precision as confirmed by measurements. In addition, a software tool to reconstruct wavefronts was developped. The main concept of the reconstruction algorithm is based on Zernike polynomials to transform the measurement’s raw data into wavefront shapes. Error sources were analyzed using Monte-Carlo technique and lead to a maximum error of $0.2\lambda$. Mainly the construction of a reference, which is a crucial factor to the accuracy of the measurements, caused problems due to environmental conditions. The measurements of the wavefront were further analyzed in order to obtain the seed’s beam quality factor $M^2$. The analysis shows values of $M^2$ in the order of 30 down to 11.6 when excluding low intensity outliers which represent 5% of the measured data.

Outlook

The current layout of the diagnostic beam line, where the sensor was build up, did not allow to insert a point source reasonable far away of the sensor. The reference measurement can be further improved with a longer diagnostic branch. The generation of a plane wave would reduce the error of the measurements.

Furthermore, the diagnostic beam line uses all in all eight mirrors which induce errors to the wavefront. These errors could be eliminated by a straight beam line with only one mirror deflecting the seed out of the sFLASH beam line in the tunnel. With a measurement of the wavefront directly in front of the undulator, the numerical simulation of beam propagation could be omitted. Thus, deriving more reliable results,
9 Summary

as the simulation process might induce errors. Unfortunately, a wavefront measure-
ment in front of the undulators is technically very difficult. There is practically no
possibility to build an extra experimental setup directly before the undulators as
the existing environment in the tunnel must not be disturbed.
Bibliography


[42] Carl Zeiss AG, Germany


Bibliography


Acknowledgements

I am very grateful to my advisor Prof. Dr. Jörg Roßbach for giving me the opportunity to prepare my thesis at DESY. He introduced me to the field of accelerator physics, whereupon I decided to work in this field. I value the thought-provoking impulses and the support he gave me throughout my thesis.

More than anyone else, I want to thank my second advisor Dr. Velizar Miltchev for his steady support. In the many fruitful discussions I gained knowledge in diverse fields of accelerator physics. He patiently helped to prepare my work and whenever I had problems, I could ask him for advice.

Special thanks to the specialists of the HHG source, Theophilos Maltezopoulos and Manuel Mittenzwey. They provided valuable help in the experimental setup of the Hartmann sensor and spent much time supporting me during my measurements. Moreover, they were always willing to listen to my many questions.

For their helpful advices concerning laser physics, I want to thank Marek Wieland and Armin Azima. The interesting discussions with them were very promoting during my thesis.

I am grateful to Jörn Bödewadt for his support. He supported me with important hints concerning computer problems, especially ZEMAX. But he also gave many important advices to my theoretical work.