Erzeugung und Transport von Doppelpaket-Elektronenstrahlen im FLASH Linearbeschleuniger

(Generation and transport of double-bunch electron beams in the FLASH beamline)

von

Carlos Manuel Entrena Utrilla

geboren am

04. Mai 1989

Master-Arbeit im Studiengang Physik

Universität Hamburg

2014

1. Gutachter: Dr. Bernhard Schmidt

2. Gutachter: Prof. Jörg Rossbach

Contents

Zusammenfassung

Der Freie-Elektronen Laser in Hamburg (FLASH) ist Teil des Forschungszentrum Deutsches Elektronen-Synchrotron (DESY). Sein Linearbeschleuniger erzeugt Elektronenpakete hoher Qualität mit einer maximalen Energie von 1.2 GeV, die kurze, hochintensive Röntgenblitze mit einer Wellenlänge von 4.2 nm bis 45 nm im SASE Prozess produzieren. Unter anderem, diese Eigenschaften machen FLASH zu einer der weltweit führenden Freie-Elektronen Lasern neben Linac Coheren Light Source (beim SLAC in Standford, USA), FERMI@Elettra in Trieste (Italien) und SACLA (Japan).

Neben der Erzeugung hochintensiver Röntgenblitze für verschiedene Anwendungen besteht für mehrere Experimente Bedarf an der Beschleunigung zweier Elektronenpaketen mit einem finalen zeitlichen Abstand von einigen hundert Femtosekunden. Diese zwei Paketen werden innerhalb eines Abstands von einigen zehnteln Pikosekunden von einander auf der Photokathode erzeugt und in der gleichen Periode der beschleunigenden HF (Hochfrequenz) Felder durch den Beschleuniger transportiert. Diese Experimenten sind von Bedeutung für zwei farbiger FEL für Anrege-Abfrage-Experimente und für die externe Injektion von Elektronen in den zukünftigen Experimenten in teilchenstrahlgetribener Plasmabeschleunigung im Projekt FLASHForward, das voraussichtlich Anfang 2016 beginnen wird.

Diese Masterarbeit analysiert die longitudinale Strahldynamik dieser Paketen, von der Erzeugung bis zum Transport und der Kompression durch den Beschleuniger. Es wird gezeigt, wie ein Arbeitspunkt für ein beliebiges Kompressionszenario (Form und Spitzenstrom der Paketen, und Endabstand zwischen diesen) mit verschiedenen numerischen Simulationen gefunden werden kann, wie die Elektronen experimentell erzeugt und durch das aktuelle Layout des Beschleunigers transportiert werden können. Dies wurde in einem ersten Experiment am Ende Mai 2014 demonstriert.

Abstract

The Free Electron Laser in Hamburg (FLASH) is part of the Deutsches Elektronen-Synchrotron (DESY) research center. Its linear accelerator produces high-quality electron bunches of up to about 1.2 GeV that are used in its undulator to generate short, intense, high-brilliance soft-X ray pulses with a wavelength from 4.2 nm to 45 nm with the SASE process. This characteristics make FLASH a leading facility worldwide in photon science and linear accelerator technologies, along with the Linac Coherent Light Source (in SLAC, Standford, USA), the FERMI@Elettra in Trieste (Italy) and SACLA (Japan).

For several reasons, there is a substantial interest to accelerate two electron bunches with a final temporal distance of several hundreds of femtoseconds. These two bunches are generated on the photocathode within picoseconds from each other and accelerated within the same RF bucket (the same period of the RF (radio-frequency) accelerating fields). These experiments are of interest for two-color FEL for pump-probe experiments, and for the external injection of electrons in the future particle-driven plasma wakefield accelerator experiment, called FLASH-Forward, which will start in early 2016.

This work analyzes the longitudinal dynamics of said double-bunches, from generation on the photocathode to the transport and compression through the linac. It is shown how a working point for a desired compression scenario (shape and final current of the bunches, and final distance between them) can be found with different numerical tracking procedures, and how the electrons can be experimentally generated and transported through the accelerator in the current layout, which was confirmed in a proof-of-concept experiment in late May 2014.

Chapter 1.

Introduction

DESY (Deutsches Elektronen-Synchrotron) is one of the most important research centers in Germany and one of the World's learders in particle accelerator research. As such, it is home to several such machines, including the Free-Electron Laser in Hamburg (FLASH). FLASH is a 315 m-long facility that is pioneering the Free-Electron Laser (FEL) technology. It was the first facility to achieve single-pass soft x-ray lasing in 2005 [1], opening up a new area in photon science with its new, short, highly brilliant pulses. The current layout of FLASH is shown in Figure 1.1 [2].

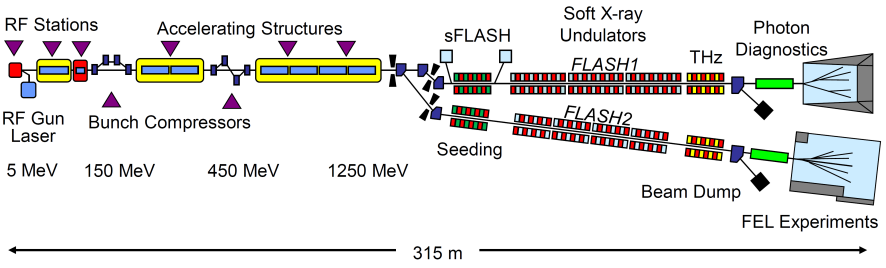

Figure 1.1.: Current layout of FLASH with all its undulators (current and planned) and design bunch energies.

The beamline of FLASH includes the injector, containing the laser-driven RF gun, a booster section and two bunch compressor [3], and the linear accelerator (linac), which finally feeds the electron bunches to the SASE undulators, which generate the laser beams for the FEL experiments.

The injector [3] generates the electrons and gives the bunch the first acceleration. The electrons are generated in a Cs_2Te photocathode with an UV laser of 262 nm. The electrons are then accelerated in the RF gun, a 1.5-cell L-band cavity powered by a 5 MW klystron. The RF has a frequency of 1.3 GHz and is operated in the TM_{010} mode. After that, a solenoid (not shown in the figure) focuses and matches the beam to the booster section (the first accelerating module).

After the booster [4], a four-cavity module (shown in red) operated at 3.9 GHz (3rd harmonic of the RF of the other modules) after the first module linearizes the longitudinal phase space in order to provide a better, more regular shape in the electron bunches through a more uniform compression. The bunches are compressed using two magnetic chicanes. They are then accelerated up to 1.25 GeV with seven TESLA-type modules, each consisting of eight 9-cell superconducting cavities at 1.3 GHz. Before reaching the undulators, the bunches go through an energy collimator (also called "dogleg") to eliminate off-energy particles [5].

The high-quality beams from FLASH are used in the undulators to generate SASE-FEL radiation. The FLASH1 undulator section consists of six 4.5 m-modules with a gap of 12 mm, and undulator period of 25 mm and a peak magnetic field of 0.48 T [5]. Other experiments include the seeded FLASH experiment (sFLASH), a 10 m-long, variable-gap undulator [6] which in 2012 produced the first results in seeded FEL [7] and the THz undulator, a beamline designed to produce undulator radiation from the mid to the far infrared [8].

There are also two major upgrades planned. The FLASH2 project is an additional FEL undulator line which will use the bunches provided by the FLASH linac. The beamline is under construction and the first user experiments are expected in 2015 [2]. The second project is the FLASHForwardproject, which plans to study beam-driven plasma acceleration with bunches from FLASH. The experiments are planned to start in early 2016.

On this work, the generation and transport of two electron bunches within the same RF bucket in the FLASH linac will be studied. These bunches would be used for the FLASHForward experiment. For future reference along the chapters, Figure 1.2 shows the linac with the terminology that will be used in this work for each element.

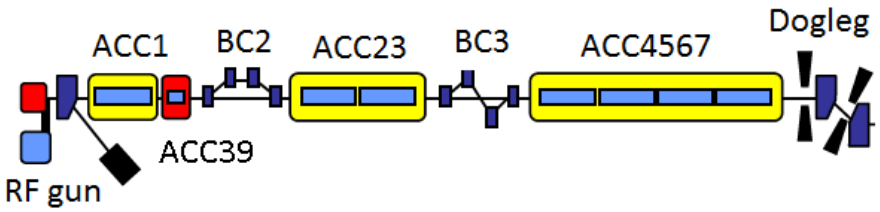

Figure 1.2.: Linac with used denominations (figure adapted from Figure 1.1)

Chapter 2.

Task definition

The task at hand is to generate two electron bunches on the cathode of the photoinjector with a delay of some tens of picoseconds, and then transport them through the machine in the same RF bucket and bring them within some tens of micrometer distance (some hundreds of femtoseconds) after the two-stage compression. There are two scenarios in which this double-bunch beams are desirable: external injection studies in particle-driven plasma acceleration and two-color FEL experiments.

In particle-driven plasma acceleration, the accelerating fields are created by an electron bunch (the driver) being shot through a gas target. The gas is ionized either by the field of the bunch or with a laser. The bunches then displace the electrons in the plasma, creating huge electric fields on the wake of the bunch. These wakefields can achieve gradients in the order of several GV/m [9] and are therefore very well suited for particle acceleration with unprecedent gradients.

One key to the problem is injecting the electrons in this wake for them to be accelerated. Several approaches are currently being studied at DESY and other research centers around the world, like the SLAC National Accelerator Laboratory in California (USA). One of the possible solutions is the external injection, which consists of shooting another bunch shortly after the driver through the plasma. This bunch, called witness bunch, sits on the wake of the driver bunch and can experience the before mentioned accelerating gradients. In order for it to be accelerated, the witness bunch needs to be at a determined distance from the driver in which the generated wakefields are accelerating for electrons. This distance is given by the properties of the driver bunch (charge, current, beam size, etc) and those of the plasma (plasma density, plasma wavelength, etc) and is usually around some tens of micrometers.

The other interesting application is the two-color FEL experiments. The FEL in FLASH uses Self-Amplified Spontaneous Emission (SASE) to generate coherent light pulses. In SASE, an electron bunch is shot through the undulator, a periodic structure of successive permanent magnets arranged so that it creates an approximate sinusoidal magnetic field on the beam's path. This field makes the bunch oscilate, emitting EM radiation with the wavelength of equation (2.1) (on the beam axis). Under favorable conditions, the EM wave (faster than the electrons) moves a distance of λ_l relative to the electrons after one undulator period, so that the interaction between the electron bunch and the wave can be sustained and transfer of energy from the bunch to the wave occurs. The radiation generated interacts with the electron bunch and creates a periodic modulation of the longitudinal electron density, creating microbunches with the same length as the radiation wavelength. This density modulation generates the coherent FEL radiation. This leads to higher intensity, which in turn leads to stronger microbunching, which generates exponential growth of the radiated intensity [10].

$$\lambda_l = \frac{\lambda_u}{2\gamma^2}(1 + \frac{K^2}{2}) \tag{2.1}$$

Equation 2.1 describes the FEL wavelength λ_l as a function of the undulator period λ_u, the relativistic gamma factor of the electrons γ, and the undulator parameter $K = eB_0\lambda_u/2\pi m_o c$. A two-color FEL would consist of a FEL beam with two different wavelenghts. According to equation (2.1), the laser's wavelength can be selected by fine tuning the bunch's energy or by varying the period or the gap of the undulator, so there are several ways of obtaining two colors in the same laser. One way would be to use the same bunch in two undulators with different undulator period, like the method applied in [11], and other way consists in having two bunches with different energies going through the same undulator one after each other. This approach, which is also being studied at the Linac Coherent Light Source (LCLS) in SLAC [12], is the one of interest here. The resulting two-color laser beams could then be used for pump-probe experiments in time-resolved spectroscopy, for example.

Since FLASH operates with single bunches (meant in contrast with this double-bunch case), generating and transporting two bunches in the same RF bucket presents several difficulties:

- The injector laser pulse needs to be separated into two equivalent laser pulses with a varying delay of some tens of picoseconds.
- A proper working point for the accelerator needs to be determined. The working point is defined as the voltages and phases of the RF in the accelerating modules and the currents in the bunch compressors.
- The system needs to easily produce different final energy differences and different final distances between the bunches to adapt to the two possible uses (external injection in plasma and two-color FEL)
- The ability of the current equipment for beam diagnostics needs to be tested for the double-bunch case.

These items will be adressed during this work.

Chapter 3.

Determination of the working point

This chapter explains the process for the determination of the working point, i.e. the voltages and phases of the accelerating modules, the RF parameters. This working point is required in order to obtain a specific final particle distribution after the second compressor. The process is semianalytical, that is, it combines analytical calculations with numerical simulations, and is described in [13]. First, the evolution of the distribution through the accelerator is calculated analytically (without collective effects), obtaining an initial set of RF parameters. This set is then iteratively optimized to take into account collective effects with numerical tracking, initially just of the longitudinal distribution and finally with full 3D simulations. The final result is a set of RF parameters that produces the wanted result after the transport of the bunches.

While the method presented in [13] is generalized to a machine with any number of accelerating modules and compressors, here the method for a FLASH-like machine (the two-compressor, two-accelerating-module scheme) will be presented.

3.1. Semianalytical model

3.1.1. Analytical solution for two-stage compression without collective effects

The first step of the determination of the working point is analytical without collective effects. This analytical model is introduced and thoroughly developed in [13], here the particular case for the FLASH linac will be explained.

First of all, an initial particle distribution is required. This distribution has to be generated with numerical tracking, for example with the code ASTRA [14], which can simulate the emission of particles from the photocathode and is very well suited for sub-relativistic bunches, like those found in the RF gun. This distribution would be the one entering the first accelerating module, so in the FLASH case it could be tracked until after the solenoid in the gun. At this point, the reference energy is about 5 MeV and the particles can already be considered relativistic.

Figure 3.1 shows an example of a simulated particle distribution of a single bunch after the RF gun. The head of the bunch is on the right (positive values of s_0). The current profile is the projection of the phase space distribution onto the longitudinal axis.

The relative position along the bunch, noted as s_0, is defined as the position relative to the reference particle, which is placed in the middle of the distribution. The relative energy deviation, also measured relative to the energy of the reference particle, is defined as the dimensionless

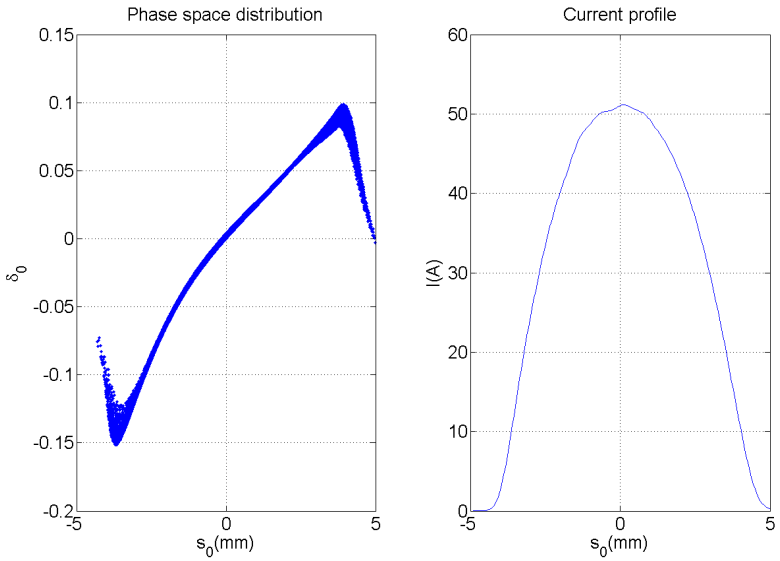

Figure 3.1.: Phase space plot and current profile of a single 1 nC bunch after the gun

quantity $\delta_0 = (E_0 - E_0^0)/E_0^0$, where the upper index 0 indicates that the magnitude is that of the reference particle. The lower index 0 indicates that the magnitudes are those at the beginning of the accelerator (before ACC1).

The phase space distribution will now be approximated as a third order polynomial (a Taylor polynomial up to third order) around the center of the distribution $s_0 = 0$ (3.1). This approximation works better around the center of the distribution and is less precise in the head and tail of the bunch. Note that $\delta_0(0) = 0$ by definition, and that all derivatives are with respect to the initial position s_0, which will be used as independent variable throughout this section and all the functions depend on it.

$$\delta_0(s_0) = \delta_0'(0)s_0 + \frac{\delta_0''(0)}{2}s_0^2 + \frac{\delta_0'''(0)}{6}s_0^3 \tag{3.1}$$

This approximation allows for the distribution to be described using just the vector $\vec{\delta_0} = (\delta_0', \delta_0'', \delta_0''')$. The particle density follows the changes in these parameters. The evolution of $\vec{\delta_0}$ in the accelerators and the compressors can be calculated analytically for a specific set of RF parameters, resulting in a final distribution (final vector) $\vec{\delta_2} = (\delta_2', \delta_2'', \delta_2''')$, where the subindex 2 indicates that those parameters are after the second bunch compressor. This calculation can be inverted, so that for a fixed $\vec{\delta_2}$ it results in a specific set of RF parameters. In order to do that, the equations describing the evolution of $\vec{\delta_0}$ through the accelerator need to be presented.

In the accelerating modules, relavisic electrons have an energy gain proportional to the cosine of the phase angle between the particle's position, s, and the position of maximum energy gain. This leads to a position-dependent energy variation, with no changes in the position if we ignore collective effects (3.2). ΔE is the energy gain, V is the maximum voltage (also called "on-crest" voltage), ϕ the phase of the reference particle, k the wave number of the RF and e is the electron charge. Note that individual particles are no longer taken into account, and these are just functions depending on s. For the first three accelerating modules, ACC1, ACC39 and ACC23, the equations are:

$$\begin{aligned}\Delta E_1 &= eV_1 \cos(ks_0 + \phi_1)\\ \Delta E_{39} &= eV_{39} \cos(3ks_0 + \phi_{39})\\ \Delta E_{23} &= eV_{23} \cos(ks_1 + \phi_{23})\end{aligned} \tag{3.2}$$

In the bunch compressors, an equation as simple as the former can be used. It is shown in [15] that the change in position in a bunch compressor can be approximated by equation (3.3), where δ is the relative energy deviation, and r_{56}, t_{56} and u_{56} are the first, second and third order momentum compaction factors (R_{56}, T_{566} and U_{5666} respectively, see Appendix A) of the compressor. In contrast with the behaviour in the accelerating modules, the bunch shows in the compressors an energy-dependent variation in the position, with no changes in the energy. There is one equation for each compressor. The subindices indicate if the magnitudes are after the first (1) or second (2) bunch compressor.

$$\begin{aligned}s_1 &= s_0 - (r_{561}\delta_1 + t_{561}\delta_1^2 + u_{561}\delta_1^3)\\ s_2 &= s_1 - (r_{562}\delta_2 + t_{562}\delta_2^2 + u_{562}\delta_2^3)\end{aligned} \tag{3.3}$$

Equations (3.2) and (3.3) describe the evolution of $\vec{\delta_0}$ for a given set of RF parameters and bending radii in the compressors (in C and S chicanes (see Appendix A) the momentum compaction factors are related to the bending radius and can be analytically approximated from it [15]). Here instead it is intended to calculate the RF parameters that produce a specific $\vec{\delta_2}$. In this case it is necessary to know the design energies after (or before) the bunch compressors and the variation of the s coordinate in them. This variation is called the inverse compression factor Z, which is defined as follows:

$$\begin{aligned}Z_1 &= \frac{\partial s_1}{\partial s_0}\\ Z_2 &= \frac{\partial s_2}{\partial s_0}\end{aligned} \tag{3.4}$$

Where s_i is the longitudinal coordinate after the i-th bunch compressor (i is just either 1 or 2), and s_0 is the longitudinal coordinate of the initial distribution. This means that the particle that had a position s_0 in the initial distribution will have a position $s_i(s_0)$ after the i-th compressor. The global inverse compression factor would be defined as the compression factor at the end of the linac, in this case Z_2. With this definition, the compression factors are dimensionless quantities.

With the compression factors, the design energies, and the fixed bending radii in the compressors, the only remaining unknowns are the RF settings. In FLASH there are six RF parameters for the compression: four for ACC1 and ACC23 and other two for the third harmonic module ACC39. The compression factors and the energies result in 2 equations for each compressing stage, that is 4 equations, shown in (3.5), where V_2 and ϕ_2 are the RF parameters for ACC23, $V_{1,1}$ and $\phi_{1,1}$ are those for ACC1 and $V_{1,3}$ and $\phi_{1,3}$ are for ACC39. The upper index 0 indicates that the values are refered to $s = 0$.

$$\begin{aligned}
\delta_1^0(V_{1,1},\phi_{1,1},V_{1,3},\phi_{1,3}) &= 0\\
s_1'^0(V_{1,1},\phi_{1,1},V_{1,3},\phi_{1,3}) &= Z_1^0\\
\delta_2^0(V_{1,1},\phi_{1,1},V_{1,3},\phi_{1,3},V_2,\phi_2) &= 0\\
s_2'^0(V_{1,1},\phi_{1,1},V_{1,3},\phi_{1,3},V_2,\phi_2) &= Z_2^0
\end{aligned} \tag{3.5}$$

These equations have a pretty straightforward interpretation: the center of the distribution (the reference particle) has always the design energy, hence $\delta^0 = 0$, and the compression in $s = 0$ is equal to the design (inverse) compression factors. The equations have to be derived from (3.2) and (3.3) and the definition of $\delta = (E - E^0)/E^0$. For example, the first two of them (those for the first compressor) are derived in (3.6). Note that $\delta_1'^0$ is left as it is a parameter of the distribution after the first compressor.

$$\begin{aligned}
\delta_1^0 &= \frac{E_1 - E_1^0}{E_1^0} = \frac{E_0^0 + eV_{1,1}\cos(\phi_{1,1}) + eV_{1,3}\cos(\phi_{1,3})}{E_1^0} - 1 = 0\\
s_1'^0 &= 1 - r_{561}\delta_1'^0 - 2t_{561}\delta_1^0\delta_1'^0 + 3u_{561}(\delta_1^0)^2\delta_1'^0 = 1 - r_{561}\delta_1'^0 = Z_1^0
\end{aligned} \tag{3.6}$$

So far there are 4 equations but 6 RF parameters total. The remaining two other equations to complete system are given by the final shape of the compression in the middle of the distribution. That is, the first and second derivative of the inverse compression factors Z_2' and Z_2''. These two allow for careful selection of the final compression, determining where the maximum (or minimum) of the compression will be placed and how the compression behaves along the bunch. The role of these derivatives will be explained later on in Section 3.2.3, but for now they produce the remaining two equations needed to complete the system, which are equations (3.7).

$$\begin{aligned}
s_2''^0(V_{1,1},\phi_{1,1},V_{1,3},\phi_{1,3},V_2,\phi_2) &= Z_2'^0\\
s_2'''^0(V_{1,1},\phi_{1,1},V_{1,3},\phi_{1,3},V_2,\phi_2) &= Z_2''^0
\end{aligned} \tag{3.7}$$

These equations need again to be calculated from (3.3) but will not be shown here due to their length. The final resulting system is then:

$$\begin{aligned}
\delta_1^0(V_{1,1},\phi_{1,1},V_{1,3},\phi_{1,3}) &= 0\\
s_1'^0(V_{1,1},\phi_{1,1},V_{1,3},\phi_{1,3}) &= Z_1^0\\
\delta_2^0(V_{1,1},\phi_{1,1},V_{1,3},\phi_{1,3},V_2,\phi_2) &= 0\\
s_2'^0(V_{1,1},\phi_{1,1},V_{1,3},\phi_{1,3},V_2,\phi_2) &= Z_2^0\\
s_2''^0(V_{1,1},\phi_{1,1},V_{1,3},\phi_{1,3},V_2,\phi_2) &= Z_2'^0\\
s_2'''^0(V_{1,1},\phi_{1,1},V_{1,3},\phi_{1,3},V_2,\phi_2) &= Z_2''^0
\end{aligned} \tag{3.8}$$

This system has as unknowns the six RF parameters ($V_{1,1}$,$\phi_{1,1}$,$V_{1,3}$,$\phi_{1,3}$,V_2,ϕ_2) and as free parameters the eight parameters that form the compression scenario: the design energies in the

compressors (E_1, E_2), the bending radii in the compressors (r_1,r_2), the first inverse compression factor Z_1, and the second inverse compresion factor and its derivatives (Z_2,Z_2',Z_2''). This set of eight parameters (energies, radii and inverse compression factors and derivatives) is going to be called the compression scenario from now on. The solution of this system provides the theoretical parameters needed in the accelerator in order to achieve a desired bunch after the compression for a given initial distribution. This solution is done in full detail step by step in [13].

The system in (3.8) is essentially the application of a non-linear operator A_0, which from a vector of the RF parameters $x = (V_{1,1}, \phi_{1,1}, V_{1,3}, \phi_{1,3}, V_2, \phi_2)$ gives the compression parameters vector $f_0 = (E_1, E_2, Z_1, Z_2, Z_2', Z_2'')$. That is $A_0(x) = f_0$. Solving the system for an unknown vector of RF parameters is then the inversion of A_0, that is:

$$x_0 = A_0^{-1}(f_0) \tag{3.9}$$

This is the base for the iterative procedure for the inclusion of the collective effects into the calculation with numerical particle tracking, which is explained in Section 3.1.2.

Application to the double-bunch case

Figure 3.2 shows an example of a particle distribution with a double bunch after the gun.

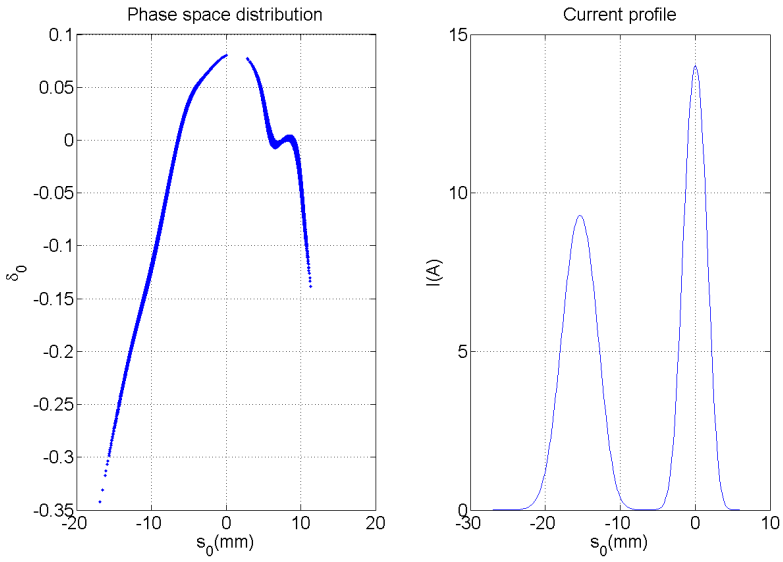

Figure 3.2.: Distribution after the RF gun for 65 ps delay on cathode and 180 pC charge on each bunch.

In order to describe the bunches separately, having two of them instead of one would double the number of parameters needed. That means having 7 parameters for the initial distribution (3 for the first bunch and 4 for the second one, for which δ_0 will not necessarily be zero) and other 7 for the final one, making a total of 14 parameters. However, there are only 6 RF parameters available for tweaking, so the system would be undetermined.

One possibility would be to describe both bunches separately but to a lower order. Since second order would require still 10 parameters in total, the approximation has to be reduced to first order for each bunch, leaving 3 parameters on the initial distribution (one for the first bunch and two for the second one) and 3 for the final one, for a total of 6 parameters. The distribution would then be described by the energies of the bunches, their compression factors, and the initial distance between them, without possiblity to include the derivatives of the compression factor for the final fine tuning. This option produces neat results for the analytical calculation, but the resulting set of RF parameters proved not to be a good starting point for the subsequent numerical tracking.

The better solution to the problem is to just apply the same process as for the single bunch. In the calculation, there were almost no requirements on the distribution, only that the reference energy was high enough for the particles to be considered relativistic. The distribution was just approximated up to third order near the center, leaving the head and the tail of the distribution unconsidered. With a double bunch, the distribution can be considered as a bigger bunch with a gap in the middle, and it can still be approximated up to third order around $s_0 = 0$. This approximation will be less precise than with just one bunch: since the distribution is wider, there will be a bigger part of it unconsidered. Nevertheless, the only consequence of this is that the method requires more iterations in the numerical tracking to achieve a required precision. These results are discussed in Section 3.2.3.

3.1.2. Iterative procedure for the solution with collective effects

The analytical method provides an initial set of RF parameters that were calculated using an approximation for the core of the distribution and neglecting collective effects. These parameters need to be tuned in order to take into account the neglected effects and the rest of the distribution, which can be done iteratively.

As it was mentioned, the analytical solution to the problem without collective effect could be treated as the inversion of a non-linear operator A_0 (3.9). The general problem, however, includes collective effects and therefore the operator will not be A_0, but a more general (unknown) version of it A_x. This operator, which cannot be analytically inverted and is only realized through numerical particle tracking, gives the same compression vector $f_0 = (E_1, E_2, Z_1, Z_2, Z_2', Z_2'')$ for a set of RF parameters $x = (V_{1,1}, \phi_{1,1}, V_{1,3}, \phi_{1,3}, V_2, \phi_2)$ different than the result of the analytical calculation x_0:

$$A_x(x) = f_0 \tag{3.10}$$

The iterative procedure is then the following: each iteration starts with a set of RF parameters x_n, which is either the result of the analytical solution in the first iteration x_0 or the result of the previous iteration. The resulting compression vector $f_n = A_x(x_n)$ is calculated with particle tracking. The residual in the resulting parameters (the difference between the resulting and the desired set) is then used to obtain the correction to f_0 for the next iteration $g_{n+1} = A_0(x_n) + f_0 - f_n$. This g_{n+1} is in turn used to calculate the new RF parameters x_{n+1} analytically as $x_{n+1} = A_0^{-1}(g_{n+1})$. The vector x_{n+1} is the input for the next iteration. This process is repeated until the selected convergence criteria on the resulting compression vector f_n is met. The iteration is summarized in the next equation:

$$x_{n+1} = A_0^{-1}[A_0(x_n) + f_0 - A_x(x_n)] \tag{3.11}$$

The vectors f_i are the resulting compression vectors of each iteration and the vectors g_i are the corrections to f_0, so that the final set of RF parameters on the N-th iteration are $x_N = A_0^{-1}(g_N)$, where x_N produce the desired compression through numerical tracking $A_x(x_N) = f_0$ within the chosen precision.

In order to reduce the long iteration time due to the time-expensive full 3D particle tracking, the inclusion of the collective effects into the solution is done in two steps. First, a routine in MATLAB (by *MathWorks*, version R2012b was used) is used for fast tracking with wakefields of the longitudinal phase space (1D tracking). This routine called RFTweak, uses fast particle tracking with wakefields [19] to follow the behaviour of the longitudinal phase space of a particle distribution through an accelerator [18]. For the double-bunch distribution case, this process takes in most cases about five iterations to achieve a precision better than 1% for all the compression parameters. The CPU time required is less than five minutes in just one processor.

The second step is finally the full 3D tracking. These simulations are done using the ASTRA code for the tracking through the whole linac except for the compressors, where the CSRtrack code is used. This is due to CSRtrack taking into account the effects of the coherent synchrotron radiation [16] generated by the bunch during the bending in the compressors. The same iterative procedure is applied here using as starting point the final RF parameters produced by the 1D tracking. In this case, however, the simulations take far longer time and computational power: one iteration requires about four and a half hours in 64 processors to track 250000 particles from before ACC1 until just after BC3, where the compression parameters are measured. The whole convergence requires around ten iterations.

While the main reason for this two-step separation is the reduction of the total computing time when compared to just iterating with full 3D tracking, the double-bunch case has a more noticeable effect. For a single bunch, the analytical calculation based on the compression of the densely populated center of the bunch is a very good starting points for the subsequent numerical optimization. With two bunches, this step considers mainly the gap between the two bunches, with the two density maxima well outside this area. Specially the influence of the wakefield of the leading bunch on the trailing one is severe. Therefore, it is very useful to have the 1D tracking as a first step in the optimization of the RF parameters, since it requires very little time per iteration and the convergence behaviour of the procedure can be quickly checked in each different case.

3.1.3. Estimation of RF Tolerances

While [13] shows an analytical estimation of the RF tolerances, that method has proven to give wrong results in the double-bunch case. This is due to its use of the estimated analytical compressions for a determined set of RF parameters, which, due to the huge effect of wakefields on the second bunch, are very different from the real values. For example, where the 3D simulations predict the desired compression, the analytical estimation would produce overcompression (the second bunch taking over the first). The RF tolerances can however be estimated from simulations.

Let Θ be the maximum allowed variation in the global compression due to a voltage variation ΔV in one of the accelerating modules:

$$\left|\frac{\Delta C}{C}\right| \leq \Theta \tag{3.12}$$

Neglecting second order terms, that inequality can be written as:

$$|\Delta V \partial_V C| \leq \Theta |C| \tag{3.13}$$

Where $\partial_V C$ is compact notation for the derivative of C with respect to V. Now the tolerance is defined as the maximum allowed relative variation in the voltage $\Delta V/V$:

$$\theta_V = \frac{\Delta V}{V} = \frac{\Theta}{V}\left|\frac{C}{\partial_V C}\right| \tag{3.14}$$

The tolerance can be written with respecto to Z instead of C with $C^{-1}\partial_V C = -Z^{-1}\partial_V Z$. The same reasoning can be done for the phases, and the tolerances are defined as:

$$\begin{aligned} \theta_V &= \frac{\Theta}{V}\left|\frac{Z_N}{\partial_V Z_N}\right| \\ \theta_\phi &= \Theta\left|\frac{Z_N}{\partial_\phi Z_N}\right| \end{aligned} \tag{3.15}$$

This tolerances give upper boundaries for the possible changes in the voltages and phases. This means, that the relative change in the compression will be less than Θ if the relative change in the voltage is less than θ_V *or* if the relative change in the phase is less than θ_ϕ.

Since the analytical estimation of these derivatives does not produce reliable results, it is better to estimate them via numerical simulations. This is done by calculating the resulting compression for voltages or phases around the calculated optimized RF parameters, and then approximating the derivatives to first order as $\partial_V Z = \Delta Z/\Delta V$ and $\partial_\phi Z = \Delta Z/\Delta\phi$. The results are shown in the next Section.

3.2. Method application

This section will show the application of the method explained on the previous section. The selection of the starting distribution and the compression scenario are discussed and an example of results from the numerical simulations will be shown.

3.2.1. Simulations of the RF Gun

The first step of the process is to simulate the generation of the bunch on the photocathode and the injection with the RF gun. ASTRA's generator program can create a wide array of particle distributions, including the simulation of the emission from the cathode. It also takes into account the fields generated by the particles (and the mirror charges on the cathode) and the effect of these onto the distribution, making it ideal for subrelativistic particles.

In the FLASH case, the laser pulse has a temporal gaussian profile with a σ of 6.3 ps and a relatively flat transverse profile with a variable size (see Section 4.1.1). These features are available by default in ASTRA, so the two bunches can be generated by just adding two distributions with the proper charges and the desired delay between them. An example of the generated input distribution for the simulations is shown in Figure 3.3. The first pulse to arrive at the cathode is on the right. Note that the reference $\Delta t = s = 0$ is in the middle of the first bunch. The reference does not need to be placed in the middle of the distribution until the process for the determination of the working point starts. In the gun simulations, the position of the reference only influences the accelerating phase in the gun.

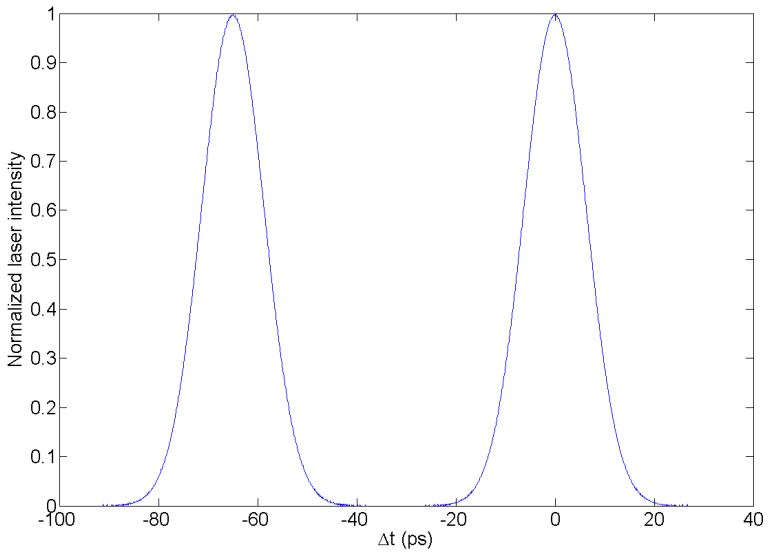

Figure 3.3.: Two identical lasers pulses for the generation of two bunches with the same charge.

The RF gun consist of a 1.5 cell L-band cavity operated at 1.3 GHz on the TM_{010} mode, followed by a solenoid with a field of 0.163 T to focus and match the beam to the booster section [3]. These elements are both simulated in ASTRA with the help of external files containing the field profile, which are scaled to the desired field values. ASTRA then simulates the emission of the particles from the cathode and the transport of these until the end of the gun. One of these simulations takes about one and a half hours in one processor, so it is easy to generate a wide array of possible scenarios. The following figures show the resulting longitudinal distribution after the simulation, for several different charges: 180 pC, 100 pC and 50 pC. The delay at the cathode was kept constant at 65 ps. The head of the distributions is always on the right.

Figures 3.4, 3.5 and 3.6 show the influence of the charge of the bunches on the particle distribution after the gun, as long as the other properties of the laser pulse remain the same. Note that the maximum energy would be in the point between the bunches, and as the charges

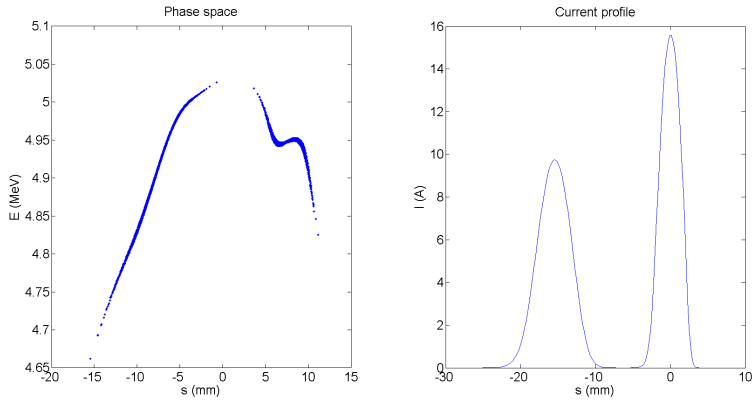

Figure 3.4.: Distribution after the RF gun for 65 ps delay on cathode and 180 pC charge on each bunch.

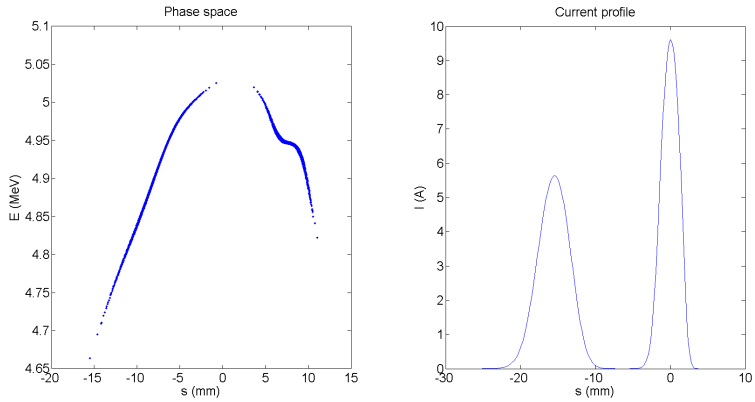

Figure 3.5.: Distribution after the RF gun for 65 ps delay on cathode and 100 pC charge on each bunch.

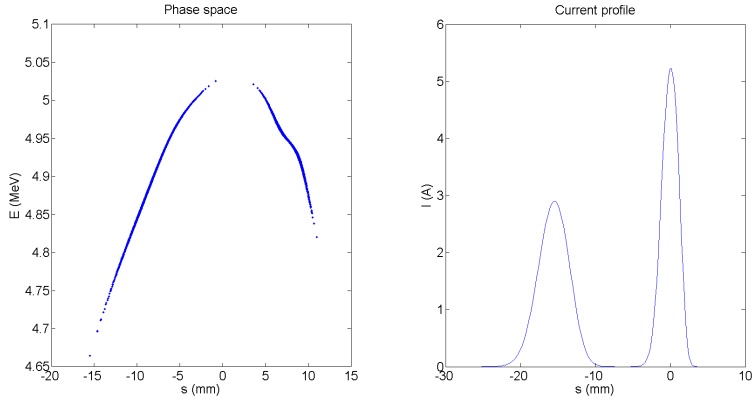

Figure 3.6.: Distribution after the RF gun for 65 ps delay on cathode and 50 pC charge on each bunch.

become lower, the phase space starts to look more like a cosine wave with the maximum in the middle. The deformation from this cosine-like shape are due to space charge fields: the head of the distribution "sees" the whole charge of the bunches behind it, which accelerate it and creates that characteristic S shape. The tail, on the other hand, sees the charge of the bunch in front of it, so it is decelerated, creating a slight deformation in the center of the trailing bunch.

The two bunches can have different charges, the pulse stacker, explained in Section 4.1, is designed to allow arbitrary charge ratios. Figure 3.7 shows the distribution of a beam with a 180 pC driver bunch and a 100 pC witness bunch compared to the case in Figure 3.5 with 100 pC in both bunches. For the case with the higher charge on the leading bunch, the trailing one has lower energy and the bunches are slightly more separated.

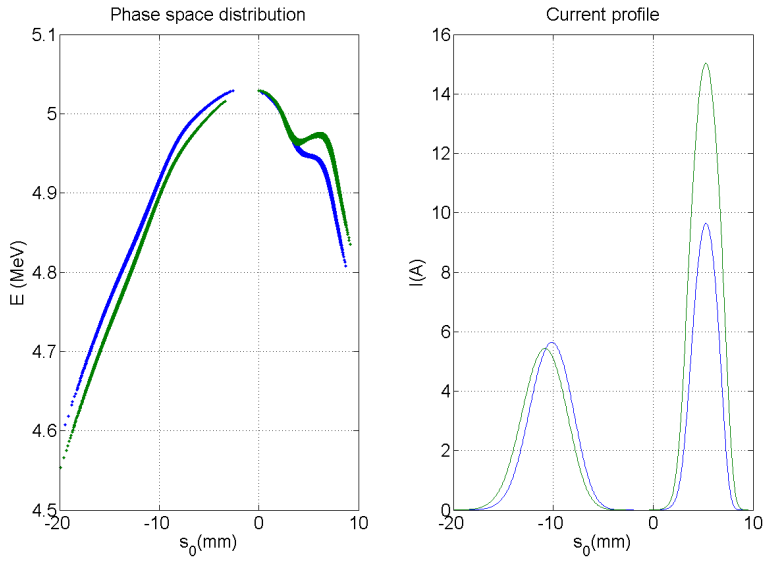

Figure 3.7.: Distributions after the RF gun for 65 ps delay on cathode, with 100 pC on both bunches (blue) and 180 pC on the first bunch and 100 pC on the second (green)

From now on, all the examples use the distribution shown in Figure 3.4. Other distributions may give different results, but the reasoning remains the same.

3.2.2. Selection of the compression scenario

As mentioned at the end of Section 3.1.1, the compression scenario is defined by the design energies in the compressors E_1 and E_2, the fixed bending radii in them r_1 and r_2, the inverse compression factors Z_1 and Z_2 and the derivatives of the inverse global compression factor Z_2' and Z_2''. The double-bunch case adds also another degree of freedom: the initial delay between the pulses on the cathode. The selection of the compression scenario has to be based not only on the desired final bunch, but also on the technical constraints of the accelerator in the form of maximum voltages in the accelerating modules.

Initial delay between laser pulses

Different initial delays on the cathode will result in different distributions after the RF gun, and therefore in different RF parameters required for the same compression scenario. The delay should be chosen as low as possible, in order to make the uniform compression of the distribution easier. If the delay is too big, the first bunch needs to sit too far off-crest in order for the second bunch to also be on the flank and not on the crest (or even on the other flank), which would result in the wrong energy chirp for the bunch and therefore the wrong compression. On the other hand, smaller delays mean higher effect of the wakefields from the first bunch onto the

second, which are already fairly noticeable since the particles are not relativistic yet, or the two bunches could even blend in the RF gun, thus resulting in just one long, deformed bunch instead of two. Therefore, a compromise needs to be reached.

For the smaller delays, the bunches are brought together during the injection on the gun. This defines a minimum value for the possible delays. As a reference, the bunch lengths from the cases in Section 3.2.1 are approximately 8 mm for the first bunch and 10 mm for the second one, so a distance between peaks of at least 9 mm is needed to avoid the overlapping of the bunches.

For the higher delays, on the other hand, there is also a maximum value, which is given by the possible phases for injection in the gun. Figure 3.8 shows the resulting energy after the gun for a single particle as a function of the injection phase, simulated with ASTRA. From the setup of ASTRA, the phase 0 is by definition the phase of maximum energy gain and particles arriving on a later time will have positive phases. There is only a certain window of phases which are acceptable before the energy spread on one of the bunches becomes too high, or even some particles are lost. From the plot, it can be seen that at a phase of around 70° the particle would not even be accelerated in the gun. However, the maximum phase would be at around 50°, where the particle would have an energy of only 3 MeV, already more than 25% less than the maximum energy gain. This translates into a maximum delay of approximately 100 ps (with the RF frequency of 1.3 GHz).

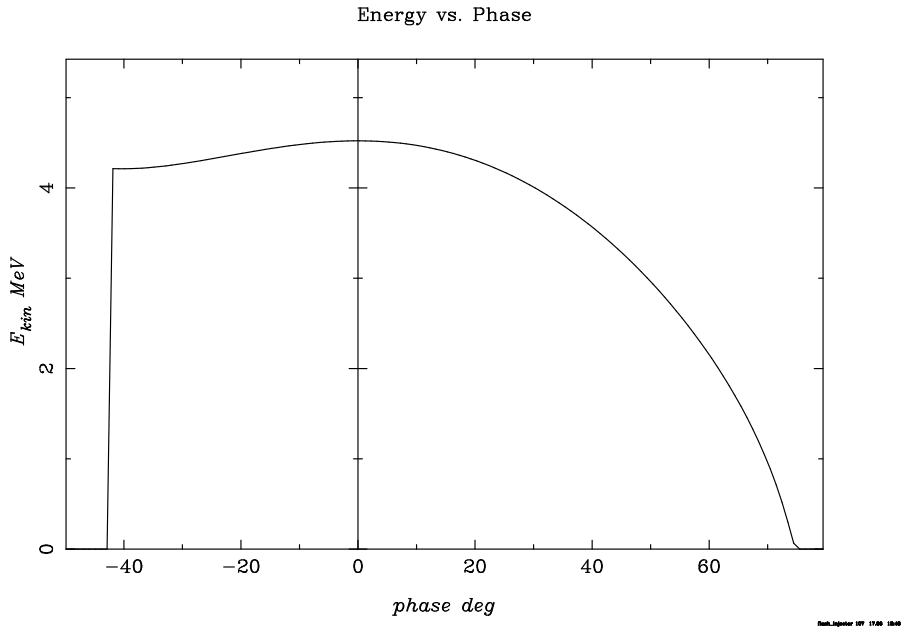

Figure 3.8.: Energy gain in the gun as a function of the RF phase (imported from ASTRA).

Figure 3.9 shows the resulting distance between the bunches after the RF gun as a function of the initial delay on the cathode. The distance is defined as the distance between the peaks in the current profile and is the result of the ASTRA simulations for the case in which both bunches have a charge of 180 pC. For clarity, the distance is displayed in picoseconds. The bunches would start to overlap at around 30 ps, so the minimum acceptable initial delay would be around 37.6 ps.

The figure shows that the distance after the gun is a linear function of the initial delay. The slope of 0.78ps/ps indicates that the bunches are slightly compressed during the acceleration in

the gun. This is due to velocity bunching: since the particles are not relativistic, the particles that have the maximum energy do not have an on-crest phase on injection, but rather a phase slightly on the flank, that guarantees that when they are accelerated to c, they sit on-crest. That means that particles injected after these experience a higher accelerating field at the beginning, thus being accelerated faster and having a shorter flight time than those in front.

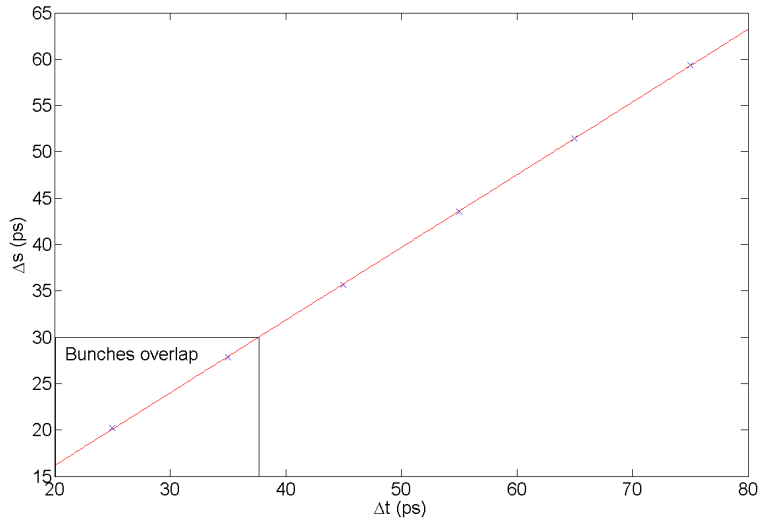

Figure 3.9.: Distance between bunches after the gun Δs as function of the initial delay between laser pulses Δt

Generating a plot like this with the desired charges can help determine which delay is the best in each case. The lower delays produce just one bunch, while higher delays will not make it through the gun. Other considerations for the selection of the delay on the cathode include the maximum compression required, which will be limited by the technical constraints of the accelerator, and the final energy difference between the bunches. Due to the off-crest acceleration, both bunches will have different energies, and this difference may be important for example for the two-color FEL.

Energies in the compressors

The possible energies are given by the maximum accelerating voltages achievable in the accelerating modules. The maximum achieved voltages in FLASH are approximately 161 MV in ACC1, 20 MV in ACC39 and 360 MV in ACC2 and ACC3 together. However, the standard operating energies are around 145 MeV in BC2 (the first bunch compressor) and 445 MeV in BC3 (the second one), so choosing energies around these values may be more desirable.

Another consideration with two bunches is that the distribution is longer, so the bunches need to sit further down the flank in order to have more or less the same energy chirp in both bunches, so that the compression is the same. This will cause the maximum achievable energies to be lower. Due to its higher frequency, the same length increase in the bunch translates in more phase difference in ACC39 than in other modules, compared to standard operation. This greatly reduces the maximum effective voltage of ACC39.

For the simulations shown in Section 3.2.3, energies of $E_1 = 135$ MeV (energy in BC2) and $E_2 = 450$ MeV (energy in BC3) are chosen. This values leave some margin for corrections with respect to the maximum voltages of the accelerating modules mentioned before.

Bending radii

The choice of the bending radius will affect both the required voltages in the accelerating modules and the overall shape of the bunch. Lower bending radii will generate higher amounts of synchrotron radiation, which may affect the quality of the bunch (transverse size, divergence, etc), but will also lower the voltage requirements in the previous modules. In C and S-chicanes, the R_{56} is proportional to the square of the bending angle [15], which means that the lower the radius (the higher the angle), the higher the R_{56}. Since the compression is being approximated by equation (3.3), a lower beding radius will translate, through the higher R_{56}, in lower requirements for the energy chirp in order to achieve the same compression. That means that the bunches can be transported more on-crest, which in turn means more possible energy for the same accelerating voltage, which may be desirable.

However, there are several ways to optimize the choice of bending radii in the bunch compressors. For example, a higher bending radius in the first compressor will lower the voltage required in the second accelerating modules, since it will need a higher energy chirp from the first accelerating and thus the contribution to the total energy chirp from the second modules will need to be lower.

In FLASH, the bending radii need to be between 1.4 m and 1.93 m for the first bunch compressor (BC2) and between 5.3 m and 16.8 m [17] for the second (BC3). In the double-bunch case, choosing the lowest radii possible is desirable, since the voltage constraints are already limiting the available energy. Fortunately, in standard operation in FLASH, the bunch compressors are set to almost the lowest radii, so choosing them will help when trying to do operate the two bunches simultaneously with other modes.

For this example, the bending radii are chosen as the minimum for BC2 ($r_1 = 1.4$ m), to further lower the voltage requirements in the first two accelerating modules, and as the standard FLASH value for BC3 ($r_2 = 6$ m).

Compression factors and shape of the compression

While the total compression factor can usually be defined as the ratio between the final and the initial peak currents in the single-bunch case, this definition is only applicable because the reference particle sits in the middle of the bunch. In the double-bunch case, the reference particle was placed in the middle of the distribution, far from the two current peaks. Therefore, it is more useful to define the total compression factor as the ratio between the initial and final distances between the bunches. For example, in the case shown in Figure 3.4 with an initial delay of 65 ps, the distance after the gun is 15.4 mm. If a final distance of 100 μm is required, the global compression factor should be chosen as $C_2 = 100\ \mu\text{m}/15.4\ \text{mm} = 154.25$.

As for the first compression factor, it presents more possibilities. A lower compression factor in BC2 is preferable, in order to reduce the effects of self-fields and wakefields until the energy is higher. However, if the first compression is too low, it may require too much compression in the second stage and conflict with the voltage constraints (due to higher requirements in energy

chirp). A lower compression in the first stage may also be difficult to achieve due to voltage constraints on the 3rd harmonic module. If the energy chirp needs to be lower, the bunches need to sit more on-crest in the first acceleration, and then the 3rd harmonic module needs to have a higher effect in order to linearize the distribution.

Figure 3.10 shows the required voltages in ACC1 and ACC39 as a function of the first compression factor. The values all come from the 1D-simulation iterations and are normalized to the approximate maximum voltages (161 MV for ACC1 and 20 MV for ACC39). In this case the compression was chosen to be uniform ($Z_2' = Z_2'' = 0$).

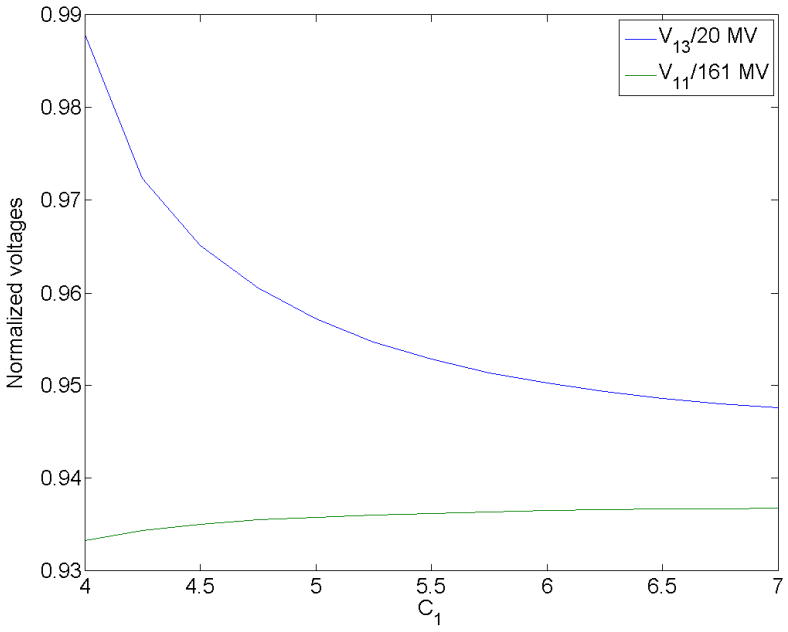

Figure 3.10.: Voltage in ACC1 (green line) and ACC39 (blue line) normalized to maximum value as function of the first compression factor (C_1).

As Figure 3.10 shows, the voltage requirements for ACC39 are lower, the higher the first compression is, while for ACC1 is the other way around. However, ACC1 has plenty of margin, while ACC39 already requires a voltage around the maximum for $C_1 = 4$. In order to leave some margin for corrections in the 3D simulations and for the adjustment of the global parameters (Z_2' and Z_2''), the compression factor is chosen to be $C_1 = 6$. The results of the numerical trackings later on show that, while close to the maximum, the required voltage in ACC39 is still achievable.

Finally, the shape of the compression should be chosen to actually achieve the desired bunch shape. By defining the first and second derivatives of the (inverse) compression factor, the compression in the middle of the bunch is being approximated by a second degree polynomial, which means that there can be a maximum or a minimum in the middle. This will be determined by the sign of the second derivative Z_2'': a negative value will make the inverse compression Z_2 to have a maximum in the middle (a minimum for the compression factor C_2), while a positive value will mean a minimum for Z_2 (a maximum for C_2). A maximum in C_2 should be chosen when the total compression of the distribution needs to be higher than the compression in the bunches, and this will be decided by the desired final currents and the desired final distance between the bunches.

One way to choose the value of Z_2'' can be the following: if Z_2' is set to 0 for the moment, the inverse compression around the reference $s = 0$ is approximated by equation (3.16).

$$Z_2(s) = Z_2(0) + \frac{Z_2''(0)}{2} s^2 \tag{3.16}$$

Now Z_2'' can be chosen so that the compression in a specific point has a desired value, say, the compression in the middle of the driver bunch. If the driver bunch has an initial peak current of 15 A and a final peak current of 2.5 kA is needed, the compression factor in the position of the driver C_d should be around 160. Z_2'' is then given by just solving equation (3.16):

$$Z_2'' = \frac{2(1/C_d - Z_2(0))}{s_d^2} \tag{3.17}$$

Where s_d is the initial position of the driver's peak current. The resulting value for the driver's current will of course be different if Z_2' is chosen different than zero, but this is a good starting point for the value of Z_2''. The number can be adjusted later on if need be.

As for the first derivative Z_2', it can decide where the extremum of the compression actually is along the bunch. This is helpful in fine-tuning the final distance between the bunches, though choosing a high value for it would deform the bunches from the desired shape, so it is better to leave it at value close to zero (between -1 and 1).

The effect of these two parameters will be discussed more in depth in the next section with the 1D tracking procedure.

3.2.3. 1D simulations

The 1D simulations are done in MATLAB with the help of RFTweak [18], a routine that uses fast particle tracking with wakefields [19] to follow the behaviour of the longitudinal phase space of a particle distribution through an accelerator. These calculations are really fast when compared to 3D particle tracking: an initial step, that calculates the effects of wakefields and self-fields scaled from a source file to the particular particle distribution, takes one to five minutes. After that, each simulation with a new set of RF parameters takes about 0.4 s in just one processor. It is a computationally-cheap procedure to check the desired compression scenario. As it will be shown later, the results from the 3D simulations, regarding the phase space distribution of the bunches and the current profile, are not very different from those from 1D. Due to this, the most part of the studies can be done with the fast 1D tracking, leaving the time-consuming 3D tracking for the final optimization step of the RF parameters and the estimation of the RF tolerances.

Figure 3.11 shows the typical results form the 1D tracking, from the starting gun distribution up to after the second compression stage (after BC3), again with the example distribution (65 ps delay and 180 pC charges). The target and the resulting compression scenarios are shown in Table 3.1, and the bending radii were those chosen in the previous section. This is the result of just ten iterations with the iterative procedure of Section 3.1.2, taking less than five minutes in one processor.

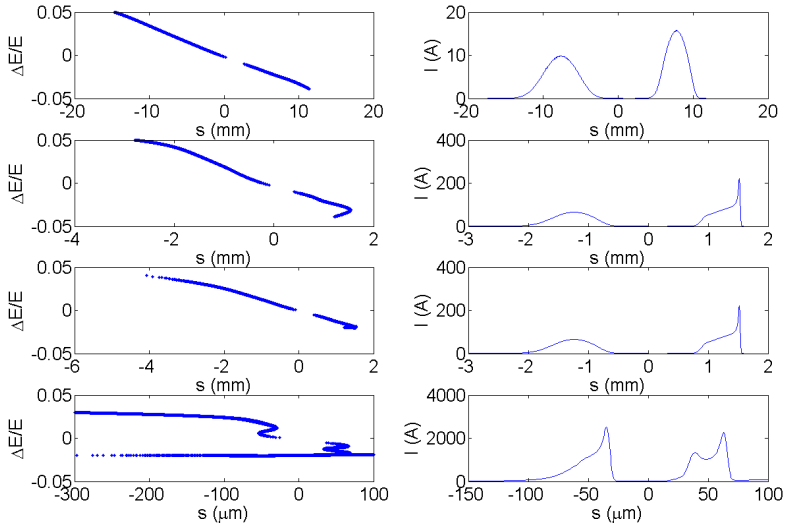

Figure 3.11.: Results of 1D tracking for the example distribution. Phase space plot and current profile before BC2 (first row), after BC2 (second row), before BC3 (third row) and after BC3 (fourth row).

	Target CS	Resulting CS
E_1	135 MeV	135 MeV
E_2	450 MeV	450 MeV
C_1	6	6
C_2	154.24	154.27
Z_2'	0.8 m^{-1}	0.8004 m^{-1}
Z_2''	-7.8 m^{-2}	-7.794 m^{-2}

Table 3.1.: Target compression scenario for the example and resulting values after 10 iterations.

The compression scenario was chosen to have a final distance of 100 μm and a final peak current in the driver of around 2.5 kA (with Z_2'' calculated as in equation (3.17)). This example shows the good convergence of the iterative procedure: in 10 iterations, the difference between the desired values and the resulting ones is less than 0.1% for all the parameters. The distance, measured between the first current peak of the driver and the current peak of the witness, is indeed 100 μm, and the peak current is around the desired value. There are, however, some unordinary features in Figure 3.11, namely the very sharp peaks in the current profiles between BC2 and BC3, and that deformation of the phase space from $s = 50$ μm to $s = 100$ μm after BC3. These features, while fairly small, are still there and should be clarified with the help of the full 3D tracking. Notice as well that after the second compression, both bunches show a very long tail in phase space, but it corresponds to a small number of particles, as show in the current profile.

The RF parameters are also well within technical constraints. These are shown in Table 3.2 along with the analytical RF parameters of the starting point. The biggest difference between

the analytical and the iterated set is in the voltage of ACC39 V_{13}, which increases about 5% after the iterations. This is consistent with the parameter V_{13} being the most sensitive one.

	V_{11}	ϕ_{11}	V_{13}	ϕ_{13}	V_{23}	ϕ_{23}
Analytical	149.78 MV	9.0°	18.05 MV	187.74°	332.77 MV	18.81°
1D tracking	150.88 MV	9.94°	19.13 MV	189.98°	334.23 MV	19.82°

Table 3.2.: Analytical RF parameters and resulting set after 1D iterations.

With the 1D simulations some aspects of the compression scenario can be discussed in more detail, namely the effect of the compression shape parameters Z_2' and Z_2'', and the fine tuning of the final distance and energy spread.

Effect of Z_2'' on the bunches

As mentioned in Section 3.2.2, the second derivative of the inverse compression factor Z_2'' determines whether there is a maximum or a minimum in the compression near the center of the distribution (or wherever the reference $s = 0$ is placed). This is helpful because not always will a uniform compression along the bunch be desired. If Z_2'' is calculated as in equation (3.17), the compression in $s = 0$ will be a maximum or minimum depending on the value chosen for the compression of the driver C_d. In the case shown in Figure 3.11 there is little difference between the global compression and that of the driver, thus resulting in a small value for Z_2''. Since the shape of the compression curve is just approximated by the two derivatives, this small value will have little effect compared for example to the case $Z_2'' = 0$. However, bigger values have very noticeable results.

Figure 3.12 shows the effect of the value of Z_2'' on the final current profile. The same initial distribution used in Figure 3.11 is iterated for different values of Z_2'', with the rest of the compression scenario equal to that in Table 3.1. The blue curve shows the result for $Z_2'' = -7.8$ m^{-2}. The green curve shows the result for $Z_2'' = 500$ m^{-2}. In this case, the positive value of Z_2'' will make the compression in $s = 0$ a maximum, so the compression quickly decreases for positions away from the reference. This will cause the bunches to be further appart from each other and to have lower peak currents. On the other hand, the red curve shows the result for $Z_2'' = -500$ m^{-2}. This case has negative sign, so it will result in a compression minimum in $s = 0$, as it was the case with $Z_2'' = -7.8$ m^{-2}. However, for the higher value of Z_2'' the compression factor increases quickly for positions away from the reference, thus bringing the bunches closer together and generating far higher peak currents and lower distance.

Note that changing the value of Z_2'' produces similar results to changing the overall compression. However, Z_2'' should be used for fine tuning of the final distance, in which case the value of Z_2'' would not be changed as much as 500 m^{-2}. The big variation was used to better ilustrate the effect.

Choosing the proper shape for the compression curve along the distribution is key in order to have the desired final distribution, since both the distance between the bunches, and the peak currents depend strongly on how the compression behaves along the distribution.

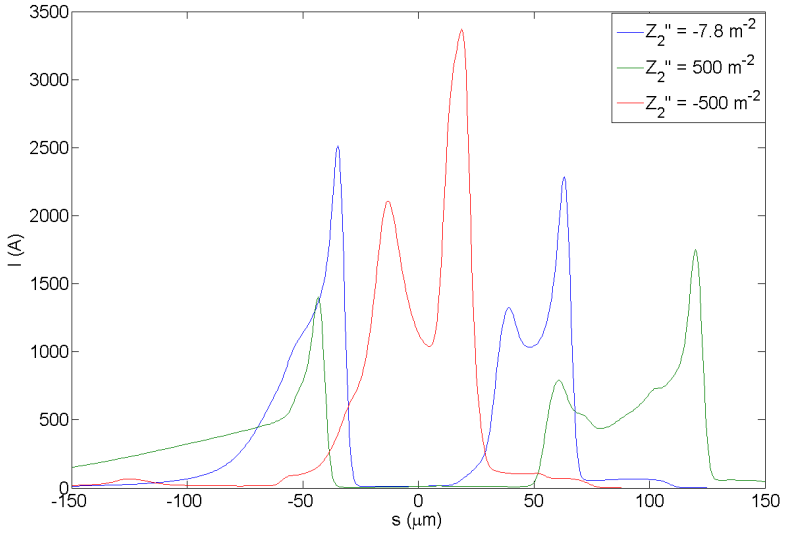

Figure 3.12.: Final current profile for the same distribution, with the compression scenario of Table 3.1, except for different values of Z_2'' indicated in the legend.

Fine tuning of the final distance with Z_2'

On contrast with Z_2'', the first derivative of the inverse compression factor Z_2' will determine where the extremum (maximum or minimum) of the compression is placed. If $Z_2' = 0$, the extremum will be at the reference point $s = 0$. If Z_2' and Z_2'' have the same sign, the extremum will shift to the positive values of s, and to negative values if they have opposite signs. This can be used to fine tune the final distance between the bunches in the range of several μm. Once all the other compression parameters have been selected, shifting the extremum of the compression slightly will return different distances between the current peaks.

This effect is shown in Figure 3.13. The same example distribution (65 ps delay, 180 pC charges) is used for the same compression scenario of Table 3.1 except for different values of Z_2'. By shifting the position of the minimum of compression, the distance between the highest current peaks is shifted from around 80 μm for $Z_2' = 0$ to the target 100 μm for $Z_2' = 0.8$.

Of course, moving the extremum has also other effects. The peak currents are different, and the current profiles change slightly. This effects will be higher for higher values of Z_2', so leaving it at small values ($|Z_2'| < 1$ for example) and adjusting the value of Z_2'' or even Z_2, accordingly will mitigate this deformations. For higher changes in the final distance, either another compression scenario or even a higher initial delay on the cathode should be used.

Figure 3.14 summarizes the effect of both derivatives of the global inverse compression factor Z_2. The figure shows how Z_2 varies with the initial position in the distribution s_0. The red curve would correspond to uniform compression ($Z_2' = Z_2'' = 0$), the green curve shows the maximum in Z_2 created by a negative value of $Z_2'' = -500$ and the blue curve shows the displacement of the maximum by a value of Z_2' different than zero (here $Z_2' = 0.8$). As it was the case in the previous subsection, a big value for Z_2'' has been chosen in order to make the effect more clear.

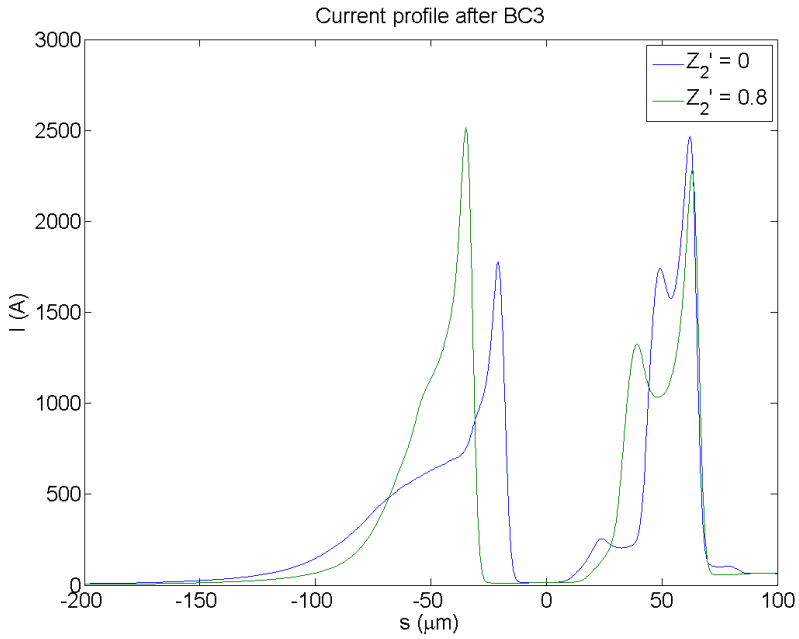

Figure 3.13.: Final current profile for the same distribution, with the compression scenario of Table 3.1, but for different values of Z_2' indicated in the legend.

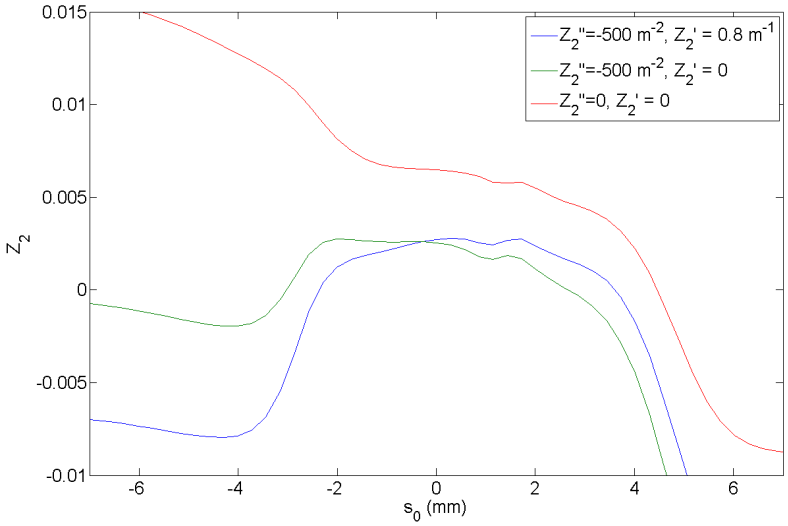

Figure 3.14.: Global inverse compression factor Z_2 as function of the initial position on the distribution s_0 with the compression scenario of Table 3.1 except for the values of Z_2' and Z_2'' indicated in the legend.

Fine tuning of the final energy spread with the initial delay

Apart from the distance between the bunches, another property of interest is the difference in energy between both bunches. This is important mainly because the energy collimator in FLASH has only an acceptance of 3% [5], so the energy spread cannot be higher than that. Apart from this, the fine tuning of the energy spread can be important in the two-color FEL experiments,

since it lets the user determine the difference in wavelength of the two generated colors (equation (2.1)).

Fine-tune this energy spread without changing the final current profile can be done with the initial delay on the cathode. Having a different delay means changing the difference in the accelerating voltage that the bunches experience, which in turn changes the final energy spread. The compression scenario has to be adapted so that the final distance and currents are the same (by changing the global compression C_2 and the compression shape parameters Z_2' and Z_2'') but the energy spread would be slightly different. Figure 3.15 shows two different distributions, with different delays on the cathode but same charges, which are set up so that the compression in the end is the same.

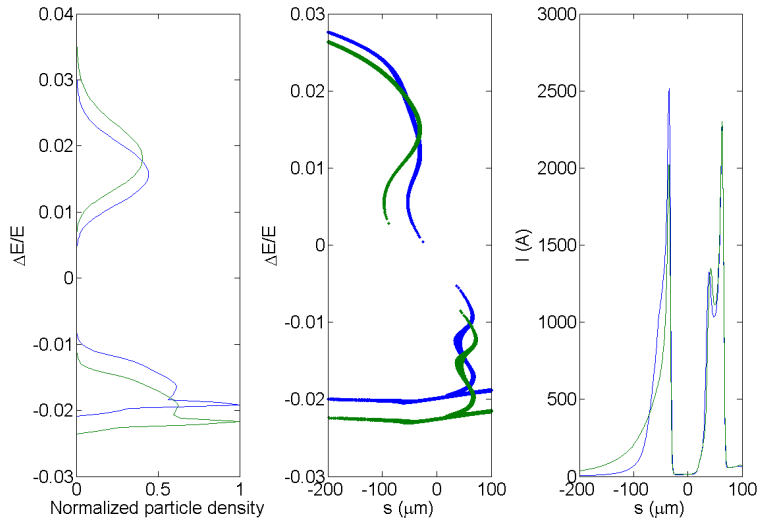

Figure 3.15.: Results for the compression scenario of Table 3.1 but delays on cathode of 65 ps (blue) and 75 ps (green). Energy projection (left), phase space (center) and current profile (right).

Note that the current profiles after BC3 are basically equal, with only a slight difference in the profile of the trailing bunch, due to the difference in the effect of wakefields from the first bunch onto the second, due to the different initial distance. The most noticeable difference is the different energy spread between the bunches. The 65 ps case generates a difference in energy of 14.6 MeV, while the 75 ps case has an energy difference of 17.1 MeV, for an average energy for the whole distribution of 450 MeV. This can be done for the whole array of possible initial delays. Figure 3.16 shows the simulated final energy difference (in MeV) as function of the delay on cathode (in ps), for the same bunch shape and distance between bunches, along with a linear fit for the data. For the smaller delays, the bunches may be brought together, but the general shape of the current profile remains the same.

The energy difference can then be fine-tuned in a range of about 10 MeV for this particular case. However, this is restricted by the compression. Different compressions will yield different energy differences, with lower compressions giving the highest spreads, since the phase differences will be higher. There is also the possiblity to transport the bunches off-crest in the accelerating modules after the compression, but this could result in unwanted behaviour in the energy collimator later on. In the end, anyway, the limit is the energy acceptance of the collimator, so the energy difference should never exceed 3%.

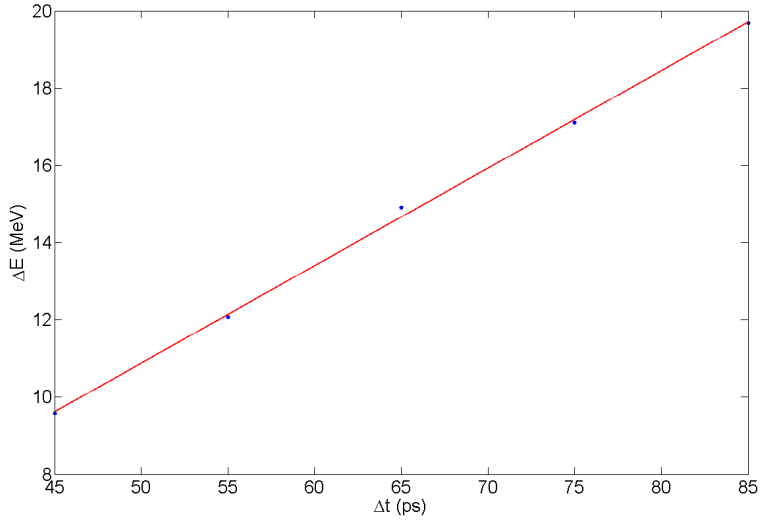

Figure 3.16.: Variation of the energy difference between bunches after BC3 with the initial delay on cathode.

Effect of the charges of the bunches

Changing the charge of the bunches has little effect on which final current profiles can be achieved, provided that the charge ratio remains the same. Figure 3.17 shows the result of having two distributions with same delays on cathode but different charges iterated for the same compression scenario of Table 3.1, that is, the compression is the same, and each case have different final RF parameters. Both results show very similar features, with the main difference begin the lower currents and the shorter bunches generated for lower charges. This is due to the higher effect of the wakefields of the higher charges. In 3D simulations, lower charges also generate bunches with smaller transversal size.

3.2.4. RF tolerances

The 1D simulations were also used in the estimation of the RF tolerances. As mentioned in Section 3.1.3, in order to calculate the dimensionless tolerances from equation (3.15), the derivative of the inverse global compression with respect to the voltages and phases needs to be estimated. This can be done numerically by tracking the bunches through the linac for the same set of RF parameters but with a small variation in one of the voltages ΔV or in one of the phases $\Delta\phi$. For small enough variations, the change of the compression with the voltage or the phase can be approximated by a linear function, and the slope will be the derivate.

Figure 3.18 shows the relative variation of the total compression with the variation of the voltage in ACC1, as calculated with RFtweak. An interval of ± 0.05 MV was taken around the optimized value, and all the other RF parameters were kept constant. There was no noticeable uncertainty observed in the values after several numerical simulations. The plot has the expected behaviour: the compression increases for higher voltages, since it generates a bigger enery chirp, and the points can be well approximated with a linear function, since the interval taken was small enough, even though the behaviour is not completely linear, as expected. The derivative can be

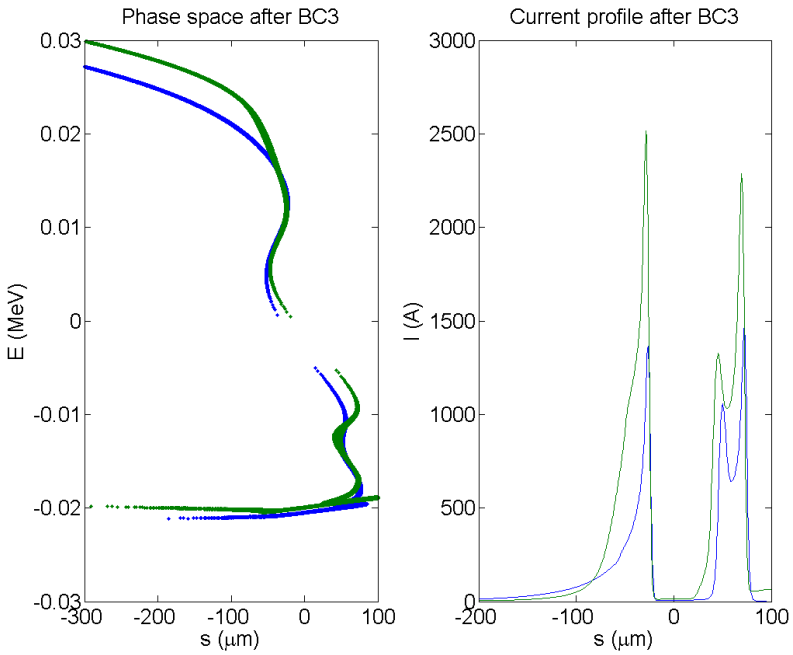

Figure 3.17.: Final phase space and current profile for the same compression scenario but charges of 180 pC (green) and 100 pC (blue) on both bunches

approximated by the slope of the linear fit of the inverse compression vs the voltage (not that of Figure 3.18).

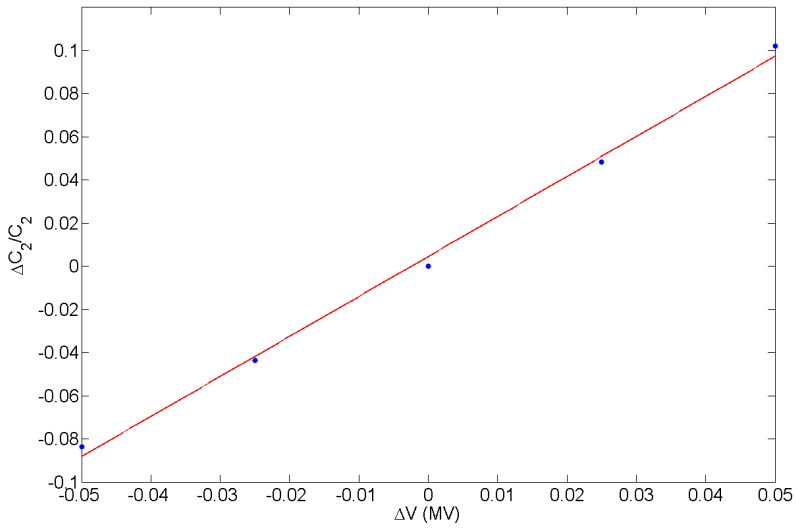

Figure 3.18.: Relative variation of the global compression with the voltage in ACC1 as calculated with RFTweak and linear fit.

This process can be repeated for all the RF parameters. For the example case shown here, and for a maximum accepted variation in the compression of 10% ($\Theta = 0.1$) it results in the tolerances in Table 3.3. Doing the calculation with full 3D simulations produces the same results.

	ACC1	ACC39	ACC23
$\theta_V(10^{-4})$	3.6	14.7	25.5
θ_ϕ (deg)	0.009	0.02	0.1

Table 3.3.: Estimated tolerances in all the RF parameters for a maximum tolerance of 10% for the global compression.

The module with the highest constraints is ACC1 in both relative variation in voltage and in variation in phases. This is due to this module having the biggest effect in the compression in both compressors: a change in the energy chirp given by ACC1 will change the compression in BC2, which in turn will change the energy chirp given by ACC23. The voltage tolerance in ACC1 is just above the stability of the RF (approximately $2 \cdot 10^{-4}$ [20]). The phase tolerance, however, is about half the phase stability, which is about 0.02° [21]. This would mean a variation in the compression of up to about 20% for this particular case. The tolerance can be increased by for example reducing the compression in BC2: choosing a compression of $C_1 = 5$ instead of $C_1 = 6$ brings the tolerance up to 0.0011 and reducing the total compression to $C_2 = 128.5$ (final distance of 120 μm for this initial delay) brings it up to 0.0013. So the phase tolerance in ACC1 can be maximized by changing the compression scenario. This should be taken into account when selecting the compression and the initial delay: since the final distance and currents can more or less be chosen with the different parameters of the scenario, the selection of the initial delay and compressions should take into account the possible tolerances in ACC1.

Note that, compared to the tolerances for a single bunches, like for example those calculated in [13], the double-bunch case has higher restrictions in the RF parameters.

3.2.5. 3D simulations

The 3D simulations are done with two different tracking codes: ASTRA for the accelerating modules and straight sections, and CSRtrack for the bunch compressors. As explained before, CSRtrack keeps track of the coherent synchrotron radiation (CSR) effects during the compression, making it ideal for these parts. Moreover, ASTRA has some problems when it comes to displacing the particle distribution from the original longitudinal direction, due to the way the intra-bunch fields are calculated, so it should not be used in bunch compressors.

In CSRtrack, the lattice of the bunch compressors is defined as just a series of dipoles, the strength of the dipoles is given through their bending radius, and the field is automatically scaled to the bunch's energy. CSRtrack has several different methods to calculate the self-forces in the bunches. The simples of them, the "projected" method, approximates the 3D distribution by a smooth line charge density, neglecting transverse dimensions of the distribution [22]. Note that this approximation is only used for the calculation of the self-charges, but the bunches are still tracked in three dimensions. This method produces very reliable results in a very short time (several minutes in a single processor). The other methods are more realistic, since they take into account the full three-dimensional nature of the distribution, but one run can take up to a week to complete in 64 processors. Due to this long time, the "projected" method is better for the iterative procedure from 3.1.2.

On the other hand, the tracking with ASTRA can be a little more complicated. As it was said before in Section 3.2.1, ASTRA scales the fields in the cavities from external field files, which have to be provided manually. Apart from that, the lattice in ASTRA is more complex, since all the quadrupoles from the accelerator need to be included. The quadrupoles require no external file, but the correct strength needs to be defined. A proper optic is therefore required for the simulations, or else the particles can be lost at some point of the lattice.

Figure 3.19 shows the result from the tracking until after BC3 together with the results from the 1D simulations for comparison. The starting distribution is the same one used for the examples in the 1D simulations, shown in Figure 3.11, which had 65 ps delay on the cathode and both bunches with 180 pC charge. The optimization took about 10 iterations, each of them taking around five hours in 64 processors. The resulting RF parameters are shown in Table 3.4, along with the starting set from the analytical calculations and the results from the 1D tracking.

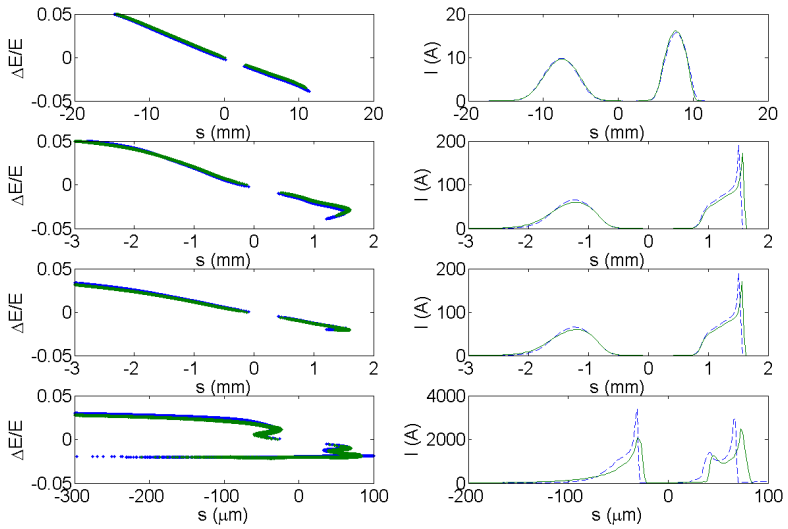

Figure 3.19.: Tracking results from the 1D simulations (blue) and the 3D simulations (green). Phase space (left) and current profile (right) before BC2 (first row), after BC2 (second row), before BC3 (third row) and after BC3 (fourth row).

	V_{11}	ϕ_{11}	V_{13}	ϕ_{13}	V_{23}	ϕ_{23}
Analytical	149.78 MV	9.0°	18.05 MV	187.74°	332.77 MV	18.81°
1D tracking	150.88 MV	9.94°	19.13 MV	189.98°	334.23 MV	19.82°
3D tracking	150.93 MV	9.17°	19.14 MV	187.57°	333.09 MV	18.72°

Table 3.4.: Analytical RF parameters and resulting set after 10 iterations with 1D and 3D tracking.

As it can be seen, the differences between both results are minimal. 3D tracking produces a slightly "smoother" distribution, getting rid of some of the unusual features produced by RFtweak in the phase space after the second compression. The biggest difference is in the final current profile after BC3, where the distance between the bunches is not exactly the same. On the big picture, however, both results are essentially equal. This makes the 1D tracking the best

option for the selection and evaluation of the desired compression scenario, and the 3D tracking can just be used for the final optimization of the RF parameters.

	Selected CS	Resulting CS
E_1	135 MeV	135 MeV
E_2	450 MeV	450 MeV
C_1	6	5.9987
C_2	154.24	153.73
Z_2'	0.8	0.7983
Z_2''	-7.8	-8.4655

Table 3.5.: Selected compression scenario and resulting values after 10 iterations with 3D simulations.

Chapter 4.

Experimental realization

Besides the simulation studies, an experimental verification of the concept of double-bunch generation and compression was done. For this, the laser system that generates the electrons in FLASH needs to be adapted for the generation of two bunches with adjustable delay and charges. Second, the usability of the diagnostics systems in the accelerator had to be validated before compression studies could be done. Due to schedule restrictions, there was no available beam time with FLASH for the compression studies, and only a proof-of-concept experiment was possible in an 8-hour shift. This chapter shows all the work regarding the experimental part.

4.1. Pulse stacker

The pulse split-and-delay system (also callsed "pulse stacker" from now on) is the first component for the generation of double-bunch beams. The normal laser beam is split with the help of a polarizer, each beam is driven through paths of different length and then brought together, resulting in two pulses of identical spatial characteristics (profile shape, transversal size, pulse length) on the same path with a variable delay between them. The properties of the pulses were characterized with a streak camera and an optical powermeter.

4.1.1. Structure and working principle

The structure of the pulse stacker is shown in Figure 4.1. The system is installed in the beam path of the standard injector laser system in FLASH, the so called Laser2. This is a Nd:YLF laser with a wavelength of 1047 nm, which is frequency-doubled twice for a final wavelength in the UV range of 262 nm [23]. The laser has a Gaussian profile in the longitudinal direction with a size of 6.3 ps (sigma) and a Gaussian transverse profile that is cut in size with the Beam Shaping Aperture (BSA in Figure 4.1), a motorized iris with apertures of several sizes from 0.05 mm to 10 mm.

The schematic in Figure 4.1 shows how the splitter works. The beam comes from the bottom side of the picture with 100% linear polarization parallel to the plane of incidence (P polarization) and follows the blue path. It goes though a zero-order half-wave plate (see Apendix B.1) on a motorized mount. This plate can turn the polarization to any degree, so that there is a P component and an S component (perpendicular to the plane of incidence) and is controlled remotely. The beam then goes on to the first beam splitter plate (see Apendix B.2), where the P component of the polartization is transmitted and the S component is reflected. The S component of the beam follows the dotted blue path and has usually the shortest path, which makes

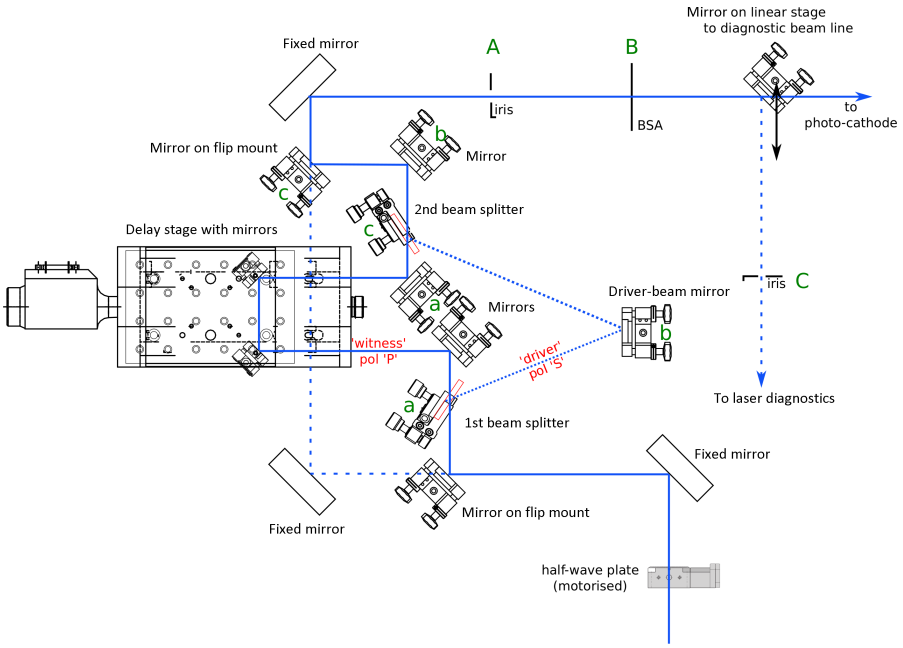

Figure 4.1.: Schematic of the structure of the pulse stacker (the figures of the components were taken from their respective catalogues).

it the pulse that will generate the driver bunch, and will from now on be called "driver beam". The P component (now called "witness beam" in contrast) follows a longer path and passes an optical delay line, consisting of two mirrors mounted on a motorized linear stage. Both beams are brought together on the second beam splitter plate and are inmediately redirected with two mirrors to the standard beam path for the injector laser, which then proceeds to an iris and the BSA, on to the photocathode.

The system has further components for diagnostic purposes. A movable mirror (shown top right in the schematic) can divert the beam towards the diagnostics equipment (powermeter, streak camera, etc). There are also two remotely controlled shutters (not shown for clarity) that allow for each beam to be blocked independently, in order to characterize them separately. Finally, two mirrors on respective flip mounts can be moved down to allow the laser to follow the standard path (dashed blue line) and not go through the pulse stacker. These mirrors have to be operated manually.

The whole system had to be aligned to follow the same path as Laser2 and hit the cathode in the same area of interest. Each beam was adjusted separately, using one mirror for each aperture. The corresponding mirrors and apertures are shown in the schematics with the same green letter: a capital letter for the apertures and a lower case letter for the mirror. Note that the beam splitters are also mirrors for one of the beams, so they were used as such for the alignment.

The properties of the pulse stacker were measured with a streak camera for the delay stage and a powermeter for the intensity of the pulses. The alignment was checked with the help of the so-called "virtual cathode", a Ce:YAG crystal (a scintillator) that is placed at a distance equal to that of the photocathode [24].

4.1.2. Streak camera measurements

The main property of the pulse stacker is the possibility to select the delay between both pulses. This is achieved with the linear stage in the witness beam's path. However, two different characteristics were measured with the streak camera: the delay between the driver beam and the standard arrival time of Laser2, and the delay between the driver and the witness beams as a function of the position in the delay stage. The streak camera is a FESCA-2000 by *Hamamatsu*, which has a temporal resolution of about 200 fs. This resolution is of the same order than the jitter betwee the arrival time of the laser and the global reference of the accelerator (about 200 fs rms (root mean square) [23]) and it is about two orders of magnitude lower than the delays that are of use in this setup.

Delay between driver beam and standard Laser2 arrival time

The delay between the driver and the standard arrival trime is due to the additional beam path through which the laser needs to go in the pulse stacker. This is important to measure, since the arrival time of the laser at the cathode needs to be synchronized with the RF reference, which sets the time reference for all the other subsystems in the accelerator. First, the delay was calculated through measurement of the lengths of the beam paths with a ruler, estimating a delay of be about 400 ps. Using this expected value, the streak camera was set to a proper time window that would allow for both beams to be seen. The window was chosen to be 500 ps, which has a time-step per pixel of about 1.12 ps. Finally, since the setup of the stacker makes it impossible to observe the standard beam path and the driver beam from the stacker simultaneously, the time reference was set with Laser2 and the beam was measured before diverting the beam through the pulse stacker and measuring the driver.

Examples of the raw images of the streak camera can be seen on the left of Figures 4.2 for Laser2 and 4.3 for the driver beam. The raw pictures are signal from the (512x512) pixels of the camera represented in a color code: lower detected intensity is represented by the blue color, and higher detected intensity goes through green and yellow to finally saturate at red. The vertical axis is the time axis (time goes downward in this case), while the horizontal axis is position of the beam in the camera. Due to the jitter between the arrival time of the laser and the reference of the camera, a total of 20 images were taken for each beam.

In order to process this images, the 2-dimensional picture is projected onto the time axis, giving a profile like the one shown to the right of the pictures. The profile is then numerically fit with a nonlinear least-squares method to a Gaussian function plus a constant to account for background noise, resulting in the red curve shown overlapping the projected profiles in Figures 4.2 and 4.3. Once the fitted profile has been extracted from each picture, the results are averaged for each set of images. Both averaged profiles are shown in Figure 4.4. Since the scope of these measurements was not to characterize the intensity, the profiles are normalized to the intensity of Laser2 alone. Note that the driver beam has about 75% of the intensity of Laser2, due to the laser being split in two and one of the beams being blocked. The resulting delay between the pulses, defined as the distance between the intensity peaks, and length of the pulses are shown in Table 4.1. The values shown as errors are the standard deviation of the 20 values.

The results are completely in agreement with the expected values. The delay was expected around 400 ps, which is well within one standard deviation of the measured value. The error in the measurement of the delay is generated by the jitter between the trigger of the streak camera

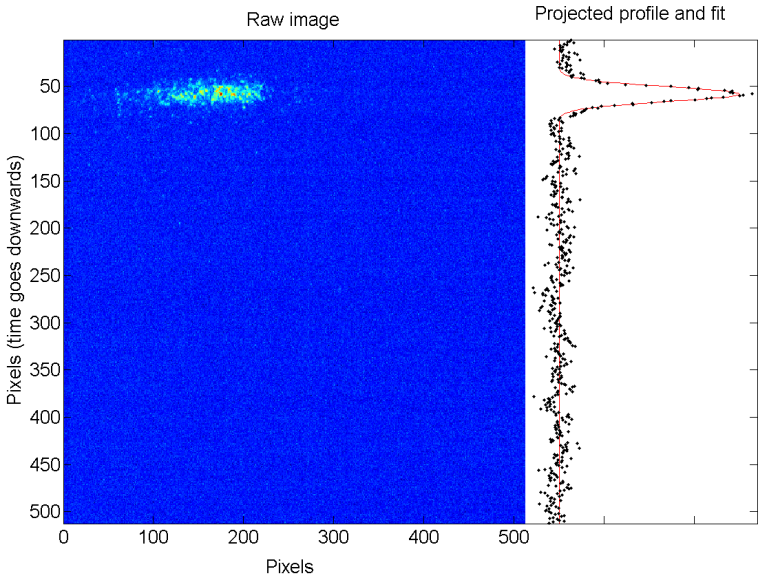

Figure 4.2.: Raw picture of Laser2 with the streak camera, and projected, normalized profile with fit.

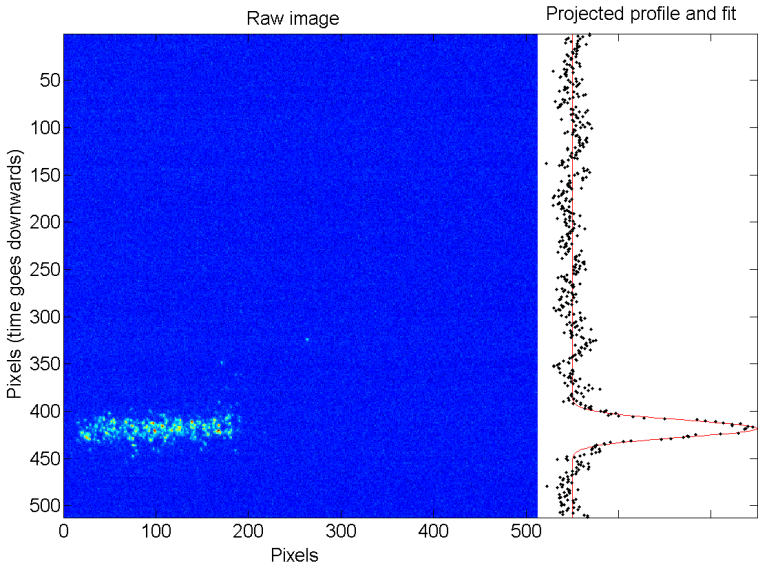

Figure 4.3.: Raw picture of the driver beam with the streak camera, and projected, normalized profile with fit. Time reference is the same as in Figure 4.2.

Delay	(408 ± 19) ps
Pulse length (Laser2)	(6.4 ± 0.4) ps
Pulse length (driver)	(6.3 ± 0.3) ps

Table 4.1.: Delay between Laser2 and the driver beam and measured bunch lengths.

and the laser pulse. This effect is increased by the fact that both pulses were measured separately, so the jitter affected each pulse independently, increasing the error.

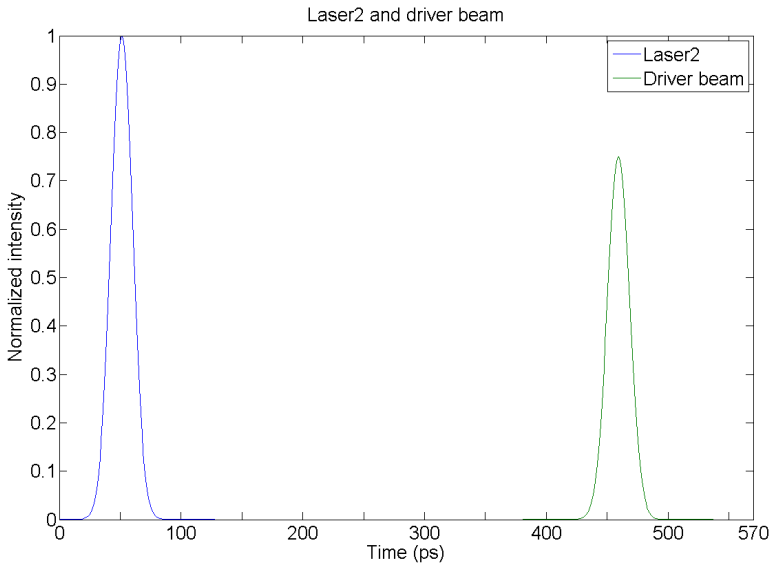

Figure 4.4.: Average longitudinal profiles of Laser2 (blue) and the driver beam (green).

On the other hand, both pulses have essentially the same length, since both the difference in the values and the standard deviation of the measurements is lower than the time resolution of 1.12 ps provided by the chosen time window. The pulse length (one sigma) also turns out to be around the expected value of 6.3 ps.

Delay between driver and witness beams

The delay between the driver and the witness beams can be fine-tuned thanks to the linear stage in the witness's beam path, a table moved by a remotely controlled motor. The linear stage is a precision translation stage from *Physik Instrumente (PI) GmbH*, model M-405.2S. It has a minimum step of 0.1 μm and a repeatability of 0.2 μm. The total travel range is 50 mm which corresponds to a range of variation in the delay of about 340 ps (the beam traverses the extra path twice), which is far more than is going to be used.

First, the relation between the position of the table in its own rails and the position of the motor in the software (number of steps) was calibrated. This was done by manually measuring with a micrometer the table's distance from the end of the rails. The minimum distance is not zero, due to the electrical end-switch that prevents the table from colliding with the motor. The results are then fit to a staight line. The points measured and the linear fit are shown in Figure 4.5. The error bars are too small to be shown.

After that, the delay between the pulses was measured with the streak camera for several positions of the table. The image and data processing is the same as for measuring the delay between the driver and the standard arrival time, except that this time the function for the fit was the sum of two gaussian profiles plus the background constant. Figure 4.6 shows an example of raw image with the projected profile and fit, for the case of 104.4 ps delay. The fitting procedure works well, so long as the starting parameters are well selected. The fitting works properly even when the pulses start to overlap, as is the case in Figure 4.7. Note that the intensity ratio between the pulses is not the same in both images. This was due to a small alignment error in the linear stage, that caused the witness beam to arrive with a different angle when the table is move, so that

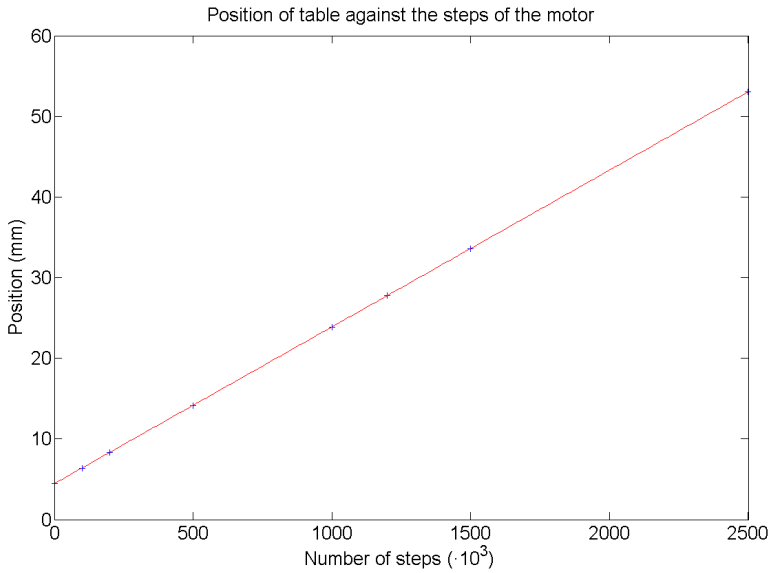

Figure 4.5.: Calibration of the table's position in the linear stage with the steps of the motor.

a different intensity went through the aperture. However, the intensities are not important in this particular measurements, so the error could be neglected and the gain in the different pictures was adjusted to account for saturation. The alignment error was corrected for the experiment with the accelerator.

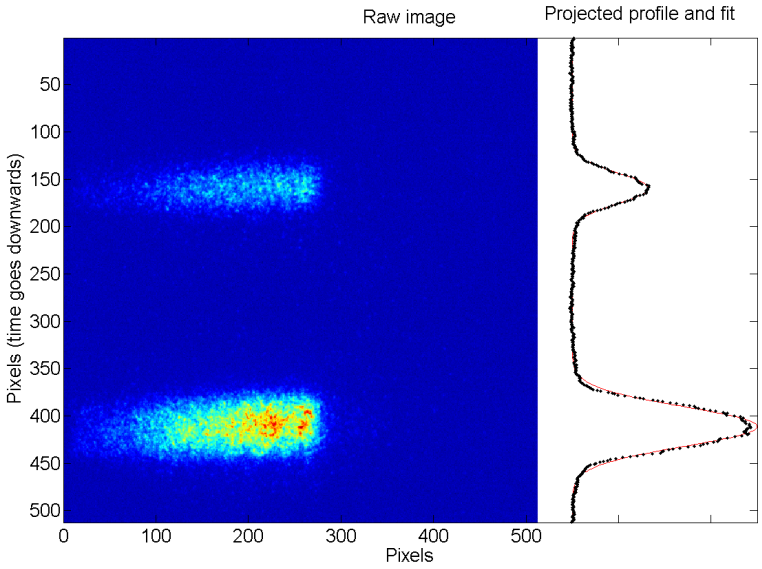

Figure 4.6.: Raw picture of both laser pulses with the streak camera for 104.4 ps delay, and projected, normalized profile with fit.

The delay was measured for a total of 9 positions, corresponding to a range of delays of about 100 ps, with 20 measurements per position, for a total of 180 measurements. The results are shown in Figure 4.8, along with the linear fit. The error bars are once again one standard deviation. The fit works perfectly well, as it was expected, except for one point at the position 32.6 mm. This measurement was taken with a smaller time window, so the measurement was

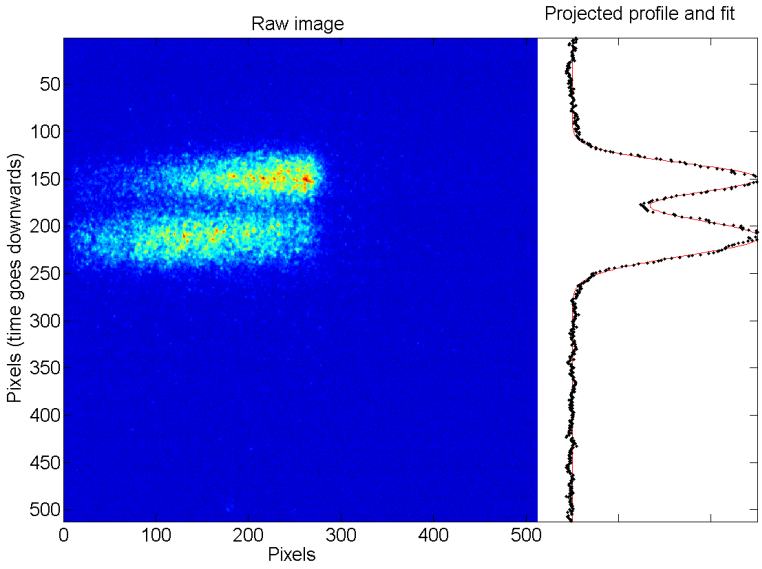

Figure 4.7.: Raw picture of both laser pulses with the streak camera for 25.8 ps delay, and projected, normalized profile with fit. Note how well the fitting works even for overlapping pulses.

more sensitive to the jitter in the arrival times. The measurement could not be repeated since the streak camera was not available again due to maintenance.

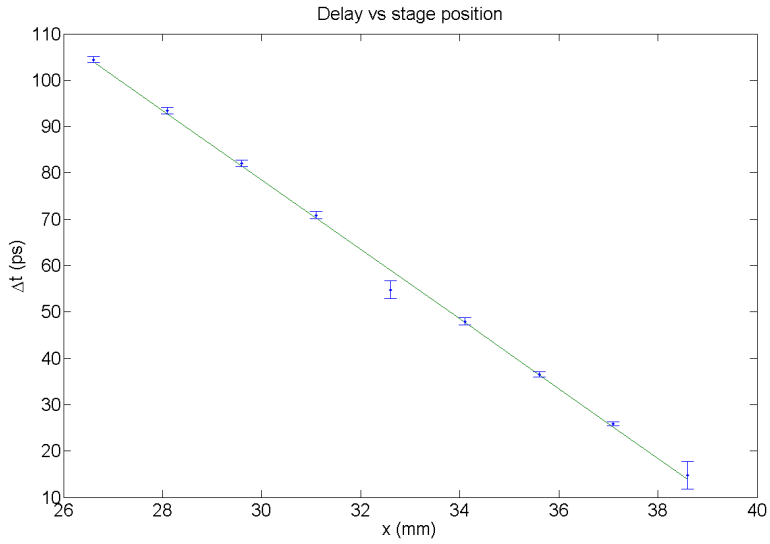

Figure 4.8.: Delay between the pulses Δt as a function of the position x of the linear stage.

The linear fit corresponds to the following equation:

$$\Delta t[\text{ps}] = 303.7 - 7.5x[\text{mm}] \tag{4.1}$$

These results show that the delay can be indeed fine-tuned to a desired value with very good precision. The linear stage behaves as expected and the range of delays is more than enough.

The fit shows a maximum delay of 303.7 ps, so with a range in the variation of the delay of approximately 340 ps, the path in the linear stage could be shortened enough to reverse the role of the beams and make the witness beam have the shortest path. This may be of use at some point, but it is usually better to keep the beam with the fixed path on the driver role, to allow for better synchronization with the global reference of the accelerator.

4.1.3. Intensity measurements

The second main property of the pulse stacker is the posibility to select the intensity ratio between the beams. This is achieved by rotating the polarization of the incoming beam with a half-wave plate (model RZQ 2.15 L from *Bernhard Halle Nachfolger*, which is controlled by a rotational stage from *Owis*, model DMT 40. This motor has an unlimited angle of rotation (does not stop at one complete 360° turn) with a repeatability of less than 0.07°.

The first step of the measurements is, as it was before with the linear stage, the calibration of the angle of rotation with the steps of the motor. The angles were measured with the reference scale that is printed on the motor and are a perfectly linear function of the steps. The error bars are too small to be shown.

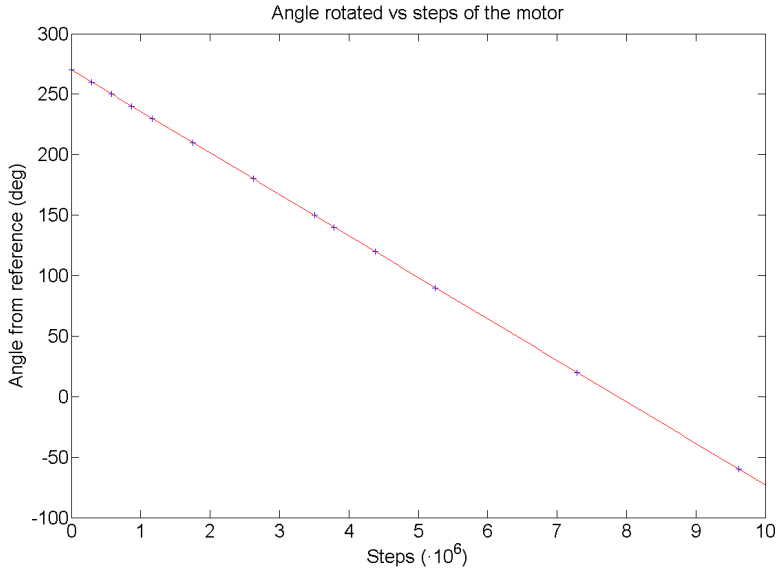

Figure 4.9.: Calibration of the rotational stage with the steps of the motor and linear fit.

After that, the intensity of the beams for different angles was measured. The intensity was measured with a powermeter from *Coherent*, model J5-09. This is a pyroelectric detector, which measures the rate of change of the detector's temperature instead of the temperature itself. This results in reponse times lower than nanoseconds, but still higher than the delays used in this work, so both pulses will show as one.

The typical response from the powermeter is shown in Figure 4.10. Each point represents the voltage output (in mV) for each sampling time. The total sampling time is around 2 ms. The signal shows a linear rise in the voltage, followed by an exponential decay back to the background levels. Note the huge amount of background noise shown in the signal, with a measured rms of 0.4 mV. For the data analysis, two numerical fits are done, again with a non-linear least-squared method. The first fit follows the linear rise of the signal, from a predetermined starting

point, and the second describes the exponential decay until the end of the sampling window. The signal's maximum, which is the value of interest, is then determined by calculating the crossing point between both fitted lines. Example of the fits can be seen in Figure 4.10.

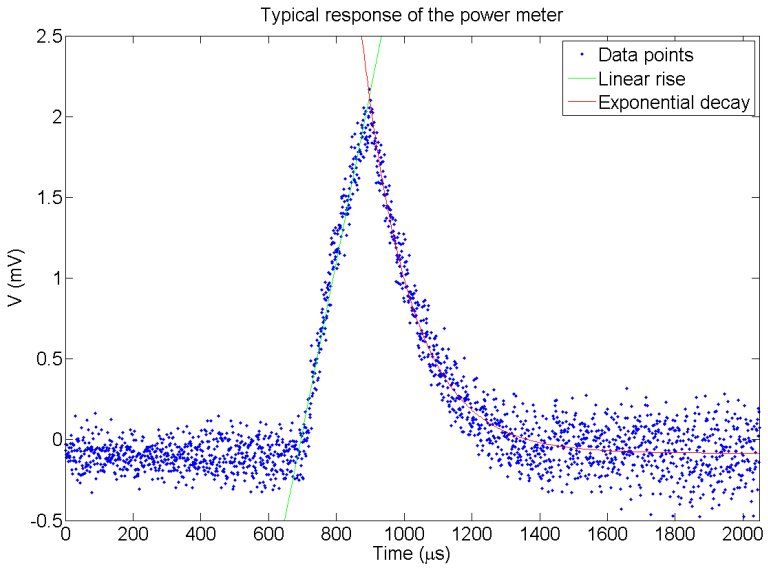

Figure 4.10.: Powermeter reading, with linear signal rise and exponential decay and their respective fits.

The measurement was done by alternatively blocking the beams with the installed electronic shutters and taking readings from both the beams together and separated for a determined angle of the waveplate. In order to minimize the effect of the noise, ten measurements were taken for each angle and beam, averaging them for the analysis. The results of the measurements are shown in Figure 4.11. The error bars were calculated with the 95% confidence intervals of the parameters of the numerical fits. The points have been joined with straight lines just to facilitate the reading of the plot, the lines do not intend to describe any behaviour or numerical fit.

Shortly after the experiment was concluded, the manufacturer of the waveplate (*Bernhard Halle Nachfolger Optische Werkstätten GmbH*) confirmed that the provided waveplates were defective, having the wrong thickness. This is the only explanation to the results from Figure 4.11. For comparison, Figure 4.12 shows the expected behaviour of the intensity of the beams for this setup.

Ideally (see Appendix B.1), if the beam comes in with 100% P-polarization, for the reference angle (0° with the waveplate's fast axis) there should only be the witness beam. After a 45° rotation the polarization should have been rotated 90°, having now full S polarization and only the driver beam. This repeats itself with a period of 90° in the rotation, which corresponds to a rotation of 180° in the polarization, or only a phase shift for linear polarization like this case. The total sum of the intensities should remain constant for the whole process.

In the measured case, however, there is no such periodicity. Taking as reference the only place where the witness beam seems to be reduced to almost zero (about 90°), there should be a position with no driver beam at ±45° from there, which is not, and the polarization seems to not be rotated at all for angles above 120°. A waveplate with an irregular surface or a wrong cut would produce this irregular pattern for the rotation of the polarization, since the dependence of the phase-shift for the slow component with the rotated angle would not necessarily be constant. Moreover, since zero-order waveplates are build by putting together two waveplates with the

slow axis of one matching the fast axis of the other (see Appendix B.1), the wrong thickness on the second plate could also completely compensate the phase-shift on the first, thus negating the rotation of the polarization vector.

Apart from that, the total intensity does not seem to be constant. The sum of the signals from the separated beams closely follows the signal of both beams together, except for when one of the signals turns negative due to background noise being too high for the analysis to give a proper value. However, this could be explained by the driver beam experiencing more absorption during its path (for example while being reflected in the beam splitters or on the mirror of its path).

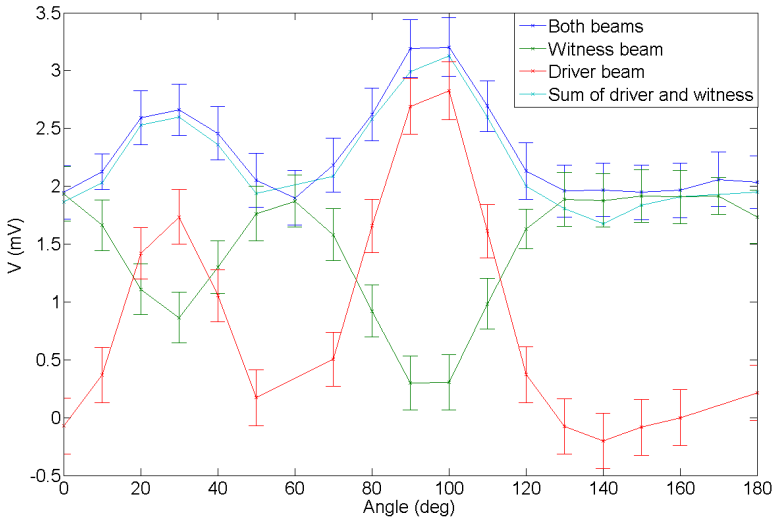

Figure 4.11.: Measured voltage from each beam as a function of the rotated angle of the waveplate.

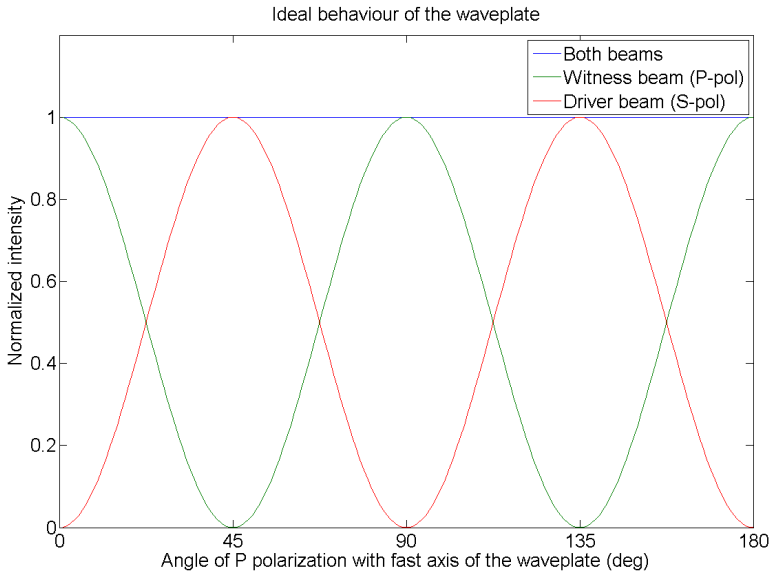

Figure 4.12.: Ideal behaviour of the intensity of the beams with the rotated angle of the waveplate.

In conclusion, the waveplate needs to be replaced for future experiments if the pulse stacker is to work as designed. However, these measurements helped choose a value for the angle in

which both beams had the same intensity, so that the generated charge was approximately the same in the proof-of-concept experiment.

4.2. Proof of concept experiment

Due to time constraints in the schedule of FLASH, only one 8-hour shift was assigned for the test of the generation and transport of the double-bunch electron beams. Since the time was very limited, the objective of the shift was to test the generation of the double bunches, their transport through the accelerator and the use of the current diagnostics equipment for the characterization of the bunches, leaving exact compression studies for further shifts down the line. Some simulated cases were prepared for the test, but in the end there was no time to set up the accelerator for them, so they were postponed until further beam time was available.

4.2.1. Generation or the double-bunch electron beam

The study showed the proper funcioning of the pulse stacker. For the proof-of-concept, the waveplate was set to a charge ratio between the bunches of approximately 1, and the delay between the bunches was set to 65 ps at the start, but it was changed during the experiment to check the behaviour of the compression with the delay.

The first step to check the functionality of the pulse stacker was to verify the proper alignment of the beam with the cathode. This was done, as mentioned in 4.1.1, with the virtual cathode. First, a picture of the Laser2 without the stacker was taken, and then both beams from the stacked are verified independently.

The next step was to verify the generation of electrons on the cathode. This was done with a so-called gun phase scan. These scans measure the charge of the bunch that is transmitted through the gun from the cathode, for different phases of the gun with respect to the global reference. Their purpose is two-fold: first, the phase can be optimized to generate the maximum charge, and second, the charge generated by each bunch can be directly measured. First, a scan was done for Laser2. Then the stacker was put in place, and a scan was done for each of the beams respectively and finally for the stacked beams together.

Figure 4.13 shows the gun phase scan for Laser2. The charge shows a maximum in which the highest number of particles are transported, with a more or less flat region of constant transported charge, and then quickly decreases. This shape is pretty similar to that of the simulated energy gain as a function of the gun phase done with ASTRA in Figure 3.8. The peak on the left-hand side slope is due to space charge effects: at that point, the field of the electrons is enough to accelerate some electrons to the point where they can be transported in spite of the RF in the gun not being able to.

Figure 4.14 shows the phase scan done for the driver beam, which is the one with a fixed beam path, and then Figure 4.15 shows the gun phase scan for both pulses together. Note that the delay between the pulses was increased to about 200 ps (95° phase difference) in order to be able to see both curves overlapping. This delay was not used further on. It is a very good sign that the curve in Figure 4.15 is essentially two typical curves from a gun phase scan overlapping. This indicates that both pulses have the same characteristics as Laser2 alone, which was also confirmed with the measurements with the streak camera.

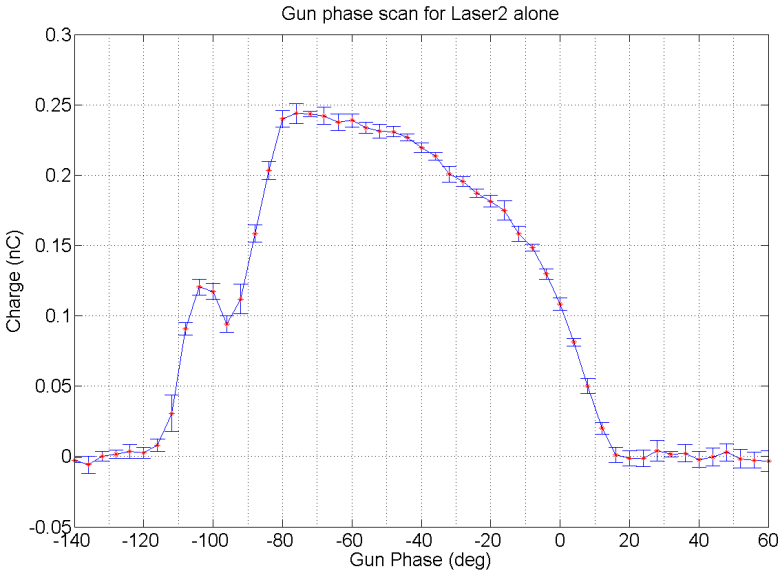

Figure 4.13.: Gun phase scan of Laser2 without pulse stacker.

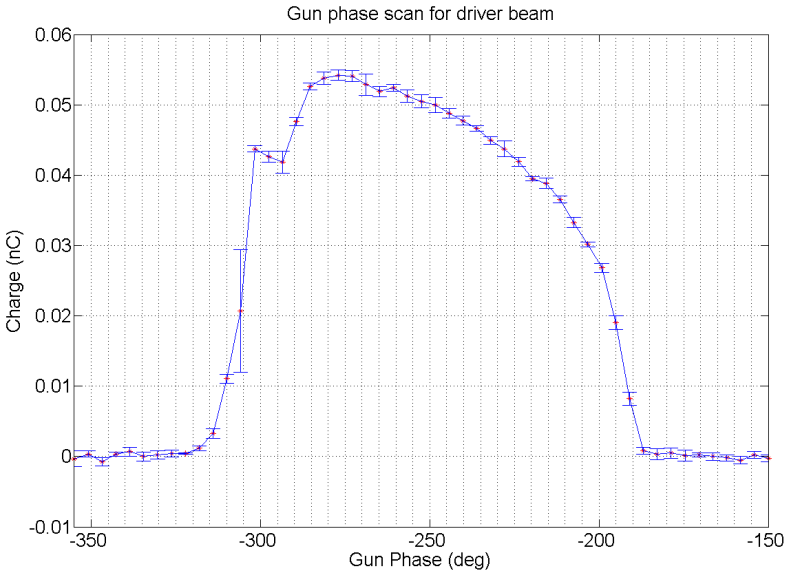

Figure 4.14.: Gun phase scan of the driver beam.

With the charge generated in the gun, the functioning of the pulse stacker is verified. The next step was to shift the laser arrival time to match that of Laser2 alone. This is needed in order to adjust the phases of the RF properly in reference to the bunches and because the measuring window of some instruments is smaller than the measured delay of (408 ± 19) ps. To this end, the gun phase scans help in determining the real difference in arrival time. This is done by measuring the difference between the zero-crossing phase (the phase for which there starts to be transmitted electrons, seen from the right side of the plots) for Laser2 and that of the driver beam. In this case, the measured delay between the standard arrival time of Laser2 and that of the driver beam was 205° or 438.5 ps, which is just within the error (one standard deviation) of the value measured with the streak camera. This difference between measurements corresponds

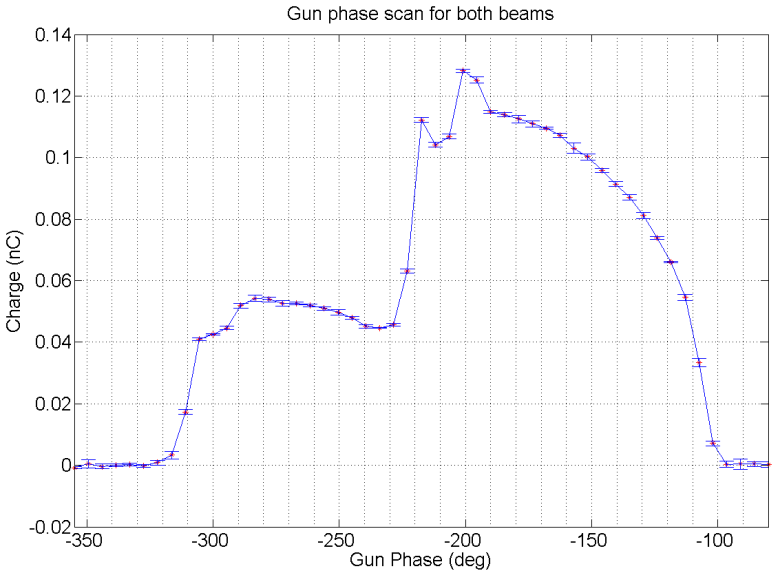

Figure 4.15.: Gun phase scan of the double-pulse beam.

to a difference in path length in the stacker of 6 mm, which could have easily been accidentally created while adjusting the laser or moving the flip-mount mirrors.

The shifting of the arrival time of the laser was done by changing the phase of the laser while using an optical cross-correlator [23] to ensure that the laser remained stable. This was done for the first time during this experiment and took a little more than an hour to be completed. The successful shifting of the laser arrival time can be checked with another gun phase scan of the driver beam. Figure 4.16 shows the gun phase scan for the driver beam after shifting the arrival time and the scan done for Laser2 alone as reference (Figure 4.13). Each curve has been normalized to their maximum charge to ease the comparison of the plots. The zero crossing of both curves was measured to be at 16°, so the arrival time was accurately shifted within less than 2 ps (assuming an error of 1° in the zero-crossing measurement). With the laser timing properly adjusted, the stacker is ready to generate bunches for transport through the accelerator.

4.2.2. Transport of double-bunch electron beams within one RF-bucket

Once the double-bunches have been generated, the next objective of the experiment was to transport them to the end of the linac and complete measurements of the longitudinal phase space.

This process takes several steps. To begin with, the on-crest phases in the accelerating modules need to be determined for the bunches. The most sensitive way to do this is by determining the minimum transition time through a magnetic chicane. This is measured with two Bunch Arrival time Monitors (BAMs) [25] placed before and after the chicane respectively. The RF phases in FLASH are only relative values and are not completely reproducible, so the on-crest phases need to be determined each time that reliable absolute phases are needed . This step took about 1 hour. After that, a new optics (the strengths of all the magnets that focus the bunch) for the accelerator was set up. The optics from the previous shift only had transmission up to the Dogleg (the energy collimator), which did not allow for longitudinal phase space measurements

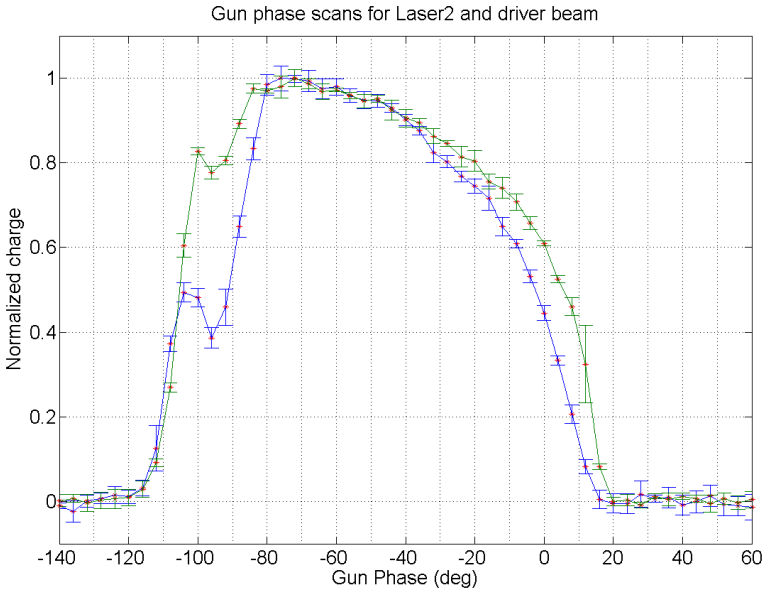

Figure 4.16.: Gun phase scan of the Laser2 (blue) as reference and for the driver beam after shifting the arrival time (green).

but allowed the BAM measurements. A new optics was chosen for the standard energy profile with transmission until the Transverse Deflecting Structure (TDS).

Finally, the TDS was the instrument used for the longitudinal phase space measurements. The TDS in FLASH [26] uses a transverse deflecting EM-field to shear the beam and translate the longitudinal distribution into a transversal one (in this case, vertical). This vertical distribution is then imaged using an optical transition radiation screens (OTR), which consists of a screen for the generation of light by transition radiation and a CCD camera to record the event.

For the pictures with the TDS, the accelerator was set toa moderate compression of the bunches, since the initial delay on the cathode was too big to fit on the screen. The camera produced raw pictures as shown in Figure 4.17. The analysis of that picture is shown in Figure 4.18, containing the calibrated phase space distribution (top left), the projected energy deviation (projection of the phase space plot on the vertical axis, top right), the longitudinal current profile (projection of the phase space plot on the horizontal axis, bottom left) and the slice energy spread (bottom right). The total charge (both bunches combined) was 720 pC and the final energy approximately 700 MeV.

The bunches in Figure 4.18 are overlapping, due to the non-optimized working point of the accelerator. Also, the energy spread is about $\pm 1\%$, since the compression was low and the bunches were transported mostly on-crest. Notice as well the long tail created by the witness bunch, which is barely compressed compared to the driver. These problems can be avoided with an optimization of the working point, as presented in the previous chapter. On the other hand, both bunches are transported succesfully and were imaged by the TDS without any trouble, which serves as proof of concept and counts as a succesful completion of the experiment.

In order to confirm that it was indeed two bunches and not one extremely deformed bunch, separate measurements of the bunches were taken by blocking one optical path. The measurements can be seen in Figures 4.19 and 4.20. The structure of the results is the same as that of Figure 4.18. They are however hard to compare, since the calibration changes the references.

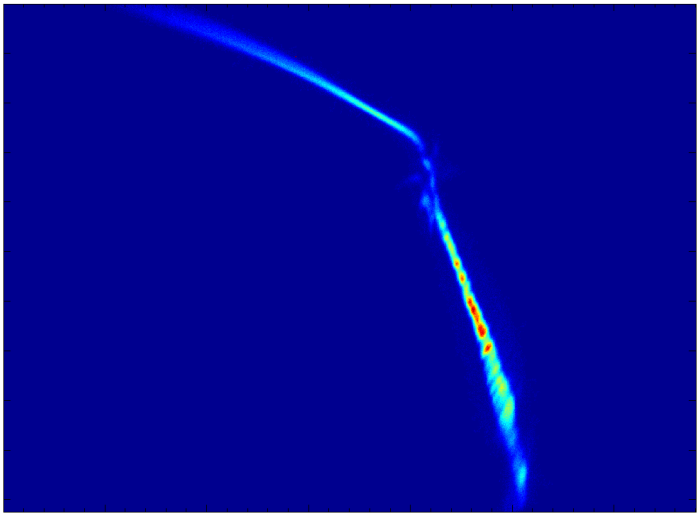

Figure 4.17.: Raw camera picture from TDS

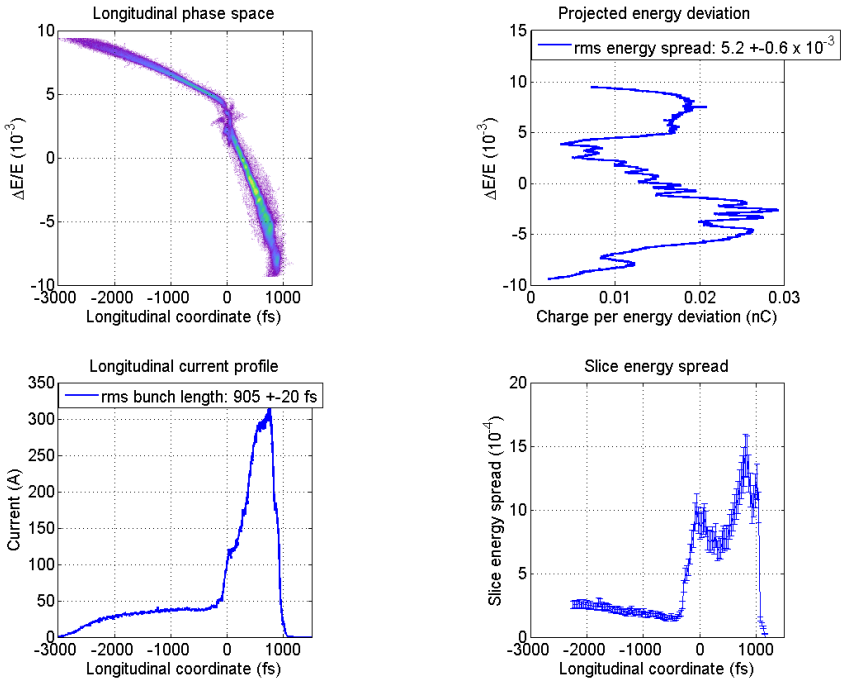

Figure 4.18.: Calibrated measurement from TDS of both bunches together.

Figure 4.21 shows the superposition of the current profiles, with the adjusted time axis, from the three measurements: both bunches together, and the driver and witness independently. It can be seen that the profile originally taken is indeed that of the two bunches. Note that the profile of each bunch closely resembles the correspondent part of the double-bunch profile, but there are some differences. The head of the witness bunch is slightly deformed in the double bunch, which can also be seen in Figure 4.18. Also, the current of the head is higher in the double bunch than in the driver beam alone, and in the tail it is lower in the double-bunch than

in the witness bunch alone. This can easily be explained by space charge effects in the gun: in the double bunch, the witness' charge would "push" the tail of the driver forward, and the driver's would push the witness' head backwards, thus creating a higher current in the driver and a smaller current in the witness. This was also discussed in Section 3.2.1.

Due to the very limited time given to the experiment, no other measurements could be taken, since the calibration of the TDS measurements takes some time and the process to set up the accelerator to the proper state took the most part of the 8 hours. However, the possibility to change the final distance and energy difference between the bunches by varying the initial delay on the cathode was also verified.

In conclusion, the measurement with the TDS showed the following:

- The pulse stacker works as designed in generating two equivalent laser pulses at the desired delay.
- The double electron bunches can be generated at the gun without complications.
- Both bunches can be transported successfully through the accelerator in the same RF bucket.
- The bunches can be successfully compressed and the generated compression is stable.
- The current diagnostics equipment in FLASH is perfectly able to measure and characterize the double bunches.

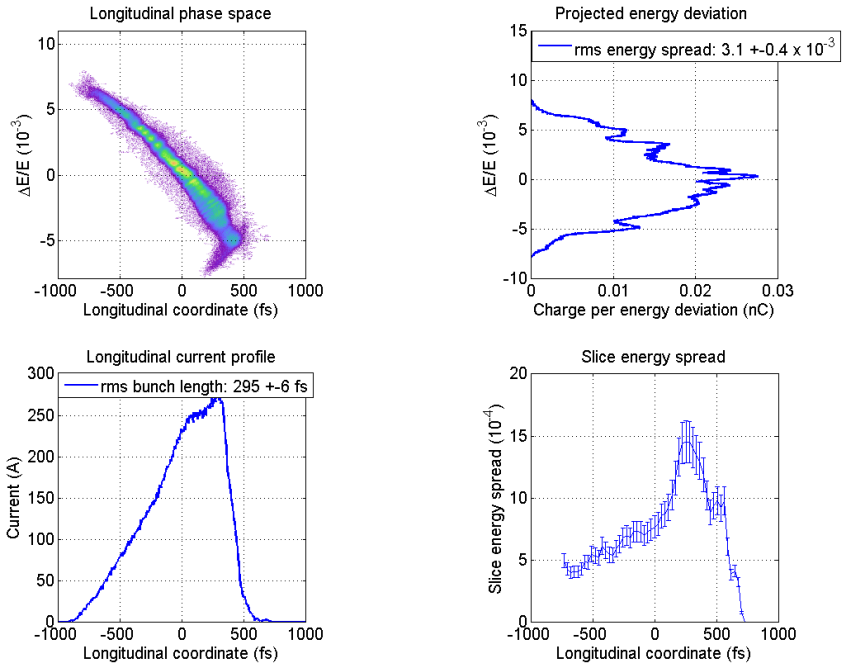

Figure 4.19.: Calibrated measurement from TDS of the driver bunch.

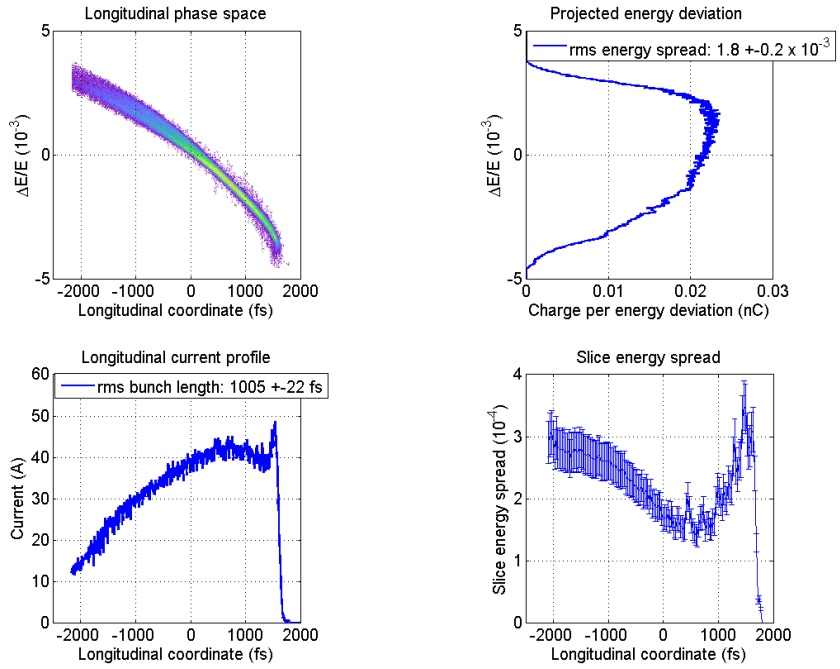

Figure 4.20.: Calibrated measurement from TDS of the witness bunch.

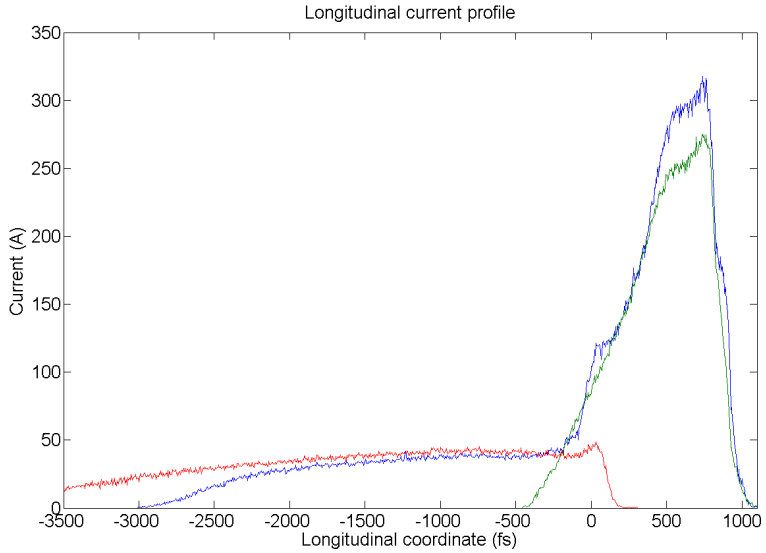

Figure 4.21.: Measured current profiles of the driver (green), the witness (red) and both bunches together (blue).

Chapter 5.

Summary and discusion

With the successful completion of the proof of concept, this work has completed most of its main objectives. From the *Task Definition* chapter, these were:

- The separation of the laser pulse into two equivalent laser pulses with a delay of some tens of picoseconds.
- The determination of a working point for the accelerator for the desired compression scenario.
- The test of the transport of double bunches in the same RF bucket and of the ability of the diagnostics equipment at FLASH to characterize these bunches.

The working point of the accelerator was determined with an iterative procedure based on numerical particle tracking using first simulations of the longitudinal dynamics and then confirming these with full 3D simulations with ASTRA and CSRtrack. The simulations showed the flexibility of the procedure to generate diverse final current profiles after the compression, all within the technical constraints of the accelerator. The system to experimentally produce the double bunches was also explained and characterized. Finally, the setup was tested in a experiment in the accelerator with very promising results.

On the other hand, the pulse stacker needs to have a new waveplate in order to work properly, so the possibility to arbitrarily select the charge of each bunch still needs to be tested. The transport of the double bunches was a good proof-of-concept, but the results from the simulations need to be properly verified with complete compression studies, which, taking into account the busy schedule of FLASH, will probably not be happening before 2015. The transversal dynamics need also to be studied for each different application, since both SASE FEL and plasma-wakefield acceleration need very specific bunch properties.

In conclusion, a method for the determination of the working point for the generation of an arbitrary compression scenario was adapted for the double-bunch case and its results were tested with particle tracking simulations. The possibility to generate double-bunch electron beams in FLASH was also demonstrated, with generation of double bunches on the photocathode and the transport of these double-bunch electron beams through the accelerator.

Acknowledgements

I would like to thank the following people, without whom this work would not have been the same:

First and foremost, thanks to Steffen Wunderlich, for his endless help and patience throughout this last year of work. His advice and insight made this work possible.

Thanks to my advisor Bernhard Schmidt, for giving me the opportunity to work in DESY and for his many pieces of advice, always very useful.

Special thanks to Igor Zagorodnov and Martin Dohlus for their help with MATLAB. Their libraries and routines saved me countless hours of work.

Thanks to the rest of the FLA group for providing their insight when it was needed and for the candid atmosphere at work.

Thanks to the Deutscher Akademischer Austausch Dienst (DAAD) and "la Caixa" for their fantastic scholarship. They granted me this wonderful opportunity and made possible for me to study in Hamburg during these two years.

Appendix A.

Basics of bunch compressors

A bunch compressor works by creating a curved beam path for the particles to go through. The particles are bent with dipole magnets and then brough back to the path of the linear accelerator. Particles with higher energy will be bent less and therefore follow a straighter, shorter path than the particles with lower energy. This path difference is the cause of the compression: if the more energetic particles are behind in the bunch, they will catch up with the head of the bunch. The following is an excerpt from [15]

The path difference for particles with relative energy deviation δ can be approximated by the following equation, which is essentially equation 3.3 as well:

$$\Delta s \approx R_{56}\delta + T_{566}\delta^2 + U_{5666}\delta^3 \tag{A.1}$$

The linear term R_{56} is usually called linear longitudinal dispersion or, in analogy with circular accelerators, momentum compaction factor, which was the term chosen here.

The two bunch compressors in FLASH are two of the simplest compressors used in linear accelerators: BC2 is a C-chicane (seen in Figure A.1) and BC3, an S-chicane.

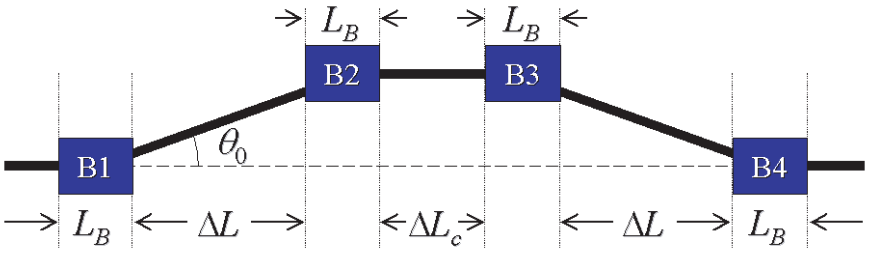

Figure A.1.: Simple C-chicane [15]

In a C-chicane, the additional path length with respect to a straight trajectory is approximated by:

$$\Delta s \approx \theta_0^2(\Delta L + \frac{2}{3}L_B) \tag{A.2}$$

Where θ_0 is the bending angle, as shown in Figure A.1. By expanding θ_0^2 in terms of small δ with $\theta^2 = \theta_0^2/(1+\delta^2) = \theta_0^2(1-2\delta+3\delta^2-4\delta^3+...)$, the compression coefficients are pretty clear. The linear term is then:

$$R_{56} = -2\theta_0^2(\frac{1}{2}(L_T - \Delta L_c) - \frac{4}{3}L_B) \tag{A.3}$$

where L_T is the total length of the chicane projected in the accelerator's line and it was assumed that the head of the bunch is at $z < 0$ (this last condition simply changes the sign of R_{56}). The other orders (zero, second and third) can be approximated from R_{56} pretty easily:

$$\Delta s \approx -\frac{R_{56}}{2},\ T_{566} \approx -\frac{3}{2}R_{56},\ U_{5666} \approx 2R_{56} \tag{A.4}$$

An S-chicane can be though of as two C-chicanes in succession where the second one bends to the other side, so as to form an S shape in projection. The parameters for the S-chicane can be approximated by the same process. The first momentum compaction factor R_{56} is, for example:

$$R_{56} \approx -2\theta_0^2(\frac{1}{2}(L_T - \Delta L_c) - 2L_B) \tag{A.5}$$

Where the parameters are the equivalents of those from the C-chicane. The second and third order parameters can also be approximated from R_{56} in the same manner. Since the bending radius and the bending angle are uniquely related, all the parameters of the compressor can be calculated from the bending radius for C and S-chicanes, which was used in the method exposed in Section 3.1.1.

Appendix B.

Optical principles of the pulse stacker

The following is a brief explanation of the working principles of waveplates and thin film polarizers. A more detailed, formal review of these can be found in [27].

B.1. Zero-order half-wave plates

A half-wave plate is an optical retarder: a device that changes the polarization of the light travelling through it. These plates are made out of a birefringent material with uniaxial anisotropy, that is, a material in which the refractive index depends on the polarization and propagation direction of the incoming light, but that has an axis of symmetry, called optic axis, along which the light propagates normally independently of the polarization. For light not travelling along the optic axis, the components of the polarization parallel and perpendicular to the optic axis have different refractive indices n_o and n_e. The component perpendicular to the optic axis will have an *ordinary* refractive index n_0, while the other component will have an *extraordinary* refractive index n_e determined by the angle of indicence.

A retarder is then built by cutting a birefringent material by a plane containing the optic axis, thus assuring that the indicent light experiences birefringence. For normally incident light, the polarization components will experience different propagation speeds within the crystal, which generates a phase shift between the components. The axis with the lowest refractive index is called the fast axis, while the other one is the slow axis. The generated phase shift $\Delta\phi$ is then:

$$\Delta\phi = \frac{2\pi\Delta n L}{\lambda} \tag{B.1}$$

Where λ is the light's wavelength, $\Delta n = |n_e - n_o|$ is the birefringence and L is the path length inside the crystal, equal to the crystal thickness for normally incident light.

In a half-wave plate, the parameters are chosen so that the phase shift is half a wavelength, $\Delta\phi = \pi$. If the angle between the incident polarization and the fast axis is θ, the angle at the exit will be $-\theta$. In fact, if the incident light wave can be decomposed in the fast component $\hat{f}$ and the slow component $\hat{s}$ as:

$$\vec{E}_{in} = E(\cos\theta\,\hat{f} + \sin\theta\hat{s})e^{i(kz-\omega t)} \tag{B.2}$$

Where k is the light's wave number, ω is the frequency, t is the time and z is the direction of propagation. Introducing a phase shift of π corresponds to a phase term $e^{i\pi} = -1$ in the $\hat{s}$ component, which results in:

$$\vec{E}_{out} = E(\cos\theta\hat{f} - \sin\theta\hat{s})e^{i(kz-\omega t)} = E(\cos(-\theta)\hat{f} + \sin(-\theta)\hat{s})e^{i(kz-\omega t)} \tag{B.3}$$

So the angle between the final polarization and the fast axis $\hat{f}$ is $-\theta$, and the polarization has been rotated an angle of 2θ.

In a real waveplate, however, the thickness of the plate cannot be so small as to have only a shift of half a wavelength and it would rather be one half plus an integer number of wavelenghts. In order to have only half a wavelength, a *zero-order* waveplate is manufactured with two waveplates stacked together, with the fast axis of the first one aligned with the slow axis of the second one and viceversa. The second waveplate is then compensating the phase shift generated by the first one, and the thickness of the plates are chosen so that they differ by a small enough amount, so that the final phase-shift is indeed one half of a wavelength.

B.2. Thin film polarizer

A thin film polarizer is a kind of polarizer used to separate the P polarization (parallel to the plane of incidence) from the S polarization (perpendicular to the plane of incidence). They are usually operated at the Brewster's angle, but can also be designed for other angles of incidence.

When light is incident at the Brewster's angle of a medium, the component of the polarization that is parallel to the plane of incidence (P polarization) is perfectly transmitted without reflection. In fact, from the Fresnel equations, the reflectance for P-polarized light in a medium-interface is:

$$R_p = \left(\frac{n_1\cos\theta_t - n_2\cos\theta_i}{n1\cos\theta_t + n_2\cos\theta_i}\right) \tag{B.4}$$

Where n_1 and n_2 are the refractive indices of the first and the second medium respectively, θ_i is the angle of incidence and θ_t is the angle of transmission. At the Brewster's angle, the angles of incidence and transmission follow the condition:

$$\theta_t = 90^\circ - \theta_i$$

And so the numerator in the reflectance is:

$$n_1\cos(90^\circ - \theta_i) - n_2\cos\theta_i = n_1\sin\theta_i - n_2\cos\theta_i \tag{B.5}$$

This can be proven to be zero using the relation in (B.5) in Snell's law of refraction:

$$n_1 \sin\theta_i = n_2 \sin\theta_t = n_2 \sin(90^\circ - \theta_i) = n_2 \cos\theta_i \tag{B.6}$$

And so, the reflectance is zero and the P component of the polarization is fully transmitted. The reflected light is only part of the S component (perpendicular to the plane of incidence) and is completely polarized, while the transmitted light has P polarization and a small proportion of S polarization.

In a thin film polarizer, a special optical coating uses intereference effects to provide a nearly perfect 50-50 split in the polarization, by neutralizing the transmitted S component, but that means that they only work in a narrow band of wavelengths around the design wavelength.

Erklärung

Hiermit bestätige ich, dass die vorliegende Arbeit von mir selbständig verfasst wurde und ich keine anderen als die angegebenen Hilfsmittel - insbesondere keine im Quellenverzeichnis nicht benannten Internet-Quellen - benutzt habe und die Arbeit von mir vorher nicht einem anderen Prüfungsverfahren eingerecht wurde. Die eingereichte schriftliche Fassung entspricht der auf dem elektronischen Speichermedium. Ich bin damit einverstanden, dass die Masterarbeit veröffentlicht wird.

Hamburg, 10. Juli 2014.

Carlos Manuel Entrena Utrilla

Bibliography

[1] J.R.Schneider, *FLASH - from accelerator test facility to the first single-pass soft x-ray free-electron laser*, J. Phys. B: At. Mol. Opt. Phys. 43 (2010) 194001 (9pp)

[2] DESY's website for FLASH, *flash.desy.de*

[3] S. Schreiber, *The injector of the VUV-FEL at DESY*, Proceedings of FEL 2005, p. 545-548

[4] S. Schreiber, B. Faatz, J. Feldhaus, K. Honkavaara, R. Treusch, M. Vogt, *Status of the FLASH facility*, Proceedings of FEL 2012, p. 37-40

[5] S. Schreiber, *First lasing at* 32 nm *of the VUV-FEL at DESY*, Proceedings FEL 2005, p. 12-18

[6] J. Bödewadt *et al.*, *sFLASH - First results of direct seeding at FLASH*, Proceedings of FEL 2010, p. 330-335

[7] C. Lechner *et al.*, *First direct seeding at 38 nm*, Proceedings of FEL 2012, p. 197-199

[8] M. Gensch *et al.*, *New infrared undulator beamline at FLASH*, Infrared Physics & Technology 51 (2008) 423-425

[9] I. Blumenfeld *et al.*, *Energy doubling of 42 GeV electrons in a metre-scale plasma wakefield accelerator*, Nature 445, 741-744 (2007)

[10] S. V. Milton *et al.*, *Exponential Gain and Saturation of a Self-Amplified Spontaneous Emission Free-Electron Laser*, Science 292, 2037 (2001)

[11] R. Prazeres, F. Glotin, C. Insa, D.A. Jaroszynski, J.M.Ortega *Two-color operation of a Free-Electron Laser and applications in the mid-infrared*, Eur. Phys. J. D 3, 87–93 (1998)

[12] P. Emma *et al.*, *First lasing and operation of an angstrom-wavelength free-electron laser*, Nature Photonics 4, 641 - 647 (2010)

[13] I. Zagorodnov, M. Dohlus, *Semianalytical modeling of multistage bunch compression with collective effects*, Phys. Rev. STAB 14, 014403 (2011)

[14] K. Flöttmann *ASTRA version 3.0 user manual* DESY

[15] M. Dohlus, T. Limberg, P. Emma, *ICFA Beam Dynamics Newsletter* 38, 15 (2005)

[16] M. Dohlus, T. Limberg, *CSRtrack: faster calculation of 3-D CSR effects*, Proceedings of FEL 2004, p. 18-21

[17] P. Castro *Beam trajectory calculations in bunch compressors of TTF2*, Technical note DESY, 2003.

[18] M.Dohlus, private communication, work not yet published.

[19] M. Dohlus, K. Flottmann, C. Henning, *Fast particle tracking with wakefields*, arXiv.org prepint 2012, arXiv:1201.5270 [physics.acc-ph]

[20] H. Schlarb, Ch. Gerth, W. Koprek, F. Loehl, E. Vogel, *Beam based measurements of RF phase and amplitude stability at FLASH*, Proceedings of DIPAC 2007, p. 307-309

[21] F. Löhl *et al.*, *Electron bunch timing with femtosecond precision in a superconducting Free-Electron Laser*, Phys. Rev. 104, 144801 (2010)

[22] M. Dohlus, T. Limberg, *CSRTRACK Version 1.2 User's Manual*, DESY, 2007

[23] S. Schreiber *et al.*, *Operation of the FLASH photoinjector laser system*, Proceedings of FEL 2011, p. 507-510

[24] S. Schreiber *et al.*, *Experience with the photoinjector laser at FLASH* Proceedings of FEL 2006, p. 590-593

[25] F. Loehl, K. Hacker, H. Schlarb, *A sub-50 femtosecond bunch arrival time monitor system for FLASH* Proceedings of DIPAC 2007, p. 262-264

[26] M. Röhrs, C. Gerth, H. Schlarb, B. Schmidt, P. Schmüser *Time-resolved electron beam phase space tomography at a soft x-ray free-electron laser* Phys. Rev. STAB 12, 050704 (2009)

[27] E. Hecht, *Optics*, 4th. Ed, (2001), Addison Wesley