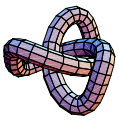
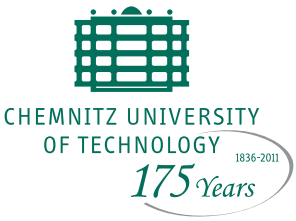




Regularization of an autoconvolution problem occurring in measurements of ultra-short laser pulses

Diploma thesis

submitted to **Faculty of Mathematics**

at **Chemnitz University of Technology**

in cooperation with **Max-Born-Institute**

for Nonlinear Optics and Short Pulse Spectroscopy, Berlin

in fulfillment of the requirements

for the degree of

Dipl.-Math. techn.

By: Daniel Gerth

Supervisors: Prof. Dr. Bernd Hofmann

Dr. Günter Steinmeyer

Chemnitz, September 9, 2011

CONTENTS

LIST OF FIGURES

LIST OF ALGORITHMS

CHAPTER

ONE

INTRODUCTION

Table of contents

1.1 Motivation

As children grow up, they are asked a lot of times what they want to become after finishing school. I remember answering with jobs like pilot, train driver or engineer. In secondary school I took part in an advanced course in physics, where among other subjects we dealt with lasers. I was fascinated by the physical background and the huge variety of applications, so at that time I decided to focus on lasers as my future profession. I cannot recall the exact reasons, but I lost track of that and started studying *Technomathematik*, which can be briefly described as mathematical studies with strong focus on application engineering and industry. During a five month-long stay in Finland I discovered my special interest in a mathematical field called Inverse Problems, and being close to the end of my study time, decided to write my thesis in this subject. Therefore I joined Prof. Hofmann's group, who among others provided the topic described here. Because of the connection between both my preferred mathematical and physical field and the possibility to solve a real problem occurring in current research, the decision to accept it as the final work of my studies was rather easy.

The problem is provided by the Max-Born-Institute for nonlinear optics and short pulse spectroscopy, Berlin. Introducing a new method to measure ultra-short laser pulses, it has been discovered that only an autoconvolved signal of the desired pulse is measurable. The task of this work is the derivation of the convolution equation from the physical processes, its analysis and finally the reconstruction (deconvolution) of the unknown laser pulse.

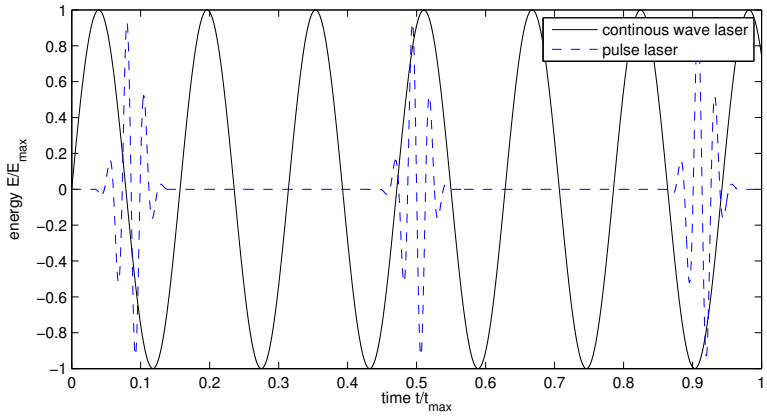
1.2 Background

Inverse Problems is a relatively young branch of mathematics which started to develop in the second half of the 20th century after the publications of A. Tikhonov (cf [34], [35]) and is closely connected to applied problems and the rapidly accelerating processing power of computers. This allows to simulate more and more complex problems. The classical problem in numerical mathematics is to determine the effect of a given causality by numerical simulation of the process. Given a mathematical model that approximates the real world problem sufficiently, the effect can be calculated whenever all necessary parameters affecting it are known precisely enough. However, sometimes there is not enough information on at least one parameter. In this context, a parameter could be a real or complex valued number, but in most cases, including this work, a function describing a physical quantity is meant. In this case the connection between cause and effect can be utilized backwards. By measuring the effect quantity, it is often possible to identify the missing parameters. Therefore those situations are called *identification problems* and form the first class of Inverse Problems which can be roughly outlined by determining the cause from the effect. Unfortunately, in most situations this is much more difficult than the direct or forward problem. Inevitably all measured data carry small errors which may lead to huge errors in the estimated cause. But this is only one aspect of an ill-posed problem (see section 2.4). To overcome these additional difficulties special techniques for the inversion of a direct problem are necessary.
For more information and examples see for example the books of Hofmann [15, chapter 2], Engl et al. [7], Groetsch [12] or Scherzer [31].

'Laser' is an acronym for *Light Amplification by Stimulated Emission of Radiation*. In other words, a laser is a device that emits electromagnetic radiation using the principle of stimulated emission of photons (see for example [28, chapter 6]). The first laser was built in 1960 by Theodore Maiman [25]. The light amplification was realized in a ruby crystal (aluminium oxide with chromium, $Al_2O_3 : Cr$) resulting in red light pulses of 694 nanometres wavelength. In opposition to continuous wave lasers, which were developed shortly after, pulse lasers emit light in short discrete time intervals, see Fig. 1.1.

Since 1960, lasers have taken a rapid development. Many other materials have been discovered to induce the stimulated emission, resulting in a huge variety of possible

Figure 1.1: Example of a pulse laser signal and a continuous wave laser signal



wavelengths. While the energy generated by lasers has increased, the length of laser pulses has decreased as shown in Fig. 1.2.

A laser pulse is described by the time-dependent electric field $E(t)$ it generates. Its complex valued Fourier-transformed signal $\varepsilon(\omega)$ can be separated into absolute values $|\varepsilon(\omega)|$ and phase $\varphi(\omega)$ with $\varepsilon(\omega) = |\varepsilon(\omega)|e^{i\varphi(\omega)}$, see Fig. 1.3. The phase consists of a constant part φ_0 and the integral over the so called group delay $GD(\omega)$,

$$\varphi(\omega) = \varphi_0 + \int_{-\infty}^{\omega} GD(\hat{\omega})d\hat{\omega}. \tag{1.1}$$

In other words, the group delay is the derivative of the phase,

$$GD(\omega) = \frac{d\varphi(\omega)}{d\omega}. \tag{1.2}$$

The particular interest in a laser pulse lies in its group delay, because it is the important part for the pulse duration and its shape. The constant offset can often be neglected. We will later see that this solves some issues occurring from the properties of the problem.

Meanwhile lasers have found many applications in medicine, industry and everyday life. Because of the huge variety of lasers, only few examples can be given here. They are a core part of every CD, DVD or Blue-Ray drive; they are used in eye surgeries or to remove tattoos and to measure distances or times. They can have enough energy to cut steel, and perhaps even enough to start a nuclear fusion.

Figure 1.2: Development of pulse durations

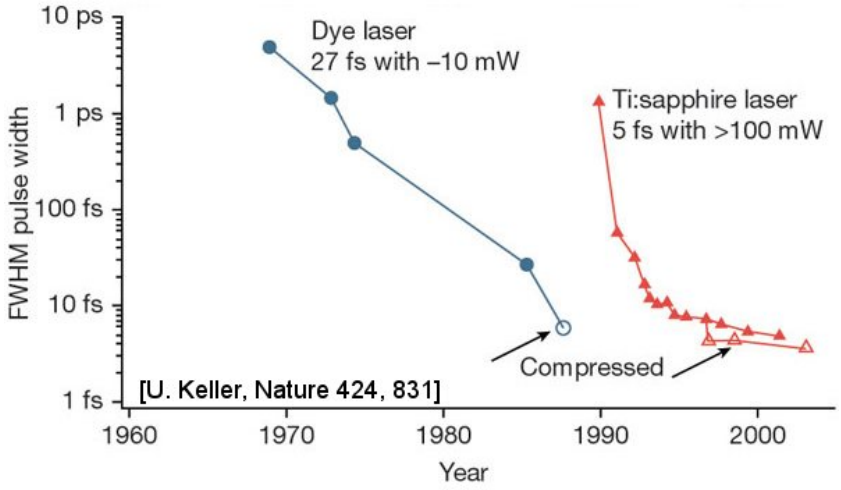
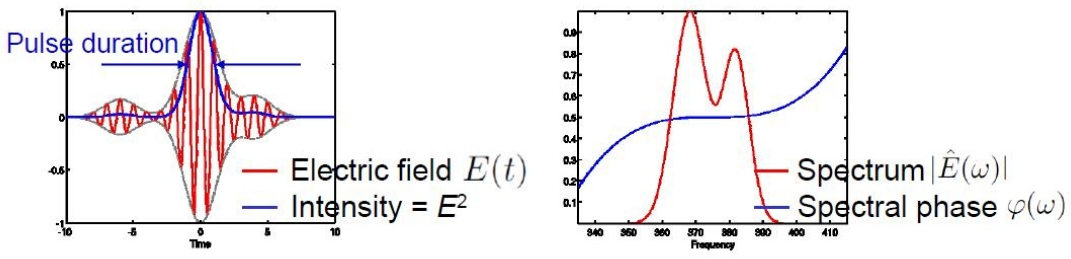


Figure 1.3: Laser pulse representation in time (left) and Fourier domain (right)



1.3 Introduction to nonlinear optics

To understand the physical processes involved in this work, a closer look on nonlinear optics is necessary. Therefore a short summary of chapter 1.1 of the book of Boyd [2] is presented in this section.
The polarization $P(t)$ describes the relative tendency of a charge distribution (e.g. electrons and protons in an atom) to be distorted from its normal shape by an external electric field $E(t)$. A light field interacting with a molecule or atom causes a displacement of the charge carriers in the system. In the simplest case of a single electron bound to a core, the displacement of the electron by the light field creates a dipole moment which is measurable as polarization. Assuming that the applied light field is much weaker than the field between the electron and the core, the polarization is proportional to the light field. This is the case in linear optics. If the light field is of the same magnitude as the electron-core-field (which could only be achieved since the invention of lasers) the polarization is no longer proportional, but still a light field results from the polarization. According to Boyd, 'the reason why the polarization plays a key role in the description of nonlinear optical phenomena is that a time-varying polarization can act as the source of new components of the electromagnetic field. For example, the wave equation in nonlinear optical media often has the form

$$\nabla^2 \tilde{E} - \frac{n^2}{c^2}\frac{\partial^2 \tilde{E}}{\partial t^2} = \frac{4\pi}{c^2}\frac{\partial^2 \tilde{P}^{NL}}{\partial t^2}. \tag{1.3}$$

This equation expresses the fact that, whenever $\frac{\partial^2 \tilde{P}^{NL}}{\partial t^2}$ is nonzero, charges are being accelerated, and according to Larmor's theorem from electromagnetism accelerated charges generate electromagnetic radiation.'[1] This resulting field then also contains powers of the frequencies of irradiating field.
Within this section, t denotes a time dependency. For a laser beam with the (angular) frequency ω ($\omega = 2\pi f$ where f is the real frequency) and a wave vector $\vec{k}$ the electric field will be represented as

$$\tilde{E}(t) = \mathcal{E} e^{-i\omega t + i\vec{k}\vec{r}} + c.c.$$

where the tilde indicates a quantity rapidly varying in time, $\mathcal{E}$ is a complex vector describing the amplitude, phase and polarization of the beam. The abbreviation c.c. stands for the complex conjugated term (as a simple example, $a = b + c.c.$ means $a = b + \bar{b}$). Finally, $\vec{r} = (\xi, \eta, \zeta)^T$ is a vector describing spatial properties. By applying Taylor's formula around $E(t) = 0$ the polarization $\tilde{P}(t)$ can be expressed as a power series in the field strength $E(t)$,

$$\tilde{P}(t) = \epsilon_0[\chi^{(1)}\tilde{E}(t) + \chi^{(2)}\tilde{E}^2(t) + \chi^{(3)}\tilde{E}^3(t) + \dots]. \tag{1.4}$$

Here $\chi^{(1)}$ is the linear susceptibility and $\chi^{(2)}$, $\chi^{(3)}$ are the second- and third-order nonlinear susceptibilities, respectively. If the polarization in (1.4) only depends on the $\chi^{(1)}$-term, the material is called linear. Otherwise it is called nonlinear.

[1] [2, p.3f]

Definition 1.1:
A medium is called $\chi^{(n)}$-medium, if the n-th term in (1.4) dominates and the other terms can be neglected.

A second effect necessary for understanding the main topic is the *optical Kerr effect* (cf, for example [30]). The refracting index of a medium is not constant, but varies slowly depending on the intensity of the applied electric field. For centrosymmetric materials (in terms of crystallogy, centrosymmetry 'refers to a space group which contains an inversion center as one of its symmetry elements. In such a space group, for every point (x,y,z) in the unit cell there is an indistinguishable point (-x,-y,-z)'[2]), the refractive index writes as

$$n(E) = n_0 + n_2 I(\omega) = n_0 + n_2 |E(\omega)|^2. \tag{1.5}$$

Thus each frequency of a laser pulse is refracted differently.

For some reasons, one is interested in the detailed shape of ultrashort laser pulses. Knowing a pulse helps to adjust the laser to create even shorter pulses with applications in medicine, material processing and to study events acting on short timescales. The shorter the pulse is, the shorter the observable period of time becomes. Usually a gaussian pulse shape is desired to avoid local maxima which perturb spectroscopic measurements. One is also interested in the particular shape of a laser pulse to enhance established laser systems. Changing one component of the system and then determining the pulse again shows precisely the effect of that component.

The fundamental limit for measuring pulses is electronics, as they can only be used to record events in the order of picoseconds (10^{-12} seconds). Since this limit cannot be overcome, techniques are necessary which work on the pulses themselves. They key in all approaches is the application of a nonlinear material.

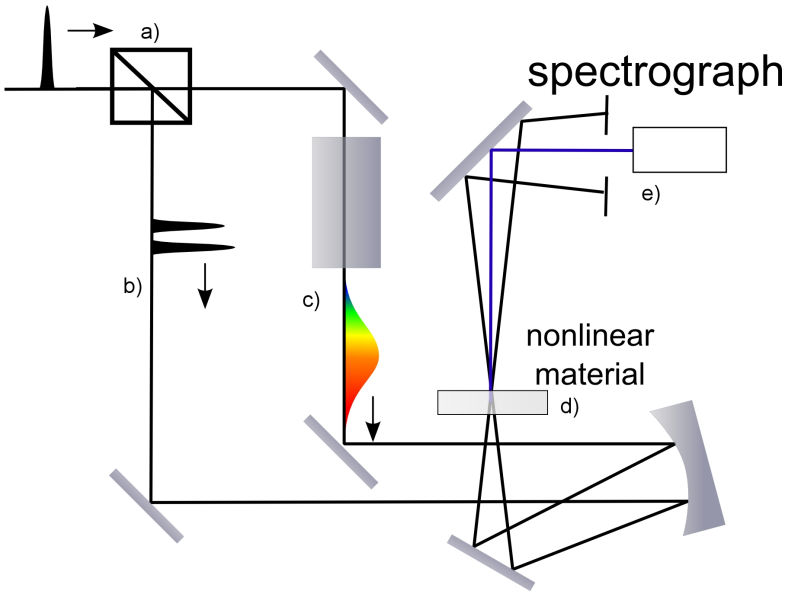
Several possible solutions like *autocorrelation* [6] and *FROG* [37] have been found which sample the pulse by itself, convert time to space and make use of the frequency doubling. Neither of them were totally satisfying. A third way, called *SPIDER*, will be the basic idea for the measurement of the pulses. For a review of all these methods and additional references see [32].

The abbreviation *SPIDER* stands for **S**pectral **P**hase **I**nterferometry for **D**irect **E**lectric-field **R**econstruction and was proposed by I. A. Walmsley and C. Iaconis [19], [42].
A possible realization of SPIDER is shown in Fig. 1.4.

The unknown ('fundamental') pulse is split into two beams in a). In the first beam b) it is duplicated with a time delay, while in the second beam c) it is strongly chirped which means, in simple words, that it is stretched so its pulse duration is increased significantly. This effect is caused by the frequency-dependent optical density of a medium. Therefore

[2] http://en.wikipedia.org/wiki/Centrosymmetry

Figure 1.4: possible SPIDER setup



the particular frequency components have a different speed within the medium. It is called *continuous-wave* (cw) *beam*, because it is so much longer compared to the fundamental pulse that relatively to its short pulse duration the fundamental pulse 'sees' more or less only one frequency of the chirped beam. Both beams b) and c) are then combined in d) in a $\chi^{(2)}$ medium. In e) a spectrograph finally measures the intensity of the resulting signal at certain frequencies.
But *SPIDER* has flaws as it can only measure a relative group delay. Because of the processes in a $\chi^{(2)}$-medium [2], a frequency doubling of the SPIDER-signal with respect to the fundamental pulse occurs. Therefore, like all $\chi^{(2)}$-based ideas, it is limited to visible and the near-infrared spectral range. This follows, because transparent media (media which let relevant frequencies pass) start to absorb at shorter wavelengths (200-300nm). A frequency doubling (or bisection of the wavelength) of 400 nm is therefore critically.
This limit is relaxed for $\chi^{(3)}$-based characterization methods, where a four wave mixing process takes place. This has been applied in *SD-FROG* [37] already and will now be explained for SD-SPIDER.

1.4 SD-SPIDER

SD-SPIDER (SD standing for self diffraction) has been introduced by the research group 'Solid State Light Sources' led by Dr. Günter Steinmeyer as subdivision of division C 'Nonlinear Processes in Condensed Matter' at Max-Born-Institute for Nonlinear Optics and Short Pulse Spectroscopy, Berlin. The approach was presented at the conference CLEO 2010 [29].

The basic setup of *SD-SPIDER* is similar to conventional *SPIDER* (see Fig. 1.4), but instead of a $\chi^{(2)}$-medium a $\chi^{(3)}$-medium, namely Bariumfluoride BaF_2 throughout this work, is used. This leads to a fundamental change in the behaviour of the nonlinear gating process. An electric field consisting of the three electric fields with one frequency component each (in particular the laser beams incident in the $\chi^{(3)}$ medium) can be written as

$$\tilde{E}_{(3)}(t) = E_1 e^{-i\omega_1 t} + E_2 e^{-i\omega_2 t} + E_3 e^{-i\omega_3 t} + c.c. \tag{1.6}$$

The resulting polarization (see (1.4))

$$\tilde{P}(t) = \chi^{(3)} \tilde{E}^3_{(3)} \tag{1.7}$$

then consists of 16 different combinations of the original frequencies. For the full list and associated complex amplitudes see [2, p.12].

Because the two laser beams from Fig. 1.4 are not interchangeable, two possible implementations of SD-SPIDER exist. In the first one, two photons from the chirped beam (c) are mixed with one replica photon. This version is called *upshift* SD-SPIDER, referring to the sign of the frequency shift induced by the chirped beam in the four-wave mixing process as can bee seen in the bottom part of Fig. 1.6. This figure shows the k-vector diagram of both implementations and illustrates the phase matching. Upshift SD-SPIDER allows analytic retrieval of the pulse shape and was successfully demonstrated in [29]. However, the obtained signal strength is quite limited.
The second implementation of SD-SPIDER, referred to as *downshift* SD-SPIDER, allows higher signal strength. There, two photons from the replica beam (b) are mixed with one photon from the chirped beam (see top part of Fig. 1.6). We then have

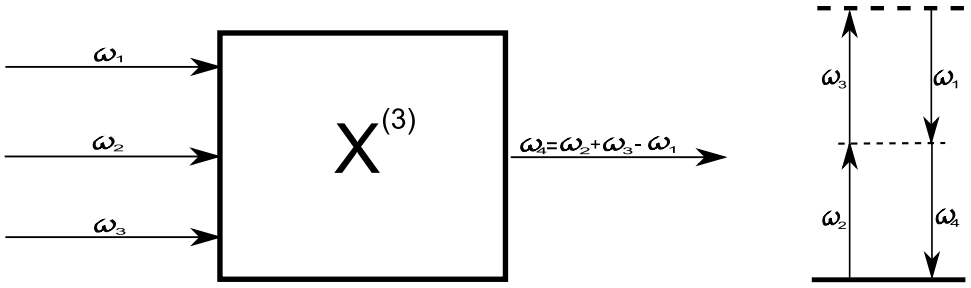
$$\omega_4 = \omega_2 + \omega_3 - \omega_1, \tag{1.8}$$

where $\omega_2 = \omega_3 =: \omega$ is the frequency of the replica pulses and $\omega_1 =: \omega_{cw}$ is the frequency of the chirped, quasi continuous-wave beam. Therefore, as shown in Fig. 1.7 the frequencies ω are shifted back by ω_{cw}. This can also be interpreted as the law of energy conservation. One photon of frequency ω_2 and one photon of frequency ω_3 are destroyed and two photons of frequency ω_1 and ω_4 are created as illustrated in Fig. 1.5.

Including the complex amplitudes we derive the polarization

$$\tilde{P}_{SD}(t) = 6\chi^{(3)} E_1 E_2 \overline{E_3} e^{-i(\omega_2+\omega_3-\omega_1)t} \tag{1.9}$$

Figure 1.5: Law of energy conservation in the four-wave mixing process



from plugging (1.6) into (1.7).

For the description of the self-diffraction effect, the particular frequency components of the pulses are considered as plane waves. If two plane waves from each pulse respectively interfere, a periodical pattern of intensity minima and maxima is formed. Because of the optical Kerr-effect (1.5) this pattern generates a periodical refractive index (refractive index grating) within the material. Another wave of one of the pulses is refracted at this index, resulting in the SD-signal. Because photons of every frequency contained in the laser pulse create a grating with one photon from the cw-pulse and at each of this resulting gratings photons of every frequency contained in the laser pulse are refracted, the SD-signal is the convolution of the fundamental pulse. In other words the SD-signal is the sum of all combinations $E_1(\omega_2)E_2(\omega_2 - \omega_3 + \omega_{cw})e^{i\omega_{\mathrm{cw}}t}$. Writing this as an integral, one can identify the autoconvolution effect. Since the fundamental pulse is shorter and thus much stronger than the cw-pulse, it is the one that is refracted primarily.
The measurements themselves are made by a spectrometer measuring the Fourier-transformed intensity of the electric field for certain frequencies using a CCD-camera. The phase of the complex signal is then retrieved via the Takeda algorithm [33].

Because deconvolution is nontrivial, downshift SD-SPIDER has only been proposed in theory yet, but needs to be demonstrated. It will be introduced throughout this work and associated papers.

1.5 Equation

The derivation of the convolution equation starts from the nonlinear wave equation

$$\nabla \times (\nabla \times E) + \frac{n^2}{c^2}\partial_t^2 E = -\mu_0 \partial_t^2 P_{\mathrm{NL}}(E), \tag{1.10}$$

Figure 1.6: k-vector diagram showing phase matching of the four-wave mixing process

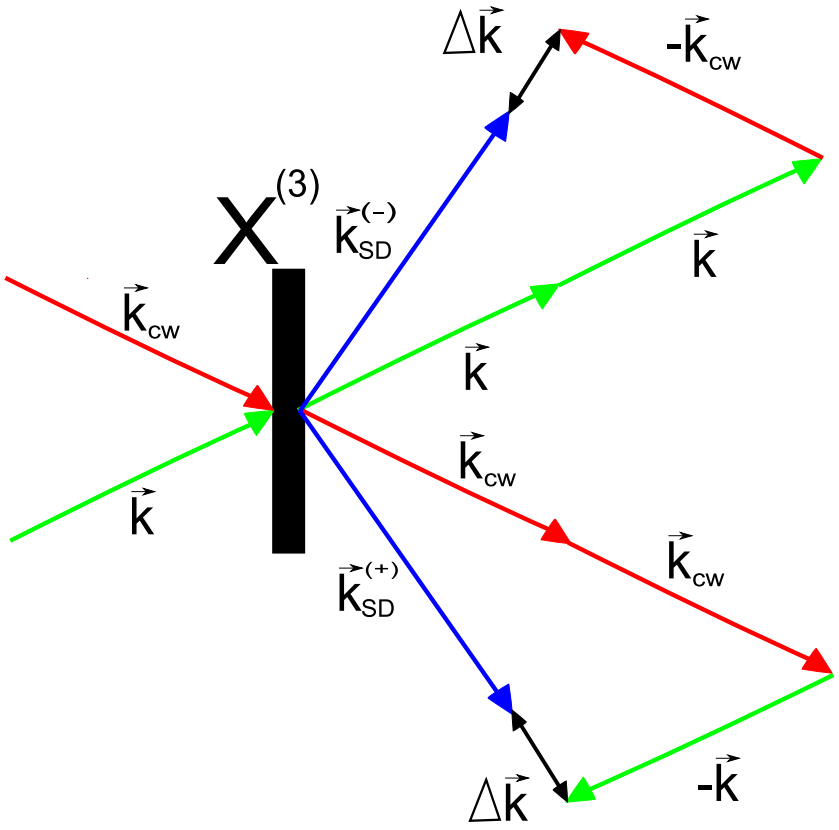
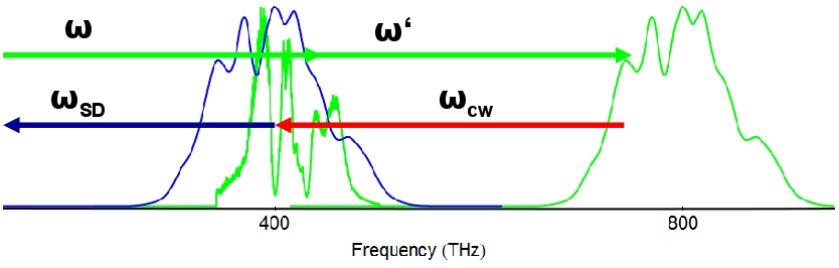


Figure 1.7: Frequency shift of SD-SPIDER signal with respect to the fundamental pulse



where P_{NL} is the nonlinear polarization and $E = \sum_{j=1}^{3} E_j$ the sum of all three electric fields inducing the SD-wave in the four-wave mixing process. Applying standard approximations (isotropic non-absorbing media, undepleted pump fields, loose focusing, plane wavefronts, slowly varying envelope approximation), we arrive at an equation for the generated electric field

$$E_4(\omega) = \frac{\mu_0 c\omega}{2n} \int_0^L \hat{P}_{\mathrm{NL}}(z,\omega) e^{-i\vec{k}_4\vec{r}} d\zeta, \tag{1.11}$$

where c is the speed of light, μ_0 the vacuum permeability (for for numerical values see section 3.2), $n = n(\omega)$ the refractive index and L is the length of the BaF_2-crystal. In our case, the polarization $\hat{P}_{\mathrm{NL}}$ reads

$$\begin{aligned}&\hat{P}^{(3)}(r,\omega)\\ &= \int\int\int \chi^{(3)}(\omega,\omega_1,\omega_2,\omega_3) 6 \overline{\mathcal{E}}(r,\omega_1)\mathcal{E}(r,\omega_2)\mathcal{E}(r,\omega_3)\delta(\omega-\omega_1-\omega_2-\omega_3)d\omega_1 d\omega_2 d\omega_3.\end{aligned} \tag{1.12}$$

The electric fields to be considered are
first cw-beam as beam 1

$$E(r,\omega_1) \simeq \mathcal{E}^{\mathrm{cw}}(\omega_1)\delta(\omega-\omega_{\mathrm{cw}})e^{i\vec{k}_{\mathrm{cw}}\vec{r}} + \overline{\mathcal{E}^{\mathrm{cw}}}(\omega_1)\delta(\omega+\omega_{\mathrm{cw}})e^{-i\vec{k}_{\mathrm{cw}}\vec{r}}, \tag{1.13}$$

second fundamental pulses as beam 2 and 3

$$E_p(r,\omega_j) = \begin{cases} \mathcal{E}_p(\omega_j)e^{i\vec{k}_p\vec{r}} & \omega > 0 \\ \overline{\mathcal{E}_p}(\omega_j)e^{-i\vec{k}_p\vec{r}} & \omega < 0 \end{cases}, j = 2,3. \tag{1.14}$$

Now only those waves propagating in the direction of the conventional (downshift) SD are of interest, such that only the special polarization given by

$$\hat{P}^{(3)}(r,\omega) \propto \begin{cases} \int_0^\infty\int_0^\infty \chi^{(3)}\overline{\mathcal{E}^{\mathrm{cw}}}\mathcal{E}_p(\omega_2)\mathcal{E}_p(\omega_3)\delta(\omega+\omega_{\mathrm{cw}}-\omega_2-\omega_3) & \\ \qquad e^{i(-\vec{k}_{\mathrm{cw}}+\vec{k}_p+\vec{k}_p)\vec{r}} d\omega_2 d\omega_3 & \omega > 0 \\ \int_0^\infty\int_0^\infty \chi^{(3)}\mathcal{E}^{\mathrm{cw}}\overline{\mathcal{E}_p}(\omega_2)\overline{\mathcal{E}_p}(\omega_3)\delta(-|\omega|-\omega_{\mathrm{cw}}+\omega_2+\omega_3) & \\ \qquad e^{i(\vec{k}_{\mathrm{cw}}-\vec{k}_p-\vec{k}_p)\vec{r}} d\omega_2 d\omega_3 & \omega < 0 \end{cases} \tag{1.15}$$

is to be considered. Both integrals start at zero since in either process the two photons out of the pulse spectrum have to enter with the same sign. By simplification we get

$$\hat{P}^{(3)}(r,\omega) \propto \begin{cases} \int_0^\infty \chi^{(3)}\overline{\mathcal{E}^{\mathrm{cw}}}\mathcal{E}_p(\omega_2)\mathcal{E}_p(\Delta\omega)e^{i(-\vec{k}_{\mathrm{cw}}+\vec{k}_p+\vec{k}_p)\vec{r}}\Theta(\Delta\omega)d\omega_2 & \omega > 0 \\ \int_0^\infty \chi^{(3)}\mathcal{E}^{\mathrm{cw}}\overline{\mathcal{E}_p}(\omega_2)\overline{\mathcal{E}_p}(\Delta\omega)e^{i(\vec{k}_{\mathrm{cw}}-\vec{k}_p-\vec{k}_p)\vec{r}}\Theta(\Delta\omega)d\omega_2 & \omega < 0 \end{cases}, \tag{1.16}$$

with $\Delta\omega = \Delta\omega(\omega, \omega_{\mathrm{cw}}, \omega_2) = \omega + \omega_{\mathrm{cw}} - \omega_2$ and

$$\Theta(\omega) = \begin{cases} 1 & \omega \geq 0 \\ 0 & \omega < 0 \end{cases} \tag{1.17}$$

the Heaviside step-function guaranteeing that no argument with negative sign/photon with negative frequency enters. We thus have

$$\hat{P}^{(3)}(r,\omega) \propto \begin{cases} \int_0^{\omega+\omega_{\mathrm{cw}}} \chi^{(3)} \overline{\mathcal{E}^{\mathrm{cw}}} \mathcal{E}_p(\omega_2) \mathcal{E}_p(\Delta\omega) e^{i(-\vec{k}_{\mathrm{cw}}+\vec{k}_p+\vec{k}_p)\vec{r}} d\omega_2 & \omega > 0 \\ \overline{\hat{P}^{(3)}(r,|\omega|)} & \omega < 0. \end{cases} \tag{1.18}$$

This makes sense since it has the same structure as required for the spectrum in order to generate the real-valued field in the time domain. Plugging 1.18 into (1.11) we get

$$E_{\mathrm{SD}}(L,\omega) = \frac{\mu_0 c\omega}{2n} \int_0^L \int_0^{\omega+\omega_{\mathrm{cw}}} \chi^{(3)} \overline{\mathcal{E}^{\mathrm{cw}}} \mathcal{E}_p(\omega_2) \mathcal{E}_p(\Delta\omega) e^{i(-\vec{k}_{\mathrm{cw}}+\vec{k}_p+\vec{k}_p)\vec{r}} d\omega_2 e^{-i\vec{k}_{\mathrm{SD}}\vec{r}} d\zeta \tag{1.19}$$

Applying Fubini's law[3] we have

$$E_{\mathrm{SD}}(L,\omega) = \frac{\mu_0 c\omega}{2n} \int_0^{\omega+\omega_{\mathrm{cw}}} \chi^{(3)} \overline{\mathcal{E}^{\mathrm{cw}}} \mathcal{E}_p(\omega_2) \mathcal{E}_p(\Delta\omega) \int_0^L e^{i(-\vec{k}_{\mathrm{cw}}+\vec{k}_p+\vec{k}_p-\vec{k}_{\mathrm{SD}})\vec{r}} d\zeta d\omega_2. \tag{1.20}$$

With $\vec{\Delta k} = -\vec{k}_{\mathrm{cw}} + \vec{k}_p + \vec{k}_p - \vec{k}_{\mathrm{SD}}$ and $\vec{r} = (\xi, \eta, \zeta)^T$ we get

$$E_{\mathrm{SD}}(L,\omega) = \frac{\mu_0 c\omega}{2n} \int_0^{\omega+\omega_{\mathrm{cw}}} \chi^{(3)} \overline{\mathcal{E}^{\mathrm{cw}}} \mathcal{E}_p(\omega_2) \mathcal{E}_p(\Delta\omega) e^{i\vec{\Delta k}_\xi \xi + \vec{\Delta k}_\eta \eta} \int_0^L e^{i\vec{\Delta k}_\zeta \zeta} d\zeta d\omega_2. \tag{1.21}$$

It is

$$\begin{aligned} \int_0^L e^{i\vec{\Delta k}_\zeta \zeta} d\zeta &= \left[\frac{1}{\vec{\Delta k}_\zeta} e^{i\vec{\Delta k}_\zeta \zeta} \right]_0^L \\ &= \frac{1}{\vec{\Delta k}_\zeta} (\sin(\vec{\Delta k}_\zeta L) - \sin(0) + i(-\cos(\vec{\Delta k}_\zeta L) + \cos(0))) \\ &= \frac{1}{\vec{\Delta k}_\zeta} (2\sin(\vec{\Delta k}_\zeta \frac{L}{2}) \cos(\vec{\Delta k}_\zeta \frac{L}{2}) + 2i \sin^2(\vec{\Delta k}_\zeta \frac{L}{2})) \\ &= \frac{2}{\vec{\Delta k}_\zeta} e^{i\vec{\Delta k}_\zeta \frac{L}{2}} \sin(\vec{\Delta k}_\zeta \frac{L}{2}) = L e^{i\vec{\Delta k}_\zeta \frac{L}{2}} \frac{\sin(\vec{\Delta k}_\zeta \frac{L}{2})}{\vec{\Delta k}_\zeta \frac{L}{2}}. \end{aligned} \tag{1.22}$$

[3] we will later see that the integrand is continuous

Defining $sinc(z) := \frac{sin(z)}{z}$ and plugging (1.22) into (1.21) we arrive at an equation

$$E_{\mathrm{SD}}(\omega) = \begin{cases} \int_0^{\omega+\omega_{cw}} K(\omega,\hat{\omega})\mathcal{E}_p(\hat{\omega})\mathcal{E}_p(\omega+\omega_{cw}-\hat{\omega})d\hat{\omega} & \omega > 0 \\ \overline{\hat{E}_4(L,|\omega|)} & \omega < 0 \end{cases}, \tag{1.23}$$

with the kernel

$$K(\omega,\hat{\omega}) = \frac{\mu_0 cL}{2}\frac{\omega}{n(\omega)}\chi^{(3)}(\omega,-\omega_{cw},\hat{\omega},\omega+\omega_{cw}-\hat{\omega})\overline{\mathcal{E}}^{cw}e^{i(\Delta\vec{k}_\xi\xi+\Delta\vec{k}_\eta\eta+\Delta\vec{k}_\zeta\frac{L}{2})}sinc(\Delta\vec{k}_\zeta\frac{L}{2}). \tag{1.24}$$

To improve readability we set $y(\omega) := E_{\mathrm{SD}}(\omega)$ and $x(\hat{\omega}) := \mathcal{E}_p(\hat{\omega})$. Both y and x are complex functions which are composed of absolute values $|\cdot|$ and phase $\varphi_\cdot$ so that we have $y(\omega) = |y(\omega)|e^{i\varphi_y(\omega)}$ and $x(\hat{\omega}) = |x(\hat{\omega})|e^{i\varphi_x(\hat{\omega})}$. Furthermore it is assumed that x is nonzero only between a lower boundary frequency $\hat{\omega}_l$ and an upper boundary frequency $\hat{\omega}_u$, in mathematical terms $supp(x) = [\hat{\omega}_l, \hat{\omega}_u]$. From (1.8) follows immediately $supp(y) = [2\hat{\omega}_l - \omega_{\mathrm{cw}}, 2\hat{\omega}_u - \omega_{\mathrm{cw}}]$.
Until now, the frequency dependency of the k-vectors has not been mentioned. In fact, we have

$$\Delta\vec{k}(\omega,\hat{\omega},\omega_{\mathrm{cw}}) = -\vec{k}_{cw}(\omega_{cw}) + \vec{k}_p(\hat{\omega}) + \vec{k}_p(\omega+\omega_{cw}-\hat{\omega}) - \vec{k}_{SD}(\omega,\omega_{cw},\hat{\omega}). \tag{1.25}$$

If the pulses and the cw-wave are incident with a crossing angle 2α, the direction of their k-vectors is

$$\frac{\vec{k}_{cw}}{||k_{cw}||_2} = (\sin\alpha, 0, \cos\alpha), \tag{1.26}$$

$$\frac{\vec{k}_p}{||k_p||_2} = (-\sin\alpha, 0, \cos\alpha) \tag{1.27}$$

and their euclidean norm is

$$||\vec{k}_{...}(\tilde{\omega})||_2 = \frac{n(\tilde{\omega})\tilde{\omega}}{c}. \tag{1.28}$$

The k-vector of the self-diffracted wave is

$$\vec{k}_{SD} = \frac{n(\omega)\omega}{c}\frac{-k_{cw}(\omega_{cw}) + k_p(\hat{\omega}) + k_p(\omega+\omega_{cw}-\hat{\omega})}{||-k_{cw}(\omega_{cw}) + k_p(\hat{\omega}) + k_p(\omega+\omega_{cw}-\hat{\omega})||_2}. \tag{1.29}$$

The refractive index is calculated from

$$n^2(\tilde{\omega}) = 1 + \chi^{(1)}\tilde{\omega}. \tag{1.30}$$

The Sellmeier equation (cf [30]) then gives

$$n^2(\tilde{\omega}) = 1 + \frac{B_1\lambda^2(\tilde{\omega})}{\lambda^2(\tilde{\omega}) - C_1} + \frac{B_2\lambda^2(\tilde{\omega})}{\lambda^2(\tilde{\omega}) - C_2} + \frac{B_3\lambda^2(\tilde{\omega})}{\lambda^2(\tilde{\omega}) - C_3}, \tag{1.31}$$

Table 1.1: Sellmeier coefficients for BaF_2

B_1	B_2	B_3	C_1	C_2	C_3
0.6434	0.5058	3.8261	0.0578	0.1097	46.3864

with the wavelength $\lambda = \frac{c}{f}$ corresponding to $\mu m = 10^{-6}m$. Therefore, for given ω

$$\lambda(\omega) = \frac{c}{\frac{\omega}{2\pi}} \frac{1}{\mu m} = \frac{2\pi c 10^{-6} m}{\omega}. \tag{1.32}$$

The Sellmeier coefficients B_j and C_j are experimentally determined parameters. For the BaF_2 crystal used in the experiment see table 1.1 (measured by Malitson [26]).
The parameter $\overline{\mathcal{E}}^{cw}$ is a complex value which is constant throughout an experiment, but unknown and can not be measured. As it only scales the absolute values of the SD-SPIDER signal and adds a constant value to its phase, we set $\overline{\mathcal{E}}^{cw} := 1$ in this work.

According to Boyd [2], $\chi^{(3)}$ can be approximated as

$$\chi^{(3)}(\omega_4, \omega_3, \omega_2, \omega_1) = A\chi^{(1)}(\omega_4)\chi^{(1)}(\omega_3)\chi^{(1)}(\omega_2)\chi^{(1)}(\omega_1) \tag{1.33}$$

for ionic crystals where A is some proportionality constant. Throughout this work, it is assumed that $A = 1$. At least for the BaF_2 crystal this can be regarded as a good approximation since fluorine has a high electro-negativity such that it will attract an electron from the Barium. Calculating the refractive index via (1.31), we get $\chi^{(3)}$ from (1.33) using (1.30) to evaluate

$$\chi^{(1)}(\tilde{\omega}) = n^2(\tilde{\omega}) - 1. \tag{1.34}$$

For the mathematical analysis of the problem we will assume that the kernel is continuous. The following remark shows that this applies to the physical kernel.

Remark 1.2:
Using a BaF_2 crystal, the kernel $K(\omega, \hat{\omega})$ in (1.24) is continuous for $\min\left\{\frac{\hat{\omega}_l}{2\pi}, \frac{2\hat{\omega}_l - \omega_{cw}}{2\pi}\right\} > f_3^c$ with $f_3^c \approx 4.0608 \cdot 10^{13} Hz$.
It can be seen immediately that a discontinuity may only appear in the calculation of the refractive index in (1.31). Because $B_k > 0$ for $k = 1, 2, 3$, only the denominators have to be investigated as they must not be zero. Singularities occur if $\lambda^2(\omega) - C_k^2 = 0$. With λ from (1.32), the critical frequencies f_k^c are

$$f_k^c = \frac{2\pi c 10^6 m}{C_k}.$$

Evaluating this for all three terms, we get

$$f_1^c \approx 3.2589 \cdot 10^{16} Hz$$
$$f_2^c \approx 1.7171 \cdot 10^{16} Hz$$
$$f_3^c \approx 4.0608 \cdot 10^{13} Hz.$$

We will restrict the problem to the case $\min\{\frac{\hat{\omega}_l}{2\pi}, \frac{2\hat{\omega}_l - \omega_{cw}}{2\pi}\} > f_3^c$. The laser cannot produce a signal with frequencies of the order of f_1^c or f_2^c. Thus, the kernel is continuous.

For the mathematical analysis, the notation is changed to $s := \omega$, $t := \hat{\omega}$, $k(\cdot,\cdot) := K(\cdot,\cdot)$. Because of the symmetry, only the case $s = \omega > 0$ will be considered furthermore. Thus, equation (1.23) can be written as

$$\int_0^{\tilde{s}+s_{cw}} \tilde{k}(\tilde{s},\hat{t})\hat{x}(\hat{t})\hat{x}(\tilde{s}+s_{cw}-\hat{t})d\hat{t} = \tilde{y}(\tilde{s}). \tag{1.35}$$

It is $s \in [2t_l - s_{cw}, 2t_u - s_{cw}]$ and $\hat{t} \in [t_l, t_u]$. Substituting $\hat{s} := \tilde{s} + s_{cw}$ and defining $\hat{x}(\cdot) := \tilde{x}(\cdot - s_{cw})$; $\hat{k}(\cdot,\cdot) := \tilde{k}(\cdot - s_{cw}, \cdot)$ we arrive at an operator equation

$$[F(\hat{x})](\hat{s}) = \int_0^{\hat{s}} \hat{k}(\hat{s},\hat{t})\hat{x}(\hat{t})\hat{x}(\hat{s}-\hat{t})d\hat{t} = \hat{y}(\hat{s}), \tag{1.36}$$

where $\hat{s} \in [2t_l, 2t_u]$ and $\hat{t} \in [t_l, t_u]$. To improve readability, the intervals can be transformed onto $[0,2]$ and $[0,1]$ respectively by substituting $\hat{s} = s(t_u - t_l) + 2t_l$, $\hat{t} = t(t_u - t_l) + t_l$, $d\hat{t} = (t_u - t_l)dt$,

$$\int_0^{s(t_u-t_l)+2t_l} (t_u - t_l)\hat{k}(s(t_u-t_l)+2t_l, t(t_u-t_l)+t_l)\hat{x}(t(t_u-t_l)+t_l)\hat{x}(s(t_u-t_l)+2t_l-t(t_u-t_l)-t_l)dt$$
$$= \hat{y}(s(t_u - t_l) + 2t_l). \tag{1.37}$$

Finally, by defining $y(s) := \hat{y}(s(t_u - t_l) + 2t_l)$, $x(t) = \hat{x}(t(t_u - t_l) + t_l)$ and $k(s,t) := (t_u - t_l)\hat{k}(s(t_u - t_l) + 2t_l, t(t_u - t_l) + t_l)$ the convolution equation takes the form

$$\int_0^s k(s,t)v(s-t)v(t)dt = y(s). \tag{1.38}$$

where

$$x : [0,1] \subset \mathbb{R} \mapsto \mathbb{C},$$
$$y : [0,2] \subset \mathbb{R} \mapsto \mathbb{C},$$
$$k : \mathbb{R}^+ \times \mathbb{R}^+ \mapsto \mathbb{C}.$$

1.6 Measurement setup

In the beginning of this work, only measurements of the phase of the SD-pulse of $\varphi_y(s)$ and the absolute values of the fundamental pulse $|x(t)|$ were available.
Intuitively, one might think that measuring $|y(s)|$ would also be necessary so that the complete SD-signal y can be used for the reconstruction of the fundamental pulse (for the algorithm used in section 3.3 it is absolutely necessary to know the complete SD-pulse). Another aspect that emphasizes the necessity of measuring $|y(s)|$ is that in any inverse problem one is interested in having as much information as possible about the unknown objects. In fact, without the knowledge of $|y(s)|$, the unknown pulse is not even identifiable, in other words a reconstruction is not possible because the phase $\varphi_x(t)$ essentially influences $|y(s)|$. Unfortunately, this cannot be shown analytically for all phases φ_x because an integral of the form $\int_{x_0}^{x_1} e^{f(x)}dx$ generally does not have a primitive.
Instead we will use a simple example. Let $k(s,t) \equiv 1$, $y(s) = |y(s)|e^{i\varphi_y(s)}$, $x(t) = |x(t)|e^{i\varphi_x(t)}$ with $|x(t)| \equiv 1$ and a linear phase trend $\varphi_x(t) = at + b$ $(a, b \in \mathbb{R})$. Then

$$y(s) = \int_0^s k(s,t)x(t)x(s-t)\,dt \tag{1.39}$$

$$= \int_0^s e^{i(\varphi_x(y)+\varphi_x(s-t))}\,dt. \tag{1.40}$$

Switching to polar form this can be written as

$$y(s) = \int_0^s \cos((as+b)+(a(s-t)+b)) + i\sin(at+b+a(s-t)+b)\,dt \tag{1.41}$$

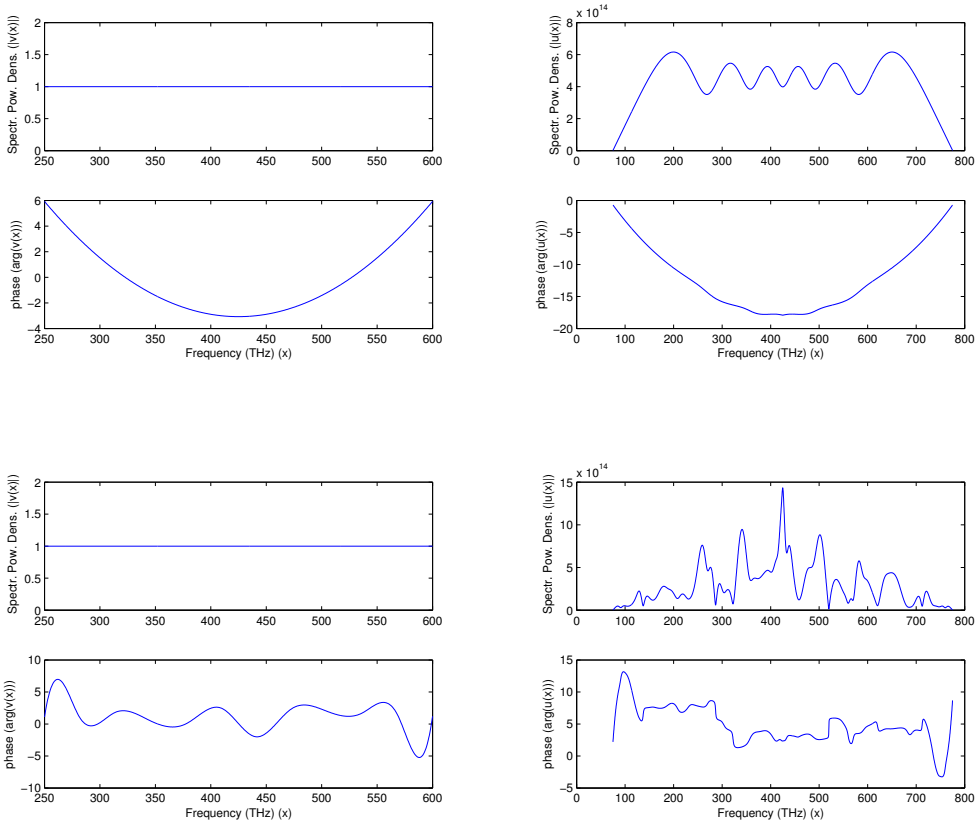
$$= \cos(as+2b)\int_0^s 1\,dt + i\sin(as+2b)\int_0^s 1\,dt. \tag{1.42}$$

Now

$$|y(s)|^2 = \cos^2(as+2b)\left(\int_0^s 1\,dt\right)^2 + \sin^2(as+2b)\left(\int_0^s 1\,dt\right)^2 = s^2. \tag{1.43}$$

Note that the sum of the squares of the integrand in (1.41) equals one for all s.
Even in its simplest form the phase of x has an influence on the absolute values of y and vice versa $|y(s)|$ carries information about the sought-after phase. In general, in case of a nonlinear phase which is to be expected from the measurements and a non-trivial kernel this influence will obviously become unpredictable, therefore it is inevitable to

Figure 1.8: Comparison between fundamental pulse and SD-pulse. Left column: fundamental pulse. Right column: autoconvolved pulse. Although the absolute values of the fundamental pulses are the same, the convolved pulses have different absolute values



measure $|y(s)|$.
To further prove this, Figure 1.8 shows two pulses x and their numerically computed autoconvolution signal, where $|x| \equiv 1$ and the kernel is set to equal one. In the upper part of a figure the absolute values are shown, in the bottom part the phase is plotted. This confirms the theoretically predicted influence.
Writing (1.38) as

$$y(s) = \int_0^s |k(s,t)|e^{i\varphi_k(s,t)}|x(t)|e^{i\varphi_x(t)}|x(s-t)|e^{i\varphi_x(s-t)}dt, \tag{1.44}$$

it can be seen that the phase of y depends on the phase of the fundamental pulse x and the kernel phase. Since the absolute values of y are influenced by the phase of x, also the

Figure 1.9: Dependencies between input and output

$$\begin{array}{ccccc} \varphi_x & \rightarrow & \varphi_y & \leftarrow & \varphi_k \\ & \searrow & & \swarrow & \\ |x| & \rightarrow & |y| & \leftarrow & |k| \end{array}$$

kernel phase must influence $|y(s)|$. Fig. 1.9 shows how the functions in (1.44) depend on each other.

CHAPTER

TWO

MATHEMATICAL ANALYSIS

Table of contents

This section deals with the theoretical analysis of the nonlinear ill-posed autoconvolution equation presented in the previous chapter. We can formulate (1.38) as an operator equation

$$Fx = y, \tag{2.1}$$

where $F : D(F) \subset X \mapsto Y$ is a nonlinear operator mapping from a domain $D(F)$ of a Banach space X into a Banach space Y. In this work we will consider the cases $\{X = C[0,1], Y = C[0,2]\}$, $\{X = L^2[0,1], Y = C[0,2]\}$ and $\{X = L^2[0,1], Y = L^2[0,2]\}$, where $C[a,b]$ is the Banach space of continuous *complex valued* functions over a real interval $[a,b]$ with the norm

$$\|x\|_{C[a,b]} = \max_{t\in[a,b]} |x(t)|$$

and $L^2[a,b]$ the Hilbert space of quadratically integrable *complex valued* functions over a

real interval $[a, b]$ with the norm

$$\|x\|_{L^2[a,b]} = \int_a^b x(t)\overline{x(t)}dt = \int_a^b |x(t)|^2 dt.$$

Several mathematical properties of the equation will be analyzed to draw conclusions on various effects of ill-posedness, for example a discontinuous dependence of solutions x from the measured data y. In other words, it will be shown that small errors in the data may lead to huge errors in the solutions and regularization methods are necessary for the stable reconstruction of the pulses. In detail, we deal with the equation

$$\int_0^s k(s,t)x(s-t)x(t)dt = y(s) \qquad (0 \le t \le 1, 0 \le s \le 2) \tag{2.2}$$

from (1.38).

For the functions occurring in (2.2) we have the following assumption.

Assumption 2.1:

The complex functions x have a support in $[0, 1]$ but will be defined on the whole real axis such that $x : \mathbb{R} \mapsto \mathbb{C}$ with $x(t) = 0$ if $t \notin [0, 1]$. Consequently the functions y are also complex with a support $supp(y) \subset [0, 2]$, but analogously we define $y : \mathbb{R} \mapsto \mathbb{C}$ with $y(s) = 0$ if $s \notin [0, 2]$.

The complex valued kernel function $k = k(s, t)$ is continuous for all $s, t \in [0, 2]$. Therefore the absolute values are bounded by a real constant K,

$$\|k\|_{C([0,2]^2)} \le K < \infty \tag{2.3}$$

or in L^2 formulation

$$\|k\|_{L^2([0,2]^2)} \le \tilde{K} < \infty. \tag{2.4}$$

We can extend the kernel to the whole real plane $k : \mathbb{R}^2 \mapsto \mathbb{C}$ by setting $k(s, t) = 0$ for $s \notin [0, 2]$ or $t \notin [0, 2]$.

Remark 1.2 shows that this assumption applies to the physical problem.

2.1 Some remarks on preceeding publications

The problem under consideration seems to be rather new since we could not find convincing assertions on the autoconvolution problem (2.2) in the literature under the condition that $k(s, t) \neq 1$ and x, y are complex valued functions.

Existing publications consider the real valued autoconvolution equation, indicated by the index $\mathbb{R}$,

$$F_{\mathbb{R}}x = y, \tag{2.5}$$

$$\int_0^s x(s-t)x(t)dt = y(s) \qquad (0 \le t \le 1, 0 \le s \le 1 \text{ or } 0 \le s \le 2), \tag{2.6}$$

where analogously to (2.1) the operator $F_{\mathbb{R}} : D(F) \subset X \mapsto Y$ maps between Banach spaces X of real functions defined on $[0,1]$ and Banach spaces Y defined on $[0,1]$ or $[0,2]$, respectively. In the papers [11], [10] and [17] equation (2.6) was under consideration under the conditions stated above. Different domains $D(F)$ and regularization approaches for the stable approximate solution were discussed there. In [18] the ill-posedness character of the autoconvolution equation was analyzed and inversion rates have been derived using Fourier transform techniques. Lamm and Dai presented in [5] a local regularization theory for the inverse autoconvolution problem and confirmed it numerically. Another regularization approach based on Lavrent'ev regularization was successfully applied to the autoconvolution equation (2.6) in [21] by Janno. Integral equations of the first or third kind involving autoconvolution integrals were discussed in [39], [40], [41] by von Wolfersdorf. Autoconvolution models accur in various fields, see for example spectroscopy [1], stochastics [38] and image reconstruction [4].
A special technique to analyze properties of the ill-posed autoconvolution equation is to transform (2.6) into a well-posed integral equation of the second kind through two times differentiation as done in [20]. Then this attains the form

$$y''(s) = 2x(0)x'(s) + \int_0^t x'(s-t)x'(t)dt \qquad (0 \le s \le 2).$$

This equation has a unique solution $x'(t)$, and it can be shown that equation (2.6) has two solutions x_1 and $x_2 = -x_1$ if $y \in W_2^4[0,1]$ and $y(0) = y'(0) = y''(0) = 0$.
Unfortunately this method fails if a nontrivial kernel is involved. Differentiating

$$y(s) = \int_0^s k(s,t)x(s-t)x(t)dt$$

with respect to s, we have

$$y'(s) = k(s,s)x(0)x(s) + \int_0^s (k'(s,t)x(s-t) + k(s,t)x'(s-t))x(t)dt,$$

which is equivalent to

$$y'(s) = k(s,s)x(0)x(s) + \int_0^s (k'(s,s-t)x(t) + k(s,s-t)x'(t))x(s-t)dt,$$

resulting in a second derivative

$$\begin{aligned} y''(s) = {} & k'(s,s)x(s)x(0) + k(s,s)x'(s)x(0) + k'(s,0)x(s)x(0) + \\ & + \int_0^s k''(s,s-t)x(s-t)x(t) + k'(s,s-t)x'(s-t)x(t) + \\ & + k'(s,s-t)x(s-t)x'(t) + k(s,s-t)x'(s-t)x'(t)dt. \end{aligned}$$

This is not an equation of the second kind. Hence this transformation technique can not be applied to our case. However, for complex functions without a kernel it would be a valid strategy.

2.2 Properties of the forward operator

2.2.1 Continuity

Continuity is a basic property of the forward operator F. It can be expected for practically relevant mathematical models.

Definition 2.2:
An operator $F : D(F) \subset X \mapsto Y$ is called continuous in a point $x \in D(F)$ if for $x + h \in D(F)$ $\|h\|_X \to 0$ implies $\|F(x+h) - F(x)\|_Y \to 0$. F is called continuous in $D(F)$ if it is continuous in all points $x \in D(F)$.

Proposition 2.3:
Under assumption 2.1 the operator $F : D(F) \subset X \mapsto Y$ from (2.1) is continuous for $F : D(F) = L^2[0,1] \mapsto L^2[0,2]$, $F : D(F) = L^2[0,1] \mapsto C[0,2]$ and $F : D(F) = C[0,1] \mapsto C[0,2]$. The image $y = F(x)$ is a continuous complex function in any of these cases.

Proof. At first, the case $F : L^2[0,1] \mapsto C[0,2]$ will be considered. Because of the kernel, there is no inequality

$$\|F(x_1) - F(x_2)\|_{C[0,2]} \leq (\|x_1\|_{L^2} + \|x_2\|_{L^2[0,1]})\|x_1 - x_2\|_{L^2[0,1]}$$

as used in [11]. Instead, we have for fixed $x \in L^2[0,1]$ and an increment $h \in L^2[0,1]$ with

$h(t) = 0$ for $t \notin [0,1]$

$$
\begin{aligned}
\|F(x+h) - F(x)\|_{C[0,2]} &= \max_{s\in[0,2]} |[F(x+h)](s) - [F(x)](s)| \\
&= \max_{s\in[0,2]} \left| \int_0^s k(s,t)((x(s-t)+h(s-t))(x(t)+h(t)) - x(s-t)x(t))dt \right| \\
&= \max_{s\in[0,2]} \left| \int_0^s k(s,t)h(s-t)h(t)dt + \int_0^s k(s,t)h(s-t)x(t)dt + \int_0^s k(s,t)x(s-t)h(t)dt \right| \\
&\leq \underbrace{\max_{s\in[0,2]} \left| \int_0^s k(s,t)h(s-t)h(t)dt \right|}_{(1)} + \underbrace{\max_{s\in[0,2]} \left| \int_0^s k(s,t)h(s-t)x(t)dt \right|}_{(2)} + \underbrace{\max_{s\in[0,2]} \left| \int_0^s k(s,t)x(s-t)h(t)dt \right|}_{(3)}
\end{aligned}
\tag{2.7}
$$

To show continuity in point x, (1), (2) and (3) must vanish for $\|h\|_{L^2[0,1]} \to 0$. With the Cauchy-Schwarz-inequality, and $s \in [0,2]$ it is

$$
\begin{aligned}
(1) &\leq \|k\|_{L^2([0,2]^2)} \|h\|_{L^2[0,2]} \|h\|_{L^2[0,2]} \leq K\|h\|^2_{L^2[0,1]}, \\
(2) &\leq \|k\|_{L^2([0,2]^2)} \|h\|_{L^2[0,2]} \|x\|_{L^2[0,2]} \leq K\|x\|_{L^2[0,1]} \|h\|_{L^2[0,1]}, \\
(3) &\leq \|k\|_{L^2([0,2]^2)} \|x\|_{L^2[0,2]} \|h\|_{L^2[0,2]} \leq K\|x\|_{L^2[0,1]} \|h\|_{L^2[0,1]}.
\end{aligned}
$$

Because $x(t) = 0$ for $t \notin [0,1]$, we have $\|x\|_{L^2[0,2]} = \|x\|_{L^2[0,1]}$ and analogously the norms of h remain the same on $L^2[0,1]$. From assumption 2.1 we get $\|k\|_{L^2([0,2]^2)} < K$. Thus, all three terms in (2.7) go to zero for $\|h\|_{L^2[0,1]} \to 0$ and the continuity is shown in every point $x \in L^2[0,1]$. If a weaker norm in Y is chosen, the property of continuity remains as well as if a stronger norm in X is taken. Thus, starting from $F : D(F) = L^2[0,1] \mapsto C[0,2]$, F is also continuous for $F : D(F) = L^2[0,1] \mapsto L^2[0,2]$ and $F : D(F) = C[0,1] \mapsto C[0,2]$.
It remains to prove that $F(x)$ is continuous for any $x \in L^2[0,1]$.
For all $x \in L^p[0,1]$ $(1 \leq p < \infty)$ we define the *modulus of continuity*

$$
\mu_x^{L^p}(\Delta) := \sup_{|h|\leq\Delta} \left(\int_0^1 |x(t+h) - x(t)|^p dt \right)^{1/p} = \sup_{|h|\leq\Delta} \|x(\cdot + h) - x(\cdot)\|_{L^p[0,1]} \leq \infty.
$$

The property of *continuity in the mean* (cf [3, p. 4]) ensures

$$
\lim_{\Delta\to 0} \mu_x^{L^p}(\Delta) = 0 \qquad \forall x \in L^p[0,1] \quad (1 \leq p < \infty). \tag{2.8}
$$

Because we have for $0 < s < s + h \leq 2$

$$\begin{aligned}|[F(x)](s+h) - [F(x)](s)| &= \left| \int_0^{s+h} k(s+h,t)x(s+h-t)x(t)dt - \int_0^s k(s,t)x(s-t)x(t)dt \right| \\ &= \left| \int_0^s (k(s+h,t)x(s+h-t) - k(s,t)x(s-t))x(t)dt + \int_s^{s+h} k(s+h,t)x(s+h-t)x(t)dt \right| \\ &\leq K\|x\|_{L^2[0,1]}\|x(\cdot+h) - x(\cdot)\|_{L^2[0,1]} + K\|x\|^2_{L^2[0,1]}h, \end{aligned} \tag{2.9}$$

(2.8) ensures $|[F(x)](s+h) - [F(x)](s)| \to 0$ for $h \to 0$ and $F(x)$ is continuous. □

Remark 2.4:

If the preimage space X and image space Y are not defined on the same real interval, the requirement $x(t) = 0$ for $t \notin [0,1]$ in the domain of F is essential. Otherwise continuity can not be expected. Let for example

$$h_n(t) = \begin{cases} \frac{1}{n}, & t \in [0,1] \\ \infty, & t \in (1,2]. \end{cases}$$

Then we have $\|h_n\|_{L^2[0,1]} \to 0$ for $n \to \infty$, but $\|F(x+h_n)\|_{L^2[0,2]} = \infty \quad \forall n \in \mathbb{N}$, in other words F is not continuous.

2.2.2 Injectivity

Another important property is injectivity. If an operator F is injective, for given y the equation $y = Fx$ has only one solution x. For the real case (2.6) it has been shown in [11] that $F_{\mathbb{R}}$ is injective if $y \in R_0^+ := \{y \in C[0,1] : y \geq 0, max\{s : y(\xi) = 0 \; \forall \xi \in [0,s]\} = 0\}$ for positive functions $x(t) > 0 \; \forall t \in [0,1]$. A similar result can be derived in the general case (2.2), but injectivity is lost as with x also $-x$ is a solution and for complex functions the domain restriction $x(t) > 0$ can not be maintained.

A lemma sometimes called Titchmarsh's theorem will be the basis for analysis on the injectivity.

Lemma 2.5:

For $f \in L^2[0,1]$ and $g \in L^2[0,1]$ let γ $(0 < \gamma \leq 1)$ exist, such that

$$\int_0^s f(s-t)g(t)dt = 0 \qquad (0 \leq s \leq \gamma).$$

Then there exist numbers $\alpha, \beta \in [0,1]$ with $\alpha + \beta \geq \gamma$, $f(t) = 0$ a.e. in $t \in [0,\alpha]$ and $g(t) = 0$ a.e. in $t \in [0,\beta]$.

Proof. See for example [36]. □

At first we want to discuss injectivity in the case that the kernel only depends on the first variable, $k(s,t) = k(s)$.

Proposition 2.6:
Let $k(s,t) = k(s)$ with $k(s) \in C[0,2]$, $|k(s)| \neq 0 \forall s \in [0,2]$ and $supp(y) = [\delta_1, 2-\delta_2] \subset [0,2]$, $0 \leq \delta_1, \delta_2 \leq 1$. Then the autoconvolution equation (2.2) has two solutions x_1 and $x_2 = -x_1$ with $supp(x_1) = supp(x_2) = [\delta_1/2, 2 - \delta_2/2]$.

Proof. At first let $\delta_1 = \delta_2 = 0$. If the kernel does not depend on the variable t we have for $x_1, x_2 \in X$

$$\begin{aligned}[F(x_1)](s) - [F(x_2)](s) &= \int_0^s k(s)x_1(s-t)x_1(t) - k(s)x_2(s-t)x_2(t)dt \\ &= \int_0^s k(s)(x_1(s-t) - x_2(s-t))(x_1(t) + x_2(t))dt.\end{aligned}$$

If $|k(s)| \neq 0 \quad \forall s \in [0,2]$, then

$$[F(x_1)](s) - [F(x_2)](s) = 0 \iff \int_0^s (x_1(s-t) - x_2(s-t))(x_1(t) + x_2(t))dt = 0 \qquad (0 \leq s \leq 2). \tag{2.10}$$

By lemma 2.5 the integral in (2.10) is zero if either $x_1(t) = x_2(t)$ for all $t \in [0,1]$ or $x_1(t) = -x_2(t)$ for all $t \in [0,1]$.
Now let $\delta_1, \delta_2 > 0$. Then by lemma 2.5 we have for a non-vanishing kernel $k(s)$

$$0 = \int_0^s k(s)x(s-t)x(t)dt \ \forall s \in [0, \delta_1/2] \iff x(t) = 0 \text{ a.e in } [0,a] \text{ and } x(t) = 0 \text{ a.e in } [0,b]$$

with $a + b \geq \delta_1$. Since a function can have only one support it is obviously $a = b \geq \delta_1/2$. On the other hand $a = b > \delta_1/2$ would yield $supp(y) \neq [\delta_1, 2 - \delta_2]$ and we have $a = b = \delta_1/2$. Finally

$$0 = \int_{2-s}^2 k(s)x(s-t)x(t)dt \qquad (0 \leq s \leq \delta_2/2)$$

is equivalent to

$$0 = \int_0^s k(s)x(s-(-t+2))x(-t+2)dt \qquad (0 \leq s \leq \delta_2/2)$$

and hence by analogous arguments to the lower boundary case we get $x(t) = 0$ for $t \in [2 - \delta_2/2, 2]$. □

Remark 2.7:

From Proposition 2.6 it becomes clear that for the real valued case with $F : D(F) \subset L^2_{\mathbb{R}}[0,1] \to L^2_{\mathbb{R}}[0,2]$ injectivity can be achieved by adding a constraint $x(t) > 0 \; \forall t \in [0,1]$. In the complex valued case no such restriction is possible. If the whole autoconvolution signal $y \in L^2[0,2]$ can be measured there are no ambiguities if $y(s) = 0$ for $s \in [0, \delta]$ as in the case where also y is only known for $0 \leq s \leq 1$, cf [11, p.360].

Now let the kernel be also dependent on the variable t and $R(F)$ denote the range of F, $R(F) := \{y \in Y : \exists\, x \in X$ such that $y = F(x)\}$. Then we have the following proposition.

Proposition 2.8:

For given data $y \in R(F)$ the operator F in (2.1) has at least two solutions x_1 and $x_2 = -x_1 \in X$.

Proof. Let $x_1, x_2 \in X$.Then

$$\begin{aligned} 0 = [F(x_1)](s) - [F(x_2)](s) &= \int_0^s k(s,t)(x_1(s-t)x_1(t) - x_2(s-t)x_2(t))dt \\ &= \int_0^s k(s,t)(x_1(s-t) - x_2(s-t))(x_1(t) + x_2(t))dt + \\ &\quad + \int_0^s k(s,t)x_1(s-t)x_2(t)dt - \int_0^s k(s,t)x_2(s-t)x_1(t)dt. \end{aligned} \tag{2.11}$$

□

The first integral is zero if $x_1(t) = x_2(t)$ or $x_1(t) = -x_2(t)$ for all $t \in [0,1]$. In these cases the other two integrals cancel each other out and $F[x_1] = F[x_2]$.

Although the author assumes that there exist only two solution, it is an open problem if additional solutions for equation (2.11) exist.

For the recovery of the SPIDER-signal this non-injectivity does not matter. Since for complex $x(t) = |x(t)|e^{i\varphi_x(t)}$, $-x(t) = |x(t)|e^{i\varphi_{-x}(t)}$ with $\varphi_x(t) = \varphi_{-x}(t) - \pi$ and we are only interested in the group delay (see (1.2)), both solutions are identical with respect to our interest. In particular there is no need to regularize the solutions concerning injectivity as both solutions are acceptable.

2.2.3 Weak closedness

Definition 2.9:
An operator $F : D(F) \subset X \mapsto Y$ is called *weakly (sequentially) closed* if for any sequence $\{x_n\}_{N=1}^{\infty} \subset D(F)$, weak convergence $x_n \rightharpoonup x_0$ in X and $F(x_n) \rightharpoonup y_0$ in Y imply $x_0 \in D(F)$ and $y_0 = F(x_0)$.

Proposition 2.10:
The operator F from (2.1) under assumption 2.1 is weakly (sequentially) closed for $F : L^2[0,1] \mapsto L^2[0,2]$.

Proof. The proof follows the ideas from Theorem 2 in [11, p. 361f], but the different situation requires to carry out the proof completely here.
For $z_n \rightharpoonup z_0$ in $L^2[0,2]$ it is necessary and sufficient that $\{z_n\}_{n=0}^{\infty}$ is bounded in $L^2[0,2]$ and that for all $0 \leq \tau \leq 1$ it holds that $\int_0^{\tau} z_n(t)dt \to \int_0^{\tau} z_0(t)dt$ as $n \to \infty$ (cf [24, p 151]. The L^2-norms of the sequence $\{x_n\}_{n=1}^{\infty}$ are then uniformly bounded by a constant $C < \infty$.
We have

$$\|F(x)\|_{L^2[0,2]} = \int_0^2 \int_0^s k(s,t)x(s-t)x(t)dtds.$$

With Schwarz' inequality and $\|x\|_{L_2[0,2} = \|x\|_{L^2[0,1]}$ because of $supp(x) = [0,1]$ it is

$$\begin{aligned}\|F(x)\| &\leq K \int_0^2 \|x\|_{L^2[0,1]}\|x\|_{L^2[0,1]}ds \\ &\leq 2K\|x\|^2_{L^2[0,1]},\end{aligned}$$

showing that F and thus $\{F(x_n)\}$ is bounded.

To complete the proof, it is sufficient to show that for all $\tau \in [0,2]$

$$\begin{aligned}\Theta_n = \Theta_n(t) := &\left| \int_0^\tau [F(x_n)](s) - [F(x_0)](s) ds \right| \\ =| \int_0^\tau [&\underbrace{\int_0^s k(s,t)(x_n(s-t) + x_0(s-t))(x_n(t) - x_0(t))dt}_{(1)} + \\ &+ \underbrace{\int_0^s k(s,t)x_n(s-t)x_0(t)dt}_{(2)} - \underbrace{\int_0^s k(s,t)x_0(s,t)x_n(t)dt}_{(3)}]ds|\end{aligned} \tag{2.12}$$

tends to zero as $n \to \infty$. Integrals (2) and (3) are linear functionals in x_n for fixed s and x_0. Because $x_n \rightharpoonup x_0$

$$\begin{aligned}(2) \int_0^s k(s,t)x_n(s-t)x_0(t)dt &\to \int_0^s k(s,t)x_0(s-t)x_0(t)dt, \\ (3) \int_0^s k(s,t)x_0(s-t)x_n(t)dt &\to \int_0^s k(s,t)x_0(s-t)x_0(t)dt,\end{aligned}$$

so both terms cancel each other out for $n \to \infty$. If $x_n \rightharpoonup x_0$ then then there exists a sequence $\{\Delta_n\}_{n=1}^\infty$ so that $x_n = x_0 + \Delta_n$ and $\Delta_n \rightharpoonup 0$ as $n \to \infty$. For the first integral we have then

$$\begin{aligned}&\int_0^s k(s,t)(x_n(s-t) + x_0(s-t))(x_n(t) - x_0(t))dt \\ &= \int_0^s k(s,t)(2x_0(s-t) + \Delta_n(s-t))\Delta_n(t)dt \\ &= \int_0^s k(s,t)2x_0(s-t)\Delta_n(t)dt + \int_0^s k(s,t)\Delta_n(s-t)\Delta_n(t)dt.\end{aligned} \tag{2.13}$$

□

From $\Delta_n \rightharpoonup 0$ follows the convergence of integral (1) in (2.12) to zero. Therefore F is weakly continuous and since $L^2[0,1]$ is weakly closed the operator F is weakly (sequentially) closed.

2.2.4 Compactness

Compactness of an operator is of special interest as it provides information about the ill-posedness of the operator. In the appendix of [9], Engl, Kunisch and Neubauer proved the following proposition.

Proposition 2.11:
Let F be a (nonlinear) compact, continuous and weakly closed operator. Moreover, assume that $y_0 = F(x_0)$ and that there exists an $\varepsilon > 0$ such that $y = F(x)$ has a unique solution for all $y = F(D(F)) \cap U_\varepsilon(y_0)$. If there exists a sequence $\{x_n\} \subset D(F)$ satisfying $x_n \rightharpoonup x_0$ but $x_n \nrightarrow x_0$, then F^{-1} (defined on $F(D(F)) \cap U_\varepsilon(y_0)$) is not continuous in y_0.

In other words, if the operator is compact and there is such a sequence x_n (which, as mentioned in Remark 4 of the cited paper, exists if F is defined on a non-compact domain) , a small error in the data y^δ can lead to huge errors in the solution x^δ.

Definition 2.12:
A subset $S \subset X$ is called *relatively compact*, if any sequence $\{x_n\} \subset X$ has a subsequence which converges in X.

Definition 2.13:
An operator $F : X \mapsto Y$ is called *compact*, if it maps any bounded subset of X into a relatively compact subset of Y.

A comprehensive discussion of compactness for the real-valued, kernel-free autoconvolution equation (2.6) can be found in [11, p.357ff]. There it has been shown theoretically and exemplarily that the operator F is not compact in the general cases $F_{\mathbb{R}} : L^2[0,1] \mapsto L^2[0,1]$ and $F_{\mathbb{R}} : C[0,1] \mapsto C[0,1]$, but by considering specific domains $D(F) \subset X$ compactness can be achieved. Namely, the operator is compact if $F_{\mathbb{R}} : D(F) \subset L^2[0,1] \mapsto L^2[0,1]$ and $D(F)$ contains only non-negative, non-decreasing functions or if $F_{\mathbb{R}} : D(F) \subset C[0,1] \mapsto C[0,1]$ contains only monotone functions. Neither of these conditions can be used for complex functions. It is hard to show compactness without using examples. Because a kernel could essentially influence the behaviour of the operator, all examples would have to be verified for all possible kernels. We will at first consider the trivial kernel and then try to draw conclusions about the nontrivial case.

Proposition 2.14:
Let $k(s,t) \equiv 1$. Then the operator F in (2.1) is not compact for $F : C[0,1] \mapsto C[0,2]$ or $F : L^2[0,1] \mapsto L^2[0,2]$.

Proof. Example 4 in [11, p. 359] shows non-compactness for the real valued problem (2.5,2.6) for $F_{\mathbb{R}} : L^2[0,1] \mapsto L^2[0,1]$ and $F_{\mathbb{R}} : C[0,1] \mapsto C[0,1]$. The same example applies

to the situation when the whole data is measured. This a special case of our problem (2.1,2.2) with $k(s,t) \equiv 1$ and $x(t), y(s) \in \mathbb{R}\ \forall t \in [0,1], s \in [0,2]$.
Let $x_n(t) := \sin(nt)$. We obtain

$$y_n(s) := [F(x_n)](s) = -\frac{s\cos(ns)}{2} + \frac{\sin(ns)}{2n} \qquad (0 \le s \le 2; n = 1,2,...)$$

and compare these functions with $x_0(t) = 0$ $(0 \le t \le 1)$ and $y_0(s) := [F(x_0)](s) = s$ $(0 \le s \le 2)$. Because we still consider x_n on the interval $[0,1]$ the properties are the same as in [11, Example 4, p. 359] and we have $\|x_n\|_{C[0,1]} = 1$, $\|x_n\|_{L^2[0,1]} \le 1$, $\lim_{n\to\infty} \|x_n\|_{L^2[0,1]} = \frac{1}{2}\sqrt{2} \neq 0$ and $\lim_{n\to\infty} \int_0^\tau x_n(t)dt = 0 = \int_0^\tau x_0(t)dt$ for all $0 \le \tau \le 1$. Thus we have weak convergence $x_n \rightharpoonup x_0$ as $n \to \infty$ (cf [24, 151]), but norm convergence $\|x_n\|_{L^2[0,1]} \to \frac{1}{2}\sqrt{2} \neq \|x_0\|_{L^2[0,1]} = 0$. Consequently there is no subsequence x_{n_k} strongly convergent in $L^2[0,1]$. For the functions y_n we now have to consider the new interval $[0,2]$ instead of $[0,1]$. We then have $\|y_n\|_{L^2[0,2]} \le \|y_n\|_{C[0,2]} \le 2$, but still

$$\int_0^\tau \left(-\frac{s\cos(ns)}{2} + \frac{\sin(ns)}{2n}\right) ds \to 0 \qquad \forall 0 \le \tau \le 2$$

and weak convergence $y_n \rightharpoonup y_0$ in $L^2[0,2]$ as $n \to \infty$. From
$\lim_{n\to\infty} \|y_n\|_{L^2[0,2]} = \frac{1}{\sqrt{3}} \neq 0 = \|y_0\|_{L^2[0,2]}$ it becomes clear that the autoconvolution operator F from (2.1) does not map every bounded subset of $L^2[0,1]$ or $C[0,1]$ onto a relatively compact subset in $L^2[0,2]$ or $C[0,2]$, respectively. Hence F is not compact whenever these function spaces are considered. □

Since the autoconvolution equation is not compact for the trivial kernel $k(s,t) \equiv 1$, we do not expect it to be compact for a nontrivial kernel like the one occurring in the physical problem. The following remark will give another justification for this.

Remark 2.15:

Let assumption 2.1 hold true. From lemma 2.3 we have $y \in C[0,2]$ even if $x \in L^2[0,1]$. It is therefore sufficient to investigate the case $y \in C[0,2]$. By the the Arzelá-Ascoli theorem (e.g [24, p 163]), a subset $D \subset C[a,b]$ is relatively compact if and only if it is uniformly bounded and equicontinuous.
Let therefore $F : D(F) \subset X \mapsto C[0,2]$ with $X \in \{L^2[0,1], C[0,1]\}$ and a domain $D(F) := \{x \in X : \|x\|_X \le M\}$ of functions bounded by a real constant $M < \infty$. It is easy to see that for a continuous kernel with $K := \max_{s,t\in[0,2]^2} k(s,t)$ from assumption 2.1

the range of F is bounded,

$$\begin{aligned}\sup_{s\in[0,2]}|y(s)| &= \sup_{s\in[0,2]}\left|\int_0^s k(s,t)x(s-t)x(t)dt\right| \\ &\leq \sup_{s\in[0,2]}\left|K\|x\|_{C[0,1]}\|x\|_{C[0,1]}\int_0^s dy\right| \\ &\leq 2KM^2.\end{aligned} \tag{2.14}$$

If $x \in L^2[0,1]$ only, then using Cauchy-Schwarz-inequality we derive from (2.14)

$$\sup_{s\in[0,2]}|y(s)| \leq KM^2.$$

Hence $F(x)$ is bounded in both cases. However, equicontinuity cannot be shown without further restrictions on the domain of F. It is

$$\begin{aligned}|y(s+h)-y(s)| &= \left|\int_0^{s+h} k(s,t)x(s+h-t)x(t)dt - \int_0^s k(s,t)x(s-t)x(t)dt\right| \\ &= \left|\int_0^s k(s,t)(x(s+h-t)-x(s-t))x(t)dt + \int_s^{s+h} k(s,t)x(s+h-t)x(t)dt\right| \\ &\leq K\|x\|_{C[0,1]}\int_0^s |x(s+h-t)-x(s-t)|dt + K\|x\|^2_{C[0,1]}h \\ &\leq KM\int_0^s |x(s+h-t)-x(s-t)|dt + KM^2h\end{aligned} \tag{2.15}$$

for $x \in C[0,1]$ or

$$|y(s+h)-y(s)| \leq KM\|x(\cdot+h)-x(\cdot)\|_{L^2[0,1]} + o(h) \tag{2.16}$$

respectively for $x \in L^2[0,1]$. Neither the remaining integral in (2.15) nor the norm $\|x(\cdot+h)-x(\cdot)\|_{L^2[0,1]}$ in (2.16) converge to zero equally for all functions $x \in C[0,1]$ or $x \in L^2[0,1]$.
Hence F is not expected to be compact. It follows immediately that compactness can be achieved if the domain $D(F)$ is chosen in a way that (2.15) or (2.16) ensure the equicontinuity of the functions y.

Compactness of F can of course always be forced by choosing a compact domain $D(F)$. Since F is continuous, the Tikhonov theorem (cf [15, p.24]) provides the compactness of F and the solutions x depend continuously on the right-hand side y (cf [8, section 4]). An example of a compact domain is shown below.

Definition 2.16:
We define a domain of bounded total variation

$$D^{TV} := x \in L^2[0,1] : |x(t)| \leq M \quad (0 \leq t \leq 1) \qquad \bigvee_{t=0}^{1} x(t) \leq M$$

where M is a real constant.

Lemma 2.17:
The set D^{TV} from defintion 2.16 is a compact subset of $L^p[0,1] \quad (1 \leq p < \infty)$.

Proof. See [11, p.356]. There the theorem has been proven for real functions, but the lemma and the lemmata required for the proof follow with the same arguments for complex functions. □

Proposition 2.18:
The operator F in 2.1 is compact in both cases $F : D^{TV} \subset L^2[0,1] \mapsto C[0,2]$ or $F : D^{TV} \subset L^2[0,1] \mapsto L^2[0,2]$.

2.3 Derivatives of the forward operator

For the computation of the regularized solution a derivative-based algorithm will be used. Therefore the derivative of the operator F in (2.2) has to be examined. We will only consider $F : L^2[0,1] \mapsto L^2[0,2]$ furthermore.

Definition 2.19:
A linear operator $A : H_1 \mapsto H_2$ is called *Hilbert-Schmidt operator* if, for an arbitrarily chosen complete orthonormal system $\{u^{(i)}\}_{i=1}^{\infty} \subset H_1$, $\sum_{i=1}^{\infty} \|Au^{(i)}\|_2^2 < \infty$.

Lemma 2.20:
Any Hilbert-Schmidt operator is compact. The integral operator A defined by $[Ax](s) = \int_0^1 k(s,t)x(t)dt$ with a quadratically integrable kernel $\int_0^1 \int_0^1 |k(s,t)|^2 dsdt < \infty$ is a Hilbert-Schmidt operator.

Proof. See [15, p.28]. There it has been mentioned that this lemma also applies to Volterra operators

$$\int_0^s k(s,t)x(t)dt = y(s). \tag{2.17}$$

□

Definition 2.21:
A bounded linear operator $F'(x_0)$ is called *Gâteaux-derivative* of an operator F in a point $x_0 \in X$, if

$$\lim_{\tau \to 0} \frac{\|F(x_0 + \tau h) - F(x_0) - F'(x_0)\|_Y}{\tau} = 0 \tag{2.18}$$

for $x_0 + \tau h \in X$, $\tau \in \mathbb{R}$.

Definition 2.22:
A bounded linear operator $F'(x_0) \in \mathcal{L}(X, Y)$ is called *Fréchet-derivative* of an operator F in a point $x_0 \in X$, if

$$\lim_{h \to 0} \frac{\|F(x_0 + h) - F(x_0) - F'(x_0)h\|_Y}{\|h\|_X} = 0 \qquad \forall h \in X \tag{2.19}$$

for all $x_0 + h \in D(F)$.

Lemma 2.23:
The operator F in (2.1) has a Gâteaux-derivative in a point $x_0 \in D(F)$ of the form

$$[F'(x_0)h]](s) = \int_0^s (k(s,t) + k(s, s-t))x_0(s-t)h(t)dt \qquad \forall x_0 \in X. \tag{2.20}$$

The Fréchet-derivative coincides with (2.20) and is a compact operator for every $x_0 \in X$.

Proof. We have

$$\begin{aligned}
&[F(x_0+\tau h)](s) - [F(x_0)](s) = \\
&\quad \int_0^s k(s,t)(x_0(s-t) + \tau h(s-t))(x_0(t) + \tau h(t))dt - \int_0^s k(s,t)x_0(s-t)x_0(t)dt \\
&\quad = \int_0^s k(s,t)x_0(s-t)x_0(t)dt + \tau \int_0^s k(s,t)h(s-t)x_0(t)dt \\
&\qquad + \tau \int_0^s k(s,t)x_0(s-t)h(t)dt + \tau^2 \int_0^s k(s,t)h(s-t)h(t)dt - \int_0^s k(s,t)x_0(s-t)x_0(t)dt.
\end{aligned} \tag{2.21}$$

The first term equals $[F(x)](s)$. This one and the last term cancel each other out. The second and third term can be combined by substituting $\hat{t} := s - t$ in the second term.

Switching the integral limits and resubstituting $t = \hat{t}$ it is

$$\int_0^s k(s,t)x_0(t)h(s-t)dt = -\int_s^0 k(s,s-\hat{t})x_0(s-\hat{t})h(\hat{t})d\hat{t}$$
$$= \int_0^s k(s,s-t)x_0(s-t)h(t)dt.$$

Combining this with (2.21), we have for the middle integrals

$$\int_0^s k(s,t)h(t)x_0(s-t)dt + \int_0^s k(s,t)h(s-t)x_0(t)dt$$
$$= \int_0^s (k(s,t)+k(s,s-t))x_0(s-t)h(t)dt.$$

Thus, the limit from definition 2.18

$$\lim_{\tau\to 0}\frac{\|F(x_0+\tau h)-F(x_0)-F'(x_0)\|_Y}{\tau}$$
$$= \lim_{\tau\to 0}\frac{\|\tau\int_0^s (k(s,t)+k(s,s-t))x_0(s-t)h(t)dt + \tau^2\int_0^s k(s,t)h(s-t)h(t)dt - F'(x_0)h\|_Y}{\tau}$$

is zero if $[F'(x_0)h](s) := \int_0^s (k(s,t)+k(s,s-t))x_0(s-t)h(t)dt$, which is obviously linear in h. One can easily see that $F'(x_0)$ also satisfies (2.20). For the Fréchet-derivative it remains to prove that $F'(x_0)$ is bounded. It is

$$\|F(x_0)h\|^2_{L^2[0,2]} = \int_0^2 \overline{\int_0^s (k(s,t)+k(s,s-t))x_0(s-t)h(t)dt}\int_0^s (k(s,t)+k(s,s-t))x_0(s-t)h(t)dtds$$
$$\leq \int_0^2 \left|\int_0^s (k(s,t)+k(s,s-t))x_0(s-t)h(t)dt\right|^2 ds. \tag{2.22}$$

With the constant K from assumption 2.1 and by using Hölder's inequality we have

$$\begin{aligned}
\|F'(x_0)h\|^2_{L^2[0,2]} &\leq \int_0^2 \left(\sqrt{\int_0^s |(k(s,t)+k(s,s-t))x_0(s-t)|^2 dy} \sqrt{\int_0^s |h(t)|^2 dt} \right)^2 ds \\
&\leq \int_0^2 \left(\sqrt{\int_0^s |2Kx_0(s-t)|^2 dt} \sqrt{\int_0^s |h(t)|^2 dt} \right)^2 ds \\
&\leq \int_0^2 4K^2 \int_0^s |x_0(s-t)|^2 dt \int_0^s |h(t)|^2 dt ds \\
&\leq 4K^2 \int_0^2 \int_0^s |x_0(s-t)|^2 dt \int_0^s |h(t)|^2 dt ds \\
&\leq 4K^2 \int_0^2 \int_0^2 |x_0(s-t)|^2 dt ds \int_0^2 |h(t)|^2 dt \\
&\leq 8K^2 \|x_0\|^2_{L^2[0,1]} \|h\|^2_{L^2[0,1]}. \qquad (2.23)
\end{aligned}$$

For any fixed x_0 we have $\|x_0\|_{L^2[0,1]} < \infty$ and $F'(x_0)$ is bounded.
Because $x_0 \in L^2[0,1]$ and the kernel $\tilde{k}(s,t) := (k(s,t)+k(s,s-t))x_0(s-t)$ is quadratically integrable, $F'(x_0)$ is a Hilbert-Schmidt operator of Volterra type (2.17) and thus compact for all $x_0 \in D(F)$. □

Since we will also need the adjoint of the Fréchet derivative for the iterative regularization algorithm, it will be presented in its analytic form here, too.

Definition 2.24:
Let $A : X \mapsto Y$ be linear operator mapping between Hilbert spaces H_1 and H_2. The adjoint operator $A^* : H_1 \mapsto H_2$ is defined by

$$\langle Ax, f \rangle = \langle x, A^* f \rangle \quad \forall x \in X, \forall f \in H_2.$$

Here $\langle \cdot, \cdot \rangle$ denotes the scalar product in $L^2[0,2]$.

Lemma 2.25:
The adjoint $F'(x_0)^*$ of the Fréchet derivative from lemma 2.23,
$F'(x_0) : L^2[0,1] \mapsto L^2[0,2]$ takes the form

$$[F'(x_0)^* h](s) = \int_s^2 \overline{(k(s,t)+k(s,s-t))}\ \overline{x(s-t)} h(s) ds. \qquad (2.24)$$

Proof. We have to show that

$$\langle F'(x_0)h, u\rangle = \langle h, F'(x_0)^* u\rangle \qquad \forall h, u \in L^2[0,1].$$

It is

$$\langle F'(x_0)h, u\rangle = \int_0^2 \int_0^s (k(s,t) + k(s, s-t))x_0(s-t)h(t)dt\overline{u(s)}ds.$$

Applying Fubini's theorem we get

$$\begin{aligned}\langle F'(x_0)h, u\rangle &= \int_0^2 \int_t^2 (k(s,t) + k(s, s-t))x_0(s-t)h(t)\overline{u(s)}dsdt \\ &= \int_0^2 h(t) \overline{\int_t^2 \overline{(k(s,t) + k(s, s-t))x_0(s-t)}u(s)dsdt} \\ &= \langle h, F'(x_0)^* u\rangle.\end{aligned}$$

□

2.4 Ill-posedness

Based on the results of the previous sections we can now draw conclusions on the ill-posedness of the operator F from (2.1) and the associated linear problem

$$F'(x_0)x = \tilde{y}, \tag{2.25}$$

where $F'(x_0)$ is the Fréchet derivative of F in x_0 (see lemma 2.23).
There exist various definitions of ill-posedness of which some will be discussed here. The most well-known definition was first formulated by Hadamard [13].

Definition 2.26:
An operator F of an operator equation $F(x) = y$, $F : X \mapsto Y$ is called well-posed if

- *(Existence)* for every $y \in Y$ there exists $x \in X$ with $F(x) = y$
- *(Uniqueness)* for every $y \in Y$ there exists exactly one $x \in X$ with $F(x) = y$
- *(Stability)* the inverse operator F^{-1} is continuous

If at least one of these properties is violated, F is called ill-posed.

In lemma 2.8 we have shown that the uniqueness condition always fails. But also the existence condition does not hold. Let for example $x \in L^2[0,1]$ with $|x| \neq 0$ in $[0, \varepsilon]$ and

$k(s,t) \equiv 1$ for simplicity. Then

$$F[x](s) = \int_0^s x(s-t)x(t)dt \qquad (0 \le s \le \varepsilon_2)$$

is monotonically increasing for some $0 < \varepsilon_2 < \varepsilon_1$. Thus, any strictly monotonically decreasing function y is not part of the range of F. Examples showing discontinuity of the inverse operator in the cases $F : L^2_{\mathbb{R}}[0,1] \mapsto L^2_{\mathbb{R}}[0,1]$ and $F : C_{\mathbb{R}}[0,1] \mapsto C_{\mathbb{R}}[0,1]$ can be found in [15, p. 39] or [11, p. 357]. We will present an example for the case $F : L^2[0,1] \mapsto L^2[0,2]$ later. If F is defined on a compact domain $D(F)$, the Tichonov-theorem (cf [15, p.24]) ensures that F^{-1} is continuous if F is injective over $D(F)$. Although Hadamard's definition applies to all operator equations, it is more suitable for linear equations. Since for nonlinear operators the properties strongly depend on the local behaviour of the operator around the points x_0, none of Hadamards conditions can be expected to hold in general, but for certain points some might be true. Therefore for nonlinear equations local analysis of the properties is more appropriate. Because both uniqueness and existence are usually hard to maintain, we will use the following definition of ill-posedness that focuses on the stability which is the most important property for the regularization.

Definition 2.27:
We define a nonlinear operator equation (2.1) to be locally ill-posed in x_0 if, for arbitrarily small $\rho > 0$ there exists a sequence $\{x_n\} \subset B_\rho(x_0)$ satisfying the condition

$$F(x_n) \to F(x_0) \text{ in } Y \text{ as } n \to \infty, \text{ but } x_n \nrightarrow x_0 \text{ in } X. \tag{2.26}$$

Otherwise the equation is called locally well-posed.

This condition is similar to Hadamard's instability condition, but to emphasize this again Hadamard is of global character whereas Def. 2.27 only considers a ball $B_\rho(x_0)$ around the solution x_0 with radius $\rho > 0$.

Proposition 2.28:
The autoconvolution equation (2.2) is everywhere locally ill-posed.

Proof. We will use the basic idea from the proof of Prop. 2.3. in [18, p.422].
Let $\hat{x} \in D(F) \subset L^2[0,1]$ and $x_n := \hat{x} + q_n$ (n=1,2,...). The perturbations q_n are selected from the family of functions

$$\Psi_\beta(t) = \frac{r\sqrt{1-2\beta}}{t^\beta}, \qquad 0 \le \beta \le \frac{1}{2},$$

where $q_n := \Psi_{\frac{1}{2}-\frac{1}{n}}(t)$ for $t \in [0,1]$ and $q_n(t) := 0$ if $t \notin [0,1]$. For the functions Ψ_β we have $\|\Psi_\beta\|_{L^2[0,1]} = r$ for all parameters $0 \le \beta \le \frac{1}{2}$ and

$\|\Psi_\beta * \Psi_\beta\|_{L^2[0,2]} \leq \sqrt{2}\|\Psi_\beta * \Psi_\beta\|_{C[0,2]} \leq \sqrt{2}r^2(1-2\beta)\pi 2^{1-2\beta} \to 0$ as $\beta \to \frac{1}{2}$. We have

$$F[x_n](s) = \int_0^s k(s,t)(\hat{x}(s-t) + q_n(s-t))(\hat{x}(t) + q_n(t))dt$$

$$= \int_0^s k(s,t)\hat{x}(s-t)\hat{x}(t)dt + \int_0^s (k(s,t) + k(s,s-t))\hat{x}(s-t)q_n(t)dt + \int_0^s k(s,t)q_n(s-t)q_n(t)dt.$$

Because $k(s,t), k(s,s-t) \in L^2[0,2]$ and $q_n \rightharpoonup 0$ the second and third integral vanish in $L^2[0,1]$ as $n \to \infty$. Hence $F(x_n) \to F(\hat{x})$ in $L^2[0,2]$. Thus, although the sequence $\{x_n\}$ does not converge, its images $\{F(x_n)\}$ do converge. Hence this sequence satisfies condition (2.26). Because $\hat{x}$ was chosen arbitrarily, F is everywhere locally ill-posed. □

This proposition shows the necessity of regularizing the autoconvolution equation. We will use an algorithm based on the linearized operator equation. For this we have the following lemma.

Lemma 2.29:

The inverse operator $F'(x_0)^{-1}$ of the linearized autoconvolution operator $F'(x_0)x = \tilde{y}$ is discontinuous for every $x_0 \in L^2[0,1]$.

Proof. The Fréchet derivative $F'(x_0)$ is compact for all $x_0 \in L^2[0,1]$ (cf lemma 2.23). Because of Proposition A.3 and Remark 4 in [9, p. 538f] $F'(x_0)^{-1}$ is not continuous. □

This is a very remarkable result. We could not find other examples of an operator which is non-compact but has an everywhere compact Fréchet-derivative.

CHAPTER

THREE

REGULARIZATION

Table of contents

Solutions to inverse problems of the type $F(x) = y$ as in (2.1) are, especially in the case of nonlinear problems, usually retrieved through an optimization approach, where the residual

$$||F(x) - y||_Y^2 \tag{3.1}$$

is minimized. The index δ denotes measured and therefore noisy data. It is often presumed that there is an estimate $||y - y^\delta|| \leq \delta$ for some preferably small $\delta > 0$, but we do not have such a δ for our problem. Due to the ill-posedness of the considered problems, solutions gained with method (3.1) are very bad or even useless as the algorithms do not necessarily converge to anything near to the actual solution. Even if they do, the result is often very inaccurate and consists of a lot of spikes. To overcome these effects, the functional (3.1) is extended to

$$||F(x) - y||_Y^2 + \alpha\Omega(x), \tag{3.2}$$

where α is the regularization parameter and $\Omega(x)$ a regularization functional. In most cases $\Omega(x)$ takes into account a-priori information of the solution, for example smoothness conditions $\Omega(x) = ||x'||_X^2$ for the first derivative $x'(t) := \frac{dx(t)}{dt}$, $\Omega = ||x''||_X^2$ for the

second derivative $x''(t) := \frac{d^2x(t)}{dt^2}$ or a total variation constraint $\Omega(x) = \|x\|_{BV}$. Another common functional is $\Omega(x) = \|x - x_0\|_X^2$, where a reference function x_0 or a good approximation of it is known a-priori. Another possible choice is $\Omega(x) = \|x\|_X^2$, known as Tikhonov regularization (cf [15]).
The role of α is to tune the balance between x being a low residual solution and being smooth. If α is too small, it fails to inhibit the ill-posedness effects. Is it chosen too large, the solution does not approximate the data anymore. Many mathematically justified rules to choose α properly have been introduced. Two of them and a new approach will be explained in detail in section 3.4.

3.1 Discretization

As computers cannot deal with infinite dimensional information and measurements are only available for certain frequencies, a discretization of the problem is necessary. This can also be seen as a first regularization step. The larger the number of supporting points, the more ill-posedness effects are amplified. Therefore, even if measurements are provided on a fine grid, they should be taken at a much more coarse grid. This of course also reduces the computational effort significantly. Vectors and matrices resulting from a discretization will be underlined.

The function values of x are to be reconstructed at N supporting points t_n which will be chosen equidistantly in an interval $[t_l, t_u]$. The discrete signal is denoted by

$$\underline{x} = (x_n)_{n=1}^N = (x(t_n))_{n=1}^N = (|x(t_n)|e^{i\varphi_x(t_n)})_{n=1}^N \tag{3.3}$$

for $n = 1 \dots N$ and $t_n = t_l + (n-1)\Delta t$ with $\Delta t = \frac{t_u - t_l}{N-1}$.
The notation for the output signal is analogous,

$$\underline{y} = (y_m)_{m=1}^{2N-1} = (y(s_m))_{m=1}^{2N-1} = (|y(s_m)|e^{i\varphi_y(s_m)})_{m=1}^{2N-1} \tag{3.4}$$

for $m = 1 \dots 2N-1$ and $s_m = 2t_l - t_{cw} + (m-1)\Delta t$.
The kernel takes the form

$$\underline{K} = k_{m,n} = k(s_m, t_{cw}, t_n), \qquad m = 1, 2, \dots, 2N-1; \;\; n = 1, 2, \dots, N. \tag{3.5}$$

From (1.24) we get

$$k(s_m, t_{cw}, t_n) = \frac{\mu_0 cL}{2} \frac{s_m}{n(s_m)} \chi^{(3)}(s_m, -t_{cw}, t_n, s_m + t_{cw} - t_n) e^{i(\Delta\vec{k}_\xi \xi + \Delta\vec{k}_\eta \eta + \Delta\vec{k}_\zeta \frac{L}{2})} sinc(\Delta\vec{k}_\zeta \frac{L}{2}) \tag{3.6}$$

where the refractive index n is calculated from the Sellmeier equation (1.31) using the values in table 1.1, the k-vectors are calculated from (1.25)-(1.29) (with $\omega := s_m$, $\hat{\omega} := t_n$, $\omega_{cw} := t_{cw}$) and

$$\chi^{(3)}(s_m, -t_{cw}, t_n, s_m + t_{cw} - t_n) = \chi^{(1)}(s_m)\chi^{(1)}(-t_{cw})\chi^{(1)}(t_n)\chi^{(1)}(s_m + t_{cw} - t_n) \tag{3.7}$$

with $\chi^{(1)}(\cdot)$ from (1.34). The remaining parameters are either natural constants (μ_0, c) or constants depending on the measurement setup (L, ξ, η).

The autoconvolution operator F is discretized using the rectangular rule. That way,

$$y_m = \sum_{j=1}^{N} k(s_m, t_{cw}, t_j) x(t_j) x(s_m + t_{cw} - t_j) \Delta t. \tag{3.8}$$

Because of the finite support of x, $x(s_m + t_{cw} - t_j) = 0$ for $s_m + t_{cw} - t_j < t_l$ and $s_m + t_{cw} - t_j > t_u$. The complete operator can be written as a multiplication of a matrix $\underline{F}(\underline{x}) \in \mathbb{C}^{2N-1 \times N}$ with the vector $\underline{x}$,

$$\underline{F}\,\underline{x} = \Delta t \begin{pmatrix} k_{1,1}x_1 & 0 & \dots & 0 & 0 \\ k_{2,1}x_2 & k_{2,2}x_1 & \dots & 0 & 0 \\ & \ddots & \ddots & & \vdots \\ k_{N-1,1}x_{N-1} & k_{N-1,2}x_{N-2} & \dots & k_{N-1,N-1}x_1 & 0 \\ k_{N,1}x_N & k_{N,2}x_{N-1} & \dots & k_{N,N-1}x_2 & k_{N,N}x_1 \\ 0 & k_{N+1,1}x_N & \dots & k_{N+1,N-1}x_3 & k_{N+1,N-1}x_2 \\ \vdots & & \ddots & \ddots & \\ 0 & 0 & \dots & k_{2N-2,N-1}x_N & k_{2N-2,N}x_{N-1} \\ 0 & 0 & \dots & 0 & k_{2N-1,N}x_N \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N-1} \\ x_N \end{pmatrix} = \underline{y}. \tag{3.9}$$

To increase computational efficiency, $\underline{F}$ can be split into two matrices $\hat{\underline{F}} \in \mathbb{C}^{2N-1 \times N}$ and $\underline{K} \in \mathbb{C}^{2N-1 \times N}$ such that

$$\underline{F} = \hat{\underline{F}} \circ \underline{K} \tag{3.10}$$

where $\circ$ denotes element by element multiplication. Thus $\hat{\underline{F}}$ takes the x-dependent entries and $\underline{K}$ consists of the kernel values $k_{m,n}$ and the factor Δt. This way, $\underline{K}$ has to be computed only once for a given discretization grid which saves a lot of time. The convolution

$$\underline{F}\,\underline{x} = \underline{y} \tag{3.11}$$

itself is then rather fast as it is only takes N^2 multiplications to calculate $\underline{F}$ and a matrix times vector multiplication to get the convolved signal $\underline{y}$.
If the measurements are not spaced equidistantly, a constant factor Δt is not applicable. The information about the spacings can then be included in the matrix $\underline{K}$ element by element.

The Fréchet derivative $F'(x)$ (2.23) is discretized in a similar way. The m-th entry $(\underline{F}(\underline{x})h)_m$ of the derivative applied to a function $\underline{h}(t)$ is

$$(\underline{F}'(\underline{x})\underline{h})_m = \sum_{j=0}^{N} (k(s_m, t_{cw}, t_j) + k(s_m, t_{cw}, s_m + t_{cw} - t_j)) x(s_m + t_{cw} - t_j) h(t_j) \Delta t. \tag{3.12}$$

Being a linear operator, $\underline{F}'(\underline{x}) \in \mathbb{C}^{2N-1 \times N}$ can of course be written in matrix form. Using the matrices from (3.10) we have

$$\underline{F}'(\underline{x})\underline{h} = \hat{\underline{F}} \circ (\underline{K} + \hat{\underline{K}})\underline{h} \tag{3.13}$$

with $\hat{\underline{K}} = \Delta t \cdot k(s_m, t_{cw}, s_m + t_{cw} - t_n)$, $m = 1, 2, \ldots, 2N-1$, $n = 1, 2, \ldots, N$. Implementing this, only $2N^2$ kernel evaluations are needed for each grid and problem.
Its adjoint $\underline{F}'(\underline{x})^* \in \mathbb{C}^{N \times 2N-1}$ reads as

$$(\underline{F}'(\underline{x})^*\underline{h})_n = \sum_{j=0}^{2N-1} \overline{(k(s_j, t_{cw}, t_n) + k(s_j, t_{cw}, s_j + t_{cw} - t_n))}\,\overline{x(s_j + t_{cw} - t_n)}h(t_n)\Delta t. \tag{3.14}$$

It has not to be generated separately, because $\underline{F}_x^* = \underline{F}_x^H$ with $\cdot^H$ standing for transposed and complex conjugated matrix.

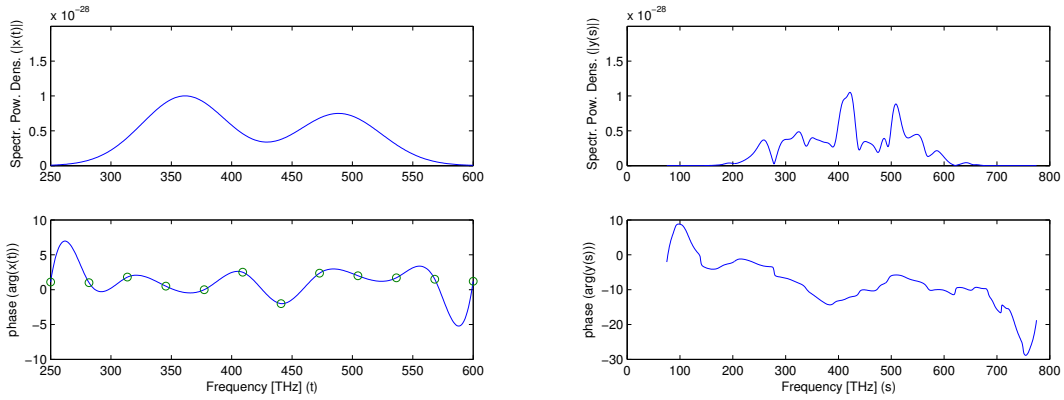
3.2 Simulation

In order to test the algorithm and describe its properties, artificial data had to be created. As standard example, the fundamental pulse was chosen as in Fig. 3.1 with a lower boundary of the frequency $t_l = 250THz$ and an upper boundary $t_u = 600THz$. This ensures that the kernel is continuous (cf remark 1.2). The frequency of the cw-pulse was set to be $t_{cw} := \frac{1}{2}(250THz + 600THz) = 425THz$. The absolute values are calculated explicitly for each frequency t from

$$|x(t)| = (e^{-\frac{(t-0.85t_{cw})^2}{\sigma^2}} + 0.748e^{-\frac{(t-1.15t_{cw})^2}{\sigma^2}}) \cdot 10^{-28} \tag{3.15}$$

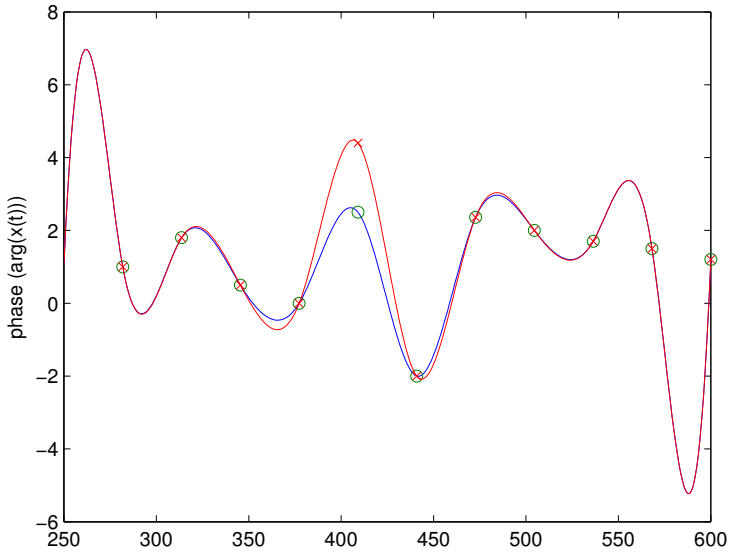
with $\sigma := 50THz$. The phase was derived from interpolation of the spline defined by the twelve points $[(t_1, 1.1)\ (t_2, 1)\ (t_3, 1.8)\ (t_4, 0.5)\ (t_5, 0)\ (t_6, 2.5)$
$(t_7, -2)\ (t_8, 2.36)\ (t_9, 2)\ (t_{10}, 1.7)\ (t_{11}, 1.5)\ (t_{12}, 1.3)]$ with $t_j = t_l + (j-1)\frac{t_u - t_l}{11}$.
Fig. 3.1 shows this pulse which we will later refer to as pulse I. The main problem with it is the rather large phase at lowest and highest frequencies, where at the same time the absolute values are very small. As it can be seen in (3.9), at the edges of the domain the convoluted signal depends on only a few values, in other words it carries few information about the input pulse. Therefore, it is more sensitive to the phenomenon of incorrectness there than in the middle of the frequency interval. The complex valued exponential function is 2π-periodical. This causes additional ambiguities. Let $\varphi_x(t_0) = 0$. Then any $\tilde{\varphi}_x(t_0) = k2\pi$, $k \in \mathbb{Z}$, cannot be distinguished from $\varphi_x(t_0)$. Therefore errors are likely occur when the phase jumps between different periods. This is illustrated in the left image of Fig. 3.3, where the phase of the pulse was plotted using only one period. In the other figures, the phase is built from this representation by adding or substracting 2π to $\varphi_x(t_{j+1})$ whenever the difference between two neighbouring values $\varphi_x(t_j)$ and $\varphi_x(t_{j+1})$ $(j = 1 \ldots N-1)$ is lower than $-\pi$ or greater than π, respectively.

Figure 3.1: Standard fundamental pulse for the simulations and its autoconvolved pulse. In the upper panels the absolute values are plotted while the lower panel shows the phase



To create artificial data for the measurements, the pulse sampled at $\tilde{N}$ points is convolved using formula (3.11), resulting in the autoconvolution signal y at $2\tilde{N}-1$ supporting points which is split into absolute values $|\underline{y}|$ and phase $\underline{\varphi}_y$. The absolute values $|\underline{x}|$ of the fundamental pulse were again sampled at $N << \tilde{N}$ points to complete the measurement situation. The supporting points $t_1, t_2, \ldots, t_N$ corresponding to N are the points where the unknown phase is to be reconstructed. The reason why N must be much smaller than $\tilde{N}$ (also N must not be a divisor of $\tilde{N}$) is an effect called *inverse crime*. In case of an inverse crime, the reconstruction from artificially created data will always be good, regardless of the ill-posed character of the problem. This effect does not occur for real measurements.
Because measurements of real life processes are never exact, a normaly distributed error of magnitude δ is added to the data so that $|\underline{y}|^\delta \sim \mathcal{N}(|\underline{y}|, (\delta|\underline{y}|)^2)$, $\underline{\varphi}_y^\delta \sim \mathcal{N}(\underline{\varphi}_y, (\delta\underline{\varphi}_y)^2)$ and $|\underline{x}|^\delta \sim \mathcal{N}(|\underline{x}|, (\delta|\underline{x}|)^2)$. This is equivalent to a relative error of $\delta\%$. Finally, $|\underline{y}|^\delta$ and $\underline{\varphi}_y^\delta$ are interpolated at the $2N-1$ supporting points corresponding to the convolved frequencies of $t_1, t_2, \ldots, t_N$. Thus a complete set of measurements is available that avoids the inverse crime. To compute the kernel (3.6), the following numerical values have been used: $\mu_0 = 4\pi \cdot 10^{-7} \frac{kg\ m}{A^2 s^2}$, $c = 299792458 \frac{m}{s}$, $\alpha = 0.02 rad$, $\xi = 10^{-5} m$, $\eta = 10^{-5} m$, $L = 10^{-4} m$. Once $|\underline{y}|^\delta$, $\underline{\varphi}_y^\delta$, $|\underline{x}|^\delta$ and the kernel matrices $\underline{K}$ and $\underline{\tilde{K}}$ from (3.10) and (3.13) respectively have been computed, the regularization algorithm can be started. The parameters $\tilde{N}$, N, δ and the regularization parameter α that will be introduced in the next section will be given for every figure showing reconstructed phases. Instead of the angular frequencies $\omega = 2\pi f$, the real frequency f will be marked at the x-axis to improve readability.

Figure 3.2: comparison of the two standard phases for simulations



To show the difficulties resulting from the periodic behaviour of the pulses, we introduce a second pulse II where in comparison to the first phase only the sixth supporting point was changed. Instead of $(t_6, 2.5)$, $(t_6, 4.4)$ was used so that this value is now larger than π. The resulting phase in comparison with the original one is shown in Fig. 3.2. The absolute values remain the same.

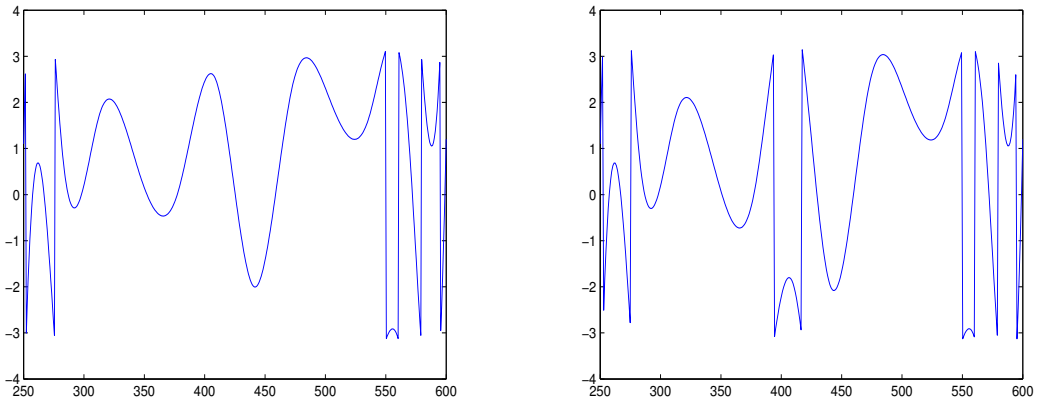
To point out the different periodical behaviour, Fig. 3.3 shows both phases plotted in the period $[-\pi, \pi]$. The increased value of the right phase in comparison to the phase on the left leads to jumps in the middle part of the phase in this representation. This will cause serious problems as we will see later.

3.3 Regularization algorithm

To realize the minimization, only iterative methods were taken into account. Attempts with MATLAB®-solvers failed as well as Landweber-Methods (see [22, chapter 2,3]). Better results could be observed with Newton-type methods, especially with the Levenberg-Marquardt algorithm (LMA) [23], [27] on which the presented algorithm is based. We will minimize

$$\|\underline{F}\,\underline{x} - \underline{y}\|_2^2 \tag{3.16}$$

Figure 3.3: standard phases plotted in the $[-\pi, \pi]$-period showing jumps which will prove to cause significant problems for the reconstruction



as the discretized analogon to (3.1). The euclidean norm is denoted as $\|\cdot\|_2$.
One explanation of the LMA is that it interpolates between the Gauss-Newton algorithm (GNA) and the Landweber iteration (LWI) (also called method of steepest descent) to globalize convergence of the GNA.
Given an iterate $\underline{x}_k$, the next one is obtained by

$$\underline{x}_{k+1} = \underline{x}_k + \gamma(\underline{F}'(\underline{x}_k)^*\underline{F}'(\underline{x}_k) + \alpha_k I)^{-1}\underline{F}'(\underline{x}_k)^*(y^\delta - \underline{F}(\underline{x}_k)), \tag{3.17}$$

I being the identity matrix. For the iteration procedure see for example [22, chapter 4]. The parameter α_k balances GNA and LWI. For $\alpha_k \to \infty$, (3.17) turns into the Landweber iteration and reads as

$$\underline{x}_{k+1} = \underline{x}_k + \gamma\underline{F}'(\underline{x}_k)^*(\underline{y}^\delta - \underline{F}(\underline{x}_k)). \tag{3.18}$$

On the other hand, for $\alpha_k \to 0$, (3.17) becomes

$$\underline{x}_{k+1} = \underline{x}_k + \gamma(\underline{F}'(\underline{x}_k)^*\underline{F}'(\underline{x}_k))^{-1}\underline{F}'(\underline{x}_k)^*(\underline{y}^\delta - \underline{F}(\underline{x}_k)), \tag{3.19}$$

the Gauss-Newton algorithm. Since the GNA converges much faster than the LWI when the iterates are close to a minimum usually $\alpha_k \to 0$ is chosen. Several methods how to chose α_k have been introduced (cf [22, chapter 4]), but here a rather simple approach

$$\alpha_k = \alpha_0 \cdot q_\alpha^k \tag{3.20}$$

is taken, where α_0 is the initial value and $0 < q_\alpha < 1$ is a damping factor.

The purpose of the step length γ in (3.17) is simply to ensure a decrease in the residual, so that

$$\|\underline{y}^\delta - \underline{F}(\underline{x}_{k+1})\|_2 \leq \|\underline{y}^\delta - \underline{F}(\underline{x}_k)\|_2. \tag{3.21}$$

Similar to α_k, $\gamma = \gamma_0 \cdot q_\gamma^l$ with $\gamma_0 = 1$, $0 < q_\gamma < 1$ and
$l = \min\{l \in \mathbb{N} : \|\underline{y}^\delta - \underline{F}(\underline{x}_{k+1})\|_2 < \|\underline{y}^\delta - \underline{F}(\underline{x}_k)\|_2\}$. It turned out that in most cases $l = 0$.
The choice of the initial value $\underline{x}_0$ will be discussed later.
Finally, the iteration is stopped either if no significant decrease in the residual is gained,

$$\|\underline{y}^\delta - \underline{F}(\underline{x}_{k+1})\|_2 \geq q\|\underline{y}^\delta - \underline{F}(\underline{x}_k)\|_2, \tag{3.22}$$

$q < 1$ but q close to one or if a maximal number of iterations $k_{\max}$ has been performed.
The main flaw of this algorithm is that it only minimizes $\|y^\delta - F(x)\|^2$, and therefore has no smoothing properties. Fig. 3.4 shows the best regularization (for the choice of α_0, see section 3.4). Although the solutions fit the original pulse quite well, there are a lot of spikes that should not appear for a real pulse as they have to be smooth. The effect gets worse when the reconstruction is performed on a larger amount of supporting points as illustrated in Fig. 3.5. In both cases all data carried a relative error of 1%.

Algorithmus 3.1 Calculating a regularized solution with LMA minimizing (3.16) by minimizing (3.23) in each iteration

input: measured data $\underline{y}^\delta$ and $|\underline{x}|^\delta$, initial value $\underline{\varphi}_0$, α_0, k_{max}.
set $k := 0$, $l := 0$, $\underline{x}_0 = |\underline{x}|^\delta e^{i\underline{\varphi}_0}$, $\alpha = \alpha_0$
while $\|\underline{y}^\delta - \underline{F}\underline{x}_k\|_2 < q\|\underline{y}^\delta - \underline{F}\underline{x}_k\|_2$ AND $k < k_{max}$ **do**
  solve $(\underline{F}'(\underline{x}_k)^*\underline{F}'(\underline{x}_k) + \alpha I)p = \underline{F}'(\underline{x}_k)^*(\underline{y}^\delta - \underline{F}(\underline{x}_k))$
  while $\|\underline{F}(\underline{x}_k + \gamma\underline{p})\|_2 > \|\underline{F}(\underline{x}_k))\|_2$ **do**
    $l = l + 1$
    $\gamma = q_\gamma^l$
  end while
  set $\underline{x}_{k+1} = \underline{x}_k + \gamma\underline{p}$, $l = 0$, $\alpha = \alpha \cdot q_\alpha$, $k = k + 1$
end while
return $\underline{x}_{\text{end}}$

Because the phase is periodic, it can jump between different periods. Being assembled from the $[-\pi, \pi]$-period, this occurs if caused by errors in the solution, the critical difference indicating a change of the period in the assembling algorithm is reached where there should not be such a difference. An example of this effect is given in Fig. 3.6. The reconstruction of the phase is rather good except for the boundaries, but at about $\omega = 2600THz$ the phase jumps from one period to the one below where that should not happen.

Figure 3.4: Best regularized solution retrieved with Alg. 3.1 showing non-smooth reconstruction of both absolute value and phase
$\tilde{N} = 467$, $N = 130$, $\delta = 1\%$, $\alpha_0 = 3.1421$

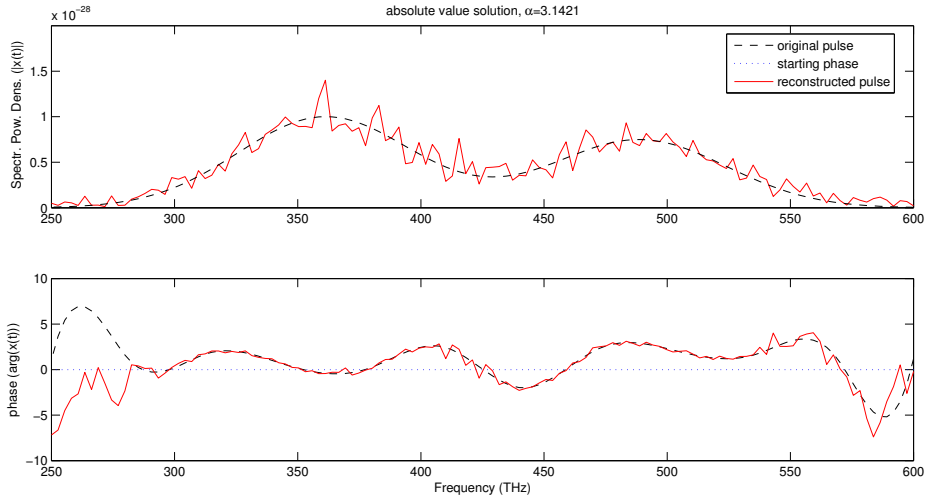


Figure 3.5: Best regularized solution retrieved with Alg. 3.1 on a larger amount of supporting points. The absolute values show much higher spikes than the phase. $\tilde{N} = 867$, $N = 330$, $\delta = 1\%$, $\alpha_0 = 7.5775$

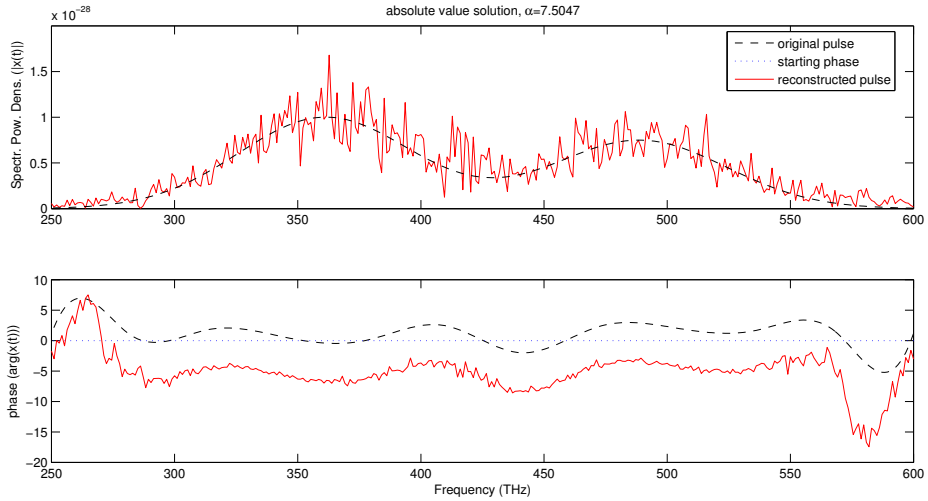
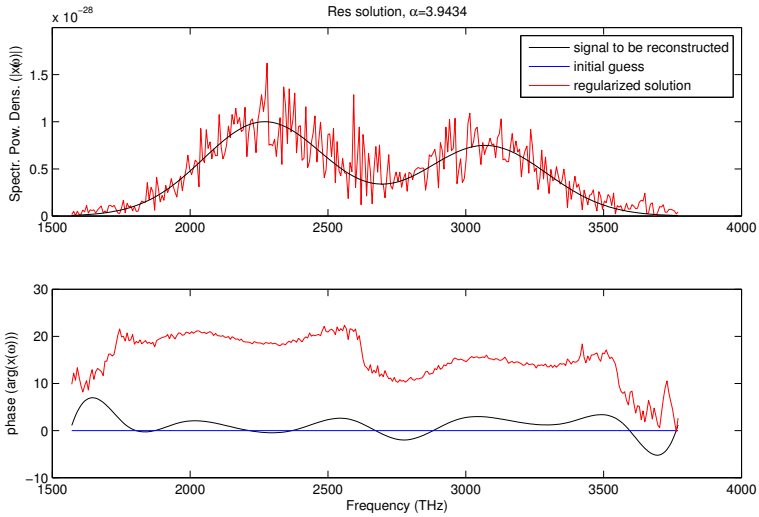


Figure 3.6: jump in the period of the regularized phase, $\tilde{N} = 867$, $N = 330$, $\delta = 1\%$, $\alpha_0 = 3.9434$



To get smoother regularized solutions, we will analyze a second explanation for the LMA. As described in [22, chapter 4], the LMA minimizes (3.16) within a trust region. This is equivalent to minimizing the linearized functional

$$\|\underline{y}^{\delta} - \underline{F}(\underline{x}_k) - \underline{F}'(\underline{x}_k)\underline{z}\|_2^2 + \alpha_k\|\underline{z}\|_2^2 \tag{3.23}$$

around the current iterate $\underline{x}_k$ where α_k is the corresponding Lagrange parameter. This can be interpreted as adding a penalty term $\alpha_k\|\underline{z}\|_2^2$ to the linearized functional, or as Tikhonov-regularization on it. From the latter one, the idea was derived to add a smoothing penalty to (3.23) by changing the functional to

$$\|\underline{y}^{\delta} - \underline{F}(\underline{x}_k) - \underline{F}'(\underline{x}_k)\underline{x}_k\underline{z}\|_2^2 + \alpha_k\|\underline{L}\underline{z}\|_2^2 \tag{3.24}$$

where $\underline{L}\underline{z}$ represents the second derivative of $\underline{z}$.
The second derivative at a point $z(t_j)$ can be approximated as

$$z''(t_j) = \frac{z(t_{j-1}) - 2z(t_j) + z(t_{j+1})}{\Delta t^2}. \tag{3.25}$$

Because Δt is of a magnitude of $10^{14}Hz$ and all t_j are distributed equidistantly, Δt^2 has been excluded from the matrix $\underline{L}$. Thus $\underline{L}$ becomes a tridiagonal $N \times N$-matrix with 2 on the main diagonal and -1 on its side diagonals. The factor Δt^2 is then included in the parameter α_k, such that $\alpha_k = \frac{\tilde{\alpha}_k}{\Delta t^2}$ where $\tilde{\alpha}_k$ is the real regularization parameter. However

α_k will be the parameter used in the simulation as its numerical values are much more handy than those of $\tilde{\alpha_k}$.

Using this matrix $\underline{L}$, the linearized functional is now minimized in every step around the current iterate, but the penalty oppresses non-smooth solutions depending on the magnitude of α_k. The iteration procedure (3.17) thus becomes

$$\underline{x}_{k+1} = \underline{x}_k + \gamma(\underline{F}'(\underline{x}_k)^*\underline{F}'(\underline{x}_k) + \alpha_k\underline{L}^*\underline{L})^{-1}\underline{F}'(\underline{x}_k)^*(\underline{y}^\delta - \underline{F}(\underline{x}_k)). \tag{3.26}$$

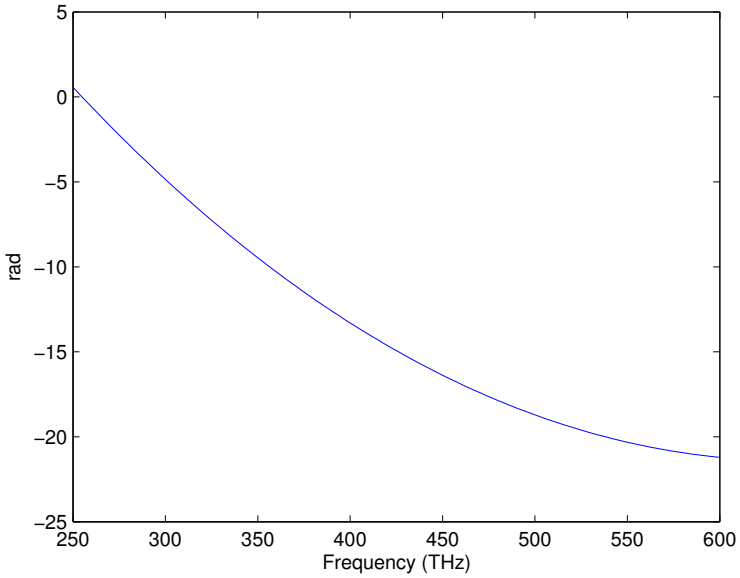
In simulations, a deacreasing sequence of α_k leads to a loss of smoothing properties as the regularization term in (3.24) vanishes for $k \to \infty$. Therefore, $\alpha = \alpha_k$ is chosen constant for all iterations. The algorithm to find a regularized solution for a given parameter α is shown in Alg. 3.2

Algorithmus 3.2 Calculating a regularized solution using a smoothing penalty minimizing (3.24) in each iteration

input: measured data $\underline{y}^\delta$ and $|\underline{x}|^\delta$, initial value $\underline{\varphi}_0$, α, k_{max}.
set $k := 0$, $l := 0$, $\underline{x}_0 = |\underline{x}|^\delta e^{i\underline{\varphi}_0}$
while $\|\underline{y}^\delta - \underline{F}\underline{x}_k\|_2 < q\|\underline{y}^\delta - \underline{F}\underline{x}_k\|_2$ AND $k < k_{max}$ **do**
  solve $(\underline{F}'(\underline{x}_k)^*\underline{F}'(\underline{x}_k) + \alpha\underline{L}^*\underline{L})\underline{p} = \underline{F}'(\underline{x}_k)^*(\underline{y}^\delta - \underline{F}(\underline{x}_k))$
  while $\|\underline{F}(\underline{x}_k + \gamma\underline{p})\|_2 > \|\underline{F}(\underline{x}_k))\|_2$ **do**
    $l = l + 1$
    $\gamma = q_\gamma^l$
  end while
  set $\underline{x}_{k+1} = \underline{x}_k + \gamma\underline{p}$, $l = 0$, $k = k + 1$
end while
return $\underline{x}_{\text{end}}$

As any iterative algorithm, the LMA requires an initial value $\underline{x}_0$ to start the iteration. Although the LMA is robust, it only finds a local minimum, not necessarily the global minimum. Therefore, the choice of the starting point is crucial. Because the absolute value of the fundamental pulse is available as measurement, it is obvious to use them as the initial guess, setting $|\underline{x}_0| := |\underline{x}|^\delta$. Thus, it remains to find a good approximation for the unknown phase. Comparing convolved phases φ_y from simulations with their fundamental phases φ_x (see Fig. 3.1), a certain alikeness can be seen. A simple analytic example will show that at least at the boundaries the phase of x can be guessed from

Figure 3.7: Phase of the kernel for $s = 2*\pi*475THz$ and $2*\pi*250THz \leq t \leq 2*\pi*600THz$



the phase of y. Taking equation (1.44) and assuming that for small s $\varphi_x(s-t) \approx \varphi_v(t)$,

$$\begin{aligned} y(s) &= \int_0^s |k(s,t)|e^{i\varphi_k(s,t)}|x(t)|e^{i\varphi_x(t)}|v(s-t)|e^{i\varphi_x(s-t)}dt \\ &= \int_0^s \underbrace{|k(s,t)||x(t)||x(s-t)|}_{=:|\tilde{k}(s,t)|} e^{i(\varphi_k(s,t)+\varphi_x(t)+\varphi_x(s-t))}dt \\ &\approx \int_0^s |\tilde{k}(s,t)|e^{i(\varphi_k(s,t)+2\varphi_x(t))}dt. \end{aligned} \tag{3.27}$$

To get an idea of the phase of the kernel, Fig. 3.7 shows $\varphi_k(s,t)$ for $s = 2\pi \cdot 475THz$ and $2\pi \cdot 250THz \leq t \leq 2\pi \cdot 600THz$. From this and (3.27), the idea was taken to choose

$$\varphi_x^0(t) := \frac{1}{2}P(\varphi_y(s)) - l(s,t) \tag{3.28}$$

with the projector $P : [2t_l - t_{cw}, 2t_u - t_{cw}] \mapsto [t_l, t_u]$ mapping the phase of y onto the interval of x and a correction term $l(s,t)$. In discrete version, P just takes every second

entry of $\underline{\varphi}_y$,

$$P(\underline{\varphi}_y) = P(\varphi_{y_1}, \varphi_{y_2}, \varphi_{y_3}, \dots, \varphi_{y_{2N-3}}, \varphi_{y_{2N-2}}, \varphi_{y_{2N-1}}) = (\varphi_{y_1}, \varphi_{y_3}, \dots, \varphi_{y_{2N-3}}, \varphi_{y_{2N-1}}). \tag{3.29}$$

The purpose of the function $l(s,t)$ is to lessen the influence of the kernel. To compare different starting phases, α and δ are set to zero, resulting in a standard Gauss-Newton algorithm. Although the LMA is less sensitive to the initial value than the GNA, it is obvious that the better the result for GNA, the better the result will be for the LMA. Four possible choices for the initial phase and the resulting reconstructed pulse are shown in Fig. 3.8. In part (a) the initial phase was set to zero, resulting in a very bad reconstruction. The first initial guess using a projection of the phase φ_y without a kernel correction and the resulting reconstructed phase is shown in (b). Although the solution roughly follows the fundamental pulse, it is still too inaccurate. The final two figures show reconstructions where phase of the kernel was included in the choice of the initial guess. In (c), instead of only projecting φ_y on the domain of x, its linear approximation $l(s)$ has been subtracted. We have

$$\varphi_x^0(t) := \frac{1}{2}(P(\varphi_y(s)) - l(s)) \tag{3.30}$$

with the least squares approximation l,

$$l(s) = \frac{1}{2}(as + b), \qquad \|\underline{\varphi}_y(s) - (as + b)\|_2 = \min_{\tilde{a},\tilde{b}\in\mathbb{R}}\{\|\underline{\varphi}_y(s) - (\tilde{a}s + \tilde{b})\|_2\}. \tag{3.31}$$

The resulting reconstruction is very accurate. A similar result was obtained in (d) where l was taken directly from the kernel phase shown in Fig. 3.7 with formula (3.28). The phase correction was computed for a fixed $s = 2\pi \cdot 475THz$ and all t_j $(j = 1 \dots N)$

$$l(t) = \frac{1}{2}\underline{\varphi}_k(2\pi \cdot 475THz, t). \tag{3.32}$$

The difference of the reconstructed phase and the original one in (d) is approximately 3π for all values. This corresponds to the non-injectivity of the autoconvolution equation (cf section 2.2.2) leading to an offset of π. The remaining 2π result from the 2π-periodicity of the phase, meaning both phases are identical. But as mentioned in section 1.2, a constant difference between reconstructed and optimal solution does not matter. Evaluating Fig. 3.8, it becomes clear that a good solution always has both good reconstruction of the absolute values and the phase, which is quite surprising because in all figures, the initial guess for the absolute values was the exact solution. This will be used as a strategy to find optimal regularized solutions. Some other starting phases like a constant non-zero function or linear guesses have been tested but no satisfying results could be obtained. Figure 3.9 shows the same series for the second fundamental pulse. For the zero-guess the algorithm stopped after a few iterations, resulting in the rather straight reconstruction that is far away from the actual solution. Again, the projection of φ_y alone (b) did not lead to a good result. Surprisingly, the two initial phases with kernel correction also

failed. An explanation can be given analyzing Fig. 3.10. The period jump in the second pulse leads to a lot of spikes in the solution. Although the reconstruction roughly follows the true solution, every spike leads to a 2π-jump when the phase is assembled. Understanding that there is no optimal choice for the starting phase, the zero-function will be taken as the initial guess for the phase. There exist regularization parameters $\alpha > 0$ for which the algorithm converges to a good solution. The zero function is also a good choice because there will be no period jumps in the beginning.

3.4 Regularization parameter

There are several approaches how to choose α in an optimal way. These can be classified into three groups

- a-priori $\alpha = \alpha(\delta)$
- a-posteriori $\alpha = \alpha(y^\delta, \delta)$
- heuristic methods $\alpha = \alpha(y^\delta)$.

A-priori information are required for theoretical purposes like convergence analysis. For practical problems either no information on δ or only a coarse estimate is available. Therefore only heuristic methods will be taken into consideration here. Unfortunately, because of the Bakushinsky-veto (cf for example [16, p. 131]), there exists no convergent regularization method where the choice of α only depends on y^δ if the inverse operator F^{-1} is unbounded. However, for practical problems they are usually sufficient.
Three rules will be examined here, namely the principle of quasioptimality, the popular L-curve method, and a new strategy inspired by observations from simulations.
All three methods are based on the same principle. Regularized solutions x_α are calculated for a series of parameters α. The best solution is derived by minimizing a functional ψ over all x_α. The difference lies in the choice of ψ.

The principle of quasioptimality (QO) was introduced by Tikhonov/Glasko in 1968 (cf [15, p. 96] and references therein). The idea is to find the regularization parameter α^*, where the regularized solutions x_α change least,

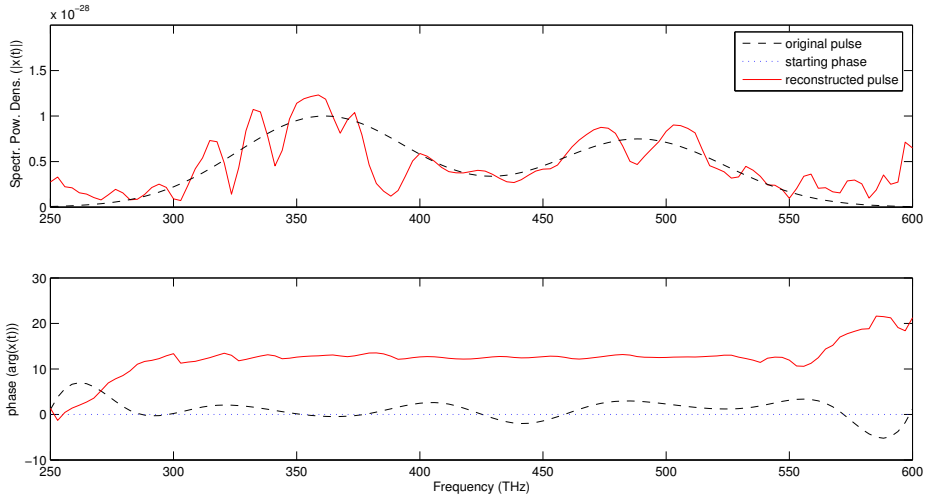
$$f(\alpha) = \|\alpha \frac{dx_\alpha}{d\alpha}\| \to \min . \tag{3.33}$$

In praxis, $x_\alpha = x_{\alpha_l}$ is computed for a sequence of α_l, usually $\alpha_l = \alpha_0 * 2^{-l}$ for $l = 1..l_{\max}$ and $\alpha_0 >> 0$. The functional (3.33) can then be replaced by

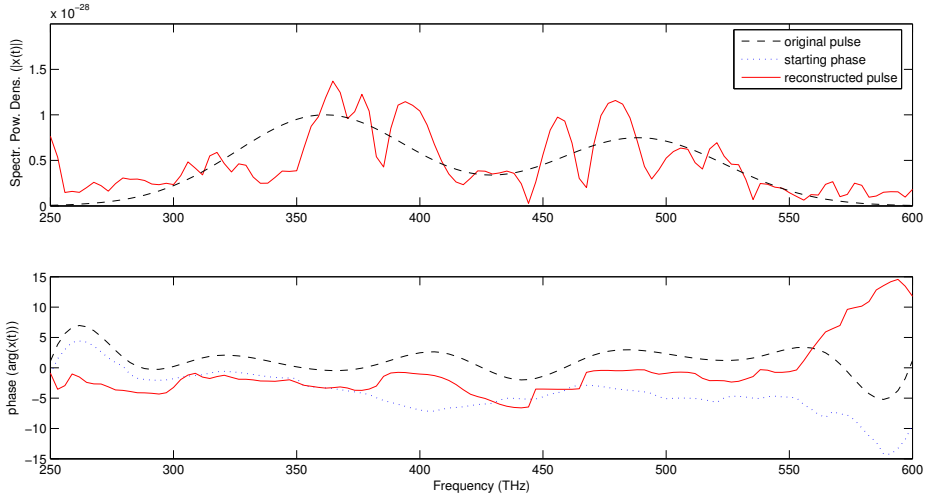
$$\psi_{\mathrm{QO}}(\alpha) = \|x_{\alpha_{l+1}} - x_{\alpha_l}\| \to \min \tag{3.34}$$

which is easy to compute.

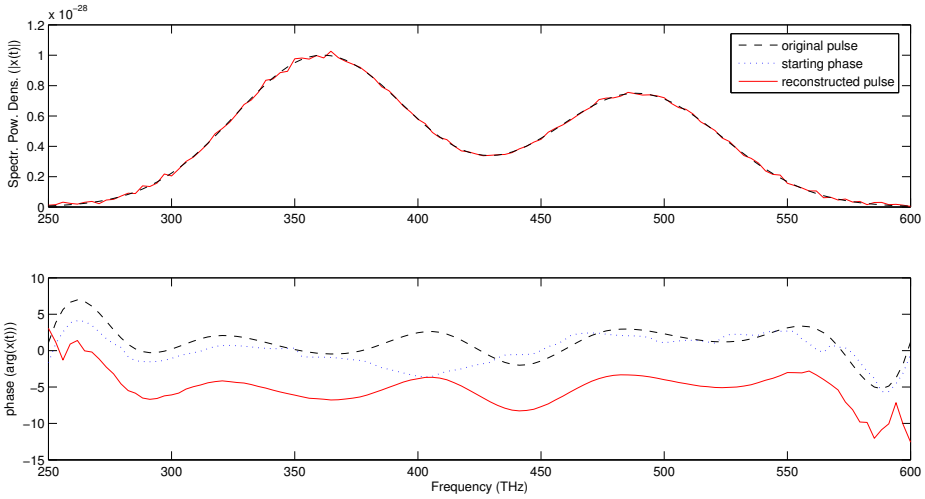
Figure 3.8: Comparison of different starting phases for the reconstruction of the first fundamental pulse, $\delta = 0$



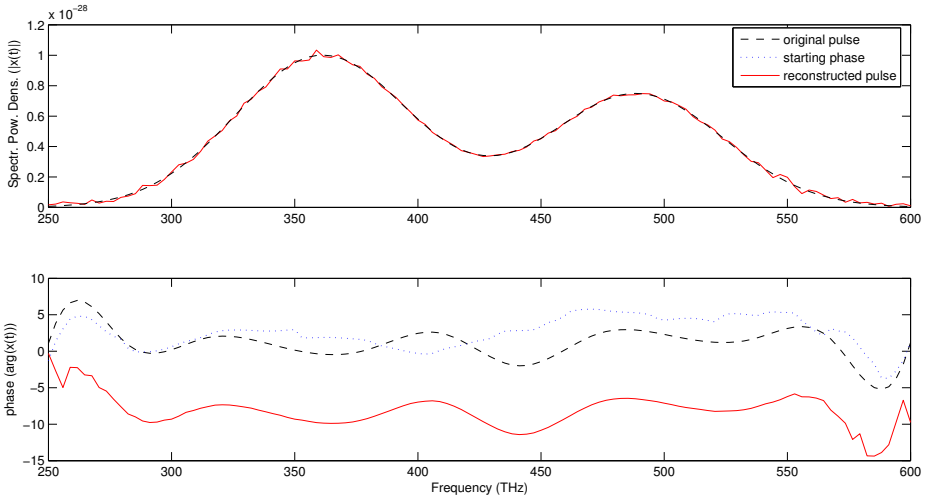
(a) constant zero initial guess, resulting in bad overall reconstruction



(b) projected SD-phase as initial guess, in the middle the reconstruction roughly follows the original pulse with a 2π-offset



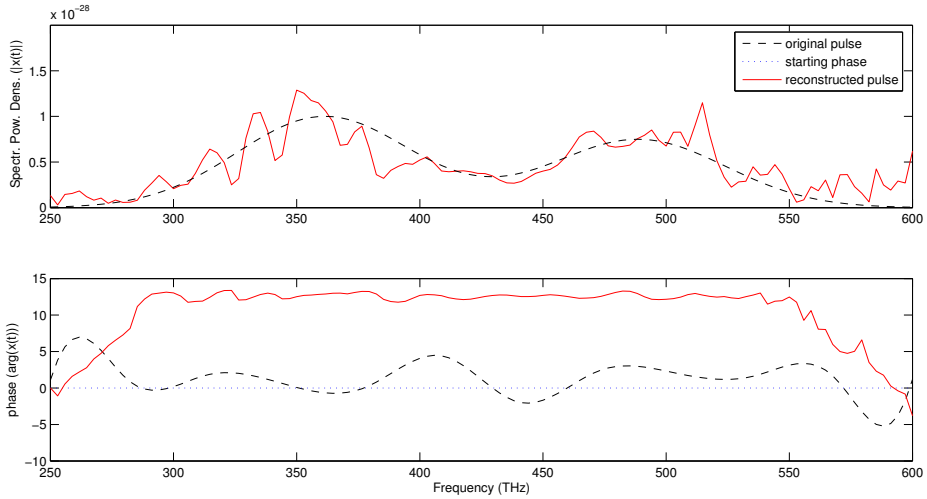
(c) projected SD-phase minus its linear approximation as initial guess, resulting in a near perfect reconstruction



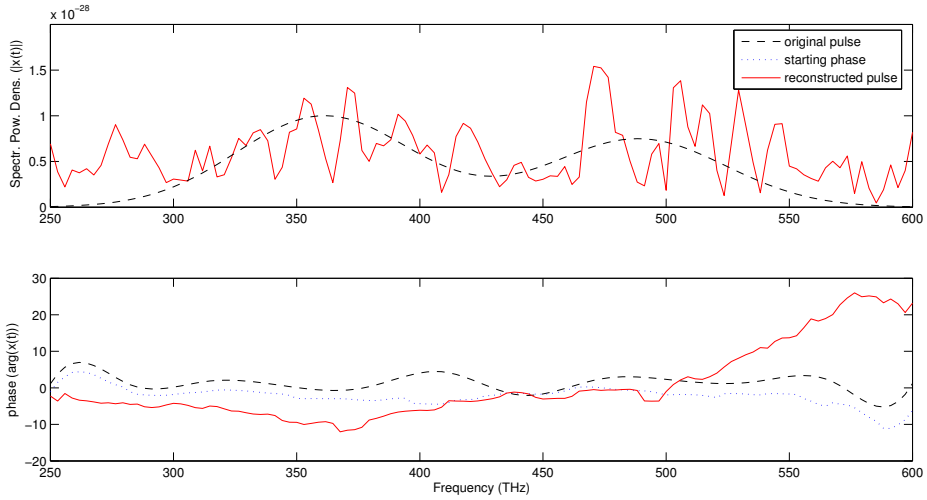
(d) projected SD-phase minus the kernel phase as initial guess, resulting in a near perfect reconstruction with an offset of 3π corresponding to the non-injectivity of the problem

Figure 3.8

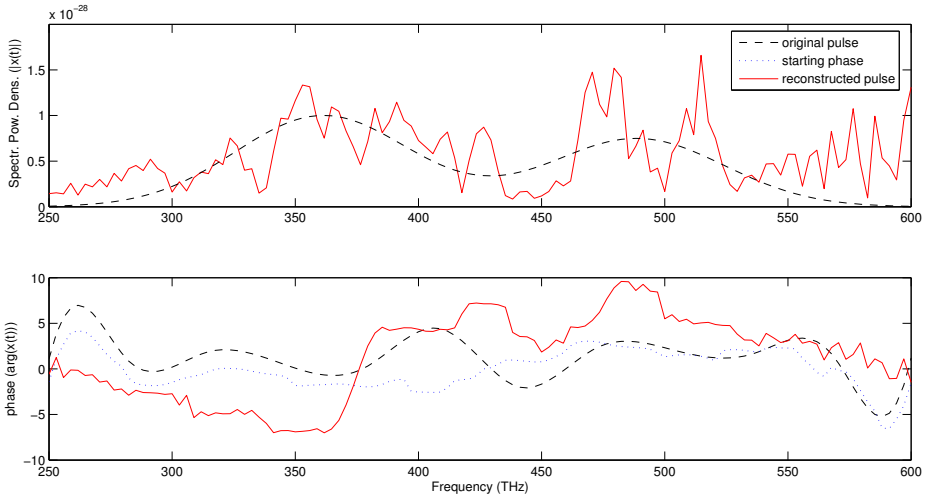
Figure 3.9: Comparison of different starting phases for the reconstruction of the second fundamental pulse, $\delta = 0$



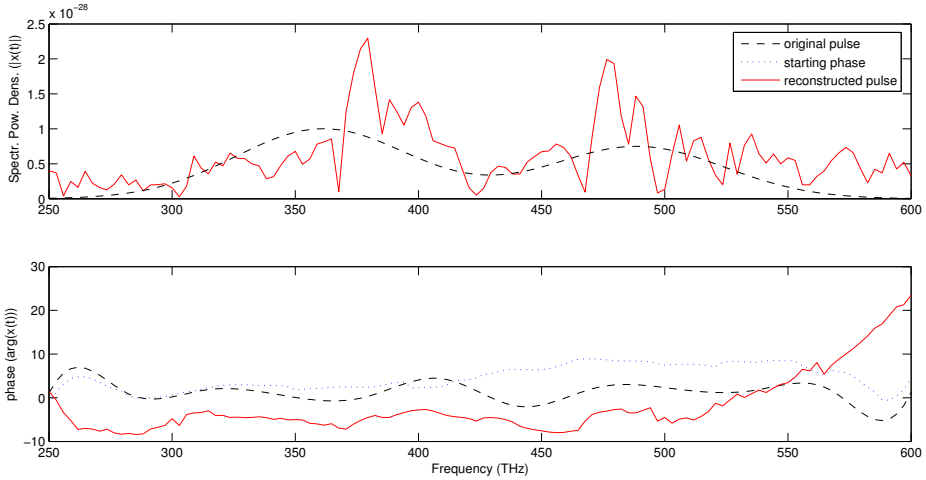
(a) constant zero initial guess, resulting in bad overall reconstruction



(b) projected SD-phase as initial guess, the reconstruction is now absolutely unfit



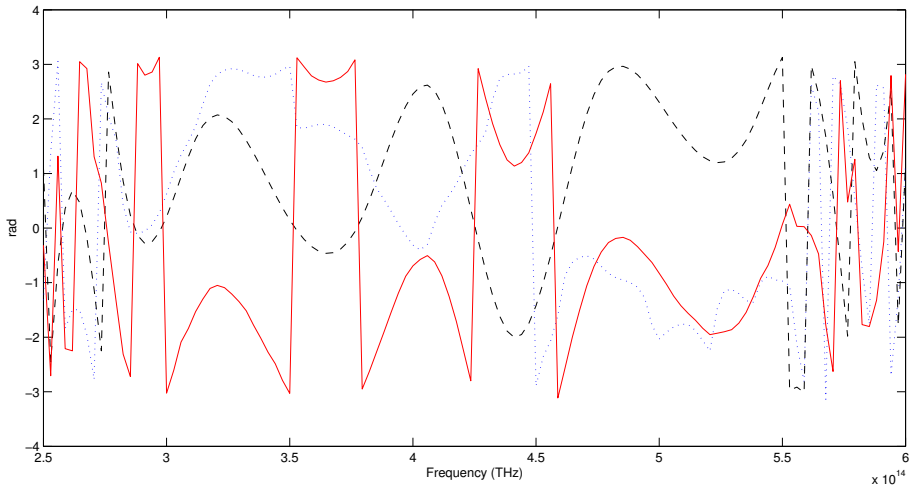
(c) projected SD-phase minus its linear approximation as initial guess, surprisingly the result is far off the fundamental pulse



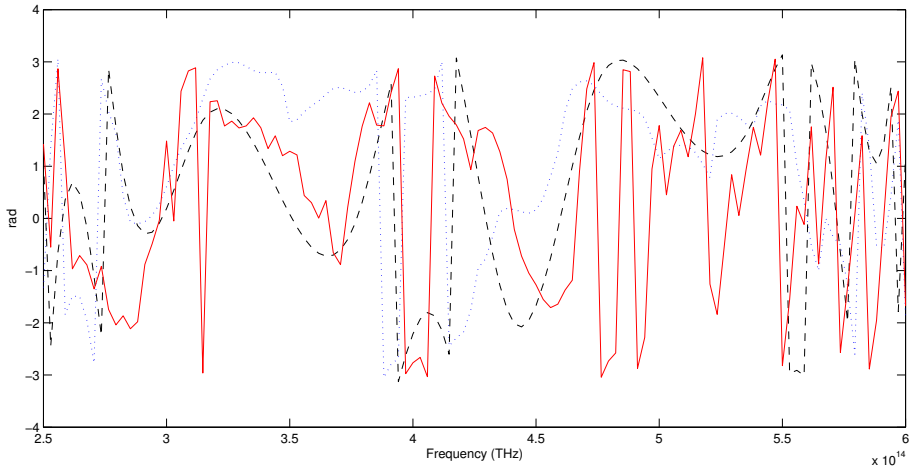
(d) projected SD-phase minus the kernel phase as initial guess, best reconstruction achieved for this phase without regularization. It is much worse than the solution for the first phase.

Figure 3.9

Figure 3.10: Comparison of the reconstructions Fig. 3.8(d) and Fig. 3.9(d) in the $[-\pi, \pi]$ period



(a) first fundamental pulse, the middle part is reconstructed perfectly. The starting phase (blue dotted line) has a single jump in this part.



(b) second fundamental pulse, the double jump in the middle in both fundamental and starting phase leads to a very spiky, inaccurate reconstruction.

The L-curve method (LC) was introduced by P. Hansen [14]. It is motivated by the observation that for small α the residual $\|y^\delta - F(x_\alpha)\|$ hardly changes while the penalty $\|Lx_\alpha\|$ increases. On the other hand, for large α $\|Lx_\alpha\|$ remains almost constant while $\|y^\delta - F(x_\alpha)\|$ increases. Both values are plotted in a log-log-coordinate system often resulting in a curve shaped like the letter L. The optimal solution for a L-curve $L(\alpha) = (log\|y^\delta - F(x_{\alpha_l})\|, log\|Lx_{alpha_l}\|)$ generated from a sequence $\alpha_0 > \alpha_1 > ... > \alpha_{l-1} > 0$ and corresponding solutions x_{α_l} is to be found in the corner of the L, where a point C is called corner, if

- $L(\alpha)$ is concave in a neighbourhood of C
- the tangent on $L(\alpha)$ has a slope of -1 in C.

This is equivalent to

$$\psi_{\mathrm{LC}}(\alpha) = \|Lx_{\alpha_l}\| \cdot \|y^\delta - Fx_{\alpha_l}\| \to \min \tag{3.35}$$

which holds the requirement for the L-curve solution.

As a third alternative, a new approach is presented here, making use of the specific situation that a part of the solution, namely the absolute values, are available as measurements. It can be observed from calculations with artificial data that, whenever the phase is reconstructed badly, the absolute values of x are also badly shaped, see for example Fig. 3.8. On the other hand, if the phase is approximated well, so are the absolute values. In section 1.6 it has been shown that the phase of x influences both phase and absolute values of y. If φ_x does not converge to the actual solution φ_x^*, neither $|y|$ nor φ_y can converge to fit the data y^δ. Because the complete pulse x is adjusted in each iteration, $|x|$ tries to compensate the effects of a diverging phase and thus also diverges. The rule to choose the regularization parameter here is simply to take the solution x_α that approximates the given data $|x^\delta|$ best. For a sequence α_l and corresponding solutions x_{α_l}

$$\psi_{\mathrm{AV}}(\alpha) = \||x_{\alpha_l}| - |x^\delta|\| \to \min \tag{3.36}$$

This rule will be called absolute value method (AVM). It proved to be the most reliable of the presented methods.

The optimal solution was calculated using Alg. 3.3. After artificial data has been created as explained in section 3.2, algorithm 3.2 is executed for a series of regularization parameters α_l. These parameters are obtained from

$$\alpha_l = \alpha_0 \cdot q^l \qquad l = 1, 2, \ldots l_{max} \tag{3.37}$$

where α_0, q and l_{max} have to be chosen appropriately. From the resulting series of reconstructed pulses the the final solution is gained via the criteria explained in this section. The algorithm can be performed again in a close neighbourhood of the best regularization parameter α^* to improve the result.

Algorithmus 3.3 Finding an optimal regularization parameter

create artificial data
input: α_0, q_α, l_{MAX}
set $l := 1$
while $l < l_{\mathrm{MAX}}$ **do**
  compute x_α for $\alpha = \alpha_0 q^l$ using Alg. 3.2
  set $l = l + 1$
end while
compute QO solution x_α^{QO} by finding the minimum of ψ_{QO}
compute L-curve solution x_α^{LC} by finding the minimum of ψ_{LC}
compute absolute value x_α^{AV} solution by finding the minimum of ψ_{AV}
optional: repeat algorithm in a neighbourhood of the solutions
return x_α^{QO},x_α^{LC},x_α^{AV}

3.5 Reconstruction results from artificial data

Using Alg. 3.3, reconstructed pulses could be calculated. All computations have been performed using MATLAB®R2007a. The regularization has been performed for fundamental pulse II and three magnitudes of the error δ. For each δ, the results obtained from the three parameter choice rules from section 3.4 are presented in Fig. 3.11, Fig. 3.12 and Fig. 3.13 for $\delta = 0.1\%$, $\delta = 1\%$ and $\delta = 5\%$, respectively. Part (a) shows the L-curves for each error. None of them is in any way shaped like an L. The corresponding solution, shown in part (b) of the figures is therefore absolutely not acceptable. The L-curve method is not a good choice for our problem. One reason might be that the L-curve method is usually applied when a functional

$$\|\underline{\underline{F}}\,\underline{x} - \underline{y}\|_2^2 + \alpha\|\underline{\underline{L}}\,\underline{x}\|_2^2 \tag{3.38}$$

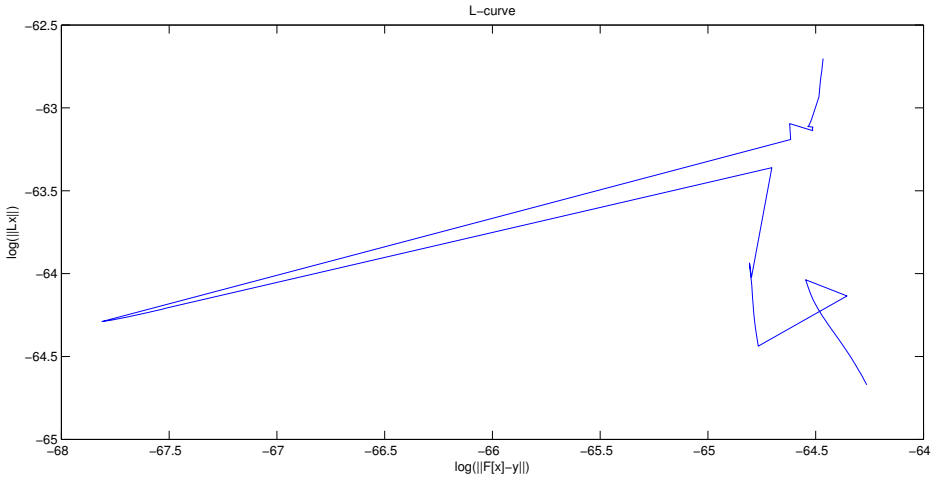
is minimized. However, we minimize only $\|\underline{\underline{F}}\,\underline{x} - \underline{y}\|_2^2$ and apply the smoothing penalty within the minimization. The solutions computed with the quasi-optimality criteria are shown in part (c) of the figures. In all three cases the results are very bad. Thus, also quasi optimality is not a good choice. Instead, the absolute value method proved to produce the best results as shown in part (d) and should be used for this problem. As expected, the solutions become worse the higher the error δ is. For $\delta = 10\%$, even the absolute value method could not produce an acceptable result anymore. Errors in the solution occur at the boundaries depending on the magnitude of the error. For $\delta = 0\%$, only one part is not reconstructed well. Increasing δ to 1%, the reconstruction becomes worse especially at the left side of the phase. For the high error, the reconstruction is very inaccurate. However, the middle part is still reconstructed fairly well. There are two main reasons for the problems at the boundaries. The first one is that the autoconvolution equation carries only few information at the beginning and the end of the interval, as can be seen in the discretization (3.9). For example the first value y_1 only depends on x_1, while y_N depends on all x_j, $j = 1 \ldots N$. Also, the absolute values

are very small at the boundaries compared to the maximum. The second reason for the bad reconstruction is that the hose fundamental phase has high values for low and high frequencies, leading to jumps in the $[-\pi, \pi]$ interval (see Fig. 3.3). As we have seen, those jumps always lead to additional problems.

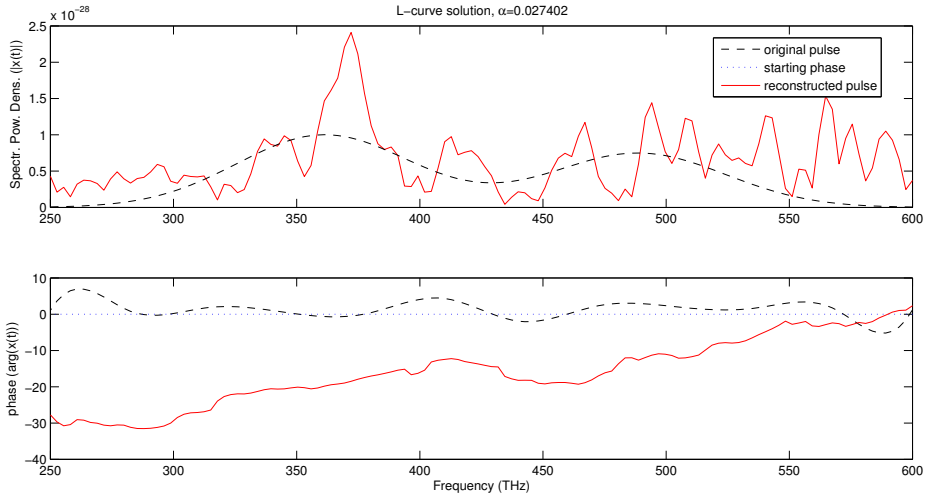
3.6 Regularization with measured data

This section was supposed to present reconstructed phases retrieved from real data. Unfortunately this could not be achieved yet. Although measurements for $\underline{\varphi}_y^\delta$, $|\underline{y}|^\delta$ and $|\underline{x}|^\delta$ have been provided, all calculated solutions for $\underline{\varphi}_x$ were unacceptable as the absolute values did not converge at all.
A very likely reason for this is that no magnitude for absolute values $|\underline{y}|^\delta$ and $|\underline{x}|^\delta$ is known. Only the shape of the curves and thus the relative magnitudes can be measured. Consequently the magnitude of the step p in Alg. 3.2 is unknown and the step length cannot be controlled. Several attempts to scale the absolute values with constant factors failed. Moreover, it is known that the signal of y is much weaker than the signal of x. This behaviour could not be observed when y was convolved from x with a roughly estimated magnitude. This might be a sign of an error in the modelling of the problem. Anyway, we do not know the value of the constant $\overline{\mathcal{E}}^{cw}$ in (1.24) which alone might explain the observation.
An additional problem results from the length of the measurements. Sampling $|\underline{x}|^\delta$ at N supporting points t_n, the convolved signals have to be sampled at resulting $2N-1$ supporting points (see section 3.1). Real data for $\underline{y}$ is available in an interval from $292THz$ to $608THz$ while $|\underline{x}|^\delta$ is known in the interval from $258THz$ to $576THz$. Since the coninuous-wave frequency is estimated as $390THz$, $\underline{y}$ would have to be available for frequencies from $126THz$ to $761THz$. Truncating $|\underline{x}|^\delta$ to fit the frequencies of $\underline{y}$, the unknown phase can only be reconstructed between $341THz$ and $499THz$. Since significant magnitudes of $|\underline{x}|^\delta$ have to be truncated, the algorithm tries to compensate this by increasing the absolute values at the boundaries. As can be seen in the discretization of the operator (3.9), the boundaries influence every single convolved value. Consequently, the whole regularization failed.
Least of all problems is that the cw-frequency cannot be given precisely, for example we have $f_{cw} = 390THz \pm 20THz$.

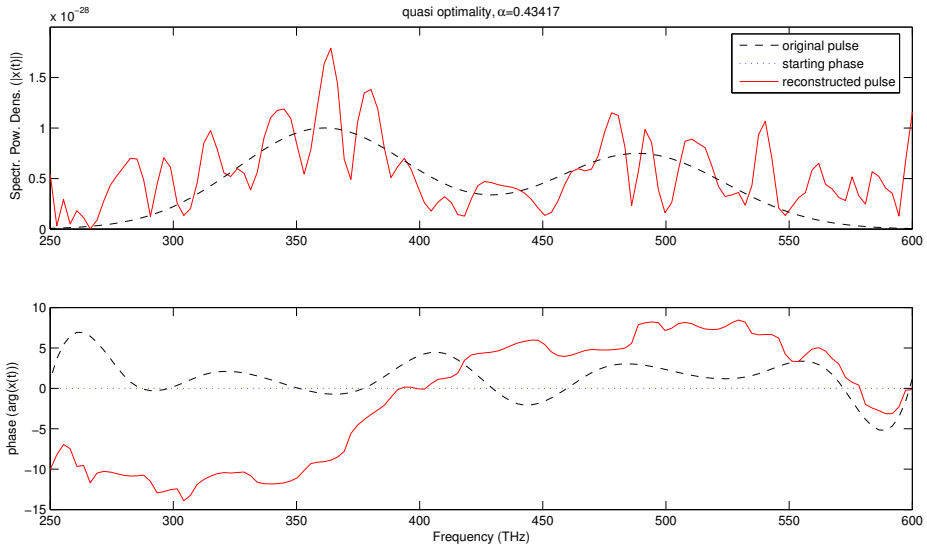
Figure 3.11: Regularization results for $\delta = 0.1\%$, $\tilde{N} = 467$, $N = 130$



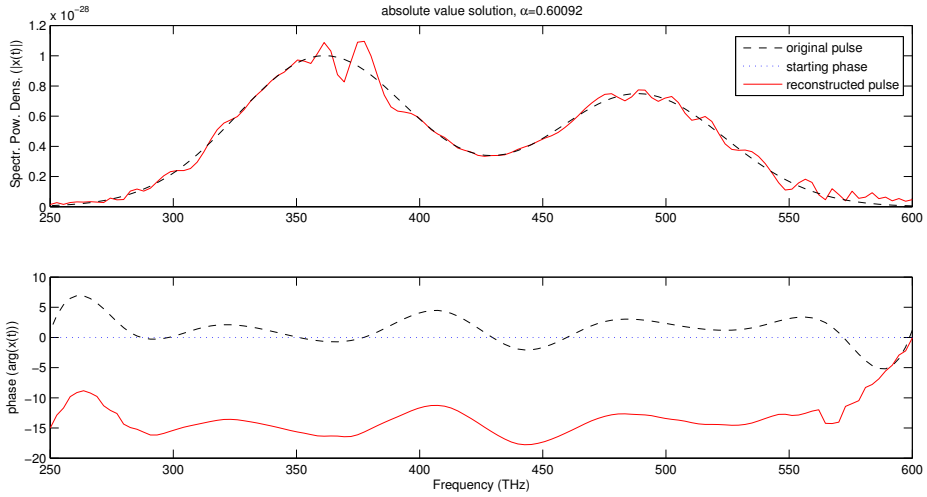
(a) L-curve, not shaped like an L



(b) L-curve solution, $\alpha^* = 0.027402$. As expected from the L-curve itself, the corresponding solution failed.



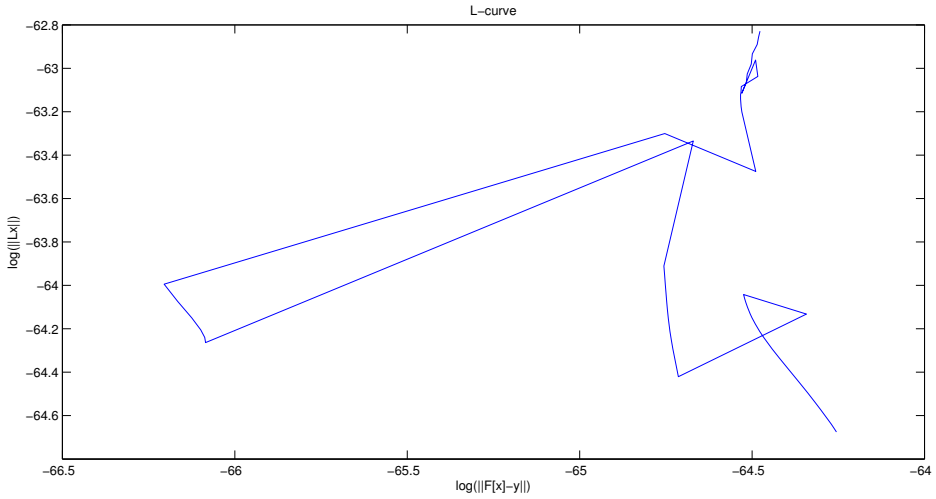
(c) quasi optimality solution, $\alpha^* = 0.43417$, showing a very bad reconstruction



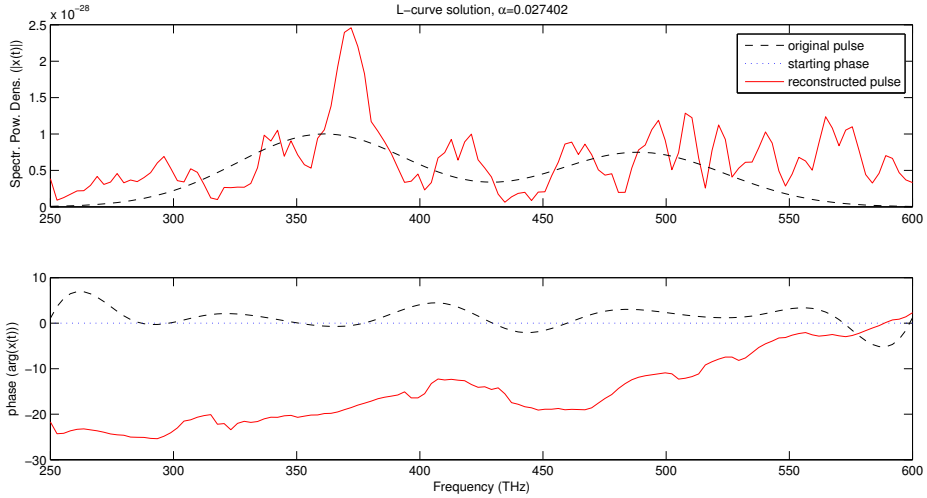
(d) absolute values solution, $\alpha^* = 0.60092$. Except for the right boundary, the reconstruction is very accurate. The absolute values are also reconstructed better at the left boundary than on the right side.

Figure 3.11

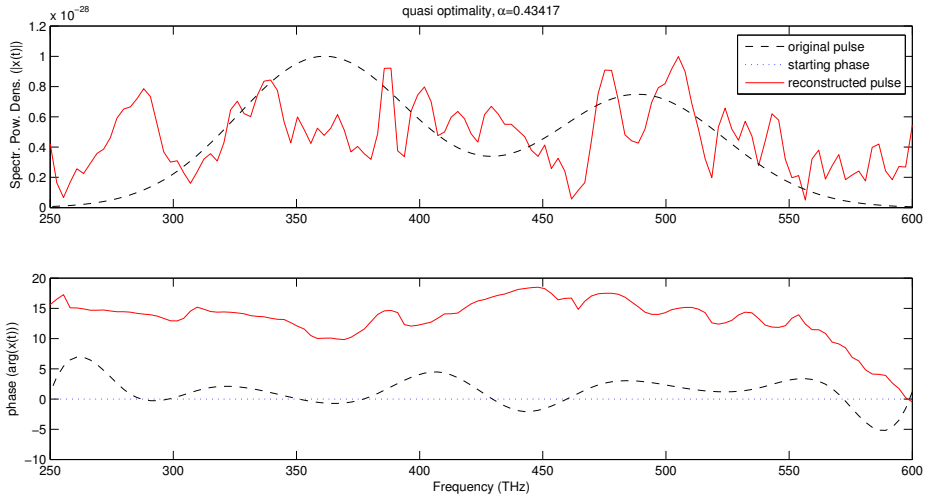
Figure 3.12: Regularization results for $\delta = 1\%$, $\tilde{N} = 467$, $N = 130$



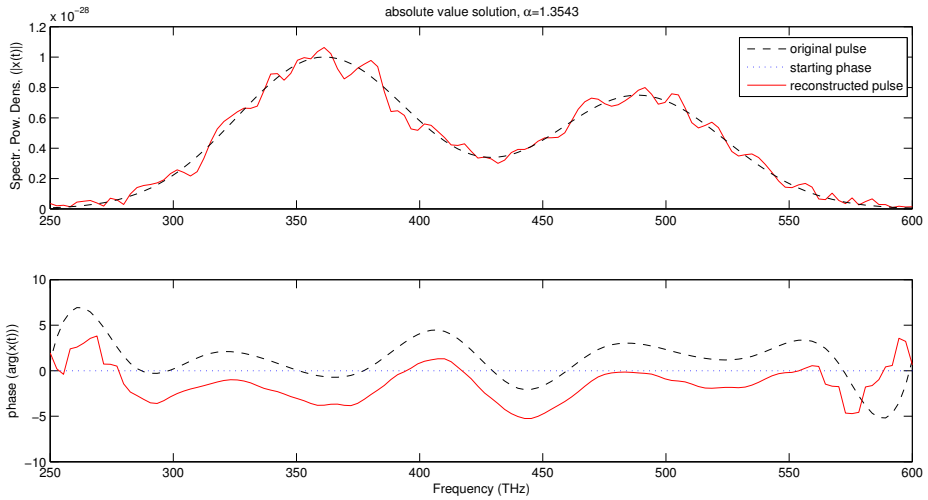
(a) L-curve, again not shaped like an L



(b) L-curve solution, $\alpha^* = 0.027402$. As expected from the L-curve itself, the corresponding solution failed.



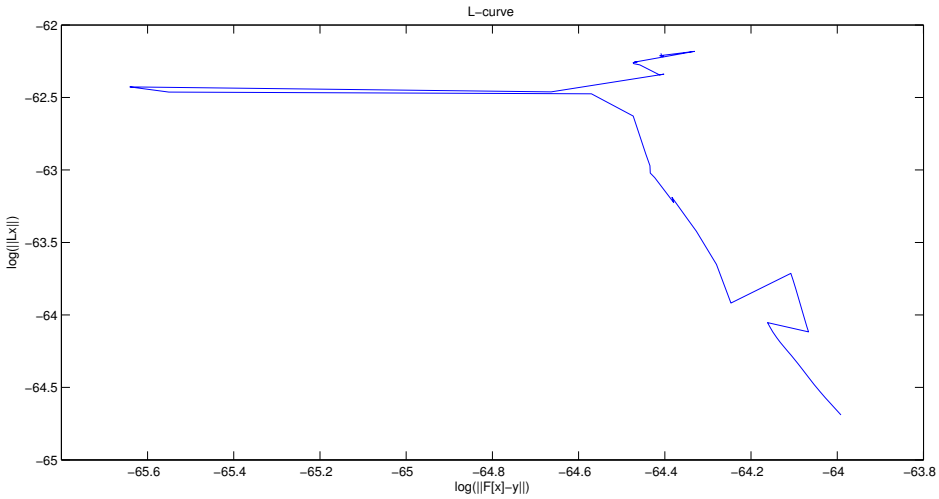
(c) quasi optimality solution, $\alpha^* = 0.43417$, showing a very bad reconstruction



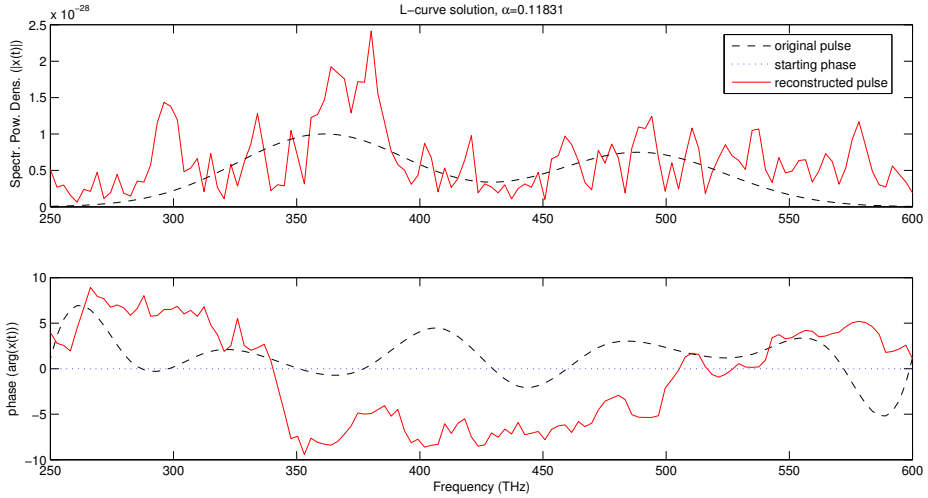
(d) absolute values solution, $\alpha^* = 1.3543$. The reconstruction now has problems at both boundaries, but the middle part is still very accurate

Figure 3.12

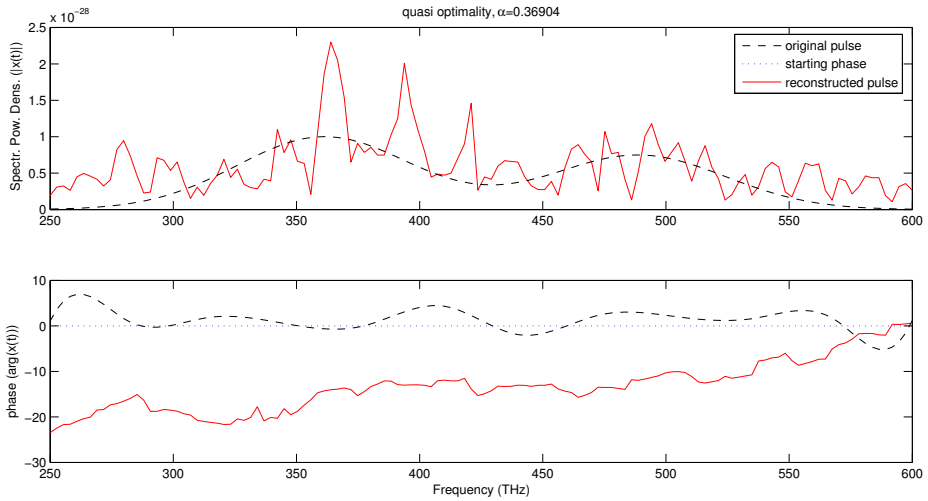
Figure 3.13: Regularization results for $\delta = 5\%$, $\tilde{N} = 467$, $N = 130$



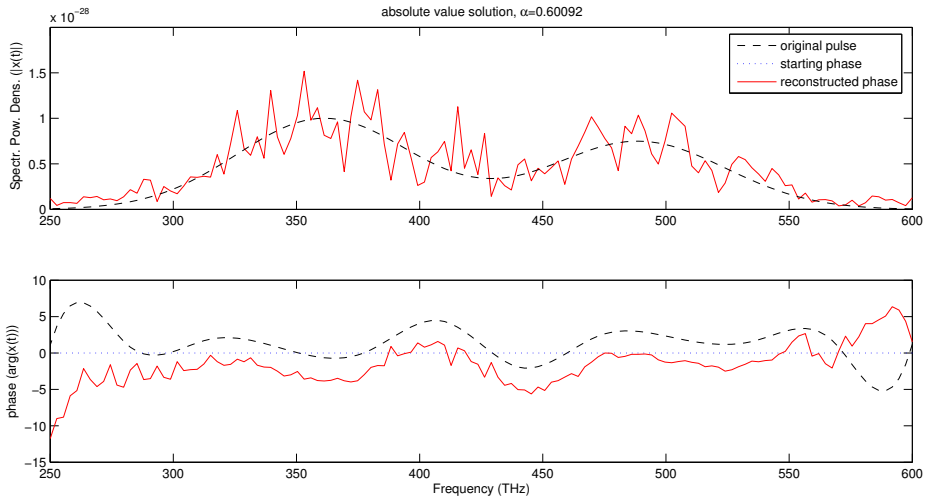
(a) L-curve, again not shaped like an L



(b) L-curve solution, $\alpha^* = 0.11831$. As expected from the L-curve itself, the corresponding solution failed.



(c) quasi optimality solution, $\alpha^* = 0.36904$, showing a very bad reconstruction



(d) absolute values solution, $\alpha^* = 0.60092$. Both absolute values and phase are incorrect at the boundaries. Both reconstructions show lots of spikes, which are more severe for the absolute values. The middle part of the phase is reconstructed well considering the high error in the data.

Figure 3.13

CHAPTER

FOUR

AN APPROPRIATE REGULARIZATION STRATEGY

In chapter 3 a regularization approach has been introduced and explained in great detail. This section will summarize the the main results. We deal with the equation

$$Fx = y, \tag{4.1}$$

$$\int_0^s k(s,t)v(s-t)v(t)dt = y(s) \qquad (0 \le t \le 1,\ \ 0 \le s \le 2). \tag{4.2}$$

We have $y(s) = |y(s)|e^{i\varphi_y(s)}$ and $x(t) = |x(t)|e^{i\varphi_x(t)}$. Discretizing y, x and equation (4.2), we arrive at the problem

$$\underline{F}\,\underline{x} = \underline{y}, \tag{4.3}$$

$$\underline{F}\,\underline{x} = \Delta t \begin{pmatrix} k_{1,1}x_1 & 0 & \dots & 0 & 0 \\ k_{2,1}x_2 & k_{2,2}x_1 & \dots & 0 & 0 \\ & \ddots & \ddots & & \vdots \\ k_{N-1,1}x_{N-1} & k_{N-1,2}x_{N-2} & \dots & k_{N-1,N-1}x_1 & 0 \\ k_{N,1}x_N & k_{N,2}x_{N-1} & \dots & k_{N,N-1}x_2 & k_{N,N}x_1 \\ 0 & k_{N+1,1}x_N & \dots & k_{N+1,N-1}x_3 & k_{N+1,N-1}x_2 \\ \vdots & & \ddots & \ddots & \\ 0 & 0 & \dots & k_{2N-2,N-1}x_N & k_{2N-2,N}x_{N-1} \\ 0 & 0 & \dots & 0 & k_{2N-1,N}x_N \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N-1} \\ x_N \end{pmatrix} = \underline{y}. \tag{4.4}$$

Analogously we derive the discretized Fréchet-derivative $\underline{F}'(\underline{x})$. See section 3.1 for further information. From real measurements or by creating them artificially (see section 3.2) we obtain the data necessary to solve the complex valued autoconvolution problem (4.2). We then have $\underline{y}^\delta = |\underline{y}|^\delta e^{i\underline{\varphi}_y^\delta}$ and $\underline{x}^\delta = |\underline{x}|^\delta e^{i\underline{\varphi}_x}$, where $|\underline{y}|^\delta$, $\underline{\varphi}_y^\delta$ and $|\underline{x}|^\delta$ are known and

$\underline{\varphi}_x$ will be the solution.
To compute $\underline{\varphi}_x$, we use a Levenberg-Marquardt-type algorithm to minimize

$$||\underline{F}\,\underline{x}^{\delta} - \underline{y}^{\delta}||_2^2 \tag{4.5}$$

by minimizing in each iteration the linearized functional

$$||\underline{y}^{\delta} - \underline{F}(\underline{x}_k) - \underline{F}'(\underline{x}_k)\underline{x}_k\underline{z}||_2^2 + \alpha||\underline{L}\underline{z}||_2^2. \tag{4.6}$$

The term $\underline{L}\underline{z}$ in the smoothing penalty $\alpha||\underline{L}\underline{z}||_2^2$ approximates the second derivative of $\underline{z}$. Including the spacing term Δt^2 in the parameter α, L attains the tridiagonal form

$$L = \begin{pmatrix} -2 & 1 & 0 & \dots \\ 1 & -2 & 1 & \dots \\ 0 & \ddots & \ddots & \ddots \end{pmatrix}. \tag{4.7}$$

We then have the iteration procedure

$$\underline{x}_{k+1} = \underline{x}_k + \gamma(\underline{F}'(\underline{x}_k)^*\underline{F}'(\underline{x}_k) + \alpha\underline{L}^*\underline{L})^{-1}\underline{F}'(\underline{x}_k)^*(\underline{y}^{\delta} - \underline{F}(\underline{x}_k)). \tag{4.8}$$

The step control γ ensures a decrease in the residual so that $||\underline{y}^{\delta} - \underline{F}(\underline{x}_{k+1})||_2 \leq ||\underline{y}^{\delta} - \underline{F}(\underline{x}_k)||_2$. Then $\gamma = \gamma_0 \cdot q_\gamma^l$ with $\gamma_0 = 1, 0 < q_\gamma < 1$ and $l = \min\{l \in \mathbb{N} : ||\underline{y}^{\delta} - \underline{F}(\underline{x}_{k+1})||_2 < ||\underline{y}^{\delta} - \underline{F}(\underline{x}_k)||_2\}$. As the starting point $\underline{x}_0$ we propose $\underline{x}_0 = |\underline{x}|^{\delta}e^{i\underline{0}}$, the measured absolute values and a zero phase. The algorithm is stopped either if no significant decrease in the residual is gained,

$$||\underline{y}^{\delta} - \underline{F}(\underline{x}_{k+1})||_2 \geq q||\underline{y}^{\delta} - \underline{F}(\underline{x}_k)||_2, \tag{4.9}$$

$q < 1$ but q close to one or if a maximal number of iterations $k_{\max}$ has been performed. The parameter α is constant for all iterations to ensure smooth solutions. The optimal solution, denoted as $\underline{x}^*$, is retrieved by calculating regularized solutions $\underline{x}^*_{\alpha_j}$ for a series of regularization parameters $\alpha_k = \alpha_0 \cdot q_\alpha^j$, $j = 1, 2, \dots j_{max}$ with $\alpha_0 \in \mathbb{R}$ and $0 < q_\alpha < 1$. Here $\underline{x}^*_{\alpha_j}$ denotes the solution retrieved from minimizing (4.5) using (4.6) for fixed α_j. In most cases it is sufficient to choose $\alpha_0 = 500$, $0.5 \leq q_\alpha \leq 0.95$ and $50 < j_{max} < 120$. The final result $\underline{x}^*$ is then retrieved by minimizing

$$||\,|\underline{x}^*_{\alpha_j}| - |\underline{x}|^{\delta}\,||_2 \tag{4.10}$$

over all j, $j = 1, \dots, j_{max}$. In other words we choose the regularized solution which approximates the measured data for the absolute values best, making use of the special situation that with the absolute values a part of the solution is provided as data. Algorithm 4.1 illustrates the whole procedure.

Algorithmus 4.1 Complete algorithm for the deautoconvolution

input: measured data $\underline{y}^\delta$ and $|\underline{x}|^\delta$, initial value $\underline{\varphi}_0$, α_0, k_{max}, j_{max}, q_α, q.
set $\underline{x}^* := \underline{0}$
for $j := 0$ to $j_{max} - 1$ **do**
  set $k := 0$, $l := 0$, $\underline{x}_0 = |\underline{x}|^\delta e^{i\underline{\varphi}_0}$, $\alpha = \alpha_0 \cdot q_\alpha^j$
  while $\|\underline{y}^\delta - \underline{F}\underline{x}_k\|_2 < q\|\underline{y}^\delta - \underline{F}\underline{x}_k\|_2$ AND $k < k_{max}$ **do**
    solve $(\underline{F}'(\underline{x}_k)^*\underline{F}'(\underline{x}_k) + \alpha\underline{L}^*\underline{L})\underline{p} = \underline{F}'(\underline{x}_k)^*(\underline{y}^\delta - \underline{F}(\underline{x}_k))$
    while $\|\underline{F}(\underline{x}_k + \gamma\underline{p})\|_2 > \|\underline{F}(\underline{x}_k))\|_2$ **do**
      $l = l + 1$
      $\gamma = q_\gamma^l$
    end while
    set $\underline{x}_{k+1} = \underline{x}_k + \gamma\underline{p}$, $l = 0$, $k = k + 1$
  end while
  if $\| |\underline{x}^*_{\alpha_j}| - |\underline{x}|^\delta \|_2 \leq \| |\underline{x}^*| - |\underline{x}|^\delta \|_2$ **then**
    set $\underline{x}^* := \underline{x}^*_{\alpha_j}$
  end if
end for
return $\underline{\varphi}_{\underline{x}^*}$

Regularized solutions obtained with Alg. 4.1 have already been presented in part (d) of Fig. 3.11, Fig. 3.12 and Fig. 3.13 at the end of the previous chapter. Here some more results will be given, but the figures will no more show the absolute values of the solution $\underline{x}^*$, only its phase $\underline{\varphi}_{\underline{x}^*}$. Figures 4.1 and 4.2 show the reconstruction of fundamental pulse I from section 3.2 for a relative error of 0.1% and 1%, respectively. As mentioned in chapter 3, constant offsets of $k \cdot \pi$, $k \in \mathbb{N}$ can be neglected. They result from the 2π-periodicity of the phase and the non-injectivity of the problem (see 2.2.2). To show results for a different phase than the ones introduced in chapter 3, artificial data has been created using the absolute values from (3.15) and a quadratic fundamental phase. A special property of this phase is its high level of smoothness since its second derivative is constant for all frequencies. As we penalize the second derivative in (4.6), a very large regularization parameter can be chosen without overregularizing the solutions. This ensures no spikes occur in the regularized solutions. Figure 4.3 shows a satisfying reconstruction from noiseless artificial data. As with all pulses, problems occur at the boundaries. In Fig. 4.4, these problems increase as the data now carry a relative error of 1%. However, the middle part is still following the fundamental pulse relatively well. This remains even if the relative error of the data is increased to 5% as done in Fig. 4.5. While the right boundary is still reconstructed in an acceptable way, the reconstruction failed for low frequencies up to $350THz$.
With these examples we have demonstrated algorithm 4.1 to be suitable for the numerical deautoconvolution of (4.3), thus solving (4.2) in its discretized version.

Figure 4.1: Reconstructed phase for pulse I, $\delta = 0.1\%$, $\alpha = 0.11259$.
The phase is reconstructed well with an offset of 2π. Only at the right boundary the curves do not match while the left boundary is reconstructed in an acceptable way.

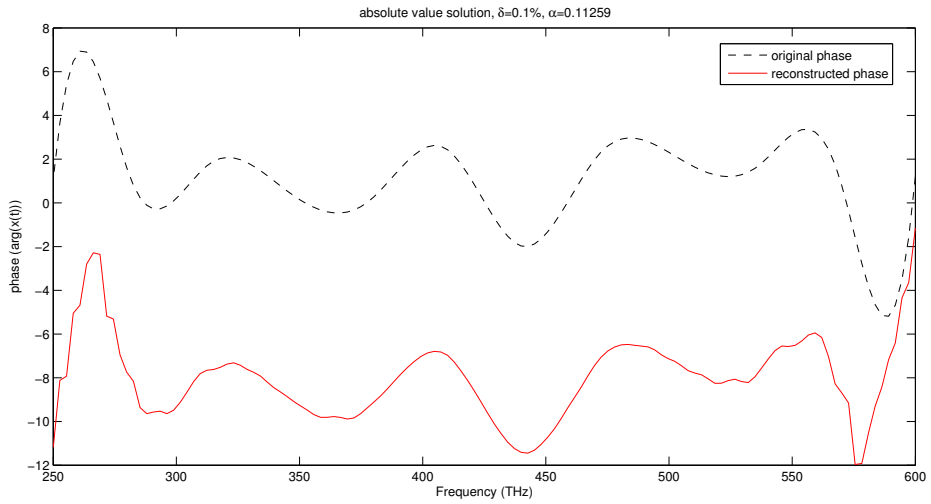


Figure 4.2: Reconstructed phase for pulse I, $\delta = 1\%$, $\alpha = 0.10562$.
The phase is still reconstructed well with an offset of 4π. But now the reconstruction failed at both boundaries and the middle part is less smooth in comparison the the relative error of 0.1%.

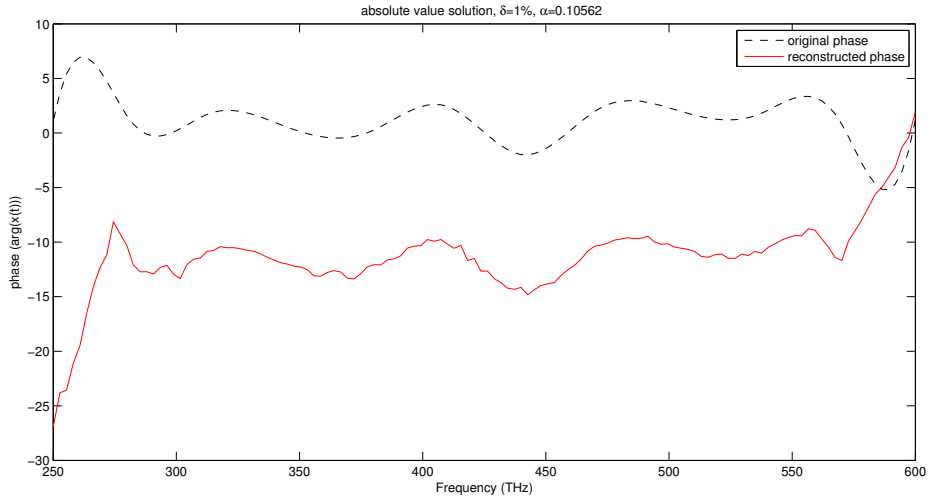


Figure 4.3: Reconstructed phase for a quadratic fundamental phase, $\delta = 0\%$, $\alpha = 197.9121$. Besides a few values for small frequencies the reconstructed phase matches the original one almost perfectly.

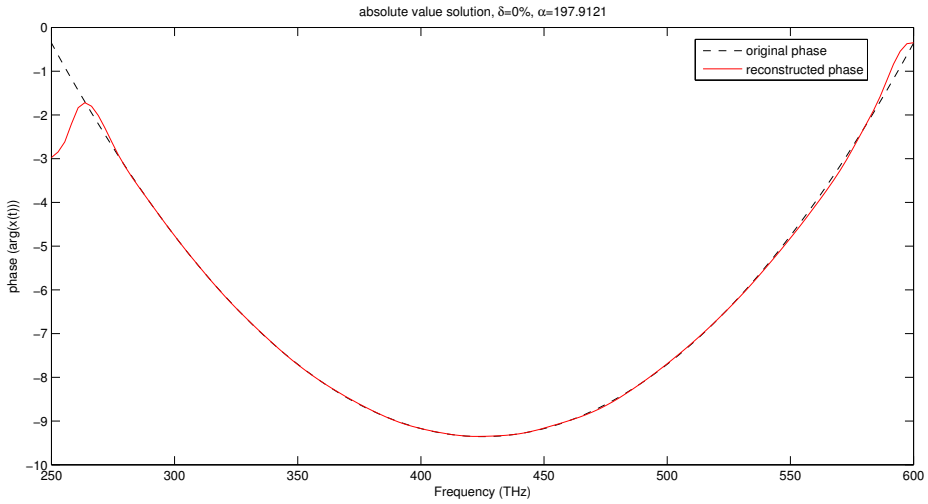


Figure 4.4: Reconstructed phase for a quadratic fundamental phase, $\delta = 1\%$, $\alpha = 197.9121$. Again, the boundaries provide the largest problems. Because the fundamental phase is very smooth, spikes in the solution can be suppressed very effectively. Thus the reconstruction of the middle part is fairly accurate considering the relative error of 1% in the data.

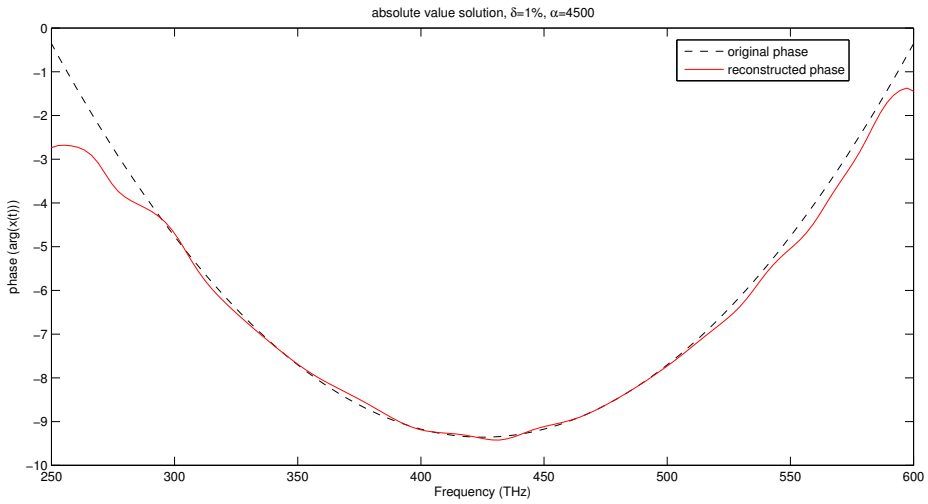
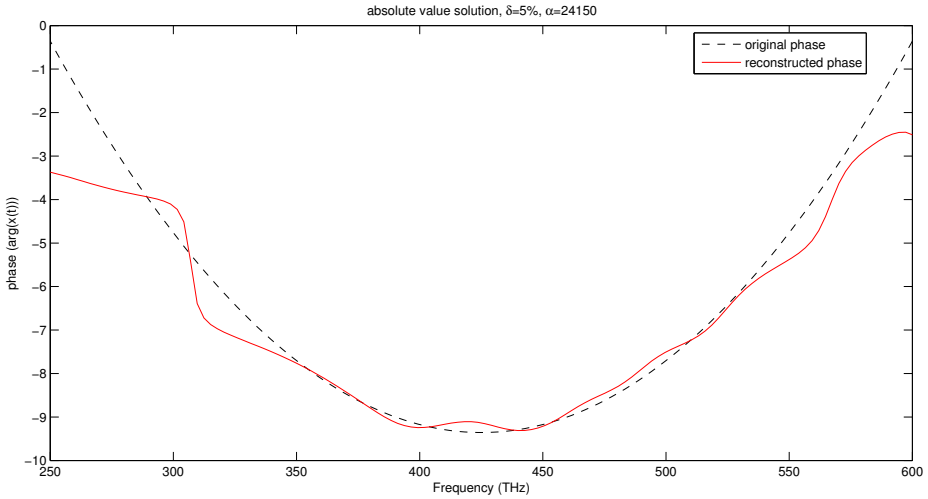


Figure 4.5: Reconstructed phase for a quadratic fundamental phase, $\delta = 5\%$, $\alpha = 24150$. Due to the large relative error of 5% in the data a large regularization parameter is necessary to ensure an acceptable reconstruction. The reconstructed values for the central frequencies are still close to the original ones while especially at the left boundary the phases do not match anymore.



BIBLIOGRAPHY

[1] J. Baumeister, *Deconvolution of appearance potential spectra*, in Direct and Inverse Boundary Value Problems, Proc. Conf. Oberwolfach, 1991, Frankfurt am Main : Lang, 1991, pp. 1–13. 21

[2] R. P. Boyd, *Nonlinear optics*, Academic Press, San Diego, 2 ed., 2003. 5, 7, 8, 14

[3] P. L. Butzer and R. J. Nessel, *Fourier analysis and approximation I*, Basel: Birkhaeuser, 1971. 23

[4] K. Choi and A. Lanterman, *An iterative deautoconvolution algorithm for nonnegative functions*, Inverse Problems, 21 (2005), pp. 981–995. 21

[5] Z. Dai and P. Lamm, *Local regularization for the nonlinear inverse autoconvolution problem*, SIAM J. Numer. Anal., 46 (2008), pp. 832–868. 21

[6] J.-C. Diels and W. Rudolph, *Ultrashort laser pulse phenomena: fundamentals, techniques and applications on a femtosecond timescale*, Academic Press, San Diego, 1996. 6

[7] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of inverse problems*, Kluwer Acad. Publ., 2000. 2

[8] H. W. Engl, B. Hofmann, and H. Zeisel, *A decreasing rearrangement approach for a class of ill-posed nonlinear integral equations*, J. Int. Equations Appl. 31

[9] H. W. Engl, K. Kunisch, and A. Neubauer, *Convergence rates for Tikhonov regularization of non-linear ill-posed problems*, Inverse Problems, 5 (1989), pp. 523–540. 29, 38

[10] G. Fleischer, R. Gorenflo, and B. Hofmann, *On the autoconvolution equation and total variation constraints*, ZAMM, (1999), pp. 149–159. 21

[11] R. Gorenflo and B. Hofmann, *On autoconvolution and regularization*, Inverse Problems, 10 (1993), pp. 353–373. 21, 22, 24, 26, 27, 29, 30, 32, 37

[12] C. W. Groetsch, *Inverse problems in the mathematical sciences*, Vieweg, 1993. 2

[13] J. Hadamard, *Le Problem de Cauchy et les equations aux dirivees partielles lineaires hyperboliques*, Hermann, Paris, 1923. 36

[14] P. C. Hansen, *Analysis of discrete ill-posed problems by means of the l-curve*, SIAM Review, 34 (1992), pp. 561–580. 58

[15] B. Hofmann, *Regularization for applied inverse and ill-posed problems*, Teubner, 1986. 2, 31, 32, 37, 40, 52

[16] ——, *Mathematik inverser probleme*, Teubner, 1999. 52

[17] ——, *Ill-posedness and local ill-posedness concepts in Hilbert spaces*, Optimization, (2000), pp. 219–238. 21

[18] B. Hofmann and G. Fleischer, *On inversion rates for the autoconvolution equation*, Inverse Problems, 12 (1996), pp. 419–435. 21, 37

[19] C. Iaconis and I. Walmsley, *Spectral phase interferometry for direct electric-field reconstruction of ultrashort laser pulses*, OPTICS LETTERS, 23 (1998), pp. 792–794. 6

[20] J. Janno, *On a regularization method for the autoconvolution equation*, ZAMM, 77 (1997), pp. 393–394. 21

[21] ——, *Lavrent'ev regularization of ill-posed problems containing nonlinear near-to-monotone operators with appliction to the autoconvolution equation*, Inverse Problems, 16 (2000), pp. 333–348. 21

[22] B. Kaltenbacher, A. Neubauer, and O. Scherzer, *Iterative regularization methods for nonlinear ill-posed problems*, de Gruyter, 2008. 44, 45, 48

[23] K. Levenberg, *A method for the solution of certain problems in least squares*, Quart. Appl. Math., 2 (1944), pp. 164–168. 44

[24] L. Ljusternik and W. Sobolev, *Elemente der Funktionalanalysis*, Zürich: Harri Deutsch, 1976. 27, 30

[25] T. Maiman, *Stimulated optical radiation in ruby*, Nature, 187 (1960). 2

[26] I. Malitson, *Refractive properties of bariumfluoride*, J.Opt.Soc.Am, 54 (1967(?)), p. 628. 14

[27] D. Marquardt, *An algorithm for least-squares estimation of nonlinear parameters*, SIAM J. Appl. Math, 11 (1963), pp. 431–441. 44

[28] D. Meschede, *Optik, Laser, Licht*, Vieweg+Teubner, 2008. 2

[29] Optical Society of America, *Self-diffraction SPIDER*, Conference on Lasers and Electro-Optics, 2010. 8

[30] B. Saleh and M. C. Teich, *Grundlagen der Photonik*, WILEY-VCH, 2008. 6, 13

[31] O. Scherzer et al., *Variational methods in imaging*, Springer, 2009. 2

[32] G. Steinmeyer, *A review of ultrafast optics and optoelectronics*, J. Opt A: Pure Appl. Opt, 5 (2003). 6

[33] M. Takeda, H. Ina, and S. Kobayashi, *Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry*, J. Opt. Soc. Am., (1982), pp. 156–160. 9

[34] A. Tikhonov, *On the stability of inverse problems*, Dokl. Akad. Nauk SSSR, 39 (1943), pp. 195–198. 2

[35] ——, *On solution of ill-posed problems and the regularization method*, Dokl. Akad. Nauk SSSR, 151 (1963), pp. 501–504. 2

[36] E. C. Titchmarsh, *The zeros of certain integral functions*, Proc. London. Math. Soc., 1 (1926), pp. 283–302. 25

[37] R. Trebino, *Frequency resolved optical gating: the measurement of ultrashort pulses*, Kluwer, 2002. 6, 7

[38] L. v. Wolfersdorf, *A regularization procedure for the auto-correlation equation*, Math. Meth. Appl. Sci., 24 (2001), pp. 1073–1088. 21

[39] ——, *Autoconvolution and special functions*, Integral Transforms and Special Functions, 19 (2008), pp. 677–686. 21

[40] ——, *Autoconvolution equations and special functions ii*, Integral Transforms and Special Functions, 21 (2010), pp. 295–306. 21

[41] L. v. Wolfersdorf and J. Janno, *A general class of autoconvolution equations of the third kind*, Journal for Analysis and its Applications, 24 (2005), pp. 523–543. 21

[42] I. Walmsley, *Few-cycle laser pulse generation and its applications*, Springer-Verlag, 2004, pp. 265–289. 6

CONCLUSIONS

In chapter 1, section 1.3 some fundamentals of nonlinear optics have been summarized. Using these, the principle of *downshift SD-SPIDER* could be explained in detail and especially the autoconvolution effect was pointed out. From the physical processes the autoconvolution equation was derived in section 1.5. Consisting of complex valued functions and a nontrivial kernel, it turned out to be an equation which has not been studied before. Therefore, after some remarks on the identifiability of the problem, existing publications on the real valued, kernel free autoconvolution problem were mentioned. Chapter 2 dealt with the mathematical analysis of the problem. Formulating it as an operator equation, it could be verified that the operator is continuous and weakly closed, but neither injective nor compact, see subsections 2.2.1, 2.2.3, 2.2.2 and 2.2.4, respectively. In section 2.4 several aspects of ill-posedness have been analysed. Although non-compactness is a hint for less ill-posed behaviour, it could be proven that the operator is everywhere locally ill-posed and regularization techniques are necessary for the reconstruction of the unknown function. However, the Fréchet derivative of the forward operator is always compact and thus everywhere ill-posed as pointed out in sections 2.3 and 2.4. This is a very special case as no other example of such a behaviour was mentioned in the literature. Finally, in chapter 3 a regularization approach for the problem was presented. At first, the equation and all occurring functions were discretized. Artificial data were created to test the reconstruction properties in section 3.2. Using a Levenberg-Marquardt type algorithm, solutions could be obtained which carried too many spikes to be acceptable, although they followed the solution quite well. This problem had been solved by adding a smoothing penalty within the algorithm, resulting in algorithm 3.2 as the proposed method for the reconstruction. It turned out that no optimal starting phase can be chosen, but a constant zero phase is always a suitable guess. This is a result of the fact that the algorithm is very sensitive with respect to the periodical behaviour of the sought-after phase. Finally, three possible rules for the choice of the regularization parameter have been introduced and numerically tested in section 3.4. The so called absolute value method established in this work proved to be the most reliable alternative. As shown in the next section, it produced acceptable solutions for relative errors up to five percent. After that it has been explained why at the moment no reconstruction from real measured data is possible in section 3.6. Finally, chapter 4 summarized the previous chapter and provided the regularization approach proposed to solve the autoconvolution problem discussed in this work.

STATUTORY DECLARATION

Statutory Declaration: I declare that I have authored this thesis independently, that I have not used other than the declared sources or resources and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Daniel Gerth

ACKNOWLEDGEMENTS

I would like to express my gratitude to all those who helped me on the way to finishing this thesis. I am deeply indebted to Prof. Dr. Bernd Hofmann from Chemnitz University of Technology. Not only by criticizing and patiently revising the individual parts of this work again and again he was the most important academical support during the last eleven months, but also his encouraging words and the always friendly atmosphere helped me a lot. Thank you very much.
I want to thank Dr. Günter Steinmeyer and his team from MBI Berlin for the warm welcome and the help provided during my stays in Berlin from November to December 2010 and in February 2011. It was a pleasure working in your department. Especially I have to thank Sebastian Koke and Simon Birkholz for answering all my questions about the physical processes and giving me an insight into the world of nonlinear optics.
From Chemnitz University of Technology I have to mention Dr. René Schneider who, although not involved in this work, was the most important contact person in the past years. Out of my fellow students Heiko Weichelt and Martin Köhler are the ones whom I would like to dedicate my deepest gratitude to. Thank you for the friendship that established soon after we started university. Thank you for your help provided, support given, solace spent, jokes made, beer drunk,etc. This listing could be extended to fill several pages.
Leaving the academic circles, I arrive at my family and friends. Thanks to Franziska, Gregor and Michael for all the years spent together. I want to thank my aunts, uncles, cousins and friends who have at least for a while accompanied me during the time since I first took breath. Special thank goes to my grandparents Eberhard, Gisela and Ria, whom I will present this work as proudly as the first lines I sketched on paper long time ago. I bemoan the thought that I will never have the chance to express my deepest thank to my grandfather Reiner, who left this world too early. Most of all my heartfelt gratitude goes to my father Uwe and my mother Andrea. Without their love, advise and support I would have never arrived at this point.