

LUND UNIVERSITY

FYSM60

Manipulation and application of laser light using
spatial light modulation

Author:

Albin HEDSE

Thesis submitted for M.Sc.

Project duration: 8 months

Department of Physics

Division of Atomic Physics

Supervisors

Cord ARNOLD

Robin WEISSENBILDER



LUND
UNIVERSITY

June 2020

Populärvetenskaplig introduktion: Färgblödning i hög hastighet

Färgblödning, eller mer formellt kromatisk aberration, uppstår när ljus som består av mer än en våglängd, till exempel vitt ljus, inte kan fokuseras till samma punkt av en lins, såsom illustreras i Figur 1. I detta fall orsakas det av att ljus av olika våglängder bryts olika beroende på materialet det passerar igenom. Ett väldigt välkänt exempel av detta är avbildat på albumomslaget till Pink Floyds "The Dark Side of the Moon", som visar en stråle av vitt ljus som skickas genom ett glasprisma vilket får dess färgkomponenter att spridas ut.

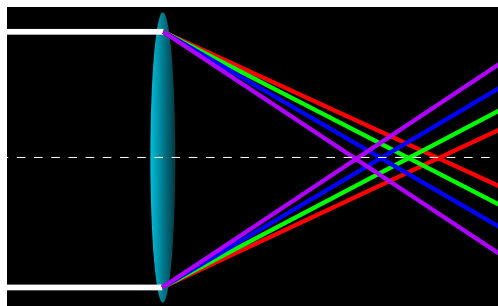


Figure 1: En illustration av kromatisk aberration, där ljus av olika färger fokuseras i olika punkter, vilket innebär att fokuset blir utspritt.

Detta åtgärdas oftast med användningen av en akromatisk lins, det vill säga en kompositlins som består av två eller flera linser av olika material och kurvatur som är konstruerad för att fokusera ljus i samma punkt oberoende av dess våglängd. Den kromatiska aberration som kommer att diskuteras här är däremot inte densamma som färgblödning under vanliga omständigheter, utan kommer att vara den som påträffas inom attofysik.

Så vad är attofysik? Attofysik är ett forskningsområde där reaktioner och processer på attosekundsskalor (10^{-18} sekunder) undersöks med hjälp av ultrakorta laserpulser. Dessa tidsskalor är för korta för att en tillräckligt kort laserpuls ska kunna genereras av en vanlig laser, därför måste de istället genereras i en gas genom svängningarna av en stark, fokuserad, laserpuls. Då skapas skarpt ultraviolettera pulser som har en pulslängd på hundratals attosekunder, så kallade attosekundpulser. Dessa pulser lider dock av en väldigt stark kromatisk aberration när de fokuseras trots att akromatiska instrument används för detta.

Detta sker på grund av att de olika våglängderna som genereras påverkas olika mycket av drivlaserns intensitet och har olika spridningsvinklar, så kallad divergens, som en konsekvens av detta. Det är alltså på grund av att olika våglängdskomponenter har annorlunda divergens som kromatisk aberration sker när attosekundpulser fokuseras. För att åtgärda detta så krävs det då att den drivande laserns intensitet jämnas ut över hela generationsområdet, så att intensiteten är likadan överallt. Detta hade inneburit att skillnaden i divergens hade minskats skarpt mellan de olika attosekundpulskomponenterna. Detta hade i sin tur möjliggjort högkvalitativa fokus.

Syftet med detta projekt är alltså att försöka utveckla en metod som möjliggör bra fokusering av attosekundpulser. Detta görs här genom manipulering av drivlaserns vågfront med hjälp av optiska instrument. Målet är alltså att kunna korrigera optiska aberrationer i drivlasern och manipulera den på ett sådant sätt att en jämn intensitet uppnås för att den kromatiska aberration som förhindrar fokusering av attopulser ska kunna elimineras.

Abstract

Attosecond pulses obtained via high harmonic generation often suffer from chromatic aberration. This is inherent from the generation process as a consequence of the spatial profile of the driving laser pulses. Consequently, the focus positions and divergences of consecutive harmonics vary, leading to chromatic aberration which prevents high-quality refocusing as is critically required for applications.

The aim of this project is to develop a way of diminishing this chromatic aberration through spatial phase manipulation of the driving laser pulses. This is done through use of a spatial light modulator in conjunction with a phase-retrieval and beam-shaping program which was developed during the project. Harmonic spectra and driving beam foci are recorded with an extreme ultraviolet spectrometer and an infrared-sensitive camera, respectively.

It is shown that the spatial light modulator in conjunction with the written program can successfully correct aberrated beams, leading to an increase in peak intensity and a decrease in beam width near the focus. However, reducing the divergence of generated harmonics through beam shaping turned out to be complicated.

It was concluded that the ability to measure the average harmonic intensities and peak intensities of the driving field over several iterations during the experiment is a necessary step. Additionally, different types of focus-improving techniques, such as other phase-retrieval algorithms and Zernike optimization could be used in future iterations to potentially improve focus correction, which would assist beam shaping.

Acknowledgements

I would like to thank my supervisor, Cord Arnold, for the time spent in the lab with me, the knowledge shared with me and, perhaps most importantly, his patience with me. I have learned a lot about optics and science in general, both in theory and in practice, thanks to you.

I would also like to thank my co-supervisor, Robin Weissenbilder, for teaching me about the laboratory, the discussions we had and all the help during the experimental work. I wouldn't have nearly as good an understanding of the laboratory and its procedures without you.

Of course, I also need to thank many other people as well. Chen Guo at the atomic physics department, who I would like to thank for working in the laboratory with us at several points and for helping me figure out the electronics of the setup. Åke Johansson, the system administrator at atomic physics, for being immensely helpful with all things computer-related. I also have to thank Samuel Bengtsson who, apart from being one of the friendliest people I have met, also spent a lot of time with me discussing the spatial light modulator and how to use it in the lab.

Apart from these people, there are many others at the department of atomic physics who have been incredibly helpful and welcoming during the course of my project. I would therefore also like to extend thanks to everyone I haven't already mentioned at the department of atomic physics as well for taking an interest in my project, helping me, all the discussions we've had and of course, the fun chats during the coffee breaks. You helped make this a very fun and interesting experience and I have learned much from all of you.

Last but not least, I also want to thank my friends and family for all the fun we've had and for your support during the past years. I'm a very lucky person to have met you all, you're great and I couldn't have done it without you!

Contents

Populärvetenskaplig introduktion	iii
Abstract	iv
Acknowledgements	v
Acronyms	vii
1 Introduction	1
2 Theory	2
2.1 High Harmonic generation and Attosecond pulses	2
2.1.1 High Harmonic generation: The semi-classical model and its consequences	2
2.1.2 High Harmonic generation: Harmonics and attosecond pulses	5
2.2 Fourier optics	7
2.2.1 Free-space light propagation	7
2.2.2 The Fourier transform property of converging lenses	9
2.2.3 Sampling for discrete Fourier transforms	10
2.3 Phase-retrieval algorithms	11
2.3.1 The Gerchberg-Saxton algorithm	12
2.3.2 Input-output algorithms	13
2.3.3 Three-plane intensity algorithm	17
2.3.4 Multi-image Gerchberg-Saxton algorithm	17
2.3.5 Spatial phase perturbation algorithm	17
2.3.6 GS/HIOA	20
2.3.7 Zernike Polynomials	21
2.4 Phase modulation with spatial light modulators	22
3 Implementation	24
3.1 Experimental Setup 1	24
3.2 Experimental Setup 2	26
3.3 Experimental Setup 3	27
3.3.1 Harmonic generation, pre-generation chamber	27
3.3.2 Harmonic generation, generation chamber	27
3.3.3 Harmonic generation, experiment	27
3.4 Program for phase retrieval	30
3.4.1 Fill-fixing program	31
3.4.2 Main program	31
4 Results	34
4.1 Experimental sections 1 and 2	34
4.2 Experimental section 3	38
5 Conclusion and outlook	45
References	46
Appendix: Full HHG spectra	48

Acronyms

CCD Charge-coupled device.

GS/HIOA Hybrid Gerchberg-Saxton and hybrid input-output algorithm.

GSA Gerchberg-Saxton algorithm.

HHG High harmonic generation.

HIOA Hybrid input-output algorithm.

I/O Input-output.

IFT Inverse Fourier transform.

LCoS Liquid crystal on silicon.

MMAA Multiple-measurement averaging algorithm.

O/O Output-output.

SLM Spatial light modulator.

SPPA Spatial phase perturbation algorithm.

XUV Extreme ultraviolet.

1 Introduction

What is a harmonic? A harmonic refers to a signal which has a frequency which is an integer multiple of some reference frequency. The first thing that may come to mind when harmonics are brought up might be music, where acoustic harmonics are often used in the composition of songs. Here, however, the harmonics of interest are instead harmonics of light. In 1961 it was first observed that strong laser light could cause the emission of light harmonics [1]. The process of generating several light harmonics of higher orders is referred to as High harmonic generation (HHG). Similarly to the generation of harmonic overtones in music instruments, light harmonics are often very weak compared to their driving field. Unlike acoustic harmonics, however, in HHG there is a tendency for light harmonics to form "plateaus" of similar intensity up to a certain point, known as a cutoff [2]. This means that harmonic pulses from HHG consist of many consecutive, but similarly intense, odd harmonics. These harmonics are generated in time with the oscillations of the driving laser field, which is very quick even for light with longer wavelengths than visible light! As an example, infrared light has frequencies between 1-400 THz [3]. This makes high harmonic pulses excellent for pump-probe experiments aimed at observing processes which occur at very short time scales [4]. In fact HHG is very important within the field of attosecond (10^{-18} s) physics. Attosecond physics today makes use of HHG to produce ultrashort pulses of Extreme ultraviolet (XUV) light with large bandwidths, that is, light constituted by a large variety of frequencies, like an extremely energetic version of white light. Attosecond pulses are interesting because of their ability to probe effects at very short time scales, such as the ionization time delay of an atom [5]. Attosecond pulses do however have a problem with chromatic aberration when they are refocused. This simply means that the light which is generated in HHG can not be refocused into a small spot easily, but rather smears out in space around where one might expect the focus to be. Normally this can happen when any light with a large bandwidth passes through a lens which is not made to accommodate a bandwidth that large. In the case of HHG, however, this is not only caused by the component wavelengths being very different from one another, as there are optics which can, to some extent, adjust for this. It is also caused by their different spatial wavefronts as they are generated. This effectively means that most generated harmonics will focus in different places, with several harmonics not forming a focus after generation at all. These harmonics are referred to as having a "virtual" focus, because they propagate as though their focus lies before the point at which they were generated. This variation in focus positions makes it tough to refocus these harmonic pulses, even when disregarding the difference in wavelength. It has been suggested that this can be mitigated through controlling the spatial intensity profile of the driving laser [6] and that is what this project sets out to do.

The basic aim of the project is therefore relatively simple: namely to investigate the possibility of reducing this aforementioned chromatic aberration by manipulating the spatial intensity profile of the driving field. This will be carried out by shaping the spatial phase of the driving laser by using a Spatial light modulator (SLM) such that the wavefront causes the beam to propagate in such a way that an intensity profile better suited for low-divergence HHG is obtained. The SLM in essence functions by slowing down different parts of an incoming beam differently. This manipulates the wavefront of the beam and can be used to control, to some degree, how a beam of light develops in terms of intensity as it propagates through space. Here, this type of spatial beam shaping has been used to modify the beam

profile used in driving HHG. In order to then observe the divergence of these harmonics, they were recorded with an XUV photon spectrometer.

Additionally, the project will aim to develop a method to use the SLM for wavefront correction, that is, using an SLM to compensate for optical aberrations in a system. This is similar to shaping a beam with the SLM, but requires one to know the spatial phase of the incoming light. This can be found through use of a phase-retrieval algorithm, that is, an algorithm which calculates the phase of a beam given that its intensity profile can be measured in at least two planes. One can then use the SLM to give the beam a flat wavefront which ideally improves the beam's focus. This may also be useful in order to be able to implement the beam shaping mentioned previously, as a phase meant for beam shaping may not interact in a desired way with an aberrated phase. Using a phase-retrieval algorithm in conjunction with an SLM for the purpose of wavefront correction has been done before [7], the novelty here lies in using it in conjunction with the beam shaping properties of the SLM in the process of HHG. The phase-retrieval algorithm used will be evaluated based on its improvement of the focus, which most clearly shows itself with the increase in peak intensity.

To summarize, the main aim of this project is to use an SLM to reduce aberrations and shape the driving beam used in HHG. This was done in order to diminish the chromatic aberration observed upon the refocusing of attosecond pulses. This was carried out through the development of a phase-retrieval algorithm and integrating it and an SLM into a currently existing experimental setup. HHG was performed and divergence measurements of harmonics were taken and evaluated.

2 Theory

2.1 High Harmonic generation and Attosecond pulses

2.1.1 High Harmonic generation: The semi-classical model and its consequences

HHG in a gas is fairly well described with a semi-classical model which is structured in three steps. Firstly, the driving laser field distorts the atomic potential, causing it to oscillate periodically along with the laser field. When the atom's potential then is strongly distorted, tunnel ionization becomes possible. Secondly, immediately after ionization, the electron is accelerated away from the atom, but as the electric field oscillates in the other direction, the electron is accelerated back towards the atom. Thirdly it may recombine with the parent ion and emit a photon with an energy equivalent to the ionization energy of the atom and the electron's kinetic energy upon return. This process is illustrated in Figure 2 and is a fairly intuitive image to someone who is familiar with atomic physics. It can also be used to provide a derivation of the energy of the harmonics.

If we treat the electron immediately after ionization classically, we can get rather simple, Newtonian, expressions for the electron's motion. This is largely accomplished by neglecting any effect not due to the electric field of the laser, which we can afford to do due to the field strength of the laser. The expression for the electron's acceleration can in this classical

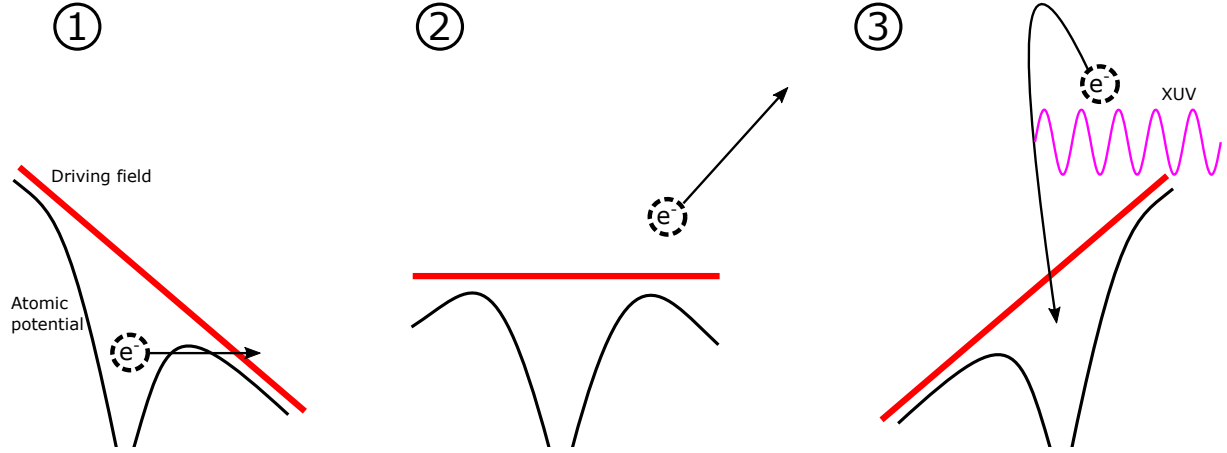


Figure 2: This figure illustrates the semi-classical three-step model of HHG. Step 1 is the tunnel ionization of an electron out of a distorted atomic potential. Step 2 is the acceleration of the electron after its exit from parent atom and 3 is the recapture of the electron as the atomic potential has shifted due to the external driving field, leading to the emission of XUV radiation.

model be described as:

$$\frac{d^2x}{dt^2} = -\frac{eE_0}{m_e} \sin(\omega t) \quad (1)$$

where x is the position of the electron, t represents time, e is the elementary charge, m_e is the electron mass, E_0 is the electric field strength of the laser, and ω is the angular frequency of the laser field. We also have a few initial conditions to set into place, namely that the time at which the electron is ionized is $t = t_0$ and that $x(t_0) = 0$ and $\frac{dx(t_0)}{dt} = 0$. If we then integrate Equation 1 with respect to time with these parameters in mind, we get:

$$\frac{dx}{dt} = \frac{eE_0}{m_e\omega} [\cos(\omega t) - \cos(\omega t_0)], \quad (2)$$

and subsequently:

$$x = \frac{eE_0}{m_e\omega^2} [\sin(\omega t) - \sin(\omega t_0) - \omega(t - t_0) \cos(\omega t_0)]. \quad (3)$$

Equations 2 and 3 give us the ability to classically calculate the kinetic energy of the electron as it recombines and its recombination time respectively. The kinetic energy, using Equation 2 is given by the expression:

$$E_k = \frac{1}{2}m \left(\frac{dx}{dt} \right)^2 = \frac{e^2E_0^2}{2m\omega} [\cos(\omega t) - \cos(\omega t_0)]^2 = 2U_p[\cos(\omega t) - \cos(\omega t_0)]^2. \quad (4)$$

Where U_p is the ponderomotive energy. The maximum kinetic energy of the electron can, using Equation 4, be calculated to be $3.17U_p$ [8]. Since intensity is proportional to the square of the electric field strength, this also gives that the ponderomotive energy is proportional to the laser intensity [3]. With this, it becomes apparent that both recombination time and kinetic energy upon recombination are dependent upon the ionization time. Additionally, if one plots the frequency of emitted photons at recombination as a function of return time,

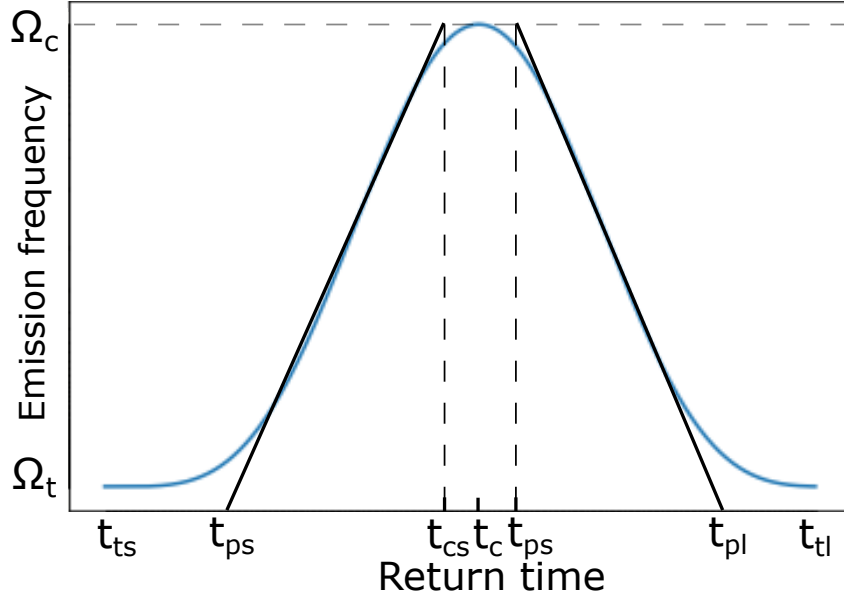


Figure 3: This graph shows the emission frequency caused by HHG as a function of electron return time. The two black lines represent the linear approximations of the recombination time as given by Equation 5. Where t_{ts} and t_{tl} are the threshold return times for short and long trajectories, t_{ps} and t_{pl} are the approximated threshold return times for the linear approximation of the short and long trajectories respectively. Similarly, t_c is the actual return time for the cutoff frequency with t_{cs} and t_{cl} being the approximated cutoff times for the short and long trajectories respectively.

as is illustrated in Figure 3 one can observe that each frequency, apart from the maximum, is represented twice, one for a shorter, and one for a longer return time. These are referred to as the 'short' and 'long' electron trajectories which represent identical energies but for different recombination times. For a significant portion of these two trajectories, emission frequency can be approximated as a linear function of return time [6]. For our purposes, we will use a simple function with recombination time as a function of frequency:

$$t_i(\Omega) = t_{pi} + (t_{ci} - t_{pi}) \frac{\Omega - \Omega_t}{\Omega_c - \Omega_t}, \quad (5)$$

where i refers to the short or long trajectory depending on which one is being considered, t_{pi} and t_{ci} refer to the threshold return times and cutoff return times respectively for the different linearly approximated trajectories, whereas Ω , Ω_c and Ω_t represent the frequency, cutoff frequency and threshold frequency of our emissions, respectively. Most of these quantities are illustrated in Figure 5. If we then assume there are no other significant effects on the emission times, we can treat $t_i(\Omega)$ as a delay of our emissions and integrate Equation 5 with respect to $\Omega - \Omega_t$ and treat this as an expression for the spectral phase [6]:

$$\Phi_i(\Omega) = \Phi_i(\Omega_t) + t_{ti}(\Omega - \Omega_t) + \frac{t_{ci} - t_{ti}}{2} \frac{(\Omega - \Omega_t)^2}{\Omega_c - \Omega_t}. \quad (6)$$

From our previous equations of motion, we know that the recombination times are independent of the energy of the laser field, we also know that the energy threshold of the photons is determined by the ionization potential of the atom itself and the maximal kinetic energy

of the electron, giving the expression $\hbar\Omega_c = (3.17U_p + U_I)$, where \hbar is the reduced Planck constant and U_I represents the ionization potential. This means that the second term in Equation 6 is independent of the intensity of the driving laser. One can then use the following expression:

$$\Omega_c - \Omega_t = 3.17 \frac{U_p}{\hbar} = 3.17 \frac{e^2 I \lambda^2}{8\pi^2 c^3 \hbar \epsilon_0 m_e}, \quad (7)$$

where I is the intensity of the driving field, c is the speed of light in vacuum and ϵ_0 is the vacuum permittivity [8]. Using Equation 7 in the third term of Equation 6, one gets that the third term is inversely proportional to the intensity, with a similar magnitude, but opposite sign between the short and long-trajectory versions [6]. Additionally, the first term is proportional to the laser intensity for the long trajectory and zero for the short, with a constant of proportionality several orders of magnitude larger than that of the third term, causing the first-order term to outweigh it in long-trajectory phase calculations [6]. In effect, this means that differences in the driving laser intensity have a much larger and more visible effect on the phase of long-trajectory emissions relative to the short-trajectory ones. An effect of the third term is that the dipole phases of different harmonics are affected differently by intensity. This means that there is, via dipole coupling, a direct link between the spatial intensity profile of the driving field and the spatial phase profile of the generated harmonics. Since the frequency of the radiation in each harmonic is different, this means that if the driving intensity is not uniform in the generation volume, different harmonics will have different divergences. Thus their phase and thereby their wavefront, will be impacted differently by a non-homogeneous intensity in the generation volume. This is what causes the difference in focal positions at generation. In practice, this leads to chromatic aberration when one attempts to refocus these harmonics, as illustrated in Figure 4. Thus, if one intends to refocus the harmonic pulses at a later time, one might wish to use an intensity distribution which is as flat as possible when generating harmonics. This would lead to a flat phase difference between harmonics and would ideally give all harmonics similarly positioned foci, reducing the chromatic aberration, hence increasing the quality of the focus.

2.1.2 High Harmonic generation: Harmonics and attosecond pulses

Given equations 1, 2 and 3 above, one can see that for a constant driving field, or at least a several-cycle pulsed driving field, one could anticipate an emitted pulse of light for every half-cycle of the driving field with a sign inversion between each half-cycle pulse. If one wants to look at the interference between these two pulses, one gets the following expression for the electric field:

$$\mathbf{E}(t) = \mathbf{E}_1(t) - \mathbf{E}_1\left(t - \frac{T}{2}\right), \quad (8)$$

where $\mathbf{E}(t)$ represents the total electric field, $\mathbf{E}_1(t)$ represents the electric field of a single half-cycle pulse and T represents the time it takes to one complete pulse cycle of the driving field. We also want to use the spectral amplitudes, given by FT, defined as follows [8]:

$$\mathbf{A}(\Omega) = \int \mathbf{E}(t) \exp(-i\Omega t) dt. \quad (9)$$

Then we insert Equation 8 into Equation 9 and define $\mathbf{A}_1(\Omega)$ so that we get:

$$\mathbf{A} = \mathbf{A}_1(\Omega) - \int \mathbf{E}_1\left(t - \frac{T}{2}\right) \exp[-i\Omega t] dt.$$

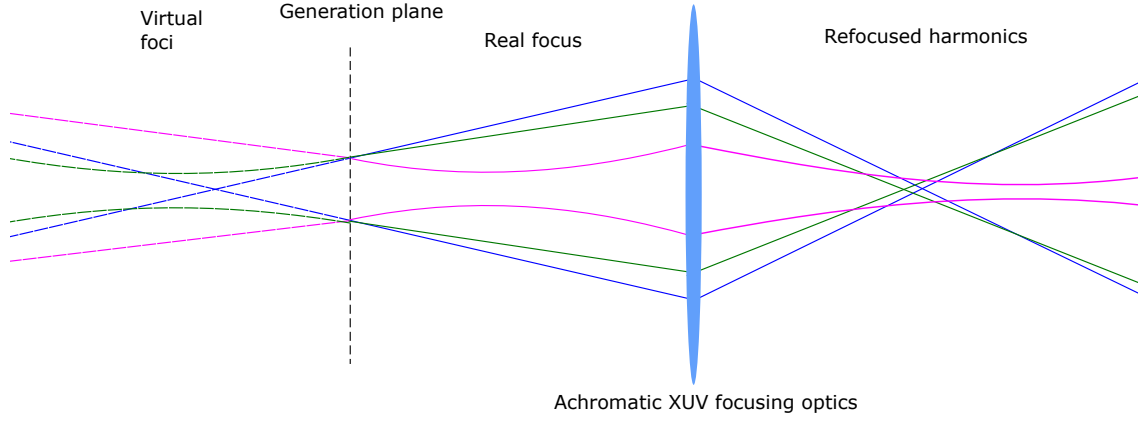


Figure 4: This figure illustrates an attempt at achromatic refocusing of several harmonics from HHG. The differently colored lines represent the beam widths of different harmonics. When generated, due to the volume of generation being the same for most harmonics, the beam width of the harmonic pulse upon generation is fairly uniform. However, this means that if all generated harmonics have different phase fronts, their foci will all be located in different planes along the optical axis of the system, as illustrated. This also means that an attempt to achromatically refocus these beams by means of some advanced optics, here represented by a lens-like object, will result in different harmonics being focused in different spots, leading to chromatic aberration.

We can, using our knowledge of the function of the electric field, get:

$$\mathbf{A}(\Omega) = \mathbf{A}_1(\Omega) - \int \mathbf{E}_1(t) \exp\left[-i\Omega t + \frac{T}{2}\right] dt = \mathbf{A}_1(\Omega) \left(1 - \exp\left[-i\Omega \frac{T}{2}\right]\right).$$

If we look at what this means in terms of physical frequencies by looking at the absolute square of the spectral amplitude, we get:

$$|\mathbf{A}(\Omega)|^2 = 4|\mathbf{A}_1(\Omega)|^2 \sin^2\left(\frac{\Omega T}{4}\right)$$

We can see that this gives $|\mathbf{A}(\Omega)|^2 = 0$ for $\frac{\Omega T}{4} = n\pi$ where n is an integer, and maximal values when $\frac{\Omega T}{4} = \frac{(2n+1)\pi}{2}$. This in turn means that when the frequency $\Omega = \frac{4n\pi}{T} = 2n\omega$, $|\mathbf{A}(\Omega)|^2 = 0$, meaning that even harmonics of the driving field are not emitted. We can also see that the values of Ω at the maxima of $|\mathbf{A}(\Omega)|^2$ are $\Omega = \frac{(2n+1)2\pi}{T} = (2n+1)\omega$, which corresponds to odd harmonics of the driving field. This pattern can be further generalized for larger amounts of pulses [8] and explains why only odd orders of harmonics are observed for this type of HHG, this is reflected in spectrometer measurements of the emitted harmonics as well. Additionally, the fact that the emission of these pulses occurs during a fraction of a cycle of the driving field gives them a duration in the low femtoseconds or high attoseconds with some chirp. This also means that each pulse of the driving field produces a pulse train of harmonic pulses whose envelope is related to the amplitude of the driving field.

2.2 Fourier optics

2.2.1 Free-space light propagation

When one considers the propagation of light, one of the simplest possible cases is that of a plane wave, which has a complex amplitude $U(x, y, z) = A \exp[-i(k_x x + k_y y + k_z z)]$ [3], where x , y and z represent spatial coordinates, i is the imaginary unit, A is the complex envelope of the wave, and k_x , k_y and k_z are the x , y and z components, respectively, of the wavevector \mathbf{k} . The wavenumber, which is the magnitude of the wavevector, is given by $k = \frac{2\pi}{\lambda}$, with: $k = \sqrt{k_x^2 + k_y^2 + k_z^2}$. If we think of z as the "main axis" of our system, we can define the z -component in terms of its angles in the x and y direction relative to the z -axis, $\theta_x = \sin^{-1}(k_x/k)$ and $\theta_y = \sin^{-1}(k_y/k)$ respectively, this definition will be useful later. If we then consider the complex amplitude in the plane $(x, y, 0)$, it is in effect identical to a harmonic function $f(x, y) = A \exp[-i2\pi(\nu_x x + \nu_y y)]$ where $\nu_x = \frac{k_x}{2\pi}$ and $\nu_y = \frac{k_y}{2\pi}$, partially illustrated in Figure 5. Given the above results, we can now use the angles which were defined before to give us the expression: $\theta_x = \sin^{-1} \frac{1}{\nu_x \lambda}$ and $\theta_y = \sin^{-1} \frac{1}{\nu_y \lambda}$, which for sufficiently small angles θ_x and θ_y , can be approximated to be $\theta_x \approx \frac{1}{\nu_x \lambda}$ and $\theta_y \approx \frac{1}{\nu_y \lambda}$. This approximation is valid when we deal with light which is well collimated enough to follow the paraxial approximation [3], such as a Gaussian beam, and this condition will be assumed to be fulfilled from this point on.

The knowledge that any plane wave has a corresponding 2-dimensional harmonic function is highly useful. The reason for this is that we to a fairly good degree can approximate most paraxial waves as a superposition of plane waves with a small angular spread relative to the optical axis. By extension, this means that in a given plane, a paraxial wave can be considered as a superposition of many harmonic functions. Because of the fact that both the FT and the Inverse Fourier transform (IFT) deal with decomposing a function into its harmonic wave constituents and the inverse, this becomes the basis of Fourier optics.

Now, an important part of optics is describing how light propagates, one of the most basic examples of this being a wave propagating in free space between two planes, separated by some distance d . If we assume our wave to be monochromatic with some wavelength λ , this means our whole system is shift-invariant and can therefore be completely characterized by its own transfer function [3]. We can then set the source plane to be at $z = 0$ and the target plane at $z = d$, with complex amplitudes $f(x, y) = U(x, y, 0)$ and $g(x, y) = U(x, y, d)$ respectively. Since we want to find the transfer function of the system we can use the function $f(x, y) = A \exp[-i2\pi(\nu_x x + \nu_y y)]$ from before, which we know at $z = 0$ corresponds to the harmonic decomposition of the plane wave $U(x, y, z) = A \exp[-i(k_x x + k_y y + k_z z)]$. This means that our target plane at $z = d$ is $g(x, y) = A \exp[-i(k_x x + k_y y + k_z d)]$. Since the transfer function is simply the factor by which our harmonic source function is multiplied by to produce the target function, we can get it via division of the two functions. This yields

$$\mathbf{H}(\nu_x, \nu_y, d) = \frac{g(x, y)}{f(x, y)} = \frac{U(x, y, d)}{U(x, y, 0)} = \exp[-ik_z d] = \exp\left[-i2\pi d \sqrt{\lambda^{-2} - \nu_x^2 - \nu_y^2}\right],$$

which is known as the transfer function of free space.

The transfer function of free space can be further simplified if we take into consideration that we are working within the paraxial approximation. We can then use the same approximation

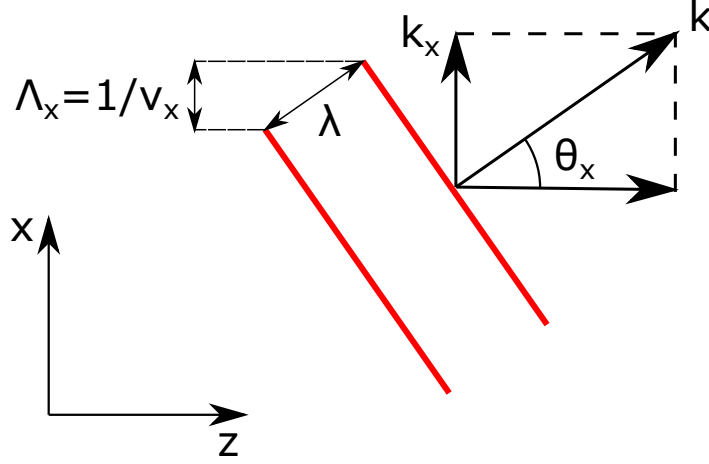


Figure 5: This figure illustrates the decomposition of the wavevector of a plane wave, depicted in red, in two dimensions. k_x represents the projection of the wavevector onto the x axis, and Λ_x refers to the spatial period of the light wave along the x axis. It is the frequency of this period in space which is referred to as the spatial frequency ν_x . This frequency can then be used to describe the harmonic function which characterizes a plane wave in space passing through a single two-dimensional plane.

as before to relate the plane wave angles and spatial frequencies, such that $\theta^2 = \theta_x^2 + \theta_y^2 \approx \lambda^2(\nu_x + \nu_y)$ which gives us $\mathbf{H}(\nu_x, \nu_y, d) \approx \exp \left[-i \frac{2\pi d}{\lambda} \sqrt{1 - \theta} \right]$, which with a Taylor expansion becomes $\mathbf{H}(\nu_x, \nu_y, d) \approx \exp \left[-i \frac{2\pi d}{\lambda} \left(1 - \frac{\theta^2}{2} + \dots \right) \right]$. We then neglect all orders over the second, which is also justifiable due to the paraxial assumption, leaving us with the so-called Fresnel approximation of the free space transfer function:

$$\mathbf{H}(\nu_x, \nu_y, d) \approx \exp[-ikd] \exp[i\pi d \lambda(\nu_x^2 + \nu_y^2)] \quad (10)$$

This simplified transfer function can then be used to calculate the propagation of more complicated waves than a plane wave. However, since we normally work with the complex amplitude of a wave, we need to decompose it into its harmonic wave components by using the FT. The mathematical definition of the FT is as follows:

$$F(\nu_x, \nu_y) = \iint_{-\infty}^{\infty} f(x, y) \exp[i2\pi(\nu_x x + \nu_y y)] dx dy,$$

Where we use that $f(x, y) = U(x, y, 0)$, such that we perform the operation $\mathcal{F}U(x, y, 0) = F(\nu_x, \nu_y)$ where \mathcal{F} signifies the FT. Now that the decomposition of the harmonic has been acquired, we may simply multiply it by the Fresnel propagator with d set to an appropriate distance and the λ of our wave, leaving us with $\mathbf{H}(\nu_x, \nu_y, d) \cdot F(\nu_x, \nu_y)$. Now, in order to find out what the wave looks like after propagation to our target plane we need to reconstitute the complex wave function from its harmonic components. This is done via IFT, which is

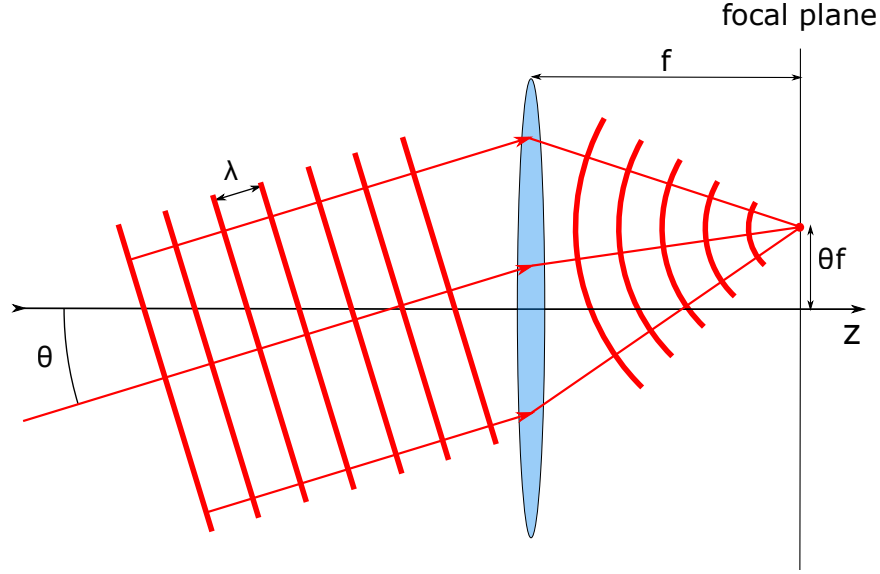


Figure 6: This figure illustrates how a plane wave incident on a thin lens focuses in the focal plane of said lens. Notably, this figure is only accurate when small angles relative to the optical axis of the lens, here z , are considered.

defined as:

$$f(x, y) = \iint_{-\infty}^{\infty} F(\nu_x, \nu_y) \exp[-i2\pi(\nu_x x + \nu_y y)] d\nu_x d\nu_y.$$

If we perform the IFT on the expression $\mathbf{H}(\nu_x, \nu_y, d) \cdot F(\nu_x, \nu_y)$ into the above equation, this leaves us with our complex amplitude $g(x, y) = U(x, y, d)$ calculated by the operation

$$g(x, y) = U(x, y, d) = \mathcal{F}^{-1}\{\mathcal{F}[U(x, y, 0)] \cdot \mathbf{H}(\nu_x, \nu_y, d)\} \quad (11)$$

where \mathcal{F}^{-1} represents the IFT. This is the expression for free space Fresnel propagation.

2.2.2 The Fourier transform property of converging lenses

In practice, if one wishes to move from the far field to the near field of a beam, this can be carried out using a converging lens. Mathematically, this can be simulated through a spatial FT of the electric field and as such is used in the algorithms discussed in the coming sections. This can be understood best, if one first considers what the FT represents and the effect of a lens on a plane wave separately.

If we consider a very basic scenario where there is only a plane wave incident on a thin converging lens at some small angle θ relative to the optical axis, the focus will be located in the focal plane. The focal plane lies a distance away from the lens equivalent to its focal length f in the direction of propagation. The focus in the focal plane will be located an orthogonal distance $\theta \cdot f$ from the optical axis, dependent on the orientation of the incoming light [3]. This is illustrated in Figure 6. For example, if the angle of incidence of a plane wave

on a lens is $(\theta_x, \theta_y, 0)$, where the z axis represents the optical axis which forms an orthogonal basis with the x and y axes, is focused by a converging lens centered at $(0,0,0)$ the focus will be located at the coordinates $(\theta_x f, \theta_y f, f)$. We then consider that for small angles, the angle of an incident plane wave along an axis orthogonal to propagation x can be described with a spatial frequency $\theta_x = \lambda \nu_x$. Here λ represents the wavelength of the light and ν_x represents the spatial frequency decomposition of the light wave when projected onto the x -axis. If we then apply this to a paraxial wave, like a Gaussian beam, we can decompose this into many plane waves at the lens, given that we are indeed dealing with a well-collimated beam with low divergence. For the decomposed plane waves, the above shows us that for the point (x, y, f) , where the focus of a component plane wave incident on a lens at angles θ_x and θ_y is located in the focal plane, we have: $x = \theta_x f = \lambda \nu_x f$ and $y = \theta_y f = \lambda \nu_y f$ respectively. We call the function which describes the complex amplitude of the paraxial wave at some point a distance d before the lens $f(x, y)$. This function has an FT which is proportional to the complex amplitude in the focal plane, the function of which we call $g(x, y)$. When we evaluate the FT of $f(x, y)$ at $\nu_x = \frac{x}{\lambda f}$ and $\nu_y = \frac{y}{\lambda f}$ respectively, we get:

$$g(x, y) \propto F\left(\frac{x}{\lambda f}, \frac{y}{\lambda f}\right).$$

Through use of Fresnel propagation of the individual component waves one gets the complex amplitude [3]:

$$g(x, y) = \frac{i}{\lambda f} \exp[-ik(d + f)] \exp\left(i\pi \frac{(x^2 + y^2)(d - f)}{\lambda f^2}\right) F\left(\frac{x}{\lambda f}, \frac{y}{\lambda f}\right),$$

where $k = \frac{2\pi}{\lambda}$ is the wave number. This gives the following relation between the intensity in the focal plane, which is just $|g(x, y)|^2$, and the FT:

$$I(x, y) = \frac{1}{(\lambda f)^2} \left| F\left(\frac{x}{\lambda f}, \frac{y}{\lambda f}\right) \right|^2 \quad (12)$$

This relationship between the FT and the complex amplitude is highly useful for mathematically simulating the propagation of light.

2.2.3 Sampling for discrete Fourier transforms

When considering a continuous FT as above, one considers all of two-dimensional space in both the source plane and the FT plane, however, for a discrete FT like the ones used for digital image or signal processing this is not feasible. The fidelity with which one can describe a signal or image by discrete sampling is determined by to what degree it is sampled. Under-sampling will generally give rise to artifacts, if not completely distorting the signal beyond understanding, this makes it important to understand how much sampling is appropriate. This is what will be discussed in this section.

To use an example, while an analog image, such as one taken with a film camera, is continuous, a digital image is discrete, that is, sampled. This can be thought of as subdividing the analog image into M by N sections, where each section contains a value associated with that portion of the image, for example a number which represents the color in that area. Each sample

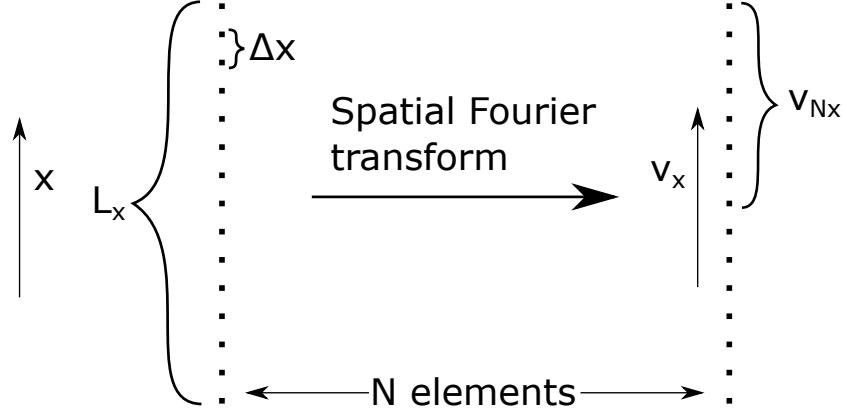


Figure 7: This figure illustrates the relevant quantities relevant to interpolation in a discrete FT from a matrix in the spatial domain to one in the spatial frequency domain, here simplified to one dimension.

makes up a point in an M by N matrix where the distance between each point in the different dimensions can be referred to as Δx and Δy . The inverse of these sampling distances is what is called the sampling frequency, which will come up later. It is possible to deduce the length of this plane in each dimension by simply multiplying the amount of samples in each direction with its associated sampling distance, like $L_x = \Delta x N$ and $L_y = \Delta y M$. In order to include every part of the original image when sampling, the sampled width must be strictly larger than the image one wishes to sample. Of course, simply having a wide sampled image does not necessarily give any information if it has not been sampled enough, meaning that details are lost. Thus analogously, there is also a need to define a sampling frequency which is high enough to include all the information we want. The expression for this is given by the Nyquist-Shannon sampling theorem, which tells us that the sampling frequency must strictly be more than twice as large as the bandwidth of the image we wish to resolve [9]. This yields an interesting result, namely that the absolutely highest frequency which can be described by a given discrete system, also known as the Nyquist frequency, in each dimension is given by:

$$\nu_{Nx} = \frac{1}{2\Delta x}, \quad \nu_{Ny} = \frac{1}{2\Delta y}. \quad (13)$$

These quantities are illustrated in Figure 7 in one dimension. Knowledge of the depicted quantities together with the fact that the matrices in each domain consist of N elements means that it is possible to calculate the spacing in the frequency domain. When this is combined with Equation 12 it becomes possible to calculate the physical size of the Fourier plane when used for optical propagation.

2.3 Phase-retrieval algorithms

In order to be able to correct the wavefront of the laser, one must first know what its wavefront looks like. This would normally be done with some type of wavefront detector, such as a Shack-Hartmann wavefront sensor [10]. Here, we instead use a number of algorithms which have been designed specifically for wavefront reconstruction by means of phase retrieval from amplitude measurements.

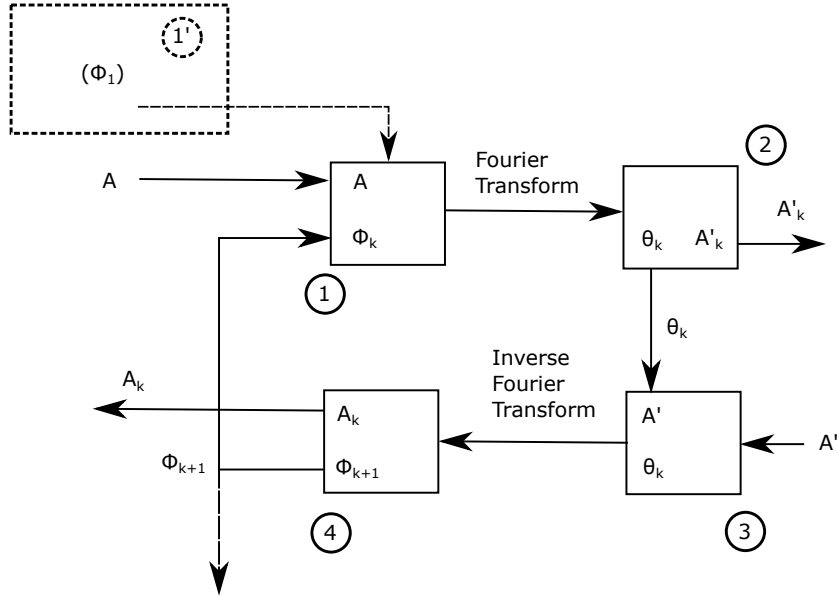


Figure 8: This figure illustrates how an iteration of the GSA is performed, with an iteration number of k . A , A' represent the measured amplitude in the source and target planes respectively, and their counterparts A_k and A'_k refer to the magnitude of the amplitude given by propagation. ϕ_k and θ_k refer to the calculated phase in a given iteration in the source and target planes, respectively, and ϕ_1 refers to the random phase used in the first iteration of the algorithm, indicated by 1'. Step 1 is where the source amplitude is multiplied with the phase from the previous iteration (or ϕ_1) and an FT is performed. Step 2 is the extraction of the amplitude from the transformed system, and step 3 is the multiplication of the phase from step 2 with the measured target plane amplitude and the performance of an IFT on this system. Step 4 finishes the cycle by extracting the amplitude and phase from this step and passing the phase on for the next iteration of the system.

2.3.1 The Gerchberg-Saxton algorithm

The first algorithm employed here, and possibly the oldest, is known as the Gerchberg-Saxton algorithm (GSA) [11]. The GSA is in principle rather simple and requires only two inputs: the image/beam amplitude in its image plane and in a plane to which it is Fourier related, such as its Fourier plane. The algorithm in its first iteration takes the image plane amplitude and applies a random phase across the image and performs an FT on it. The phase of the resulting simulated Fourier plane is then extracted and applied to the measured Fourier plane amplitude and an IFT is performed, transforming it back into the image plane. At this point, the algorithm starts over but uses the phase given by the previous iteration instead of a randomly generated one. This process is repeated, and will decrease the discrepancies between the transformed images and the input images until the error reaches a minimum [11]. The phase can be extracted and used as reference for wavefront correction.

The GSA has the advantage that it can be mathematically proven that the error from one

iteration to the next can never worsen [11], with error here defined as:

$$\frac{\sum_n^{MN} |A_{kn} - A_n|^2}{\sum_n^{MN} |A_n|^2},$$

where M and N are the dimensions of the amplitude matrices, and A_n and A_{kn} are the measured and calculated amplitudes at the n th point in the matrix, respectively. This guarantees that the algorithm will never diverge, but not that it will converge to the correct solution every time [12, 13]. As such, the GSA makes for a good baseline algorithm for phase retrieval, but it is not necessarily the best one available. The algorithm in its k th iteration is described in Figure 8, and is carried out entirely using the below equations:

$$f_k(x, y) = |f(x, y)| \cdot e^{i\phi_k}, \quad (14)$$

where $f_k(x, y)$ represents the current iteration of the complex amplitude of the source plane, $|f(x, y)|$ represents the measured source amplitude and ϕ_k represents the currently calculated phase in the source plane. This is propagated to the target plane via FT, giving the below step:

$$F_k(u, v) = \mathcal{F}[f_k(x, y)] = |F_k(u, v)| \cdot e^{i\theta_k}, \quad (15)$$

where $F_k(u, v)$ is the complex amplitude in the target plane after propagating $f_k(x, y)$ via FT and θ_k is the calculated phase in the target plane. The phase is then extracted and combined with the measured focal plane amplitude, giving the below:

$$F'_k(u, v) = |F(u, v)| \cdot e^{i\theta_k}, \quad (16)$$

where $F'_k(u, v)$ is the adjusted focal amplitude, and $|F(u, v)|$ is the measured focal plane amplitude. An IFT is then used to bring us to the last step:

$$f'_k(x, y) = \mathcal{F}^{-1}[F'_k(u, v)] = |f'_k(x, y)| \cdot e^{i\phi_{k+1}}, \quad (17)$$

where $f'_k(x, y)$ is the adjusted source plane field amplitude and ϕ_{k+1} is the phase to be used as input in the following iteration of the algorithm. In the very first step, ϕ_1 is usually set to be a random phase profile. This is done in order to avoid running into problems which are related to convergence for centrosymmetric amplitude profiles [11].

2.3.2 Input-output algorithms

Input-output algorithm refers not only to a specific algorithm, but also to a family of algorithms, all of which are dedicated to phase retrieval. The three main algorithms will be discussed here, these are: the titular Input-output (I/O), the Output-output (O/O) and the Hybrid input-output algorithm (HIOA) [14] which is the one used in this work. The key difference between these algorithms and the GSA lies not only in what is done in the fourth step of the algorithm, which is the main difference, but also in how they make use of the three initial steps. That is, the FT of the image plane, the phase extraction in the Fourier plane and the IFT of the adjusted Fourier plane. The GSA functions like a closed loop until the desired phases in either plane can be extracted at the end. The input-output algorithms instead form what could be called open algorithms with a first and a last step. These

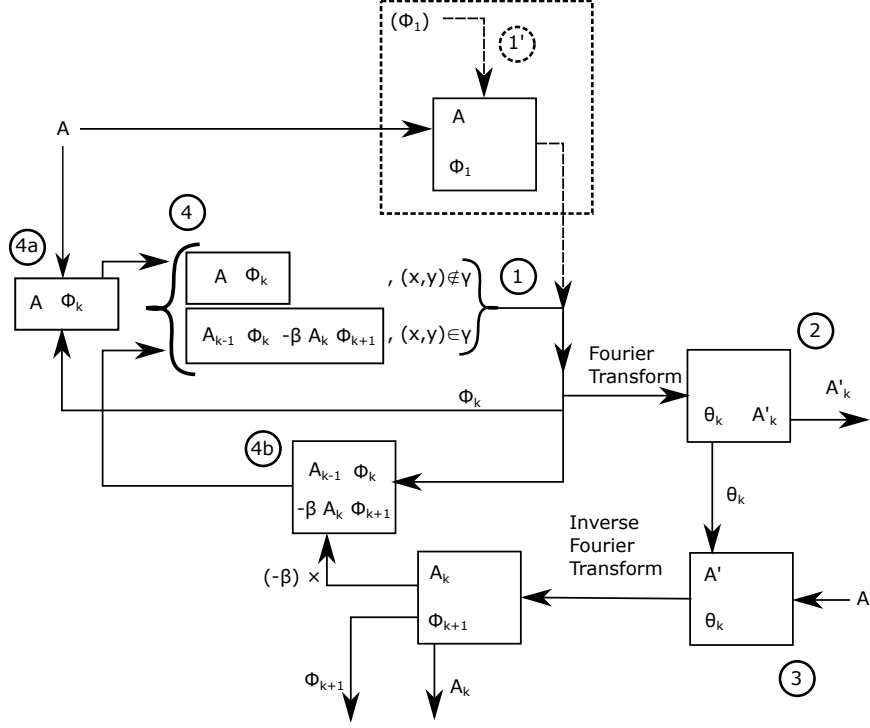


Figure 9: This figure illustrates an iteration of the I/O algorithm. It is carried out similarly to an iteration of the GSA, with exception of the last step, which uses values from several inputs and previous points in the iteration, which makes it appear far more complicated, but it only makes use of the same quantities as the GSA with the exception of the constant β and the set γ . Steps 1-3 are carried out more or less identically to the ones in the GSA as described in Figure 8. Step 4 has here been divided into two steps, 4a, and 4b. 4a represents the step where the measured source amplitude is combined with the phase of the previous step to form the amplitude in places where the measured amplitude is not zero. 4b represents a step similar to step 4 in the GSA, with the exception that it subtracts the result of the IFT from the input of the previous iteration and sets this as the value of the amplitude where the source amplitude is zero.

algorithms use the outputs to varying degrees in order to drive the inputs of further iterations.

The I/O utilizes the same first steps as the GSA. But when the next iteration is started, rather than simply applying the phase from the IFT to the measured amplitude it instead uses the input for the previous iteration as well. This new input is subtracted with the output of the previous iteration, with some prefactor β determined by circumstances of use, but usually set to < 1 [14], where the original input amplitude is equal to zero. The use of the previous iteration's input in combination with the output is what gives this algorithm its name. This is carried out by the algorithm in replacing the fourth step with the one below:

$$f_{k+1}(x, y) = \begin{cases} |f(x, y)| \cdot e^{i\phi_k}, & (x, y) \notin \gamma, \\ f_k(x, y) - \beta f'_k(x, y), & (x, y) \in \gamma \end{cases}$$

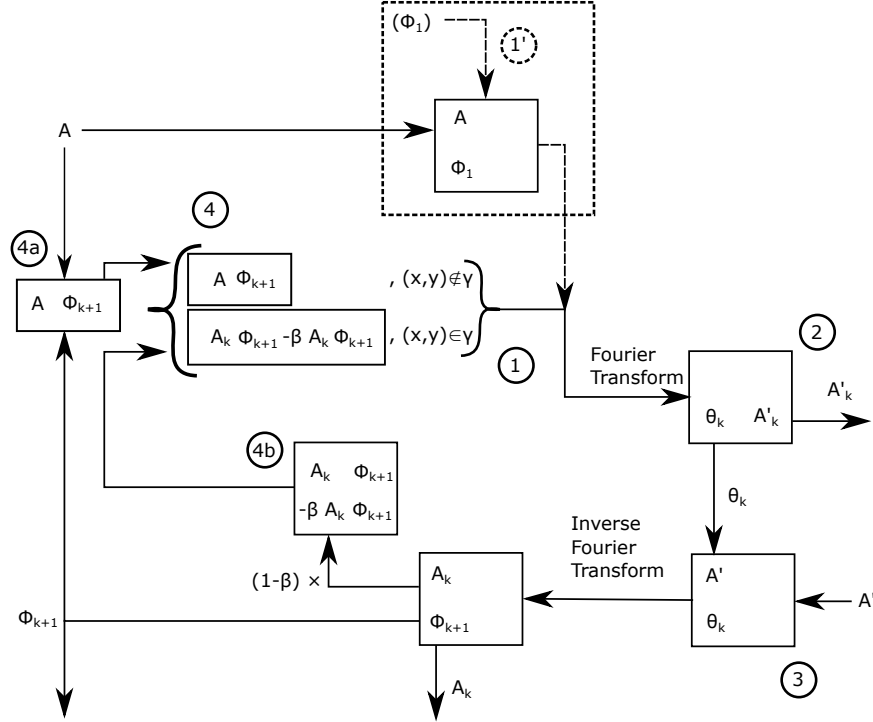


Figure 10: Illustrated here is a diagram of the O/O algorithm. It is structured somewhat similarly to the I/O as described in Figure 9 with a split in step 4, except it does not make use of values from previous iterations. Step 4a is identical to step 4 from the GSA, but only applies to a subset of the source plane, where 4b returns a diminished amplitude of the IFT output in places where the amplitude would simply be zero in the GSA.

Where β is a scaling factor chosen to be between 0.5 and 1 [17], and where γ refers to the set of points in the source plane where the measured field amplitude is zero. What this changes in each iteration of the algorithm is illustrated in Figure 9.

The O/O is similar in structure to the I/O with the exception that one replaces every instance of the previous iteration's input with its output. This means that if one chooses the prefactor $\beta = 1$ one gets an algorithm which is functionally identical to the GSA [14]. The algorithm is illustrated in Figure 10, and the below equation replaces step four in the GSA:

$$f_{k+1}(x, y) = \begin{cases} |f(x, y)| \cdot e^{i\phi_{k+1}}, & (x, y) \notin \gamma, \\ f'_k(x, y) - \beta f'_k(x, y), & (x, y) \in \gamma \end{cases}$$

The above algorithms each have their advantages and disadvantages relative to one another. For example, similarly to the GSA, the O/O may stagnate without reaching a solution but may otherwise be faster than the I/O.

The HIOA is an algorithm which synthesizes the two methods of I/O and O/O into one and in doing so, attempts to transcend their weaknesses. What this means in practice, is that the HIOA uses as its input for a new iteration the phase from the output multiplied with the

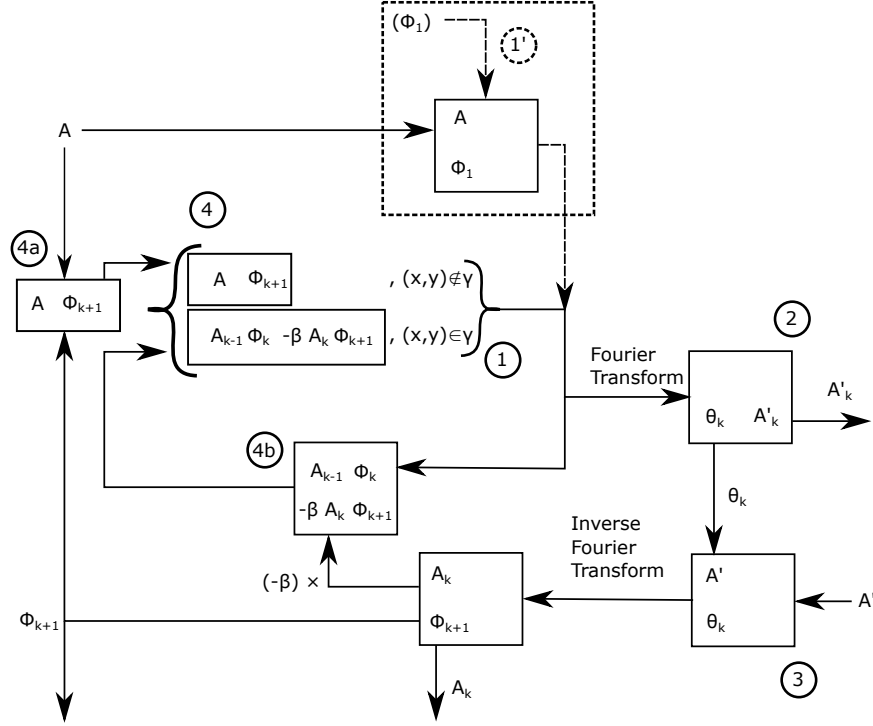


Figure 11: Illustrated here is a flow diagram of the HIOA. Its step 4a is identical to that of the O/O algorithm, and its step 4b is identical to that of the I/O algorithm, but is otherwise identical in structure to either of them.

original input amplitude, as in the GSA and O/O, but where the original amplitude is zero, it uses the previous iteration's input subtracted with the output multiplied with the factor β as before. The point of this, in contrast to previous algorithms, is to introduce a sign flip between different iterations in areas where errors are more persistent, so as to avoid getting stuck in local error minima with persistent errors. The HIOA is the most efficient phase retrieval algorithm introduced so far [14] but it has its share of issues as well. It is hard to judge the accuracy of the HIOA in the same way as the GSA, because while it can be good at retrieving the wavefront, due to its nature as an open algorithm, it is more concerned with generating outputs which will yield a correct phase in the end, rather than gradually minimizing the error. This means that the error, as defined for the GSA, may not be as good of an indicator of the accuracy of the wavefront here.

The HIOA is identical to the GSA with the exception of step 4, as illustrated in Figure 11, which for the HIOA is replaced by the below, which can be seen to be a mixture of both previous algorithms:

$$f_{k+1}(x, y) = \begin{cases} |f(x, y)| \cdot e^{i\phi_{k+1}}, & (x, y) \notin \gamma, \\ f_k(x, y) - \beta f'_k(x, y), & (x, y) \in \gamma \end{cases} \quad (18)$$

2.3.3 Three-plane intensity algorithm

The three-plane intensity measurement algorithm is based on the GSA, however, while the GSA relies on propagating back and forth between two planes through use of the FT, the three-plane algorithm uses intensity measurements in three planes and propagates between them. This propagation can be done in any way which is appropriate for the distance propagated. It is here done with a Fresnel transfer function propagator as described in Section 2.2.1 above. The algorithm starts in a plane of choice and propagates to a second plane, where the phase is extracted and multiplied with the measured plane amplitude, similarly to the GSA, this is done again, so that the third and last plane is reached. The third and last plane is propagated all the way back to the first position, without stopping at the second position between them, and upon this the cycle starts over. This modified cycle is illustrated in Figure 12. Despite its simplicity, effectively only being a single step added to the GSA, this algorithm can be seen to be more effective in some circumstances than the regular GSA insofar as it has been shown to reproduce more complicated wavefronts with smaller errors without stagnation, which is seemingly caused by the fact that it requires more information as input [15]. Propagation is done within the Fresnel approximation, as has been assumed in Section 2.2.1. This gives the possibility of using Equation 11 for purposes of propagation, which we do, with the exception that the definition for $\mathbf{H}(\nu_x, \nu_y, d)$ is not given by Equation 10, and is instead given by:

$$\mathbf{H}(\nu_x, \nu_y, z) = e^{-i\lambda z\pi(\nu_x^2 + \nu_y^2)}$$

Which is the Fresnel approximation of the free-space transfer function with the first term removed, as it only contributes a flat phase in the direction of propagation, and can therefore be neglected here [9].

2.3.4 Multi-image Gerchberg-Saxton algorithm

An algorithm in a similar vein as the previous one, the Multiple-measurement averaging algorithm (MMAA) carries out propagation from the source plane to several target planes in parallel [16]. Each of these planes function similarly to the GSA in that the algorithm simply propagates back and forth. This time, however, when the planes are propagated back to the source, the phase and amplitude is averaged in the source plane. This new source is now propagated in parallel to each target again, this is illustrated in Figure 13. This algorithm has the drawback of running more calculations in the background, making it more costly. What makes it so appealing, though, is how robust it is to camera noise for applications where this is an issue. Functionally, this is nearly identical to the three-plane intensity measurements, only that propagation is carried out several times in parallel with propagation to slightly different positions.

2.3.5 Spatial phase perturbation algorithm

The next algorithm to be discussed is the Spatial phase perturbation algorithm (SPPA). The SPPA is fundamentally nearly identical to the GSA with one exception. This exception is the "phase perturbation" which gives the algorithm its name. What this means in practice is that at the end of each iteration of what normally would be the GSA, a random phase profile, multiplied by some factor $0 < \alpha < 1$ is added to the phase. There should also be some factor

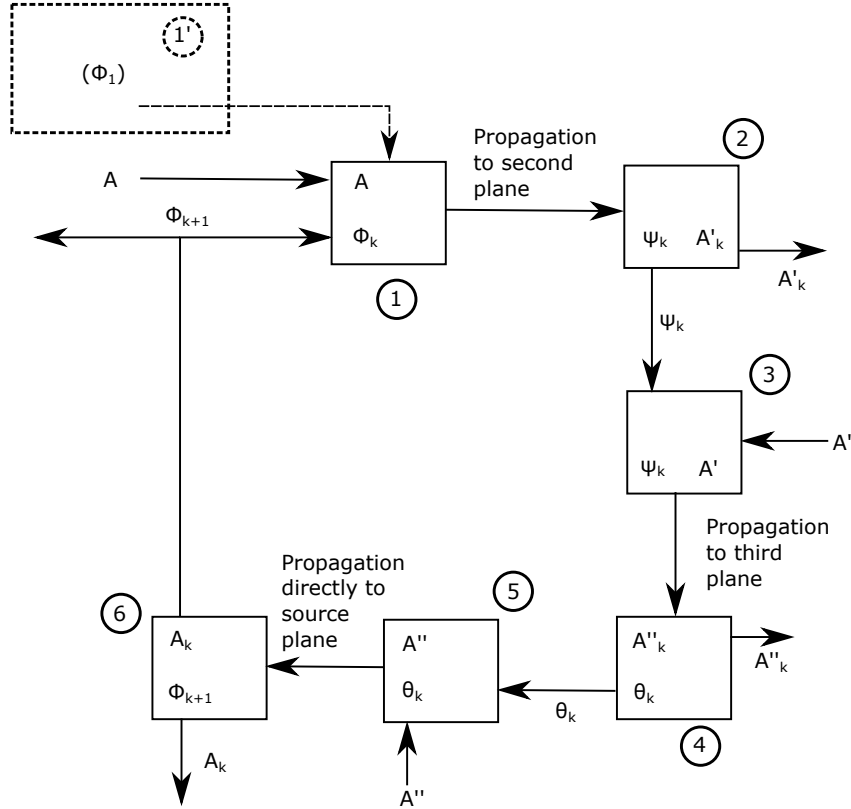


Figure 12: Shown here is a diagram of the three-plane intensity algorithm. It is structured similarly to the GSA described in Figure 8 but with more steps, and instead of simply using the FT directly, propagation between several planes is carried out. Step 1 consists of propagating the input phase and amplitude to a second plane, and step 2 consists of extracting the phase from that plane, to be combined with its measured amplitude in step 3. This procedure is then repeated for the second to third plane propagation and from the third to first plane in steps 3-6.

dependent on the iteration number such that this random phase steadily decreases to zero over several iterations, leaving the last iterations to be calculated like the normal GSA. In effect, this small random phase addition have the effect of "scrambling" the FT of the image in each iteration to a decreasing degree. This seems counterintuitive at first, but in most cases this causes an eventual convergence to a smaller error than the regular GSA [17]. The reason for this lies in the fact that the error of the GSA cannot increase between iterations, as mentioned above. Consequentially, the GSA is at risk of getting stuck in a local error minimum from which it cannot get out. This behaviour of reaching a local minima and staying there is usually what is meant when the GSA is referred to as having stagnated.

The point of the SPPA is thus to run the regular GSA with a mechanism which makes it possible for the algorithm to "shake itself loose" from local minima and other possible causes of stagnation. The perturbation must, however, gradually decrease to zero at some point if the algorithm is to be able to converge at all. This algorithm may take many iterations before it converges as it needs a long, gradual decrease of the perturbation for best results.

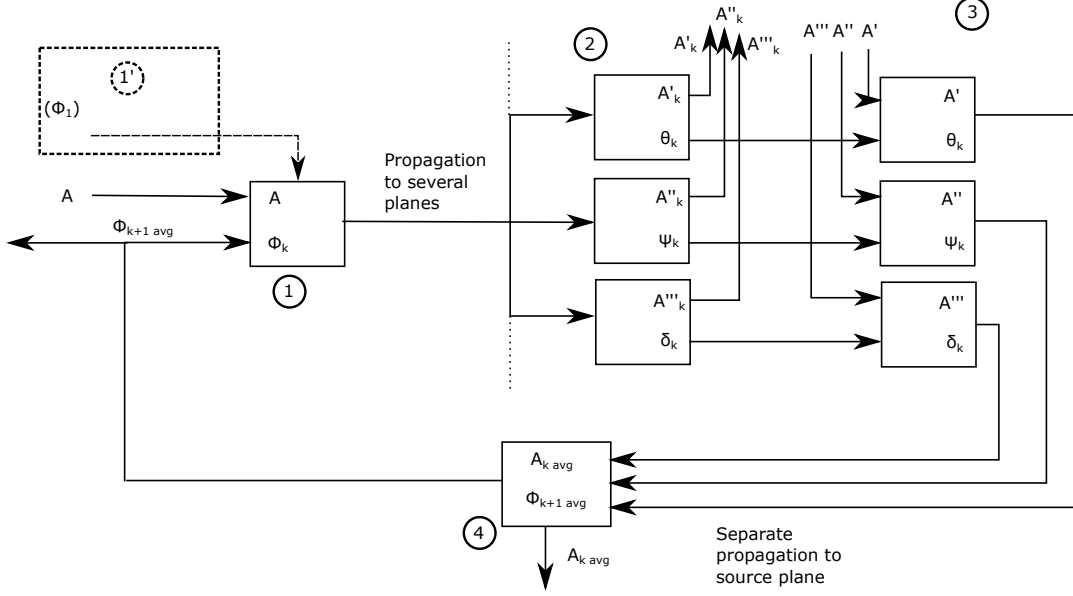


Figure 13: This figure shows a diagram of the MMA. Step 1 of this algorithm is similar to the GSA, except that parallel propagation from the source plane to a given amount of planes in space is performed instead of an FT. Step 2 consists of the extraction of the phase from all these separate planes and multiplying them with their respective measured amplitudes. Step 3 then consists of propagation of all separate planes back to the source plane. Step 4 then consists of taking the average phase from this propagation and combining it with the measured source amplitude.

The algorithm has, however, been shown to converge to errors much lower than that of either the GSA or the HIOA on their own, and is more robust to random noise in the amplitude measurements [17]. This algorithm can be used on its own, but the spatial perturbation component can be implemented as a component in most other algorithms. It is important to note, however, that this may not be suitable or helpful for all algorithms. The HIOA for example, has its own mechanism for avoiding stagnation and is thus not likely to benefit from this spatial perturbation and it may in fact be harmful to convergence.

The SPPA is similar to the GSA except for the last few steps, and for one step, the difference is purely notational. The exact differences may be noted by comparing the below equations to their equivalents equations 16 and 17.

$$f'_k(x, y) = \mathcal{F}^{-1}[F'_k(u, v)] = |f'_k(x, y)| \cdot e^{i\phi'_k}$$

Where ϕ'_k is used in place of ϕ_k from equation 16, and the last step is replaced by:

$$f_{k+1}(x, y) = |f(x, y)| \cdot e^{i\phi_{k+1}} = |f(x, y)| \cdot e^{i(\phi'_k + \Delta\phi_k)}$$

Where $\Delta\phi_k$ is described by:

$$\Delta\phi_k = \alpha \cdot \left(1 - \frac{k}{K}\right) \cdot (\text{random phase from } 0 \text{ to } 2\pi)$$

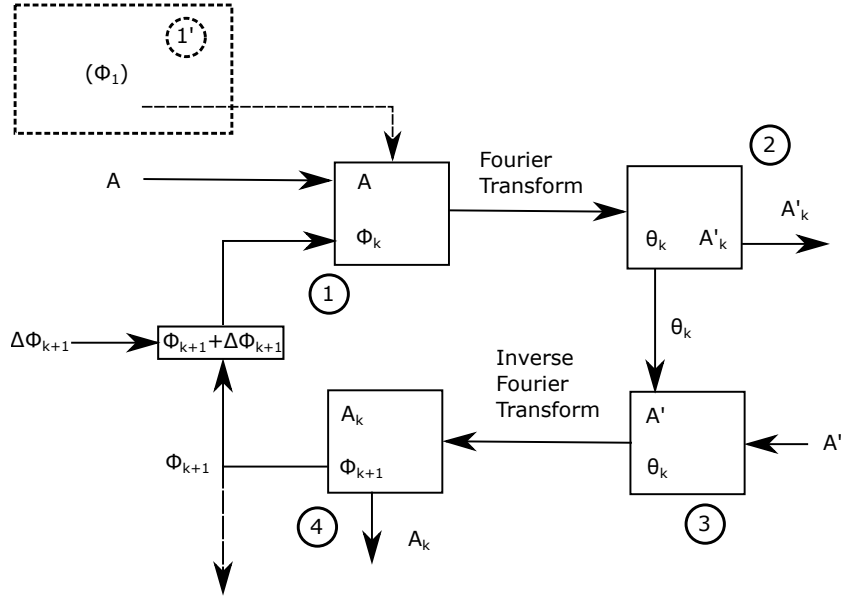


Figure 14: This diagram illustrates the SPPA, which is identical to the GSA, as described in Figure 8, with the exception of a random, decreasing, phase which is added to the outgoing phase in step 4.

Where α is a scaling factor, and K is the amount of iterations for which the perturbation is used in the algorithm.

This algorithm has been shown to have a better convergence than either the HIOA or the GSA [17], but it has the drawback of needing to run through at least K iterations before one can rely on the phase or amplitude output. Depending on what the appropriate magnitude K has, it may take very long for the algorithm to converge relative to other methods.

2.3.6 GS/HIOA

The Hybrid Gerchberg-Saxton and hybrid input-output algorithm (GS/HIOA) is, as the name implies, a combination of the GSA and the HIOA. What this means in practice is that the algorithm is run in cycles of some amount n of iterations. The first out of every n iterations in the algorithm is run as a GSA iteration and the remaining $n - 1$ iterations in the cycle is run as iterations of HIOA. It could hypothetically be run with a different balance of GSA to HIOA iterations, but literature seems to indicate that one GSA iteration per cycle is optimal [17]. The cycle size appears to have an effect on the error size, with a cycle with too few iterations often not converging as well as possible, whereas extending the cycle too much only causes it to converge more slowly [17]. It thus appears preferable to start with a larger cycle size, as opposed to a smaller one. This algorithm seems, according to literature, to converge to a smaller error more quickly than any other algorithm tested, at the expense of being slightly less robust to noise than the SPPA above [17]. This algorithm has the drawback that it can only be run in some integer number of cycles, which may make it difficult to estimate how many cycles should be used if one does not monitor the error in real

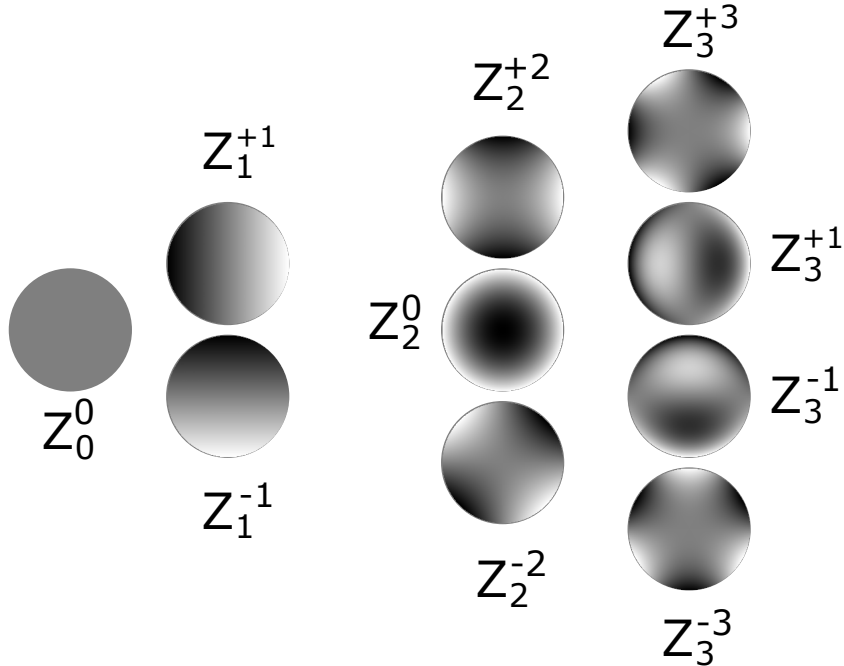


Figure 15: Illustrated here are the lowest orders of Zernike polynomials. The lowest order, Z_0^0 corresponds to a flat surface, orders $Z_1^{\pm 1}$ correspond to tilted wavefronts, orders $Z_2^{\pm 2}$ represent astigmatism and Z_2^0 corresponds to the wavefront of a defocused beam. The orders $Z_3^{\pm 1}$ corresponds to optical coma and $Z_3^{\pm 3}$ is trefoil.

time. The algorithm may also take a few cycles before the calculated error actually decreases below that calculated by either the HIOA or the GSA for the same amount of iterations. Ultimately though, this seems to be one of the stronger options for phase retrieval, as it seems faster than both SPP and MMAA due to its apparently faster convergence [17], and a much shorter computation time of each iteration, respectively.

2.3.7 Zernike Polynomials

Zernike polynomials are a specific set of polynomials which form an orthogonal basis over a circular area, meaning they can be used to take on any value over a given circular area, similarly to how the Cartesian coordinates (x, y and z) can be used in regular three-dimensional space. They are of special interest in optics, as they are defined in such a way that individual polynomials describe different types of optical aberrations [18]. Some of the low-order Zernike polynomials are illustrated in Figure 15. A consequence of this is that any wavefront which is calculated over a circular area, such as a Gaussian beam, can be decomposed into Zernike polynomials. It was used in conjunction with one of the above algorithms in the following way: a full phase reconstruction is done via an algorithm. The resultant phase is afterwards decomposed over the radius of the beam into Zernike polynomials. This decomposition may then be used to approximately reconstruct the wavefront for the SLM. The benefit of doing this is that the Zernike-approximated wavefront does not have odd discontinuities in the phase front which may cause speckles or other odd behavior when the beam is focused. The

fact that the polynomials are orthogonal means that one could hypothetically use only a few Zernike polynomials instead of the full wavefront in order to partially compensate for optical aberrations.

The advantage of the Zernike polynomials, as mentioned before, is that they correspond to well-known aberrations of an optical system, such as astigmatism, spherical aberration and comas. This means that when one uses them to compensate for a reconstructed wavefront, one can potentially identify different types of aberrations and use this knowledge to examine the setup. Using the knowledge of what type of error occurs, one can use this to correct the experimental setup itself, if possible.

2.4 Phase modulation with spatial light modulators

As stated in the introduction, this project will aim to perform a divergence minimization of attosecond pulses through modulation of the spatial phase of the driving laser. This phase modulation will allow for the shaping of the beam into an intensity profile more suited for generation. The driving field which is used for HHG commonly has a Gaussian intensity profile. If one also takes into account that the beam intensity affects the phase of the generated harmonics, as discussed in Section 2.1.1, using a beam with a Gaussian intensity distribution will give rise to an increased curvature in the wavefront of these harmonics. This could theoretically be ameliorated by instead using a so-called "flat-top" beam which ideally has a constant intensity across the entire beam profile. With an equal spatial intensity distribution, the divergence across the generating volume should be minimized. Furthermore, the focal positions (virtual or real) of consecutive harmonics would be closer together.

Being able to use a flat-top beam profile would require that the original Gaussian beam is reshaped. This reshaping could be done by a phase modulation of the beam wavefront by use of an SLM. An SLM is a device whose main purpose it is to manipulate the phase of incoming light based on its spatial position. An SLM often has a system of "pixels", similar to those on a computer screen, as its active surface. The kind of SLM used here is a Liquid crystal on silicon (LCoS) SLM and is described in Figure 16. The refractive index change along the extraordinary axis of the liquid crystals allows the operator to fine-tune the optical path distance through the cells, which allows fairly precise control of the phase of incoming light. By controlling the SLM via a computer, one can then shape the wavefront of the reflected beam. This will upon sufficient propagation of the beam, such as after focusing it with a lens, cause it to take on an intensity profile which is associated with such a phase. There is, however, an omitted step from the above procedure, namely the requirement of a beam with a relatively flat phase being used as input. Otherwise, the applied phase may fail to have the desired effect. Thus, in addition to using an SLM to shape the beam, it will also be used here to correct for aberrations in the beam wavefront.

The SLM itself effectively interacts with the computer as a screen which can display greyscale images in 8-bit, that is pixels with color values from 0 to 255 which corresponds to a scaling level from black to white. The greyscale value which corresponds to a phase shift of 2π depends on the wavelength used. The phase imparted by the SLM scales linearly starting from 0, and knowing this makes it possible to convert from phase to a greyscale image. When an image is projected onto the SLM "screen" from the computer, it uses the value of each

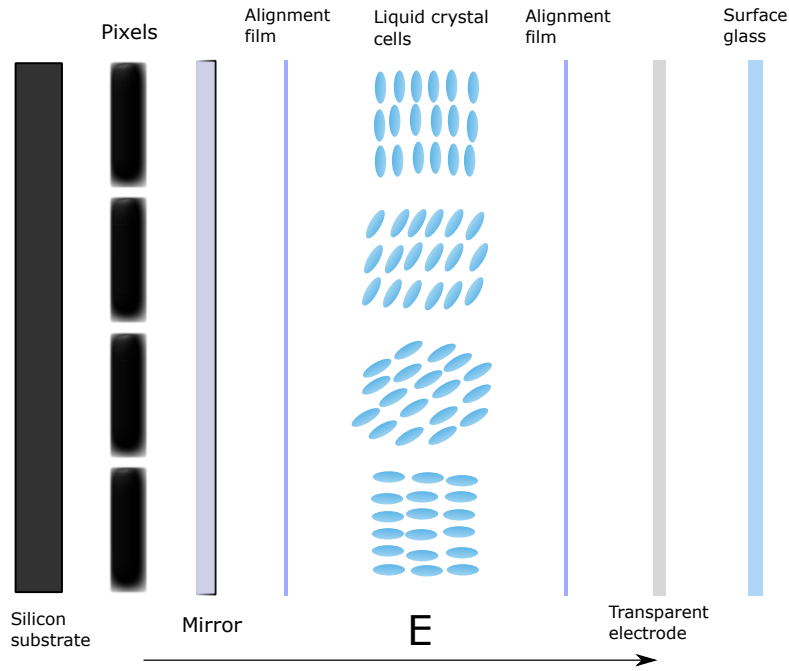


Figure 16: This figure illustrates a cross-section of the surface of an LCoS SLM. The layers displayed from left to right corresponding to their position from bottom to surface in the active aperture of the SLM. The pictured small cells of nematically orientated and uniaxially birefringent liquid crystals are positioned on a silicon substrate and arranged such that each cell is on top of one of the pixels shown to the left of the mirror. Phase change is induced by changing the voltage between the pixels and the transparent electrode which causes the molecules to rotate and the refractive index in the cell to change. The direction of the applied electric field is described by the arrow labeled E . The change in refractive index itself has to do with a gradual rotation of the individual molecules in the liquid crystal cells as illustrated from top to bottom in the figure. This change in each of the "pixels" effectively amounts to adding a certain amount of phase to the reflected light. For the refractive index change to be fully effective, the incident light must be linearly polarized such that the light encounters the full extraordinary refractive index of the liquid crystals. This is necessary in order to gain the largest possible modulation depth out of the SLM. The purpose of the mirror is to ensure that as much light as possible is reflected back out of the SLM, such that the phase manipulation has any effect. The alignment film is located on each side of the liquid crystal cells to help keep them in place.

pixel to determine the voltage it applies over each liquid crystal cell. This in turn decides the phase it imparts on the incoming light. The phase profile which is calculated must be interpolated such that it fits onto the surface of the SLM and translated into a greyscale picture format, it is then projected onto the SLM screen as an image.

There are numerous things which must be taken into account when working with an SLM, many of which are related to its surface structure, which is illustrated in Figure 16. One of these is the fact that not all of the light incident on the active surface of the SLM will reflect as intended. Some of this is due to loss of power, which happens mostly due to absorption in the transparent electrodes and scattering off the liquid crystals [19]. For the SLM used here,

which is a Hamamatsu X13139-02, the light utilization efficiency is 97 % [20], meaning that 3% of the incoming power is lost when the SLM is used. There is also light which is reflected off the SLM without reaching the liquid crystal cells, which will from now be referred to as 'zeroth order' light. The amount of power which goes into the zeroth order varies from device to device generally taking up between 5-20 % of the incoming power [21]. The zeroth order can be separated from the properly refracted light by applying a blazed grating to the phase profile projected onto the SLM as this will tilt the wavefront and cause them to diverge from one another.

Another type of loss which must be taken into account are the higher orders which are diffracted due to the discrete structure of the SLM surface, that is, the grid formed by the crystal cells and pixels illustrated in Figure 16. These higher orders go out in nearly all directions from the SLM surface and may need to be taken into account. These, however, often do not take up as much power as the zeroth order and are in practice separated by themselves due to their large angles of divergence relative to the main beam.

The SLM has an additional limitation in that it cannot be used for beam intensities which are too high. This has to do with the temperature tolerance of the liquid crystal cells, as an incoming beam with too high power will cause the liquid crystal structure to warm up, which would at best decrease efficiency temporarily, and at worst permanently damage the cells. Hence, one must be careful to choose appropriate beam size, average power and peak power when using the SLM.

3 Implementation

The calculation of the phase profile, its interpolation and its conversion into an image is all done with a program that was developed in the programming language LabVIEW. The algorithm used below for phase retrieval was the GS/HIOA, but other algorithms were considered and tried before this one was decided upon. All image processing required to get the camera images suitable for FT in the algorithms, that is zero-padding, amplitude interpolation and amplitude normalization, is also carried out in LabVIEW. Additionally, MATLAB has been used as a benchmark comparison for individual sections of the code in order to verify that they function correctly and for performing certain tasks within the program. These tasks include interpolation, two-dimensional phase unwrapping using a function from the MATLAB central file exchange [22, 23] and a Zernike polynomial decomposition function, also from MATLAB central file exchange [24].

3.1 Experimental Setup 1

The first stage of the experiment was set up on an optical table as illustrated in Figure 17¹. The purpose of this stage of the experiment was mainly to test the SLM with a beam which was completely safe for the SLM. All of the potential functions of the phase-retrieval program were tested here in multiple iterations, including all of the above described phase-retrieval

¹All experimental setup figures were in large part made with use of the image library ComponentLibrary for optical components by Alexander Franzen [25].

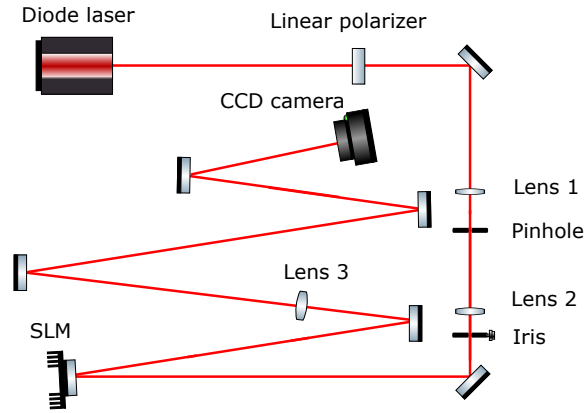


Figure 17: The optical setup as used with the diode laser. The diode has a central wavelength of approximately 800 nm. The beam is first linearly polarized so as to reach maximum effectiveness with the SLM. It is then passed through an adjustable telescope composed of two lenses, lens 1 and lens 2, in addition to a pinhole. Lens 1 has a focal length of 5 cm and lens 2 has a focal length of 10 cm, and the pinhole diameter is 0.02 mm, these components are all positioned in adjustable mounts in an opto-mechanical cage system. All three telescope components can be adjusted along the optical axis, the pinhole can also be moved in its mount transversally. The telescope is set up in order to clean up and magnify the beam by removing higher spatial frequencies, to ensure that the beam is somewhat circular. This is done in order to make it similar to the beam later used in the actual experiment. The spatial filtering by the pinhole gives rise to an Airy pattern, which is cut at the first minimum with the iris placed after the telescope. After this, the light is reflected off the SLM and then passes through lens 3, which has a focal length of 1 m. A Charge-coupled device (CCD) camera is placed in the focus to record the amplitude of the beam in the Fourier plane. Both the SLM and the camera are connected to a computer which controls inputs to both devices and stores the camera recordings.

algorithms. The possibility of beam shaping was also tested here, with phase vortices and holograms being applied to the SLM for this purpose. While the focus of this setup was on testing the SLM, it was also used as a way of getting used to the other equipment, such as the CCD camera and its settings, which would also be instrumental in the final experiment. In essence, this part of the experimental work was a staging ground for further uses of the SLM and the related setup. Most, if not all of the iterative work on programs and algorithms were performed at this stage, with the remaining steps mostly consisting of application of methods developed here.

A picture was taken of the beam amplitude at the SLM and the lens (the image plane) for use in calculations. For the multi-image and multiple-intensity measurement algorithms, multiple pictures were taken around the focus as well. The algorithm was tested until results

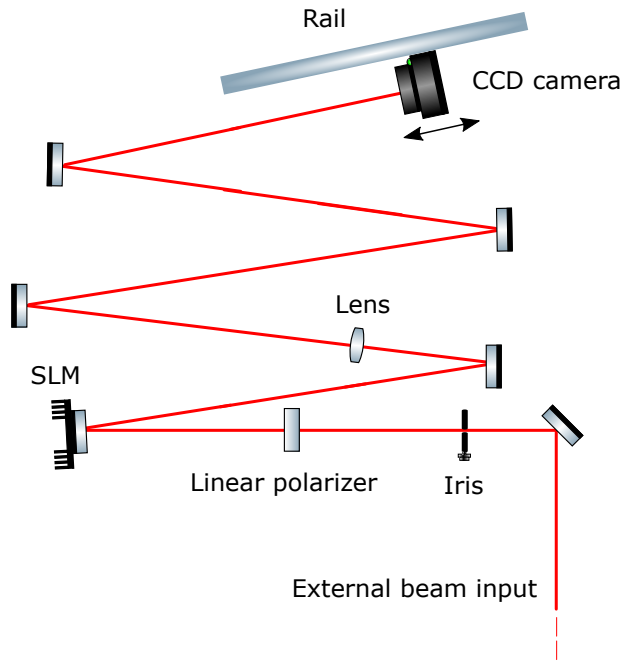


Figure 18: The optical setup for the second set of experiments. Now with an external beam from the ultrashort pulse laser coupled in rather than a diode laser, with an iris for partially blocking the beam and a linear polarizer before the SLM to ensure that maximal efficiency is reached. The lens used here has a focal length of 1 m. The only component which is entirely new to this setup is the rail, which had a ruler mounted on top of it for measurements, along with a movable base for the camera, to allow for free movement of the camera near the beam focus.

from the wavefront correction program were deemed good enough in order to move one to the ultrashort pulse laser and the HHG beamline.

3.2 Experimental Setup 2

The second stage of the experiment still used the same simple arrangement with some changes, as illustrated in Figure 18. The purpose of this stage was primarily to measure the results of wavefront correction applied to a beam from the ultrashort pulse laser. The laser used here and in the following laboratory setup is a Titanium:Sapphire laser with a central wavelength of 800 nm, emitting pulses at approximately 20 fs with an up to 5 mJ pulse energy at a repetition rate of 1 kHz. A weak copy of the main beam was used in these experiments. The telescope and spatial filter were therefore no longer necessary in the setup. Since the beam was larger than the camera CCD chip, in order to take a picture of the far field, a camera objective was used and a scale was made for the far field picture with millimeter-spaced graph paper. In the near field, the camera was placed in the beam path on a base which could be moved along a rail in order to easily take images in the proximity of the geometrical focus.

Care was taken to minimize optical aberrations caused by the setup. The SLM itself is a source of aberrations in the system due to its surface not being perfectly optically flat.

The camera was then positioned in what appeared to be the best focus of the beam without any additional phase on the SLM. The algorithm was run based on an input from this position and afterwards images were taken near the focus. This was done for several initial positions in the vicinity of the best focus to verify whether the peak intensity did increase as a result of the algorithm being used or not.

3.3 Experimental Setup 3

3.3.1 Harmonic generation, pre-generation chamber

In the third stage of the experiment, we wanted to investigate the effects of using the SLM for the process of HHG. This means that it is of interest to be able to observe HHG both with and without the SLM, and also to use the SLM to shape the driving beam in order to observe the effect this may have on the generated harmonics. We are primarily looking for the impact on the divergence of the harmonics by wavefront shaping.

The experiment is illustrated in Figure 19. An iris further upstream was used to put improve on the incoming beam by partially closing it. Irises 1 and 2 were then used to align the incoming beam into the generation chamber. The dual angled mirrors in the SLM in-coupling area were placed on a magnetic base, which made them simple to remove and replace. This allowed for easy iteration between working with and without SLM in the beam.

3.3.2 Harmonic generation, generation chamber

After the driving beam passes through the above system, it enters the generation chamber, as illustrated in Figure 20, where harmonics were observed via the XUV spectrometer. The pick-up mirror was used to align the setup each time the SLM was coupled in via the twin mirrors on the magnetic base, as seen in Figure 19, were inserted to ensure that the beam path was the same as before.

3.3.3 Harmonic generation, experiment

A pulse energy of 0.5 mJ, corresponding to an average power of 0.5 W, was considered safe use without damaging the SLM. After harmonics started being generated and observed via an XUV spectrometer, the aperture was adjusted to optimize for the cutoff and spatial shape of the harmonic spectrum. A spectrometer measurement of HHG which was used in the results is illustrated in Figure 21. A spectrometer measurement was taken, along with an image of the beam profile of the driving field. The aperture was then fully opened and another image of the focus was taken. In the following, the SLM was coupled in and the power was measured after the SLM. The focus was recorded without the aperture. An optical plane vortex applied on the SLM was used for alignment purposes and in order to center the beam on the SLM. HHG was carried out with the SLM without shaping the beam.

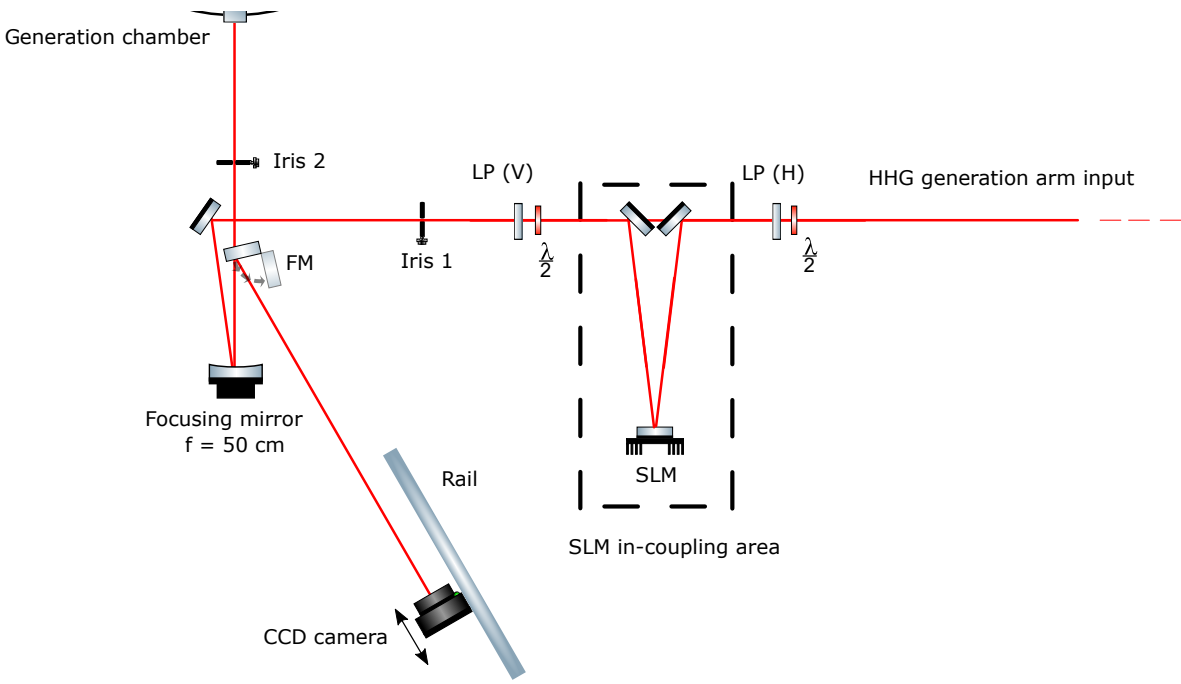


Figure 19: An illustration of the setup for HHG outside the generation chamber. The beam input used here was the direct generation arm input of the external setup using the Titanium:Sapphire ultrashort pulse laser. The start of the setup consists of an adjustable half-wave plate ($\frac{\lambda}{2}$) linear polarizer (LP (H)), to set the polarization of the incoming beam to be linear and horizontally oriented for the SLM, the half-wave plate can also be used to adjust the incoming beam power. The area surrounded by a dashed line marked "SLM in-coupling area" marks where the SLM was coupled in via two mirrors mounted on a magnetic base. This meant that they could be easily removed to couple the SLM out of the system. Immediately after the SLM area, another half-wave plate and linear polarizer (LP (V)) were placed to ensure maximal conversion efficiency from horizontal to vertical polarization. The HHG chamber requires vertically polarized light for generation, so these were adjusted to simply rotate horizontally polarized light from the SLM into vertically polarized light without any loss of power. Additionally, an iris (iris 1) was positioned before the focusing mirror which focused the incoming beam for harmonic generation in the generation chamber. The beam would then pass through the second iris (iris 2) before entering the generation chamber. The focal length of the focusing mirror used was 50 cm. There was also a flip mirror (FM) set up between the second iris and the focusing mirror, to make it possible to view the focused beam outside of the generation chamber. The camera and rail was set up near the geometrical focus.

The next step was to observe the possible impact of the SLM on the divergence of HHG. For this purpose, three scenarios with similar peak intensities were implemented. These different scenarios involved using similar peak intensities, but different beam profiles, which would be achieved in various ways described below. It was necessary that the beams in the scenarios had similar peak intensities, as critical aspects of HHG, for example the observed cutoff, depend strongly on the intensity. In practice, this was done by comparing the harmonic

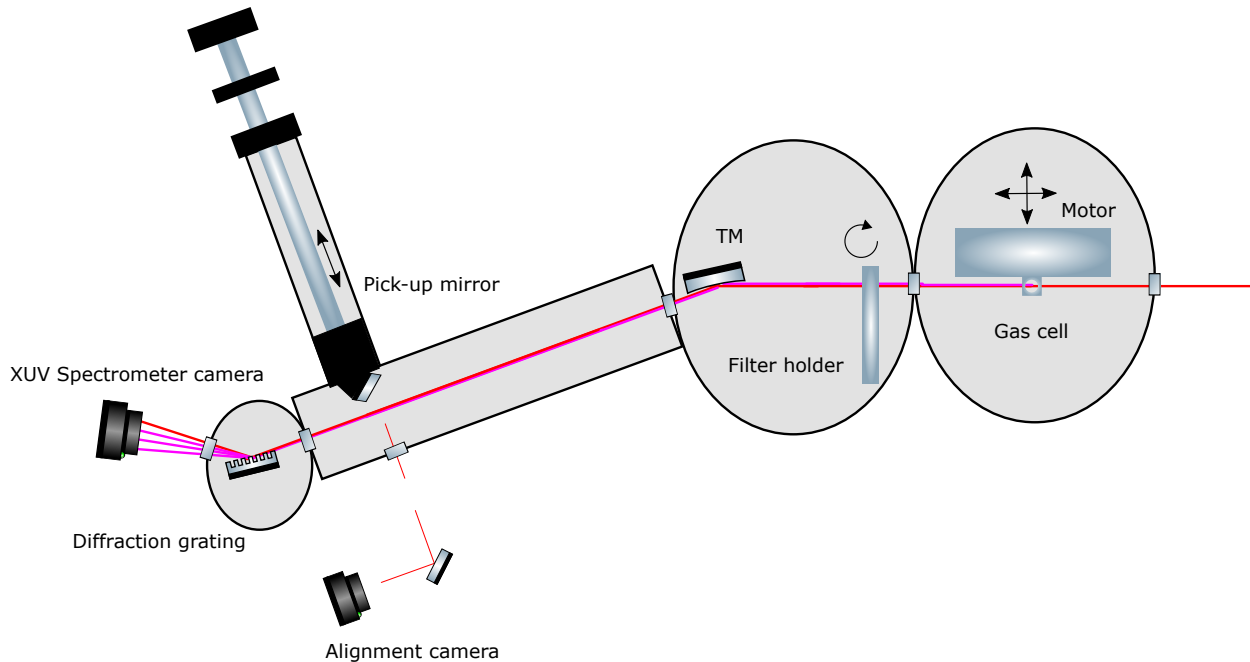


Figure 20: A simplified illustration of the HHG setup inside the generation chamber and the following chambers. This figure functions as a continuation of Figure 19. The first components in the system are the gas cell, where HHG takes place, and the motor it is attached to. This motor allows for the position of the gas cell position to be adjusted precisely in three dimensions. In the next chamber, a "rotatable" filter holder is positioned which holds several filters of different materials in different slots which can be rotated between, with some slots unfilled, leaving empty holes allowing for free passage of light if desired. The light which passes through the empty slot or filter is then reflected off a toroidal mirror (TM) toward a diffraction grating which is at a grazing angle relative to the incoming light. At the end of the system is an XUV spectrometer camera. Additionally, between the TM and the grating is a manually adjustable pick-up mirror which can be inserted into the beam path in order to reflect light out to an alignment camera, which can be used to realign the pre-generation chamber system whenever it is altered.

cutoffs in the different scenarios. The first scenario to be considered, was a flat-top beam, which was generated by imposing a phase pattern consisting of a π -shift outside a given radius on the SLM. This was tested for a few different radii and compared to two scenarios which did not use the SLM. The second scenario kept the beam power the same as what was used with the SLM, but peak intensity was lowered by partially closing iris 2, thereby lowering the beam power and altering the intensity distribution. In the third scenario, the beam power was lowered such that the peak intensity was identical. These scenarios are illustrated in Figure 22. The divergences of the harmonics generated in these scenarios were compared in order to determine which led to the lowest harmonic divergence. The described

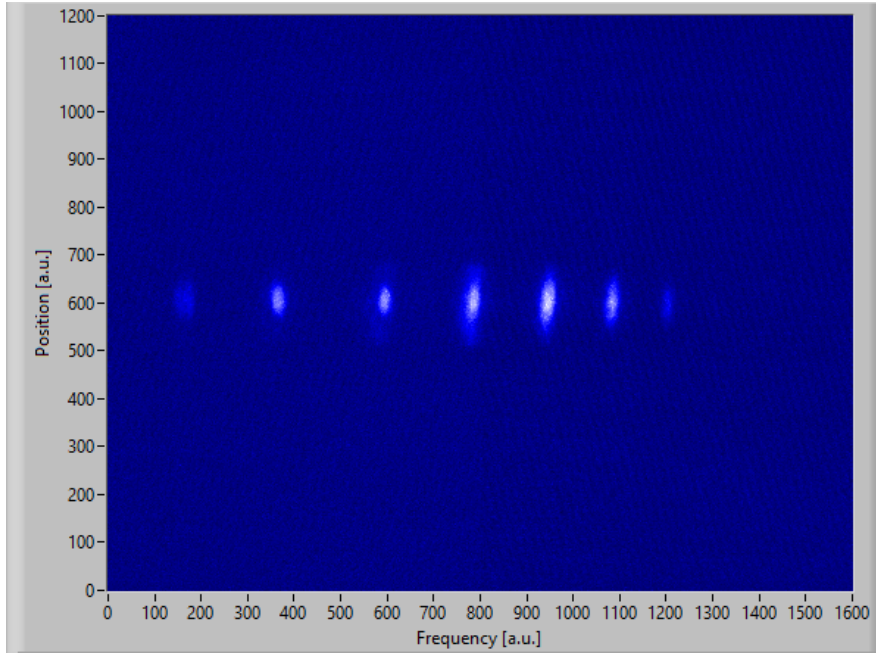


Figure 21: This figure displays a harmonic measurement as taken directly from the XUV spectrometer, as described in Figure 20. This figure was used in the generation of Figure 32, and can be found in the Appendix as Figure 41. In addition, all other XUV spectrometer measurements taken may be found in the Appendix.

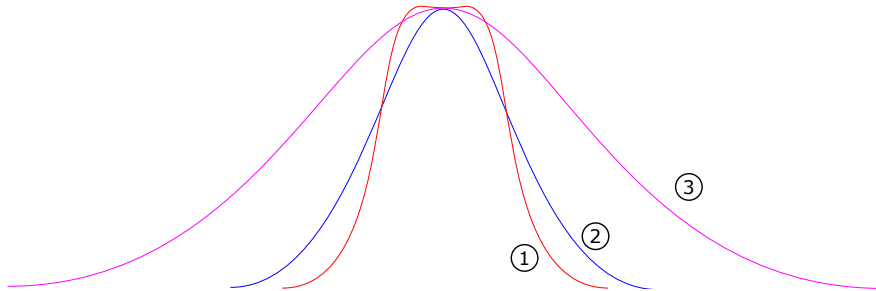


Figure 22: An approximate illustration of the three spatial intensity distribution scenarios considered. The distribution marked 1 (red) represents the spatial intensity distribution of the flat-top case, and scenarios 2 (blue) and 3 (violet) represent the beam having been adjusted by decreasing the beam power and closing an aperture, respectively.

measurements were repeated for an incoming beam power of 0.75 W instead, as this gave stronger harmonics while still being below the damage threshold of the SLM. Measurements of the foci were also taken, all with the camera in the same position, in order to be able to compare them.

3.4 Program for phase retrieval

While the core loop of a phase-retrieval algorithm (here the GS/HIOA, as stated above) does not usually have a complex structure, integrating this into an experimental setup requires some additional components. For the entirety of the process, two separate programs are used to go from pre-experiment intensity measurements to phase manipulation in-experiment. The

smaller of these is rather simple, it simply performs the task of intensity profile interpolation and, if necessary, zero-padding for the purposes of matching the dimensions of the intensity measurements in the near and far field. This program will henceforth be referred to as the "fill-fixing program" or the "filling fixer". The other program, which performs most other tasks relevant to experiments such as the phase-retrieval itself, is the only one actually used during experiments. This program will henceforth be referred to as the "main program" or simply the "program". As an additional note, all two-dimensional matrices are generated or zero-padded such that their column and row numbers are even, and all code is written in LabVIEW and MATLAB. MATLAB code is integrated in the program via so-called MATLAB nodes, which allows the code to give inputs to and take outputs from, MATLAB code, provided a linked instance of MATLAB allowed to run in parallel. This is generally avoided, but is used in a few places where it significantly simplifies or speeds up the code.

3.4.1 Fill-fixing program

The fill-fixing program was made with the knowledge that the spatial width of the plane to which propagation is simulated via FT is dependent on the size in which the source plane is sampled. The optical FT properties of the lens as discussed in Section 2.2.2, more specifically that $\nu_x = \frac{x}{\lambda_f}$ $\nu_y = \frac{y}{\lambda_f}$, can be used in conjunction with Equation 13. This effectively gives us an expression for the side length of our optical FT provided we know the sample spacing in our original picture. Given that the discrete FT also has the same dimensions as its source plane, the amount columns and rows can be used in combination with the side length to determine the sample spacing in this plane as well.

The purpose of the filling fixer is thus to calculate the spatial dimensions of a given plane's FT or vice versa. This is done by using the appropriate focal length and wavelength in combination with the desired spacing and side length of the plane whose dimensions are fixed. This can be adjusted at will but in the experiment the near field plane's sample spacing in both directions was set to 3.69 μm , equivalent to the pixel size of the camera used, in dimensions 4096 by 4096 to give a good resolution in the far field plane as well. For the interpolation, the coordinates of the plane to be rescaled must be known. These coordinates are a matter of physical measurement though, either via having a beam on a camera chip directly, or using graph paper and calculating the distance corresponding to the size of a pixel. Once all required information is entered, the program can be used. Depending on the settings, the input for the source plane and the target plane can be interpolated such that either or both planes can be rescaled onto a new matrix with a correct amount of columns and rows with appropriate grid spacing based on the inputs given. The program then outputs the rescaled plane as a new image with the desired dimensions and is then ready to be used in calculations. There is also an option to perform thresholding of values on the inputs and outputs, if this is desired. The program does need to be readjusted depending on which plane one wishes to base the scaling on, but this can be done relatively simply.

3.4.2 Main program

The main program has several functions, the primary of which is phase retrieval by means of one of the many algorithms explored in Section 2.3. The program has several options which determine what tasks are performed during the runtime. These will be explained as well.

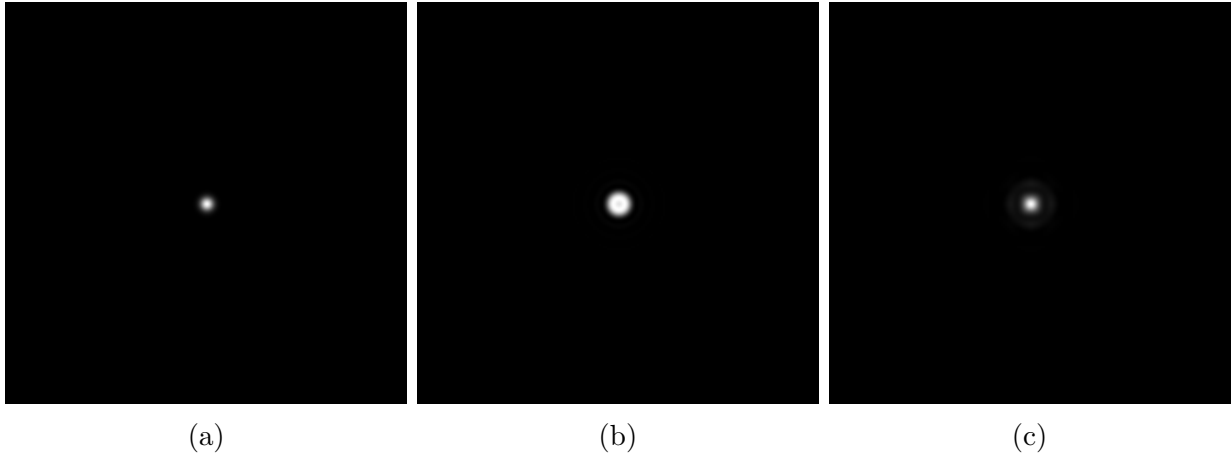


Figure 23: These figures illustrate simulated spatial intensity profiles with and without phase modification. In (a) and (b) the focus of a beam with flat phase and one with an optimized radial π -shift are displayed, respectively. In (c), the spatial intensity profile displayed in (b) is shown after propagation a short distance away from focus. It can be seen to have taken on a similar shape to the profile from (a). This is the reason it is important to remain near the focus of a π -shifted beam if a flat-top intensity distribution is required, as it is only maintained very near the focus. Additionally peak intensity is decreased drastically when a π -shift is applied, with the peak intensity of (b) being 29% compared that of (a). It is also noteworthy that the peak intensity in (c) is higher than that of (b), meaning that the flat-top beam will have so-called "bright spots" which are separate from the flat-top focus.

The program will be described in a chronological order, but as some features are mutually exclusive with one another or togglable, this section will not be entirely linear.

The first inputs which are processed by the program in the case of the GS/HIOA are the source (at the lens) and target (at the focus) planes. The source plane in this case must be an image which has been processed by the fill-fixing program first and the appropriate resolution must be selected. The input image for the target plane can either be taken directly from the camera or an already taken image can be chosen. Both planes can then be thresholded and the intensities in the planes are centered, to ensure that the phase is recovered properly. The target plane is zero-padded such that the dimensions of the two planes match. The power in the planes are normalized with respect to one another.

It is then time for the GS/HIOA itself. This algorithm section of the program takes a few inputs, including the above mentioned amplitude planes, which are still entirely positive and real. Another input is the amount of iterations for which the algorithm should be run, which effectively decides the runtime of the algorithm. The "loop size" is another input which must be given, it determines how many iterations should be in each cycle of the GS/HIOA, as described by the value n in Section 2.3.6. Another input includes the algorithm constant β as seen in Equation 18, which can be changed to alter algorithm behaviour. In the experiments β was left at 0.7. The last input required is a random phase profile, representing ϕ_1 in Figure 8.

After the algorithm is done, the wavelength of the light, the step distance in the source plane, the dimensions of the source plane, and the distance between the lens and the SLM are used

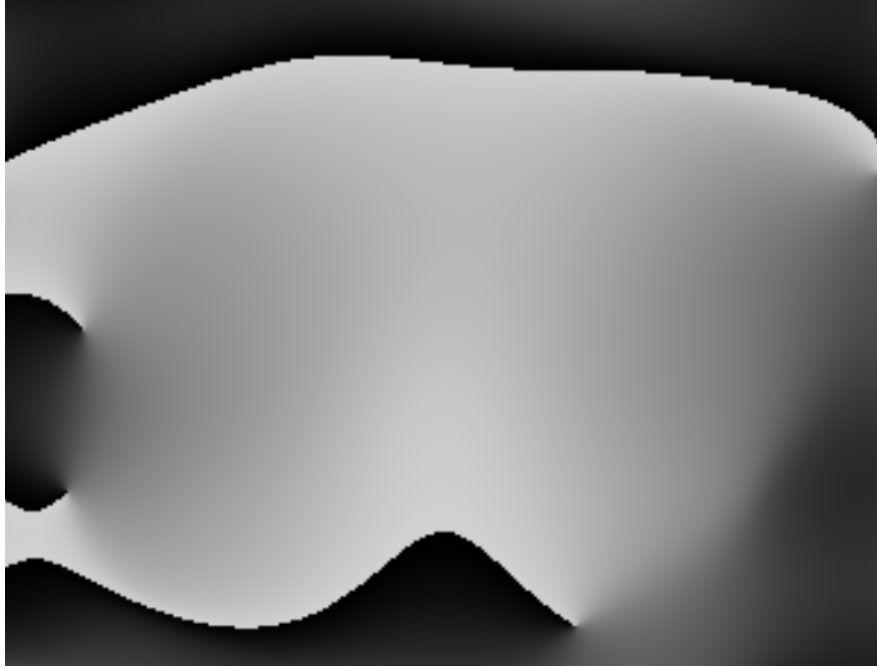


Figure 24: Pictured is the spatial phase profile used on the SLM for the correction of the beam shown in Figure 28. Sharp edges in the phase profile are caused by wrapping of the phase. If one compares the phase profile with the Zernike polynomials in Figure 15, the increasing phase from the center outward in the vertical direction, along with decreasing phase in the horizontal direction out from the center, seems to suggest that there is significant astigmatism affecting the beam.

as inputs to propagate the beam. The algorithm output in the source plane is propagated along the lens-SLM distance in order to compensate the difference in phase of the beam at the SLM relative to the lens for accurate phase correction. This is performed using the modified Fresnel propagation method described in Section 2.3.3. The conjugated plane profile is then applied to the SLM.

The plane profile obtained from the algorithm must be re-scaled for the actual size of the SLM and converted to 8-bit. After this, the phase is added to a flatness correction phase profile, which compensates for the fact that the SLM is not optically flat. This correction was provided by the manufacturer of the SLM. An example of a spatial phase generated by the program which was used for wavefront correction in an experiment is shown in Figure 24.

There is also a separate branch of the program at this point which reads the output immediately after it has been interpolated, but before it is combined with the phase correction pattern. This branch of the program performs a two-dimensional phase unwrapping [22, 23]. This is followed by a decomposition of the wavefront into Zernike polynomials with MATLAB code [24] into an amount of orders of polynomials given as an input. This is done in order to observe if contributions from any specific optical aberrations that stand out can be observed in the decomposition. This was originally intended to be used as the actual output onto the SLM. Mostly, however, the normally calculated phase is used for this purpose. In addition to

this, the magnitudes of the different polynomials are output as a bar chart on the front panel.

Now outside of the main section of the program, only two portions remain before output. The first of these is the spherical phase addition feature, which can be toggled on or off. If toggled on and given the phase matrix converted into 8-bit numbers and a radius of curvature, it adds a wavefront to the phase output equivalent to one of that radius of curvature. This can be used if one wants to move the focus position along the optical axis via the SLM. The second of these two features is the radial π -shift. This simply adds a π -shift in terms the calibration number to the phase outside of a given input radius in the phase matrix. The effect of such a π -shift on a spatial intensity profile at focus and after numerical propagation is illustrated in Figure 23.

4 Results

4.1 Experimental sections 1 and 2

When the results of the phase-retrieval algorithm are evaluated, the main object of interest is the change in peak intensity as a result of the wavefront applied on the SLM. Other measurements of interest include the beam width, quantified as the radius within which approximately 86% of the power is located. The shape of the beam profile, which is rather qualitative, can also help indicate whether or not the wavefront has been corrected. Therefore, these are the main categories which will be observed for the first portion of the results.

Table 1: This table contains the beam widths, where the wavefront of the beam has been conjugated with help of the SLM and without conjugation. These are referred to as corrected and uncorrected, respectively. Additionally, the definition of width used here is the radius within which 86% of the beam energy is located, calculated via integration of the spatial intensity profile. Focus is defined as the position the corrected beam peak intensity is maximal. For comparison, the Rayleigh range for this focus was approximated to be 7 mm.

Distance from focus [cm]	Corrected width [μm]	Uncorrected width [μm]
-1	119	146
0	100	144
1	118	137
2	148	148

For the sake of clarity, when position is referred to here, this refers to the measurement of the camera position via the rail, with a smaller distance being closer to the lens position and vice versa, the rail and setup are shown in Figure 18. The beam passing through the experimental setup in this set of experiments was observed to be astigmatic. Other optical aberrations were likely also present, but astigmatism was deemed to be dominant, so the designated "focus" of the system was assumed to be at the point of least distortion, i.e. the point between the sagittal and tangential foci of the beam. This point was estimated, but was difficult to pinpoint. What could be pinpointed, however, was the position where the

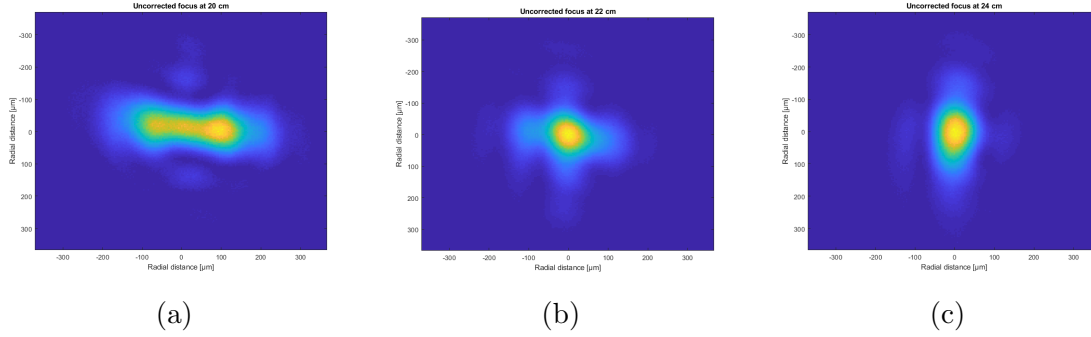


Figure 25: Images of the uncorrected beam profile taken near the point of least confusion. (a), (b) and (c) are taken in positions 2 cm before, at, and after the focus at 22 cm, respectively. As can be seen in (a) the beam has essentially split into two overlapping peaks in the horizontal direction, whereas in (c) no such corresponding separation can be seen in the profile, although most intensity is vertical along the center. Important to note is that the intensities here are normalized with respect to the peak intensity in each figure. The peak intensities of each figure can be seen in the last column of Table 2.

focus with highest peak intensity was observed, which was at 22.0 cm on the rail.

The beam widths of the beam in this instance was calculated for both the corrected and uncorrected beam near the improved focus position and are listed in Table 1. The width is here only calculated for positions relatively close to the center as the uncorrected beam. When beam distortions become too significant, the radius containing 86% of the power is not a useful metric anymore. This is more obvious for positions closer to the lens, as illustrated in Figure 25. Instead, peak intensity will be used to compare beams. These are recorded for beam corrections at multiple positions around the focus in Table 2. If one compares the results of the corrected beam for correction at 22 cm in tables 1 and 2, there appears to be a correlation for higher corrected beam peak intensities and smaller beam widths of said beams. For this reason, the beams will instead be compared based on their peak intensities. The results of Table 2 are plotted in Figure 26 as well.

Table 2: This table contains the peak beam intensities at different positions, as measured by the camera, with position set to zero where the best focus was achieved. U refers to the uncorrected beam, while the peak number refers to the position where the reference plane for wavefront correction was located. All peaks have been normalized with respect to the best optimized peak.

Pos. [cm]	Peak -1 [a.u.]	Peak 0 [a.u.]	Peak 1 [a.u.]	Peak 2 [a.u.]	U. peak [a.u.]
-2	0.63297	0.72601	0.66984	0.30134	0.36044
-1	0.71380	0.87057	0.75824	0.56679	0.51233
0	0.74847	1.0	0.86422	0.7292	0.62295
+1	0.62686	0.82491	0.90403	0.67814	0.66647
+2	0.45472	0.76141	0.84689	0.78926	0.65519
+3	N/A	0.50427	0.60147	0.75360	0.58339
+4	N/A	0.32625	0.48132	0.53480	0.43663

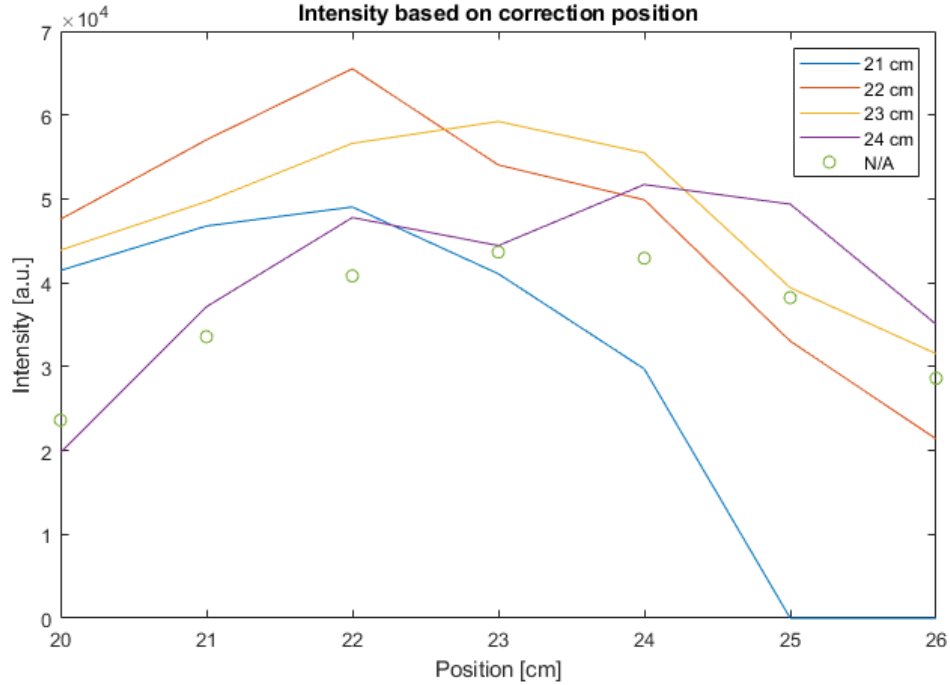


Figure 26: Illustrated here are the plotted results from Table 2. The curve indicated by points refers to the uncorrected beam, whereas the lines refer to corrected beams, separated by the position at which their reference plane was located.

A comparison of the beam width for the corrected and uncorrected beam as referred to in Table 1 at the focus can be seen in Figure 28. It seems rather clear that the beam profile has become much more close to a Gaussian beam than the original beam, indicating that the program indeed has been successful at correcting for some aberrations. In addition, the peak beam intensity seems to have increased in the beam with over 50% by correction in this optimal case, when comparing the highest peak intensities of uncorrected and corrected beams in Table 2. The next interesting point in Table 1 is at 2 cm from the focus (+2 cm), where the calculated widths are roughly equal. An intensity profile comparison in the horizontal and vertical direction can be observed in Figure 27, where it can be seen that the uncorrected beam is sharper in the horizontal direction. This is to be expected, as this seems to be the sagittal focus of the beam, and the uncorrected beam can also be seen to be wider in the vertical direction. A slight difference in width for the corrected beam in the two axes can also be observed, with the vertical component appearing slightly wider than its horizontal counterpart. This is likely due to that the wavefront correction of the beam is not perfectly flattened and is still mildly aberrated.

As can be seen in Figure 26 and Table 2, there is no data recorded for peak 21 at positions 25 and 26 cm. This is due to a combination of a rapid drop in intensity and the failure of the beam profile to form a single Gaussian peak, instead splitting, similarly to the profile as seen in Figure 25a. This was worsened at 20 cm where it was not possible to get any functioning correction of the beam at all. This seems to be tied to the fact that the uncorrected beam is split into two overlapping peaks nearer the tangential focus, but no solid conclusion has been reached about this yet.

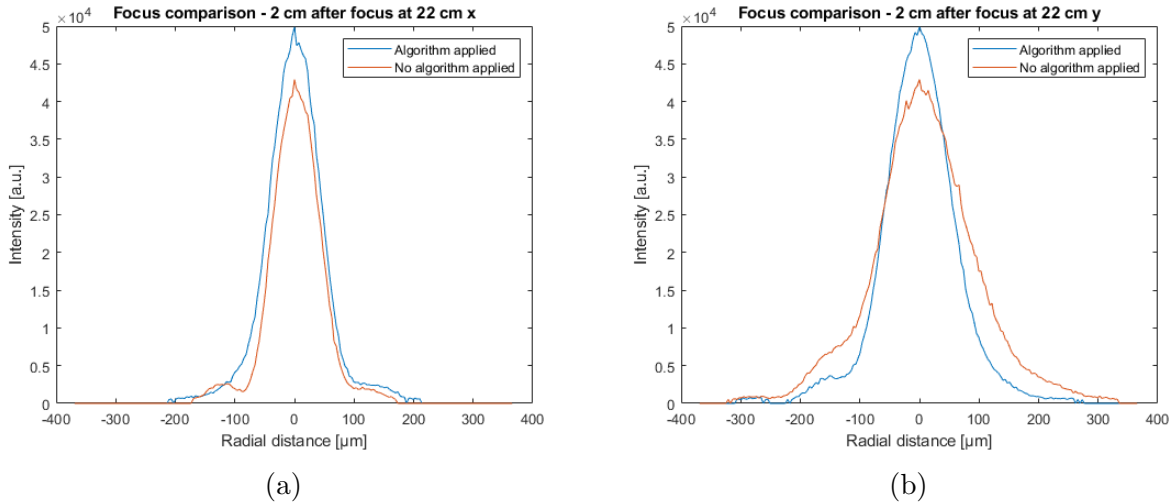


Figure 27: These graphs show the intensity distribution at 2 cm after the focus (24 cm) in the horizontal and vertical plane of the beam corrected at 22 cm in (a) and (b) respectively. A comparison with the uncorrected beam intensity can also be seen in each figure.

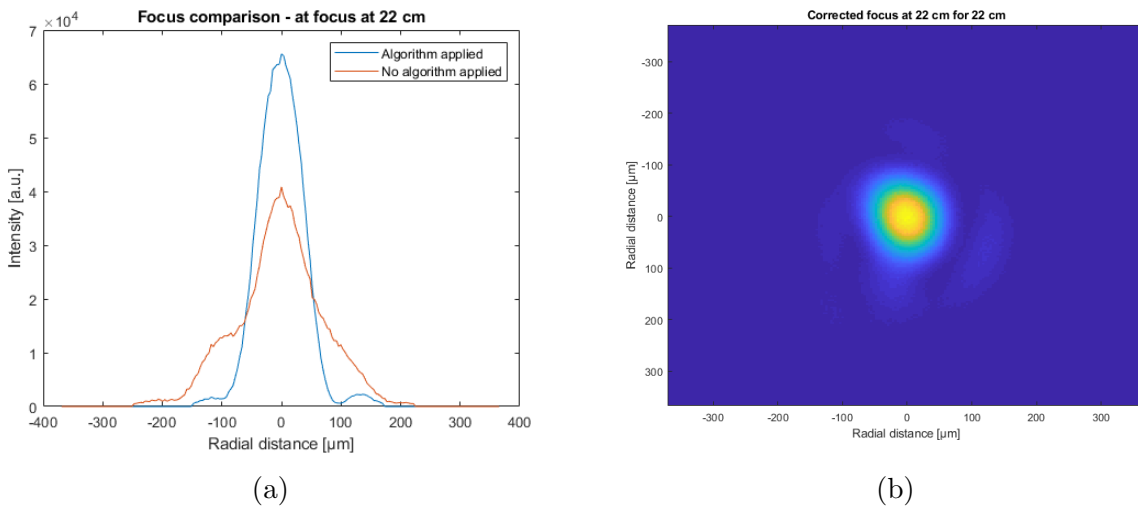


Figure 28: These figures show the intensity for the corrected beam at 22 cm, whose width is plotted in Table 1. (a) shows a cut in the horizontal plane of the intensity as compared with the intensity shown in Figure 25b. (b) shows the two-dimensional intensity profile of the corrected beam. The only difference between the curves seen in (a) is the addition of a correction pattern on the SLM.

Additionally, it can be observed in Figure 26 that the peak achieved intensity reached to depend on where the correction of the beam seems to occur. This makes sense, as the GS/HIOA presumes the target plane intensity to be in the focus of the lens, meaning that the algorithm yields better results more closely to this point. This is to be expected, as the algorithm expects the reference target plane to be in the Fourier plane. The general trend of the corrected beams which maintain a relatively Gaussian shape, that is beams corrected at 0, +1 and +2 cm, all approach peak intensities similar to the uncorrected beam some

distance away from the focus. This can be more clearly seen further away from the lens, due to the more homogeneous intensity profile of the uncorrected beam there. This similarity in peak intensity after a certain point is caused by the fact that the shape of the wavefront has a larger impact on the intensity when the beam is focused because there the component wavevectors of the wavefront play an important role in determining focus size, as was discussed from Section 2.2.2. As the wavevector components of the beam's wavefront become spatially dispersed away from the focus, their impact on the beam shape and therefore the wavefront correction, becomes less meaningful.

All results are taken with the same filter settings on the camera in order for them to be comparable. It would have been possible to use less filters to view intensities below what could be measured here, but this would have resulted in camera saturation for the highest peaks, with the highest peak of the best corrected beam (0 cm) very nearly saturating the camera with the settings used. Additionally, it is important to note that the program did not produce a phase which was capable of improving the focus every time, and that the grade of improvement this correction phase would bring could vary. However, for the planes where a relatively good wavefront correction could be achieved, that is 0, +1 and +2, the program was estimated to yield a correction which gave an improvement in peak intensity on the first try more often than not, approximately 5 out of 7 times. These tests were also done with the program running the recommended iteration sizes of 40 iterations per loop [17], but with only 2 loops or 80 iterations in total, which is in general less than recommended. It is possible that this impacted results as well, but this is not certain. Since the aberrations we correct for are static, it is not of ultimate importance that the correction employed here works perfectly every time. Ultimately, from the results of this section of the experiment it does appear that the program, in conjunction with the SLM, is capable of reducing aberrations.

4.2 Experimental section 3

At this point, in the HHG section of the experiment, the driving beam focused almost without aberrations without the application of an SLM. Therefore, it was tough to achieve an improvement that goes beyond only correcting for aberrations introduced by the SLM itself. In this part of the experiment, however, the point of using the SLM was mainly for it to be used as a beam shaping device in order to manipulate the divergence of the harmonics rather than wavefront correction of the driving beam. The results of interest here are therefore the divergence of the generated harmonics, the intensity profile of the foci and the spectral measurements, all for the different settings with and without the SLM.

The addition of the SLM and the two mirrors required to couple it in, as illustrated in Figure 19 reduced the measured power at the focusing mirror from 0.5 to 0.4 W. This can be attributed to the use of the silver-coated mirrors (approximately a 4% loss) and the SLM itself. This also led to a drop in peak intensity at the focus, as is illustrated in Figure 29, with little change in the actual shape of the profile, as mentioned above. The beam power was then further decreased in all cases due to the adjustment of iris 1, as described in Figure 19. The peak intensity was also affected by the radius of the π -shift in the experiment. A comparison between the peak intensities in the focus of beams with different settings and with no SLM coupled in can be seen in Figure 30. The peak intensity can be seen to decrease for decreasing π -shift radius, which is to be expected, as a larger part of the beam is then

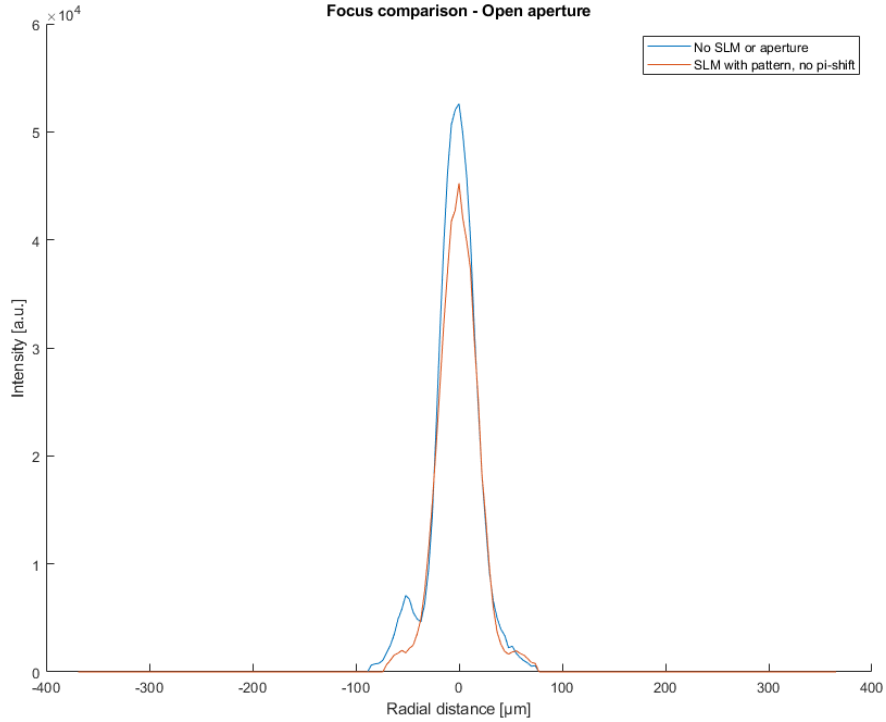


Figure 29: This graph illustrates the intensity profile of the focused beam used for HHG with and without the SLM. This intensity measurement is taken with all irises completely open, which is not the conditions the rest of the experiment was conducted under.

π -shifted to interfere destructively in the focus. It can also be seen in Figure 30a that larger radii are more similar to cases with identical aperture sizes without the SLM, as mostly unshifted light is let through the iris. It also be seen in Figure 30 that the intensity profile of the π -shifted beams are not much flatter than a Gaussian beam. This is likely due to the fact that all images are taken in the same spot and have not accounted for the possibility of the focus moving when a π -shift was applied. This means that the image may not have been taken in the correct plane for a flat-top.

Comparisons of the intensities of the integrated harmonic spectra themselves for 0.5 W can be seen in Figure 31. These integrated spectra are taken from the direct measurements of the XUV spectrometer at 0.5 W, which can be found in the Appendix. The intensities of the harmonics can here be seen to not match especially well in most scenarios. This is likely due to the change in peak intensity between scenarios, as illustrated by Figure 30.

For the integrated spectrum when a 0.75 W input was used, as seen in Figure 32, there is good intensity agreement for several harmonics between the three different scenarios with the π -shift set at 3.5 mm. The harmonic near 325 eV in particular showed rather good agreement in general, and was picked for a divergence comparison. A divergence comparison was done by spatially integrating over single harmonics of the full spectral measurements, all of which can be seen in the Appendix, such that the spatial width could be extracted. The vertical dimension was here unaffected by the spectrometer grating and therefore the beam width in the vertical dimension is used as an indication of the divergence. The harmonic which can

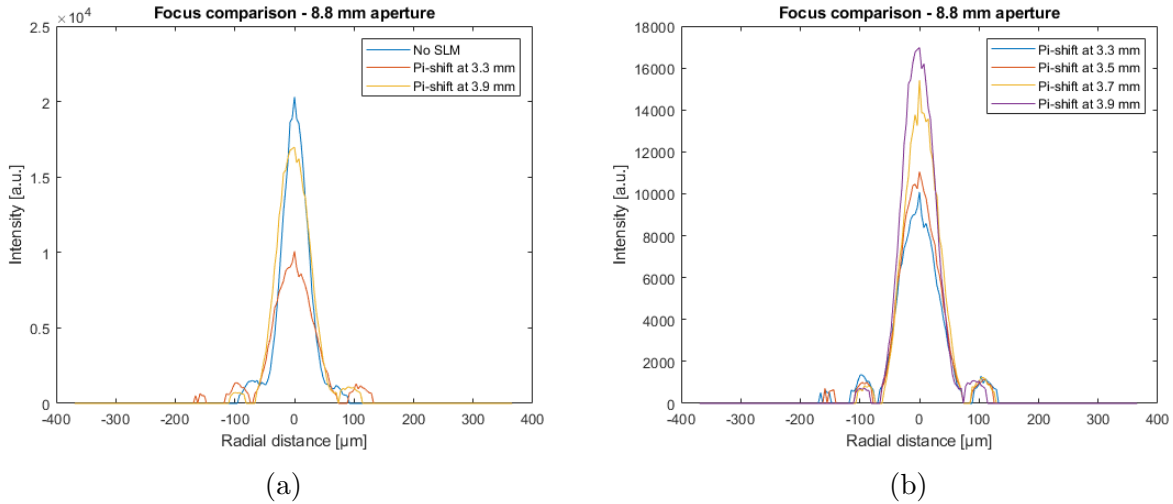


Figure 30: These figures show the comparison of the focal intensity for an iris setting which was used for HHG. (a) shows a comparison with the SLM coupled out and the largest and smallest π -shift radii, whereas (b) shows a comparison of all the π -shift radii used.

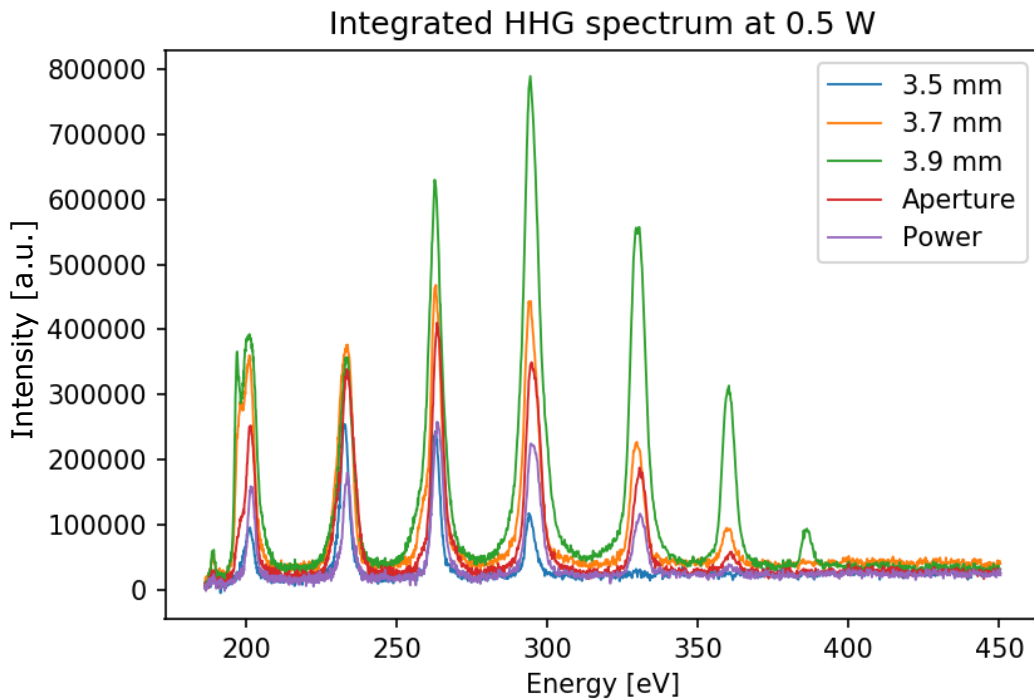


Figure 31: This figure shows the integrated HHG spectrum with the intensities of the different harmonics for an incoming beam with 0.5 W power. "Power" and "Aperture" refer to the spectra which were measured when the beam power and aperture sizes were adjusted to match the cutoff of the same beam, respectively. The mm distances refer to the radii of π -shifts used. The x-axis of the figure is given in electronvolt, based on their position on the camera, as light has here passed a diffraction grating, causing them to separate spatially by wavelength.

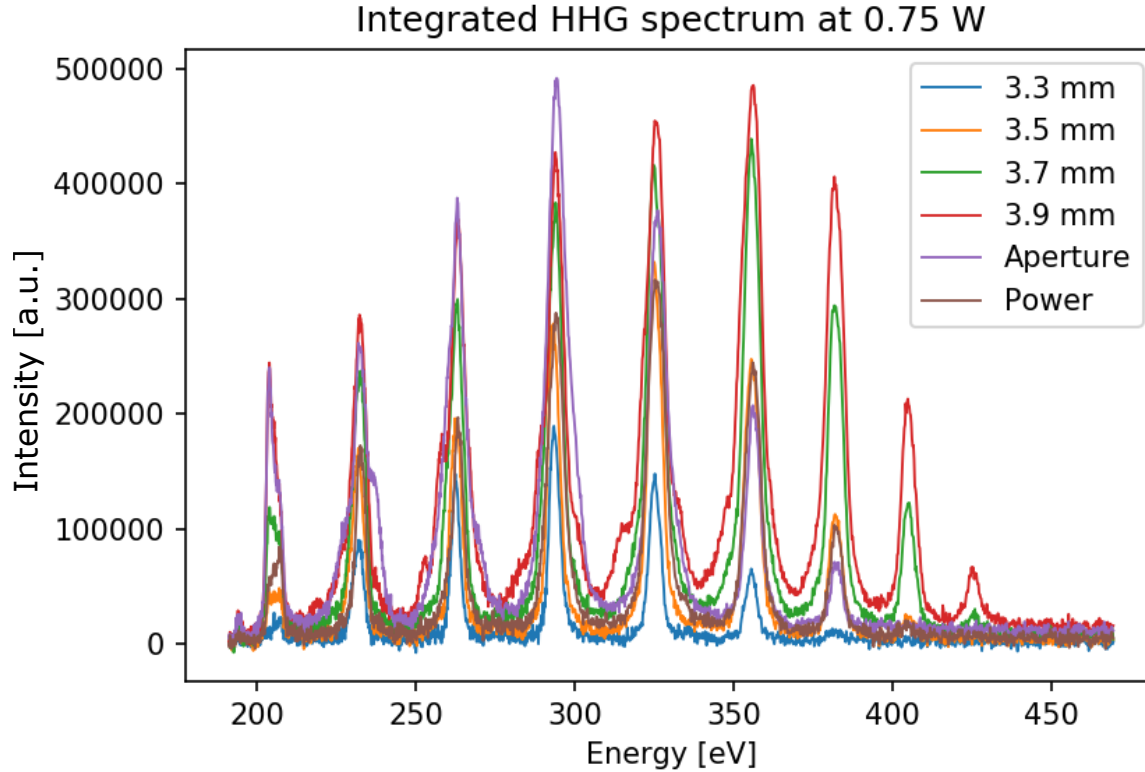


Figure 32: This figure shows the integrated HHG spectrum with the intensities of the different harmonics for an incoming beam with 0.75 W power. The labels in the legend refer to the method used in acquiring the measurement, as in Figure 31. The x-axis of the figure is again given by position on the camera converted into eV via the diffraction grating prior to the spectrometer camera.

be seen in Figure 32 at approximately 325 eV, which was determined to be harmonic 21 of the driving field, was used for this purpose.

The width was determined by fitting a Gaussian to a spatially integrated harmonic. The values which were acquired from this were tabulated in Table 3, and the integrated harmonics can be seen in Figure 33. As can be seen in Table 3, the only fitted beam using the SLM which appears to be less wide is the one using a 3.3 mm radius π -shift. However, as can be seen in Figure 33 this is because the peak of this harmonic is a lot lower than all the other harmonics, additionally indicating that peak intensity was significantly different in this case. The 3.5 and 3.7 mm radii, which are the most similar in peak intensity to the non-SLM cases are both wider, and given that the profile for 3.7 mm is not Gaussian, it is likely to be even wider than the fit suggests. Thus, it seems like the divergence of the harmonics is reduced the most when using a previously functioning aperture size and adjusting the power for similar cutoff to the 3.5 mm shift, rather than using a radial π -shift or adjusting the aperture.

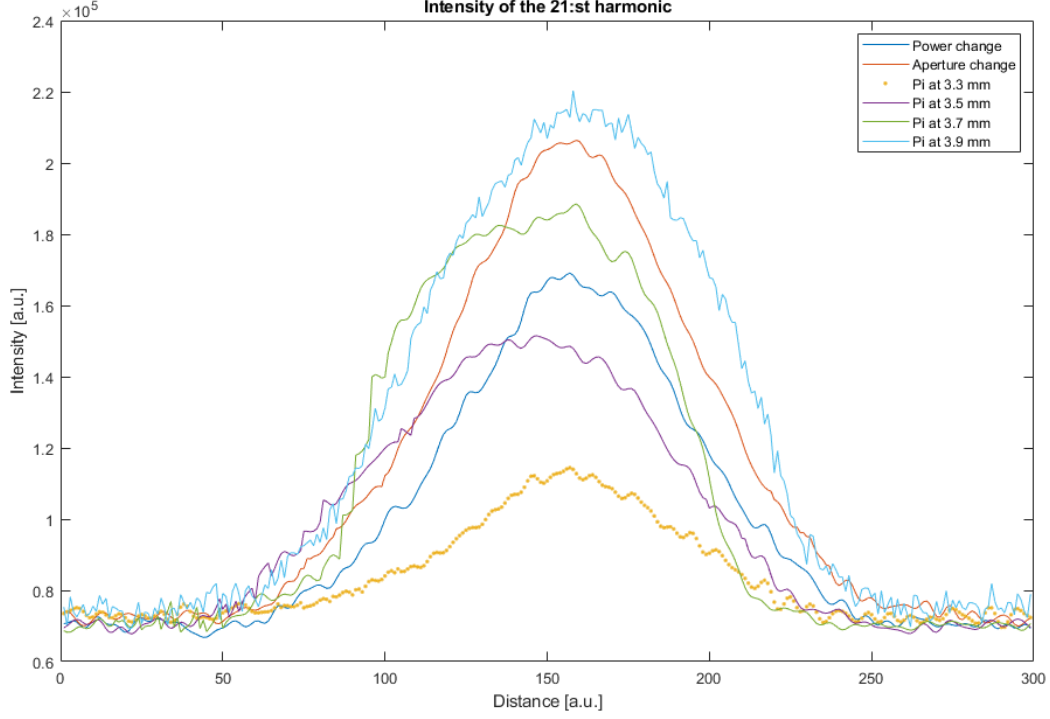


Figure 33: This figure contains the integrated intensity profile across the divergence axis of the 21:st harmonics for an input of 0.75 W. The data has here been low-pass filtered before integration to reduce noise.

Table 3: This table contains the fitted width of the 21:st harmonic for the input beam of 0.75 W. The column on the left describes the scenarios which corresponds to the width on the right, this column also corresponds to the cases as indicated in the legend of Figure 33. Width units corresponds to the same arbitrary units used in Figure 33.

Scenario considered at harmonic 21	Fitted width [a.u.]
Power change	84.5
Aperture change	88.6
3.3 mm π -shift	77.7
3.5 mm π -shift	98.5
3.7 mm π -shift	92.0
3.9 mm π -shift	95.9

As discussed in Section 2.1.1, it is not only the divergences of the harmonics for different cases which are of interest, but also the relative divergence of consecutive harmonics in a given case. Therefore, a comparison of the widths of different harmonics was made. This comparison was made in the 0.75 W input, 3.5 mm radius case, as this was the case with the most similarly intense harmonics, in addition to their divergence profiles seemingly being the most Gaussian out of all the SLM-using cases. The results of this are tabulated in Table 4 and intensity profiles are shown in Figure 34. From these results, it appears that the divergence is not consistent from harmonic to harmonic. In Figure 34 a plateau-like structure

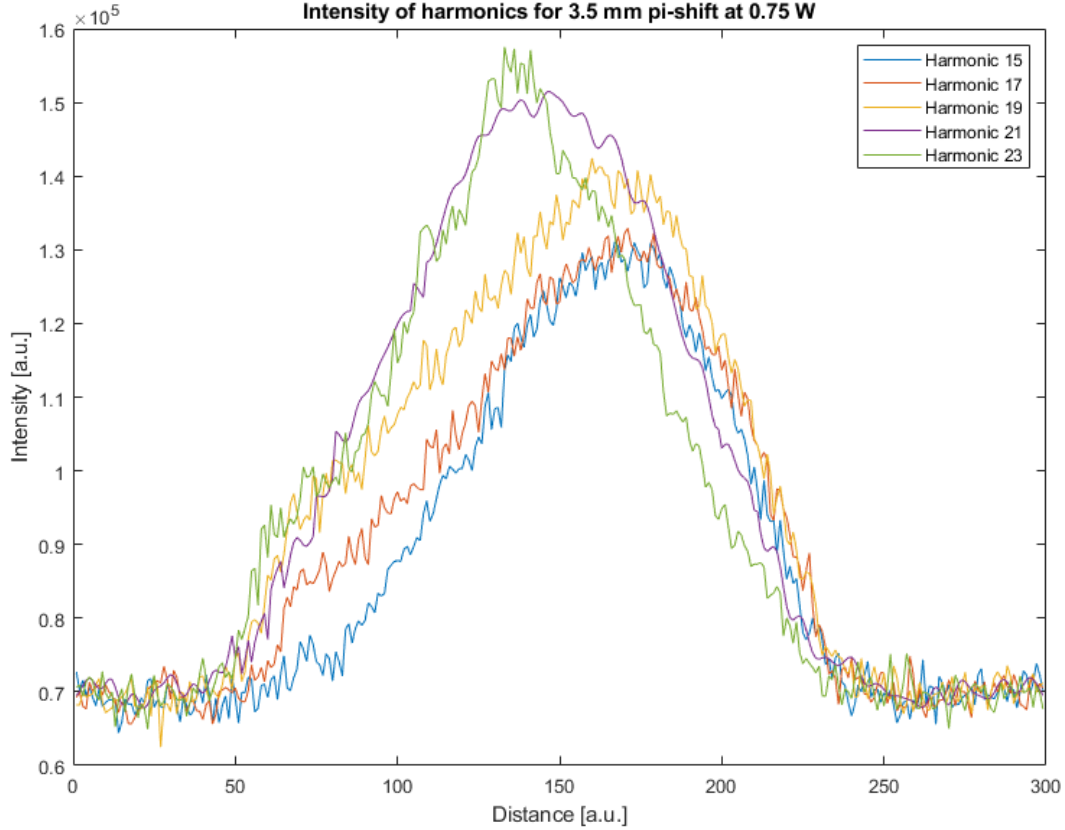


Figure 34: This figure contains the integrated intensity profile across the divergence axis of the harmonics in the case of a 3.5 mm radius π -shift and a beam input of 0.75 W. The data has here been low-pass filtered before integration to reduce noise.

can be seen on the left side of the beam profile of harmonic 17 and 19, which were partially removed in order for the fit to be possible, but regardless, harmonics 21 and 23 were rather well-fitted and did still not show very consistent widths. It thus seems that it was not possible to achieve identical harmonic divergences with a 3.5 mm radial shift. It appears that this is the case for the other diameters of π -shift as well, given that their spatial intensity profiles are similar, with different peaks, as illustrated in Figure 30. This seems to be confirmed at a glance at the figures in the Appendix as well.

To summarize, there does not seem to be a lower divergence for π -shifted beams when compared to their counterparts without SLM. In general, out of these, the beam with its power adjusted for somewhat similar cutoffs seems to give a better divergence than other cases. After the initial alignment, the focus measurement and gas cells were left as they were, to avoid requiring realignment. Because it is possible that both different π -shifts and different aperture sizes moved the focus somewhat, this will have impacted the effectiveness of HHG for different beams as well. Given that the intensity profiles in Figure 30 don't appear to be significantly flatter than a Gaussian distribution, this may be the case. This is something which has to be taken into consideration in any future attempts. It also appears that the relative divergence of harmonics does not disappear for any of the π -shifted beams

Table 4: This table contains the width of the fit of different harmonics in the case of a 3.5 mm radius shift and a 0.75W input beam. Width is in the same arbitrary units as are used in Figure 34

Harmonic number	Fitted width [a.u.]
15	86.2
17	100.6
19	110.3
21	98.5
23	94.2

either. Thus, it is possible that a phase other than a radial π -shift could be used as a way to form a more flat intensity profile. It is also possible that this lack of flatness could be adjusted for by finding (if possible) a plane where the spatial intensity distribution is flatter. This was not done here, as all focal measurements were taken in the same plane. Additionally, the fact that an extended gas target, like the gas cell, was used likely had an effect on the divergence as well.

5 Conclusion and outlook

In this project, a program has been written for the purpose of beam shaping and wavefront correction in conjunction with an SLM. Several different phase-retrieval algorithms were tested. An algorithm was then chosen and the program, together with the SLM, was used to correct a beam wavefront. Afterwards, HHG was performed with the SLM used for beam shaping with the aim of producing a flat-top spatial intensity distribution to drive HHG. HHG was carried out with and without the SLM in several different scenarios.

The primary outcomes of this appears to be that the phase-retrieval program and SLM was shown to successfully improve the focus for a beam with some obvious optical aberration. Additionally, it was shown that the SLM in conjunction with the program can shape an incoming beam. This in turn appears to have had some effect on the spatial intensity distribution of the harmonics, which was expected. However, the desired reduction in divergence of the XUV harmonics was not observed within the experimental work. The phase-retrieval program also did not correct the beam significantly when a beam with an already nearly diffraction limited spot was used.

With regards to future development, different phase-retrieval algorithms could be used in the experiment, such as the SPPA. Moreover, one could also use Zernike optimization for focus optimization in future iterations. That is, one could find Zernike coefficients for a number of aberration orders that maximize the focal intensity to assist wavefront correction. If possible, one also could try using a gas jet instead of a gas cell, as this likely affects the impact of the driving field spatial intensity profile on the harmonics. Furthermore, one has to take into consideration several factors, such as the possibility that the flat-top intensity profile is located in a different plane relative to a regular focus. Additionally, a method to quantitatively compare the intensity to drive HHG with different generation conditions would be needed. The ability to control the focus position, focus shape and exact position of the gas jet and intensity is a requisite for further experiments.

References

- [1] P. A. Franken, A. E. Hill, C. W. Peters, and G. Weinreich (1961) Generation of Optical Harmonics, *Physical Review Letters* Volume 7, Issue 4, 118-119
- [2] Li X. F., L’Huillier A., Ferray M., Lompre L. A., Mainfray G. (1989) Multiple-harmonic generation in rare gases at high laser intensity. *Physical Review A* Volume 39, Issue 11, 5751–5761
- [3] Bahaa, E. A. Saleh and Malvin, Carl Teich, (2007), Fundamentals of Photonics, Third Printing. Chapters 2 and 4, Appendices A and B *John Wiley and Sons*. ISBN 978-0-471-35832-9.
- [4] Paul P. M., E. S. Toma, Pierre Breger, Geneviève Mullot, F. Augé, Philippe Balcou, Harm Geert Muller and Pierre Agostini (2001) Observation of a train of attosecond pulses from high harmonic generation. *Science* Volume 292, Issue 5522, 1689-92
- [5] Kroon D. (2016) Attosecond Interferometry: Techniques and Spectroscopy, *Department of Atomic Physics, Lund University* ISBN 978-91-7623-688-8.
- [6] Wikmark H., et al. (2019) Spatiotemporal coupling of attosecond pulses. *Proceedings of the National Academy of Sciences*, Mar. 116 (11) 4779-4787;
- [7] Lifa Hu, Li Xuan, Yongjun Liu, Zhaoliang Cao, Dayu Li, and QuanQuan Mu (2004) Phase-only liquid-crystal spatial light modulator for wave-front correction with high precision, *Optics Express* 12, 6403-6409
- [8] L’Huillier A. (2019) Current trends in atomic physics, eds. A. Browaeys, T. Lahaye, T. Porto, C. S. Adams, M. Weidemüller, and L. F. Cugliandolo, Chap. 8 “Generation of high-order harmonics and attosecond pulses” *Oxford University Press*
- [9] Voelz D. G. (2011) Computational Fourier Optics: A MATLAB Tutorial, chapters 2.1 and 5.1, accessed via *SPIE digital library*, <https://www.spiedigitallibrary.org/ebooks/TT/Computational-Fourier-Optics-A-MATLAB-Tutorial/eISBN-9780819482051/10.1117/3.858456?SSO=1> on 21-01-2020
- [10] R. G. Lane and M. Tallon (1992) Wave-front reconstruction using a Shack–Hartmann sensor, *Applied Optics* Vol.31, Issue 32, 6902-6908
- [11] Gerchberg R. W. and Saxton W. O. (1972) A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik*, Vol. 35, (No. 2), 237-246
- [12] Meng F., Zhang D., Wu X., and Liu, H. (2009) A comparison of iterative algorithms and a mixed approach for in-line x-ray phase retrieval, *Optics communications* 282(16), 3392–3396. <https://doi.org/10.1016/j.optcom.2009.05.027>
- [13] Osherovich E. (2012) Numerical methods for phase retrieval, <https://arxiv.org/abs/1203.4756v1>
- [14] Fienup J. R. (1982) Phase retrieval algorithms: a comparison. *Applied Optics*, Vol. 21, Issue 15, 2758-2769

- [15] Bruel L. (30 May 2003) Numerical phase retrieval from beam intensity measurements in three planes. *Proceedings SPIE 4932, Laser-Induced Damage in Optical Materials: 2002 and 7th International Workshop on Laser Beam and Optics Characterization*
- [16] Described to me and discussed in e-mail correspondence with Dr. David O'Dwyer of ASML research department in Veldhoven, the Netherlands
- [17] Guo C., Liu S. and Sheridan J. T. (2015) Iterative phase retrieval algorithms. I: optimization. *Applied Optics*, Vol. 54, Issue 15, 4698-4708
- [18] F. Zernike (1934) Beugungstheorie des Schneidenverfahrens und Seiner Verbesserten Form, der Phasenkontrastmethode, *Physica*, Vol. 1, Issue 8, 689–704.
- [19] Hamamatsu Opto-semiconductor handbook, Chapters 1 and 12, from https://www.hamamatsu-news.de/hamamatsu_optosemiconductor_handbook/, accessed on 07-05-2020
- [20] Hamamatsu LCoS-SLM modules, X13268/X13139 series, from https://lcos.hamamatsu.com/resources/pdf/ssd/x13268_series_etc_kacc1213e.pdf, accessed on 07-05-2020
- [21] Rosales-Guzmán C. and Forbes A. (2017) How to shape light with spatial light modulators. *SPIE Spotlight eBooks*, Vol. SL30, Accessed from: <https://www.spiedigitallibrary.org/ebooks/SL/How-to-Shape-Light-with-Spatial-Light-Modulators/eISBN-9781510613010/10.1117/3.2281295> on 21-01-2020
- [22] M. A. Herraiez, D. R. Burton, M. J. Lalor, and M. A. Gdeisat (2002) Fast two-dimensional phase-unwrapping algorithm based on sorting by reliability following a noncontinuous path, *Applied Optics*, Vol. 41, Issue 35, 7437-7444
- [23] M. F. Kasim (2017) Fast 2D phase unwrapping implementation in MATLAB, https://github.com/mfkasim91/unwrap_phase/ Accessed February 06, 2020
- [24] Paul Fricker (2020) Zernike polynomials Central File Exchange, Retrieved February 08, 2020 <https://www.mathworks.com/matlabcentral/fileexchange/7687-zernike-polynomials>
- [25] Franzen A. (2006) ComponentLibrary, accessed via <http://www.gwoptics.org/ComponentLibrary/> on 18-04-2020, ComponentLibrary by Alexander Franzen is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported License, more information can be found at <https://creativecommons.org/licenses/by-nc/3.0/legalcode>, accessed on 18-04-2020

Appendix: Full HHG spectra

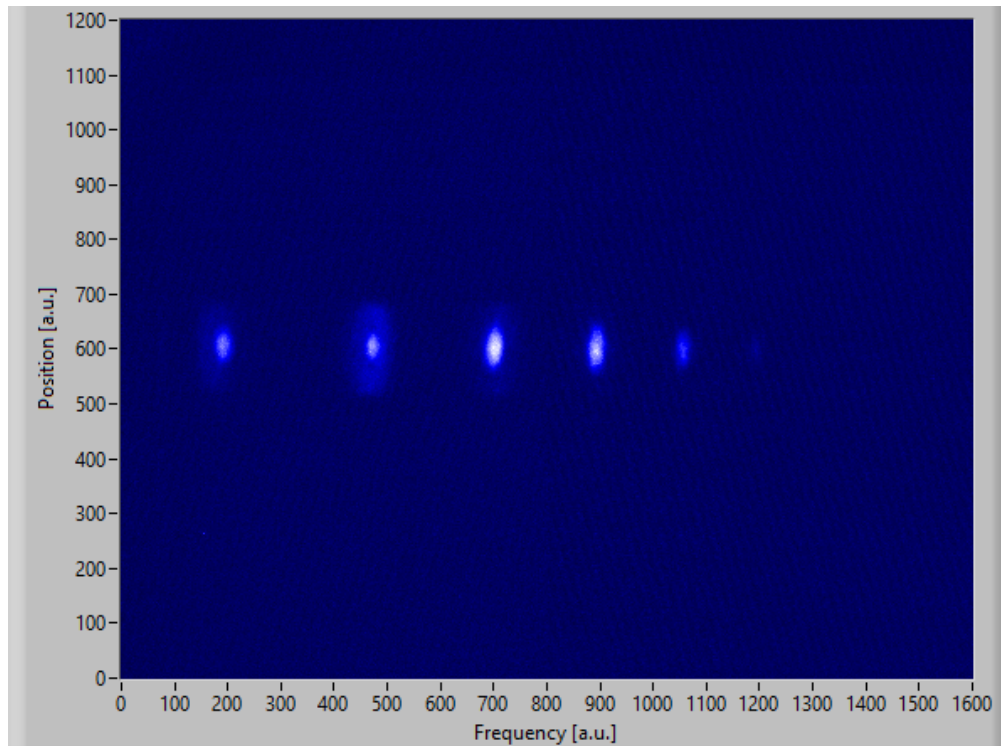


Figure 35: 0.5 W input with aperture adjusted.

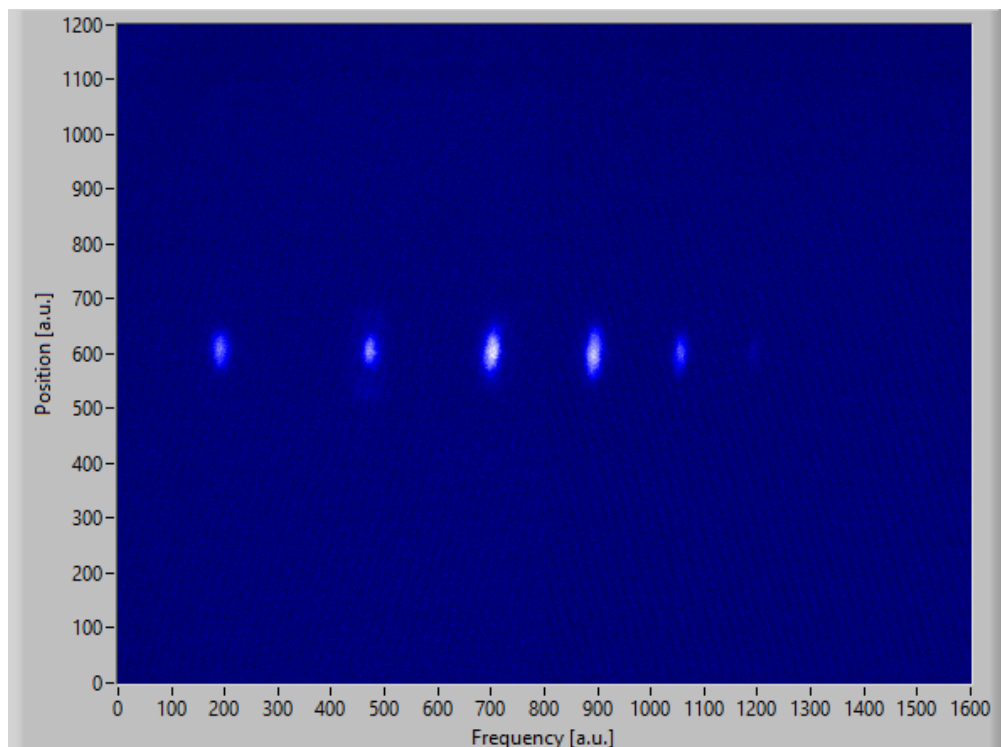


Figure 36: 0.5 W input with power adjusted.

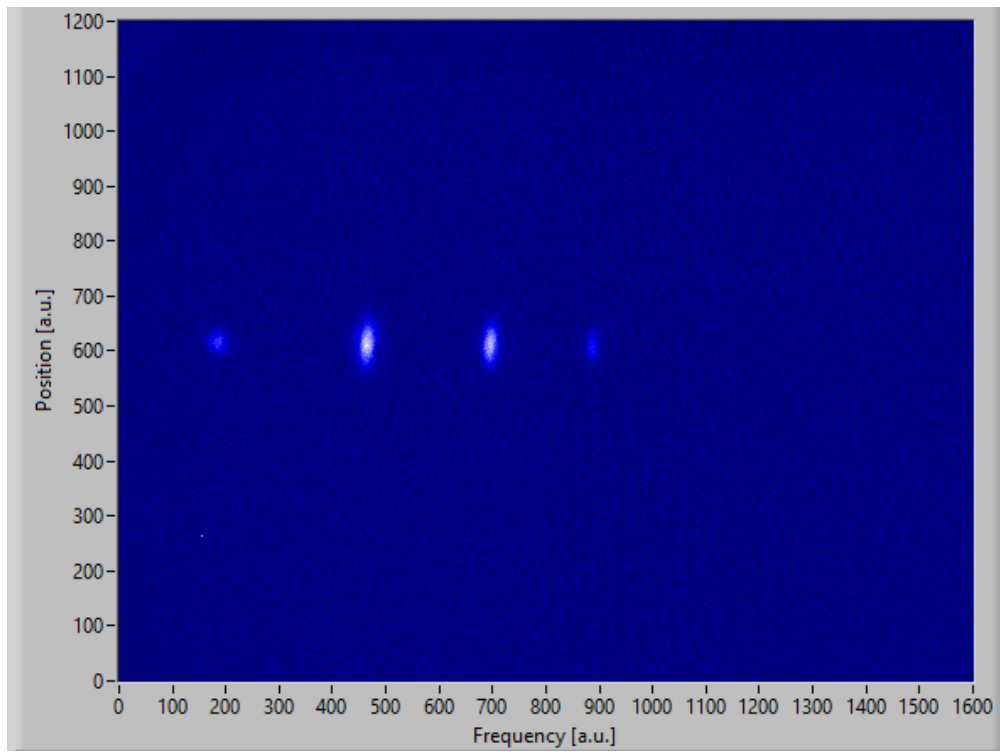


Figure 37: 0.5 W input with 3.5 mm shift.

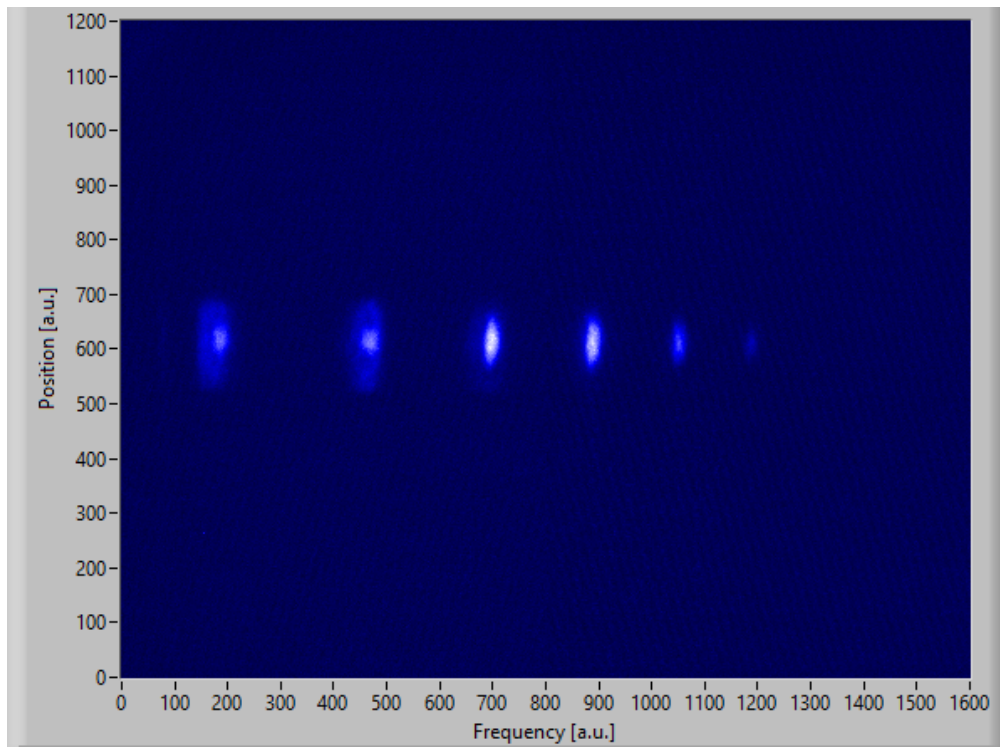


Figure 38: 0.5 W input with 3.7 mm shift.

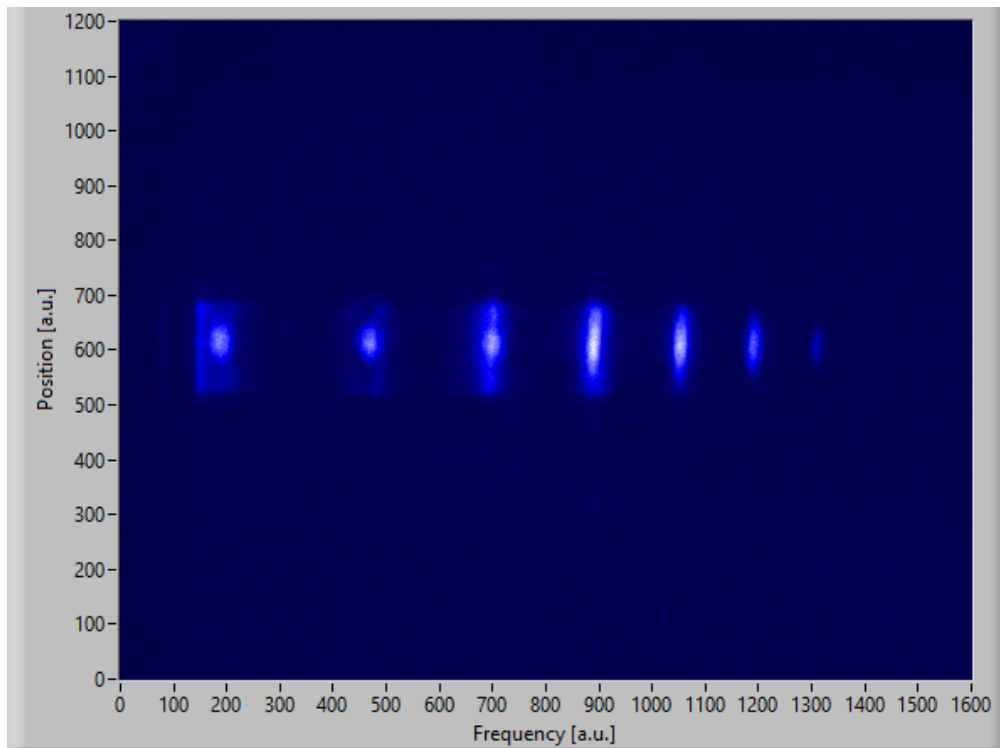


Figure 39: 0.5 W input with 3.9 mm shift.

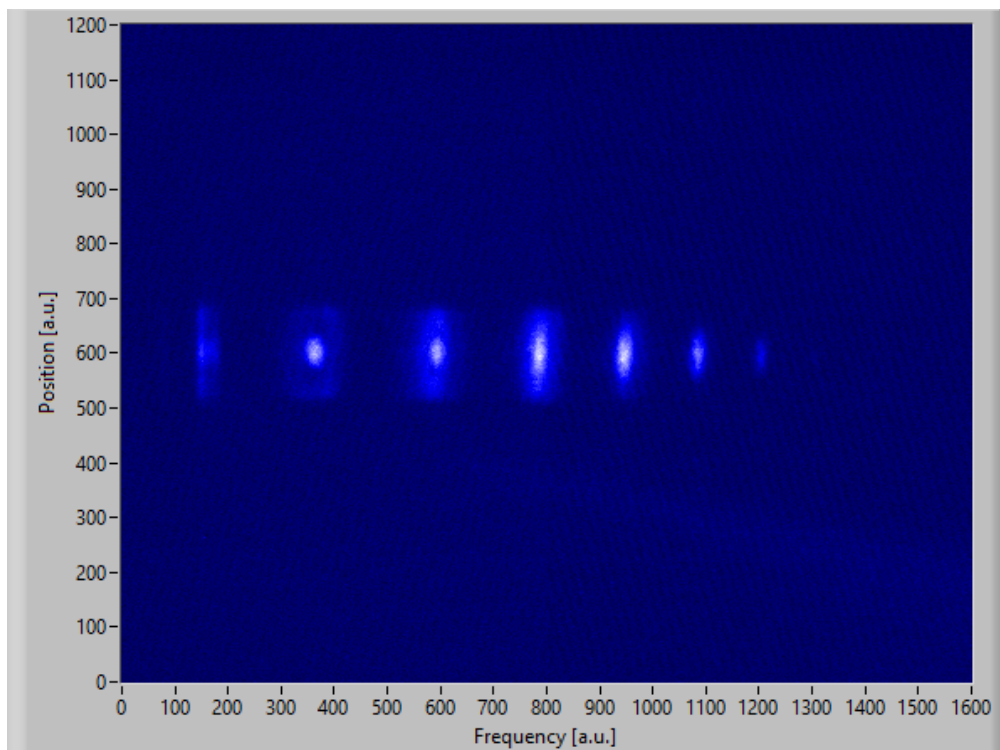


Figure 40: 0.75 W input with aperture adjusted.

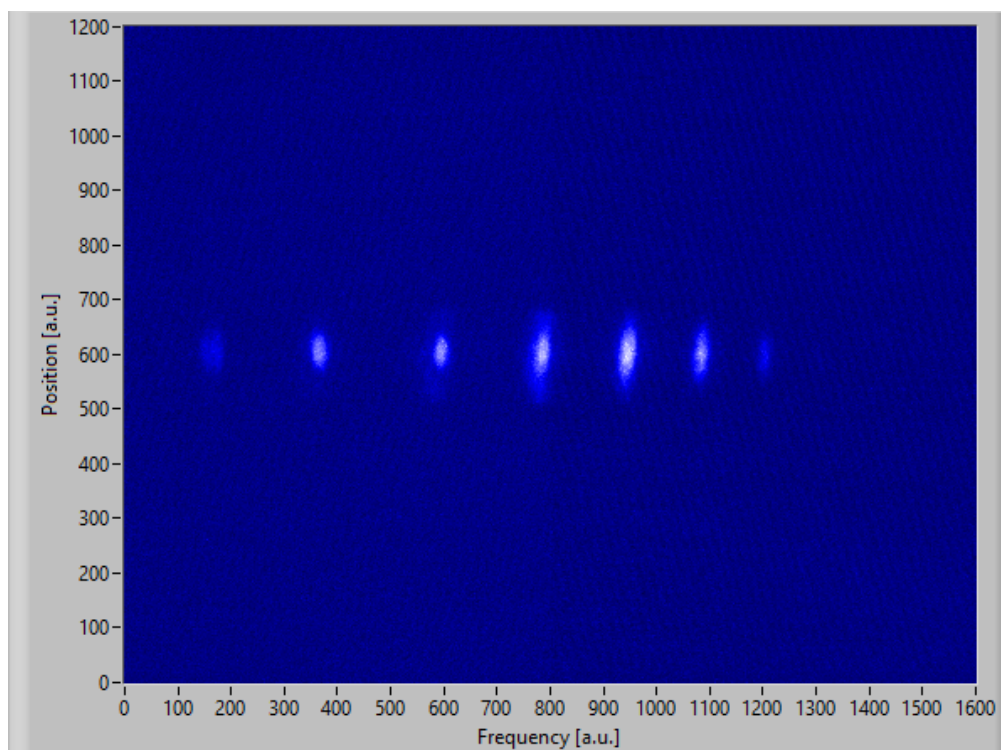


Figure 41: 0.75 W input with power adjusted.

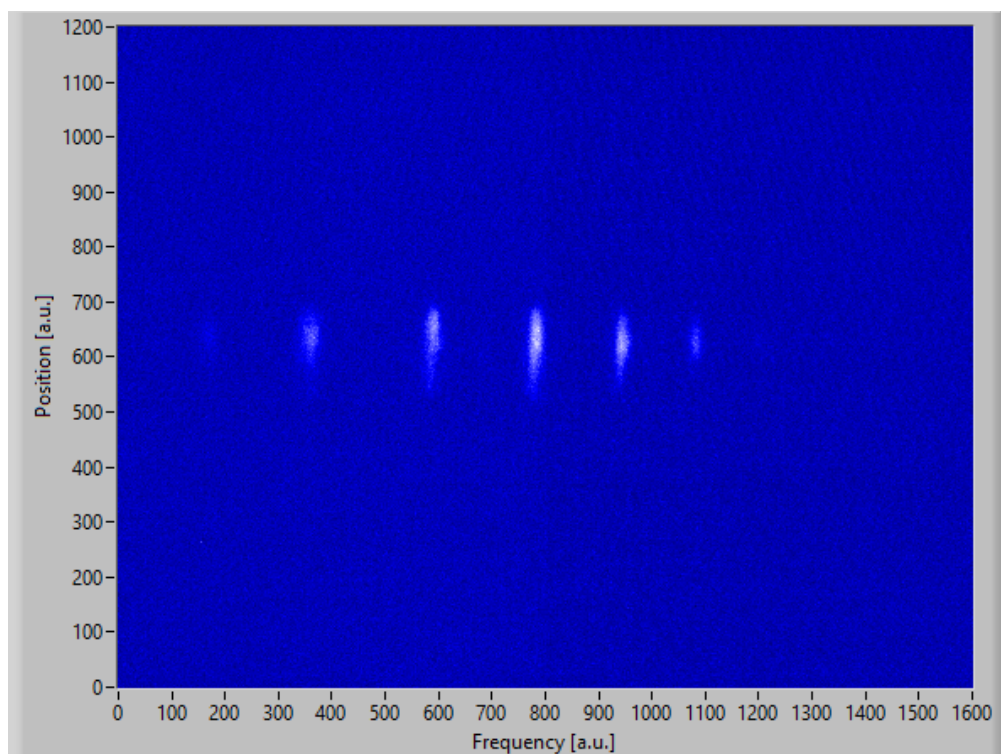


Figure 42: 0.75 W input with 3.3 mm shift.

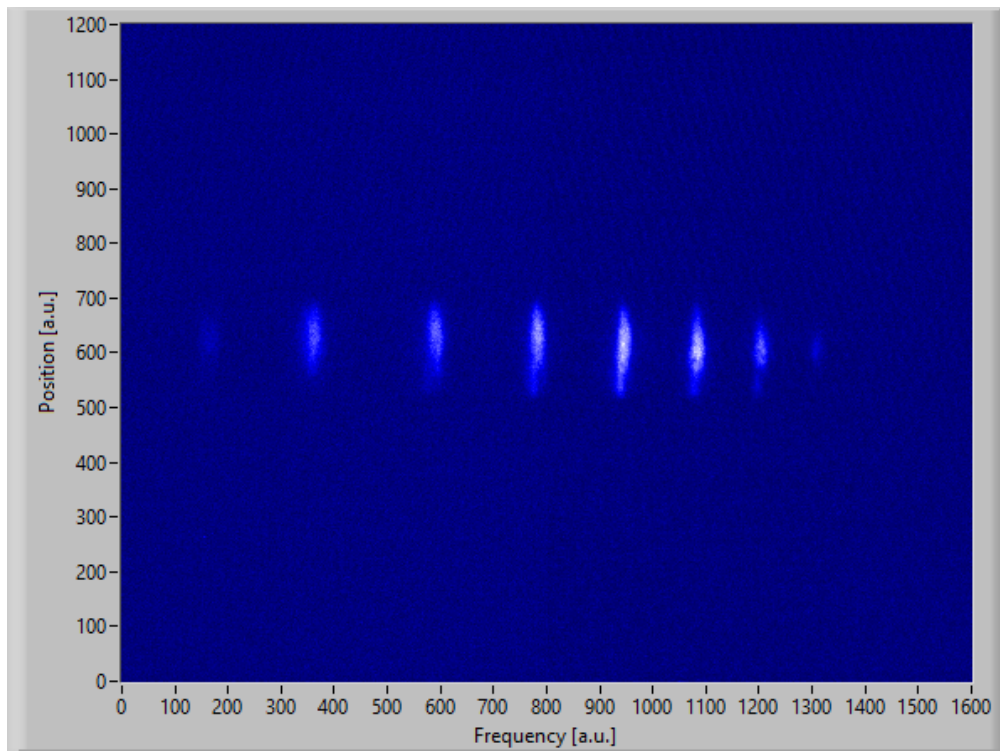


Figure 43: 0.75 W input with 3.5 mm shift.

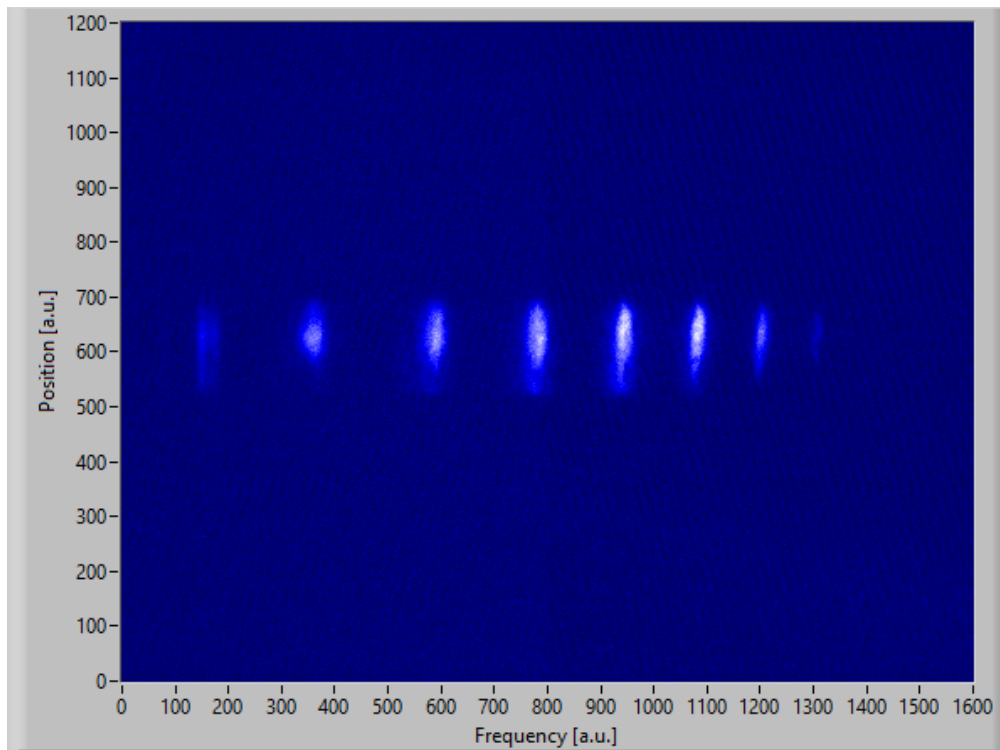


Figure 44: 0.75 W input with 3.7 mm shift.

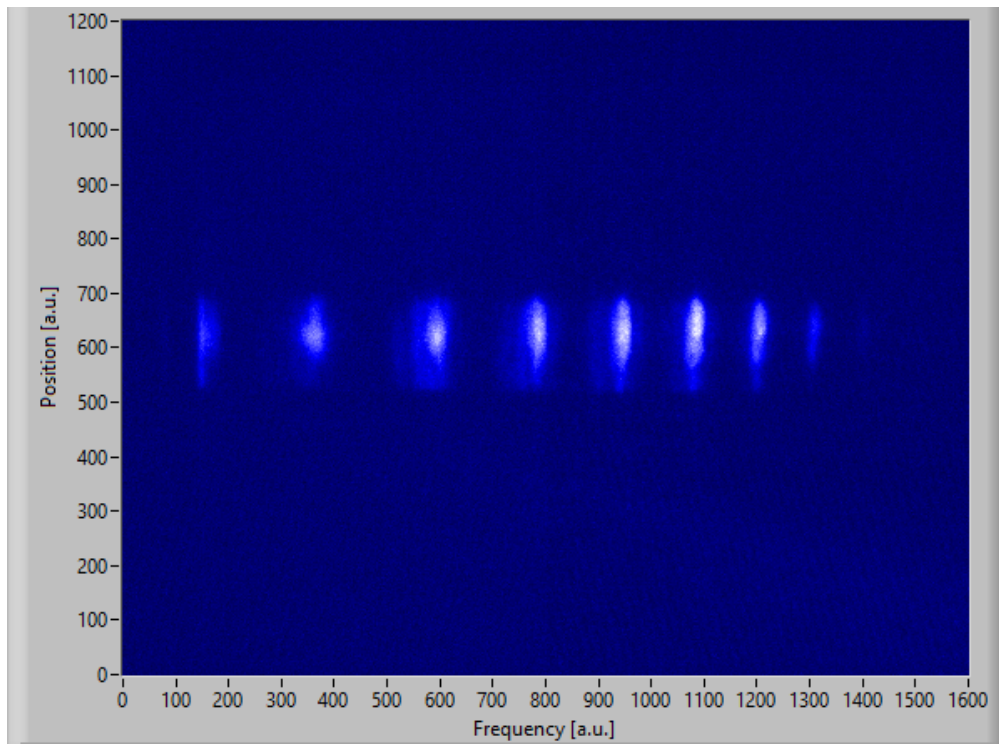


Figure 45: 0.75 W input with 3.9 mm shift.