

Generation of highly-energetic electrons in laser-plasma wakefields

Citation for published version (APA):

Trines, R. M. G. M. (2003). *Generation of highly-energetic electrons in laser-plasma wakefields*. [Phd Thesis 1 (Research TU/e / Graduation TU/e), Applied Physics and Science Education]. Technische Universiteit Eindhoven. <https://doi.org/10.6100/IR562748>

DOI:

[10.6100/IR562748](https://doi.org/10.6100/IR562748)

Document status and date:

Published: 01/01/2003

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Generation of Highly-Energetic Electrons in Laser-Plasma Wakefields

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van de
Rector Magnificus, prof.dr. R.A. van Santen, voor een
commissie aangewezen door het College voor
Promoties in het openbaar te verdedigen op
woensdag 26 maart 2003 om 16.00 uur

door

Raoul Milan Guido Monique Trines

geboren te Eindhoven

Dit proefschrift is goedgekeurd door de promotoren:

prof.dr.ir. T.J. Schep
en
prof.dr. F.W. Sluijter

Copromotor:
dr. L.P.J. Kamp

Druk: Universiteitsdrukkerij Technische Universiteit Eindhoven

CIP-DATA LIBRARY TECHNISCHE UNIVERSITEIT EINDHOVEN

Trines, Raoul Milan Guido Monique

Generation of Highly-Energetic Electrons in Laser-Plasma Wakefields / by Raoul Milan Guido Monique Trines. - Eindhoven : Technische Universiteit Eindhoven, 2003. -

Proefschrift.

ISBN 90-386-1655-4

NUR 925

Trefwoorden: plasmagolven / deeltjesversnellers / elektronenoptica

Subject headings: plasma waves / collective accelerators / particle beam injection / particle optics

Contents

1	Introduction	7
1.1	Particle acceleration in plasma	7
1.2	Important issues in laser wakefield excitation	11
1.2.1	Geometric diffraction of the laser pulse	11
1.2.2	Irregularities in the wakefield	13
1.3	Injection, acceleration, and extraction of the electron bunch	14
1.3.1	Injecting the bunch into the plasma wave	14
1.3.2	Acceleration and extraction of the bunch	15
1.3.3	Controlling the bunch quality	15
1.4	Wakefield excitation	16
1.5	After the acceleration	17
1.6	Contents of this thesis	18
1.6.1	Bunch quality control	18
1.6.2	Plasma wave breaking	19
1.6.3	Raman scattering and fast electron generation	20
1.6.4	Charged particle optics	21
2	Fluid description of laser-plasma interaction	25
2.1	Plasma fluid equations	25
2.2	Separation of time scales	27
2.3	Wakefield potential	30
2.4	Charged particles in the plasma	34
3	Energy spread in plasma-based acceleration	37
3.1	Introduction	37

3.2	Single electron dynamics	38
3.3	Beam loading	43
3.4	Numerical analysis	45
3.5	Simulation results	48
3.6	Accuracy requirements	54
3.7	Summary and conclusions	56
4	Fast electron generation through breaking of a plasma wave	59
4.1	Introduction	59
4.2	Wave breaking in one dimension	61
4.2.1	Wave breaking in a cold plasma	61
4.2.2	Wave breaking in a warm plasma	65
4.3	Wave breaking in two or three dimensions	66
4.4	Wave breaking in the quasi-static approximation	67
4.5	Non-linear waves in a cold plasma	74
4.6	Excitation of 1-D plasma waves	76
4.7	Fast particle generation as wave breaking	77
4.8	Simulations	79
4.9	Behaviour of the broken wave	82
4.10	Conclusions	84
5	Stimulated Raman scattering	89
5.1	Introduction	89
5.2	Dispersion relation	90
5.3	Scattering off electron modes	92
5.4	Growth rates	94
5.5	Extensions	96
5.5.1	Thermal plasma	96
5.5.2	Finite pulse length	96
5.5.3	Frequency mismatch	97
6	Enhancing fast electron production through suppression of Raman backscattering	101
6.1	Introduction	101

6.2	Effect of Raman scattering on electron yield	102
6.3	Effects of laser and plasma parameters on RBS and RFS	105
6.4	Simulations	107
6.4.1	Simulation setup	108
6.4.2	Effect of Raman backward scattering	110
6.4.3	Effect of Raman forward scattering	113
6.4.4	Plasma density profile	115
6.4.5	Laser pulse envelope	117
6.4.6	Pulse length preserving chirp	119
6.4.7	Bandwidth preserving chirp	121
6.5	Summary and conclusions	123
7	Magnetic field modelling in particle optical devices	127
7.1	Introduction	128
7.2	General field description	131
7.2.1	Basic equations	131
7.2.2	Harmonic potentials	132
7.2.3	Introducing boundary conditions	133
7.3	Calculating the field from measurements	137
7.3.1	Using measurements at the boundary	137
7.3.2	Using measurements not at the boundary	138
7.3.3	Experimental test of the presented theory	141
7.4	Application to charged particle optics	142
7.4.1	Charged particle Hamiltonian	143
7.4.2	Calculating transfer functions	144
7.4.3	Incorporating the field description into existing results	145
7.5	Conclusions	146
A	The code XOOPIC	151
A.1	Description of the code	151
A.1.1	General outline	152
A.1.2	Fields	153
A.1.3	Particles	154

A.1.4	Boundary conditions	155
A.2	Benchmarks	156
A.2.1	Dispersion relations	156
A.2.2	RFS seeding and beat-wave excitation	157
A.2.3	Growth rates	158
Samenvatting		161
Curriculum Vitæ		165
Dankwoord		167

Chapter 1

Introduction

1.1 Particle acceleration in plasma

Ever since the discovery of X-rays by W.C. Röntgen in 1895 and the discovery of the electron by J.J. Thompson in 1897, physicists have been striving to accelerate charged particles to the highest energies possible. Along the road to high-energy physics, one finds such milestones as the Van de Graaff generator (1929), the first cyclotron by E.O. Lawrence in 1931, the first betatron by D. Kerst in 1940, and the first linear accelerator (linac) by W. Hansen in 1947. At present, accelerator development has found its culmination in such superstructures as the Centre Européenne des Recherches Nucléaires (CERN) in Geneva, the Deutsches Elektronen-Synchrotron (DESY) in Hamburg, and the Stanford Linear Accelerator Center (SLAC) in Menlo Park, California.

The most defining characteristic of the accelerators operating at the forefront of high-energy physics is the sheer size of the structures involved. The SLAC linac measures 3.8 kms, the largest storage ring of DESY measures about 7 kms, and the largest ring of CERN measures no less than 27 kms and stretches over two countries. Building accelerators this size costs billions of euros, and not many governments are willing to spend such amounts on fundamental research in high-energy physics. A clear example of this is the termination of the Superconducting Super Collider Project (SSC) by the US government in 1993.

Conventional accelerators need to be so large, because they accelerate particles using electromagnetic (EM) fields in vacuum. The largest electric field strength that can be sustained in vacuum before breakdown occurs, is 10 MV/m continuously or up to 100 MV/m pulsed. So in order to reach ultrahigh particle energies in a conventional accelerator, one needs to increase the acceleration length, i.e. the size of the apparatus. This does not only set a limit on the highest energy that can be obtained ever, it also prevents an average university or research center from building and operating an accelerator that reaches beyond say 1 GeV

because of sheer cost. Therefore, many results in highest-energy physics simply cannot be reproduced because there are only a few accelerating structures in the world operating in the necessary energy range. This situation is of course less than satisfactory.

A remedy has been sought in replacing vacuum by plasma as a medium to accelerate particles in [1–3]. Electric fields in a plasma can reach values up to 100 GV/m, i.e. three to four orders of magnitude beyond what can be realized in vacuum [4–6]. The basic idea behind plasma-based acceleration is that a longitudinal EM wave is excited in the plasma, which is subsequently used to accelerate electrons just as a radio-frequency (RF) wave accelerates particles in a linac. Since the excitation of such a wave implies the occurrence of charge separation in a plasma, this idea was hardly deemed feasible early on, but it was found that a small but intense source propagating through the plasma could indeed expel a significant fraction of the plasma electrons from its path. If this source moves fast enough, the much heavier ions will remain more or less immobile, so directly behind the pulse there will be a spot having a positive total charge. The strong restoring Coulomb forces push the electrons back quickly, so quickly that behind the positive spot there appears a spot with too many electrons, and thus a net negative charge. This process is repeated a number of times, so an oscillation of the plasma electrons is induced that follows the moving source. In other words, a plasma wave, also called the *wakefield*, is realized. The frequency of this wave is the *electron plasma frequency* ω_p , which is the resonant frequency for unforced plasma electron oscillations. The plasma frequency is given by $\omega_p = \sqrt{e^2 n_0 / (\epsilon_0 m_e)}$, where e denotes the unit charge, n_0 the background electron density, and m_e the electron rest mass. In this thesis, the plasma electron density ranges between 10^{17} and 10^{19} cm $^{-3}$, corresponding to a plasma frequency from $3 \cdot 10^{12}$ rad/s to $6 \cdot 10^{13}$ rad/s. As the phase velocity of the plasma wave is often close to the speed of light, the *plasma wavelength* λ_p is defined as $\lambda_p = 2\pi c / \omega_p$.

Various types of sources can be used to excite a wakefield in a plasma. For example, an electron bunch at relativistic speed can be used. In this scheme, two electron bunches are injected into a plasma. The length of each bunch is shorter than λ_p . The transverse Coulomb forces of the first drive the wakefield that subsequently accelerates the second. Alternatively, one can use a single bunch that is longer than λ_p . In that case, the head of the bunch creates the wakefield that accelerates the tail. This method of plasma-based acceleration goes by the somewhat confusing name of *plasma wakefield acceleration* (PWFA) [7–9]. The PWFA scheme has been popular since the early days of plasma-based acceleration, because electron bunches of ultrahigh energy were readily available at that time. PWFA is still studied today, most notably in the E-157 and E-162 experiments at SLAC [10,11]. However, a major disadvantage of PWFA is that when the leading bunch loses energy in driving the wakefield, its velocity decreases. This means

that the velocity of the wakefield decreases as well. As a result, the maximum energy of the trailing bunch cannot exceed twice the value of the energy of the leading bunch.

To overcome this limitation, Tajima and Dawson proposed to use a laser pulse of high intensity to drive the plasma wave, since laser pulses do not slow down when losing energy to the plasma [12]. A laser pulse can propagate through a plasma provided that its carrier frequency ω_0 is larger than the plasma frequency ω_p . If the plasma has $\omega_p < \omega_0$ for a given laser pulse, the plasma is called *underdense* to that pulse, otherwise it is called *overdense*. The pulse excites a plasma wave in the following way. When a laser pulse propagates through a plasma, the leading order response of the plasma electrons is a rapid quivering in the fast oscillating EM fields of the pulse. However, this response averages to zero in the long run. After averaging, a net force on the electrons remains, called the *ponderomotive force* \mathbf{F}_p . For a laser pulse with electric field envelope E_0 , the ponderomotive force is given by

$$\mathbf{F}_p = -\frac{e^2}{2\gamma m_e \omega_0^2} \nabla(E_0^2),$$

where γm_e is the relativistic electron mass. Since the pulse's intensity, and thus E_0^2 , is usually highest at the centre of the pulse and decreases towards the sides, the direction of the ponderomotive force is outwards. It thus pushes the plasma electrons out of the path of the laser pulse, and excites a plasma wave. Note that the laser pulse should be shorter than a plasma wavelength for this scheme to work properly, since otherwise \mathbf{F}_p is too small while the plasma has too much time to adapt itself to the presence of the laser pulse. The parameter regime in which this condition is satisfied, i.e. excitation of the wakefield is dominated by the ponderomotive force, is called the *ponderomotive regime*.

As will be explained in Section 1.4, pulses longer than a plasma wavelength may still excite a wakefield. In that case however, the wakefield is the result of an instability called *Raman forward scattering*. This instability drives a fast plasma wave that happens to have approximately the same frequency, wavelength and phase velocity as a wakefield excited by the ponderomotive force, even though it originates from entirely different processes. In this regime, the back-action of the plasma on the pulse is very important, and one can say that the laser pulse modulates itself while propagating through the plasma. Therefore, this regime is called the *self-modulated regime*.

If a laser pulse loses energy to the plasma, its amplitude decreases rather than its group velocity. This also decreases the amplitude of the wakefield, which may be problematic. However, the efficiency of the energy transfer is fairly low for a strongly underdense plasma, so this does not impose a severe limitation contrary to the case of PWFA. The disadvantage of this is, though, that a very intense laser pulse is needed to drive a decent plasma wave. In the early 1980's, sufficiently strong lasers were not yet available. In order to drive a plasma wave of reasonable

amplitude using laser pulses, the *beat wave acceleration* (BWA) scheme has been invented [13,14]. In this scheme, two long, overlapping laser pulses are used having frequencies that differ by the plasma frequency ω_p . Once these are injected into the plasma, a plasma oscillation is driven resonantly. This way, a large amplitude plasma wave can be excited even with pulses of moderate intensity. Research into BWA is carried out to this day, for example at UCLA [15,16].

The situation changed with the invention of *chirped pulse amplification* (CPA) in 1985 [17]. The CPA scheme is based on the fact that a short laser pulse has a broad colour spectrum. In this scheme, a short, low-intensity laser pulse is sent through a dispersive medium, causing the various colours in its spectrum to be separated in time, and its duration to increase. Conservation of energy dictates that its amplitude decreases. The stretched pulse is then sent through an amplifier, which can handle longer pulses much better than short ones. It is then amplified by a factor up to 10^7 . After amplification, the pulse is recompressed using a double grating, all its colours are rejoined and its peak power rises skyhigh. This way, a peak power of $10^{19} - 10^{20}$ W/cm² can be reached for a pulse length of only 50-100 fs. Laser systems capable of delivering such pulses are commonly denoted as Table Top Terawatt or T³ lasers. A single pulse from such a laser is sufficiently intense to drive a plasma wave with a decent amplitude, without having to resort to beat-wave schemes. Particle acceleration in a plasma wave driven by a single laser pulse is called *laser wakefield acceleration* (LWFA) [1].

In this thesis, we will only concern ourselves with laser wakefield acceleration, and not with PWFA. This will be obvious in Chapters 5 and 6, where we deal with instabilities typical for laser-plasma interaction. However, the sources driving the wakefield in Chapters 2 to 4 will also taken to be laser pulses rather than bunches of particles, so we can safely assume that the wakefield has a constant phase velocity.

Once a plasma wave has been excited some way or another, the next step is to inject a bunch of electrons into the wave and have the wave accelerate that bunch. The dynamics underlying electron acceleration in a plasma wave are not unlike those of a surf board that is accelerated by a big sea wave. Initially, the bunch of electrons is injected into the plasma wave at a speed below the phase velocity of the wave. Then they slip backwards against the wave, while the wave exerts a forward force on the particles. When the initial electron velocity has been sufficiently high, the bunch will acquire a velocity larger than that of the plasma wave, and start to slip forwards with respect to the wave. As the plasma wave has a velocity already close to the speed of light, the velocity of the bunch cannot be much larger than that of the wave, and it may take a while before the bunch outruns the wave. Until that happens, the wave exerts a strong accelerating force on the bunch, so the electrons may reach very high final energies.

Electrons can be injected into the plasma in basically two ways: by external and

by internal injection. In external injection, the bunch is produced outside the plasma, e.g. by means of a photocathode, and subsequently injected into the plasma. This has the advantage that the mean energy is well-defined and both the energy spread and the transverse emittance are small. Unfortunately, it is hard to control the transverse position of the bunch and its phase with respect to the plasma wave, whereas the bunch tends to be too long with respect to the plasma wavelength. In internal injection, the electrons are taken from the plasma itself, either by wave breaking of the wakefield [4,18–20], or by colliding several laser pulses inside the plasma [21,22]. At the moment of collision in the latter scheme, a fraction of the plasma electrons may receive a sufficient kick to inject them into the wakefield excited by the biggest of the colliding pulses. With internal injection, the transverse position and the phase with respect to the plasma wave are easy to control, and bunches are often much shorter than a plasma wavelength. A disadvantage of internal injection is, that often the longitudinal energy spread is very large, up to 80% of the mean energy of the bunch.

Due to the highly unstable nature both of the plasma and of the laser-plasma interaction, there are many issues to be dealt with in LWFA, the most important of which are discussed in the next few sections.

1.2 Important issues in laser wakefield excitation

As mentioned above, the immediate advantage of LWFA is that the maximum attainable accelerating field in a plasma can be three to four orders of magnitude larger than what can be obtained in vacuum. This means that the maximum energy gain in a plasma-based accelerator is several orders of magnitude larger than what can be obtained using a conventional accelerator of the same size. But the price to pay is that plasma-based acceleration is a complicated affair for a number of reasons, which are discussed in this section. For more information, see also Kruer [23], or Liu and Tripathi [24]

1.2.1 Geometric diffraction of the laser pulse

In order to generate a wakefield with sufficient amplitude, the laser pulse needs to be focused down to a small spot to increase its peak intensity. Unfortunately, this also leads to large geometrical diffraction of the pulse. In other words, the pulse widens considerably over only a small distance behind its focus, and its intensity drops accordingly. For this reason, the excitation of a wave in a homogeneous plasma is effectively limited by the *Rayleigh length* $z_R = 2\pi r_0^2/\lambda_0$, where r_0 is

the focal spot radius and λ_0 the laser wavelength. For example, in the case of a laser pulse having $\lambda_0 = 800$ nm which is focused down to $r_0 = 10$ μm , we have $z_R = 0.78$ mm. On the other hand, the acceleration length needed for a significant energy gain is usually several centimeters or even tens of centimeters, i.e. several tens or hundreds of Rayleigh lengths. So in order to sustain a decent laser pulse intensity over the total acceleration length, the pulse has to be guided in much the same way as light is guided in fiber optics. Several methods to do this will be listed here.

Relativistic self-focusing When a very intense laser pulse interacts with a plasma, the transverse quiver velocity of the plasma electrons approaches the speed of light, and the electron mass increases locally, i.e. where the pulse is. Now the plasma frequency decreases with increasing electron mass, so a local depression in the plasma frequency is obtained. It can be shown that the optical density of the plasma increases with decreasing plasma frequency and thus increases with increasing pulse intensity. Therefore, the optical density will be higher at the center of the pulse than at its sides. Therefore, the path of the pulse starts to behave like a glass fiber, nicely guiding the laser pulse and inhibiting diffraction. See, among others, Refs. [24,25] for details.

Preformed plasma channel Since the plasma frequency decreases with decreasing plasma electron density, the optical density increases with decreasing electron density. Therefore, a channel of lower density in an otherwise homogeneous plasma also acts as a fiber, and can be used to guide a laser pulse [26,27]. Such a preformed channel can for example be generated by another laser pulse.

Capillary discharge wave guide This type of wave guide consists of a gas-filled capillary with metal walls. The gas is ionized by means of an electrical discharge. If tuned properly, the plasma density is lowest at the central axis and increases towards the walls, i.e. a profile similar to that of a preformed channel. The pulse guiding is therefore based on the same mechanisms [28–30].

Glass capillary tube The capillary is filled with gas at low pressure. When a laser pulse passes through, its head is supposed to ionize the gas and create a suitable density profile to guide the tail [31–35].

Of these methods, the preformed plasma channel and the discharge waveguides look the most promising for laser-plasma acceleration, since these are stable and reproducible processes. Relativistic self-focusing is an inherently unstable process with unpredictable outcome, whereas glass capillaries usually only last a single pulse.

1.2.2 Irregularities in the wakefield

For controlled acceleration of an electron bunch, it is important that the excited plasma wave provides a stable environment. This means that it should not contain (too many) irregularities. Two important causes of irregularities that can still occur when proper pulse guiding is employed are *wave breaking* and *parametric instabilities*.

Wave breaking is a process during which the local electron velocity becomes larger than the local phase velocity of the wave. In a regular wave, the plasma electrons are slower than the wave, and only oscillate back and forth when the wave passes. For intense plasma waves however, the plasma electrons may overtake the wave itself during the oscillation. When this happens, a fraction of the plasma electrons gets trapped in a crest of the wave, and is dragged along over many plasma wavelengths. The trapping of electrons by a plasma wave may disturb the structure of the plasma wave and therefore hurt the acceleration process. Wave breaking will be treated extensively in Chapter 4.

Parametric instabilities are caused by the fact that if a laser pulse with characteristic frequency ω_0 interacts with a plasma with characteristic frequency $\omega_p < \omega_0$, electromagnetic radiation at the sum and difference frequencies is generated. This process needs small disturbances in the plasma density to start from, and it is said that the laser light *scatters* from such disturbances. As it happens, the original laser light and the scattered radiation drive the density disturbances between them, and an instability is realized: the larger the density perturbations the more scattering, and the more scattering the faster the perturbations grow. An in-depth treatment of parametric instabilities can be found in Chapter 5.

Regularity of the plasma wave is especially important during the acceleration stage, i.e. when the electron bunch already resides in the wakefield and is being accelerated to high energies. At this stage, the occurrence of irregularities needs to be avoided. Fortunately, both wave breaking and parametric instabilities can be prevented by taking the plasma density sufficiently low, such that the laser frequency is much larger than the plasma frequency and the laser pulse is shorter than one plasma wavelength. A plasma density of about 10^{17} cm^{-3} is commonly employed for the acceleration stage. This has the added benefit that the group velocity of the laser pulse, and thus the phase velocity of its wakefield, approaches the speed of light c for decreasing density, leading to higher electron energies after acceleration.

However, things are different in case of internal electron injection, when the interaction of the laser pulse with the plasma needs to result in the capture and subsequent acceleration of plasma electrons. In case a single pulse is used, the capture of plasma electrons is realized by pushing the pulse's wakefield until wave breaking is reached. As wave breaking becomes easier for decreasing wave velocity, and the wave velocity decreases with increasing plasma density,

the plasma density will be much higher here, in the range of $10^{19} - 10^{20} \text{ cm}^{-3}$. Since the plasma wavelength for such densities is about $5 \text{ }\mu\text{m}$, while the laser pulse length can hardly be below $15 \text{ }\mu\text{m}$, these processes clearly take place in the self-modulated regime. Since wakefield excitation occurs through parametric instabilities in this regime, these are to be stimulated rather than avoided. Also, as we will see later, the amount of electrons obtained in internal injection can be optimized through stimulation and/or suppression of specific instabilities, while the influence of laser and plasma parameters on the injection process mainly acts through these instabilities.

1.3 Injection, acceleration, and extraction of the electron bunch

1.3.1 Injecting the bunch into the plasma wave

In a conventional accelerator, the wavelength of the RF wave is several cm, while in a plasma accelerator, this wavelength is of the order of 0.1 mm . Since in acceleration physics, both the dimensions and the position of an electron bunch should be viewed relative to the dimensions of the wavecrest in which the bunch is to be trapped and accelerated, tuning these bunch parameters is much harder in plasma-based than in conventional acceleration. This poses serious requirements on timing and transverse position of the bunch, as well as on its longitudinal and transverse dimensions.

As we will see in Chapter 3, the mean energy and relative energy spread of a bunch of electrons after acceleration is very sensitive to changes in the injection phase. Also, it is important that the bunch length does not exceed a single plasma wavelength. Similar requirements apply to the transverse dimensions and positioning of the bunch. The requirements on the tilt, i.e. the angle between the directions of laser pulse and injected electron bunch, are not very lenient either.

There are several ways to get around the problem of injection accuracy. For example, the electron bunch can *precede* the laser pulse instead of following it. Since the bunch has only a moderate energy on injection, the laser pulse is able to overtake the bunch, which is subsequently scooped up by the pulse's wakefield. As shown by A.G. Khachatryan [36], the bunch is greatly compressed in the process, and the sensitivity to errors in the injection phase is reduced considerably.

Another option is to accelerate electrons taken from the plasma itself instead of injecting electrons from the outside. This has the advantage that both the injection phase and the bunch length are as desired by default, and controlling these is no longer an issue. Breaking of the wakefield may cause a considerable fraction of the plasma electrons to get trapped and accelerated, as mentioned be-

fore. Wave breaking and plasma electron trapping in the ponderomotive regime is investigated in Chapter 4, while Chapter 6 deals with particle trapping in the self-modulated regime. A disadvantage to wave breaking is, that it is not a very controlled process, so the electrons thus accelerated exhibit a rather large energy spread. To remedy this, several injection schemes involving multiple laser pulses have been proposed (cf. Section 1.1). The basic idea is that the wakefield is excited by a strong pump pulse, and that a small probe pulse having a slightly different frequency is arranged to collide with the pump, either head-on or sideways. At the moment of collision, the laser pulses beat together to drive a slow plasma oscillation, which is just sufficient to inject a small amount of plasma electrons into the wakefield of the pump. For such colliding pulse schemes, the resulting energy spread of the injected electrons is much lower than in a wave breaking scheme.

1.3.2 Acceleration and extraction of the bunch

In conventional accelerators, the RF wave can be reasonably controlled, ensuring “clean” acceleration. In addition, the waveguide for the RF wave is designed in such a way that each time the bunch is (nearly) out of phase with the wave, the wave undergoes a phase jump such that its phase again matches that of the bunch. In plasma based acceleration, wave propagation is much less controlled, and introducing phase jumps is not possible. Therefore, acceleration is limited to a single so-called dephasing period, after which it is decelerated. This requires proper extraction of the bunch at approximately the dephasing length. Extraction at exactly the dephasing length is not an easy task, since the dephasing length scales with the plasma density n_e as $n_e^{-3/2}$, and n_e is usually not known very accurately.

Another problem is that for a moderately to strongly underdense plasma, the dephasing length may be several tens of centimetres, so guiding of the pulse over a single dephasing length may not be possible at all. As a result, the acceleration length is often determined from entirely different considerations than whether dephasing is reached or not.

1.3.3 Controlling the bunch quality

Apart from the issues that arise while providing a proper environment for the electrons to be accelerated in, there is also the issue of *bunch quality*. Bunch quality is usually defined as its *brightness*: the total bunch charge divided by the volume it “occupies” in six-dimensional phase space (three spatial and three momentum dimensions). In other words, a bunch has good quality if it contains much charge compressed in a small spatial volume with little spread in both

longitudinal and transverse momenta. In many cases, it is not only required that the accelerated bunch has a high mean energy, but also a good quality. Controlling bunch quality in plasma-based acceleration is again much harder than in conventional acceleration, since in the former case the dimensions of the bunch are much larger with respect to the wavelength of the accelerating wave. Therefore, different bunch electrons may follow much different paths through phase space during acceleration, increasing the total volume occupied by the bunch and deteriorating bunch quality.

1.4 Wakefield excitation

As explained in Section 1.1, a laser pulse can excite a wakefield in a plasma by means of the ponderomotive force it exerts on the plasma electrons. However, ponderomotive excitation does not work in all circumstances, and other mechanisms contribute to wakefield excitation as well. This depends on whether the laser pulse is shorter or longer than a single plasma wavelength.

The regime for which the laser pulse length is shorter than the plasma wave length is called the *ponderomotive regime*. In this regime, there is little or no coupling between the fast laser oscillations at frequency ω_0 and the slow plasma response at frequency ω_p , and the plasma wave is mostly driven by the ponderomotive force of the laser pulse envelope. Also, since the energy transfer from the laser pulse to the plasma is not very efficient in this regime, one can often neglect any changes to the laser electromagnetic fields while it traverses the plasma. Numerical calculation of the plasma wave driven by the ponderomotive force of a laser pulse shows that the wave is most efficiently driven when the pulse length is between one quarter and one half of the plasma wave length. For longer pulses, the efficiency of ponderomotive excitation quickly drops to zero.

Behaviour is different if the laser pulse length is several times the plasma wave length. This regime is called the *self-modulating regime*. In this regime, the effect of the ponderomotive force on the plasma is small to negligible. On the other hand, there is a strong coupling between the frequencies ω_0 and ω_p . This coupling leads to the excitation of parametric instabilities as introduced in Section 1.2.2 and treated in more detail in Chapter 5, such as stimulated Raman scattering (SRS), stimulated Compton scattering (SCS), and stimulated Brillouin scattering (SBS). Of these, SRS will be the most important when studying the interaction between a laser pulse and a cold, underdense plasma. Through the strong perturbation of the plasma density, SRS also leads to strong modulation of the laser pulse envelope, hence the name of the regime.

Even though there is hardly any ponderomotive excitation of a plasma wave in the self-modulated regime, sufficiently intense laser pulses are observed to excite plasma wake waves here. This can be attributed to the presence of the Raman

forward scattering (RFS) instability. The plasma wave associated with RFS happens to have frequency ω_p and wave number approximately ω_p/c . These are similar to those of the wake wave excited ponderomotively by the laser pulse. Therefore RFS and the wake wave enhance each other, and the latter reaches an amplitude far beyond what could be obtained by ponderomotive excitation alone. Under the right circumstances, the wake wave can even serve to trap and accelerate plasma electrons in this regime as well. The way in which Raman instabilities affect the excitation of the laser wakefield and the trapping of plasma electrons will be investigated in depth in Chapter 6.

1.5 After the acceleration

So far, we have concentrated on issues that relate directly to the behaviour of the electron bunch when it is still inside the accelerator. However, what happens to the bunch after it has left the accelerator is no less important. This final stage consists of guiding the accelerated bunches to their final destination, by means of a beam guiding system that takes the bunch from the accelerator to an experimental setup, or a storage ring, where high-energy particles are stored for many hours.

The guiding of charged particles is often done using purely magnetic fields. These are preferred over electric or electromagnetic fields, since the guiding can be best controlled if static fields are used, and strong magnetostatic fields are easier to produce than strong electrostatic fields. The magnetic fields will not exert any work on the particles during this stage, and as such the bunch guiding stage is often regarded separately from the acceleration stage. However, the bunch quality issues that come with this stage are the very same as those important for the acceleration stage. Therefore, this last stage needs to be explored with equal care: a poorly constructed beam guiding system can easily destroy the bunch quality that the accelerator designers have taken so much pains to achieve. Bunch quality control definitely does not stop the moment the bunch leaves the accelerator.

Accelerated particle beams are usually guided using magnetic multipoles: a configuration of alternating magnetic north and south poles, grouped evenly around the beam pipe, pointing inward. A magnetic dipole is used for changing the direction of propagation, a quadrupole (or usually a combination of quadrupoles of opposite strength) for focusing the beam. Since the desired multipole symmetry is only obtained in the middle of the multipole, and not at its entry and exit faces, small sextupoles or octupoles are often used there to correct lower order aberrations from this symmetry.

In order to tune the multipoles along the path of the beam for optimal performance, it is vital to have knowledge not only of the central field, but also of the

behaviour of the fringe fields. Especially for the dipoles and quadrupoles, so the corrective sextupoles and octupoles can be tuned. This is often done by developing the magnetic fields near the fringes into powers of the transverse radius r , and determining the coefficients. A major disadvantage of this is, that the number of terms grows so explosively, that only lower order terms can be taken into account. Higher order terms, which are usually insignificant if a beam guiding element is passed once, but can become very important if it is passed many times, as happens in a storage ring, are often neglected because taking these into account would complicate matters too much.

As it happens, the fields of magnetic or electric multipoles display a remarkable amount of symmetry, which can be exploited to give a full yet compact description of these fields without having to resort to a power series expansion at all. Such a description will be presented in Chapter 7.

1.6 Contents of this thesis

In this thesis, the following four topics are investigated. In Chapter 3 a study on bunch quality control during acceleration is presented; in Chapter 4 wave breaking and its role in plasma electron trapping are investigated; in Chapters 5 and 6 stimulated Raman scattering and its role in plasma electron trapping are scrutinized, and in Chapter 7 the description of magnetic multipole fields in charged particle optics is given a thorough treatment. Each topic will be discussed briefly.

1.6.1 Bunch quality control

There are many aspects of controlled acceleration, i.e. acceleration while retaining bunch quality, that are worth investigating: minimizing the energy spread, transverse dimensions, or aperture angle, or maximizing the energy and the total charge. Of all these aspects, we decided to concentrate on the longitudinal energy gain and spread as a function of acceleration length, injection and extraction phase, bunch length and plasma wavelength (plasma density). To this end, simulations have been performed in which a bunch of charged particles is injected at low energy into a prescribed linear plasma wave with phase velocity $v_\varphi \lesssim c$ and associated Lorentz factor $\gamma_\varphi = 1/\sqrt{1-v_\varphi^2} = 100$. The simulations were continued until the particles were completely out of phase with the wave and the mean particle energy started to decrease. The energy spread of the bunch was continually monitored during acceleration. We aimed at minimizing the energy spread, and tried to pinpoint which conditions would give a decent energy gain in that case.

As it turned out, there is more to controlled acceleration than simply aiming for a large energy gain: a maximal energy gain is often only obtained at the expense of a very large energy spread, while sacrificing a small fraction of the energy gain may bring us into a parameter regime where the energy spread is fairly small. This shows that concentrating on energy gain alone, as was done in the early days of plasma-based accelerator research, is too one-sided. Nevertheless, a parameter window has been identified for which a low energy spread can be combined with a high energy gain. Unfortunately, both energy gain and spread are very sensitive to small errors in the injection time of the bunch versus that of the laser pulse. This is a motivation to investigate alternative bunch injection schemes in which the bunch is injected at the correct phase by default, such as internal injection by means of wave breaking.

1.6.2 Plasma wave breaking

Plasma wave breaking has been investigated as a mechanism for generating fast electrons for our bunch in the ponderomotive regime. A 1-D analytical model for plasma oscillations behind a driving source has been developed. The occurrence of singularities in the model serves to determine the conditions under which a 1-D plasma wave will start breaking. The model has been extended by including a “sheet” of trapped electrons preceding the point of wave breaking at a very short distance. The wakefield of this sheet will decrease the amplitude of the laser wakefield (beam loading), and at a sufficiently high charge density of the sheet, this beam loading will keep the wakefield from any further breaking. Therefore plasma electron trapping through wave breaking is a self-limiting process. The model also allows to calculate the amount of charge that will be trapped at wave breaking.

The extended model has been used to express the intensity threshold for trapping and the amount of trapped electrons as a function of intensity and group velocity of the pulse. Numerical calculations have been performed to support the analytical results, and also to investigate the influence of the pulse length and envelope shape on fast electron production. As the amount of fast electrons can be derived from the wakefield amplitude, which is an easier quantity to evaluate numerically, this amplitude has been calculated and used as a measure of the amount of charge that can be extracted. It has been found that $L \sim \lambda_p/2$ is the optimal pulse length for wakefield excitation, and that ponderomotive excitation hardly plays a role for $L \geq \lambda_p$. It has also been found that pulse envelopes having long “tails” do not excite plasma waves efficiently, while envelopes with steep flanks, especially a steep front, are much better suited for this purpose. In fact, a steep pulse front, i.e. the pulse intensity increases from 0 to its peak value over less than a plasma wavelength, is as essential to driving an intense plasma wave as the peak pulse amplitude.

1.6.3 Raman scattering and fast electron generation

In the self-modulated regime, i.e. when the laser pulse is longer than one plasma wavelength, Raman scattering becomes very important in laser-plasma interaction, while ponderomotive excitation fades into the background. Wakefield excitation, wave breaking, and fast electron generation still occur, but the mechanisms behind them are fundamentally different from those in the ponderomotive regime. Research in this regime needs therefore to be approached from a different angle.

Wave breaking and fast electron production in the self-modulated regime have been investigated using the particle-in-cell (PIC) code XOOPIC [37]. This fully relativistic code solves the full set of Maxwell equations in one spatial and three velocity dimensions, and calculates the particle motion using the Lorentz force. External electromagnetic fields such as a laser pulse are introduced through boundary conditions imposed at the boundaries of the simulation box. Using this code, the effect of various laser and plasma parameters on fast electron production has been investigated. Also, the role of Raman forward and backward scatter has been explored.

In line with earlier results, it has been found that Raman forward scatter drives the laser wakefield, and that Raman backward scatter “heats” the plasma, causing many slow plasma electrons to be injected into the wakefield. At laser intensities slightly above the electron trapping threshold, such injection increases the amount of high-energy electrons produced in laser-plasma interaction. However, at intensities well beyond this threshold, RBS and the amount (charge) of injected electrons increase to an extent, that the wakefield is heavily damped, and fast electron production suffers. Therefore, suppression of RBS at higher laser intensities is deemed favourable to the production of fast electrons.

Suppression of RBS can be achieved by enhancing RFS growth [38]. RFS can be enhanced by adding a small satellite pulse at the Stokes frequency $\omega_0 - \omega_p$ to the main pulse. In addition, it has been found that a number of laser and plasma parameters, such as plasma density profile, laser pulse shape, and laser frequency chirp, influence the balance between RBS and RFS, and thus indirectly fast electron generation. These results can be used to predict the effect of such parameters on the amount of fast electrons.

As an example, the findings from the previous paragraphs will be used to resolve the “chirp controversy”. Laser chirp (position dependent laser frequency) has been found to influence the amount of high-energy electrons generated in laser-plasma interaction quite strongly, both in simulations [39], and experiments [40]. Chirp is usually induced by detuning the double grating compressor in the CPA scheme (cf. Section 1.1). This not only causes the laser frequency to become position-dependent, but it also increases the pulse length. In the simulations as well as the experiments, positive chirp (frequency increases from front to back) seemed to cause a much higher fast electron yield than negative chirp. However,

an incorrect definition of chirp was used in the simulations (the pulse length was kept fixed when chirp was added to the pulse), while the influence of chirp as seen in the experiments has later been explained from modulations of the pulse envelope caused by the laser optics [41], rather than from the position-dependent carrier frequency. The influence of chirp, in both the correct and incorrect implementation, and envelope modulation on RBS, RFS, and fast electron generation has been studied with XOOPIC. From the results, it can be concluded that linear laser chirp will not increase the yield of fast electrons, and that any previously observed increase of this yield has to be attributed to other causes.

1.6.4 Charged particle optics

As mentioned above, the problem in describing the (magnetic) fields of multipole devices is that expansion to a few orders in the transverse coordinates already involves more terms than one can keep track of. To overcome this problem, we approached the issue from a different angle. In our approach, we exploited many of the inherent properties of a magnetic field in vacuum, and of its mathematical description. We introduced a (harmonic) scalar potential for the field, and split it up into multipole contributions; these are shown to be harmonic themselves. Then we coupled this mathematical description to the boundary conditions provided by field measurements as they are done in practice. We wrote each multipole contribution as a linear combination of spline-like harmonic functions, and showed how the spline coefficients could be derived directly from the measured values of the magnetic field. This allows us to write not only the magnetic field in terms of these values, but also related quantities, like the multipole strength and the magnetic vector potential.

The description of the vector potential in terms of field measurements can be extended to the description of the actual transfer map of a multipole device, i.e. the function that maps the transverse phase space at the entry face to the transverse phase space at the exit face, or a combination of those. There are various ways to do this; one of them has been described in this thesis. Applications include easy calculation of the transfer functions of single beam guiding elements or clusters of such elements, as well as calculation of generalized field gradients for any multipole contribution up to any order. There are also numerous applications in low-energy electron optics.

Since there is a vast body of existing results on analytical particle optics, we will also show that the methods developed in Chapter 7 can also be used to calculate field related coefficients wherever they appear in literature on charged particle optics. This way, the new results are an extension of existing ones, rather than an attempt to replace them.

References

- [1] E. Esarey *et al.*, IEEE Trans. Plas. Sci. **24**, 252 (1996), and references therein.
- [2] J.S. Wurtele, Phys. Fluids B **5**, 2363 (1992).
- [3] J.S. Wurtele, Physics Today July 1994, p. 33.
- [4] A. Modena *et al.*, Nature **377**, 606 (1995).
- [5] D. Umstadter *et al.*, Science **273**, 472 (1996).
- [6] K. Nakajima *et al.*, Phys. Rev. Lett. **74**, 4428 (1995).
- [7] R.D. Ruth *et al.*, Part. Accel. **17**, 171 (1985).
- [8] P. Chen *et al.*, Phys. Rev. Lett. **54**, 693 (1985).
- [9] J.B. Rosenzweig *et al.*, Phys. Rev. Lett. **61**, 98 (1988).
- [10] R. Assmann *et al.*, SLAC-PROPOSAL-E-157 (1997).
<http://www.slac.stanford.edu/grp/arb/e162/>
- [11] M.J. Hogan *et al.*, Phys. Plasmas **7**, 2241 (2000).
- [12] T. Tajima and J.M. Dawson, Phys. Rev. Lett. **43**, 267 (1979).
- [13] C. Joshi *et al.*, Nature **311**, 525 (1984).
- [14] Y. Kitagawa *et al.*, Phys. Rev. Lett. **68**, 48 (1992).
- [15] M. Everett *et al.*, Nature **368**, 527 (1994).
- [16] D.F. Gordon, W.B. Mori, and C. Joshi, Phys. Plasmas **7**, 3156 (2000).
- [17] D. Strickland and G. Mourou, Opt. Comm. **55**, 447 (1985); *ibid.* Opt. Comm. **56**, 219 (1985).
- [18] C. A. Coverdale *et al.*, Phys. Rev. Lett. **74**, 4659 (1995).
- [19] S. V. Bulanov, F. Pegoraro, A. M. Pukhov, and A. S. Sakharov, Phys. Rev. Lett. **78**, 4205 (1997); S. V. Bulanov *et al.*, Fiz. Plazmy **25**, 517 (1999) [Plasma Phys. Rep. **25**, 468 (1999)].
- [20] S. Bulanov, N. Naumova, F. Pegoraro, and J. Sakai, Phys. Rev. E **58**, 5257 (1998).
- [21] D. Umstadter, J.K. Kim, and E. Dodd, Phys. Rev. Lett. **76**, 2073 (1996).

-
- [22] E. Esarey *et al.*, Phys. Rev. Lett. **79**, 2682 (1997).
 - [23] William L. Kruer, *The Physics of Laser Plasma Interactions* (Addison-Wesley, Reading MA, 1988).
 - [24] C.S. Liu and V.K. Tripathi, *Interaction of electromagnetic waves with electron beams and plasmas* (World Scientific, Singapore, 1994).
 - [25] A. Ting, E. Esarey, and P. Sprangle, Phys. Fluids B **2**, 1390 (1990).
 - [26] P. Sprangle *et al.*, Phys. Rev. Lett. **69**, 2200 (1992).
 - [27] E. Esarey *et al.*, Phys. Fluids B **5**, 2690 (1993).
 - [28] D. Kaganovich *et al.*, Phys. Rev. E **59**, R4769 (1999).
 - [29] N.A. Bobrova *et al.*, Fiz. Plazmy **26**, 12 (2000) [Plasma Phys. Rep. **26**, 10 (2000)].
 - [30] D.J. Spence and S.M. Hooker, Phys. Rev. E **63**, 015401 (2001).
 - [31] S. Jackel *et al.*, Opt. Lett. **20**, 1086 (1995).
 - [32] M. Borghesi *et al.*, Phys. Rev. E **57**, R4899 (1998).
 - [33] F. Dorchies *et al.*, Phys. Rev. Lett. **82**, 4655 (1999).
 - [34] J.R. Davies and J.T. Mendonca, Phys. Rev. E **62**, 7168 (2000).
 - [35] B. Cros *et al.*, Phys. Rev. E **65**, 026405 (2002).
 - [36] A.G. Khachatryan, Phys. Rev. E **65**, 046504 (2002).
 - [37] J. P. Verboncoeur, A. B. Langdon, and N. T. Gladd, Comput. Phys. Commun. **87**, 199 (1995); J. Wurtele and H.-J. Lee (private communication).
 - [38] M.J. Everett *et al.*, Phys. Rev. Lett. **74**, 2236 (1995); *ibid.* Phys. Plasmas **3**, 2041 (1996).
 - [39] E.S. Dodd and D. Umstadter, Phys. Plasmas **8**, 3531 (2001).
 - [40] W.P. Leemans *et al.*, Phys. Plasmas **8**, 2510 (2001).
 - [41] W.P. Leemans *et al.*, Phys. Rev. Lett. **89**, 174802 (2002).

Chapter 2

Fluid description of laser-plasma interaction

In this chapter, the interaction of a laser pulse with an underdense plasma is studied by means of a fluid description. Equations for the evolution of the plasma electron fluid are derived, while the plasma ions are regarded as stationary. An extension is made to include externally injected electrons in the model, which are not covered by the standard fluid description.

For a more elaborate treatment of the topics presented here, see Reitsma [1].

2.1 Plasma fluid equations

In this chapter, we make the following assumptions about the plasma.

- The plasma is regarded as a cold electron fluid, including the effects of Debye shielding,
- The plasma is fully ionized,
- The ions are immobile with background density n_0 ,
- The plasma is globally charge neutral.

The first assumption does not cover electrons that are externally injected into the plasma. In general, their velocity will be too large for Debye shielding to work, and their interaction with the plasma is governed by other mechanisms.

Under these assumptions, the relativistic equation of motion for the plasma electron fluid reads:

$$\frac{\partial \mathbf{p}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{p} = -e(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \quad (2.1)$$

Here, \mathbf{E} and \mathbf{B} denote the electric and magnetic fields within the plasma. The symbols \mathbf{v} and \mathbf{p} denote the plasma electron fluid velocity and momentum, respectively. They are related according to

$$\mathbf{p} = \gamma m_e \mathbf{v}, \quad \gamma = \sqrt{1 + p^2/(mc)^2} = 1/\sqrt{1 - v^2/c^2}.$$

The EM fields are governed by Maxwell's equations:

$$\begin{aligned} \nabla \cdot \mathbf{B} &= 0, \\ \nabla \cdot \mathbf{E} &= \rho/\varepsilon_0, \\ \frac{\partial \mathbf{B}}{\partial t} + \nabla \times \mathbf{E} &= 0, \\ \frac{\partial \mathbf{E}}{\partial t} - c^2 \nabla \times \mathbf{B} &= -\mathbf{J}/\varepsilon_0. \end{aligned}$$

Here, ρ and \mathbf{J} denote the charge and current density within the plasma, given by $\rho = (n_0 - n)e$, $\mathbf{J} = -ne\mathbf{v}$, with n the plasma electron density. These equations already contain the continuity equation for n :

$$\frac{\partial n}{\partial t} + \nabla \cdot (n\mathbf{v}) = 0,$$

Using the identity

$$m_e c^2 \nabla \gamma = (\mathbf{v} \cdot \nabla) \mathbf{p} + \mathbf{v} \times (\nabla \times \mathbf{p}),$$

we rewrite (2.1) as follows:

$$\frac{\partial \mathbf{p}}{\partial t} + e\mathbf{E} + m_e c^2 \nabla \gamma = \mathbf{v} \times (\nabla \times \mathbf{p} - e\mathbf{B}). \quad (2.2)$$

Taking the curl of this equation gives

$$\frac{\partial \boldsymbol{\Omega}}{\partial t} - \nabla \times (\mathbf{v} \times \boldsymbol{\Omega}) = \mathbf{0},$$

where $\boldsymbol{\Omega} \equiv \nabla \times \mathbf{p} - e\mathbf{B}$ is called the *generalized vorticity*. This equation implies that if $\boldsymbol{\Omega}$ is zero at $t = 0$, then it will remain zero for all $t > 0$. In laser plasma interaction, if we assume that the plasma is neutral and at rest before the laser pulse enters it, we have $\boldsymbol{\Omega} = \mathbf{0}$ at $t = 0$ and thus $\boldsymbol{\Omega} = \mathbf{0}$ for all $t > 0$.

We introduce scalar and vector potentials for the electromagnetic fields in the usual way:

$$\mathbf{E} = -\nabla \phi - \frac{\partial \mathbf{A}}{\partial t}, \quad \mathbf{B} = \nabla \times \mathbf{A}.$$

In this thesis, the Coulomb gauge will be used for \mathbf{A} , i.e. $\nabla \cdot \mathbf{A} = 0$, unless otherwise indicated. In terms of these potentials, $\boldsymbol{\Omega} = \mathbf{0}$ becomes $\nabla \times (\mathbf{p} - e\mathbf{A}) = \mathbf{0}$, and the equation of motion (2.2) can be written as

$$\frac{\partial}{\partial t}(\mathbf{p} - e\mathbf{A}) = \nabla(e\phi - \gamma m_e c^2).$$

The term $-m_e c^2 \nabla \gamma$ in this equation is known as the *ponderomotive force*.

We proceed to rendering all variables dimensionless through scaling. This will allow us to better compare the order of magnitude of the various terms in the equations. First, the time is scaled with some reference frequency ω_r , to be determined later: $t \rightarrow \omega_r t$. We then proceed to scale all other quantities: $x \rightarrow \omega_r x/c$, $v \rightarrow v/c$, $p \rightarrow p/(m_e c)$, $E \rightarrow eE/(m_e \omega_r c)$, $B \rightarrow eB/(m_e \omega_r)$, $\phi \rightarrow e\phi/(m_e c^2)$, $A \rightarrow eA/(m_e c)$, and $n \rightarrow n/n_0$. Then the dimensionless equations for the plasma become:

$$\nabla \times (\mathbf{p} - \mathbf{A}) = \mathbf{0}, \quad (2.3)$$

$$\frac{\partial}{\partial t}(\mathbf{p} - \mathbf{A}) = \nabla(\phi - \gamma), \quad (2.4)$$

$$\Delta \mathbf{A} - \frac{\partial^2 \mathbf{A}}{\partial t^2} = \varepsilon^2 \frac{n}{\gamma} \mathbf{p} + \nabla \frac{\partial \phi}{\partial t}, \quad (2.5)$$

$$\Delta \phi = \varepsilon^2 (n - 1), \quad (2.6)$$

$$\frac{\partial n}{\partial t} + \nabla \left(\frac{n}{\gamma} \mathbf{p} \right) = 0. \quad (2.7)$$

Here, $\varepsilon = \omega_p/\omega_r$, where ω_p denotes the plasma frequency given by

$$\omega_p^2 = n_0 e^2 / (\varepsilon_0 m_e).$$

Convenient choices for ω_r are $\omega_r = \omega_p$ ($\varepsilon = 1$), or $\omega_r = \omega_0$ ($\varepsilon \ll 1$), where ω_0 denotes the laser carrier frequency. Propagation in free space corresponds to $\varepsilon = 0$.

2.2 Separation of time scales

The set of equations (2.3)-(2.7) completely describes the behaviour of the cold plasma electron fluid under the influence of electromagnetic fields. So far, no distinction has been made between EM fields that originate from the plasma itself, and those that originate from outside the plasma, such as the EM fields of a laser pulse that interacts with the plasma. In laser-plasma interaction, the laser EM fields usually operate on a much shorter time scale than those generated by the plasma response. This allows one to treat the evolution of laser and plasma EM fields separately.

The separation of time scales works as follows. For a laser pulse having frequency ω_0 to propagate through a plasma having plasma frequency ω_p , one must have $\omega_0 > \omega_p$. For the pulse to propagate over a reasonable distance, one must have $\omega_0 \gg \omega_p$, otherwise the pulse depletes too quickly through reflection or parametric instabilities. Under these circumstances, any quantity that evolves on the slow

time scale, i.e. with $\omega_p t$, can be considered constant on the fast time scale. Conversely, any quantity that evolves on the fast time scale only, i.e. with $\omega_0 t$, averages to 0 when observed on the slow time scale. This allows us to regard any quantity Q as the sum of a fast and a slow contribution, i.e. $Q = Q_s + Q_f$, where Q_s is given by the average $\langle Q \rangle$ of Q over the fast time scale. Then the equations (2.3)-(2.7) can be rewritten as two sets of equations, one for the fast parts and one for the slow parts.

In order to separate the evolution of fast and slow quantities, we set $\omega_r = \omega_0$ in (2.3)-(2.7) so $\varepsilon = \omega_p/\omega_0 \ll 1$, and determine the order in powers of ε of the various derivatives of fast and slow quantities. We observe that the temporal and some of the spatial derivatives of a fast quantity Q_f will be much larger than those of a slow quantity Q_s . For a laser pulse that moves in the z -direction with velocity $\simeq c$, and taking Q_f, Q_s both $\mathcal{O}(1)$, we have

$$\begin{aligned} \frac{\partial Q_f}{\partial t} &= \mathcal{O}(1), & \frac{\partial Q_f}{\partial z} &= \mathcal{O}(1), & \nabla_{\perp} Q_f &= \mathcal{O}(\varepsilon), \\ \frac{\partial Q_s}{\partial t} &= \mathcal{O}(\varepsilon), & \frac{\partial Q_s}{\partial z} &= \mathcal{O}(\varepsilon), & \nabla_{\perp} Q_s &= \mathcal{O}(\varepsilon). \end{aligned}$$

Here, we have assumed that even fast quantities evolve slowly in a direction perpendicular to that of the laser pulse.

It is worth noting that in the expression $Q = Q_f + Q_s$, Q_f and Q_s might be of very different order in ε without this being obvious. For this reason, both Q_s and Q_f will be taken to be $\mathcal{O}(1)$ and then multiplied by the correct power of ε . For example, in the expression $Q_f + \varepsilon^2 Q_s$, the fast part is $\mathcal{O}(1)$ and the slow part is $\mathcal{O}(\varepsilon^2)$.

We proceed to determine the fast and slow contributions to the various field and particle quantities. The scaling of the vector potential \mathbf{A} is considered first. From $\nabla \cdot \mathbf{A}_f = \nabla \cdot \mathbf{A}_s = 0$, we find that $\mathbf{A}_f = \mathbf{A}_{\perp,f} + \varepsilon A_{z,f} \hat{e}_z$. Assuming that \mathbf{A}_s has the same order in ε as $\mathbf{A}_{\perp,f}$, we write

$$\mathbf{A} = \mathbf{A}_{\perp,f} + \varepsilon A_{z,f} \hat{e}_z + \mathbf{A}_s.$$

A consequence of this assumption is that the slow parts of \mathbf{E} and \mathbf{B} are one order of ε smaller than the fast parts.

The next two quantities to be treated are the electron fluid momentum \mathbf{p} and Lorentz factor γ . We observe that the fast transverse electron motion consists mainly of the electron quiver motion in the fast EM fields, and set $\mathbf{p}_{\perp,f} = \mathbf{A}_{\perp,f} + \mathcal{O}(\varepsilon) \hat{e}_{\perp}$. Then we find for γ :

$$\gamma^2 = 1 + \mathbf{p}^2 = 1 + (\mathbf{A}_{\perp,f} + \mathbf{p}_s)^2 + 2p_{z,s}p_{z,f} + p_{z,f}^2 + \mathcal{O}(\varepsilon).$$

For the slow part γ_s of γ , we obtain:

$$\gamma_s^2 = \langle \gamma^2 \rangle = 1 + \langle \mathbf{A}_{\perp,f}^2 \rangle + \langle p_{z,f}^2 \rangle + \mathbf{p}_s^2.$$

We assume that the laser pulse is circularly polarized, so $\mathbf{A}_{\perp,f}^2 - \langle \mathbf{A}_{\perp,f}^2 \rangle = \mathcal{O}(\varepsilon)$. This basically means that the contributions of higher harmonics to the laser field can be neglected. If a linear polarization were chosen, the contributions of higher harmonics would be $\mathcal{O}(1)$, i.e. not negligible at all. As a result, we find that $\partial\gamma/\partial z = \mathcal{O}(\varepsilon)$, and from (2.4) it is found that the fast part of p_z is $\mathcal{O}(\varepsilon)$ as well. We then obtain the following ordering for \mathbf{p} and γ :

$$\begin{aligned}\mathbf{p} &= \mathbf{p}_{\perp,f} + \varepsilon p_{z,f} \hat{e}_z + \mathbf{p}_s, \\ \gamma &= \gamma_s + \varepsilon \gamma_f,\end{aligned}$$

where γ_s is given by

$$\gamma_s = \sqrt{1 + \mathbf{p}_s^2 + U}, \quad U = \langle \mathbf{A}_{\perp,f}^2 \rangle.$$

Inserting this in (2.4) yields $\phi_f = \mathcal{O}(\varepsilon)$ or even smaller. This implies that the plasma density oscillation will be mainly driven by the slow ponderomotive force rather than the fast quiver motion. As a consequence, the plasma density is ordered as $n = n_s + \varepsilon n_f$. Inserting this into (2.6) yields $\phi = \phi_s + \varepsilon^3 \phi_f$ for the ordering of the electrostatic potential.

At this point, we are ready to write down the slow part of the wakefield equations (2.3)-(2.7). We observe that $n/\gamma = n_s/\gamma_s + \mathcal{O}(\varepsilon)$, and obtain:

$$\frac{\partial}{\partial t}(\mathbf{p}_s - \mathbf{A}_s) = \nabla(\phi_s - \gamma_s), \quad (2.8)$$

$$\Delta \mathbf{A}_s - \frac{\partial^2 \mathbf{A}_s}{\partial t^2} = \varepsilon^2 \frac{n_s}{\gamma_s} \mathbf{p}_s + \nabla \frac{\partial \phi_s}{\partial t}, \quad (2.9)$$

$$\Delta \phi_s = \varepsilon^2 (n_s - 1). \quad (2.10)$$

Included in these equations is the slow part of the continuity equation:

$$\frac{\partial n_s}{\partial t} + \nabla \left(\frac{n_s}{\gamma_s} \mathbf{p} \right) = 0.$$

These equations describe the slow plasma response to a laser pulse with given $\mathbf{A}_{\perp,f}$. It should be noted that the dependence of γ_s on $\langle \mathbf{A}_{\perp,f}^2 \rangle$ is the only way in which the fast laser EM fields couple to the slow wakefield. Since the slow plasma response is dominated by ω_p rather than ω_0 , we set $\omega_r = \omega_p$, i.e. $\varepsilon = 1$, for the slow wakefield equations (2.8)-(2.10) in this and the following two chapters.

Regarding the fast part of the wakefield equations, it is found that the only fast quantities of $\mathcal{O}(1)$ are the fast vector potential $\mathbf{A}_{\perp,f}$ and the fast transverse momentum $\mathbf{p}_{\perp,f}$. Using (2.4) to eliminate $\mathbf{p}_{\perp,f}$ and neglecting all terms of higher order in ε , the fast wakefield equations reduce to a single equation (non-linear Klein-Gordon equation) for $\mathbf{A}_{\perp,f}$:

$$\left[\Delta - \frac{\partial^2}{\partial t^2} - \varepsilon^2 \frac{n_s}{\gamma_s} \right] \mathbf{A}_{\perp,f} = \mathbf{0}. \quad (2.11)$$

For this equation, $\varepsilon = \omega_p/\omega_0$ is retained. The evolution of the fast EM fields is coupled to the slow plasma response through the term $\varepsilon^2(n_s/\gamma_s)\mathbf{A}_{\perp,f}$. This means that the systems of fast and slow equations can only be considered as independent if ε is small, i.e. when $\omega_p \ll \omega_0$.

Usually, one assumes $\mathbf{A}_{\perp,f} = \mathbf{a}(\mathbf{r}, t) \exp[i(t - z/v_\varphi)]$, where $\mathbf{a}(\mathbf{r}, t)$ denotes the envelope of $\mathbf{A}_{\perp,f}$ evolving on the slow time scale. Inserting this into (2.11) yields an envelope equation for \mathbf{a} and the dispersion relation $1 - 1/v_\varphi^2 - (\omega_p/\omega_0)^2 n_s/\gamma_s = 0$. For 1-D calculations in the linear regime, the approximation $n_s \approx 1$ is often used. Then the dispersion relation reduces to $v_\varphi = 1/\sqrt{1 - (\omega_p^2/\omega_0^2)/\gamma_s}$, i.e. $v_g = \sqrt{1 - (\omega_p^2/\omega_0^2)/\gamma_s}$, and $\gamma_g = \gamma_s^{1/2} \omega_0/\omega_p \gg 1$. Here, v_g and γ_g are the pulse group velocity and corresponding Lorentz factor. For pulses that propagate in a plasma channel in 2-D or 3-D calculations, the finite spot size and channel width must be taken into account. This renders the dispersion relation more complicated.

In previous research [3,4], it has been shown that the envelope of a short ($L \leq \lambda_p/2$) laser pulse in the linear or weakly non-linear regime ($a_0 \leq 1$) displays very little evolution over very long distances, especially so when it propagates in a matched parabolic plasma density channel. This implies that, while $\partial_t \mathbf{a}_\perp$ and $\partial_z \mathbf{a}_\perp$ are of $\mathcal{O}(\varepsilon)$, the evolution with respect to a frame moving at the same speed as the laser pulse must be of higher order in ε . In other words, $(\partial_t + v_g \partial_z) \mathbf{a}_\perp = \mathcal{O}(\varepsilon^2)$. This ensures that the evolution of the pulse envelope after propagation over its own length is negligible. For this reason, in this and the next two chapters, the pulse envelope will not be calculated from (2.11), but prescribed directly as a function of $z - v_g t$, where v_g is obtained from the above dispersion relation. Of course, this approximation cannot and will not be made in Chapters 5 and 6, where we operate in a regime in which Eq. (2.11) plays an important role.

2.3 Wakefield potential

We consider the situation of a laser pulse propagating through a plasma with a group velocity $v_g \hat{e}_z$. We assume that the pulse is non-evolving, and that the plasma background density does not depend on the longitudinal coordinate z . Then it is convenient to change to a comoving coordinate system $\zeta = z - v_g t$, $\tau = t$. Partial derivatives to ζ and τ are given by

$$\frac{\partial}{\partial z} = \frac{\partial}{\partial \zeta}, \quad \frac{\partial}{\partial t} = \frac{\partial}{\partial \tau} - v_g \frac{\partial}{\partial \zeta}.$$

For a laser pulse that is not too intense, so the wakefield it excites is not too non-linear, the evolution of both the laser pulse and the wakefield in the comoving

frame is on a timescale much larger than ω_p^{-1} . In that case, all explicit τ dependence in the wakefield equations can be neglected, i.e. $\partial/\partial t = -v_g \partial/\partial \zeta$. This is called the *quasi-static approximation*. Under this assumption, the wakefield equations become (subscript “s” dropped):

$$-v_g \frac{\partial}{\partial \zeta} (\mathbf{p}_\perp - \mathbf{A}_\perp) = -v_g \nabla_\perp (p_z - A_z) = \nabla_\perp (\phi - \gamma), \quad (2.12)$$

$$-v_g \frac{\partial}{\partial \zeta} (p_z - A_z) = \frac{\partial}{\partial \zeta} (\phi - \gamma), \quad (2.13)$$

$$\Delta \mathbf{A} - v_g^2 \frac{\partial^2 \mathbf{A}}{\partial \zeta^2} = n \mathbf{v} - v_g \frac{\partial}{\partial \zeta} \nabla \phi, \quad (2.14)$$

$$\Delta \phi = n - 1. \quad (2.15)$$

From (2.12) and (2.13) we find $\nabla(\gamma - v_g p_z) = \nabla(\phi - v_g A_z)$, which yields on integration

$$\gamma - v_g p_z = 1 + \phi - v_g A_z. \quad (2.16)$$

We define the *wakefield potential* Ψ as $\Psi = 1 + \phi - v_g A_z$. This quantity plays an important role in the description of plasma wakefields. It can be shown that in the quasistatic approximation, a regular plasma wakefield can be described entirely in terms of the wakefield potential. Also, the dynamics of charged particles in the wakefield are governed by this wakefield potential.

The remainder of this section shall be devoted to the derivation of simplified versions of the wakefield equations that are entirely expressed in terms of Ψ , for 1-dimensional and 2-dimensional slab geometries.

1-D geometry In 1-D geometry, the Coulomb gauge reduces to $\partial A_z / \partial z = 0$. Without loss of generality, one can take $A_z = 0$. From (2.12), we find that $\mathbf{p}_\perp = \mathbf{A}_\perp$. The equation for \mathbf{A}_\perp then reads:

$$(1 - v_g^2) \frac{\partial^2 \mathbf{A}_\perp}{\partial \zeta^2} = \frac{n}{\gamma} \mathbf{A}_\perp.$$

The right-hand side of this equation is $\mathcal{O}(1)$, while the left-hand side is $\mathcal{O}(1 - v_g^2)$, where $1 - v_g^2 = \omega_p^2 / \omega_0^2 \ll 1$. Therefore, the only possible solution is $\mathbf{A}_\perp = \mathbf{0}$. As a consequence, the slow 1-D wakefield neither induces a magnetic field nor transverse motion of the plasma fluid. Then we have $\Psi = 1 + \phi$ and the wakefield equations (2.12)-(2.15) simplify to

$$\Psi = \gamma - v_g p_z, \quad (2.17)$$

$$\frac{\partial^2 \Psi}{\partial \zeta^2} = n - 1, \quad (2.18)$$

$$n(v_g - v_z) = v_g. \quad (2.19)$$

Using $\gamma = \sqrt{1 + p_z^2 + U}$ and (2.17), we express γ and p_z in terms of Φ :

$$\begin{aligned}\gamma &= \gamma_g^2 \left(\Psi - v_g \sqrt{\Psi^2 - (1 + U)/\gamma_g^2} \right), \\ p_z &= \gamma_g^2 \left(v_g \Psi - \sqrt{\Psi^2 - (1 + U)/\gamma_g^2} \right).\end{aligned}$$

Since p is a real-valued quantity by definition, these relations are only valid for $\Psi^2 > (1 + U)/\gamma_g^2$. This follows from the fact that $\partial\Psi/\partial p_z = 0$ at this point, and there is a one-to-one correspondence between p_z and Ψ for $p_z < \gamma_g v_g \sqrt{1 + U}$ only. This upper limit on p_z , as derived from the equation of motion (2.17) will prove important in Chapter 4, when wave breaking of plasma waves is investigated.

Equation (2.19) also contains a singularity, occurring when $v_z \uparrow v_g$, and thus $p_z \uparrow \gamma_g v_g \sqrt{1 + U}$. For a cold plasma, this singularity coincides with the singularity in (2.17). In Chapter 4, we will see that in case of a warm plasma, the singularity in (2.17) occurs already at lower p_z , while that in (2.19) will remain at $v_z = v_g$, meaning that both singularities will no longer coincide.

Using (2.19), we obtain the following equation for Ψ :

$$\frac{\partial^2 \Psi}{\partial \zeta^2} = \gamma_g^2 \left(\frac{v_g \Psi}{\sqrt{\Psi^2 - (1 + U)/\gamma_g^2}} - 1 \right). \quad (2.20)$$

Several limit cases may be of interest. In the limit $v_g \rightarrow 1$, Eq. (2.20) transforms into

$$\frac{\partial^2 \Psi}{\partial \zeta^2} = -\frac{\Psi^2 - (1 + U)}{2\Psi^2}.$$

In the limit of a small amplitude plasma wave, (2.20) can be linearized. We substitute $\Psi = 1 + \delta\Psi$ and neglect terms of $\mathcal{O}(\delta\Psi^2)$ and all terms smaller than $\mathcal{O}(1/\gamma_g^2)$ to obtain

$$\left(\frac{\partial^2}{\partial \zeta^2} + \frac{1}{v_g^2} \right) \delta\Psi = \frac{U}{2v_g^2}. \quad (2.21)$$

Green's function for this equation with the boundary condition $\lim_{\zeta \rightarrow \infty} \delta\Psi(\zeta) = 0$ reads

$$G(\zeta, \zeta') = H(\zeta' - \zeta) \sin[(\zeta' - \zeta)/v_g],$$

where $H(\zeta)$ denotes the Heaviside step function. The solution to (2.21) reads:

$$\delta\Psi(\zeta) = \frac{1}{2v_g^2} \int_{-\infty}^{\infty} U(\zeta') G(\zeta, \zeta') d\zeta'.$$

For a pulse of finite size, e.g. located at $-L < \zeta < 0$, the wakefield behind the pulse ($\zeta < L$) is given by

$$\begin{aligned}\delta\Psi(\zeta) &= I_s \cos(\zeta/v_g) - I_c \sin(\zeta/v_g), \\ I_s &= \frac{1}{2v_g^2} \int_{-L}^0 U(\zeta') \sin(\zeta'/v_g) d\zeta', \\ I_c &= \frac{1}{2v_g^2} \int_{-L}^0 U(\zeta') \cos(\zeta'/v_g) d\zeta' .\end{aligned}$$

In other words, a low-intensity laser pulse in a 1-D setting drives a sinusoidal wakefield, of which the amplitude depends on the integrals I_s and I_c . The wakefield will have maximum amplitude for $L/v_g \approx \pi$, i.e. when the pulse length is half a plasma wavelength. This is called *resonant laser wakefield excitation*. Also, the wakefield amplitude tends to 0 for $L/v_g > 2\pi$, since then $U(\zeta)$ is averaged over multiple periods of the plasma oscillation and I_s, I_c tend to 0. This means that plasma wave excitation through the ponderomotive force $\nabla\gamma$ does not work for pulses longer than the plasma wavelength. The fact that such pulses are still found to excite plasma waves can almost entirely be ascribed to the action of the *Raman forward scattering* instability together with *bunching* of the laser EM fields. These issues will be treated in detail in Chapters 5 and 6.

2-D geometry In 2-D slab geometry, we take x to be the transverse coordinate and drop all y -dependence from (2.12)-(2.15). The resulting equations are still quite complicated, so we simplify them further by taking the limit $v_g \rightarrow 1$. Since this comes down to assuming that $(1 - v_g^2)\partial^2/\partial\zeta^2 = 0$, this simplification is only valid for plasma waves that are far from breaking conditions. (Just before breaking, a plasma wave displays large ζ -derivatives for which the assumption does not hold; see Chapter 4.) In this limit, all wakefield quantities can be expressed in the wakefield potential Ψ as follows:

$$\begin{aligned}\eta &\equiv \frac{n}{\gamma} = \frac{1}{\Psi} \left(1 + \frac{\partial\Psi}{\partial x} \right), \\ p_x &= \frac{1}{\eta} \frac{\partial^2\Psi}{\partial\zeta\partial x}, \\ p_z &= \frac{1}{2\Psi} (1 + p_x^2 + U - \Psi^2), \\ \gamma &= \frac{1}{2\Psi} (1 + p_x^2 + U + \Psi^2).\end{aligned}$$

Equations (2.12) and (2.13) have been used to eliminate \mathbf{A} . The remaining wakefield equations reduce to a single equation for Ψ :

$$\frac{\partial^2\Psi}{\partial\zeta^2} - \frac{\partial^2\Psi}{\partial x^2} - \frac{\partial^2}{\partial\zeta\partial x} \left(\frac{1}{\eta} \frac{\partial^2\Psi}{\partial\zeta\partial x} \right) + \eta\Psi = \eta\gamma - \frac{\partial^2\gamma}{\partial x^2}. \quad (2.22)$$

For small amplitude waves ($\Psi \approx 1 + \delta\Psi$, $\gamma \approx 1 + U/2$, $\eta \approx 1$), this equation can be linearized as follows:

$$\left(1 + \frac{\partial^2}{\partial \zeta^2}\right) \left(1 - \frac{\partial^2}{\partial x^2}\right) \delta\Psi = \frac{1}{2} \left(1 - \frac{\partial^2}{\partial x^2}\right) U.$$

Green's function for this equation with boundary conditions $\lim_{\zeta \rightarrow \infty} \delta\Psi(x, \zeta) = \lim_{|x| \rightarrow \infty} \delta\Psi(x, \zeta) = 0$ reads

$$\begin{aligned} G_2(x, x', \zeta, \zeta') &= \frac{1}{2} \exp(-|x' - x|) H(\zeta' - \zeta) \sin(\zeta' - \zeta) \\ &= \frac{1}{2} \exp(-|x' - x|) G(\zeta, \zeta'; v_g = 1). \end{aligned}$$

We see that both the 2-D wakefield equation and the corresponding function of Green have the same longitudinal behaviour as in one dimension. This means that the longitudinal effect of charged particles on the wakefield can be described correctly in 1-D, so the use of the 1-D approximation in the next chapter can be justified. However, as we will see in the next section, the mechanisms behind wakefield excitation in one and two dimensions are entirely different.

2.4 Charged particles in the plasma

When charged particles are brought into the laser wake for acceleration, the wake does not only govern the motion of these particles, but the presence of the particles influences the evolution of the wake as well. In addition, the particles also influence each other, either through direct Coulomb interaction or through the way each particle modifies the wake behind itself. This section will be devoted to the study of particle-wake and particle-particle interaction, and will give justification to some common assumptions.

We start by studying the interaction between two particles, both having charge q and moving with relativistic velocities $v_1 \hat{e}_z$ and $v_2 \hat{e}_z$. From relativistic EM field theory [2], we find that the potentials for particle 1 are given by

$$\phi = \frac{kq\gamma_1}{\sqrt{\gamma_1^2(z - v_1 t)^2 + x^2 + y^2}}, \quad \mathbf{A} = v_1 \phi \hat{e}_z,$$

where $k = e/(4\pi\epsilon_0 m_e c^3)$. Note that in this section only, the Lorenz gauge¹ has been employed. The electromagnetic fields for this particle are given by

$$\mathbf{E} = -\left(\frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial y}, (1 - v_1^2)\frac{\partial\phi}{\partial z}\right), \quad \mathbf{B} = v_1 \left(\frac{\partial\phi}{\partial y}, -\frac{\partial\phi}{\partial x}, 0\right).$$

¹Ludvig V. Lorenz, Phil. Mag. Ser. 3 **34**, 287 (1867).

The Lorentz force² on particle 2 due to the fields of 1 is

$$F_z = -(1 - v_1^2) \frac{\partial \phi}{\partial z}, \quad \mathbf{F}_\perp = (1 - v_1 v_2) \nabla_\perp \phi.$$

Now let particle 1 be an electron trapped in the laser wake with $\gamma_1 \gg \gamma_g = 1/\sqrt{1 - v_g^2}$. We then find that $F_z = \mathcal{O}(1/\gamma_1^2)$, regardless of the velocity of particle 2. Things are different for \mathbf{F}_\perp . If particle 2 is another electron trapped in the wake with $\gamma_2 \approx \gamma_1$, then we have $\mathbf{F}_\perp = \mathcal{O}(1/\gamma_1)$. However, if particle 2 is a plasma electron having $\gamma_2 < \gamma_g \ll \gamma_1$ (a necessary condition to avoid wave breaking), we have $\mathbf{F}_\perp = \mathcal{O}(\gamma_1(1 - v_1 v_2))$. The ratio of the magnitudes of these forces is to leading order $2\gamma_2^2/\gamma_1^2 \ll 1$. This means that both the longitudinal force by particle 1 in general and the transverse force it exerts on other trapped particles can be neglected with respect to the transverse force it exerts on particles in the plasma fluid, provided that the average distance between plasma particles is not larger than the average distance between trapped particles. That is, the density of trapped particles should not be larger than the plasma electron density.

Assuming the density condition to be satisfied, we observe that trapped electrons in the plasma wakefield interact by modifying the wake which influences other trapped electrons, while direct Coulomb interaction can be neglected. This implies that the interaction is not symmetric: a trapped electron can only influence electrons that follow it, not those that precede it.

The influence of a bunch of trapped electrons on the wakefield can be worked into the wakefield equations as follows. We assume that the bunch evolves slowly with respect to the plasma oscillation while its velocity is approximately v_g , so the plasma response to any change in the bunch density and current is considered to happen instantaneously. We also assume that the transverse bunch current can be neglected with respect to the transverse current of the plasma fluid. This results in $\rho_b = -n_b$, $\mathbf{J}_b = (0, 0, -v_g n_b)$ for the charge and current density caused by the bunch. These contributions are then added to Eqns. (2.12)-(2.15). They result in a source term n_b , to be added to the wakefield equations (2.20) and (2.22) and their linearized counterparts. The contribution of a unit charge at ζ' or (x', ζ') to the wakefield is then $G(\zeta, \zeta')$ or $G_2(x, x', \zeta, \zeta')$ for 1-D or 2-D respectively. In this light, the choice of the boundary condition for $\zeta \rightarrow \infty$ is a direct consequence of the requirement that trapped electrons (or laser pulses for that matter) do not disturb the plasma in front of them.

We now take a closer look at how a trapped electron creates its wake. In 1-D, a “particle” is actually an (x, y) -plane with a certain surface charge density. Because of the surface charge, the E_z field experiences a jump in the z -direction. This jump does not decrease with γ , but remains constant regardless of the speed of the particle. The wakefield excited by a 1-D particle is a direct consequence

²Hendrik A. Lorentz

of the occurrence of this jump. In 2-D, a “particle” is a line in the y -direction with a certain line charge density. It can be shown that the longitudinal field E_z for such a line is of $\mathcal{O}(1/\gamma)$ and thus vanishes completely in the limit $v_g \rightarrow 1$. Therefore, the wakefield for a particle in 2-D (or 3-D for that matter) needs to be excited through a different mechanism. As follows from Jackson [2] Section 11.10, the transverse force a charged particle at ultrarelativistic speed exerts on the plasma fluid takes the form of a short transverse kick, with intensity of $\mathcal{O}(\gamma)$ and duration of $\mathcal{O}(1/\gamma)$. Thus, its net effect tends to a nonzero constant for $v_g \rightarrow 1$, while its duration tends to 0. This infinitesimal kick causes the sudden onset of the wakefield at the precise location of the particle, while leaving the wakefield in front of it undisturbed. So in 2-D or 3-D, the jump in the wakefield at the position of the particle in the longitudinal direction is actually caused by the action of the transverse fields.

Fortunately, the wakefields for particles in 1-D or 2-D have the same longitudinal ζ -dependence, even though they result from entirely different mechanisms. This justifies the use of the 1-D approximation in the wakefield calculations presented in the next chapter. It should be noted that the x - and ζ -dependence of the 2-D wakefield can be decoupled only because the transverse motion of the particle has been neglected when calculating the particle current \mathbf{J}_b , and because the plasma response is assumed to be instantaneous.

References

- [1] A.J.W. Reitsma, *Electron bunch quality in laser wakefield acceleration*, Ph. D. thesis, Technische Universiteit Eindhoven, The Netherlands, 2002.
- [2] J.D. Jackson, *Classical electrodynamics*, Third Edition (Wiley, New York, 1998).
- [3] P. Volfbeyn et al., Phys. Plasmas **4**, 3403 (1997).
- [4] E. Esarey et al., Phys. Rev. Lett. **84**, 3081 (2000).

Chapter 3

Energy spread in plasma-based acceleration

Electron acceleration in a one-dimensional plasma wave has been simulated, with emphasis on minimizing the energy spread of an accelerated electron bunch, while keeping the mean energy gain at a reasonable level. Bunch length, beam loading and the injection phase are tuned to reach this goal. The simulation results show that in a wide range of initial bunch lengths and beam loading parameters an optimum acceleration distance exists, which combines low energy spread and high energy gain. The energy spread at the optimum is found to be weakly dependent on bunch length and beam loading, while it is highly sensitive to deviations in the injection phase.

The contents of this chapter have been published in part as the article “Energy spread in plasma-based acceleration”, A.J.W. Reitsma, R.M.G.M. Trines, and V.V. Goloviznin, IEEE Trans. Plas. Sci. **28**, 1165 (2000).

3.1 Introduction

The worldwide study of excitation of strong plasma waves by using high-intensity lasers is directed towards the development of a compact laser-plasma accelerator [1,2]. The idea for such an accelerator has originally been proposed by Tajima and Dawson [3]. Until now, the emphasis, both in experimental work and simulations, has been on the high acceleration gradients and energy gain provided by such an accelerator. A very high acceleration gradient of 100 GeV/m has been demonstrated by Modena *et al.* [4]. This is three to four orders of magnitude higher than the acceleration gradients in conventional linacs. In this experiment, the electrons to be accelerated originated from the bulk plasma as a result of self-trapping in the plasma wave. For an externally injected electron bunch, an

energy gain of several MeV in a few mm has been demonstrated by Dorchie *et al.* [5]. Typically, the accelerated electrons come in a broad energy range from practically zero to some maximum value. Wilks *et al.* [6] showed that, in general, a low energy spread can be reached at the expense of energy gain. Van der Meer [7] proposed a special shaping of the electron bunch that reduces the energy spread considerably. Katsouleas *et al.* [8] developed a strategy for phasing and beam loading that minimizes energy spread without special shaping. The present work is an extension of the latter paper. In addition to the injection phase and beam loading, the bunch length as a tunable parameter has been included.

The calculations on longitudinal energy spread presented in this chapter have been performed in the one-dimensional approximation. Notwithstanding its limitations, this approximation is relevant for linear wake waves in a plasma accelerator with a hollow channel, as proposed by Chiou *et al.* [9]. For small beams (bunches) in a hollow channel, transverse beam emittance is well preserved and the transverse beam dynamics has little effect on energy gain and energy spread [10], so that we may consider the acceleration as a predominantly one-dimensional process.

This chapter is organized as follows. First, the basic Hamiltonian system for a particle in a linear wakefield is explored. The structure of the phase space for a single particle is investigated and those regions in which the particle is confined to a single crest of the wakefield are identified. From the equations of motion, the mean energy and energy spread of a short bunch of electrons are determined. Then the effect of beam loading, i.e. the way a charged particle affects the wakefield, is introduced. By combining the energy spread induced by the finite length of the bunch and the spread induced by beam loading, the acceleration length at which a minimum in the energy spread occurs is estimated. Subsequently, electron acceleration in a one-dimensional linear plasma wave is simulated. The simulations serve to determine those values for the bunch parameters that combine a large energy gain with a minimum energy spread. The acceleration length for which minimum spread occurs will be determined, allowing us to verify the analytical estimate of the optimal acceleration length. Finally, the sensitivity of the minimum in energy spread to errors in the various bunch parameters will be discussed.

3.2 Single electron dynamics

In this chapter, the dynamics of particles in a linearized 1-D wakefield will be investigated in the quasi-static approximation. As before, the wakefield potential is written as $\Psi = 1 + \delta\Psi$, and the linearized equation (2.21) for $\delta\Psi$ behind the

laser pulse ($U = 0$) reduces to

$$\left(\frac{\partial^2}{\partial \zeta^2} + k^2\right) \delta\Psi = 0, \quad (3.1)$$

where $k = 1/v_g$ denotes the wave number of the plasma wave. The solution can always be written in the form

$$\delta\Psi = -\Psi_0 \sin(k\zeta),$$

where z has been chosen for the longitudinal coordinate, $\zeta = z - v_g t$, and v_g denotes the velocity of the comoving frame of reference. The corresponding electric field is purely electrostatic and is given by

$$E = -\frac{\partial \delta\Psi}{\partial \zeta} = e_0 \cos(k\zeta), \quad (3.2)$$

where $e_0 = k\Psi_0$ is the amplitude of the wakefield. For $v_g \approx 1$, we have $k \approx 1 + 1/(2\gamma_g^2)$, where $\gamma_g = 1/\sqrt{1 - v_g^2}$. In the calculations presented in this chapter, $\gamma_g = 100$ will be taken, so we can safely use the approximation $k = 1$. This means that in scaled coordinates, one plasma wavelength equals 2π .

The equations of motion for a single electron with position ζ and momentum p are

$$\frac{d\zeta}{dt} = \frac{p}{\gamma} - v_g, \quad (3.3)$$

$$\frac{dp}{dt} = -e_0 \cos(\zeta) = \frac{\partial \delta\Psi}{\partial \zeta}, \quad (3.4)$$

where $\gamma = \sqrt{1 + p^2}$ as usual. Together with the Hamiltonian

$$H = \gamma - v_g p - \delta\Psi = \gamma - v_g p + e_0 \sin \zeta, \quad (3.5)$$

these equations constitute a Hamiltonian system.

We take a closer look at the Hamiltonian H . The momentum dependent part $\sqrt{1 + p^2} - v_g p$ reaches its minimum of $1/\gamma_g$ at $p = \gamma_g v_g$, whereas the position dependent part $e_0 \sin \zeta$ reaches its minimum of $-e_0$ at $\zeta = -\pi/2 + n \cdot 2\pi$, n integer. The points $\zeta = -\pi/2 + n \cdot 2\pi$, $p = \gamma_g v_g$ are local minima for H , and are called O-points. At these points, H equals $1/\gamma_g - e_0$. At the points $\zeta = \pi/2 + n \cdot 2\pi$, $p = \gamma_g v_g$ (n integer), the position dependent part reaches its maximum while the momentum dependent part is still minimum. These points are saddle points for H and are called X-points. A contour plot of $H(\zeta, p)$, i.e. a plot of curves along which H is constant, is given in Fig. 3.1; an O-point and an X-point are indicated. The driving laser pulse is to the right of the region depicted in the graph.

A single particle obeying the equations of motion (3.3) and (3.4) is confined to the curves $H(\zeta, p) = C$ in phase space, where C is some constant. Since $H \geq 1/\gamma_g - e_0$ for all ζ and p , we necessarily have $C \geq 1/\gamma_g - e_0$. The equation $H(\zeta, p) = C$ can be solved for p to obtain:

$$p(\zeta) = \gamma_g^2 \left[v_g(C - e_0 \sin \zeta) \pm \sqrt{(C - e_0 \sin \zeta)^2 - 1/\gamma_g^2} \right]. \quad (3.6)$$

We denote the roots by p_- and p_+ where $p_- \leq p_+$. Note that one always has $p_- \leq \gamma_g v_g \leq p_+$, so if the roots coincide, they coincide at $\gamma_g v_g$.

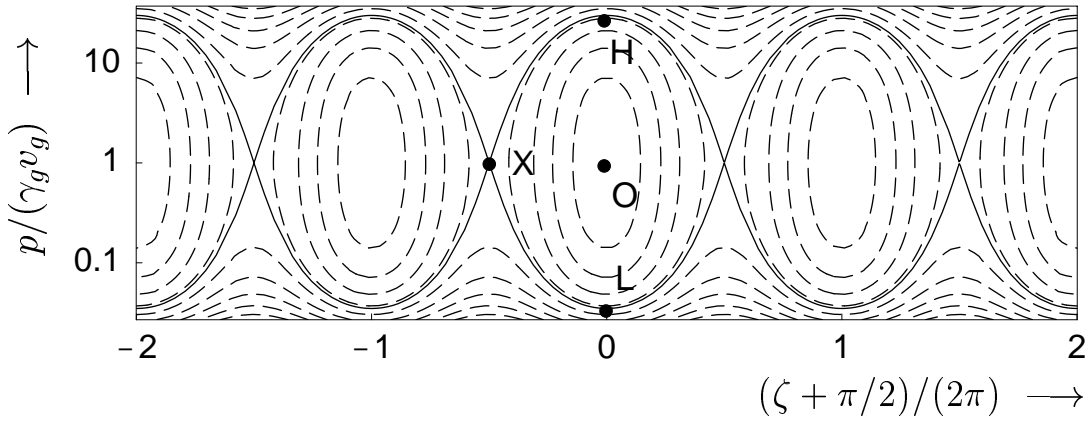


Figure 3.1: Contour plot of the Hamiltonian $H(\zeta, p)$. Particle orbits are defined by the curves of constant H . The solid curves denote the separatrices; the dashed curves denote other orbits. An O-point and an X-point are indicated. The points L and H denote the points of lowest and highest energy along a closed electron orbit in phase space. (Picture courtesy of A. Reitsma.)

We discern two cases: (i) $C > 1/\gamma_g + e_0$, and (ii) $1/\gamma_g - e_0 < C < 1/\gamma_g + e_0$. These will be treated separately.

(i) In this case, (3.6) yields two distinct solutions p_+ and p_- for any value of ζ . This implies that the curves $H = C$ in phase space are “open” curves that extend from $\zeta = -\infty$ to $\zeta = \infty$. In Fig. 3.1, these curves are near the bottom or top of the graph. A particle moving along the curve $(\zeta, p_-(\zeta))$ always has a lower velocity than the wakefield, and is called *untrapped*. A particle moving along the curve $(\zeta, p_+(\zeta))$ always has a higher velocity than the wakefield, and is called *detrapped*. Since such particles are not confined to a definite region of the wakefield, they are not considered useful for controlled acceleration.

(ii) In this case, (3.6) yields two distinct solutions p_+ and p_- only for ζ such that $C > 1/\gamma_g + e_0 \sin \zeta$, i.e. $-1 \leq \sin \zeta < (C - 1/\gamma_g)/e_0 < 1$. This is true for ζ in non-overlapping finite intervals with length $2L_C < 2\pi$ around the points

$-\pi/2 + n \cdot 2\pi$, n integer. For $\zeta = -\pi/2 \pm L_C$, we have $C - e_0 \sin \zeta = 1/\gamma_g$, and thus $p_- = p_+ = \gamma_g v_g$. This means that the curves $H = C$ are closed orbits in phase space centered around the O-points $(-\pi/2, \gamma_g v_g)$, having maximum width $2L_C$ at $p = \gamma_g v_g$. See also Fig. 3.1. For $C = 1/\gamma_g - e_0$, the curve $H = C$ reduces to a single point: the O-point itself. A particle moving along such a closed orbit is actually oscillating back and forth within a single period or *wave crest* of the wakefield, a process called *synchrotron oscillation*, and is therefore called *trapped*. The points $\zeta = -\pi/2 \pm L_C$, $p = \gamma_g v_g$ are called the *turning points* of the oscillation, since one has $d\zeta/dt = \partial H/\partial p = 0$ there. A special case is that of a particle positioned precisely at the O-point, since such a particle does not move at all with respect to the wakefield. Trapped particles can be accelerated by the wake in a controlled way. For this reason, when injecting particles into a given wakefield, the aim is to inject as many particles as possible into the regions of closed orbits in phase space, i.e. the regions where they will get trapped.

In case $C = 1/\gamma_g + e_0$, the curves $(\zeta, p_-(\zeta))$ and $(\zeta, p_+(\zeta))$ meet at the points $\zeta = \pi/2 + n \cdot 2\pi$, i.e. the X-points, but are distinct for all other ζ . These curves are called the lower and upper *separatrix*, since they separate the regions in phase space containing untrapped or detrapped particles from the regions containing trapped particles. The solid curves in Fig. 3.1 denote these separatrices.

Acceleration of a single particle in a wakefield works as follows. The aim is to inject a particle on a closed orbit in phase space, on a point with low p . After some time, the particle will have progressed to a point on the orbit with high p , and then the particle needs to be extracted before it has a chance to reach parts of the orbit at which p is low again. In other words, a particle should be injected near point L in Fig. 3.1, and extracted near point H. Usually, it is not beneficial to have a particle perform multiple synchrotron oscillations, so it should be extracted already after half an oscillation. It is also important to inject a particle on the right orbit. Orbits close to the O-point only have a small difference between lowest and highest momentum, and a particle on such an orbit will not gain much energy. Orbits that are close to the separatrices can provide a large energy gain, but a particle on such an orbit runs the risk to end up on the wrong side of the separatrix, i.e. getting un- or detrapped. This may be caused by beam loading (see below) or other effects not included in the Hamiltonian system (3.3)-(3.5).

The maximum possible energy gain along a given closed orbit can be calculated as follows. For fixed C , the functions $p_+(\zeta)$ and $p_-(\zeta)$ reach their extrema for $\cos(\zeta) = 0$, i.e. at $\zeta = -\pi/2 + n \cdot 2\pi$. At such ζ , a maximum for p_+ and a minimum of p_- occur simultaneously, and taking into account that the points $(-\pi/2, p_-)$ and $(-\pi/2, p_+)$ lie on the same orbit $H = C$, we obtain for the energy gain \mathcal{E} :

$$\mathcal{E} = \gamma_+ - \gamma_- = v_g(p_+ - p_-) = 2\gamma_g^2 v_g \sqrt{(C + e_0)^2 - 1/\gamma_g^2}.$$

We observe that the energy gain is zero for $C = 1/\gamma_g - e_0$ (particle on the O-point), while it is maximum for $C = 1/\gamma_g + e_0$, i.e. the particle's orbit is infinitesimally close to the separatrices. For $\gamma_g e_0 \gg 1$, we obtain for the maximum possible energy gain: $\Delta\gamma \approx 4\gamma_g^2 e_0$.

A closer look at the dynamics of a trapped particle reveals that if $\gamma < \gamma_g$ at injection, the time needed to reach the first turning point at $\gamma = \gamma_g$ is much shorter than the time needed to reach its maximum energy at the top of its orbit. This means that $\gamma \gg \gamma_g$ along most of the particle's trajectory. The derivations in this and the following section will be done for that part of the particle orbit for which $\gamma \gg \gamma_g$ holds.

For a particle injected at an energy $\gamma_{in} \gg \gamma_g$ and $\zeta = \zeta_{in}$, and extracted at $\zeta = \zeta_{ex}$, the energy gain is calculated from (3.5) and (3.6) to obtain [11]:

$$\mathcal{E} = \gamma_{ex} - \gamma_{in} = v_g(p_{ex} - p_{in}) + e_0(\sin \zeta_{in} - \sin \zeta_{ex}) \approx 2\gamma_g^2 e_0(\sin \zeta_{in} - \sin \zeta_{ex}). \quad (3.7)$$

The acceleration distance L_a can be estimated from the injection and extraction phases as follows. For $\gamma \gg \gamma_g \gg 1$, we have $v - v_g \approx 1/(2\gamma_g^2)$, so

$$L_a = \frac{\zeta_{ex} - \zeta_{in}}{v - v_g} \approx 2\gamma_g^2(\zeta_{ex} - \zeta_{in}). \quad (3.8)$$

The distance necessary to obtain $\zeta_{ex} - \zeta_{in} = \pi$, so the particle is completely out of phase with the wave, is called the *dephasing length* and is given by

$$L_d = 2\pi\gamma_g^2.$$

As stated in the introduction, the goal of this work is to minimize the energy spread of a bunch of electrons after acceleration, while keeping the energy gain at a reasonable level. The energy spread can be estimated as the sum of two contributions. The first one results from the finite length of the bunch, causing different particles to have different injection and extraction phases, and thus different energy gains. The second one is caused by the fact that leading particles modify the wakefield as seen by the trailing particles, a topic that will be treated in the next section.

In order to estimate the length-dependent contribution to the energy spread, a simple two-particle model is used [8]. At $t = 0$ a bunch with uniform density and length $L < \pi$ is located between $\zeta = \zeta_{in} \pm L/2$. During the acceleration process, the distance between the electrons is assumed not to change significantly, so that, at extraction, the bunch is located between $\zeta = \zeta_{ex} \pm L/2$. From (3.7) the difference in energy gain between a particle at the head of the bunch and a particle at the tail of the bunch reads:

$$\Delta\mathcal{E} = 4\gamma_g^2 e_0 \sin(L/2)(\cos \zeta_{in} - \cos \zeta_{ex}). \quad (3.9)$$

From this equation, it is obvious that the use of short bunches reduces this contribution to the energy spread. Observe that this contribution may have either positive or negative sign, depending on ζ_{in} and ζ_{ex} . This can be explained as follows. At the beginning of the acceleration, i.e. ζ close to the X-point at $\pi/2$, the head of the bunch is accelerated more than the tail, while head and tail are accelerated equally halfway, and the tail is accelerated more than the head near the end of the acceleration, i.e. ζ close to $3\pi/2$. So depending on the injection and extraction phases, the compensation for the positive energy difference obtained at the beginning may be too little, just right, or too much. As we shall see in the next section, the contribution by the beam loading always has positive sign, since the close-range particle interaction is always repulsive. This means that by proper tuning, both contributions can be made to cancel each other, leading to a reduced energy spread.

3.3 Beam loading

As discussed above, a charged particle in a plasma wakefield provides its own contribution to the wakefield and therefore influences all particles that follow it. In this section, the influence of a bunch of electrons with bunch density n_b will be studied. The wakefield equation (3.1) is extended to include $n_b(\zeta)$ as a source term to read

$$\left(\frac{\partial^2}{\partial \zeta^2} + 1\right) \delta\Psi = n_b(\zeta).$$

The wake electric field caused by this bunch is given by

$$E_b(\zeta) = -\frac{\partial \delta\Psi}{\partial \zeta} = \int_{\zeta}^{\infty} n_b(\zeta') \cos(\zeta' - \zeta) d\zeta'. \quad (3.10)$$

For a short symmetric bunch located in the region $\zeta_b - L/2 < \zeta < \zeta_b + L/2$, the wakefield behind the bunch is given by

$$\begin{aligned} E_b(\zeta) &= \cos(\zeta - \zeta_b) \int_{-L/2}^{L/2} n_b(\zeta_b + \zeta') \cos \zeta' d\zeta' \\ &\approx \cos(\zeta - \zeta_b) \int_{-L/2}^{L/2} n_b(\zeta_b + \zeta') d\zeta', \end{aligned}$$

where the approximation is valid if $L \ll \pi$ only. The total charge Q of the bunch is defined as

$$-Q = \int_{-L/2}^{L/2} n_b(\zeta_b + \zeta') d\zeta'.$$

The *beam loading factor* η is defined as the ratio of the amplitudes of the laser wakefield and the bunch wakefield, i.e. $\eta = -Q/e_0$ for a short bunch. Then

the bunch wakefield may cancel out the laser wakefield as η approaches unity, which sets a natural upper limit on the beam loading. A beam loading larger than 1 would mean that the wakefield is dominated by the fields of the bunch rather than by those of the laser. Therefore, only $\eta < 1$ will be considered in this chapter.

For a uniform bunch located between $\zeta = \zeta_{in} \pm L/2$, the bunch wakefield can be calculated from (3.10) to obtain [8]:

$$E(\zeta) = \eta e_0 \times \begin{cases} 0, & \zeta - \zeta_{in} > L/2, \\ -\frac{\sin(\zeta - \zeta_{in} - L/2)}{2 \sin(L/2)}, & -L/2 < \zeta - \zeta_{in} < L/2, \\ \cos(\zeta - \zeta_{in}), & \zeta - \zeta_{in} < -L/2, \end{cases} \quad (3.11)$$

with beam loading factor

$$\eta = 2 \sin(L/2) \frac{n_b}{e_0}. \quad (3.12)$$

In order to calculate the average energy gain and spread of the electrons in the bunch, it is more convenient to switch to coordinates $s = z - t$ and t , since the bunch electrons are approximately static in these coordinates, so the beam loading forces will be constant during the acceleration. Then we have $\zeta = s + (1 - v_g)t \approx s + t/(2\gamma_g^2)$. From (3.7) and (3.11), the energy gain of an infinitesimal slice of the bunch of width ds is given by

$$d\mathcal{E}(s, t) = \{2\gamma_g^2 e_0 [\sin(s) - \sin(s + t/(2\gamma_g^2))] - t \cdot n_b \sin(s_{in} + L/2 - s)\} ds, \quad (3.13)$$

where $s_{in} = \zeta_{in}$, since we inject at $t = 0$. The average energy gain for the bunch is given by

$$\begin{aligned} \mathcal{E}_{av}(L_a) &= \frac{1}{L} \int_{s_{in}-L/2}^{s_{in}+L/2} d\mathcal{E}(s, L_a) \\ &= 2\gamma_g^2 e_0 \frac{\sin(L/2)}{L/2} \left[\sin(s_{in}) - \sin(s_{in} + \pi L_a/L_d) - \frac{\pi L_a}{L_d} \frac{\eta}{2} \right], \end{aligned}$$

where L_a is given by (3.8) and η by (3.12). It is obvious that both the finite bunch length and the beam loading contribute to a reduction of \mathcal{E}_{av} . For both $L = 0$ and $\eta = 0$, the single particle result (3.7) is returned.

The contribution to the energy spread induced by beam loading is a consequence of the constant retarding force on an electron at the tail of the bunch due to the bunch's wakefield. This induces an energy difference between the head and the tail of the bunch. From (3.13), we find for the energy difference between the head

and the tail of the bunch:

$$\begin{aligned}\Delta\mathcal{E} &= 4\gamma_g^2 e_0 \sin(L/2)(\cos(s_{in}) - \cos(s_{in} + \pi L_a/L_d)) + L_a n_b \sin(L) \\ &= 2\gamma_g^2 e_0 [2 \sin(L/2)(\cos \zeta_{in} - \cos \zeta_{ex}) + \eta \cos(L/2)(\zeta_{ex} - \zeta_{in})].\end{aligned}$$

Note that the first contribution to $\Delta\mathcal{E}$ matches the one given in (3.9). Under certain circumstances, the contributions of finite bunch length and beam loading to $\Delta\mathcal{E}$ can compensate each other, leading to a minimized energy difference between the head and the tail of the bunch. The acceleration distance at which this minimum in energy spread occurs is denoted by L_{opt} . Using Eq. (3.8) and taking ζ_{in} , ζ_{ex} close to π , i.e. halfway the X-point at $\pi/2$ and the O-point at $3\pi/2$, the distance at which $\Delta\mathcal{E} = 0$ can be estimated as

$$\frac{\pi L_{opt}}{L_d} = \frac{\eta}{\tan(L/2)} + 2(\pi - \zeta_{in}). \quad (3.14)$$

As we will see, Eq. (3.14) gives a correct scaling for L_{opt} in terms of L , η and ζ_{in} , apart from a constant offset.

3.4 Numerical analysis

The evolution of the energy spread of a bunch of macroparticles (virtual particles that each represent a large number of electrons) in a linear wakefield as given by (3.2) has been studied in 1-D numerical simulations. The behaviour of both energy gain and energy spread has been investigated as a function of injection phase, acceleration length, bunch length, and beam loading fraction.

In the simulations, the electron bunch is represented by N macroparticles, each having charge $-Q$, mass M , momentum p_n and phase ζ_n ($n = 1 \dots N$). Mass and charge are chosen such that $Q/M = e/m_e$ (electron charge-to-mass ratio). One should think of these one-dimensional particles as charged sheets with uniform electron density in the transverse direction and infinitesimally thin in the longitudinal direction (wide-beam limit). The wakefield of the n -th macroparticle is then given by

$$E_n(\zeta) = \begin{cases} -\frac{\alpha}{N} e_0 \cos(\zeta - \zeta_n), & \zeta < \zeta_n, \\ -\frac{\alpha}{2N} e_0, & \zeta = \zeta_n, \\ 0, & \zeta > \zeta_n. \end{cases} \quad (3.15)$$

The number α is a measure for the charge of a macroparticle, related to the beam loading fraction as

$$\eta = \frac{\sin(L/2)}{L/2} \alpha.$$

With this definition, taking the limit $N \rightarrow \infty$ in Eq. (3.15) correctly reproduces the wakefield of the continuous bunch, as given in (3.11). At the phases of the macroparticles, the electric field is non-continuous (see Fig. 3.2), because the macroparticle is taken to be infinitely thin. The choice of $E_n(\zeta_n) = -\alpha e_0/(2N)$ is motivated by the fact that a macroparticle should represent a cluster of plasma electrons, which obviously also suffers from beam loading induced by its own charge, regardless of how short it is. This field is called the *self field* of the macroparticle.

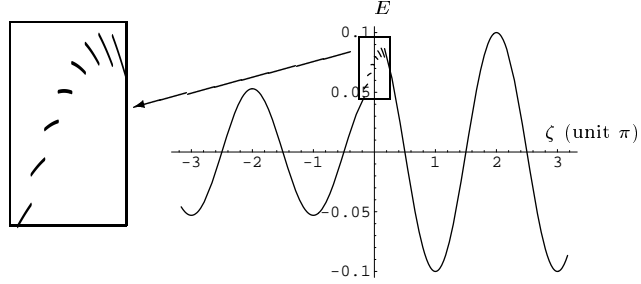


Figure 3.2: Example of electric field as a function of phase with only 7 macroparticles in the bunch, visible in the inset as discontinuities in E . Beam loading fraction: $\eta = 0.5$.

From (3.3)-(3.4), the equations of motion for the n -th macroparticle, $n = 1 \dots N$, are:

$$\frac{d\zeta_n}{dt} = \frac{p_n}{\sqrt{1 + p_n^2}} - v_g, \quad (3.16)$$

$$\frac{dp_n}{dt} = -e_0 \cos \zeta_n - \sum_{m=1}^N E_m(\zeta_n). \quad (3.17)$$

In the simulations, a leap-frog algorithm is used to integrate (3.16)-(3.17). For the electric field, the analytic expressions (3.2) and (3.15), are used, in which at every timestep all macroparticle phases are inserted. At the start of a simulation, all macroparticles have the same energy γ_0 and are uniformly distributed around the injection phase ζ_{in} :

$$\begin{aligned} \zeta_n &= \zeta_{in} + L \left(\frac{n-1}{N-1} - \frac{1}{2} \right), \\ p_n &= p_0 \equiv \sqrt{\gamma_0^2 - 1}. \end{aligned}$$

At every time step of the simulation, the mean energy $\langle \mathcal{E} \rangle$ and the energy spread $\Delta \mathcal{E}$ of the macroparticles are calculated. For the energy spread, the standard

deviation of the particle energy rather than the energy difference between the head and tail of the bunch has been used. The reason for this is that during the simulation, the head and tail of the bunch may have the same energy, while the middle of the bunch has a higher energy. This renders the difference between head and tail a poor measure of the actual energy spread of the bunch. The initial energy spread is zero, because all macroparticles have the same initial energy. The evolution of the bunch has been calculated until the first local minimum in the relative energy spread $\Delta\mathcal{E}/\langle\mathcal{E}\rangle$ is encountered. It has been verified that the first minimum is the lowest: continuing the simulation after the first minimum means that the particles in the bunch perform multiple synchrotron oscillations. Since the duration of a synchrotron oscillation is different for each particle, this implies that the particles will be spread across the wave crest while the bunch loses its coherence, resulting in a poor energy spread.

The following parameters are the same for all simulations presented in this chapter:

- $\gamma_g = 100$, i.e. the plasma is very underdense,
- $e_0 = 0.1$, i.e. the wakefield is linear,
- $\gamma_0 = \gamma_g/3$,
- Number of macroparticles: 100.

We have chosen the above set of parameters with a realistic case in mind: plasma electron density $\approx 10^{17} \text{ cm}^{-3}$, laser wavelength $1 \text{ } \mu\text{m}$, wake wavelength $100 \text{ } \mu\text{m}$, electrons of initially 17 MeV. The choice of $\gamma_0 = \gamma_g/3$ implies that the assumption $\gamma \gg \gamma_g$ is not applicable to the early stages of the acceleration. As a result, the simulated values for the energy gain and spread may deviate from the estimates made above. Also, the expression $L_a = 2\gamma_g^2(\zeta - \zeta_{in})$ for the acceleration length is no longer applicable. Instead, the acceleration length needs to be calculated from the (simulation) time t as $L_a = v_g \lambda_p t / (2\pi) \approx \lambda_p t / (2\pi)$. For a more elaborate discussion on the effects of injection at $\gamma_0 < \gamma_g$, see Andreev *et al.* [13], who used test electrons without beam loading effects, and Reitsma [14], who did include beam loading effects.

The results of a sample simulation are displayed in Fig. 3.3. The parameters for this simulation are: injection phase $\zeta_{in} = \pi$, bunch length $L = \pi/3$, and beam loading fraction $\eta = 0.8$. Graph (a) shows the mean energy $\langle\mathcal{E}\rangle$ in GeV, the relative energy spread $\Delta\mathcal{E}/\langle\mathcal{E}\rangle$, the standard deviation $\Delta\zeta$ of the phase, and the longitudinal root mean square (r.m.s.) emittance σ of the bunch, all as a function of the acceleration length L_a . Here, the r.m.s. emittance is defined as

$$\sigma^2 = \langle(\mathcal{E} - \langle\mathcal{E}\rangle)^2\rangle\langle(\zeta - \langle\zeta\rangle)^2\rangle - \langle(\mathcal{E} - \langle\mathcal{E}\rangle)(\zeta - \langle\zeta\rangle)\rangle^2.$$

The parameters have been chosen such that a minimum of $\Delta\mathcal{E}/\langle\mathcal{E}\rangle = 0.02$, a minimum of $\sigma = 0.03$, and a maximum of $\langle\mathcal{E}\rangle = 640$ MeV all coincide at an acceleration length of $d \approx 45$ cm. This result is a good representation of the aim of this work: tuning the bunch parameters such that a large energy gain and a minimum energy spread are obtained simultaneously.

In the course of the simulation, some macroparticles may lose so much velocity due to beam loading forces that they become untrapped. These are mostly macroparticles in the tail of the bunch having a phase too close to $\pi/2$, the edge of the wave crest. After acceleration, these particles will lag so far behind the bunch that they are no longer regarded as being part of the bunch. For this reason, we wish to exclude any untrapped particles from the simulation. In practice, this means that only those macroparticles that are not more than one wavelength behind the crest in which the bunch was originally injected are taken into account. As soon as a particle violates this condition, it is discarded from the simulation. The fraction of discarded particles should not become too high, but a small fraction may be sacrificed in order to improve energy gain and/or spread. The conditions of the simulations presented here are such, that only a small fraction, i.e. less than 5%, of the particles is lost on the way.

In the next section, the results of our simulations will be presented. The minimum relative energy spread and the acceleration length at which this minimum occurs will be determined, as well as the energy gain corresponding to the minimum energy spread. A parameter window will be determined for which an energy gain of hundreds of MeV coincides with a relative energy spread of only a few percent. The sensitivity of the energy spread to errors in the various bunch parameters will also be discussed.

3.5 Simulation results

In the simulations, the mean energy $\langle\mathcal{E}\rangle$ and the relative energy spread $\Delta\mathcal{E}/\langle\mathcal{E}\rangle$ have been calculated as functions of the injection phase ζ_{in} , the beam loading fraction η and the bunch length L . We have performed three series of simulations; in each series one of the input parameters is kept fixed, while the other two are varied. This has resulted in the identification of an optimal combination of parameters: $\zeta_{in} = \pi$, $\eta = 0.8$, and $L = \pi/3$ i.e. $17 \mu\text{m}$ for $\lambda_p = 100 \mu\text{m}$. This case has been studied in more detail.

It is instructive to start with the optimal case. In Fig. 3.3(a), the mean energy $\langle\mathcal{E}\rangle$ and the relative energy spread $\Delta\mathcal{E}/\langle\mathcal{E}\rangle$ are displayed for this case, as well as the standard deviation of the phase $\Delta\zeta$ and the r.m.s emittance σ , all as a function of the acceleration length L_a . Snapshots of particle phase space at various acceleration lengths for the same simulation are displayed in Fig. 3.3(b). The mean energy can be seen to increase as a function of L_a until it saturates

at $L_a \approx 48$ cm. Given a dephasing length of $L_d = \gamma_g^2 \lambda_p \approx 100$ cm, we find that saturation already sets in at less than half the dephasing length. This is caused by the fact that the dephasing length has been calculated for electrons on orbits that are very close to the separatrix in phase space (cf. Fig. 3.1), whereas electrons on orbits closer to the O-point dephase much quicker and start losing energy after that. Given the bunch length of $\pi/3$, it is inevitable that the head of the bunch lies on an orbit much closer to the O-point than the tail. As can be seen from the phase space snapshots, the head is already past $\zeta = 3\pi/2$ and therefore losing energy while the tail is still at $\zeta = \pi$ and thus a long way from dephasing. Hence the early saturation of the mean energy.

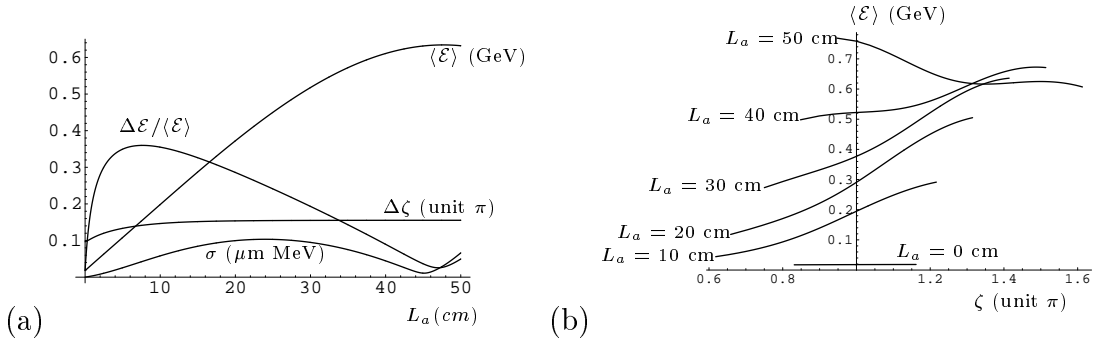


Figure 3.3: (a) Mean energy $\langle \mathcal{E} \rangle$, relative energy spread $\Delta \mathcal{E} / \langle \mathcal{E} \rangle$, phase spread $\Delta \zeta$ and longitudinal r.m.s. emittance σ as functions of acceleration distance L_a . Injection phase $\zeta_{in} = \pi$, bunch length $L = \lambda/6$, beam loading fraction $\eta = 0.8$. (b) Phase space snapshots for various acceleration distances for the same simulation.

The relative energy spread increases quickly in the early stages of the acceleration because, as can be seen from the phase space graph, the head of the bunch is already accelerated while the tail is still busy slipping backward in the wakefield without gaining energy. When the tail also starts gaining energy, the absolute energy difference between head and tail saturates, while the relative energy spread decreases because of the increase in mean energy. For L_a between 40 and 50 cm, the head and tail of the bunch have comparable energies, and a minimum relative energy spread of only 2.6% is reached at $L_a = 48$ cm. Beyond that point, the tail continues to gain energy while the head starts losing energy. Since the mean energy saturates at this length, this causes the relative energy spread to increase again. Therefore, the simulation has been stopped at 50 cm acceleration length.

From both the plot of $\Delta \zeta$ versus L_a and the phase space plots, it can be found that the bunch length increases in the early stages of the acceleration, but stays more or less constant later. Not surprisingly, this coincides with $\gamma \gg \gamma_g$ for all particles in the bunch. This justifies our earlier assumption that the bunch length does not change if the electron energy is much higher than γ_g . As for the r.m.s.

emittance σ , we find that it does not reach very large values at any moment. This is caused by the absence of momentum spread at injection, which causes the bunch to be line-like in phase space. To see a more realistic behaviour of σ , a finite momentum spread should have been used at injection.

For the same values of beam loading and bunch length, i.e. $\eta = 0.8$ and $L = \pi/3$, the relation (3.14) between the optimal acceleration distance L_{opt} and the injection phase ζ_{in} has been verified. The result is displayed in Fig. 3.4. The lower curve (straight line) represents the analytical prediction, while the upper curve denotes L_{opt} as taken from the simulations. It can be seen that for $0.7 < \zeta_{in}/\pi < 1.1$, the analytical prediction of L_{opt} is correct apart from a constant offset. However, for larger ζ , the predictions tend to be too low, while for $\zeta > 1.22 \cdot \pi$, Eq. (3.14) predicts a negative value for L_{opt} , which is obviously invalid. However, the general trend that L_{opt} decreases with increasing ζ_{in} is reproduced very well. This trend stems from the fact that the larger ζ_{in} , the closer to the O-point the bunch is injected and the shorter the time needed to reach dephasing. As we have seen above, the onset of dephasing usually defines the moment of minimum energy spread, and thus L_{opt} . The offset between the analytical prediction and the simulation results originates from the fact that the prediction does not take the first stage of the acceleration into account, during which the condition $\gamma \gg \gamma_\phi$ is not satisfied. The offset as found in Fig. 3.4 more or less agrees with the additional time needed to pass through this stage. For $\zeta_{in} > 1.1 \cdot \pi$, L_{opt} does not decrease to zero but saturates to some small but non-zero value. This region corresponds to bunches injected close to the O-point, which dephase quickly and gain little energy during their acceleration. The predicted value of L_{opt} is grossly off here as ζ_{in} , ζ_{ex} are no longer close to π , but this region is not interesting for acceleration anyway.

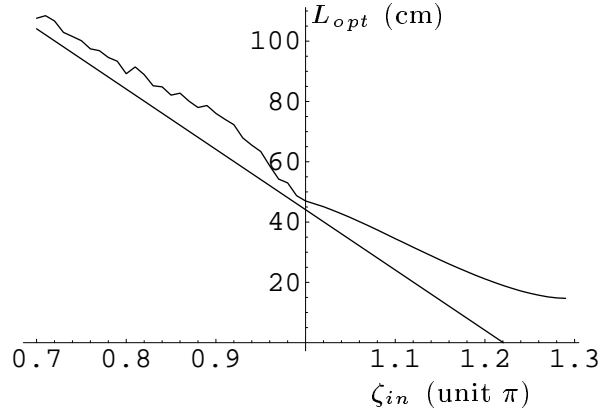


Figure 3.4: Optimal distance L_{opt} as a function of injection phase ζ_{in} . Upper curve: simulation result, lower curve: solution of Eq. (3.14).

We proceed to the series of simulations in which one of the parameters ζ_{in} , η ,

or L is kept fixed, while the other two are varied. For each combination of parameters, the simulation has been continued until the first minimum in $\Delta\mathcal{E}/\langle\mathcal{E}\rangle$ was encountered; at that instant both $\Delta\mathcal{E}/\langle\mathcal{E}\rangle$ and $\langle\mathcal{E}\rangle$ have been determined. It should be noted that this procedure may be difficult to follow in experiments, since guiding a laser pulse over a distance of say 50 cm and sustaining a wakefield of reasonable amplitude and structure at the same time is a formidable task. Therefore, the acceleration length in laser wakefield acceleration experiments is usually much shorter than L_{opt} , and determined on different grounds.

In the first series, the bunch length L is fixed at $\pi/3$. Figure 3.5 displays the relative energy spread and mean energy respectively, as functions of ζ_{in} and η . In Fig. 3.5(a), the region to the left of line A displays a large relative energy spread, which is more due to the small energy gain in this region than due to a large absolute energy spread, as is confirmed by Fig. 3.5(b). This region corresponds to injection close to the O-point, and is therefore not suitable for acceleration. The region between line A and line B displays a reasonably small energy spread, while the mean energy increases towards line B. The neighbourhood of line B looks especially suited for acceleration, since a small energy spread is accompanied by a large energy gain, with the comparatively large beam loading of at least 0.5 as an added bonus. In this region, it has been found that the acceleration lengths corresponding to maximum energy gain and minimum spread coincide. For the region to the right of line B, the mean energy is already beyond its optimum before a minimum in spread is reached. As a result, the acceleration length increases without any significant improvement in either energy gain or spread. This renders this region less suitable for acceleration than the neighbourhood of line B.

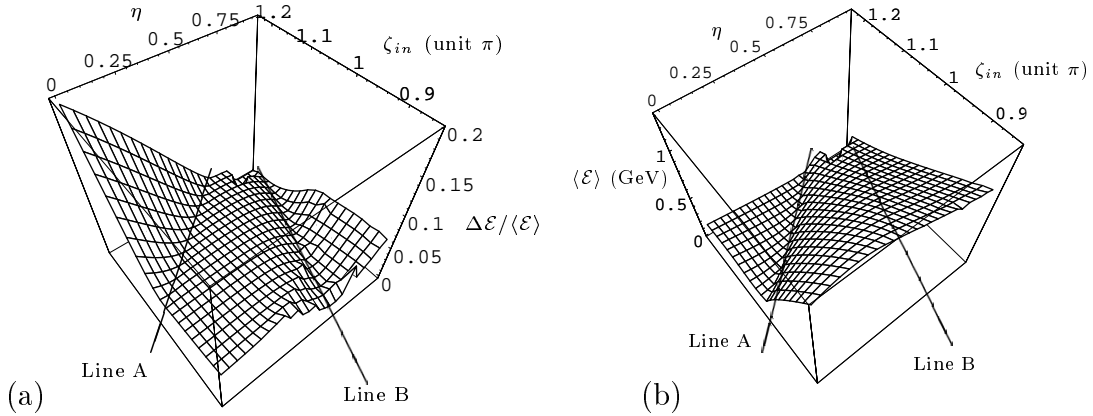


Figure 3.5: (a) Relative energy spread and (b) mean energy, both as a function of beam loading fraction and injection phase at $L = \pi/3$.

In the second series of simulations, the beam loading η is fixed at 0.2. In Fig. 3.6 the relative energy spread and mean energy are plotted as functions of bunch

length and injection phase. The region of $\zeta_{in} > \pi$ corresponds again to injection close to the O-point, and displays a small energy gain and consequently a large relative energy spread. The region of $\zeta_{in} < \pi$ corresponds to injection close to an X-point, which results in a large energy gain. However, injection too close to the X-point results in particle loss, as found from the behaviour at the front corner of Fig. 3.6(a). This loss saturates the energy gain, whereas the decrease in the energy spread results from the fact that precisely those particles that contribute most to the spread have been discarded. The energy spread is seen to increase for $L > 0.1 \cdot \pi$, as could be expected. However, the lowest energy spread is not reached for $L = 0$, but for $L \approx 0.1 \cdot \pi$. This is due to the fact that in the early stages of the acceleration, very short bunches tend to explode while slightly longer bunches are compressed, so the shortest bunches display a comparatively large energy spread. In practice, bunches are usually not that short ($0.1 \cdot \pi \simeq 5 \mu\text{m}$), but Fig. 3.6(a) shows that a low energy spread can be maintained up to $L = 0.4 \cdot \pi \simeq 20 \mu\text{m}$, at the expense of a decreasing range of suitable injection phases. In general, it is concluded that for $\zeta_{in} \lesssim \pi$ and $L < 0.4 \cdot \pi$, a decent energy gain can be obtained while the energy spread is kept at a reasonable level. For the given plasma density, this corresponds to a bunch length of about $20 \mu\text{m}$, or 70 fs, which is problematic if the bunch is to be obtained from an external source.

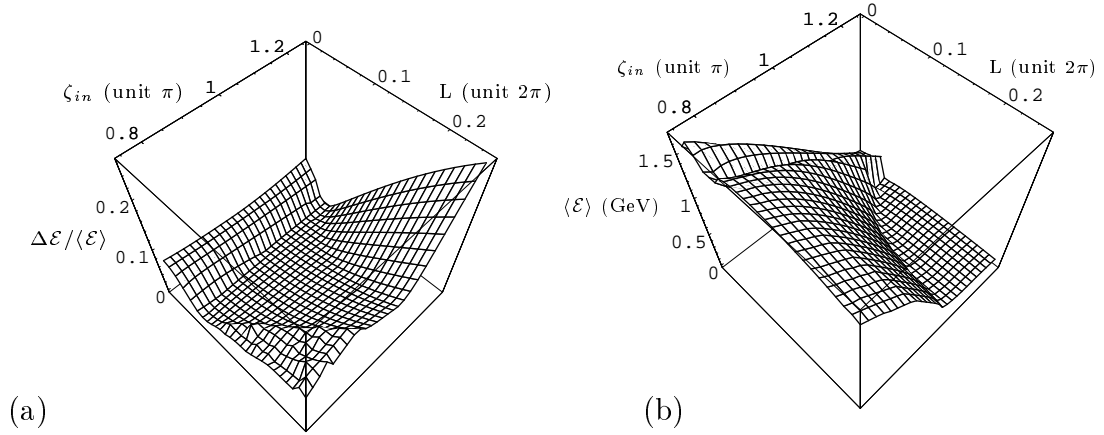


Figure 3.6: (a) Relative energy spread and (b) mean energy, both as a function of bunch length and injection phase at $\eta = 0.2$.

In the third series of simulations, the effect of changing both bunch length and beam loading at a fixed injection phase $\zeta_{in} = \pi$ is considered. This corresponds to injection halfway the O-point and the X-point. Figure 3.7 shows the relative energy spread and mean energy as functions of beam loading fraction and bunch length. Figure 3.7(a) demonstrates that a low energy spread can be reached for nearly every value of η between 0 and 1, as long as one does not attempt to put too much charge in a short bunch. If a bunch is overloaded, the tail of the bunch

will hardly feel any accelerating field due to beam loading by the bunch's head. Then the tail will be accelerated much less than the head, resulting in a large energy spread. Also, the assumption that the bunch length does not change will not be valid in this case, and as a consequence the value of L_{opt} in the simulations does not match the prediction of Eq. (3.14). For the regime corresponding to the front part of the graph, i.e. for $L > 0.4\pi \cdot \eta$, L_{opt} has been found to be in agreement with Eq. (3.14). Regarding the energy gain, Fig. 3.7(b) shows that for sufficiently large L (in order to keep the energy spread under control), the mean energy reached at L_{opt} actually increases with beam loading. This is caused by the following mechanism. If a bunch were to be accelerated without the effect of beam loading on the wakefield, its tail would overtake its head during the later stages of the acceleration, resulting in a large energy spread. This would require the acceleration to be stopped fairly soon in order to keep the energy spread low, at the expense of the energy gain. Because of beam loading however, a backward force is exerted on the tail of the bunch, preventing it from overtaking the head and thus postponing the optimal moment for extraction. Of course, the increased acceleration length leads to an increased mean energy while the energy spread remains low.

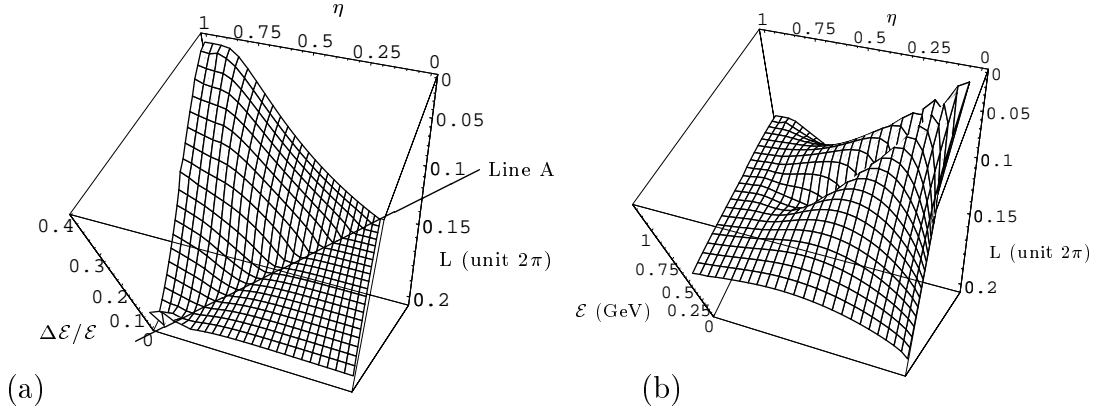


Figure 3.7: (a) Relative energy spread and (b) mean energy, both as a function of beam loading fraction and bunch length at $\zeta_{in} = \pi$.

In general, the bunch parameters should obey the following conditions:

- Injection phase closer to an X-point than to an O-point: $\zeta_{in} \lesssim \pi$,
- Bunch length not too small for the given beam loading, i.e. $L > 0.4\pi \cdot \eta$, but not too large either, $L < 0.4 \cdot \pi$.

The special case presented at the beginning of this section, $\zeta_{in} = \pi$, $L = \pi/3$, $\eta = 0.8$, obeys these conditions well, and thus displays both a large energy gain together with a small energy spread. However, this case may be difficult to

realize experimentally, since the requirements on both bunch length and optimal acceleration distance are very hard to meet. A more detailed discussion on these and other requirements on the experimental parameters will be presented in the next section.

3.6 Accuracy requirements

In the previous section, it has been shown that for the parameter combination $\zeta_{in} = \pi$, $L = \pi/3$, and $\eta = 0.8$, an electron bunch can be accelerated to high energies, while its energy spread is kept low. The practical value of this result, however, can only be determined if its sensitivity to errors in the input parameters is known. To this end, we have performed simulations in which one of the parameters ζ_{in} , L , and η has been deviated from its optimal value, while the other two have been kept fixed. The (unscaled) acceleration length has been fixed at 48 cm for all simulations, which is the optimal length for $\zeta_{in} = \pi$, $L = \pi/3$, $\eta = 0.8$, and a plasma wavelength of $100 \mu\text{m}$. The reason to keep L_a fixed is that in experiments this is also a fixed length.

In Fig. 3.8, the sensitivity of energy spread (left) and mean energy (right) to errors in the injection phase is displayed. Already for $\zeta_{in} = \zeta_{opt} \pm 0.08 \cdot \pi$, the energy spread has increased from 2% to 12%, while the mean energy ranges between 450 and 800 MeV. To keep both the energy spread and gain within reasonable bounds, one should take $|\zeta_{in} - \zeta_{opt}| < 0.04 \cdot \pi$. For a $100 \mu\text{m}$ plasma wavelength, this corresponds to a maximum tolerance of $2 \mu\text{m}$, or about 7 fs, for the timing of the electron bunch with respect to the laser pulse. This requirement will be very hard to meet.

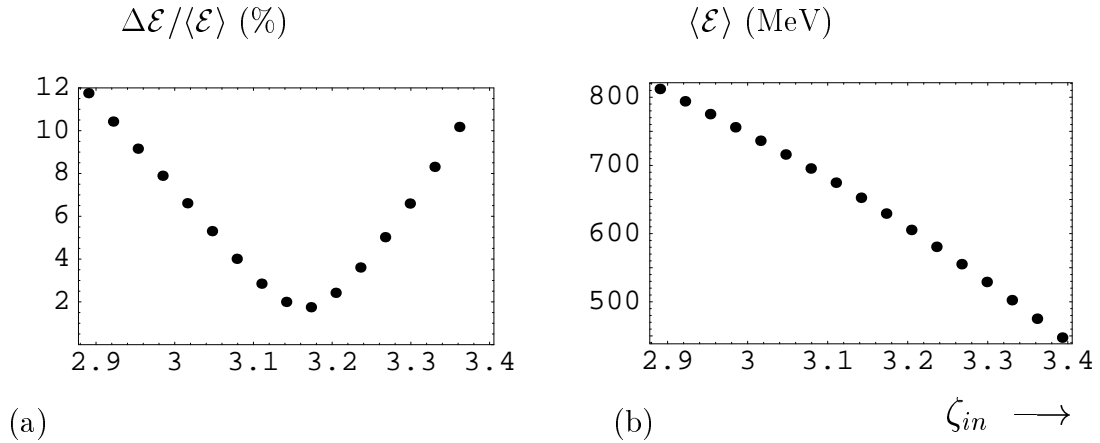


Figure 3.8: Effect of small errors in the injection phase ζ_{in} on (a) energy spread and (b) mean energy. Optimal value is $\zeta_{in} = \pi$. Both energy spread and gain are highly sensitive to errors in ζ_{in} .

Note that decreasing the plasma density in order to increase the plasma wavelength will not improve matters much, since this will also increase the necessary acceleration length, which is already hard to attain. As seen from Fig. 3.6(a), reducing the bunch length increases the window for ζ_{in} for which the energy spread remains reasonable. However, as stated above, producing a bunch with a duration of 70 fs is already difficult, leave alone reducing its length even further.

In Fig. 3.9, the sensitivity of energy spread (left) and mean energy (right) to errors in the bunch length is displayed. For $L = L_{opt} \pm 0.2$, i.e. an error of 20% in the bunch length, the energy spread increases from 2% to 10%, while the change in mean energy is insignificant. Considering the fact that $L_{opt} = \pi/3$, which corresponds to 17 μm or 55 fs for a 100 μm plasma wavelength, the real issue here is the production of very short bunches. Once this has been accomplished, a reasonable relative error in the bunch length can be tolerated.

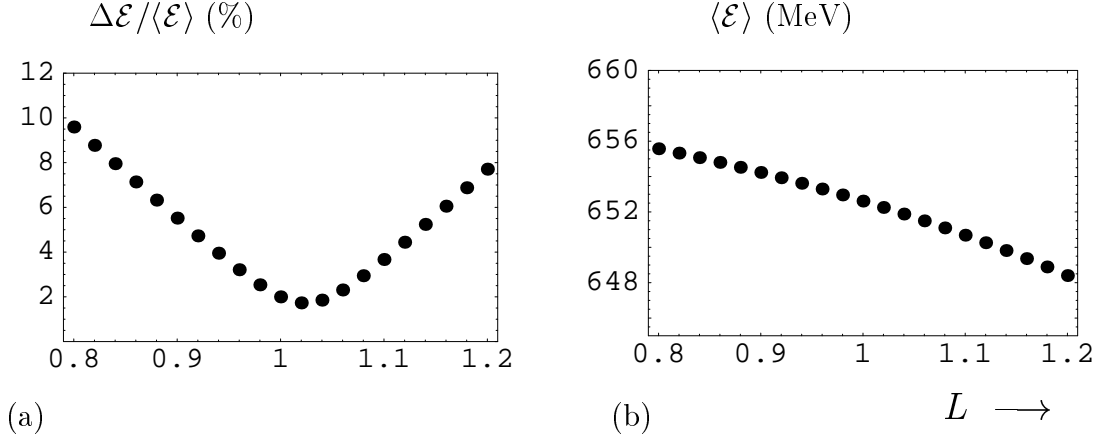


Figure 3.9: Effect of small errors in the bunch length L on (a) energy spread and (b) mean energy (right). Optimal value is $L = 1.04$. The energy spread is less sensitive to errors in L than to errors in the injection phase, and the energy gain is nearly unaffected.

In Fig. 3.10, the sensitivity of energy spread (left) and mean energy (right) to errors in the beam loading is displayed. For $0.7 < \eta < 0.85$, where $\eta_{opt} = 0.8$, there is hardly any change in the energy spread, while the energy gain decreases from 700 to 620 MeV for increasing η . The decrease in mean energy is a result of the self-retarding forces the bunch experiences due to beam loading. The anomalies seen for $\eta > 0.85$ are the result of putting too much charge in a short bunch. All in all, the beam loading can vary over a rather wide range of values, as long as its value is not taken too high, without significant adverse effects on energy spread or mean energy. Its value will usually be determined by the amount of charge one wishes to accelerate and the energy one wishes to reach.

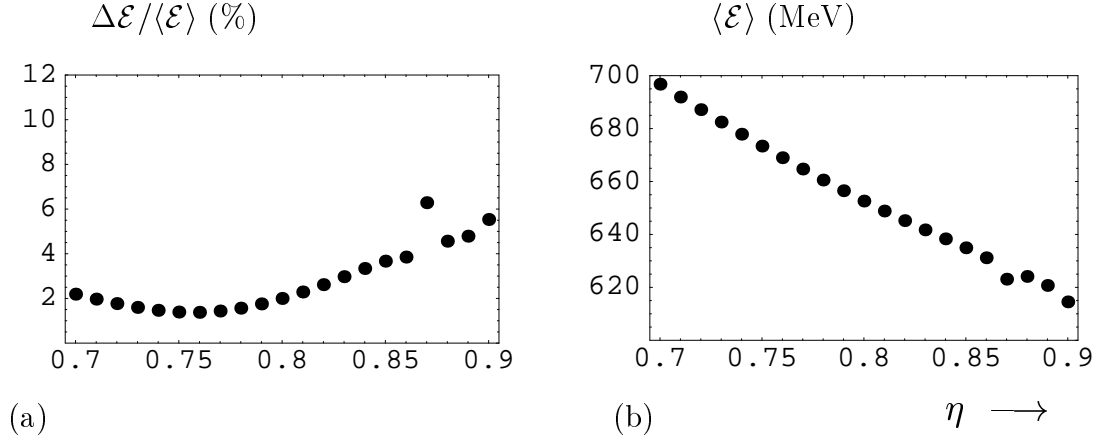


Figure 3.10: Effect of small errors in the beam loading fraction η on (a) energy spread and (b) mean energy. Optimal value is $\eta = 0.8$. The energy spread is nearly insensitive to errors in η , while the mean energy is affected to some degree. The anomalies for $\eta > 0.85$ are the result of putting too much charge in a short bunch.

It is concluded that the electron bunch needs to meet very strict requirements on its absolute length and the accuracy of its timing, while the requirements on the accuracy of length and beam loading are much more lenient. The situation is aggravated by the fact that the plasma wavelength should be much lower than the $100\ \mu\text{m}$ used in our simulations in order to bring down the acceleration length. This renders the external production and injection of suitable bunches a very demanding task. A way to get out of this predicament is to use electrons provided by the plasma itself for acceleration, instead of injecting them from the outside. This method of electron injection is called *internal injection*. Plasma electrons can be injected into a plasma wave by pushing that wave until wave breaking sets in. Wave breaking means that part of the electrons in the plasma obtain a longitudinal velocity larger than the phase velocity of the plasma wave, so they are actually injected into the wave. These electrons are injected in the form of a bunch much shorter than the plasma wavelength, located at the end of a plasma wave crest, i.e. near an X-point, so both the timing and the bunch length issues are solved simultaneously. Unfortunately, the electrons injected by wave breaking usually display a large initial energy spread, so the final energy spread may be hard to control in such schemes. Internal injection of electrons by means of plasma wave breaking will be investigated in Chapters 4 and 6.

3.7 Summary and conclusions

The dynamics of electron acceleration in a one-dimensional plasma wave has been studied, with emphasis on the minimization of the energy spread. The two

major mechanisms that contribute to the energy spread, i.e. finite bunch length and beam loading, have been shown to counterbalance each other at a certain acceleration distance L_{opt} . An analytical estimate for L_{opt} has been obtained in terms of beam loading η , bunch length L , and injection phase ζ_{in} . This estimate has been found to be in agreement with 1-D simulations. Using a particle code, we have demonstrated that a low energy spread and high energy gain can be achieved simultaneously if the parameters η , L and ζ_{in} are properly tuned. An optimum has been identified for $\zeta_{in} = \pi$, $L = \pi/3$, and $\eta = 0.8$. The sensitivity of this optimum to errors in the input parameters has been investigated. Here, it has been found that the optimum is very sensitive to errors in the injection phase, while the sensitivity to errors in the bunch length or beam loading factor is much lower. The most important requirements on the bunch are therefore that it be very short, i.e. below 100 fs, and that the timing of its injection is very demanding, i.e. within 10 fs. Since these requirements are hard to meet for externally injected bunches, internal injection of electrons may be needed. This method will be the topic of the following three chapters.

References

- [1] E. Esarey *et al.*, IEEE Trans. Plas. Sci. **24**, 252 (1996).
- [2] J.S. Wurtele, Phys. Fluids B **5**, 2363 (1993).
- [3] T. Tajima and J.M. Dawson, Phys. Rev. Lett. **43**, 267 (1979).
- [4] A. Modena *et al.*, Nature **377**, 606 (1995).
- [5] F. Dorchies *et al.*, Phys. Plasmas **6**, 2903 (1999).
- [6] S. Wilks *et al.*, IEEE Trans. Plas. Sci. **15**, 210 (1987).
- [7] S. van der Meer, CLIC Note No. 3 CERN/PS/85-86(AA), 1985.
- [8] T.C. Chiou and T. Katsouleas, Phys. Rev. Lett. **81**, 3411 (1998).
- [9] T.C. Chiou *et al.*, Phys. Plasmas **2**, 310 (1995).
- [10] T. Katsouleas *et al.*, AIP Conf. Proc. **395**, 75 (ed. S. Chattopadhyay, Arcidosso, 1996).
- [11] R.D. Ruth and A.W. Chao, AIP Conf. Proc. **91**, 94 (ed. P.J. Channell, Los Alamos, 1981).
- [12] C.B. Schroeder *et al.*, Phys. Rev. Lett. **82**, 1177 (1999).

- [13] N.E. Andreev, S.V. Kuznetsov, and I.V. Pogorelski, Phys. Rev. ST Accel. Beams **3**, 021301 (2000).
- [14] A.J.W. Reitsma, *Electron bunch quality in laser wakefield acceleration*, Ph. D. thesis, Technische Universiteit Eindhoven, The Netherlands, 2002.

Chapter 4

Fast electron generation through breaking of a plasma wave

A 1-D model for fast electron generation by a short, intense, non-evolving laser pulse propagating through an underdense plasma has been developed. This is done for pulses shorter than the plasma wavelength, so parametric instabilities can be neglected and plasma wave breaking is the dominant mechanism behind this process. First, an overview of wave breaking theory is given for both cold and warm plasmas. Next, plasma wave excitation by a laser pulse and breaking both in front of and behind the pulse are discussed. Simulations have been performed to determine the wave breaking conditions for several different pulse shapes. Fast electrons emerge as a short bunch, and the electrostatic field of this bunch is shown to limit self-consistently the amount of generated fast electrons.

Sections 4.1, 4.5-4.7 and 4.9-4.10 have been published as the article “Generation of fast electrons by breaking of a laser-induced plasma wave”, R.M.G.M. Trines, V.V. Goloviznin, L.P.J. Kamp, T.J. Schep, *Phys. Rev. E* **63**, 026406 (2001). These sections and Section 4.8 have also appeared in the Proceedings of the International Conference on Lasers 2000, Albuquerque NM, Editors V. J. Corcoran and T. A. Corcoran (STS Press, McLean VA, 2001).

4.1 Introduction

Fast particle generation, the production of multi-MeV electrons during the propagation of an intense (10^{19} W/cm²) laser pulse through an underdense plasma, is an aspect of laser-plasma interaction that has recently attracted a lot of attention, in both theoretical and experimental circles. The effect has been observed in many experiments, [1–6] as well as in numerical simulations [7,8].

At the heart of this phenomenon lies the trapping of plasma electrons by the

laser-induced wake wave. Once captured, these electrons are accelerated by the wakefield, just like in a laser wakefield accelerator. As for the mechanism behind such electron capture, one has to distinguish between the ponderomotive regime, in which the laser pulse is shorter than one plasma wavelength and there is little or no back-action of the plasma on the pulse [9,10], and the self-modulated regime, in which the pulse is (much) longer than one plasma wave period and the back-action of the plasma on the pulse is very important. In the ponderomotive regime, electron capture can be explained through breaking of the wake wave, occurring when the laser pulse intensity is strong enough to push the longitudinal electric field beyond the wave breaking threshold [2,11,12]. As we shall see in Chapter 6, wave breaking is also important in the self-modulated regime, although its role is entirely different and there are more waves involved than just the wakefield. Closer investigations of electron trapping due to wave breaking have been conducted by, among others, Bulanov et al. [11,12]. Electron capture in a plasma wave induced by a relativistic electron bunch has been studied by Schep et al. [13].

In this chapter, the behaviour of plasma waves driven by the ponderomotive force of a short ($L < \lambda_p$), intense laser pulse with a non-evolving envelope will be investigated. First, an overview of wave breaking theory will be given with emphasis on plasma waves, and wave breaking of cold and warm plasma waves in the quasi-static approximation will be discussed. Since both cold and warm plasma cases can be dealt with using the same methods, only the case of a relativistic wave in a cold plasma will be investigated in depth. For non-relativistic waves or warm plasmas, only the results will be given. Next, the excitation of a plasma wave by a non-evolving laser pulse will be investigated. Several scenarios will be explored that lead to the breaking of the plasma wave, and the necessary conditions will be derived for this to happen. These conditions have been compared to the results of simulations, in which the amplitude of the excited plasma wave has been calculated as a function of pulse intensity and shape. Finally, we show that the fast electrons emerge as a very short bunch, whose electric field will limit the amount of generated fast electrons in a self-consistent way. From this, we derive the maximal charge density of the bunch as a function of laser pulse intensity.

In order to simplify the equations, all quantities in the equations governing the laser plasma interaction will be scaled as in Section 2.1 with $\omega_r = \omega_p$: $t' = \omega_p t$, $z' = (\omega_p/c)z$, $v' = v/c$, $p' = p/(m_e c)$, $E' = eE/(m_e \omega_p c)$, $\Phi'_s = e\Phi_s/(m_e c^2)$, $A'_\perp = eA_\perp/(m_e c)$, $n' = n/n_0$. Here, $\omega_p^2 = e^2 n_0/(\varepsilon_0 m_e)$ denotes the electron plasma frequency, and other symbols have their usual meaning. For convenience, the primes will be dropped from this point on.

4.2 Wave breaking in one dimension

Wave breaking is a non-linearity occurring in the evolution of a wave that causes a wave front to steepen, until a singularity is reached. At that point, parts of the wave that started out from different wave phases collide with each other, causing the wave function to become multiple-valued. Wave breaking is a dynamical process in which an initially regular wave will develop a singularity as it propagates. The best-known form is wave breaking in the direction of propagation of the wave, which is basically a 1-dimensional phenomenon.

In the literature, there is no unambiguous definition for 1-D wave breaking in a plasma (or fluid). Different definitions apply to a cold and a warm plasma. In a cold plasma, there is a single well-defined plasma velocity for each phase $\zeta = z - vt$, and the wave's advection equation does not contain pressure or diffusion terms. In that case, wave breaking can be described in terms of the characteristic curves of the advection equation. Also, the singularity in the advection equation coincides with a singularity in the continuity equation. In the case of a warm plasma, the presence of pressure terms in the advection equation causes the singularities in this equation and the continuity equation not to coincide any more, while the presence of dispersion terms inhibits the use of characteristics.

The use of the quasi-static approximation, i.e. the wave is considered static with respect to some comoving frame of reference, again changes the definition of wave breaking. However, in this approximation the treatment is similar for cold and warm plasma.

4.2.1 Wave breaking in a cold plasma

For waves in a cold plasma (no pressure or diffusion terms), there are several ways to define 1-D wave breaking, all equivalent:

1. Intersection of characteristics of the wave equation,
2. Singularity of the transformation between Eulerian and Lagrangian coordinates,
3. Intersection of the trajectories of plasma fluid elements,
4. Occurrence of singularities in the plasma density.

Each will be discussed briefly in this section. An extensive treatment on the subject can be found in, among others, Whitham [14].

The behaviour of a 1-D cold plasma is governed by the following equations:

$$\frac{\partial p}{\partial t} + v(z, t) \frac{\partial p}{\partial z} = -E, \quad (4.1)$$

$$\frac{\partial E}{\partial z} = 1 - n, \quad (4.2)$$

$$\frac{\partial n}{\partial t} + \frac{\partial}{\partial z}(nv) = 0. \quad (4.3)$$

For the analytical description of a travelling wave, it often pays to express everything in so-called *Lagrangian coordinates* z_0 and τ . These coordinates are chosen such that they follow the wave, in contrast to the usual *Eulerian coordinates* z and t , also known as the laboratory coordinates. The Lagrangian coordinates for the wave described by the advection equation (4.1) are given by

$$t \equiv \tau, \quad z \equiv z_0 + \xi(z_0, \tau) = z_0 + \int_0^\tau d\tau' v(z_0, \tau'). \quad (4.4)$$

The space and time derivatives transform according to

$$\begin{aligned} \frac{\partial}{\partial z} &= \left[1 + \int_0^\tau d\tau' \frac{\partial}{\partial z_0} v(z_0, \tau') \right]^{-1} \frac{\partial}{\partial z_0}, \\ \frac{\partial}{\partial t} &= \frac{\partial}{\partial \tau} - v(z_0, \tau) \left[1 + \int_0^\tau d\tau' \frac{\partial}{\partial z_0} v(z_0, \tau') \right]^{-1} \frac{\partial}{\partial z_0}. \end{aligned}$$

As a result, the advection equation (4.1) reduces to

$$\frac{\partial p}{\partial \tau} = -E,$$

and it is obvious that the Lagrangian coordinates are comoving with the wave.

With these coordinates, we are well equipped to discuss the various definitions of wave breaking. We start with the first one: wave breaking occurs at the intersection of characteristics of the wave equation. This can best be understood from the advection equation (4.1). From the theory of first-order partial differential equations, the characteristic curves for this equation are given by

$$\frac{dz}{dt} = v(z, t), \quad \frac{dp}{dt} = -E. \quad (4.5)$$

Combining this with an initial condition $p(z_0, 0) = p_0(z_0)$, we find that the characteristic curves in (t, z, p) -space are given by

$$z(t) = z_0 + \int_0^t dt' v(z(t'), t'), \quad p(t) = p_0(z_0) - \int_0^t dt' E(z(t'), t'). \quad (4.6)$$

The expression for $z(t)$ is actually the inverse from (4.4). To state it differently, the transformation (4.4) is chosen such that the solution to (4.1) propagates along the curves $z_0 = C$, and from (4.6) it follows that the solution propagates along the characteristic curves. This means that the characteristic curves of (4.1) are equivalent to the curves $z_0 = C$. The transformation of Lagrangian coordinates (z_0, τ) to Euler coordinates (z, t) is therefore similar to tracing a characteristic from its starting point z_0 at $t = 0$ to a point $z(t)$ at some later time t .

Rewriting the above in terms of Lagrangian coordinates reveals that the first definition of wave breaking actually means that the curves (z_{01}, τ) and (z_{02}, τ) with $z_{01} \neq z_{02}$ map to the same point in Eulerian coordinates for some $\tau > 0$. This implies that the transformation from Lagrangian to Eulerian coordinates is no longer one-to-one, i.e. the second definition of wave breaking.

Singularity of the coordinate transformation can also be expressed in terms of its Jacobian $J(z_0, \tau)$, given by

$$J(z_0, \tau) = \det(\partial z / \partial z_0) = \det(1 + \partial \xi / \partial z_0).$$

Observe that if two intersecting characteristics have a finite difference in z_0 , all intermediate characteristics must have intersected before, so at the earliest possible time that wave breaking occurs, it concerns characteristics having an infinitesimal difference in z_0 . In other words, wave breaking occurs as soon as $\partial z / \partial z_0 = J(z_0, \tau)$ vanishes at some point.

To see that the third definition for wave breaking, i.e. the intersection of the trajectories of plasma fluid elements, is equivalent to the other two, one only needs to consider that the advection equation (4.1) also describes the trajectories of the fluid elements, so intersection of characteristics is equivalent to intersection of fluid element trajectories.

The last definition of wave breaking to be studied is the one equating wave breaking to the occurrence of singularities in the fluid density. Consider the continuity equation (4.3). In Lagrangian coordinates, this equation reads

$$\frac{\partial n(z_0, \tau)}{\partial \tau} + n(z_0, \tau) \frac{\partial z_0}{\partial z} \frac{\partial v(z_0, \tau)}{\partial z_0} = 0.$$

Provided $\partial z_0 / \partial z$ is nonzero, this leads to

$$\frac{\partial}{\partial \tau} \left(n(z_0, \tau) \frac{\partial z}{\partial z_0} \right) = 0, \quad n(z_0, \tau) = \frac{n(z_0, 0)}{\partial z / \partial z_0}.$$

This relation expresses that the amount of fluid particles in a fluid element is conserved along the flow of the element. The local fluid density is then determined by the change in volume of the fluid element, expressed by $\partial z_0 / \partial z$. We find that the density becomes singular if and only if $\partial z_0 / \partial z$ vanishes. So, a singularity in $n(z_0, \tau)$ occurs if and only if there is wave breaking.

Consider two coordinates z_{01} and z_{02} with $z_{01} < z_{02}$, and corresponding characteristics $z_1(t)$, $z_2(t)$, i.e. $z_{1,2}(0) = z_{01,2}$. Let us assume that $v(z_1(t), t) > v(z_2(t), t)$ for $t > 0$. Then we find that these characteristics intersect at $t = t_i$, where t_i is defined by

$$\int_0^{t_i} (v(z_1(t), t) - v(z_2(t), t)) dt = z_{02} - z_{01}.$$

From the definition of the characteristics, we know that each characteristic in (t, z, p) -space is embedded in the curved surface $(z, t, p(z, t))$. So at the intersection of two characteristics, we find that $p(z, t)$ becomes multiple-valued. The physical meaning of wave breaking is therefore, that two wavelets starting at different locations at $t = 0$ collide at some later time since the trailing wavelet has a larger local phase velocity than the leading one. At that point the solution may develop in two different ways. If the physical variable associated with the wave cannot take multiple values for a single (z, t) -pair, as in the case of the density of a fluid, a shock front (a sudden jump in an otherwise continuous solution) is formed, which travels with the wave. On the other hand, if the associated variable can be multiple-valued, as in the case of the plasma fluid momentum $p(z, t)$, the continuous curve representing the solution $p(z, t)$ at some fixed t will be “folded over” to represent a true multiple-valued solution. An extensive treatment of the behaviour of a cold plasma wave after breaking, involving a multistream approach, is given by Wang, Payne, and Nicholson [15].

Breaking of an unforced wave occurs if the local phase velocity increases with the value of the displacement, so a local maximum can overtake a local minimum.

Several mechanisms may prevent wave breaking from occurring. Diffusion for example dampens the short-wavelength components of the wave, thus inhibiting the formation of a steep front. Also, the concept of characteristics does not apply in the presence of diffusion, so wave breaking cannot be defined in terms of characteristics. The Burgers equation is an example of this. Dispersion usually leads to a broadening of wave packets, also preventing the steepening of wave fronts. This can be seen in the Korteweg-de Vries equation. Mechanisms like diffusion and dispersion infer the presence of higher order z -derivatives of p in the advection equation (4.1). In terms of characteristic curves, such mechanisms work by dulling any peaks in the “density” of characteristics in the (z, t) -plane, thus preventing intersection of characteristics.

A few remarks should be made concerning waves in a cold plasma. Consider a plasma wave having initial electron density $n(z_0, 0)$ and local fluid velocity $v(z_0, 0)$. Then the condition that the density be nonnegative initially requires only that $n(z_0, 0) \geq 0$ everywhere. However, as derived by Davidson [16], the requirement that the density remain nonnegative and finite for all time gives the

more restrictive conditions

$$n(z_0, 0) > \frac{1}{2}, \quad (4.7)$$

$$\left| \frac{\partial v(z_0, 0)}{\partial z_0} \right| < \sqrt{2n(z_0, 0) - 1}. \quad (4.8)$$

These will be used later when dealing with the quasi-static approximation.

4.2.2 Wave breaking in a warm plasma

Wave breaking in a warm plasma is considerably harder to describe. For one thing, the advection equation for a warm plasma contains pressure terms, and may also contain diffusion terms that inhibit the description of wave breaking in terms of characteristics. Also, the presence of pressure terms prevents the plasma density from reaching arbitrarily large values, since in the adiabatic approximation, an infinite amount of energy would be necessary to compress the plasma. Furthermore, electron-ion collisions which have been neglected so far, become important, leading to viscosity terms in the advection equation. Therefore, a definition different from those presented in the previous section is needed. To date, there are three different but partially overlapping definitions for “warm” wave breaking.

1. Wave breaking occurs when the plasma density reaches the Coffey limit [17]. The Coffey limit is an upper bound to the plasma density, beyond which thermal pressure inhibits further compression. As a result, plasma fluid elements will be reflected from the plasma wave if the density reaches the Coffey limit at some point, which is considered wave breaking. A more extensive treatment will be given later.
2. Wave breaking occurs when plasma electrons become trapped in the wave. This definition is often used together with waterbag models (see below) for the plasma electron thermal distribution. The onset of wavebreaking is then taken to be the moment that the upper waterbag surface gets trapped. However, several authors argue that there is no one-to-one correspondence between plasma electron trapping and wave breaking, especially if only a few very hot electrons are trapped while the structure of the plasma wave is preserved. See for example Bezzerides and Gitmore [18] or Bergmann and Mulser [19].
3. Wave breaking occurs when at least one of the macroscopically observable quantities, such as the density or the longitudinal fluid velocity, loses its periodicity. In a Vlasov approach, this is accompanied by trapping of

the “bulk” of the plasma electrons and loss of the periodic structure of the distribution function $f(t, x, v)$. This definition is treated in depth by Bergmann and Mulser [19]. Unfortunately, it is difficult to derive quantitative thresholds for wave breaking from it.

As it happens, most of the work on wave breaking makes use of the Coffey criterion to calculate field or density thresholds for wave breaking. Several authors (Katsouleas and Mori [20], Rosenzweig [26,27]) use particle trapping arguments to derive such thresholds. However, these thresholds can usually also be obtained by applying Coffey-type arguments to their analysis. The definition of Bergmann and Mulser is rarely seen in literature, even though the results of their Vlasov simulations [19] clearly show that neither the Coffey criterion nor particle trapping arguments give a complete description of the phenomenon of warm plasma wave breaking.

4.3 Wave breaking in two or three dimensions

Wave breaking in two or three dimensions is partly an extension of the same phenomenon in one dimension, but it also has aspects that are entirely alien to 1-D wave breaking. To understand how 2-D breaking can be an extension of 1-D breaking, we have to reconsider the meaning of characteristic curves first. A characteristic curve is actually a curve of constant wave phase in (z, t) -space, defined by the equation $z_0(z, t) = Z_0$ for some constant Z_0 . Wave breaking occurs if two such curves given by $Z_0(z, t) = Z_1$ and $z_0(z, t) = Z_2$ intersect. This concept is easily extended to two or three dimensions. For example, in two dimensions, the characteristic curves are replaced by characteristic “surfaces”, surfaces in, say, (x, y, t) -space defined by the equation $x_0(x, y, t) = X_0$. The solution to the wave equation propagates along these surfaces, and wave breaking occurs if two such surfaces having different values X_1 and X_2 at $t = 0$ intersect for some $t > 0$. For such intersection to take place, the wave fronts, i.e. the intersections of the characteristic planes with the planes of constant t , have to move in the direction of propagation for the wave. Therefore, this type of wave breaking will be referred to as *longitudinal wave breaking*.

In plasma fluid dynamics, longitudinal wave breaking in two or three dimensions is often dealt with through a simple 1-D description. Such treatment is justified if the fluid elements oscillate mostly in the direction of propagation of the wave, and if the transverse derivatives are small compared to the longitudinal ones, i.e. the wave resembles a plane wave to some extent, especially along its central axis of propagation.

It is obvious that for a wave with fronts that do not resemble planes at all, the 1-D approximation is bound to fail. This is particularly true for a wave propagating

through a plasma channel, such as is often employed in laser wakefield acceleration to keep the laser pulse from diverging. In this case, a different form of wave-breaking, namely *transverse wave breaking* [11], can occur. This works as follows. It can be shown that in a more or less parabolic channel, the wave front of a laser-induced wakefield becomes strongly curved, with maximum curvature on the axis of the channel and proportional to the distance behind the laser pulse. This implies that at a sufficient distance behind the laser pulse the curvature radius will become smaller than the amplitude of the displacement of the plasma electron fluid. In that case, wave breaking occurs not because two different characteristic surfaces intersect, but because the two halves of a single surface meet. Contrary to longitudinal breaking, which usually occurs within the first few plasma oscillations behind the laser pulse, transverse breaking may very well occur many plasma periods behind the pulse. Also, a transverse break can occur at much lower intensities than a longitudinal break, since the decrease of the curvature radius of the wave fronts also takes place for low-amplitude wake waves. Transverse wave breaking can be responsible for the injection of electrons in the wakefield in the transverse direction, many periods behind the pulse, and may be an efficient mechanism for the trapping and acceleration of a relatively small fraction of the plasma electrons.

Transverse wave breaking can happen in any situation for which the electron plasma frequency has a transverse profile with a minimum on axis. This situation occurs for example in preformed plasma channels, capillaries, and laser formed plasma channels, but also when the plasma frequency decreases due to relativistic electron mass increase. This frequency decrease increases with increasing laser intensity, so the desired frequency minimum occurs at the centre of the pulse.

Transverse wave breaking may be important in channel-guided laser-plasma interaction, since it limits the number of regular wave crests in the wakefield behind the driving pulse. In the remainder of this chapter though, wave breaking as a means of generating large quantities of fast electrons will be investigated. Since the contribution of transverse wave breaking to the number of plasma electrons injected into the wave is small according to Ref. [11], the focus will be on longitudinal wave breaking in the following sections.

4.4 Wave breaking in the quasi-static approximation

We consider the case of a wave in an infinitely extended (quasi-)homogeneous plasma, driven by a source moving with a constant group velocity v_g , such as a laser pulse or a driving bunch of particles. To study this particular case, it is often convenient to transform to the comoving coordinate system (ζ, τ) , where

$\zeta = z - v_g t$, $\tau = t$. Note the similarities with the transformation to Lagrangian coordinates (4.4). It is assumed that the evolution of the source in this frame happens on a much slower time scale than the plasma response, in which case the phase velocity of the driven plasma wave is taken to be equal to the group velocity of the source [21]. In the (ζ, τ) -system, the advection equation (4.1) with additional non-evolving source term $f(\zeta)$ reads

$$\frac{\partial p}{\partial \tau} + (v - v_g) \frac{\partial p}{\partial \zeta} + E = f(\zeta), \quad (4.9)$$

Since the right-hand side of this equation does not depend on τ , its solution is often taken to be more or less static with respect to the comoving frame. In that case, application of the *quasi-static approximation* (QSA), as introduced in Chapter 2, is justified: the $\partial p / \partial \tau$ term will often be dropped entirely from the advection equation, and resulting solutions will depend on ζ only. This approximation has been used to describe plasma waves already in the work of Akhiezer and Polovin [22].

To study wave breaking in the QSA, one cannot simply apply the above definitions for wave breaking for the following reason. Translating the approximation $\partial p / \partial \tau = 0$ back to (z, t) coordinates, we find

$$\frac{\partial p}{\partial t} + v_g \frac{\partial p}{\partial z} = 0. \quad (4.10)$$

This is an advection equation in itself, which actually replaces the original advection equation (4.1), while (4.9) without the $\partial p / \partial \tau$ term is transformed into an initial condition for (4.10). For this new advection equation, the local phase velocity is a constant, so all its characteristic curves are parallel. Under these conditions, a solution p to (4.10) cannot evolve to wave breaking conditions starting from a non-breaking situation. Therefore, most of the literature on plasma waves uses an alternative definition for wave breaking that does not match that of Whitham [14]. First note that a solution to (4.10) can be obtained by integrating (4.9) and applying an appropriate boundary condition for $\zeta \rightarrow \infty$. However, it can be shown that such integration will fail if the intensity of the source is too high, such that v reaches v_g at some point. The alternative definition of wave breaking states, that a wave is non-broken as long as (4.9) can be integrated, and that wave breaking occurs at the moment that such integration fails.

A disadvantage of the quasi-static approximation is that it may not accurately describe the behaviour of a cold plasma wave close to the threshold at which integration of (4.9) fails. Such a wave typically displays sharp spikes in the electron density at which $\partial n / \partial \tau = 0$ fails. Also, plasma electrons that were initially at rest will be “reflected” by the accumulation of charge around the break, causing $\partial p / \partial \tau = 0$ to fail. In other words, the wave may be quasi-static before and after it has broken, but not while wave breaking is in progress.

There are also arguments that speak in favour of the QSA though. The first argument is that for a cold plasma, the density for a regular wave in the QSA satisfies the conditions (4.7) and (4.8) (the latter at least for a non-relativistic wave), while a non-regular wave in the QSA violates them. This means that the QSA correctly predicts the limit for 1-D wave breaking, even though it may not correctly describe the wave's behaviour when it is breaking. The second argument is, that in case of a warm plasma, the presence of pressure terms in the advection equation inhibits the characteristics from intersecting or even from getting close to each other, so taking them to be parallel in the QSA is a reasonable approximation there.

Notwithstanding its limitations, the QSA has been used on many occasions to predict the electric field threshold at which wave breaking sets in. We denote this threshold by E_{wb} .

In the remainder of this section, we describe the procedure to obtain the wave breaking threshold for the electric field E_{wb} , using the quasi-static approximation, for the case of a relativistic wave in a cold plasma. Then it will be shown that three other cases, namely cold non-relativistic, warm non-relativistic, and warm relativistic, are just variations on the same theme. We start with expressing the equations for the plasma electron fluid in the wakefield potential $\Psi = 1 + \Phi$, with Φ the electrostatic potential, as we did in Chapter 2. We rewrite (4.3) in the QSA to obtain:

$$\frac{\partial(n(v - v_g))}{\partial\zeta} = 0, \quad n = \frac{v_g}{v_g - v}.$$

The wakefield equations then read

$$\frac{\partial\Psi}{\partial\zeta} = -E, \tag{4.11}$$

$$\frac{\partial^2\Psi}{\partial\zeta^2} = \frac{v}{v_g - v}. \tag{4.12}$$

The advection equation for the plasma wave can then be used to express Ψ in terms of either p or v . In this case (cold relativistic), we have $v = p/\sqrt{1 + p^2}$ and the advection equation (4.9) behind the driving source ($f(\zeta) = 0$) reads

$$\left(\frac{p}{\sqrt{1 + p^2}} - v_g \right) \frac{\partial p}{\partial\zeta} = -E = \frac{\partial\Psi}{\partial\zeta}, \tag{4.13}$$

from which Ψ is found to be

$$\Psi = \sqrt{1 + p^2} - v_g p. \tag{4.14}$$

Since $\partial\Psi/\partial p \neq 0$ for $p = 0$, this expression defines a one-to-one relationship between Ψ and p for not too large p . Inversion of (4.14) yields $v = v(\Psi)$ or

$p = p(\Psi)$; insertion of this into (4.12) yields an oscillatory equation for Ψ . Any oscillatory solution for Ψ to this equation represents a quasi-static relativistic longitudinal plasma wave, provided the amplitude is small enough such that $\partial\Psi/\partial p \neq 0$ everywhere. For $p = \gamma_g v_g$, (4.14) yields that $\partial\Psi/\partial p = 0$, which sets an upper limit on the amplitude of regular oscillations of Ψ .

Multiplying Poisson's equation (4.12) by (4.13) and using $\partial\gamma/\partial\zeta = v\partial p/\partial\zeta$ yields the energy equation for the wakefield:

$$\frac{\partial}{\partial\zeta} \left[\frac{1}{2}E^2 + \sqrt{1+p^2} \right] = 0. \quad (4.15)$$

We observe that p reaches its maximum value p_m for $E = 0$, while E^2 reaches its maximum value for $p = 0$. Then integration of (4.15) yields:

$$\frac{1}{2}E^2 = \sqrt{1+p_m^2} - \sqrt{1+p^2}. \quad (4.16)$$

As it happens, there are two singularities hidden in the system (4.11)-(4.13), which may both be responsible for “wave breaking”, that is the breakdown of the quasi-static solution. The first singularity occurs for $v \uparrow v_g$. The second singularity occurs when $\partial\Psi/\partial p = 0$ while $\partial p/\partial\zeta \neq 0$. From the energy equation (4.15), we find that in that case $\partial E^2/\partial\zeta = 0$ while $\partial\gamma/\partial\zeta \neq 0$, so conservation of energy is violated. Obviously, the wave solution cannot progress beyond this point.

The second singularity is often expressed as an upper limit for the density n , and is known as the *Coffey criterion* or *Coffey limit*. As a matter of fact, it is an equivalent formulation of the amplitude limit for n found by Coffey for thermal plasmas [17]. For a cold plasma, the Coffey limit translates to $v = v_g$ or $1/n = 0$, i.e. it coincides with the singularity in the plasma density. For a warm plasma however, an enthalpy term is added to (4.15), and the Coffey limit already occurs for n equals some $n_{max} < \infty$. As we will see later, this is the reason that wave breaking in a warm plasma happens at lower electric field amplitudes than in a cold plasma. See Bergmann and Schnabl [24] for details.

Since for a regular plasma wave, there exist one-to-one relations between p , v , and n , the Coffey criterion can also be expressed as $\partial\Psi/\partial v = 0$ or $\partial\Psi/\partial n = 0$.

We have seen before that in the cold relativistic case, the Coffey limit is reached for $p = \gamma_g v_g$. In other words, a plasma wave with momentum amplitude $p_m \geq \gamma_g v_g$ will break sooner or later. From this, we can derive the electric field threshold for wave breaking E_{wb} . Setting $p = 0$ and $p_m = \gamma_g v_g$ in (4.16) yields the well-known result $E_{wb} = \sqrt{2(\gamma_g - 1)}$.

We proceed to discussing wave breaking in three additional cases: cold non-relativistic, warm non-relativistic, and warm relativistic. These cases only differ

in the advection equation used for p or v , and thus the relation between Ψ and p or v . Wave breaking thresholds have been derived from Coffey's criterion.

Cold non-relativistic. This result was first obtained by Dawson [23]. In this case, the advection equation reads:

$$(v - v_g) \frac{\partial v}{\partial \zeta} = -E, \quad (4.17)$$

from which we derive the wakefield potential

$$\Psi(v) = \frac{1}{2}v^2 - vv_g.$$

Combining (4.17) with the Poisson equation (4.12) yields

$$\frac{\partial}{\partial \zeta} \left[\frac{1}{2}E^2 + \frac{1}{2}v^2 \right] = 0. \quad (4.18)$$

Using Coffey's criterion, we find that a singularity occurs for $v = v_g$, i.e. coincident with a singularity in n . A wave on the verge of breaking reaches $v = v_g$ at $E = 0$. From (4.18) it is then found that $E_{wb} = v_g$. Note that this matches the cold relativistic result in the limit $v_g \ll c$.

Warm non-relativistic. The warm non-relativistic result was first obtained by Coffey [17]. It is derived from the 1-D Vlasov equation using a waterbag velocity distribution: the density function $f(t, x, v)$ is initially constant and equal to $n_0/2v_0$ for $-v_0 < v < v_0$ where $v_0 = \sqrt{3kT/m}$, and zero otherwise. The value of v_0 is chosen such that $\int v^2 f dv$ is the same as for a Maxwellian distribution. It then follows that the value of f will always equal $n_0/2v_0$, and the evolution of the Vlasov fluid can be described in terms of its lower and upper velocity boundaries denoted by $v_{\pm}(z)$. Then the mean fluid velocity $V(z)$ is found to equal $(v_+(z) + v_-(z))/2$, and the bag boundaries are given by $v_{\pm}(z) = V(z) \pm v_0 n(z)/n_0$.

The wakefield equations are then derived by taking moments of the Vlasov equation. The zeroth moment ($\int dv$) of the Vlasov equation yields the continuity equation $n(v_g - V) = v_g$, while the first moment ($\int dv v$) yields the momentum equation, i.e. the advection equation for V . In the QSA, the advection equation reads

$$(V - v_g) \frac{\partial V}{\partial \zeta} + \frac{\beta v_g^2}{2} \frac{\partial}{\partial \zeta} n^2 = -E, \quad (4.19)$$

and the wakefield potential is given by

$$\Psi(V) = \frac{1}{2}V^2 - Vv_g + \frac{\beta v_g^4}{2(v_g - V)^2}.$$

Here, $\beta \equiv 3kT/(mv_g^2)$ denotes the temperature coefficient. Application of Coffey's criterion yields that $\partial\Psi/\partial V = 0$ for $V = v_g(1 - \beta^{1/4}) < v_g$, i.e. before a singularity in n is reached. As before, multiplying Poisson's equation (4.12) by (4.19) yields the energy equation for the wakefield:

$$\frac{\partial}{\partial \zeta} \left[\frac{1}{2}E^2 + \frac{1}{2}V^2 + \beta v_g^4 \left(\frac{1}{2(v_g - V)^2} - \frac{1}{3(v_g - V)^3} \right) \right] = 0. \quad (4.20)$$

A wave on the verge of breaking has velocity amplitude $V_m = v_g(1 - \beta^{1/4})$, reached when $E = 0$, and field amplitude E_{wb} , reached when $V = 0$. Then E_{wb} can be obtained from (4.20) to yield:

$$E_{wb} = v_g \left(1 - \frac{8}{3}\beta^{1/4} + 2\beta^{1/2} - \frac{1}{3}\beta \right)^{1/2}.$$

A similar result has been derived by Aleshin, Drofa, and Kuz'menkov [25] through slightly different methods. They also integrate the equations describing the warm plasma fluid to obtain Ψ , E and n as functions of ζ , for both warm and cold plasmas. For a wave having an amplitude larger than the wave breaking limit, they claim that the wave period diverges, leading to the propagation of a time-independent solitary wave.

Note that for $\beta \downarrow 0$, Dawson's result is reproduced.

Warm relativistic. The warm relativistic result has been investigated by various groups. The advection equation for the mean plasma velocity, denoted by v , can be derived from the relativistic Vlasov equation as before. Katsouleas and Mori [20] used a waterbag model (bag boundaries at $p = \pm\sqrt{3kT/(mc^2)}$) to derive the following wakefield potential:

$$\Psi_{KM} = \frac{1 - vv_g}{\sqrt{1 - v^2}} \left(1 + \beta \frac{1 - v^2}{(1 - v/v_g)^2} \right)^{1/2},$$

where $\beta \equiv 3kT/(mc^2)$ for the relativistic case. Then $\partial\Psi_{KM}/\partial v = 0$ for $v \approx v_g - 2\sqrt{\beta}/\gamma_g$. Since the use of Ψ_{KM} rendered integration of Poisson's equation somewhat difficult, the wave-breaking electric field was determined using

$$\tilde{\Psi}_{KM} = \lim_{v_g \rightarrow 1} \Psi_{KM} = \sqrt{x^2 + \beta},$$

where $x^2 \equiv (1 - v)/(1 + v)$. This yields for E_{wb} :

$$E_{wb} = \beta^{-1/4} \sqrt{\ln(2\gamma_g^{1/2} \beta^{1/4})}.$$

From the way E_{wb} is calculated, its value is inherently too high, and the above expression is only expected to be valid in the limit $\gamma_g\sqrt{\beta} \gg 1$.

Rosenzweig [26] starts from a different wakefield potential given by (in the limit $v_g \rightarrow 1$): $\Psi_R = x + \beta/x$ where x is defined as before. Application of the Coffey criterion yields $\partial\Psi/\partial v = 0$ for $v = (1 - \sqrt{\beta})/(1 + \sqrt{\beta})$, and E_{wb} is found to be

$$E_{wb} = \left(\frac{4}{9\beta}\right)^{1/4}.$$

In a later paper [27], Rosenzweig employs a three-fluid model to obtain the following wave breaking threshold:

$$E_{wb} = \left(\frac{\gamma_g}{9\beta}\right)^{1/4}.$$

In this model, particle trapping arguments are used to determine E_{wb} instead of $\partial\Psi/\partial v = 0$.

Finally, Sheng and Meyer-ter-Vehn [28] start from the warm relativistic electron fluid theory proposed by Newcomb [29], and proceed to derive a very complicated expression for the wakefield potential. In the limit $v_g \rightarrow 1$, they find that $\partial\Psi/\partial v = 0$ for $v \approx 1 - 2\beta + \frac{1}{3}\beta^{3/2} - \dots$, and derive $E_{wb} = (4/(9\beta))^{1/4}$, similar to Rosenzweig's first result. They also claim that if their methods (actually a variant on the Coffey criterion) are applied to Ψ_{KM} as given above, then wave breaking in the limit $v_g \rightarrow 1$ occurs at $v = (1 - \beta)/(1 + \beta)$ and $E_{wb} = (2/(9\beta))^{1/4}$. This is actually quite remarkable, since their method and that of Katsouleas and Mori are mathematically equivalent and should therefore yield the same results.

It is interesting to note that the threshold of Katsouleas and Mori and the second threshold of Rosenzweig increase with γ_g , while the first threshold of Rosenzweig and those of Sheng and Meyer-ter-Vehn do not. In the latter case, a wave with $v_g = 1$ can still break due to thermal effects. This corresponds to the notion that such a wave may still be able to trap particles, even though these particles will never be able to keep up with the wave. This works as follows. We refer to the linear plasma wave of which the phase space is depicted in Fig. 3.1. If the limit $v_g \rightarrow 1$ is taken for this wave, then the momentum $\gamma_g v_g$ of the X-points and the O-points tends to infinity, but the minimum momentum necessary for trapping (point L in the graph) tends to the finite value $(1 - 4e_0^2)/(4e_0)$, where e_0 denotes the field amplitude. Once a particle has been trapped, it will slip back with respect to the wave for all time, but the total distance along which it slips will be finite, and the particle will not leave the wave crest in which it has initially been trapped. Note that in a cold plasma, a particle needs to pass through an X-point in order to get trapped, for which an infinite amount of energy is needed. So in the limit $v_g \rightarrow 1$, a finite amplitude wave cannot trap particles having finite energy. In a warm plasma however, the finite width of the particle distribution function in phase space may cause particles to get trapped beneath an X-point, where the separatrices lie very close together, requiring less energy. So in the

limit $v_g \rightarrow 1$, a finite amplitude wave may still trap particles having finite energy. In short, E_{wb} may or may not reach arbitrarily large values in the limit $v_g \rightarrow 1$. Deciding which view is the correct one lies beyond the scope of this thesis.

In summary, all the results on wave breaking presented above were derived through the same method, even if their respective authors claimed otherwise. The only thing in which they differ is the expression for the wakefield potential Ψ derived from the advection equation. These differences arise from (i) different initial energy distributions in the Vlasov equation, (ii) different pressure terms and/or thermal contributions to the internal energy, and (iii) different approximations in the limits $v_g \rightarrow 1$ and $\beta \ll 1$. Especially in the warm relativistic case, methods and results diverge considerably. Moreover, there are serious indications that the QSA is not fit to model a complicated phenomenon as plasma wave breaking.

In the remainder of this chapter, the cold relativistic case will be explored in more depth. The excitation of a non-linear plasma wave by the ponderomotive force of a laser pulse will be investigated, conditions for wave breaking will be derived, and predictions of the extracted charge versus laser intensity will be given. It is stressed that the methods developed are applicable to all four of the above cases, even though they are only applied to a relativistic wave in a cold plasma.

4.5 Non-linear waves in a cold plasma

We start from an infinitely extended, homogeneous plasma, through which a laser pulse propagates in the z -direction. The laser and plasma frequencies are denoted by ω and ω_p respectively, and the plasma is assumed to be underdense: $\omega_p < \omega$. We denote the group velocity of the laser pulse, which is at the same time the phase velocity of the wake wave, by v_g , and the corresponding Lorentz factor by γ_g .

The following model assumptions concerning the plasma are made: the plasma is cold and temperature effects are neglected, the plasma electrons are initially at rest, and the plasma background density n_0 changes on a very large length scale so n_0 will be considered a constant. Concerning the laser pulse, we assume that it is one-dimensional and circularly polarized, and that its envelope does not change as it propagates through the plasma. As a consequence, instabilities in the pulse evolution due to Raman scattering are not investigated here. Furthermore, we assume that γ_g is sufficiently large, so all the terms of $\mathcal{O}(1/\gamma_g^2)$ or smaller can be neglected.

We separate fast and slow timescales, transform to the coordinates $\zeta = z - v_g t$, $\tau = t$, and apply the QSA: $\partial/\partial\tau = 0$. Then the slow, longitudinal motion of the electron fluid is governed by a modified version of the system (4.11)-(4.13). In the

presence of a laser pulse having envelope $U(\zeta) = \langle |A_\perp|^2 \rangle(\zeta)$, the electron Lorentz factor is given by $\gamma = \sqrt{1 + p^2 + U}$ (cf. Chapter 2), while the plasma electrons are not only driven by the electrostatic force $-E$, but also by the ponderomotive force $(1/2\gamma)\partial U/\partial\zeta$. Then the wakefield equations become

$$\frac{\partial^2 \Psi}{\partial \zeta^2} = \frac{p}{\gamma v_g - p}, \quad (4.21)$$

$$\left(\frac{p}{\gamma} - v_g\right) \frac{\partial p}{\partial \zeta} = \frac{\partial \Psi}{\partial \zeta} - \frac{1}{2\gamma} \frac{\partial U}{\partial \zeta}. \quad (4.22)$$

Integrating (4.22) yields Ψ in terms of p and $U(\zeta)$:

$$\Psi = \sqrt{1 + p^2 + U(\zeta)} - v_g p.$$

It follows that Ψ has a U -dependent lower bound: $\Psi \geq \Psi_0(U) := (\sqrt{1 + U})/\gamma_g$, with $\Psi = \Psi_0(U)$ for $p = p_{cr} := \gamma_g v_g \sqrt{1 + U}$, i.e. $v = v_g$ and $\gamma = \gamma_g \sqrt{1 + U}$. As a result, the mapping $\Psi \mapsto p$ is double-valued:

$$p_{\pm} = \gamma_g^2 \left(v_g \Psi \pm \sqrt{\Psi^2 - (1 + U)/\gamma_g^2} \right), \quad (4.23)$$

where $p_- < p_{cr}$ and $p_+ > p_{cr}$. Since we wish to study regular plasma waves first, for which the plasma fluid velocity is below the wave velocity, $p = p_-$ is taken as an initial condition for the momentum oscillation.

Using (4.21), (4.23), and $\gamma = \Psi + v_g p$, the wakefield equations can be reduced to a single equation for Ψ :

$$\begin{aligned} \frac{d^2 \Psi}{d\zeta^2} &= -\frac{\partial}{\partial \Psi} V(\Psi, U), \\ V(\Psi, U) &= \gamma = \gamma_g^2 \left[\Psi - v_g \sqrt{\Psi^2 - (1 + U)/\gamma_g^2} \right]. \end{aligned} \quad (4.24)$$

This differential equation describes anharmonic oscillations of the variable Ψ . (See Umstadter et al. [30] for a similar equation, although with a slightly different $V(\Psi, U)$, or Bulanov et al. [31] for a similar equation, with ion motion included.) It is analogous to the Newton equation for a virtual particle with “coordinate” Ψ , moving in a 1-D “potential well” $V(\Psi, U)$, where ζ is the “time” and $U(\zeta)$ acts as a “time”-dependent parameter. Continuing this analogy, we define the “velocity” of the virtual particle as $d\Psi/d\zeta$, and its “kinetic energy” as $\frac{1}{2}(d\Psi/d\zeta)^2$.

It should be noted that a maximum (minimum) of Ψ corresponds to a minimum (maximum) of the local electron density, with minimal (maximal) electron fluid momentum p and *zero* electric field. The minimum of $V(\Psi, U)$ as a function of Ψ corresponds to a point with maximal $|\partial \Psi/\partial \zeta|$, i.e. maximal electric field, *zero* electron fluid momentum p , and $n = 1$.

When $\Psi \downarrow \Psi_0(U)$, we have $\partial V/\partial \Psi \rightarrow -\infty$, $n \rightarrow \infty$, and $p_- = p_+ = p_{cr}$. It is possible that, after passing this point, $p = p_+$ for some plasma electrons, so $\Psi = \Psi_0(U)$ is a bifurcation point. These electrons then overtake the wake wave, which is called wave breaking [20,23]. It can be concluded from the equations of motion that wave breaking is bound to happen if a sufficiently large oscillation is excited in the plasma.

4.6 Excitation of 1-D plasma waves

Following the approach of the previous section, longitudinal plasma waves are described as oscillations of a virtual particle with “coordinate” $\Psi(\zeta)$, in a potential well $V(\Psi, U)$. The mechanism behind the excitation of such oscillations by strong laser pulses will be considered below.

Multiplying (4.24) by $\partial \Psi/\partial \zeta$ yields the energy equation for the virtual particle:

$$\frac{\partial}{\partial \zeta} \left(\frac{1}{2} \left(\frac{\partial \Psi}{\partial \zeta} \right)^2 + V(\Psi, U) \right) = 0. \quad (4.25)$$

As a matter of fact, this is Eq. (4.15) with the addition of the ponderomotive pressure of a laser pulse. We employ this equation to describe plasma oscillations. For simplicity, we assume that the laser intensity envelope $U(\zeta)$ is rectangular with (scaled) length L : $U(\zeta) = U_0$ for $-L \leq \zeta \leq 0$, and 0 otherwise. Then $V(\Psi, U)$ can be seen as a function of Ψ alone in two parameter regimes: $U = 0$ and $U = U_0$. Oscillations are excited by moving the virtual particle away from the bottom of the “potential well”, located at $\Psi = \Psi_m := \sqrt{1+U}$. This will be done as follows (Roman numbers correspond to those in Fig. 4.1).

(I) In front of the pulse, the plasma is at rest, so we start at $\Psi = 1 = \Psi_m$ for $U = 0$.

(II) At $\zeta = 0$, U increases to U_0 , while Ψ remains constant. Since the minimum of $V(\Psi, U_0)$ is at $\Psi = \sqrt{1+U_0} > 1$, the virtual particle will start to oscillate between $\Psi = 1$ and $\Psi = \Psi_{\text{III}}(U) > 1$, where $\Psi_{\text{III}}(U)$ denotes the other solution to $V(\Psi, U) = V(1, U)$ (cf. Fig. 4.1), and is given by

$$\Psi_{\text{III}}(U) = \gamma_g^2 \left(2 - \frac{1}{\gamma_g^2} - 2v_g \sqrt{1 - (1+U)/\gamma_g^2} \right). \quad (4.26)$$

(III) When the laser intensity drops back to zero at $\zeta = -L$, the potential energy of the virtual particle drops from $V(\Psi(-L), U_0)$ to $V(\Psi(-L), 0)$.

(IV-V) Since $(\partial \Psi/\partial \zeta)^2 > 0$ and $\Psi > 1$, an oscillation in the potential well for $U = 0$ has been excited.

The total “energy” of the oscillation equals $V(1, U_0) - V(1, 0) - (V(\Psi(-L), U_0) - V(\Psi(-L), 0))$, and since $V(\Psi, U_0) - V(\Psi, 0)$ decreases for increasing Ψ , we find that the total oscillatory “energy” at a given pulse intensity U_0 is maximal for $\Psi(-L) = \Psi_{\text{III}}$. For $\Psi(-L) < \Psi_{\text{III}}$, the total “energy” differs from the maximal “energy” as $\Delta E \sim \mathcal{O}((\Psi_{\text{III}} - \Psi(-L))^2)$.

One must note that, whereas the front edge of the laser pulse increases the potential energy of the virtual particle, the back edge decreases it. In the special case that $\Psi(-L) = 1$, the final energy of the virtual particle will be zero, and there will be no wake behind the pulse. For this reason, we define the pulse length L to be optimal, if the excited plasma oscillation has the maximal possible energy for the pulse intensity U_0 . For the remainder of this chapter, L is always assumed to be optimal.

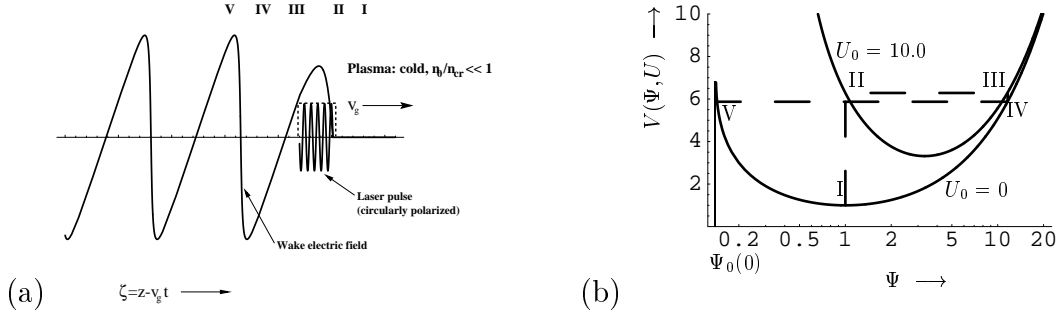


Figure 4.1: Excitation of a plasma oscillation by a rectangular pulse with $n/n_{cr} = 0.02$ and $U_0 = 10.0$, depicted in (ζ, E) -space (a) and (Ψ, V) -space (b).

If the total “energy” level of the oscillation after passage of the pulse exceeds $V(\Psi_0(0), 0) = \gamma_g$, then we will have $\Psi(\zeta) = \Psi_0(0)$ for some ζ , while $|\partial\Psi/\partial\zeta| > 0$. (This means that, in figure 4.1, the endpoint $\Psi_0(0)$ will be reached.) As mentioned above, the wave breaks at this point. For the maximal electrostatic field associated with a wave on the verge of breaking, we recover, using (4.25), the well-known result of Akhiezer and Polovin [22]:

$$E_{\text{max}} = \sqrt{2(\gamma_g - 1)}. \quad (4.27)$$

For a complete algebraic description for the relativistic wakefield excited by a square pulse, see Teychenné, Bonnaud and Bobin [32].

4.7 Fast particle generation as wave breaking

From the previous section, we know that the wake wave driven by the laser pulse breaks if $\Psi(\zeta) = \Psi_0(U_0)$ at some point, causing the production of electrons with

speed $v > v_g$. These electrons may be captured and accelerated by the wake wave itself, attaining a final energy up to $2\gamma_g^2 - 1$ as shown in Section 4.9. This mechanism is believed to be responsible for fast particle generation by a single intense laser pulse [2].

We are going to investigate two possible scenarios for reaching the wave breaking point: (i) wave breaking at the front edge of the laser pulse, and (ii) wave breaking during the first plasma oscillation behind the pulse.

In the first scenario, the envelope $U(\zeta)$ increases from 0 to its maximum value U_0 , while at the same time Ψ stays at its original value of 1. (See steps I and II in figure 4.1.) The corresponding value of $\Psi_0(U)$ increases from $1/\gamma_g$ to $(\sqrt{1+U_0})/\gamma_g$. If $\Psi_0(U)$ reaches the initial value of $\Psi = 1$ on the way, the singularity at $\Psi = \Psi_0(U)$, will be encountered right at the front edge of the pulse. The threshold for this mechanism of wave breaking is

$$U_0 = \gamma_g^2 - 1. \quad (4.28)$$

This scenario could dominate fast electron generation at high laser intensities, but it does not explain the intensity thresholds found in experiments and simulations, which are commonly one order of magnitude lower than predicted by (4.28).

In the second scenario, U_0 is assumed to be below $\gamma_g^2 - 1$, so there will be no wave breaking at the front edge of the pulse. Instead, a plasma wave will be excited that is strong enough to break behind the pulse. As shown in the previous section, this will happen as soon as the total energy of the oscillation exceeds γ_g . This situation can be reached as follows. Define Ψ_1 as the other solution for Ψ of $V(\Psi, 0) = V(\Psi_0(0), 0)$:

$$\Psi_1 = \gamma_g^2(2 - 1/\gamma_g^2)\Psi_0(0) = \gamma_g(2 - 1/\gamma_g^2). \quad (4.29)$$

Then the energy of the system will reach the wave breaking threshold if at point III in Figure 4.1, $\Psi(-L) \geq \Psi_1$.

Since L is optimal, wave breaking will occur if $\Psi_{\text{III}}(U_0) \geq \Psi_1$, or

$$U_0 \geq 2(\gamma_g - 1) - \frac{\gamma_g - 1}{4\gamma_g^2(\gamma_g + 1)} \approx 2(\gamma_g - 1). \quad (4.30)$$

For any $\gamma_g > 1$, we find that the wave breaking limit for U_0 in the second scenario is smaller than in the first scenario, so this second scenario is at least partly responsible for fast electron generation at intensities below $\gamma_g^2 - 1$. The above limit defines an upper limit to the laser intensity threshold for wave breaking. It is also in agreement with earlier theoretical results [32]. It can be expected that thermal effects will lower this intensity threshold, just as they lowered the corresponding threshold for the electric field.

As an example, we see that for $n_0/n_{cr} = 0.02$ and laser wavelength $\lambda_0 = 800$ nm, the first scenario predicts wave breaking for $U_0 \geq 49$, corresponding to a power threshold of $1.1 \cdot 10^{20}$ W/cm². The second scenario predicts a power threshold of $2.6 \cdot 10^{19}$ W/cm². Simulations by Nagashima et al. [33] predict a threshold of approximately 10^{19} W/cm² for this case. Malka et al. [4] conducted simulations with $n_0/n_{cr} = 0.05$ and $\lambda_0 = 1$ μ m, and found an intensity threshold of $10^{18} - 10^{19}$ W/cm², where our first scenario predicts $P = 2.7 \cdot 10^{19}$ W/cm², and the second predicts $P = 1.0 \cdot 10^{19}$ W/cm².

One should note that these thresholds are only valid for the case of a cold plasma and a regime for which ponderomotive excitation of the plasma wave is the dominant mechanism. Thermal effects (see above) or the presence of Raman instabilities (see Chapters 5 and 6) may lower these thresholds considerably.

4.8 Simulations

In the previous sections, it has been assumed that the laser pulse had a rectangular envelope, and expressions for thresholds for U or upper boundaries for Ψ have been derived under this assumption. Since in reality, the envelope of a laser pulse is not rectangular but rather smooth, with low-intensity tails at the front and back of the pulse, simulations have been conducted to study the influence of the pulse shape on the excitation of the plasma wave.

The idea behind the simulations is, that the electric field excited behind the laser pulse has at least one minimum behind the pulse, even in the case of wave breaking. From this minimum, the amplitude of the field can be determined. If this amplitude is below the Akhiezer-Polovin limit of $\sqrt{2(\gamma_g - 1)}$, the plasma wave will not break. If on the other hand the amplitude is above the limit, the wave will break soon after the first minimum in the field. This way, wave breaking can be diagnosed without integrating near the location of the break, where gradients are steep and numerical integration becomes unstable.

In the simulations, the equations $\partial\Psi/\partial\zeta = -E$ and $\partial E/\partial\zeta = -\partial V(\Psi, U)/\partial\Psi$ (see (4.24)) have been integrated with respect to ζ , to obtain the amplitude of the electrostatic field E behind the laser pulse. The second order Adams Bashforth method has been used, with step size $\Delta\zeta = 0.002$. The simulations have been conducted using three different pulse shapes: rectangular with width $2b$ and height U_0 , polynomial with envelope $U(\zeta) = U_0(1 - (\zeta/b)^2)^3$, and Gaussian with envelope: $U(\zeta) = U_0 \exp(-\frac{1}{2}(3\zeta/b)^2)$. This way, all pulses satisfy $L = 2b$. Two different plasma densities have been used: $n/n_{cr} = 0.02, 0.05$. For each combination of pulse shape and density, the amplitude of the electric field behind the pulse has been determined for a range of values for both U_0 and b .

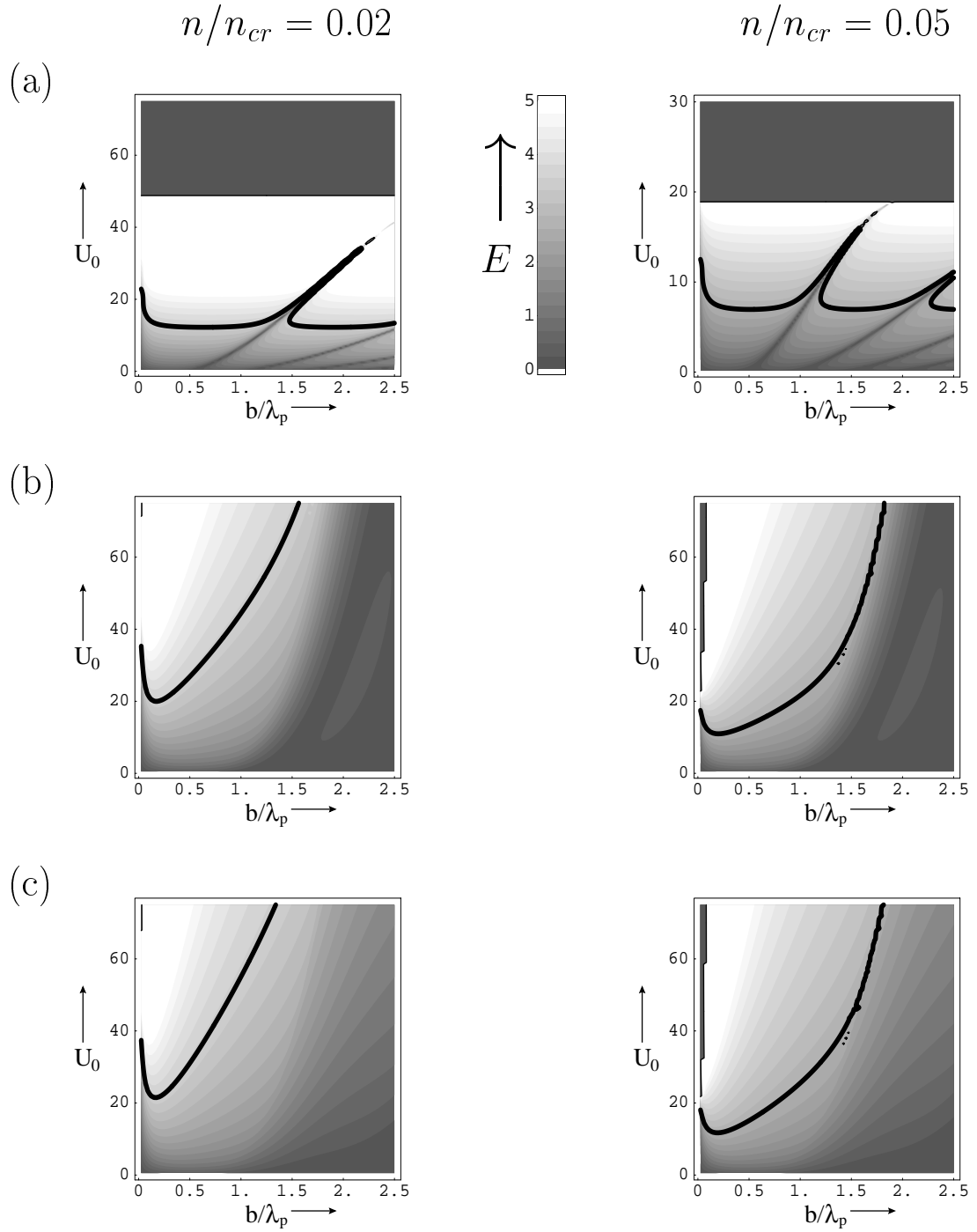


Figure 4.2: Amplitude of the electric wakefield for a pulse with (a) a rectangular envelope, (b) a polynomial envelope, and (c) a Gaussian envelope, as a function of pulse width and height. The heavy black curves denote the Akhiezer-Polovin field limit for wave breaking.

The results are shown in Figure 4.2. In the figures, the amplitude of the excited wakefield is shown as a function of length b and peak intensity U_0 . Note that all lengths are normalized, i.e. the plasma wavelength $\lambda_p = 2\pi$. Bright areas correspond to a high amplitude, while dark areas correspond to a low amplitude. The heavy black curve in each graph corresponds to the Akhiezer-Polovin limit, which separates the wave-breaking region (high amplitude, bright area above the curve) from the non-wave-breaking region (low amplitude, dark area below the curve). The black areas at the top of the figures 4.2(a) and at the top left corner of the figures 4.2(b) and (c) correspond to wave breaking at the front of the pulse. In that case, calculation of neither the wakefield nor the amplitude is possible.

From the results, it immediately follows that the minimum scaled laser intensity necessary to reach the Akhiezer-Polovin limit decreases with increasing plasma density. The main cause for this is the decrease of γ_g , and thus of this limit, with increasing plasma density, since the (scaled) electric field amplitude does not change considerably. For the same reason, intensity thresholds for wave breaking at the front of the pulse decrease as well. It is also obvious that short pulses with high intensity excite large wakefields, while long pulses with low amplitude hardly excite anything.

Regarding pulse envelope shape, there is a clear distinction between excitation by a rectangular pulse and by a pulse with a smooth envelope: excitation by a rectangular pulse is much more efficient. This follows from the fact that a rectangular pulse has infinite derivatives at the front and the back, regardless of the length of the pulse. For this reason, the behaviour of the field amplitude for constant U_0 is periodic with respect to b . The period is half a *non-linear* plasma wavelength, which can be considerably larger than λ_p for large U_0 . The “trenches” of low amplitude that can be seen in the graphs correspond to pulses having a length that is an integral multiple of the non-linear plasma wavelength. As mentioned before, such pulses have $\Psi(2b) = \Psi(0) = 1$ and as a consequence they hardly excite any wakefield at all, while slightly longer or shorter pulses may excite a considerable wakefield. Note that because of the infinitely steep front of the pulse, wave breaking at the front is inevitable for $U_0 \geq \gamma_g^2 - 1$, regardless of the pulse length.

Things are different for the pulses with smooth polynomial or Gaussian envelope. Such pulses have finite derivatives that decrease with increasing pulse length. It can be shown that the amplitude of the excited plasma wave depends mostly on the magnitude of these derivatives, rather than on the total energy carried by the pulse. Therefore, for a smooth pulse, the plasma wave amplitude decreases considerably for increasing pulse length, contrary to the case of a rectangular pulse. As soon as the pulse intensity increases from 0 to its peak value over more than a plasma wavelength, the pulse excites hardly any wakefield, i.e. ponderomotive excitation becomes extremely inefficient. Also, the conditions for wave breaking at the front of the pulse are difficult to attain for a smooth pulse. In addition, it

can be seen that plasma wave excitation by a polynomial pulse is slightly more efficient than excitation by a Gaussian pulse. This is caused by the Gaussian pulse having longer “tails” at the front and back, which decrease efficiency. But this difference is only minor, since the qualitative difference between these two envelope shapes is small.

In experiments, plasma waves will commonly be excited by a laser pulse having a more or less fixed length, intensity, and envelope shape. On the other hand, the plasma density n_0 may vary over a wide range of values, so it comes naturally to investigate the influence of a change in n_0 leaving all other parameters constant. As observed above, the Akhiezer and Polovin threshold scales approximately with $\gamma_g^{1/2} \sim n_0^{-1/4}$ for $\gamma_g \gg 1$, but this is not the most important contribution. Far more important is the fact that for a pulse with fixed unscaled length L , the scaled length $L/\lambda_p = \omega_p L/(2\pi c)$ scales with $n_0^{1/2}$, while its scaled intensity does not change with n_0 at all. From the figures, it can be seen that for $L/\lambda_p > 1$, there is a very rapid increase of the laser intensity needed to excite a breaking plasma wave, i.e. the efficiency of the laser-plasma energy transfer drops down to (almost) zero. If we assume for the moment that for fixed n_0 the critical intensity for wave breaking U_{cr} scales with $(L/\lambda_p)^2$ (it is probably even worse than that) while according to (4.30), we find that for fixed $(L/\lambda_p)^2$, U_{cr} scales with γ_g , we find that for fixed L , $U_{cr} \sim \gamma_g (L/\lambda_p)^2 \sim n_0^{1/2}$ or worse. In other words, the higher the plasma density, the less efficient the ponderomotive excitation scheme. It has been shown that efficiency is maximum for L in the neighbourhood of $\lambda_p/2$, the so-called *resonant excitation*. This is supported by our own results which show a maximum efficiency for $0 < L < \lambda_p/2$, a rapid decline in efficiency for $L > \lambda_p$, and a negligible contribution to the plasma wave for $L > 2\lambda_p$.

As it happens, laser pulses with $L > \lambda_p$ are perfectly capable of exciting a wake wave in a plasma, only not through ponderomotive excitation as described in this chapter. The main mechanism in such cases is excitation of a fast plasma wave in the plasma through so-called *parametric instabilities*, most notably *stimulated Raman scattering*, which do not play a role for $L < \lambda_p$. Excitation by means of stimulated Raman scattering will be investigated in Chapter 6.

4.9 Behaviour of the broken wave

The simple one-fluid model discussed here is suitable for the description of the plasma wave only until it breaks. For the description of the plasma after wave breaking, the model needs to be extended as follows. The plasma electrons captured by the wake wave have a velocity v obeying $v_g < v < 1$. Since $v - v_g \lesssim \mathcal{O}(1/\gamma_g^2)$, captured plasma electrons will remain virtually immobile with respect to the wakefield, and bunch up just in front of the wave breaking point. This bunch of fast electrons will be approximated by a “sheet” of charge, located

at $\zeta = \zeta_0$, just in front of the wave breaking point. As we shall see, the charge density in the sheet will increase until a certain limit has been reached; after that the charge density will remain constant. The resulting stationary situation is depicted in Figure 4.3.

This line of reasoning is supported by the simulation results of Nagashima et al., [33] which show a sharp spike in the electron density leaving the plasma and staying closely behind the laser pulse.

The final energy of plasma electrons that are captured and accelerated as a result of wave breaking can be calculated as follows. It is assumed that in the process of wave breaking, plasma particles are actually reflected by the break in the wakefield. In a Lorentz frame moving with the pulse, plasma particles in front of the pulse have originally a velocity $-v_g$. After reflection by the break, they eventually obtain a velocity v_g with respect to the comoving Lorentz frame. From relativistic addition of velocities, the final velocity v_f is found to equal $2v_g/(1+v_g^2)$, and the corresponding γ_f equals $(1+v_g^2)/(1-v_g^2) = 2\gamma_g^2 - 1$.

We proceed to study the behaviour of an already broken wakefield with an accumulation of charge just in front of the break. With the additional contribution of the charge sheet, the differential equation for Ψ becomes:

$$\frac{\partial^2 \Psi}{\partial \zeta^2} = -\frac{\partial}{\partial \Psi} V(\Psi, U) + Q\delta(\zeta - \zeta_0), \quad (4.31)$$

where Q represents the saturated surface charge density of the sheet. We find that for both $\zeta > \zeta_0$ and $\zeta < \zeta_0$, the energy equation (4.25) still holds, while the presence of the sheet causes a jump in $\partial\Psi/\partial\zeta$ at $\zeta = \zeta_0$, i.e. a jump in the electrostatic field. Note that the electric field of the sheet is cancelled by that of the background ions for $\zeta > \zeta_0$.

The scenario of plasma wave excitation is depicted in Figure 4.3.

(I-V) A plasma oscillation is excited similarly to the scenario without the charge sheet. Its energy level is higher than the energy $V(1/\gamma_g, 0) = \gamma_g$ of the end point of the potential well.

(VI) After passage of the charge sheet at $\zeta = \zeta_0$, $\partial\Psi/\partial\zeta$ has decreased by Q , and since $\partial\Psi/\partial\zeta > 0$ for $\zeta > \zeta_0$, we find that the “kinetic energy” of the virtual particle has also decreased. Since $\Psi \approx \Psi_0(0)$ at $\zeta = \zeta_0$, the “kinetic energy” just before passage of the charge sheet is given by $\frac{1}{2}(\partial\Psi/\partial\zeta)^2 = V(\Psi_{\text{III}}(U), 0) - V(\Psi_0(0), 0)$. As long as Q is small, one still has $\partial\Psi/\partial\zeta > 0$, and “kinetic energy” larger than $\Psi_0(0)$ after passage of the charge sheet, and the growth of Q will continue. However, as soon as $Q^2 = 2(V(\Psi_{\text{III}}(U), 0) - V(\Psi_0(0), 0))$, fast electron generation, i.e. increase of Q , will stop.

(VI-VII) In the stable situation where the charge sheet prevents further wave breaking, there is a regular plasma oscillation behind the sheet.

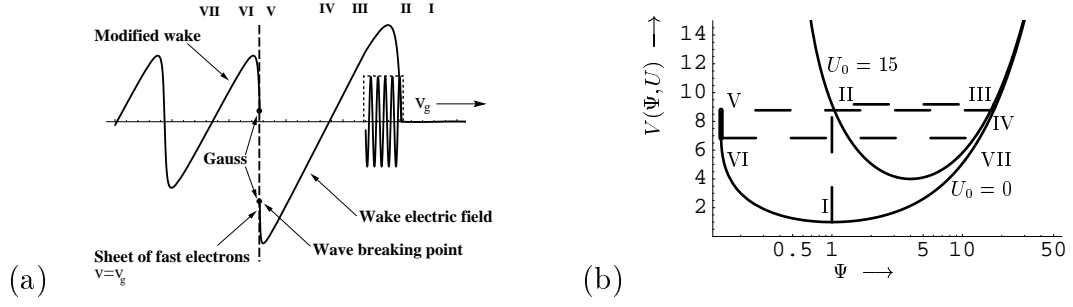


Figure 4.3: The evolution of Ψ and the total energy level of the system in the case of wave breaking and subsequent plasma electron capture, depicted in (ζ, E) -space (a) and (Ψ, V) -space (b). The laser and plasma parameters are $n/n_{cr} = 0.02$ and $U_0 = 15$.

From $Q^2 \leq 2(V(\Psi_{III}(U), 0) - V(\Psi_0, 0))$ and

$$V(\Psi_0(0), 0) = \gamma_g, \quad (4.32)$$

$$V(\Psi_{III}(U), 0) = \gamma_g^4 \left(X - v_g \sqrt{X^2 - 1/\gamma_g^6} \right), \quad (4.33)$$

where

$$X := \frac{\Psi_{III}}{\gamma_g^2} = 2 - \frac{1}{\gamma_g^2} - 2v_g \sqrt{1 - \frac{1+U}{\gamma_g^2}}, \quad (4.34)$$

we find that Q will saturate as soon as

$$Q \approx \sqrt{(U - U_{cr}) + \frac{2(U - U_{cr})}{2 + U_{cr}} + \frac{(U - U_{cr})^2}{(2 + U_{cr})^2}}, \quad (4.35)$$

defining the final charge density of the sheet for $U \geq U_0$. The total surface charge density of the generated fast electrons will be $(n_0 e c / \omega_p) Q$. (See figure 4.4.)

The behaviour of Q vs. U_0 in figure 4.4 is in good qualitative agreement with simulation results by Nagashima et al. [33].

Note that U/γ_g^2 should remain sufficiently small, otherwise wave breaking at the front of the pulse occurs, and the approximations made in the above equations are no longer valid.

4.10 Conclusions

In this chapter, plasma wave breaking and its role in fast electron production have been investigated. At wave breaking, plasma electrons get trapped in the wave itself, and bunch up just in front of the wave breaking point. There they get

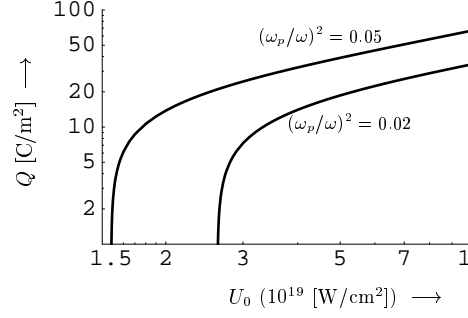


Figure 4.4: The total extracted charge density Q as a function of U_0 for $(\omega_p/\omega)^2 = 0.02, 0.05$

accelerated and emerge as fast electrons from the plasma. In general, a laser pulse moving with velocity $v_g \lesssim 1$ and corresponding Lorentz factor γ_g can generate fast electrons with energies up to $2\gamma_g^2 - 1$, which corresponds to about $1.0 * n_{cr}/n_0$ MeV. For example, from a plasma with $n_0 = 0.02n_{cr}$, electrons with an energy of up to 50 MeV can be generated, simply by injecting an intense laser pulse into it.

An overview of existing definitions for wave breaking in both cold and warm plasmas has been given, and the electric field threshold for wave breaking has been calculated in both cases for both relativistic and non-relativistic waves. It has been found that existing literature uses a number of different definitions for the phenomenon denoted as wave breaking. For a cold plasma, wave breaking is defined as intersection of the characteristic curves or planes along which the wave propagates, or equivalently the occurrence of a singularity in the plasma density. In a warm plasma, the plasma density cannot even approach a singularity because of diverging pressure terms in the advection equation. As a consequence, there is no unambiguous definition of wave breaking in this case, and the minimum value of the electrostatic field necessary for wave breaking may vary considerably depending on the definition used in its derivation.

The case of a cold relativistic wave excited by the ponderomotive force of a non-evolving laser pulse has been investigated more closely. A 1-D model has been developed to describe such excitation, from which laser intensity thresholds for the occurrence of wave breaking have been derived, both analytically and numerically. The model has then been extended to include, in a self-consistent way, the capture and acceleration of plasma electrons during wave breaking. This resulted in an analytical expression predicting the total extracted charge as a function of laser intensity. The analytical prediction for the amount of fast electrons as a function of laser intensity is found to be in qualitative agreement with the results of recent particle-in-cell simulations [33]. With only small modifications, the model can

also be used to study breaking of non-relativistic waves, or waves in a warm plasma.

The laser intensity thresholds for wave breaking as predicted by the model developed in this chapter agree with earlier results obtained through a different method [32]. The numerical simulations of wakefield excitation show that ponderomotive excitation works best for pulses with length $L \lesssim \lambda_p/2$ while its efficiency decreases sharply for $L > \lambda_p$, and there is hardly any excitation left for $L > 2\lambda_p$. Plasma wave excitation in this regime is bound to originate from different processes, such as stimulated Raman scattering.

References

- [1] C. A. Coverdale et al., Phys. Rev. Lett. **74**, 4659 (1995).
- [2] A. Modena et al., Nature (London) **377**, 606 (1995).
- [3] D. Umstadter et al., Science **273**, 472 (1996).
- [4] G. Malka et al., Phys. Rev. Lett. **79**, 2053 (1997).
- [5] A. Ting et al., Phys. Plasmas **4**, 1889 (1997)
- [6] C. I. Moore et al., Phys. Rev. Lett. **79**, 3909 (1997).
- [7] A. Pukhov and J. Meyer-ter-Vehn, Phys. Rev. Lett. **76**, 3975 (1996).
- [8] K.-C. Tzeng et al., Phys. Rev. Lett. **79**, 5258 (1997); K.-C. Tzeng, W. B. Mori, and T. Katsouleas, Phys. Plasmas **6**, 2105 (1999).
- [9] P. Volfbeyn et al., Phys. Plasmas **4**, 3403 (1997).
- [10] E. Esarey et al., Phys. Rev. Lett. **84**, 3081 (2000).
- [11] S. V. Bulanov, F. Pegoraro, A. M. Pukhov, and A. S. Sakharov, Phys. Rev. Lett. **78**, 4205 (1997); S. V. Bulanov *et al.*, Fiz. Plazmy **25**, 517 (1999) [Plasma Phys. Rep. **25**, 468 (1999)].
- [12] S. Bulanov, N. Naumova, F. Pegoraro, and J. Sakai, Phys. Rev. E **58**, 5257 (1998).
- [13] T. J. Schep, A. V. Tulupov, V. A. Bazylev, and V. V. Goloviznin, J. Phys. D **28**, 1078 (1995).
- [14] G. B. Whitham, *Linear and nonlinear waves*, John Wiley and Sons, New York, 1974.

-
- [15] Jin-Gen Wang, G. L. Payne, and D. R. Nicholson, Phys. Fluids B **4**, 1432 (1992).
 - [16] R. C. Davidson, *Methods in nonlinear plasma theory*, Academic Press, New York and London, 1972.
 - [17] T. P. Coffey, Phys. Fluids **7**, 1402 (1971).
 - [18] B. Bezzerides and S. J. Gitmore, Phys. Fluids **26**, 1359 (1983).
 - [19] A. Bergmann and P. Mulser, Phys. Rev. E **47**, 3585 (1993).
 - [20] T. Katsouleas and W. B. Mori, Phys. Rev. Lett. **61**, 90 (1988); Phys. Scr. **T30**, 127 (1990).
 - [21] N.E. Andreev *et al.*, Physica Scripta **49**, 101 (1994).
 - [22] A. I. Akhiezer and R. V. Polovin, Zh. Eksp. Teor. Fiz. **30**, 915 (1956) [Sov. Phys. JETP **3**, 696 (1956)].
 - [23] J. M. Dawson, Phys. Rev. **113**, 383 (1959).
 - [24] A. Bergmann and H. Schnabl, Phys. Fluids **31**, 3266 (1988).
 - [25] I. M. Aleshin, M. A. Drofa, and L. S. Kuz'menkov, Fiz. Plazmy **19**, 1005 (1993) [Plasma Phys. Rep. **19**, 523 (1993)].
 - [26] J. B. Rosenzweig, Phys. Rev. A **38**, 3634 (1988).
 - [27] J. B. Rosenzweig, Phys. Rev. A **40**, 5249 (1989).
 - [28] Z. M. Sheng and J. Meyer-ter-Vehn, Phys. Plasmas **4**, 493 (1997).
 - [29] W. A. Newcomb, Phys. Fluids **25**, 846 (1982).
 - [30] D. Umstadter, J. K. Kim, and E. Dodd, Phys. Rev. Lett. **76**, 2073 (1996).
 - [31] S. V. Bulanov *et al.*, Fiz. Plasmy **25**, 748 (1999) [Plasma Phys. Rep. **25**, 701 (1999)].
 - [32] Denis Teychenné, Guy Bonnaud, and Jean-Louis Bobin, Phys. Rev. E **48**, R3248 (1993).
 - [33] K. Nagashima, Y. Kishimoto, and H. Takuma, Phys. Rev. E **59**, 1263 (1999).

Chapter 5

Stimulated Raman scattering

5.1 Introduction

When a strong electromagnetic pump wave (frequency ω_0 , wave vector \mathbf{k}_0) interacts with an underdense plasma, it drives various longitudinal plasma waves having frequencies different from ω_0 . Nonlinear interaction between the pump and density perturbations in the plasma results in the excitation of electromagnetic sidebands in the EM spectrum located at the sum and difference of the pump and plasma wave frequencies. When the interaction is such, that excitation of the EM sideband excites/enhances a plasma wave, while the presence of the plasma wave enhances sideband growth in turn, the interaction process becomes unstable. Such an instability is called a *parametric instability* [1–4].

There are various types of parametric instabilities, classified according to the relation between the phase velocity v_ϕ of the plasma wave and the electron and ion thermal velocities $v_{th,e}$ and $v_{th,i}$. When the incident pump scatters off a Langmuir wave ($v_\phi \gg v_{th,e}$ so the electrons are adiabatic and the ions immobile), the instability is called *stimulated Raman scattering (SRS)*. If the plasma wave satisfies $v_\phi \approx v_{th,e}$, it will be strongly Landau damped on electrons. Then the plasma wave is no longer an eigenmode but a driven oscillation, and scattering off it is called *stimulated Compton scattering*. Finally, when the pump scatters off an ion-acoustic wave ($v_{th,i} \ll v_\phi \ll v_{th,e}$, so the electrons are isothermic and the ions adiabatic), it is called *stimulated Brillouin scattering (SBS)*.¹ In laser wakefield acceleration, the plasma is usually cold and the timescales involved are so short that ion motion can be neglected. Therefore, stimulated Raman scattering is by far the most important parametric instability in laser wakefield research, and the only one that will be dealt with in this chapter.

¹The terms Raman and Brillouin scattering have been borrowed from photon-phonon interactions in solid-state physics that show similar spectral behaviour.

In SRS, the incident pump (ω_0, \mathbf{k}_0) scatters off a Langmuir wave (ω, \mathbf{k}) , generating two EM sideband waves $(\omega_0 \pm \omega, \mathbf{k}_0 \pm \mathbf{k})$ in the process. The frequency downshifted wave is called the Stokes wave, and the upshifted wave is called the anti-Stokes wave.² Since all the EM waves involved need to be able to propagate through the plasma, SRS can only occur if $\omega_0 - \omega_p > \omega_p$, or $n_0 < n_{cr}/4$. This requirement will easily be met in underdense plasmas for which normally $n_0/n_{cr} \ll 1$ holds.

There are various types of SRS, named after the direction of the scattered EM wave with respect to the incident EM wave. One discerns Raman backscattering (RBS), side scattering (RSS), near-forward scattering (RFSS) and forward scattering (RFS). The frequencies of the associated Langmuir waves are all in the vicinity of ω_p , while the wave numbers of these waves range from about $2k_0$ for RBS to about ω_p/c for RFS. Consequently, the Langmuir wave phase velocity is lowest for RBS and highest for RFS. This means that if thermal effects come into play, the RBS Langmuir wave will feel them first. Also, the wave breaking amplitude is lowest for the RBS Langmuir wave, and highest for the RFS Langmuir wave.

In order for Raman scattering to occur in laser-plasma interaction, the incident laser pulse needs to be longer than one plasma period. This corresponds to the regime in which ponderomotive excitation of a laser wake by the laser pulse envelope ceases to play a role. However, a laser pulse can still excite a wake in this regime. The reason for that is that the role of the wake is taken over by the (fast) RFS Langmuir wave. When excited to sufficient amplitudes, this fast wave is as suitable for particle acceleration as a ponderomotively excited wake.

In this chapter, an introduction to Raman scattering theory will be given. From the dispersion relation for the coupled EM and plasma waves, the various types of Raman scattering will be derived, with their frequencies and wave numbers. Also, expressions for the growth rates of the various types will be given. Note that (almost) all the results below are taken from basic Raman theory, which implies a very long low intensity pump laser beam propagating through an infinitely extended homogeneous plasma.

5.2 Dispersion relation

From EM wave theory, the dispersion relation for the coupled EM and plasma waves can be derived using perturbation methods. Modes that are at or near a resonance in the dispersion relation will display a large growth rate and are called “resonant”. Modes that are far from such a resonance have a small growth rate and are called “non-resonant”. In our analysis we will neglect non-resonant

²The terms Stokes and anti-Stokes have been borrowed from fluorescence and Raman spectroscopy, and bear no relation to the (anti-)Stokes lines in the theory of asymptotic series.

modes, such as higher order coupled modes ($\omega_0 \pm l\omega, \mathbf{k}_0 \pm l\mathbf{k}$), $l > 1$ and integer, or modes with electrostatic sidebands, which are non-resonant in an underdense plasma.

We follow the approach of Sakharov and Kirsanov [5] to derive the dispersion relation for the ensemble of a pump laser wave and two EM sideband waves. Note that, in line with the standard practice, only the vector potentials will be normalized, while other quantities like t , \mathbf{r} , and n , are not normalized. The respective vector potentials \mathbf{A}_0 and \mathbf{A}_\pm of pump and sidebands are given by

$$\begin{aligned}\mathbf{A}_0 &= \mathbf{a}_0(\mathbf{r}, t) \exp[i(\mathbf{k}_0 \cdot \mathbf{r} - \omega_0 t)], \\ \mathbf{A}_\pm &= \mathbf{a}_\pm(\mathbf{r}, t) \exp[i(\mathbf{k}_\pm \cdot \mathbf{r} - \omega_\pm t)].\end{aligned}$$

Here, $\omega_\pm = \omega_0 \pm \omega$ and $\mathbf{k}_\pm = \mathbf{k}_0 \pm \mathbf{k}$, where ω and \mathbf{k} denote the frequency and wave vector of the Langmuir wave excited by the beating of the pump and one of the sidebands. We set $\mathbf{A}_\perp = \mathbf{A}_0 + \mathbf{A}_+ + \mathbf{A}_-$, and insert this into the wave equation for the fast part of the vector potential as derived in Chapter 2:

$$\left[\Delta - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{\omega_p^2 n_s}{\gamma_s n_0} \right] \mathbf{A}_\perp = \mathbf{0}. \quad (5.1)$$

For a circularly polarized laser pulse, we have $\gamma_s = \sqrt{1 + A_\perp^2}$. We further assume that the density perturbations are initially small, i.e. $n_s = n_0 + \tilde{n}_e$ with $\tilde{n}_e \ll n_0$. It is also assumed that $|\mathbf{A}_\pm| \ll |\mathbf{A}_0|$, so $\gamma_s \approx \gamma_0(1 + \mathbf{A}_0 \cdot (\mathbf{A}_+ + \mathbf{A}_-)/\gamma_0^2)$ where $\gamma_0 \equiv \sqrt{1 + A_0^2}$. An eikonal approximation will be used for the vector potentials, meaning that $|\nabla \cdot \mathbf{a}| \ll |\mathbf{k}_0 \cdot \mathbf{a}|$. Expanding (5.1) produces the following equations for leading order terms and first order perturbations respectively:

$$\left[\Delta - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{\Omega_p^2}{c^2} \right] \mathbf{A}_0 = \mathbf{0}, \quad (5.2)$$

$$\left[\Delta - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{\Omega_p^2}{c^2} \right] \mathbf{A}_\pm = \frac{\Omega_p^2}{c^2} \left(\frac{\tilde{n}_e}{n_0} - \frac{\mathbf{A}_0 \cdot \mathbf{A}_\pm}{\gamma_0^2} \right) \mathbf{A}_0. \quad (5.3)$$

Here, $\Omega_p = \omega_p/\sqrt{\gamma_0}$ is the plasma frequency after correction for the transverse electron quiver motion in the EM field of the pump. Note that the $\mathbf{A}_0 \cdot \mathbf{A}_\pm$ contribution at the right-hand side of (5.3) does not occur in a non-relativistic SRS analysis. This term arises from the relativistic shift in the plasma frequency due to the combined action of the pump and sideband waves. The term has the same magnitude as the first term at the right-hand side and does not vanish in the low-pump limit. Since such frequency shifts do not arise in a non-relativistic treatment, this implies that the relativistic case in the low-pump limit differs from the non-relativistic case.

The density perturbation is driven by the ponderomotive force,

$$\left(\frac{\partial^2}{\partial t^2} + \Omega_p^2 \right) \tilde{n}_e = \frac{n_0 c^2}{\gamma_0^2} \Delta (\mathbf{A}_0 \cdot \mathbf{A}_+ + \mathbf{A}_0 \cdot \mathbf{A}_-), \quad (5.4)$$

where it has been assumed that the ∇A_0^2 contribution can be neglected, i.e. the pump has infinite length and constant amplitude.

From (5.2), the well-known dispersion relation $\omega_0^2 - \Omega_p^2 = c^2 k_0^2$ for the EM pump is derived. Combining (5.3) and (5.4) and retaining only the leading order terms yields the joint dispersion relation for the first order sideband waves and the density perturbation (subscript “e” dropped):

$$\Omega_p^2 - \omega^2 = \frac{1}{2}\Omega_p^2(c^2 k^2 - \omega^2 + \Omega_p^2)\frac{A_0^2}{\gamma_0^2}\left[\frac{1}{D_-} + \frac{1}{D_+}\right], \quad (5.5)$$

where D_{\pm} are given by $D_{\pm} = c^2|\mathbf{k}_0 \pm \mathbf{k}|^2 - (\omega_0 \pm \omega)^2 + \Omega_p^2$. Resonant modes of (5.5) occur when either $D_- = 0$ (Stokes) or $D_+ = 0$ (anti-Stokes). These relations will be used below to analyze the various possible types of SRS.

It should be noted that the dispersion relation (5.5) has been derived under the implicit assumption that the pump and scattered EM waves are polarized in the same direction and scattering occurs out of the plane of polarization, i.e. $\mathbf{k}_{\pm} \cdot \mathbf{A}_0 = 0$. The reason for this is that in this case, the instability growth rate is maximum and the intensity threshold is minimum. It can also be shown that maximum growth occurs if $\Re(\omega) = \Omega_p$, \Re denoting the real part of an expression, which we assume to be the case from now on.

5.3 Scattering off electron modes

Having derived the dispersion relation (5.5), we are now ready to discuss the various possible modes of SRS off a cold electron plasma. We will treat the Stokes and anti-Stokes modes separately. For convenience, we will also express D_{\pm} in terms of the angle ϑ between the wave vectors \mathbf{k}_0 and \mathbf{k} of the EM pump and the Langmuir wave respectively: $D_{\pm} = c^2(k^2 \pm 2kk_0 \cos \vartheta) \mp 2\omega\omega_0 - \omega^2$.

Stokes wave The Stokes EM wave has frequency $\omega_- = \omega_0 - \Omega_p$, wave vector $\mathbf{k}_- = \mathbf{k}_0 - \mathbf{k}$, and satisfies $D_- = 0$, i.e.

$$c^2(k^2 - 2kk_0 \cos \vartheta) + 2\Omega_p\omega_0 - \Omega_p^2 = 0.$$

This equation can be solved for k to obtain

$$k = k_0 \cos \vartheta \pm \sqrt{k_0^2 \cos^2 \vartheta - (2\Omega_p\omega_0 - \Omega_p^2)/c^2}. \quad (5.6)$$

For k to have at least one positive real solution, one must have

$$\cos^2 \vartheta > \cos^2 \vartheta_- = \frac{2\Omega_p\omega_0 - \Omega_p^2}{c^2 k_0^2} = 1 - \frac{1 - 2\Omega_p/\omega_0}{1 - (\Omega_p/\omega_0)^2}.$$

In practice, this means that there can be scattering off the Stokes wave for $0 \leq \vartheta \leq \vartheta_-$ only, since for $\pi - \vartheta_- \leq \vartheta \leq \pi$, both solutions for k are negative. Also note that there are no real solutions for k if $\Omega_p/\omega_0 > \frac{1}{2}$, which confirms that SRS cannot exist for $n_0 > n_{cr}/4$. In the next chapter, we will use $\omega_p/\omega_0 \approx 0.15$, for which $\cos^2 \vartheta_- \approx 0.28$, i.e. $\vartheta_- \approx 58^\circ$. This guarantees that in the paraxial region, θ will be sufficiently removed from ϑ_- .

We discern the following cases:

(i) $\vartheta = 0$. In this case, we have

$$\begin{aligned} k &= k_0 \pm \sqrt{k_0^2 - (2\Omega_p\omega_0 - \Omega_p^2)/c^2} \\ &\approx k_0(1 \pm (1 - \Omega_p/\omega_0 + (\Omega_p/\omega_0)^2)), \end{aligned}$$

leading to either $k \approx 2k_0 - \Omega_p/c$ or $k \approx \Omega_p/c(1 - \Omega_p/(2\omega_0))$. In the first case, the Stokes EM wave has wave number $-(k_0 - \Omega_p/c)$, i.e. it propagates backwards with respect to the pump wave. For this reason, this case is called *Raman backscattering (RBS)*. In the second case, the Stokes EM wave has approximate wave number $k_0 - \Omega_p/c$ and propagates in the exact same direction as the pump. This case is called *Raman forward scattering (RFS)*. We shall see later that RFS also has an anti-Stokes component scattering off nearly the same Langmuir wave, while RBS only has the above Stokes component.

(ii) $0 < \vartheta < \vartheta_-$. In this case, we also have two solutions for k , one with $k > k_0 \cos \vartheta$ and one with $k < k_0 \cos \vartheta$. The direction of the Stokes EM wave associated with the first solution goes from backward to backward-sideways to sideways to forward-sideways as ϑ increases. This type of scattering is called *Raman side scattering (RSS)*. The direction of the Stokes EM wave associated with the second solution is always forward-sideways, and this case is usually disregarded as being non-resonant. However, for $\vartheta \ll \vartheta_-$ such that $k_0^2 \cos^2 \vartheta \gg (2\Omega_p\omega_0 - \Omega_p^2)/c^2$, we will show that both the Stokes and the anti-Stokes EM waves scatter from nearly the same Langmuir wave, and therefore reinforce each other. In that case (5.6) reduces to

$$k \approx \frac{2\Omega_p\omega_0 - \Omega_p^2}{2c^2 k_0 \cos \vartheta} \approx \frac{\Omega_p/c}{\cos \vartheta} \left(1 - \frac{\Omega_p}{2\omega_0}\right).$$

Since $k \ll k_0$, we find that the associated EM wave propagates nearly in the forward direction, and has approximate wave number $k_0 - \Omega_p/(c \cos \vartheta)$. Due to its similarity with Raman forward scattering, this type of scattering is called *Raman near-forward scattering (RFSS)*.

(iii) $\vartheta = \vartheta_-$. In this case, both solutions for k coincide at $k = k_0 \cos \vartheta_- = \sqrt{(2\omega_0\Omega_p - \Omega_p^2)/c^2} \ll k_0$, i.e. side scattering and near-forward scattering meet. The Langmuir wave propagates nearly sideways, and the Stokes EM wave forward-sideways.

Anti-Stokes wave The anti-Stokes EM wave has frequency $\omega_+ = \omega_0 + \Omega_p$, wave vector $\mathbf{k}_+ = \mathbf{k}_0 + \mathbf{k}$, and satisfies $D_+ = 0$, i.e.

$$c^2(k^2 + 2kk_0 \cos \vartheta) - 2\Omega_p\omega_0 - \Omega_p^2 = 0.$$

This equation can be solved for k to obtain

$$k = -k_0 \cos \vartheta + \sqrt{k_0^2 \cos^2 \vartheta + (2\Omega_p\omega_0 + \Omega_p^2)/c^2}.$$

Note that there is only one solution for k here, since the other solution has $k < 0$. The remaining solution exists for all values of ϑ , but is only considered resonant for $\vartheta \ll \vartheta_-$. We discern two cases:

(i) $\vartheta = 0$. In this case, one has $k \approx \Omega_p/c(1 + \Omega_p/(2\omega_0))$, which is nearly equal to the wave number of the Langmuir wave for the Stokes component of RFS. The associated anti-Stokes EM wave has approximate wave number $k_0 + \Omega_p/c$ and propagates in the forward direction as well. This is the anti-Stokes component of Raman forward scattering. For $\Omega_p/\omega_0 \ll 1$, both Stokes and anti-Stokes RFS are approximately resonant, and they are usually observed together, accompanied by a single Langmuir wave at $(\Omega_p, \Omega_p/c)$.

(ii) $0 < \vartheta \ll \vartheta_-$. In this case, we have

$$k \approx \frac{2\Omega_p\omega_0 + \Omega_p^2}{2c^2k_0 \cos \vartheta} \approx \frac{\Omega_p/c}{\cos \vartheta} \left(1 + \frac{\Omega_p}{2\omega_0}\right).$$

As in the case of RFS, this nearly coincides with the k corresponding to the Stokes component of RFSS for the same value of ϑ . The corresponding EM wave will also be scattered nearly forward at approximately the same angle as the Stokes EM wave for the same value of ϑ . This is the anti-Stokes component of Raman near-forward scattering. As with RFSS, both Stokes and anti-Stokes components of RFSS are approximately resonant (and can thus be observed) for $\Omega_p/\omega_0 \ll 1$ and $\vartheta \ll \vartheta_-$.

5.4 Growth rates

As stated above, SRS is an instability, which implies that the dispersion relation (5.5) must yield complex values for ω , where the imaginary part represents growth or damping in time of the EM and Langmuir waves. As the time dependence of both waves scales to leading order as $\exp(-i\omega t)$, we find that growth (damping) occurs if $\Im(\omega)$ is positive (negative). This will be done by setting $\omega = \Omega_p + \delta\omega$ where $|\delta\omega| \ll \Omega_p$, and inserting this into (5.5).

For back- and sidescatter, the dispersion relation without dissipation terms and without the anti-Stokes terms reads:

$$(\Omega_p^2 - \omega^2) [c^2|\mathbf{k} - \mathbf{k}_0|^2 - (\omega - \omega_0)^2 + \Omega_p^2] = \frac{1}{2}\Omega_p^2(c^2k^2 - \omega^2 + \Omega_p^2)\frac{A_0^2}{1 + A_0^2}. \quad (5.7)$$

We note that maximum growth occurs when the scattered EM wave is also resonant, i.e. when

$$(\Omega_p - \omega_0)^2 - c^2 |\mathbf{k} - \mathbf{k}_0|^2 - \Omega_p^2 = 0.$$

Combining this condition with (5.7) yields $\delta\omega = i\gamma$ as noted before, where the (positive real) growth rate γ is given by (note that $\Omega_p^2 = \omega_p^2 / \sqrt{1 + A_0^2}$)

$$\gamma = \frac{ck}{2\sqrt{2}} \frac{A_0}{\sqrt{1 + A_0^2}} \sqrt{\frac{\Omega_p}{(\omega_0 - \Omega_p)}} \approx \frac{ck}{2\sqrt{2}} \sqrt{\frac{\omega_p}{(\omega_0 - \Omega_p)}} \frac{A_0}{(1 + A_0^2)^{3/4}}. \quad (5.8)$$

From (5.6), it is found that γ is largest for backscattering and decreases with increasing ϑ . It is also obvious that for fixed ϑ , the mode with the larger k is clearly dominant over the mode with the smaller k , which is why the latter is usually considered non-resonant. Also note that γ reaches a maximum for $A_0 = \sqrt{2}$ and decreases as $A_0^{-1/2}$ for large A_0 , as opposed to the non-relativistic growth rate which can reach arbitrarily large values.

Growth rates for (near-)forward scattering need to be calculated in a different way from those for back- and side scattering. For the Stokes component of (near-)forward scattering, the growth rate as given by (5.8) reads $\gamma = \Omega_p A_0 / (2\sqrt{2}\gamma_0) \times \sqrt{\Omega_p / (\omega_0 - \Omega_p)}$, which is rather small compared to that for back- or side scattering. However, for a sufficiently underdense plasma, both the Stokes and anti-Stokes components of RFS are resonant simultaneously, and must both be taken into account. Then the RFS growth rate becomes:

$$\gamma = \frac{\Omega_p^2}{2\omega_0} \frac{A_0}{\sqrt{1 + A_0^2}} \approx \frac{\omega_p^2}{2\omega_0} \frac{A_0}{1 + A_0^2}. \quad (5.9)$$

This growth rate reaches its maximum for $A_0 = 1$ and decreases as A_0^{-1} for large A_0 , again different from its non-relativistic counterpart.

Lastly, McKinstrie and Bingham [6] showed that for RFS in a not very underdense plasma, the dispersion coefficient D_+ in (5.5) should be corrected for phase mismatch between the Stokes and anti-Stokes scattering processes. Using a weakly relativistic analysis, they obtain:

$$D_+ \approx 2(\omega_0 + \omega_p)(\omega - \omega_p - c^2(k - k_e)(k_0 + k_e)/(\omega_0 + \omega_p) - \delta_+),$$

where $ck_e = \sqrt{\omega_0^2 - \omega_p^2} - \sqrt{(\omega_0 - \omega_p)^2 - \omega_p^2}$ is the resonant Langmuir wave number, and $\delta_+ \approx \omega_p^4 / \omega_0^3$ is the phase mismatch coefficient. This generally leads to higher growth rates than predicted by (5.9).

5.5 Extensions

5.5.1 Thermal plasma

All the results so far have been derived under the assumption that the plasma is cold and ions are immobile. In the case of a warm plasma however, the dispersion relation (5.5) is usually derived from the Vlasov-Maxwell equations. The non-relativistic thermal dispersion relation then reads:

$$\frac{1}{\chi_e(\mathbf{k}, \omega)} + \frac{1}{1 + \chi_i(\mathbf{k}, \omega)} = c^2 k^2 A_0^2 \left[\frac{1}{D_-} + \frac{1}{D_+} \right]. \quad (5.10)$$

Here, χ_e and χ_i are the electron and ion magnetic susceptibility given by (a similar expression holds for χ_i)

$$\chi_e(\mathbf{k}, \omega) = \frac{2\omega_{pe}^2}{k^2 v_{th,e}^2} \left[1 + \frac{\omega}{k v_{th,e}} Z \left(\frac{\omega}{k v_{th,e}} \right) \right] \approx -\frac{\omega_{pe}^2}{\omega^2} \left(1 + \frac{3k^2 v_{th,e}^2}{2\omega^2} \right) + i \cdot \Im(\chi_e),$$

using $\omega/k \gg v_{th,e} = \sqrt{2T_e/m_e}$ for SRS. The function $Z(\zeta)$ is called the *plasma dispersion function*, and it is described in detail in Ref. [7]. The term $i \cdot \Im(\chi_e)$ is the contribution of Landau damping [8] to the dispersion relation. This dispersion relation is particularly useful when studying the transition from stimulated Raman scattering to stimulated Compton or Brillouin scattering, but less so for studying SRS in the cold plasma limit.

The dispersion relation (5.10) is often simplified by taking the ions to be immobile, so ω_{pi} , $v_{th,i}$ and χ_i all vanish. As for the influence of Landau damping, we note that $\Im(\chi_e) = 2\sqrt{\pi}(\omega_{pe}^2/\omega^2)\zeta^3 \exp(-\zeta^2)$ where $\zeta = \omega/(k v_{th,e})$. In the case of SRS, the pump EM wave scatters off a Langmuir wave, i.e. $\zeta \gg 1$ and the Landau contribution is negligible. With these simplifications, the dispersion relation reads:

$$\omega_k^2 - \omega^2 = \frac{\omega_p^2 c^2 k^2 A_0^2}{2} \left[\frac{1}{D_-} + \frac{1}{D_+} \right],$$

where $\omega_k^2 = \omega_p^2 + (3/2)k^2 v_{th}^2$ is the Bohm-Gross frequency. Note that in the cold plasma limit this relation is similar to Eq. (5.5).

5.5.2 Finite pulse length

The previous results are all obtained in the case that the incident pump pulse is much longer than anything else, so the spatial and temporal dependence of its envelope can be neglected. With the advent of ultrashort laser pulses, this assumption need not be valid any longer. For that reason the RBS and RFS growth rates for finite pulses are considered here.

Since we are dealing with a pulse of finite length, it makes more sense to use the term *growth* instead of growth rate (which is essentially the growth per unit of time or length). We define the comoving coordinate $\xi = x - ct$ and assume that the pulse envelope depends only on ξ : $a_0 = a_0(\xi)$. An analysis by Sakharov and Kirsanov [5] yields that starting from a seed perturbation n_1 , the RBS Langmuir wave grows according to $n(\xi) = n_1 \exp[\int_{\xi}^{\infty} q(\xi) d\xi]$. In the low intensity limit,

$$\begin{aligned} q(\xi) &= \Re(\sqrt{(\omega_0 \Omega_p / (8c^2))(a_0^2 / \gamma_0^2) - \Delta^2(\xi) / 4}), \\ \Delta(\xi) &= (k - 2k_0(\xi) + \Omega_p / c) / 2. \end{aligned}$$

For very long pulses with (nearly) constant amplitude, the ξ -dependence vanishes from the above expressions. In that case, they reduce to the familiar expression for RBS growth: $n(\xi) \sim \exp(-\gamma \xi / c) \sim \exp(\gamma t)$, where γ denotes the long pump growth rate given by (5.8). Also note that in case of a ξ -dependent pump wave number $k_0(\xi)$, RBS does not continue to grow indefinitely, but saturates as soon as the pump and RBS Stokes waves move out of the region of interaction.

For RFS, the picture is different. Again starting from a seed n_1 , the RFS Langmuir wave grows as

$$n(\xi, t) = n_1 \exp[(4(\xi + ct) / c^2 \int_{\xi}^{\infty} \Gamma_0^2(\xi') d\xi')^{1/2}] \propto \exp[\alpha(\xi) t^{1/2}],$$

which is fundamentally different from the $\exp(\gamma t)$ growth seen in the long pump limit. Also note that, since the pump and the (anti-)Stokes waves move in the same direction at approximately the same speed and beat continuously, the growth at a given value of ξ is also a function of time. Far behind the pulse, where $a_0(\xi) = 0$, the RFS instability growth behaves as $\exp[(4t/c \int_{-\infty}^{\infty} \Gamma_0^2(\xi') d\xi')^{1/2}]$. In both expressions, Γ_0^2 is given by $\Gamma_0^2(\xi) = (\Omega_p^4 / (8\omega_0^2))(A_0^2 / \gamma_0^2)$.

5.5.3 Frequency mismatch

In standard SRS analysis, one starts from the assumption that the pump is monochromatic and the plasma homogeneous. As a result, there is perfect matching of the frequencies and wave numbers of the pump, (anti-)Stokes, and Langmuir waves. In more general circumstances, this does not need to be the case at all. More often, the frequencies of the waves involved match only in a region close to the location at which the (anti-)Stokes and Langmuir waves are created. This happens, for example, for SRS in a region of non-homogeneous plasma density, or if the pulse carrier frequency changes along the pulse. Thus, while the pump generates (anti-)Stokes and Langmuir waves along all of its length, the growth of a particular combination of EM and Langmuir waves is confined to the region of (approximate) frequency matching. The instability growth stops as soon as

the waves move out of this region. Therefore, frequency mismatch limits the growth of SRS by a long pump pulse. It should be noted though, that frequency mismatch plays a lesser role for (very) short pulses, since in that case the interaction region may well be limited by the pulse length itself instead of the region of frequency matching.

The effect of frequency mismatch on RBS has been examined in the WKB approximation by Liu, Rosenbluth, and White [9]. Following their approach, we define the wavenumber mismatch

$$K(\xi) = k_0(\xi) - k_-(\xi) - k(\xi).$$

Then the peak amplification of an RBS seed \tilde{n}_0 is given by

$$\tilde{n} = \tilde{n}_0 \exp \left(\frac{2\pi\gamma^2}{vv_-|K'|} \right),$$

where v and v_- are the group velocities of the Langmuir wave and the Stokes wave respectively, γ is the standard RBS growth rate given by (5.8), and $K' = \partial K/\partial \xi$. Intuitively, this corresponds to an interaction length of $2\pi\gamma/(vv_-|K'|)$, i.e. the length necessary for the frequency mismatch to become larger than the growth rate γ . Since the bandwidth of the scattered EM wave is roughly equal to γ , RBS growth stops as soon as the frequency mismatch is larger than γ . We also see that the RBS growth tends to infinity when $K' \rightarrow 0$, since this implies that the length of the interaction region becomes infinite again.

We treat two instances of frequency mismatch in more detail. The first case is that of a non-homogeneous plasma density. Suppose the plasma density is given by $n(x) = n_0(1 + x/L)$. Then the frequency mismatch occurs because the plasma frequency ω_p changes, while the pump frequency ω_0 does not. The frequency detuning K' is given by

$$K' \approx \frac{\omega_{p0}^2}{6Lk_0v_{th}^2},$$

and the total RBS growth of a seed \mathbf{A}_s along the whole pump is given by

$$\mathbf{A}_- = \mathbf{A}_s \exp(A_0^2 k_0 L/2),$$

where A_0 denotes the pump pulse intensity. The threshold for RBS growth is then given by $A_0^2 k_0 L > 1$. This suggests that RBS growth for a long pulse can be suppressed by a steep density gradient. See Ref. [9] for details.

It should be noted that the above method for calculating RBS growth in the presence of a plasma density gradient only works for pulses that are long compared to the scale length L . In the next chapter, simulation results will be presented for which the pulse length is much shorter than L , so the density can be considered constant along the pulse and RBS growth is restricted by the pulse length rather than by frequency mismatch.

The second case is that of a pulse having a non-constant carrier frequency $\omega_0(\xi)$ and wavenumber $k_0(\xi)$, which is called *chirp*. For constant $|\partial k_0/\partial \xi|$ and constant plasma density (plasma frequency), it can be derived [5] that RBS grows as

$$\mathbf{A}_- \sim \exp \left(\frac{1}{2} \left| \frac{\partial k_0}{\partial \xi} \right|^{-1} \frac{\omega_0 \Omega_p}{c^2} \frac{A_0^2}{\gamma_0^2} \right).$$

The threshold for RBS growth is then given by

$$\left| \frac{\partial k_0}{\partial \xi} \right| < \frac{1}{2} \frac{\omega_0 \Omega_p}{c^2} \frac{A_0^2}{\gamma_0^2} = \frac{1}{2} \frac{\omega_0 \omega_p}{c^2} \frac{A_0^2}{(1 + A_0^2)^{3/2}},$$

which suggests that RBS growth can also be suppressed by chirping the laser pulse's carrier frequency. The role of chirp in RBS and RFS growth rates will receive an extensive treatment in the next chapter, especially its role in enhancing and/or suppressing RBS.

References

- [1] J. F. Drake *et al.*, Phys. Fluids **17**, 778 (1974).
- [2] William L. Kruer, *The Physics of Laser Plasma Interactions* (Addison-Wesley, Reading MA, 1988).
- [3] C. S. Liu and V. K. Tripathi, *Interaction of electromagnetic waves with electron beams and plasmas* (World Scientific, Singapore, 1994).
- [4] D. W. Forslund, J. M. Kindel, and E. L. Lindman, Phys. Fluids **18**, 1002 (1975).
- [5] A. S. Sakharov and V. I. Kirsanov, Phys. Rev. E **49**, 3274 (1994).
- [6] C. J. McKinstrie and R. Bingham, Phys. Fluids B **4**, 2626 (1992).
- [7] B. D. Fried and S. D. Conte, *The Plasma Dispersion Function* (Academic Press Inc., New York, 1961).
- [8] L. D. Landau, J. Phys. USSR **10**, 25 (1946).
- [9] C. S. Liu, M. N. Rosenbluth, and R. B. White, Phys. Fluids **17**, 1211 (1974).

Chapter 6

Enhancing fast electron production through suppression of Raman backscattering

The effect of Raman instabilities on the production of fast electrons in laser-plasma interaction has been investigated for laser intensities well above the electron trapping threshold. The results of one-dimensional particle-in-cell simulations show that in this regime the presence of Raman backscattering (RBS) hampers fast electron production, and that its suppression increases the yield of high-energy electrons (> 15 MeV). Such suppression has been realized either through deletion of all backscattered radiation from the simulations or through direct stimulation of Raman forward scattering (RFS). An increased high-energy electron yield has been observed for both methods. In addition, the influence of various laser and plasma parameters on the production of highly energetic electrons has been investigated. For each parameter, its influence on the yield of high-energy electrons can be explained from the way it affects the balance between RBS and RFS excitation in laser-plasma interaction.

This chapter will be published under the same title, by R.M.G.M Trines, L.P.J. Kamp, T.J. Schep, F.W. Sluijter, W.P. Leemans, and E.H. Esarey.

6.1 Introduction

Several recent experiments [1–7] on the interaction of intense laser pulses with underdense plasmas have demonstrated the production of energetic electrons in the self-modulated regime of the laser wakefield accelerator (LWFA) [8]. The resulting electron bunches are characterized by high charge (up to 10 nC), sub-ps duration, and an exponential energy distribution with a mean energy of tens of

MeV, and have a wide range of applications [9–12].

In this chapter, we investigate the role of RBS in high-energy electron production and the interplay between RBS and RFS in the high-intensity regime of a self-modulated LWFA [13]. In previous work [5,14–16] it has been demonstrated that the presence of RBS leads to the production of mildly energetic electrons. Since these electrons are easily trapped and accelerated by the laser wake, it has been argued that their presence, and therefore the presence of RBS, improves the yield of high-energy electrons. This has been confirmed for laser intensities slightly above the threshold for plasma electron trapping. In contrast to this, we will show that for laser intensities well above the trapping threshold, the converse is true: the presence of RBS can be detrimental to high-energy electron production. It will be demonstrated that for high laser intensities the level of RBS-induced electron trapping increases to the extent that the laser wake is severely damped. This decreases the number of electrons that are accelerated to truly high energies, even though the yield of mildly energetic electrons increases. Consequently, suppressing RBS in the high-intensity regime results in a larger plasma wake and a larger yield of highly energetic electrons. Suppression of RBS has been achieved either by periodically removing all backward-going electromagnetic waves from the simulation, or by stimulating the growth of RFS. We propose and demonstrate, via one-dimensional (1-D) particle-in-cell (PIC) simulations, an experimentally realizable RBS suppression method that uses stimulated growth of RFS by seeding the RFS Stokes wave. This results in a higher level of RFS, a larger fast wake, a lower level of RBS, and a larger amount of high-energy electrons.

In addition we investigate the way in which various laser and plasma parameters such as plasma density profile, pulse envelope shape, and laser frequency chirp, influence the production of energetic electrons. From our simulation results, it will be shown that the effect of each parameter on fast electron yield can be explained from the way it influences the balance between RBS and RFS excitation in laser-plasma interaction. Any parameter that favours RFS growth and/or suppresses RBS growth will also increase the yield of high-energy electrons. Conversely, a parameter that enhances RBS growth and/or hampers RFS growth will decrease the production of highly energetic electrons. This allows us to predict or explain the effect of a number of laser and plasma parameters on high-energy electron yield.

6.2 Effect of Raman scattering on electron yield

According to basic Raman scattering theory [17–20], RBS is a three-wave interaction, in which the incoming laser light (carrier frequency ω_0 , wave number k_0 , peak amplitude $E_0 = (m_e \omega_0 c / e) a_0$) decays into a backscattered electromagnetic

(EM) wave $(\omega_0 - \omega_p, -(k_0 - k_p))$ and a slow Langmuir wave $(\omega_p, 2k_0 - k_p)$. Here, $\omega_p = \sqrt{n_0 e^2 / (\varepsilon_0 m_e)}$, $k_p = \omega_p / c$, and n_0 is the unperturbed plasma electron density. The Langmuir wave phase velocity is approximately $\omega_p c / (2\omega_0) \ll c$ for an underdense plasma. RBS is characterized by a large growth rate. RFS on the other hand has a much smaller growth rate and involves four waves: the laser light decays in two forward-scattered EM waves, a Stokes wave $(\omega_0 - \omega_p, k_0 - k_p)$ and an anti-Stokes wave $(\omega_0 + \omega_p, k_0 + k_p)$, and a fast Langmuir wave (ω_p, k_p) with phase velocity $\sim c$ provided that $\omega_p \ll \omega_0$. Because of its high phase velocity, the RFS Langmuir wave is well suited for accelerating trapped electrons to high energy [21]. At sufficiently high laser intensities, breaking of this wave contributes heavily to the trapping and acceleration to high energies of background plasma electrons [3,21,22]. The RBS Langmuir wave has a much lower wave breaking amplitude because of its low phase velocity, and thus requires a much lower laser intensity to break. Also, its high growth rate ensures that it is repeatedly pushed to breaking conditions. Breaking of this wave only yields low-energy electrons (up to several 100 keV). When injected into the RFS Langmuir wave, part of these electrons may be accelerated to higher energies through an RBS-RFS two-stage acceleration mechanism [5,14–16]. This leads to an improved yield of energetic electrons at lower laser intensities and a decrease of the intensity threshold for electron trapping.

When the laser intensity is increased well beyond the trapping threshold, the following phenomena occur: (i) the laser intensity is such that the plasma wake (i.e. the fast RFS plasma wave) itself starts to break, resulting in large scale particle trapping even in case RBS were absent, (ii) RFS growth is sufficiently large to cause RBS and RFS to overlap in space and time even for femtosecond pulses, and (iii) RBS growth is sufficiently large to cause excessive amounts of plasma electrons to be trapped by the wake. Once trapped, these electrons will cause moderate to heavy beam loading and damping of the wake. Note that if RBS and RFS do not overlap, electron injection into the laser wake can only occur through a multi-stage acceleration mechanism involving backscatter, sidescatter, near-forward and forward scatter [15,23]. In a 1-D setting, sidescatter and near-forward scatter are absent and wake damping as a result of massive electron injection can only be observed if RBS and RFS overlap in space and time. For this reason, we have made sure that such overlap does occur in the regime under investigation. The ultimate effect of wake damping on high energy electron production is a marked decrease in the number of electrons that get accelerated to truly high energies, even though the total number of trapped electrons increases. This leads us to the speculation that the yield of high-energy particles in laser-plasma interaction can be improved if we can somehow suppress the increase in RBS that comes with an increased pump pulse intensity.

We achieve such suppression using two different methods. The first is a purely numerical method in which all backward-going radiation is removed periodically

from the simulation during a limited period while leaving the forward-going radiation untouched. This method has specifically been employed to isolate the effect of RBS on high-energy electron production. The second is RFS seeding of the pulse, which has been inspired by the notion that stimulation of RFS growth leads to a decrease in RBS through mode coupling between the RBS and RFS Langmuir waves [24]. In addition, such stimulation will lead to a larger amplitude laser wake that will be less influenced by RBS-induced beam loading.

An experimentally realizable method for RFS stimulation consists of adding a “Stokes satellite” to the main laser pulse [25], i.e. a laser pulse having the same duration and a similar envelope as the main pulse, but at the Stokes frequency $\omega_0 - \omega_p$, and peak amplitude of several percent of the peak amplitude of the laser EM field. The initial level of plasma wake excited by this method is similar to that of beat-wave excitation [21]. However, now the level of the Stokes wave is greatly below that of the pump and the laser-plasma interaction is dominated by RFS. This has the advantage over beat wave excitation that the accuracy requirements on the pump, seed, and plasma frequencies are less stringent. Also, a small level of Stokes laser seed is easier to realize experimentally than a pulse with two spectral lines of equal intensity. As we will show below, the amount of RBS found in our simulations decreases with increasing satellite amplitude.

Everett et al. [24] have presented numerical and experimental results on how a fast phase velocity plasma wave can suppress a slow plasma wave, as well as a theoretical explanation in which the amplitudes of the waves are assumed to be slowly varying compared to the plasma period. In the linear regime, the amplitude $\delta n_s/n_0$ of the normalized density perturbation of the slow (e.g. RBS) plasma wave is reduced by the factor $1 - (\delta n_f/n_0)^2(\omega_0/\omega_p)^2$, where $\delta n_f/n_0$ is the normalized density perturbation associated with the fast (e.g. RFS) plasma wave. This predicts suppression of the RBS wave when $\delta n_f/n_0 \simeq \omega_p/\omega_0$. In seeded RFS, the initial amplitude of the RFS plasma wave is given approximately by beat wave theory [21], which gives in the linear regime $\delta n_f/n_0 \simeq a_0 a_1 \omega_p \tau / 4$, where $c\tau = ct - x$ is the distance behind the head of the laser pulses, and a_0, a_1 are the scaled peak amplitudes of the pump and seed pulse, respectively. This predicts RBS suppression for $a_0 a_1 \simeq 4/(\omega_0 \tau)$. If RBS is to be suppressed within a single plasma period, i.e. $c\tau = \lambda_p$, this requires a wave amplitude product of $a_0 a_1 \simeq (2/\pi)(\omega_p/\omega_0)$. Below, we will present simulations in which complete RBS suppression is reached for $a_0 a_1 \sim 0.1$ at $\omega_p/\omega_0 \approx 0.15$, in approximate agreement with this analytical prediction.

It is stressed that RBS suppression by RFS works for both lower ($a_0 \lesssim 1.0$) and higher ($a_0 \gtrsim 1.5$) laser intensities but is only advantageous at higher intensities for the following reason. In a regime of low laser intensity, the growth of both RBS and RFS is also rather low. One consequence of this is that the laser wake (RFS Langmuir wave) will no longer break, so it will only accelerate electrons that are injected into it by the RBS Langmuir wave. Another consequence is that the

amount of electrons injected into the wake by RBS is too low to cause considerable wake damping. Suppression of RBS in the low-intensity regime will not lead to a significantly larger wake, but will decrease the number of plasma electrons eligible for acceleration by the wake. The net result is that RBS suppression at lower laser intensities hurts fast electron production, in agreement with earlier results [5,14–16] and fully supported by our simulation results below.

After having established that RBS reduces the production of high-energy electrons for laser pulse intensities well above the trapping threshold, while RFS improves it by both enhancing the laser wake and suppressing RBS, we are now ready to investigate the role of various laser and plasma parameters on both RBS and RFS growth, and the consequences of this influence for high energy electron yield. This will be done in the next section.

6.3 Effects of laser and plasma parameters on RBS and RFS

As known from previous research, a number of laser and plasma parameters influence the growth of RBS and/or RFS. For example, the effect of plasma inhomogeneity on RBS has been studied by Liu *et al.* [26]. They found that for a very long pump pulse moving up a slope in the density profile, the RBS growth rate calculated in the WKB limit decreases with increasing slope. However, this result has been derived under the assumption that the pump pulse is very long, and cannot be applied to RBS produced by ultrashort pulses. The effect of asymmetric pulse envelopes on RBS has been studied by Coverdale *et al.* [27], who showed that an asymmetric pulse leads to an asymmetric frequency spectrum, thus shifting the balance between RBS and RFS growth.

A parameter that recently received a lot of attention as a means for controlling both the growth of Raman instabilities and the yield of high-energy electrons is laser chirp, i.e. change of carrier frequency through the laser pulse. The chirp is called *positive* if the carrier frequency increases from front to back, *negative* otherwise. In experiments, chirp is usually introduced by detuning the double-pass compressor grating in the Chirped Pulse Amplification process [28]. The influence of laser chirp on RBS growth has been studied analytically by Sakharov and Kirsanov [20], and both analytically and experimentally by Faure *et al.* [29]. Both investigations yielded that the RBS growth depends on the magnitude, but not the sign, of the chirp. In addition, Sakharov and Kirsanov found that RBS is suppressed by the chirp if $\omega_0^{-1}|\partial\omega_0/\partial x| > \frac{1}{2}\omega_p/(c\sqrt{\gamma_0})a_0^2/\gamma_0^2$, where $\partial\omega_0/\partial t = -(\omega_0/k_0)\partial\omega_0/\partial x$ represents the chirp and $\gamma_0 = \sqrt{1 + a_0^2}$. The influence of chirp on RFS has also been investigated by Mori [32] and Schroeder *et al.*, [33]. In both cases it has been found that a positive (negative) chirp enhances (reduces) RFS

growth. At this point, it should be stressed that the effect of chirp on both RBS and RFS depends on the quantity $\partial\omega_0/\partial x$, rather than on the relative frequency change $\delta\omega/\omega_0$ along the whole pulse length.

Starting from an unchirped pulse with a given length and bandwidth, linear laser chirp can be applied in essentially two ways: pulse length preserving and bandwidth preserving. If the pulse length is preserved, then the fast phase ψ at $x = 0$ of the laser EM wave is given by $\psi(t) = \omega_0 t + at^2/(2t_p^2)$, where t_p denotes the root-mean-square duration of the pulse. It then follows that the bandwidth of the pulse increases by a factor $\sqrt{1+a^2}$. If the bandwidth is preserved, then ψ at $x = 0$ is given by $\psi(t) = \omega_0 t + (a/(1+a^2))t^2/(2t_p^2)$. This causes the pulse duration to increase by a factor $\sqrt{1+a^2}$. Confusingly, the quantity a/t_p^2 is often referred to as “chirp” in *both* cases. If we compare the gradient of the carrier frequency for both cases, we find that $\partial\omega/\partial t = \partial^2\psi/\partial t^2 = a/t_p^2$ for pulse length preserving chirp, while this rate equals $(a/(1+a^2))/t_p^2$ for bandwidth preserving chirp. This means that in the first case, the gradient can reach arbitrarily large values, while in the second case, the rate reaches its maximum (minimum) value of $\pm 1/(2t_p^2)$ for $a = \pm 1$, and quickly drops off to 0 for $|a| > 1$. For a more detailed discussion on laser chirp, see for example Malinovsky and Krause [31].

Since the two types of chirp have such different characteristics, they can also be expected to have very different effects on RBS and RFS growth. Pulse length preserving chirp is characterized by a large value for $\partial\omega_0/\partial x$, which leads to a significant reduction or even total suppression of RBS growth and a marked increase (decrease) of RFS growth for positive (negative) chirp. Bandwidth preserving chirp on the other hand displays only a small value for $\partial\omega_0/\partial x$ and is mainly characterized by pulse length increase, especially for large $|a|$. Due to the longer pulse-plasma interaction length, this results in a much increased level of RBS growth for this type of chirp, regardless of sign. This effect dwarfs any effect of the carrier frequency gradient, and as a result, the main effect of bandwidth preserving chirp is an increase in RBS growth independent of the sign of the chirp.

We wish to stress the point that in experiments, a laser pulse is virtually always chirped by passing it through an appropriate set of dispersive optics, so its power spectrum and thus its bandwidth are more or less preserved. This way, chirping a pulse is more a method to increase the pulse length than a method to induce a carrier frequency gradient inside the pulse. As a consequence, this gradient is nearly always small for chirped pulses in experimental situations.

The effect of chirp on fast electron yield is as follows. Since pulse length preserving chirp suppresses RBS, its effect on fast electron yield relies completely on its effect on RFS. Therefore, a positive (negative) chirp of this type is predicted to enhance (reduce) fast electron yield. The predicted effects on both RFS and fast electron yield have been observed in simulations by Dodd and Umstadter [30]

focusing on the effects of this type of chirp. Bandwidth preserving chirp on the other hand has an entirely different influence on fast electron yield. As the sign of this type of chirp hardly has an effect on either RBS or RFS, it is not expected to influence fast electron yield either. Earlier experimental results [11] seemed to indicate that a positive bandwidth preserving chirp would increase the electron yield, while a negative chirp would decrease it. However, a recent paper by Leemans *et al.*, [7] shows that any experimentally observed variation in electron yield previously attributed to chirp, is mainly the result of asymmetric pulse envelope modification occurring in the optics that also cause the chirp in the pulse. It is further presumed that (the sign of) bandwidth preserving chirp has hardly any effect because of the small carrier frequency gradient for this type of chirp, in agreement with theoretical predictions.

In conclusion, the following can be said about the effect of chirp on fast electron yield. Pulse length preserving chirp does have a significant effect on fast electron yield, but this type of chirp is usually not used in experiments. Therefore, the claims by Dodd and Umstadter [30] (who used pulse length preserving chirp) that their results are in agreement with experiments are unfounded. On the other hand, bandwidth preserving chirp as used in experiments does neither significantly suppress RBS nor significantly influence RFS. Its only important effect is an increase of RBS, and thus a decrease of fast electron yield, with the magnitude of the chirp. All in all, chirping a laser pulse should not be expected to increase fast electron yield in experiments.

In the next section, simulation results will be presented that focus on the connection between the influence of a parameter on RBS or RFS growth for laser intensities well beyond the electron trapping threshold, and the influence of that parameter on energetic electron production. We have investigated the effects of plasma density profile, pulse envelope modification, and both bandwidth and pulse length preserving chirp. In each case, we shall show that if a parameter increases RFS and/or decreases RBS, this will increase the electron yield, while if a parameter decreases RFS and/or increases RBS, this will decrease the electron yield.

6.4 Simulations

This section is organized as follows. First, the setup of our simulations is explained. In the next two subsections, the respective roles of RBS and RFS in the production of high-energy electrons for higher pulse intensities are investigated. In the case of RFS, we will concentrate mostly on its role in the suppression of RBS, since its role in enhancing the wakefield is already well-known [21,25]. In the remainder of this section, the effect of plasma density profile, laser envelope shape, and laser chirp (both pulse length and bandwidth preserving chirp) on

fast electron production are investigated. It is argued that the effect of each parameter on fast electron yield can be explained from their influence [27,29] on the balance between RBS and RFS growth during pulse-plasma interaction. This knowledge can be used to tune those parameters such that RBS growth is minimized and fast electron production is maximized.

6.4.1 Simulation setup

For our simulations, the 1-D version of the code XOOPIC [34] has been used. (For a detailed description, see Appendix A.) This code uses Yee's method to integrate the time-dependent Maxwell equations [35]. In the simulations, a laser pulse has been launched onto a slab of underdense plasma. For the simulations with pulse length preserving chirp (see below), we have used a constant plasma density $n_0 = 1.116 \cdot 10^{19} \text{ cm}^{-3} = 0.01 n_{cr}$, to match earlier simulations [30] using this type of chirp. For the other simulations, laser and plasma parameters have been chosen to match those used in experiments by Leemans *et al.* [7]. In those, two different plasma electron density profiles have been employed. The first is a rather flat-top profile given by $n(x) = n_0 \exp(-x^6/l^6)$ where $l = 539 \text{ } \mu\text{m}$ and $n_0 = 0.022 n_{cr} = 3.8 \cdot 10^{19} \text{ cm}^{-3}$. The second profile is Gaussian, $n(x) = n_0 \exp(-x^2/(2l^2))$ where $l = 530 \text{ } \mu\text{m}$ and $n_0 = 0.017 n_{cr} = 2.9 \cdot 10^{19} \text{ cm}^{-3}$. For both profiles, x denotes the longitudinal coordinate, and $n_{cr} = \varepsilon_0 m_e \omega_0^2 / e^2$ is the critical density for the propagation of an EM wave having carrier frequency ω_0 . The ions are treated as an immobile, charge neutralizing background.

Laser pulses are brought into the simulation through a time-dependent boundary condition for the electric field at the left edge of the simulation window, and move from left to right. The laser pulse has a carrier wavelength $\lambda_0 = 800 \text{ nm}$, and linear polarization with $E_z = 0$. A number of envelope shapes has been used, all derived from a standard Gaussian envelope which we will refer to as a *type 0* envelope. This type of envelope is described by $E_0 \exp(-t^2/(2t_p^2))$, where $E_0 = (m_e \omega_0 c / e) a_0$ denotes the peak amplitude, and t_p the pulse time. The standard envelope has $t_p = 30 \text{ fs}$, so the pulse has a full width at half maximum (FWHM) of 50 fs. We either use this envelope as it is, or modify it by adding either bandwidth preserving chirp, pulse length preserving chirp, or "skew" (non-symmetric deformation) to it. This results in the following three types of time-dependent boundary conditions, applied at a time when the left edge of the simulation window corresponds to a large negative value of x (\Re denotes the real part of an expression).

(i) Chirped Gaussian pulse with constant bandwidth:

$E_y(t) = \Re(E_0 / \sqrt[4]{1+a^2} \exp[-(1+ia)t^2/(2(1+a^2)t_p^2)]) \exp(-i\omega_0 t)$, where a/t_p^2 denotes the group velocity dispersion. Note that the pulse length increases with a factor $\sqrt{1+a^2}$.

(ii) Chirped Gaussian pulse with constant pulse length:

$E_y(t) = \Re(E_0 \exp[-(1 + ia)t^2/(2t_p^2)] \exp(-i\omega_0 t))$, where a has the same definition as before. In this case the bandwidth increases with a factor $\sqrt{1 + a^2}$.

(iii) Unchirped pulse with deviation from Gaussian envelope:

$E_y(t) = \Re(E_0 \exp[-t^2/(t_p^2(1 + bt/\sqrt{t_0^2 + t^2}))] \exp(-i\omega_0 t))$, where $t_0 = 1 \mu\text{m}/c$ by definition, and b is a dimensionless parameter to control the “skewness” of the pulse, $|b| < 1$. For $b < 0$, the pulse has a steep front and a gentle back, for $b > 0$ it is the other way around. Sample envelopes for $t_p = 50$ fs and $b = \pm 0.8$ are shown in Fig. 6.1.

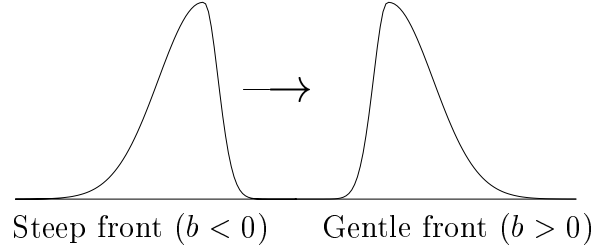


Figure 6.1: Sample “skewed” pulse envelopes for negative/positive b . The arrow denotes the direction of propagation.

For all envelope types, a_0 takes values between 0.5 and 3.0. Other parameters are $t_p = 30$ fs and $a = 0, \pm 1, \pm 5.9$ for the type (i) pulse, and $t_p = 80$ fs and $a = 0, \pm 15.4$ for the type (ii) pulse (to match the pulse in Ref. [30]). For the type (iii) pulses used in [7], one has $t_p = 50$ fs, $b = -0.52$ for the steep-front pulse and $t_p = 52$ fs, $b = 0.58$ for the gentle front pulse. For $a = 0$ ($b = 0$) all envelope shapes reduce to the standard type 0 (Gaussian) envelope.

One should note that the total energy in a type (iii) pulse varies with b as $(1 - 0.12b^2)$, whereas the total energy in type (i) or (ii) pulses is independent of a . However, for $b = 0.58$ (-0.52), the pulse energy is only 4.5% (3.5%) lower than for $b = 0$, so the differences are well within the accuracy of the values for a_0 as taken from the experiments.

A moving window has been used to follow the pulse. The simulation has been continued for either 5.33 ps or 12 ps depending on the width of the plasma slab, to allow the pulse to completely traverse the plasma. A simulation box of 0.2 mm with 5120 cells has been used (cell size 39 nm, i.e. 20-25 cells per laser wavelength), and 32 particles per cell at peak density. At the boundaries, absorbing boundary conditions have been applied, so no reflected radiation will interfere with the simulation.

The numerical removal of all left-going, i.e. backward-going, radiation mentioned above has been implemented as follows. In 1-D, the transverse EM fields can be separated into left-going $(\mathbf{E}_l, \mathbf{B}_l)$ and right-going $(\mathbf{E}_r, \mathbf{B}_r)$ parts, provided

the plasma density changes on a length scale much longer than the laser or plasma wavelength. One then has $\mathbf{E}_\perp = \mathbf{E}_l + \mathbf{E}_r$ and $\mathbf{B}_\perp = \mathbf{B}_l + \mathbf{B}_r$, where $\mathbf{E}_l = \frac{1}{2}(0, E_y - cB_z, E_z + cB_y)$, $\mathbf{E}_r = \frac{1}{2}(0, E_y + cB_z, E_z - cB_y)$, $\mathbf{B}_l = \frac{1}{2}(0, B_y + E_z/c, B_z - E_y/c)$, and $\mathbf{B}_r = \frac{1}{2}(0, B_y - E_z/c, B_z + E_y/c)$. This separation is accurate up to $\mathcal{O}(\omega_p^2/\omega_0^2)$, which is sufficient to achieve a near-complete RBS suppression while leaving the right-going fields virtually unaffected. In those simulations where field suppression has been used, the transverse EM fields have been replaced by their right-going parts once every 100 time steps during the first 2 ps. After 2 ps, field suppression has been discontinued since RBS is negligible at that stage and continuation of the suppression may eventually affect the right-going fields. This is because the transverse components of \mathbf{E} and \mathbf{B} are known at different points of the Yee mesh [35] and have to be interpolated for the calculation of the right-going fields, which introduces numerical errors.

Absent in our simulations are 2-D effects such as Raman side scatter, self-focusing, and direct laser acceleration [36], which could play an important role in certain regimes. Also note that the longitudinal beam loading of a particle in 1-D does not depend on its Lorentz factor γ , whereas its beam loading scales with $1/\gamma^2$ in 3-D [37]. However, our simulations show that the wake damping is mainly caused by slow particles ($\gamma^2 < 2$), so this issue does not affect the validity of our results.

Several benchmarks have been performed to establish the reliability of the code XOOPIC. In the weakly non-linear regime $a_0 \lesssim 1$, both the laser pulse and the scattered EM waves have to satisfy the dispersion relation $\omega^2 - \Omega_p^2 = c^2 k^2$, where $\Omega_p^2 = \omega_p^2 / \sqrt{1 + a_0^2}$ is the relativistically corrected plasma frequency. Furthermore, the backscattered wave needs to satisfy $|\omega_- - \omega_0| = \Omega_p$ and $|k_- - k_0| = \Omega_p/c$. The RBS Langmuir wave needs to satisfy $\omega = \Omega_p$ and $k = 2k_0 - \Omega_p$. A simulation involving a pulse with $a_0 = 0.5$ and a plasma $n_0 = 3.8 \cdot 10^{19} \text{ cm}^{-3}$ has been performed, and the results satisfy all of the above relations within accuracy limits.

We have also performed benchmarks to reproduce results on RFS stimulation in beat-wave experiments [21], RFS stimulation through direct RFS seeding [25], and suppression of RBS by RFS in beat-wave experiments [24], all obtained in the linear regime. Simulations have been performed with $a_0 = 0.2$ for the RFS seeding case and $a_0 = 0.1$ for both pulses in the beat-wave case. The effects described in the papers mentioned have all been reproduced. Therefore, we consider the code to be sufficiently reliable for our needs.

6.4.2 Effect of Raman backward scattering

As explained in Section 6.2, in the regime of high laser intensities the presence of RBS is expected to decrease the number of generated high-energy electrons, while its suppression is predicted to improve fast electron production. In order to isolate the effect of RBS suppression on the number of high-energy particles

we have performed two simulations for a type 0 pulse having amplitude $a_0 = 1.5$, one full simulation and one with suppression of left-going fields, as described above. The amplitude has been chosen such that the nonlinear RBS growth rate is maximum, while the nonlinear RFS growth rate is only slightly below its maximum [20]. The laser pulse has an FWHM of 50 fs, and the plasma has the flat-top density profile described above with $n_0 = 3.8 \cdot 10^{19} \text{ cm}^{-3}$. The results are displayed in Figs. 6.2 and 6.3. In both figures, the laser pulse is located roughly between $k_p(x - ct) = 120$ to 180.

Simulation results are as follows. Figure 6.2 displays plots of the transverse left-going field $E_{l,y}$ for the full simulation and the wake electric field E_x for both simulations, at $t = 1.6$ ps, i.e. when the pulse is in a region of increasing plasma density, and the RBS growth is at its peak. The corresponding detail plots of electron phase space are displayed in Fig. 6.3. The plots of $E_{l,y}$ and E_x , show that RBS and RFS are both present in the simulation at $t = 1.6$ ps, and overlap in space and time, making the RBS-RFS two stage acceleration possible. The laser wake visible in the plot of E_x is actually the RFS Langmuir wave. Ponderomotive excitation does not play a role here since the pulse length is several times the plasma wavelength. The plot of E_x clearly shows that the suppression of RBS in a simulation leads to a much higher wakefield amplitude, in agreement with our earlier prediction.

The phase space plots indicate that RBS-induced plasma heating is dominant in the full simulation, meaning that a large number of particles is available to be trapped by the wake wave. As a result, the laser wake is heavily beamloaded and has hardly any structure. In contrast, plasma heating is much less important in the simulation with suppression of left-going fields, and the laser wake has a well-defined structure, confirming our earlier speculation. It was also found that the peaks in phase space, which consist of electrons trapped in the wake, extend to much higher momenta for the simulation with suppression than for the full simulation (not shown in the phase space plots). This too points to a higher wake amplitude at wave breaking.

For both simulations particle energy spectra have been determined at $t = 2$ ps when the pulse enters the “plateau” of maximum plasma density and the amount of trapped particles begins to saturate, and at $t = 4$ ps when the pulse leaves this plateau. At 2 ps it has been found that the simulation with suppression of left-going fields contains about 15 times as many electrons having an energy of 5 MeV or more than the full simulation. Also, the simulation with suppression displays a peak particle energy of 38 MeV, compared to 19 MeV for the full simulation. Cumulative electron energy spectra are displayed in Figure 6.4b. The simulation with suppression is labelled “no E_l ”. At later times, the particle energies have increased for both simulations, especially for the full simulation. Nevertheless, the simulation without left-going fields shows 50% more electrons having an energy above 5 MeV and 6 times as many electrons having an energy

above 25 MeV than the full simulation.

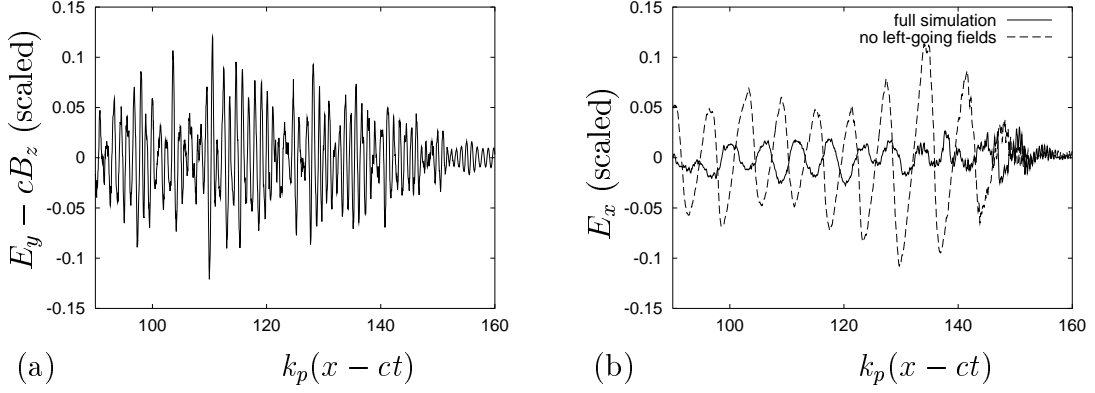


Figure 6.2: Simulation results for $a_0 = 1.5$, at $t = 1.6$ ps. The pulse is located between $k_p(x - ct) = 120$ and 180. (a) Left (i.e. backward) going transverse electric field for a full simulation. (b) Longitudinal electric (wake) field (RFS Langmuir wave) for a full simulation (solid line) and one with suppression of left-going fields (dashed line). In the full simulation, RBS and RFS overlap in space and time. Absence of RBS-induced beam loading in the other simulation leads to a higher amplitude wake field.

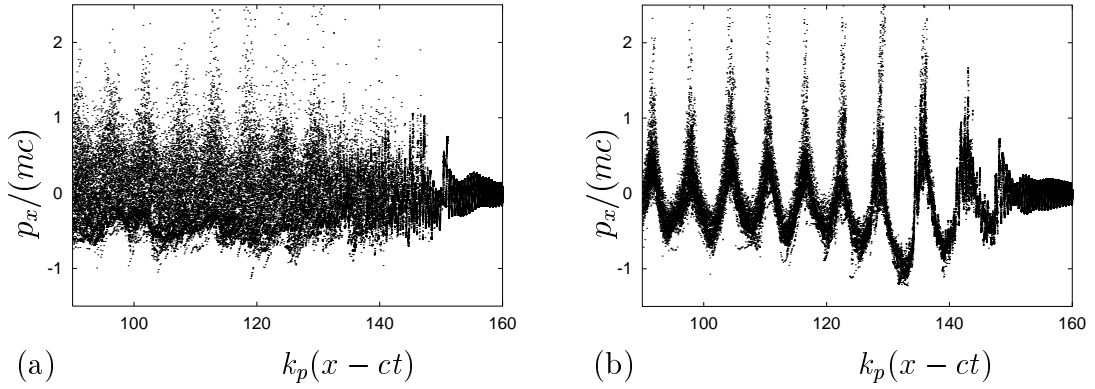


Figure 6.3: Simulation results for $a_0 = 1.5$, at $t = 1.6$ ps. Particle phase space for (a) a full simulation, and (b) one with suppression of left-going fields. Absence of RBS-induced beam loading in the latter leads to a clearer structure in phase space, which is also related to the clearer wakefield structure visible in Fig. 6.2b.

At lower laser intensities, the regime under consideration is lost. At $a_0 = 1.0$ and for the given plasma density, simulation results show that RBS growth is limited and there is little RBS-induced heating. As a result, few particles are trapped by the wake even though RBS and RFS overlapped in space, and the effect of beam loading of the wake can be neglected. Therefore, the suppression of RBS will not

lead to a larger wake amplitude but only result in a lower fast particle yield, in line with earlier results [5,14–16].

Although the direct removal of all left-going fields is a useful tool for demonstrating the effect of RBS on high-energy electron yield, RBS suppression in experiments needs to be achieved in other ways. One way, as presented in the previous section, is stimulation of RFS in laser-plasma interaction. This can be achieved by injecting into the plasma simultaneously two laser pulses having carrier frequencies that differ by ω_p [21,25]. This method will be investigated in the next subsection.

6.4.3 Effect of Raman forward scattering

In the second series of simulations, the effect of RFS on plasma wave formation and fast particle production has been investigated. Initially, the laser pulse beats with the RFS-associated Langmuir wave (ω_p, k_p) , to produce Stokes and anti-Stokes sidebands at $\omega_{\pm} = \omega_0 \pm \omega_p$ in the Fourier spectrum of the forward-going transverse fields, which are commonly used to identify RFS [2,23]. In turn, the beating between these side bands and the laser pulse resonantly drives the plasma wave, since their frequency difference is ω_p , driving it to much higher amplitudes than can be expected from ponderomotive excitation alone [16,21]. This is accompanied by the excitation of higher order (anti-)Stokes EM waves, i.e. resonance peaks at $\omega_n = \omega_0 + n\omega_p$, n integer. As the sidebands gain intensity, their mutual non-linear interaction causes sidebands at non-integer values of n to appear. This process eventually leads to the breakup of the spectrum and the laser pulse.

In order to investigate the effects of RFS more closely, simulations have been performed in which RFS has been stimulated by adding a Stokes satellite pulse to the main laser pulse, a small pulse with frequency $\omega_1 = \omega_0 - \omega_p$ and the same length and envelope shape as the main pulse. Such a satellite is expected to enhance the effect of RFS in terms of sideband growth and plasma wave excitation [21,25]. The RFS stimulation is best observed at $a_0 \sim 0.5$, since non-linear interaction occurring for larger a_0 quickly obscures the emergence of (anti-)Stokes peaks in the field spectra. Two simulations have been performed, one with a pulse with $a_0 = 0.5$ and a seed with $a_1 = 0.05$, and one without seed and a pulse with $a_0 = 0.55$, so the peak amplitudes are initially equal. The simulations have been continued for 5.33 ps, allowing the pulse to complete its transition through the plasma slab.

Since the effect of RFS seeding on (anti-)Stokes sideband growth and RFS Langmuir wave excitation has already been investigated in detail in Ref. [25], we will just summarize our results on those topics here. In both simulations, the following chain of events can be observed: emergence of the lowest order (anti-)Stokes

peaks in the spectrum and of the RFS Langmuir wave behind the pulse, emergence of higher order (anti-)Stokes peaks and further growth of the Langmuir wave, energy cascading from the main laser peak to the various (anti-)Stokes peaks, emergence of non-integer (anti-)Stokes peaks due to non-linear interaction between the various peaks, and finally loss of all structure of the transverse fields, their spectrum, and the Langmuir wave as a result of this non-linear interaction. The main difference between the two simulations is, that all these processes happen more intensely and on a much shorter timescale in the simulation with seed, than in the one without.

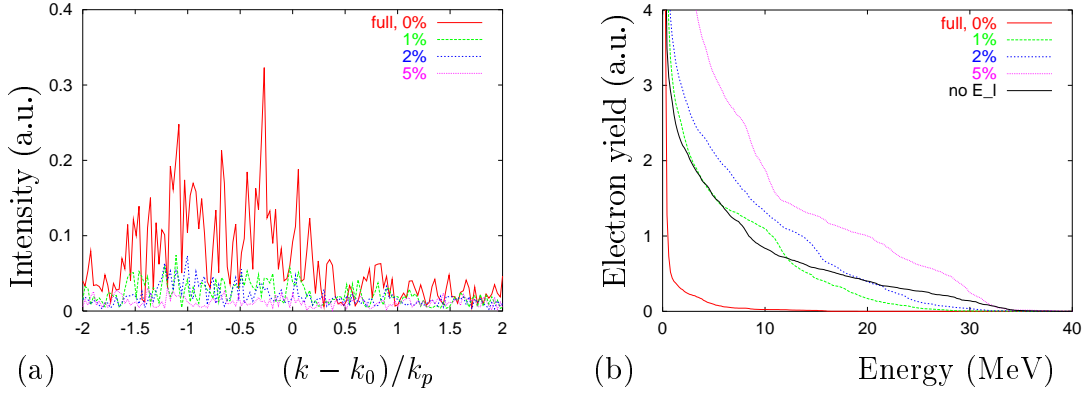


Figure 6.4: Simulation results for $a_0 = 1.5$, and a Stokes satellite of 0, 1, 2, and 5%, at $t = 2$ ps. (a) Fourier spectrum of the left-going transverse electric field, showing the decrease in RBS. (b) Cumulative particle energy spectra, showing the increase in high-energy particle yield. The curves labelled “full” and “no E_l ” correspond to the same simulations as displayed in Figs. 6.2 and 6.3.

To investigate the effect of direct RFS stimulation on RBS growth and fast electron production, we have repeated the full simulation discussed in Section 6.4.2, with a Stokes satellite pulse added to the main laser pulse. The results are displayed in Fig. 6.4. Comparing simulation results at $t = 2$ ps for $a_0 = 1.5$ and satellite levels of 0, 1, 2 and 5% of a_0 , we find that the number of high-energy electrons increases with satellite amplitude, while RBS growth decreases. For the 50 fs pulse used here, a 5% satellite is sufficient to achieve near-complete RBS suppression. Simulation results indicate that a longer pulse leads to a higher RBS level, but also that this can be compensated using a higher seed. Therefore, we conclude that adding a Stokes satellite to the main pulse is a possible way to control the amount of RBS in laser-plasma interaction experiments. Comparing the electron energy spectra of these simulations with those of the simulation without left-going fields from Section 6.4.2 (black curve in Fig. 6.4b), we find that the simulation without left-going fields produces more high-energy particles than a full simulation with 2% seed, but less than one with 5% seed. This is explained by

the fact that the seed does not only suppress RBS, but also resonantly drives the wakefield, causing a larger amplitude wakefield and therefore a larger high-energy particle yield than RBS removal alone.

Summarizing these results, we conclude that the presence of the satellite enhances the growth of RFS in the simulation, and thus the growth of the (RFS-driven) plasma wave, in agreement with earlier results [21,25]. The RFS enhancement also leads to RBS suppression, in agreement with results by Everett *et al.* [24], thus preventing heavy beam loading of the laser wake and leading to an even larger wake amplitude. Both effects ultimately lead to an increase in the yield of high-energy electrons, rendering the Stokes satellite a very effective, experimentally feasible way to stimulate fast electron production.

In the following sections, it will be investigated how the RFS/RBS growth is affected by the plasma density profile, the pulse envelope shape, and both types of laser chirp, and what the consequences are for the influence of such parameters on high energy electron production.

6.4.4 Plasma density profile

The first parameter to be investigated is the plasma density profile. As mentioned above, for a long pump moving along a plasma density ramp, the RBS growth decreases with increasing slope [26]. Unfortunately, this result cannot be applied to the case of a short pulse, i.e. much shorter than the typical length scale on which the plasma density changes. In such cases, it is expected that the RBS and RFS growth rates are determined by the local plasma density itself at the location of the pulse rather than by the slope of the density profile at that location. The total growth is then obtained by integrating the growth rate along the profile.

To investigate the influence of the plasma density profile on both RBS/RFS growth and fast electron production, we have performed two simulations in which the same laser pulse (50 fs duration, $a_0 = 1.5$) impinges upon a plasma having a piecewise linear density profile that ramps up from $n = 0$ to $n = 3.8 \cdot 10^{19} \text{ cm}^{-3}$ and then remains at that density. In the first simulation the ramp occurs over a distance of 0.6 mm, while in the second it occurs over 1.8 mm. The simulations have been continued until the pulses reached the plateau of maximum density in both cases, and particle trapping saturated.

Simulation results are displayed in Figs. 6.5 and 6.6. For both simulations, the Fourier spectrum of the backward going fields have been determined just after the onset of plasma electron trapping, since at this time the influence of RBS on particle trapping is maximum. For the steep slope case, this happens just after the pulse has reached maximum density ($t = 2.2 \text{ ps}$), while for the gentle slope this happens when the pulse is at 65% of the maximum density ($t = 4 \text{ ps}$). The spectra are shown in Figure 6.5(a): it is clear that the steep slope case exhibits

a lot more RBS than the gentle slope case. Particle energy spectra have been taken at the end of the simulations ($t = 6$ ps). They are displayed in figure 6.5b: it is obvious that once particle trapping has saturated, the pulse moving along the gentle ramp has produced much more high-energy electrons than the pulse moving along the steep ramp. It is true that the total number of trapped electrons is higher for the steep ramp case; however, nearly all of them are at the low end of the spectrum.

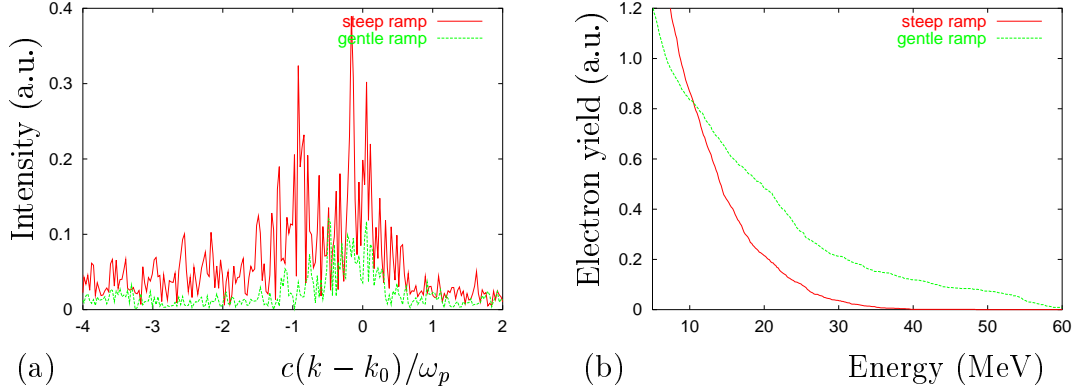


Figure 6.5: Fourier spectrum of the transverse electric field (a) and electron energy spectrum (b) for a plasma density profile with a steep ramp (red) and a gentle ramp (green). The gentle ramp case displays a higher fast particle yield than the steep ramp case. This difference is related to the lower RBS growth for a gentle ramp.

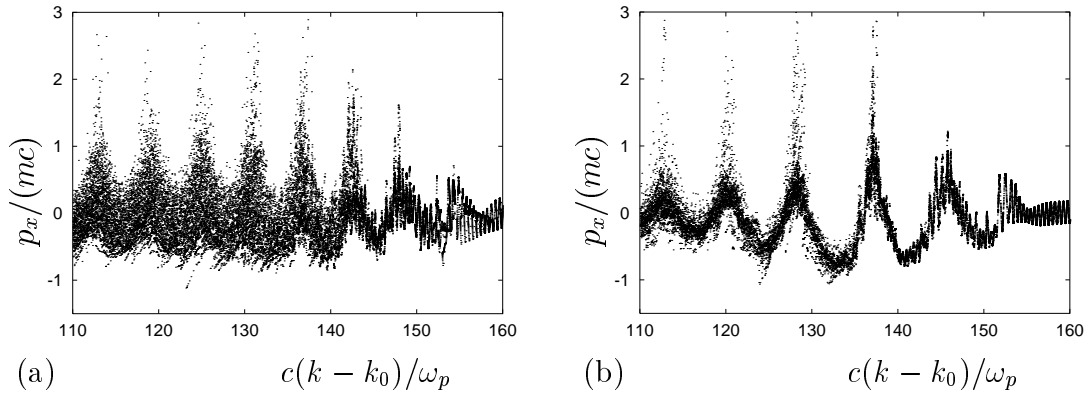


Figure 6.6: Detail plots of electron phase space for a plasma density profile with a steep ramp (a) and a gentle ramp (b). Phase space has more structure for a gentle ramp than for a steep ramp, and is less burdened by RBS-induced heating.

In Figure 6.6, detail plots of electron phase space are shown for both simulations. The steep slope case clearly displays a higher level of RBS-induced plasma heating

than the gentle slope case. This agrees with the above observation that there is more RBS excitation for the steep slope case than for the gentle slope case.

In general, it can be said that early in the simulations, the steep slope case displays a larger level of RBS than the gentle slope case, while the level of RFS is low for both cases due to the small RFS growth rate. Later in the simulations, when RFS becomes dominant and RBS has mostly disappeared, the steep slope case exhibits a higher level of RFS than the gentle slope case. The fact that the growth rates for both RBS and RFS increase with increasing plasma density may explain this: for any x between 0 and 1.8 mm, the steep slope profile has a higher local density than the gentle slope profile. Next to that, the “history” of the pulse, i.e. the profile already traversed before arriving at a certain density, is probably also important. The larger amount of highly energetic electrons seen in the gentle slope case is probably due to the fact that in both simulations, electron capture happens when RBS is still dominant. Consequently, the highest electron energy is reached in the simulation that exhibits the lowest RBS level. The difference in RFS at later times, which should work in favour of the steep slope case, seems to be much less important.

6.4.5 Laser pulse envelope

As found in experiments by Leemans *et al.* [7], the use of laser pulses with a “skew” envelope influences the yield of high-energy electrons. Experimental results presented in this paper show that a pulse with a steep front and gentle back traps more electrons than one with a gentle front and steep back. Simulations have been performed using the type (iii) laser pulse envelopes given in Section 6.4.1. Pulse parameters have been chosen to match those obtained in the experiments: $t_p = 50$ fs, $b = -0.52$ for the steep front pulse, and $t_p = 52$ fs, $b = 0.58$ for the gentle front pulse. Pulse amplitude is $a_0 = 2.6$ for both simulations. The plasma density profile is the flat-top profile from Section 6.4.1.

Simulation results are as follows. Figure 6.7 displays the cumulative electron energy spectra (a) and the Fourier spectra for the backward-going EM fields (b), both taken at $t = 2$ ps. From the energy spectra, we find that the steep front pulse generates more high-energy electrons than the gentle front pulse. The Fourier spectra reveal that the steep front pulse generates less RBS than the gentle front pulse. Figure 6.8 displays detail plots of the electron phase space at $t = 2$ ps, for a steep front pulse (a) and a gentle front pulse (b). It is clear from these plots that there is less RBS-induced plasma heating for a steep-front pulse.

The role of RFS in this case cannot readily be deduced from the simulation results. This is due to the high peak intensity of the pulses, which is well in the non-linear regime. As a result, the (anti-)Stokes peaks of RFS do not emerge clearly from the Fourier k -spectra of the forward going EM fields, and cannot

be studied. However, a linear analysis by Schroeder *et al.* [33] predicts that the steep front pulse causes a larger RFS growth than the gentle front pulse. This is supported by simulation results at low laser intensity by Fisher and Tajima [25] that indicate that a sharp leading pulse edge indeed favours RFS excitation. As a result, the steep front pulse excites a larger wake than the gentle front pulse, which also contributes to an increased fast electron yield.

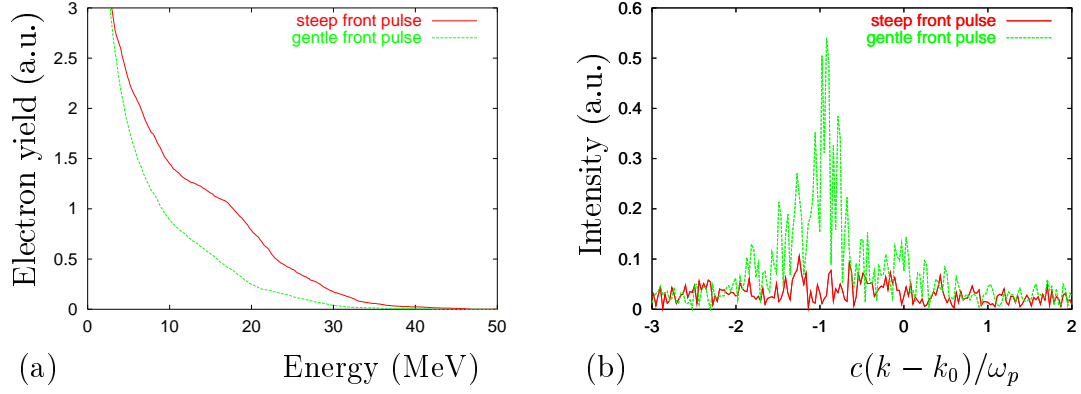


Figure 6.7: Particle energy spectrum (a) and Fourier k -spectrum of $E_{l,y}$ (b) for a pulse with a steep (red) and a gentle (green) front ($a_0 = 2.6$), at $t = 2$ ps. The steep front pulse generates less RBS and more high-energy particles than the gentle front pulse.

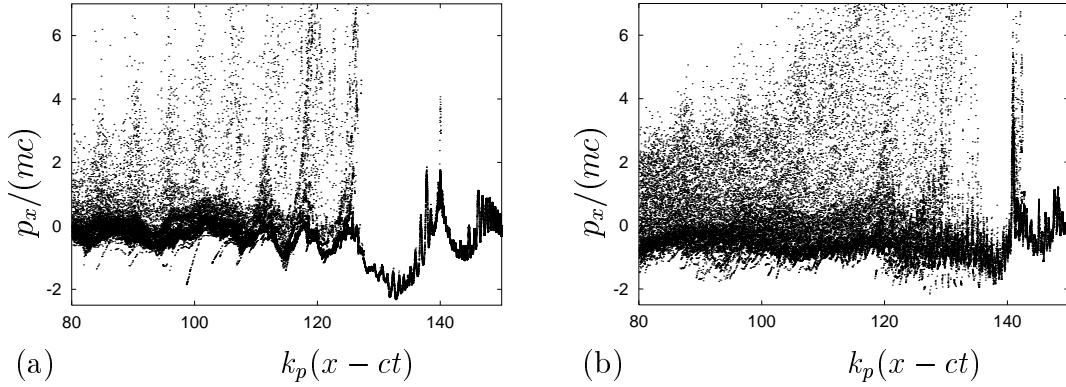


Figure 6.8: Detail plots of particle phase space for a pulse with a steep front (a) and a gentle front (b), for $a_0 = 2.6$, at $t = 2$ ps. The plot for the steep front pulse clearly exhibits more structure and less electron heating than that for the gentle front pulse.

Based on our simulation results and the theory presented in Sections 6.2 and 6.4.2, we propose the following scheme for the effect of pulse envelope shape on energetic electron yield. (i) A steep front pulse favours RFS growth, while a

gentle front pulse does the opposite; (ii) enhanced RFS growth leads to both a larger laser wake and suppression of RBS; (iii) RBS suppression leads to less beam loading of the wake and thus to an even larger wake; (iv) the combined effect of all this is a marked increase in high energy electrons for the steep front pulse as compared to the gentle front pulse, in agreement with experimental results [7].

6.4.6 Pulse length preserving chirp

As shown by Dodd and Umstadter [30], pulse length preserving laser chirp has a clear effect on both RFS growth and fast electron production. This type of chirp is characterized by its large carrier frequency gradient. In this section the effect of this type of chirp on both RBS and RFS is investigated, in order to determine whether RBS also plays a role in this case.

Several simulations with pulse length preserving chirp have been performed. For the plasma, $n_0 = 1.1 \cdot 10^{19} \text{ cm}^{-3}$ has been used, and either a constant density or a flat-top plasma slab as specified above. A type (ii) pulse envelope has been used with $a_0 = 1.0$, $\lambda_0 = 1 \text{ } \mu\text{m}$, 80 fs FWHM, and a chirp factor $a = \pm 15.4$.

Simulation results are as follows. First, the pulse with positive chirp displays a lot of erosion at the front, where its frequency is low. This causes the pulse to obtain a steep front, which is presumed to favour RFS excitation. Second, the pulse undergoes the type of breakup of its field and spectrum that points to a large RFS growth. Third, a large amplitude plasma wave is excited as a result of the high level of RFS present. Even at the moderate peak amplitude used here, some fast electron trapping can be observed.

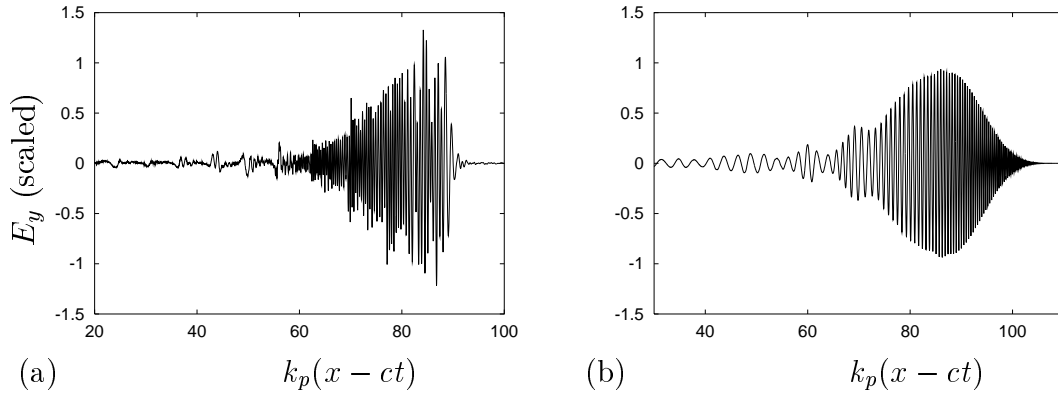


Figure 6.9: Transverse electric field at $t = 6 \text{ ps}$ for pulses with pulse length preserving chirp: (a) positive chirp, and (b) negative chirp. The pulse with positive chirp undergoes strong depletion at its front, while the pulse with negative chirp is hardly affected. In both cases, complete RBS suppression can be observed.

All these effects are much less visible, if at all, for the pulse with negative chirp. There is hardly any erosion at front or back of the pulse, pulse breakup or RFS. The absence of RFS also results in a small plasma wake; neither wave breaking nor particle trapping can be observed here. See Figs. 6.10 to 6.11.

Note that the results on both RFS growth and fast electron yield are in agreement with those from earlier simulations [30]. They can be explained from the large carrier frequency gradient, as discussed in Section 6.3.

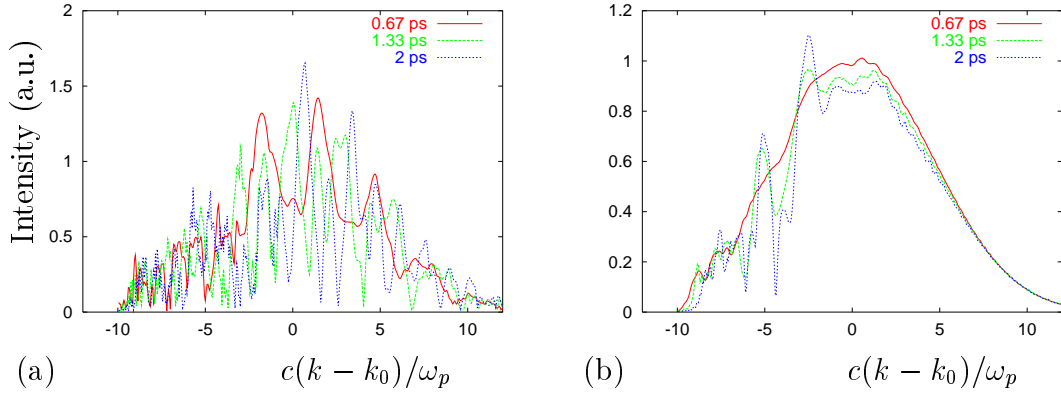


Figure 6.10: Fourier spectrum of transverse electric field for pulses with length preserving chirp: (a) positive chirp and (b) negative chirp. The spectra for positive chirp display a large RFS growth, while there is hardly a trace of RFS for negative chirp.

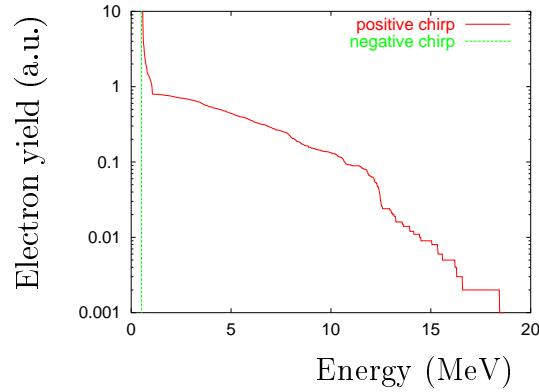


Figure 6.11: Electron energy spectrum at $t = 6$ ps for pulses with positive (red) and negative (green) pulse length preserving chirp. It is clear that positive chirp enhances fast electron yield, while negative chirp suppresses it.

As can be seen in Figure 6.9, there is hardly any RBS to be observed for either positive or negative chirp, something that is accompanied by a total absence of

any RBS-induced plasma heating. This is in agreement with the prediction by Sakharov and Kirsanov [20], that RBS will be suppressed if the gradient of the carrier frequency is sufficiently large, regardless of its sign. The absence of RBS allows us to study the effect of RFS on the plasma wave growth, without being bothered by the disturbing effects of RBS.

6.4.7 Bandwidth preserving chirp

The last parameter to be discussed is bandwidth preserving laser chirp. For this type of chirp, the Fourier power spectrum is conserved, causing the pulse duration to increase and the amplitude to decrease with chirp. As discussed in Section 6.3, the dominant characteristic of this type of chirp is the increase in pulse length, not the carrier frequency gradient. Therefore, we expect the amount of RBS, and hence the particle yield, to depend on the magnitude rather than the sign of this type of chirp. To investigate this, we have performed simulations involving pulses having a type (i) envelope with $a = \pm 1.0$ and $a = \pm 5.9$. In both cases, $a_0 = 3.0$ and FWHM = 50 fs. The plasma density profile was again Gaussian, with $n_0 = 1.1 \cdot 10^{19} \text{ cm}^{-3}$.

The following results have been obtained. In Figs. 6.12 and 6.13, the Fourier spectra of the left- and right-going transverse electric field at $t = 6$ ps are shown for both small and large chirp. We see that the spectra for pulses with small chirp ($a = \pm 1.0$) are dominated by RFS in the right-going fields and show little RBS in the left-going fields, while those for large chirp ($a = \pm 5.9$) hardly show any RFS, but display a significant amount of RBS. This behaviour agrees with the predictions made above for long pulses with a low carrier frequency gradient. Also, we observe that there is no qualitative change in the spectra when the sign of the chirp is changed while its magnitude is not. This is in agreement with the findings of Faure *et al.* [29]. In Figure 6.14, the corresponding cumulative particle energy spectra are displayed. It is clear from these spectra that particle numbers increase and particle energies decrease rapidly with increasing chirp magnitude, while they are nearly insensitive to its sign. This can be explained from the observations that RBS growth increases for increasing pulse length, i.e. increasing chirp magnitude, while the neither RBS nor RFS growth is affected by the sign of bandwidth preserving chirp. As explained before, an increase in the level of RBS rises the amount of trapped particles but strongly lowers their average energy. Also note that the number of trapped electrons naturally increases with increasing pulse-plasma interaction length.

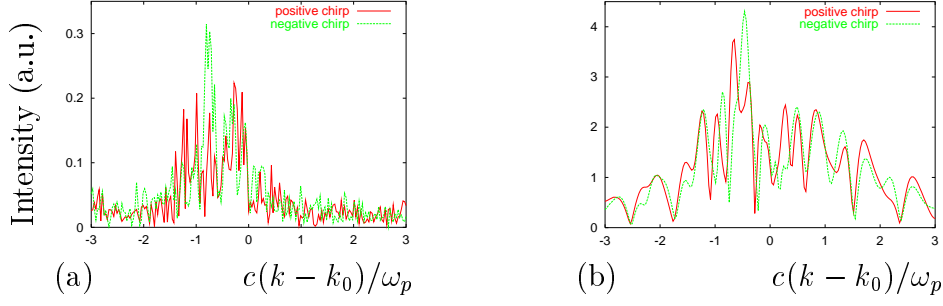


Figure 6.12: Fourier spectrum of the transverse electric field for pulses with small bandwidth preserving chirp ($a = \pm 1.0$), at $t = 6$ ps, for the backward-going (a) and forward-going (b) fields. RBS growth is limited because the pulse is still very short (~ 70 fs). The influence of the sign of the chirp on both RBS and RFS growth is small.

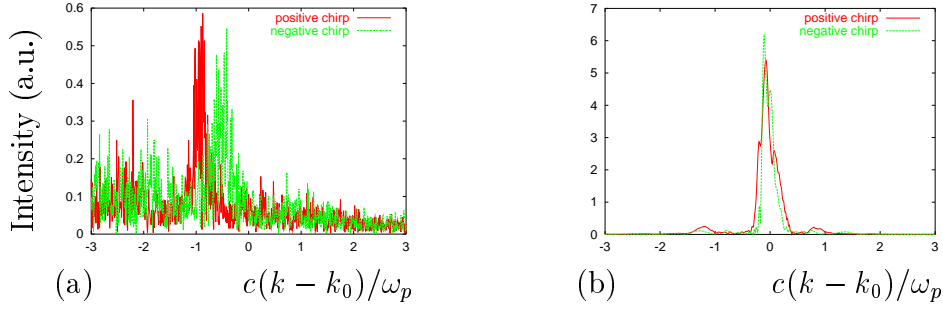


Figure 6.13: Fourier spectrum of the transverse electric field for pulses with large bandwidth preserving chirp ($a = \pm 5.9$), at $t = 6$ ps, for the left-going (a) and right-going (b) fields. RBS growth is larger in this case because of increased pulse length (~ 300 fs). Again, the influence of the sign of the chirp on both RBS and RFS growth is small.

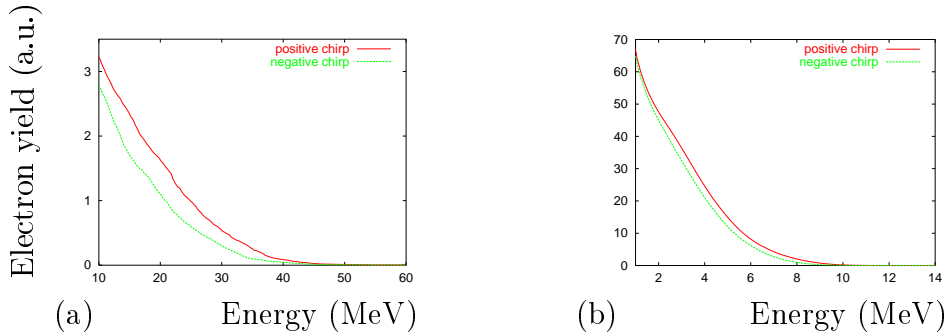


Figure 6.14: Cumulative particle energy spectrum for pulses with bandwidth preserving chirp at $t = 6$ ps, for $a = \pm 1.0$ (a) and $a = \pm 5.9$ (b). The average particle energy decreases strongly with increasing chirp magnitude, while the sign of the chirp has hardly any influence.

6.5 Summary and conclusions

A new regime for the effect of Raman instabilities on the production of fast electrons in laser-plasma interaction has been identified. The regime is characterized by laser intensities well above the threshold for electron trapping, large growth rates of the Raman instabilities, and spatial and temporal overlap of RBS and RFS. It has been demonstrated that in this regime the presence of RBS hurts the production of fast (over 15 MeV) electrons, whereas RBS suppression enhances this production. RBS suppression can be obtained by stimulating RFS growth. An experimentally feasible way to do this, is adding, to the main laser pulse, a satellite pulse at the Stokes frequency. The effect of a number of pulse and plasma parameters on fast electron production has also been investigated. It has been found that parameters that favour RFS growth and/or suppress RBS growth, ultimately lead to an increased yield of fast electrons.

In the new regime, RBS-induced plasma heating strongly increases the number of electrons that get trapped in the wake. This results in strong beam loading which reduces the wake amplitude. Although the total number of trapped electrons increases, the number of high-energy electrons decreases in this case. Under these circumstances, suppression of RBS leads to enhanced wake wave growth, and an increase of both the number of fast electrons and their energy. This is in contrast to earlier results at lower laser intensities, where RBS has been found to improve fast electron production [5,14–16].

Adding a Stokes satellite to the main laser pulse is an efficient way to enhance fast electron production. The satellite acts in two ways: it enhances the laser wake, and also causes RBS suppression. Depending on the circumstances, a satellite having 5% of the field amplitude (0.2% of the intensity) of the main pulse may already be sufficient to reach a near-complete suppression of RBS. This suppression has been explained from the fact that the seed enhances the growth of RFS, which in turn reduces the growth of RBS. Enhancing RFS through other means, such as tuning various laser and plasma parameters, gives similar results. In the simulations, a high yield of fast electrons is always accompanied by a high level of RFS and a low level of RBS.

The influence of various laser and plasma parameters on fast electron yield can be explained through their influence on RFS and RBS growth. For the plasma density profile, it has been found that a pulse climbing a gentle density increase initially excites less RBS than one meeting a steep density increase. As a result, the number of high-energy electrons generated by the pulse is higher in case of a gentle increase than that of a steep increase. Later in the simulations, the level of RFS, and thus the laser wake, is larger for a steep than for a gentle density increase. This does not seem to have a substantial influence on the final electron energy though.

The results of the simulations focusing on the effect of the shape of the pulse envelope demonstrate that an envelope with a steep front produces less RBS than an envelope with a gentle front. Although the influence of envelope shape on RFS cannot be determined from the simulations, due to the dominance of nonlinear effects at the laser intensities employed here, analytical calculations indicate that the steep front pulse excites more RFS than the gentle front pulse. Together with the fact that enhancing RFS suppresses RBS, this may well explain the difference in RBS between the pulses. The net result is a higher number of high-energy electrons produced by the steep front pulse than by the gentle front pulse, in line with experimental results [7].

The investigations on the role of chirp reveal that in order to avoid confusion, a distinction needs to be made between bandwidth preserving chirp and pulse length preserving chirp. Regarding the effect of pulse length preserving chirp, pulses with a large chirp of this type are found to generate hardly any RBS regardless of the sign of the chirp. RFS growth, and thus fast electron production, is different for pulses with positive or negative chirp, in agreement with earlier results [30,32]. For bandwidth preserving chirp, the sign of this type of chirp does not have any substantial effect on RBS or RFS growth, and thus on fast electron yield. This is in contradiction with earlier conjectures [11,30], but in agreement with recent experimental results [7] showing that differences in electron yield are likely the result of a difference in pulse envelope rather than chirp. Furthermore, RBS growth increases with the magnitude of the chirp, since the pulse length, and thus the pulse-plasma interaction region, also increases with this magnitude. This results in increased beam loading of the generated plasma wave and a drop in the yield of truly fast electrons (energy over 15 MeV). Since laser chirp as used in experiments is bandwidth preserving by nature, it is unlikely that chirping a pulse will ever increase fast electron yield in experiments.

Acknowledgments

This work has been supported by the Centrum voor Plasmafysica en Stralings-technologie, the US Department of Energy, Office of High Energy and Nuclear Physics, contract DE-AC-03-76SF0098, and the project Advanced Computing for 21st Century Accelerator Science and Technology. One of the authors (RT) acknowledges useful discussions with C. Schroeder.

References

- [1] K. Nakajima *et al.*, Phys. Rev. Lett. **74**, 4428 (1995).
- [2] C. Coverdale *et al.*, Phys. Rev. Lett. **74**, 4659 (1995).

-
- [3] A. Modena *et al.*, Nature **377**, 606 (1995).
 - [4] D. Umstadter *et al.*, Science **273**, 472 (1996).
 - [5] C. I. Moore *et al.*, Phys. Rev. Lett. **79**, 3909 (1997).
 - [6] V. Malka *et al.*, Phys. Plasmas **8**, 2605 (2001).
 - [7] W. P. Leemans *et al.*, Phys. Rev. Lett. **89**, 174802 (2002).
 - [8] E. Esarey *et al.*, IEEE Trans. Plas. Sci. **24**, 252 (1996).
 - [9] A. Reitsma *et al.*, Phys. Rev. ST Accel. Beams **5**, 051302 (2002).
 - [10] P. Catravas *et al.*, Phys. Plasmas **9**, 2428 (2002).
 - [11] W. P. Leemans *et al.*, Phys. Plasmas **8**, 2510 (2001).
 - [12] M. Tabak *et al.*, Phys. Plasmas **1**, 1626 (1994).
 - [13] R. M. G. M. Trines *et al.*, Submitted to Physical Review Letters (2002).
 - [14] P. Bertrand *et al.*, Phys. Rev. E **49**, 5656 (1994); Phys. Plasmas **2**, 3115 (1995).
 - [15] K.-C. Tzeng, W. B. Mori, and C. D. Decker, Phys. Rev. Lett. **76**, 3332 (1996); K. C. Tzeng, W. B. Mori, and T. Katsouleas, Phys. Plasmas **6**, 2105 (1999).
 - [16] E. Esarey *et al.*, Phys. Rev. Lett. **80**, 5552 (1998); E. Esarey, C. B. Schroeder, W. P. Leemans, and B. Hafizi, Phys. Plasmas **6**, 2262 (1999).
 - [17] J. F. Drake *et al.*, Phys. Fluids **17**, 778 (1974).
 - [18] William L. Kruer, *The Physics of Laser Plasma Interactions* (Addison-Wesley, Reading MA, 1988).
 - [19] W. B. Mori *et al.*, Phys. Rev. Lett. **72**, 1482 (1994).
 - [20] A. S. Sakharov and V. I. Kirsanov, Phys. Rev. E **49**, 3274 (1994).
 - [21] C. Joshi *et al.*, Nature **311**, 525 (1984).
 - [22] S. V. Bulanov, V. I. Kirsanov, and A. S. Sakharov, Pis'ma Zh. Éksp. Teor. Fiz. **53**, 540 (1991) [JETP Lett. **53**, 565 (1991)]; R. M. G. M. Trines, V. V. Goloviznin, L. P. J. Kamp, and T. J. Schep, Phys. Rev. E **63**, 026406 (2001).
 - [23] C. D. Decker *et al.*, Phys. Rev. E **50**, R3338 (1994).

-
- [24] M. J. Everett *et al.*, Phys. Rev. Lett. **74**, 2236 (1995); Phys. Plasmas **3**, 2041 (1996).
 - [25] D. L. Fisher and T. Tajima, Phys. Rev. E **53**, 1844 (1996).
 - [26] C. S. Liu, M. N. Rosenbluth, and R. B. White, Phys. Fluids **17**, 1211 (1974).
 - [27] C. A. Coverdale *et al.*, Fiz. Plazmy **22**, 685 (1996) [Plasma Phys. Rep. **22**, 617 (1996)].
 - [28] D. Strickland and G. Mourou, Opt. Comm. **55**, 447 (1985); *ibid.* Opt. Comm. **56**, 219 (1985).
 - [29] J. Faure *et al.*, Phys. Rev. E **63**, 065401 (2001).
 - [30] E. S. Dodd and D. Umstadter, Phys. Plasmas **8**, 3531 (2001).
 - [31] V. S. Malinovsky and J. L. Krause, Phys. Rev. A **63**, 043415 (2001).
 - [32] W. B. Mori, IEEE J. Quantum Electron. **33**, 1942 (1997).
 - [33] C. B. Schroeder *et al.*, submitted to Physics of Plasmas Special Issue (2003).
 - [34] J. P. Verboncoeur, A. B. Langdon, and N. T. Gladd, Comput. Phys. Commun. **87**, 199 (1995); J. Wurtele and H.-J. Lee (private communication).
 - [35] K. S. Yee, IEEE Trans. Antennas Propag. **14**, 302 (1966).
 - [36] C. Gahn *et al.*, Phys. Rev. Lett. **83**, 4772 (1999).
 - [37] A. Pukhov and J. Meyer-ter-Vehn, Appl. Phys. B **74**, 355 (2002).

Chapter 7

Magnetic field modelling in particle optical devices

So far, we have only concerned ourselves with bunch quality control during the acceleration stage. But the issue does not stop there by any means. After being accelerated, the bunch needs to be brought to its destination, which may be many metres and as many nasty turns removed from the accelerator apparatus. Or, one may wish to store the accelerated bunch in a storage ring for later use. In any case, bunch quality control is as important in this stage as it is during the acceleration stage. Therefore, the bunch guiding or storage facilities need to be designed with care: a poorly designed guiding system can easily destroy the quality that accelerator designers have taken so much pains to achieve.

As mentioned in Chapter 1, the guiding of a bunch of charged particles is done by means of magnetostatic lenses. The behaviour of a bunch of particles in such a lens can be described using the same math as used to describe the behaviour of a ray of light in an optical lens, hence the name. In order to evaluate the action of such a lens on a bunch, one needs knowledge of the magnetic field of the lens. The field is commonly described in terms of a power series expansion with respect to the transverse radius. From this expansion, the action of the lens is usually only calculated to leading order, since the inclusion of higher order contributions would lead to very complicated and unmanageable expressions. The resulting expressions are then only valid in the paraxial region, i.e. close to the central axis of the lens, while its strength is taken to be constant along its length, even though it is obvious that the strength decreases smoothly to zero at the entry and exit faces.

In this chapter, a new method for the calculation of the magnetic fields of beam guiding elements is presented. The method relates the calculation to measurement data of the magnetic field in a direct way. It can be applied to single beam guiding elements as well as to clusters of elements. The presented descrip-

tion of the magnetic field differs from the classical approach in that it does not rely on power series approximations. It is also both divergence free and curl-free, and takes fringe field effects up to any desired order into account. In the field description, pseudo-differential operators described by Bessel functions are used to obtain the various multipole contributions. Magnetic field data on a 2-dimensional surface, e.g. a cylindrical surface or a median plane, serves as input for the calculation of the 3-dimensional magnetic field. A boundary element method is presented to fit the fields to a discrete set of field data, obtained, for instance, from field measurements, on the 2-dimensional surface. Relative errors in the field approximation do not exceed the maximal relative errors in the input data. Methods for incorporating the obtained field in both analytical and numerical computation of transfer functions are outlined. Applications include easy calculation of the transfer functions of clusters of beam guiding elements, and of generalized field gradients for any multipole contribution up to any order.

This chapter has been published as the article “Modelling the fields of magneto-optical devices, including fringe field effects and higher order multipole contributions. Application to charged particle optics”, R.M.G.M. Trines, S.J.L. van Eijndhoven, J.I.M. Botman, T.J. Schep, and H.L. Hagedoorn, *Phys. Rev. ST* **4**, 062401 (2001).

7.1 Introduction

The transport of charged particle beams through particle optical devices, such as beam-transport lines, particle accelerators, and spectrometer equipment, depends strongly on the shape of their electric and magnetic fields. For this reason, a lot of effort has been put in the derivation of analytical expressions for both these fields and the trajectories of charged particles that pass through said fields. A thorough study on axisymmetric electrostatic and magnetostatic lenses has been performed by El-Kareh & El-Kareh [1]. Analytical studies on the trajectory equations for magnetic quadrupole lenses and their solutions have been performed by, among others, D. L. Smith [2] (with corrections by G. E. Lee-Whiting [3]), G. E. Lee-Whiting [4], H. Matsuda & H. Wollnik [5], H. Nakabushi & T. Matsuo [6]. A derivation of these equations using Hamilton theory has been given by H. L. Hagedoorn et al. [7], R. de Leeuw et al. [8].

The solution to the trajectory equations for a beam guiding element is generally presented as a function, called the transfer map of the element, that maps the initial location of a charged particle in phase space on its final location. Three different methods for the calculation of such maps can be distinguished. First, the map can be calculated using aberration coefficients (see G. W. Grime et al. [9] or G. W. Grime & F. Watt [10] for their definition). This method has been employed in most of the analytical work mentioned above, and has also

been used in various computer codes, e.g. TRANSPORT [11] and TURTLE [12]. Second, the map can be calculated using differential algebra, a tool developed by M. Berz [13], who implemented it in his code COSY INFINITY [14]. Finally, Lie methods can be used to calculate the map in such a way that it is always symplectic (i.e. preserving the volume occupied in phase space), regardless of the way it is truncated. These methods have been developed by A. Dragt et al. [15,16], who also implemented them in their code MARYLIE [16,17].

For the calculation of the transfer map of a magneto-optical element, an accurate description of the magnetic field of the element is essential. This is often done by expanding the scalar potential of the field in a Taylor-Fourier series, as given by, among others, M. Szilagyi [18], for straight elements having their design orbit along the z -axis. The accuracy of this description depends critically on the accuracy at which the z -dependent Taylor coefficients in the expansion, the *generalized gradients*, can be provided, especially for the so-called fringe fields near the ends of the element, where the z -dependence of the generalized gradients is the most obvious. Direct calculation of these gradients using the Biot-Savart law or on-axis field gradients is possible but not very accurate; a more accurate analytical method has been derived by M. Venturini & A. Dragt [19].

There are several disadvantages to the Taylor-Fourier expansion mentioned above. First, it leads to a culmination of power series terms in the transfer map calculations, which proves difficult to handle. Moreover, the divergence-free *and* curl-free nature of the magnetic field is not preserved once the power series is truncated [20]. Transfer map computation using series expansions in the magnetic field description usually means a trade-off between accuracy of the result and a reasonable number of terms in the expressions.

In this chapter, we formulate a different approach to the description of the fields of particle optical devices. A more detailed treatment of the work presented here can be found in the master's thesis of the author [21]. Parts of it have already been presented at the EPAC '98 [22] and PAC '99 [23] conferences. The aim of this approach is to overcome the above mentioned difficulties by using a magnetic field description that is not based on power series expansions. This way, we do not have to deal with an explosively growing number of higher order terms, which is usually the case in higher order perturbative methods.

As a first step, we give a general description of the magnetic field inside a beam guiding element and its harmonic scalar and vector potentials. This description treats each multipole contribution to the field separately, and gives the total field as a superposition of multipole contributions.

This done, we assume the magnetic field to be known on a cylindrical surface, that has the same axis as the element. We apply the above-mentioned field description to this case, in order to express the magnetic field and its potentials within the cylinder entirely in terms of the boundary values of the field on the surface. As

we do not use power series approximations here, and are able to take multipole contributions of any order into account, the accuracy of the field description is only limited by that of the boundary values. This is true for both the central field and the fringe field region of the device, since this method does not discriminate between z -dependent and z -independent fields.

Since the boundary conditions will usually arise from measurements, the next step is to give expressions for the field, given a series of measurements on either a cylindrical surface, or a median plane, taking the discrete nature of these measurements into account. The resulting field description will be shown to be insensitive to statistical noise in the measurements. Moreover, the field will be a superposition of divergence-free and curl-free terms, so it is always divergence-free and curl-free, regardless of truncation. As an example, the magnetic field of an existing quadrupole will be reconstructed from magnetic field data. It will be shown that only a fraction of the available data is needed to calculate all of the field.

This method of expressing a magnetic field in terms of measurements knows a wide range of possible applications.

One example is the reconstruction of the field of beam guiding elements (application to an existing magnetic quadrupole is presented in Section 7.3.3). Our method can also be used on clusters of elements with overlapping fringe fields, which are in fact to be treated as single elements, and can even be extended to obtain the transfer functions for a complete beam-line setup or storage ring. There are also numerous applications in low-energy electron optics.

Once the magnetic vector potential has been described in terms of field measurements, we can use it in particle optics. For this reason, we insert this vector potential in the Hamiltonian equations of motion for a charged particle in a magnetic field, and use a finite difference method to solve the system of equations numerically, in such a way that the steps in the finite difference method match the steps between the measurements. As a consequence, we have developed a tool to express the transfer function entirely in terms of these measurements.

The methods presented in this chapter are equally suitable for the calculation of generalized gradients of any order, directly from field measurements. This provides an elegant way to combine the results of this chapter with existing analytical results.

It should be noted that, although all methods have been devised to describe the fields of magneto-optical devices, they are equally applicable to electrostatic optical devices, since in the absence of free charges or currents in the interior of the device, the electrostatic potential can be expressed in the same form as the scalar potential of a magneto-optical device. Once the electrostatic scalar potential has been obtained, it can be introduced into the description of the transfer function of the device, much in the same way as the magnetic vector

potential.

7.2 General field description

7.2.1 Basic equations

From Maxwell's equations for the static electromagnetic field, a scalar potential u , and a vector potential \mathbf{A} exist for a magnetic field \mathbf{B} in a simply-connected region without free charges or currents, satisfying:

$$\nabla u = \mathbf{B} = \nabla \times \mathbf{A}, \quad (7.1)$$

$$\Delta u = 0, \quad (7.2)$$

$$\nabla \times (\nabla \times \mathbf{A}) = \mathbf{0}. \quad (7.3)$$

The vector potential \mathbf{A} will be chosen such that $\text{div } \mathbf{A} = 0$ (Coulomb gauge). Then \mathbf{A} is harmonic: $\Delta \mathbf{A} = \mathbf{0}$.

These equations are applied to the magnetic field of a magnetic multipole device, within a cylindrical surface. Cylindrical coordinates (r, φ, z) are chosen in such a way that the z -axis coincides with the central axis of the device. We assume the scalar potential u to be known at the cylindrical surface $r = R$, and introduce dimensionless coordinates $r^* = r/R$, $z^* = z/R$. Then the potential problem for the scalar potential u reads (we drop the stars for convenience):

$$\left\{ \begin{array}{ll} \frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \varphi^2} + \frac{\partial^2 u}{\partial z^2} = 0, & \\ u(1, \varphi, z) = U(\varphi, z), & -\pi < \varphi \leq \pi, \quad z \in \mathbb{R}, \\ \lim_{z \rightarrow \pm\infty} u(r, \varphi, z) = C_{\pm}, & 0 \leq r \leq 1, \quad -\pi < \varphi \leq \pi, \\ \lim_{|z| \rightarrow \infty} \frac{\partial u}{\partial z}(r, \varphi, z) = 0, & 0 \leq r \leq 1, \quad -\pi < \varphi \leq \pi. \end{array} \right. \quad (7.4)$$

Here, C_+ and C_- are constants and $U(\varphi, z)$ satisfies both $\lim_{z \rightarrow \pm\infty} U(\varphi, z) = C_{\pm}$ and $\lim_{z \rightarrow \pm\infty} \frac{\partial U}{\partial z}(\varphi, z) = 0$ for all φ . If the Fourier series expansion with respect to φ of the potential is known to contain no solenoidal (φ -independent) term, we set $C_{\pm} = 0$. The above potential problem has a unique solution for u [24]; from this solution we can derive general expressions for both \mathbf{B} and \mathbf{A} . These expressions can be used to compute \mathbf{B} and \mathbf{A} directly from magnetic field measurements, without the need to derive u first.

7.2.2 Harmonic potentials

Since we use cylindrical coordinates, we can expand a solution $u(r, \varphi, z)$ of (7.4) into a Fourier series:

$$u(r, \varphi, z) = a_0(r, z) + \sum_{m=1}^{\infty} (a_m(r, z) \cos(m\varphi) + b_m(r, z) \sin(m\varphi)).$$

The terms corresponding to a certain value of m represent the $2m$ -pole contribution to u : $m = 0$ corresponds to a solenoid, $m = 1$ to a dipole, $m = 2$ to a quadrupole, etc. For convenience, we adopt the following notation for this Fourier series:

$$u(r, \varphi, z) = (a_m(r, z), b_m(r, z))_{m=0}^{\infty},$$

using the dummy coefficient $b_0(r, z) = 0$.

Having inserted this expansion into the Laplace equation for u , we find second order partial differential equations for the coefficients $a_m(r, z)$ and $b_m(r, z)$. As u needs to be finite for $r = 0$, their formal solutions are given by:

$$\begin{aligned} a_m(r, z) &= J_m(r \frac{d}{dz}) A_m(z), \quad m = 0, 1, 2, \dots \\ b_m(r, z) &= J_m(r \frac{d}{dz}) B_m(z), \quad m = 1, 2, \dots \end{aligned}$$

The pseudo-differential operator $J_m(r \frac{d}{dz})$ is defined as [25]:

$$J_m(r \frac{d}{dz}) A_m(z) := \mathcal{F}^{-1} [J_m(i\omega r) \times (\mathcal{F} A_m)(\omega)](z),$$

where J_m denotes the Bessel function of the first kind of order m and \mathcal{F} denotes the classical Fourier integral transformation with respect to z :

$$(\mathcal{F} f)(\omega) := \int_{-\infty}^{\infty} f(z) e^{-i\omega z} dz.$$

Note that the solutions for $a_m(r, z)$ and $b_m(r, z)$ can be expressed in this form because of the boundary conditions for $|z| \rightarrow \infty$ in (7.4).

In the next section, the smooth (i.e. arbitrarily often differentiable) functions $A_m(z)$ and $B_m(z)$ will be determined from the boundary condition $u(1, \varphi, z) = U(\varphi, z)$. For convenience, we write \mathcal{J}_m for $J_m(r \frac{d}{dz})$, and introduce the dummy coefficient $B_0 = 0$. Then the general solution for $u(r, \varphi, z)$ reads:

$$u = (\mathcal{J}_m A_m, \mathcal{J}_m B_m)_{m=0}^{\infty}. \quad (7.5)$$

As follows from Fourier integral theory, this general solution will automatically obey the boundary conditions for $|z| \rightarrow \infty$.

Next, we determine a harmonic vector potential \mathbf{A} for the magnetic field \mathbf{B} using (7.1). This vector potential will be expressed solely in terms of the coefficients A_m and B_m that fix the scalar potential u for \mathbf{B} . The vector potential is not unique: adding the gradient of any harmonic scalar field to a harmonic vector potential for \mathbf{B} yields another harmonic vector potential. As long as $\nabla \times \mathbf{A} = \mathbf{B}$, there is no physical reason to prefer one choice for \mathbf{A} over another. For convenience, we choose (expressed in cylindrical coordinates)

$$\begin{aligned} A_r &= (\mathcal{J}_{m+1}B_m, -\mathcal{J}_{m+1}A_m)_{m=1}^{\infty}, \\ A_{\varphi} &= (\mathcal{J}_{m+1}A_m, \mathcal{J}_{m+1}B_m)_{m=0}^{\infty}, \\ A_z &= (-\mathcal{J}_mB_m, \mathcal{J}_mA_m)_{m=1}^{\infty}. \end{aligned}$$

From the expressions for the scalar and vector potentials, we find that both u and \mathbf{A} are completely determined by the functions A_m and B_m , which are in turn uniquely determined by the boundary conditions at $r = 1$.

One should note that we used the gauge freedom to derive a harmonic vector potential. In some cases, however, a different (non-harmonic) vector potential is more convenient. For example, the Hamiltonian equations of motion for a charged particle in a magnetic field, contain, when expressed in cylindrical coordinates, a larger number of terms containing A_{φ} . A vector potential \mathbf{A} with $A_{\varphi} = 0$ is usually employed here to get rid of these terms. Although this vector potential is not harmonic, its Fourier coefficients can nevertheless be expressed in terms of the coefficients \mathcal{J}_mA_m and \mathcal{J}_mB_m (see Ref. [19]).

7.2.3 Introducing boundary conditions

In this section, we will show how to calculate the Fourier coefficients of u , \mathbf{B} and \mathbf{A} for which we derived formal expressions in the previous section, and their various derivatives, directly from given boundary values at the cylindrical surface $r = 1$. Such boundary values result e.g. from direct magnetic field measurements, calculations using the Biot-Savart law, or spinning coil measurements.

The function $U(\varphi, z)$, from the remaining boundary condition $u(1, \varphi, z) = U(\varphi, z)$, can be expanded into a Fourier series:

$$U(\varphi, z) = (V_m(z), W_m(z))_{m=0}^{\infty}.$$

Since $\lim_{z \rightarrow \pm\infty} u(r, \varphi, z) = C_{\pm}$, we find that necessarily $\lim_{z \rightarrow \pm\infty} V_0(z) = C_{\pm}$, while $\lim_{z \rightarrow \pm\infty} V_m(z) = \lim_{z \rightarrow \pm\infty} W_m(z) = 0$ for $m > 0$. Inserting the general solution (7.5) into this boundary condition yields the equations

$$\begin{aligned} (\mathcal{J}_mA_m)(1, z) &= V_m(z), \\ (\mathcal{J}_mB_m)(1, z) &= W_m(z). \end{aligned}$$

This allows one to express the coefficient $(\mathcal{J}_m A_m)(r, z)$ in terms of V_m :

$$(\mathcal{J}_m A_m)(r, z) = \mathcal{F}^{-1} \left(\frac{I_m(\omega r)}{I_m(\omega)} (\mathcal{F} V_m) \right) (z). \quad (7.6)$$

Note that $J_m(i\omega r) = i^m I_m(\omega r)$ with I_m the modified Bessel function of the first kind of order m .

The right hand side of (7.6) is a convolution product:

$$(\mathcal{J}_m A_m)(r, z) = (g_m(r, \cdot) * V_m)(z) = \int_{-\infty}^{\infty} g_m(r, z - \zeta) V_m(\zeta) d\zeta, \quad (7.7)$$

where the basic function $g_m(r, z)$ is given by

$$g_m(r, z) = \frac{1}{\pi} \int_0^{\infty} \frac{I_m(\omega r)}{I_m(\omega)} \cos(\omega z) d\omega.$$

The notation $g_m(r, \cdot)$ indicates that evaluation of g_m for some value of z is postponed until after the convolution product is taken. Results similar to (7.7) hold for $(\mathcal{J}_m B_m)(1, z)$.

For fixed ζ , the function $g_m(r, z - \zeta)$ is the solution to (7.6) in the case that $V_m(z) = \delta(z - \zeta)$, and in this sense, it is a fundamental solution. This fundamental property of $g_m(r, z)$ will be used to full extent in the next section.

The basic result (7.7) is also obtained in Ref. [19], through slightly different methods, and used in a different context. An expansion of (7.7) into powers of r is derived in that paper in order to obtain power series expansions of the coefficients $a_m(r, z)$ and $b_m(r, z)$. Expressions similar to (7.7) describing the field of axisymmetric lenses ($m = 0$) can also be found in Ref. [1]

It can be shown that $g_m(r, z)$ is strictly positive [21, 26], so

$$\int_{-\infty}^{\infty} |g_m(r, z)| dz = \int_{-\infty}^{\infty} g_m(r, z) dz = \left. \frac{I_m(\omega r)}{I_m(\omega)} \right|_{\omega=0} = r^m. \quad (7.8)$$

Since $\mathcal{J}_m A_m$ depends linearly on V_m , this allows one to calculate the effect of errors in V_m on the calculation of $\mathcal{J}_m A_m$. We find, using (7.8),

$$|\delta(\mathcal{J}_m A_m)(r, \cdot)| \leq r^m |\delta V_m|, \quad (7.9)$$

where $\delta(\mathcal{J}_m A_m)(r, \cdot)$ and δV_m denote the maximum errors in $\mathcal{J}_m A_m(r, \cdot)$ (as a function of r) and V_m respectively on the interval $-\infty < z < \infty$ (see also Ref. [21]). This result will prove useful in the next section, where we show how to obtain V_m from measurements.

The coefficients $\mathcal{J}_k A_m$ and $\mathcal{J}_k B_m$ with k different from m , that occur in the transverse components of \mathbf{A} , and their partial derivatives with respect to r and

φ can be calculated in the same way as $\mathcal{J}_m A_m$ and $\mathcal{J}_m B_m$. This will prove useful in Section 7.4.2, where the components of \mathbf{A} and their derivatives will occur in the Hamiltonian equations of motion for a charged particle in a magnetic field. In general, we write for $k \geq m - 1$:

$$(\mathcal{J}_k A_m)(r, z) = (g_m^k(r, \cdot) * V_m)(z), \quad (7.10)$$

where

$$g_m^k(r, z) = \begin{cases} \frac{(-1)^l}{2\pi} \int_{-\infty}^{\infty} \frac{I_k(\omega r)}{I_m(\omega)} \cos(\omega z) d\omega, & k - m = 2l, \\ \frac{(-1)^{l+1}}{2\pi} \int_{-\infty}^{\infty} \frac{I_k(\omega r)}{I_m(\omega)} \sin(\omega z) d\omega, & k - m = 2l + 1. \end{cases}$$

It should be noted that the definition of g_m^k contains no physical parameters, since the input of magnetic field data is performed completely via the boundary values V_m and W_m . Therefore, the values of g_m^k can be calculated in advance, which simplifies the calculation of $\mathcal{J}_k A_m$ for a given V_m greatly.

As has been done previously for g_m , it can be shown that $\int_{-\infty}^{\infty} |g_m^k(r, z)| dz < \infty$ for $0 \leq r < 1$ and $k > m$. Analogous to (7.9), we find that

$$|\delta(\mathcal{J}_k A_m)(r, \cdot)| \leq C_k(r) |\delta V_m|, \quad (7.11)$$

thus relating the accuracy of $\mathcal{J}_k A_m$, $k > m$, to that of V_m . There are indications that $C_k(r) \sim \mathcal{O}(r^k)$, although this has not been proven yet. For $k = m - 1$ however, it can be shown that g_m^k is not integrable on $-\infty < z < \infty$, so equation (7.11) does not apply to this case.

In the case that the field description given in this section is used in charged particle optics, knowledge of both the partial derivatives (occurring in the trajectory equations) and the integrals (occurring in the solutions to these equations) of $\mathcal{J}_k A_m$ is vital. (See also Section 7.4.3.) Their calculation is rather straightforward. The partial derivatives of $\mathcal{J}_k A_m$ are given by the convolution product of the corresponding derivative of $g_m^k(r, \cdot)$ with V_m . Integration of $\mathcal{J}_k A_m$ with respect to z can also be done by integrating either g_m^k or V_m with respect to z . Moreover, since for $m > 0$, $V_m(z)$ tends to 0 for $|z| \rightarrow \infty$, integration by parts of the right hand side of (7.10) yields:

$$(\mathcal{J}_k A_m)(r, z) = (G_m^k(r, \cdot) * V_m')(z), \quad (7.12)$$

where

$$G_m^k(r, z) = \begin{cases} \frac{(-1)^l}{2\pi} \int_{-\infty}^{\infty} \frac{I_k(\omega r)}{I_m(\omega)} \frac{\sin(\omega z)}{\omega} d\omega, & k - m = 2l, \\ \frac{(-1)^{l+1}}{2\pi} \int_{-\infty}^{\infty} \frac{I_k(\omega r)}{I_m(\omega)} \frac{1 - \cos(\omega z)}{\omega} d\omega, & k - m = 2l + 1. \end{cases}$$

Using the basic functions g_m^m and g_m^{m+1} , one can calculate the components of \mathbf{A} and their derivatives directly from the boundary values V_m and W_m , without need to calculate the magnetic field \mathbf{B} or the scalar potential u first. In fact, since the functions A_m and B_m are fully determined by V_m and W_m , any quantity related to u can be calculated directly from the boundary values, if the appropriate basic function is used.

In practice, boundary values either originate from the potential u (e.g. when produced by a three-dimensional code), or from quantities like the magnetic field \mathbf{B} (e.g. when obtained from measurements). Determining V_m and W_m in the second case takes a little extra work as compared to the first. As an example, we show how to determine the functions V_m and W_m from measured values of (a component of) \mathbf{B} . We have:

$$B_r|_{r=1} = \frac{1}{R} \frac{\partial u}{\partial r} \Big|_{r=1} = \left(\frac{1}{R} \frac{\partial g_m}{\partial r}(1, \cdot) * V_m, \frac{1}{R} \frac{\partial g_m}{\partial r}(1, \cdot) * W_m \right)_{m=0}^{\infty}, \quad (7.13)$$

$$B_\varphi|_{r=1} = \frac{1}{R} \frac{\partial u}{\partial \varphi} \Big|_{r=1} = \left(\frac{m}{R} W_m, -\frac{m}{R} V_m \right)_{m=1}^{\infty}, \quad (7.14)$$

$$B_z|_{r=1} = \frac{1}{R} \frac{\partial u}{\partial z} \Big|_{r=1} = \left(\frac{1}{R} V'_m, \frac{1}{R} W'_m \right)_{m=0}^{\infty}, \quad (7.15)$$

where the components of \mathbf{B} are given in dimensional (unscaled) form. Using the boundary condition

$$B_\alpha(1, \varphi, z) = (B_{\alpha,m}^c(z), B_{\alpha,m}^s(z))_{m=0}^{\infty},$$

where $\alpha = r, \varphi, z$, we find for $m = 0, 1, 2, \dots$,

$$\begin{aligned} \mathcal{F}(V_m, W_m) &= R \frac{I_m(\omega)}{\omega I'_m(\omega)} \mathcal{F}(B_{r,m}^c, B_{r,m}^s), \\ (V_m, W_m) &= \frac{R}{m} (-B_{\varphi,m}^s, B_{\varphi,m}^c), \\ (V'_m, W'_m) &= R (B_{z,m}^c, B_{z,m}^s). \end{aligned}$$

These relations, combined with (7.10), allow one to calculate $\mathcal{J}_k A_m$ and $\mathcal{J}_k B_m$ for $0 \leq r < 1$ and all z if one component of \mathbf{B} is known at $r = 1$. Note that, in the case of B_z being known, one needs to replace g_m^k by its primitive G_m^k , like in (7.12), while in the case of B_r being known, we obtain the correct basic function using (7.6), where the denominator $I_m(\omega)$ has to be replaced by $\omega I'_m(\omega)$, resulting in a different basic function.

In general, many field-related quantities can be calculated directly by taking the convolution product of (the Fourier coefficients of) the boundary values at $r = 1$ with an appropriate basic function. This will prove to be a powerful tool. In the next section, we will treat the case of discrete boundary values at $r = 1$, and we will show how to reconstruct these boundary values if the available magnetic field data originates from another surface than $r = 1$, e.g. a median plane.

7.3 Calculating the field from measurements

7.3.1 Using measurements at the boundary

In practice, we can measure (the components of) \mathbf{B} at $r = 1$ for a discrete set of φ - and z -values, and approximate them by interpolating the measurements. Measurements of B_φ and B_z are more convenient than measurements of B_r , since the former provide direct approximations of (V_m, W_m) and (V'_m, W'_m) respectively, while the latter do not. Since piecewise constant or piecewise linear interpolations are almost always employed, the special cases of V_m being piecewise constant or linear will be considered. As the measured values for \mathbf{B} will be negligible for sufficiently large z , we assume that, for $m > 0$, $V_m(z) = 0$ for z sufficiently large. On the other hand, since $\lim_{z \rightarrow \pm\infty} V_0(z) = C_\pm$, a solenoidal contribution to u should rather be treated by fitting $B_z = \partial u / \partial z$ than by fitting u . Using (7.15), B_z can be calculated from V'_0 , which is considered to be zero for sufficiently large z .

If we assume that V_m is piecewise constant, there are pairs (λ_i, z_i) , with $\sum_i \lambda_i = 0$, such that $V'_m(z) = \sum_i \lambda_i \delta(z - z_i)$. Using (7.12), we find:

$$(\mathcal{J}_k A_m)(r, z) = \sum_i \lambda_i G_m^k(r, z - z_i), \quad (7.16)$$

where $G_m^k(r, z)$ is defined in the previous section.

If on the other hand, we assume that V_m is piecewise linear, then V'_m is piecewise constant, and $V''_m = \sum_i \lambda_i \delta(z - z_i)$, where both $\sum_i \lambda_i = 0$ and $\sum_i \lambda_i z_i = 0$. In this case we have

$$(\mathcal{J}_k A_m)(r, z) = \sum_i \lambda_i \tilde{G}_m^k(r, z - z_i), \quad (7.17)$$

where $\tilde{G}_m^k(r, z) = \int_0^z G_m^k(r, \zeta) d\zeta$.

Following from the conditions for the λ_i in the above cases, we find that the terms at the right hand side of both (7.16) and (7.17) cancel each other for $|z| \rightarrow \infty$, so $\mathcal{J}_k A_m$ tends to 0 for $|z| \rightarrow \infty$ in both cases.

Since expressions like (7.16) and (7.17) are derived in a straightforward fashion from Eq. (7.10) for specific instances of V_m , the accuracy of $\mathcal{J}_k A_m$ is completely determined by that of V_m through Eq. (7.11), and convergence of $\mathcal{J}_k A_m$ with increasing number of z -subdivisions is directly related to that of V_m .

From the integral expressions in the previous section, we find that related quantities can be derived by replacing the functions G_m and \tilde{G}_m by basic functions corresponding to these quantities, while retaining the pairs (λ_i, z_i) . In fact, these pairs determine the corresponding multipole contribution completely.

At this point, we show how to determine the various multipole contributions to the magnetic field from the values of B_z at $r = 1$, as given by (7.15). Assume

$B_z|_{r=1}$ has been measured at the points with coordinates $(\varphi, z) = (\varphi_j, w_i)$. Since V'_m is the Fourier coefficient of $\cos(m\varphi)$ in the Fourier series representation of B_z , $V'_m(w_i)$ is obtained from (“ \doteq ” denotes a numerical approximation):

$$\begin{aligned} V'_m(w_i) &= \frac{1}{\pi} \int_{-\pi}^{\pi} B_z(1, \varphi, w_i) \cos(m\varphi) d\varphi \\ &\doteq \frac{1}{2\pi} \sum_j B_z(1, \varphi_j, w_i) (\varphi_{j+1} - \varphi_{j-1}) \cos(m\varphi_j), \end{aligned}$$

where $\sum_j (\varphi_{j+1} - \varphi_j) = 2\pi$. (Note that the simple midpoint scheme used in this example can always be replaced by a more sophisticated scheme, if so desired.) We choose the piecewise constant approximation for $V'_m(z)$, i.e. $V''_m(z) \doteq \sum_i \lambda_i \delta(z - z_i)$, with $z_i = \frac{1}{2}(w_i + w_{i+1})$ and $\lambda_i = V'_m(w_{i+1}) - V'_m(w_i)$. Then $(\mathcal{J}_k A_m)(r, z)$ is determined by (7.17). Since $B_z(z) \approx 0$ for $|z|$ sufficiently large, so is $V'_m(z)$.

Since $\tilde{G}_m^k(r, z)$ is a known function, we see that $(\mathcal{J}_k A_m)(r, z)$ can directly be related to the field measurements at $r = 1$. This provides a convenient and flexible way to calculate the magnetic field and related quantities from the field measurements of any desired particle optical device. The nature of the methods outlined in this section also allows for the calculation of the field of a cluster of devices, i.e. multipoles, which can in fact be treated as a single device.

It should be noted that one needs measurements performed at $2m$ different angles φ_i at least in order to determine the $2m$ -pole contribution. In other words, the number of different angles at which measurements are taken determines the highest order multipole contributions that can be approximated from a given set of measurements.

7.3.2 Using measurements not at the boundary

In the previous sections, the multipole coefficients $\mathcal{J}_k A_m$ were calculated under the assumption that the values of the magnetic field or the magnetic scalar potential at the surface $r = 1$ were available. In many cases however, the measurements have not been performed at the surface $r = 1$, but at a plane containing the z -axis or on the z -axis itself instead. We will show that it is possible to obtain the various multipole contributions to the magnetic field in such cases, by means of a least squares method, which will be outlined below.

We will treat the case that measurements of $B_\varphi(r, \varphi, z)$ were taken in P ($P \geq 2$) planes containing the z -axis, i.e. at the points (r_k, φ_i, w_j) , $k = 1, \dots, M$, $i = 1, \dots, P$, $j = 1, \dots, N$. In this case, we are able to fit at most M different multipole contributions; in most cases, the $2m$ -pole contributions corresponding to $m = 1, \dots, M$, will be fitted. If less than M multipole coefficients are fitted, the remaining data can be used to improve the statistical properties of the fit.

We assume that $0 < r_k < 1$, $k = 1, \dots, M$; if this is not the case, then the values r_k and z_j should be suitably scaled.

As an example, we assume that B_φ originates from the field of a realistic magnetic quadrupole, where the quadrupole and sextupole contributions are dominant, and other contributions are negligible. In this case, B_φ is approximated by

$$B_\varphi(r, \varphi, z) = -\frac{2}{Rr}a_2(r, z)\sin(2\varphi) - \frac{3}{Rr}a_3(r, z)\sin(3\varphi).$$

The coefficients a_2 and a_3 are then approximated by

$$\begin{aligned} a_2(r, z; \lambda) &\doteq \sum_{l=1}^N \lambda_l G_2(r, z - z_l), \\ a_3(r, z; \mu) &\doteq \sum_{l=1}^N \mu_l G_3(r, z - z_l), \end{aligned}$$

where $z_l = \frac{1}{2}(w_l + w_{l+1})$. This corresponds to a piecewise constant approximation of the multipole coefficients of B_φ at $r = 1$. We denote the measured value of B_φ at (r_k, φ_i, w_j) by f_{kij} , and define the quantity $M(\lambda, \mu)$ by:

$$M(\lambda, \mu) = \sum_{k=1}^M \sum_{i=1}^P \sum_{l=1}^N (f_{kij} - B_\varphi(r_k, \varphi_i, w_j))^2.$$

In order for $B_\varphi(r, \varphi, z)$ to vanish for $|z| \rightarrow \infty$, λ, μ must satisfy

$$\sum_{l=1}^N \lambda_l = \sum_{l=1}^N \mu_l = 0. \quad (7.18)$$

The optimal values for λ, μ are then obtained by minimizing $M(\lambda, \mu)$ under the conditions (7.18).

There are a few remarks to be made concerning this example:

- In the above example, the assumption was made that the multipole contributions were all normal-oriented, i.e. B_φ did not contain any $\cos(m\varphi)$ terms. In this case, it is sufficient to have measurements for one angle φ_1 . In general, the orientation of the multipole contributions will be unknown, so B_φ should contain both $\sin(m\varphi)$ and $\cos(m\varphi)$ terms, and we need at least two different angles φ_1 and φ_2 , which are such that $m|\varphi_2 - \varphi_1|/(2\pi)$ is not an integer for any $2m$ -pole contribution that is assumed to occur in the magnetic field.

- In the case that the λ_l and μ_l are obtained directly from boundary values, as in the previous section, the resulting solution automatically satisfies the boundary conditions for $|z| \rightarrow \infty$ in (7.4). However, this is not true for the least squares method presented in this section; for this reason, the condition (7.18) has to be imposed explicitly. More generally, if a piecewise n -th degree polynomial approximation is employed at the boundary, the conditions

$$\sum_{l=1}^N \lambda_l z_l^k = \sum_{l=1}^N \mu_l z_l^k = 0, \quad k = 0, \dots, n,$$

have to be imposed.

- In theory, one can determine the various multipole contributions from measurements taken at only two different angles, but in practice one can use measurements at more angles in order to reduce the influence of random errors. The same can be said about the number of different values r_k , if there are M different values r_1, \dots, r_M available, and one wishes to fit the coefficients of less than M multipole contributions.

As for the different z -values: one could use a smaller number of pairs (λ_i, z_i) for the fitting process than the number of available values w_j , but this reduces the z -range where $B_\varphi(1, \varphi, z)$ is assumed to be nonzero. For this reason, one should use all pairs (λ_i, z_i) corresponding to the values z_i where $B_\varphi(1, \varphi, z)$ is assumed to be nonzero, and use the remaining data (at the z -values where statistical noise is assumed to be dominant) to improve on statistics.

- Further improvement on the statistics can be obtained by using a weighted average for the sum of squares in the definition of $M(\lambda, \mu)$, instead of the unweighted arithmetic mean used in the example. The inverse square of the relative error in the measurement f_{kij} can be used as the weight for the term $(f_{kij} - B_\varphi(r_k, \varphi_i, w_j))^2$.

As shown in the previous section, the pairs (λ_l, z_l) fully determine the corresponding Fourier coefficient of any field-related quantity, e.g. u or B_φ . Knowledge of these pairs allows one therefore to fit the corresponding multipole contribution to any field-related quantity, using the corresponding basic function. The method presented in this section allows one to obtain these pairs from measurements on the z -axis or in the plane $\varphi = \varphi_1$, instead of on the surface $r = 1$. This will be very useful in cases where it is not possible to measure on a cylindrical surface.

Comparing the methods developed in this and the previous section, we find that the method of the previous section gives better approximations for the individual multipole coefficients, while the method of this section gives a better overall approximation for the total magnetic field. This effect is caused by the way these

methods deal with higher order multipole contributions, that are assumed to be zero, but, in fact, are not. The method of the previous section yields accurate approximations of the lower order multipole coefficients, and completely neglects any higher order multipole contributions that might exist in the field; the method of this section, however, “distributes” the total contribution of higher order terms among the lower order coefficients. Although this decreases the accuracy of the approximation of the lower order coefficients, it improves the accuracy of the overall approximation of the magnetic field. For this reason, the latter method for field approximation might be preferred when using this approximation for calculating a transfer function numerically, as described in Section 7.4.2.

7.3.3 Experimental test of the presented theory

The theory developed in the previous sections has been verified using actual field measurements for a magnetic quadrupole, performed by G. Brooijmans [27]. Part of the measurements have been used as input for the calculation of the complete quadrupole field, and the outcome has been compared to the remaining measurements.

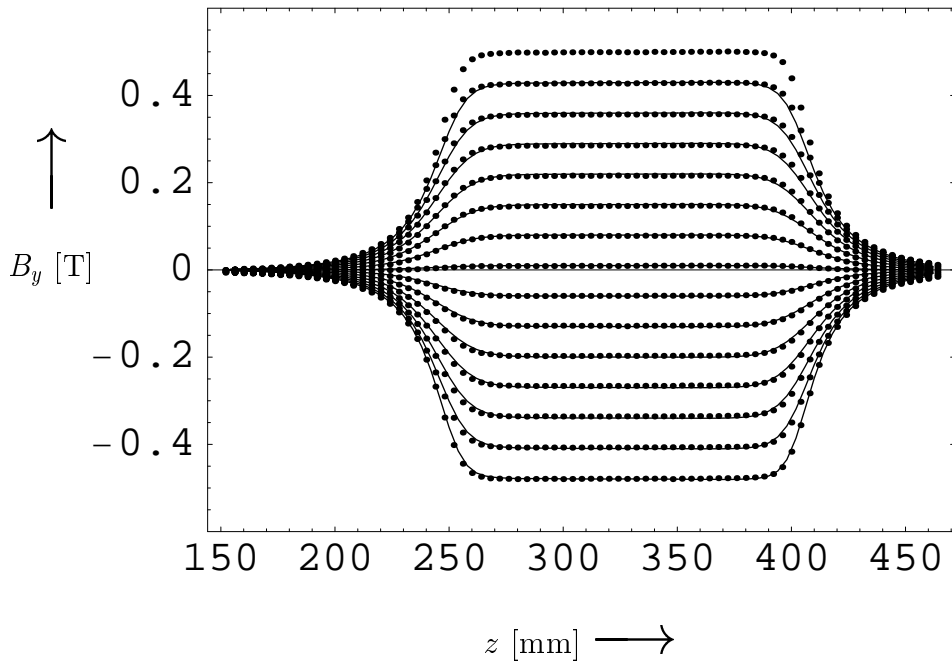


Figure 7.1: Comparison of the measured and calculated B_y as a function of z for various values of x . The dots represent the measurements, the curves the calculated field. The top row of measurements served as input for the calculations.

The quadrupole used for the measurements was normal-orientated (i.e. antisym-

metric with respect to the planes $x = 0$ and $y = 0$), and the component B_y on the plane $y = 0$ was measured, for a number of equidistant x - and z -values. Since in the center of the quadrupole, B_y turned out to depend linearly on x , it has been assumed at first, that any higher order multipole contributions could be neglected with respect to the quadrupole field. Therefore, B_y has been fitted using basic functions for $m = 2$ only. Since $B_y = B_\varphi$ for $x > 0$ and $B_y = -B_\varphi$ for $x < 0$, the basic function

$$\hat{G}_2(r, z) = \frac{1}{r} G_2(r, z) = \frac{1}{\pi r} \int_0^\infty \frac{I_2(\omega r)}{I_2(\omega)} \frac{\sin(\omega z)}{\omega} d\omega + \frac{1}{2} r$$

has been employed. Furthermore, the method of Section 7.3.1 (measurements at the boundary) has been used to fit B_y ; the outermost row of measurements provided a piecewise constant approximation of $V_2(z)$. This resulted in the following expression for B_y at $y = 0$ and $-R < x < R$:

$$B_y(x, z) = \sum_k (B_{k+1} - B_k) \hat{G}_2(x, z - z_{k+1/2}),$$

where B_k denotes the measured value for B_y at $(x, z) = (x_1, z_k)$, and $z_{k+1/2} = (z_k + z_{k+1})/2$. Finally, B_y has been calculated for all x -values at which measurements have been taken, except for the outermost one. The results are shown in Fig. 7.1.

From this figure, we find that there is a good agreement between measured and calculated values of B_y . The small differences between calculations and measurements, most obvious around the physical ends of the quadrupole at $z = 155$ and $z = 455$, arise either from parasitic dipole and/or sextupole contributions to the magnetic field, or from misalignment of the device's design orbit with respect to the grid followed by the Hall probe. A more thorough field calculation, in which these effects are included from the beginning, will take care of this.

In short, the above example shows that the methods developed here can be used to calculate the complete magnetic field of a multipole device, while only a limited amount of measurements is needed as input. By the same methods, other field related quantities, such as the vector potential \mathbf{A} , can be calculated easily from the same measurements.

7.4 Application to charged particle optics

In this section, we derive the Hamiltonian for the motion of a charged particle in a magnetic field, with z as the independent coordinate instead of the time t , and insert the expressions for \mathbf{A} obtained in Section 7.2.3. We outline a numerical method for solving the resulting system of first order non-linear ordinary

differential equations. This method will take both the Hamiltonian nature of the equations and the special nature of the components of \mathbf{A} , which are superpositions of shifted basic functions, into account. We will use this method to show that the total transfer function of the beam guiding device under consideration, can completely be expressed in terms of the field measurements at $r = R$.

Finally, we show how to incorporate the given field description in existing analytical methods that express the transfer function in terms of aberration coefficients. We outline how these coefficients can be obtained, up to any desired degree, directly from field measurements.

7.4.1 Charged particle Hamiltonian

The motion of a particle with mass m_0 and charge Q in the field of a magnetic multipole is governed by the following Hamiltonian:

$$H = \sqrt{m_0^2 c^4 + c^2((p_r - QA_r)^2 + (p_\varphi/r - QA_\varphi)^2 + (p_z - QA_z)^2)}.$$

Since the location of the particle in the z -direction is well defined, while the location in the t -direction is not, it is often desirable to make z the independent variable, instead of t . This is possible for all parts of the particle trajectory where $dz/dt \neq 0$. To obtain a Hamiltonian system for which z is the independent variable, a new canonical momentum p_t is introduced, and the Hamiltonian K for this system is chosen such that the action integral remains invariant:

$$\int (p_k dq^k + p_z dz - H dt) = \int (p_k dq^k + p_t dt - K dz).$$

This way, the canonical form of the equations of motion is preserved. The new Hamiltonian $K(q^k, p_k, t, p_t, z)$ is then obtained by solving the equation $p_t + H = 0$ for p_z :

$$K = -p_z = -QA_z - \sqrt{p_t^2/c^2 - m_0^2 c^2 - (p_r - QA_r)^2 - (p_\varphi/r - QA_\varphi)^2}, \quad (7.19)$$

where the sign of the square root has been chosen such that $dt/dz > 0$. Note that in the case that $\partial H/\partial t = 0$, the order of the system has decreased by two as a result of the exchange. A complete treatise on this method can be found in, among others, Refs. [15,28].

From the Hamiltonian K , the equations of motion for a charged particle in a magnetic field can be derived, with z as the independent variable. The magnetic field description presented in this chapter enters the equations through the vector potential \mathbf{A} . Methods for solving these equations will be presented below.

7.4.2 Calculating transfer functions

In this section, we show how to obtain the transfer function from numerical integration of the equations of motion. The system of equations derived from the Hamiltonian (7.19) will take the following form:

$$(\underline{q}', \underline{p}')(z) = \underline{f}(\underline{q}, \underline{p}, \mathbf{A}(z, \underline{q}), \frac{\partial \mathbf{A}}{\partial \mathbf{x}}(z, \underline{q})). \quad (7.20)$$

Here, $\underline{q} = (r, \varphi, t)$, $\underline{p} = (p_r, p_\varphi, p_z)$, $\mathbf{x} = (r, \varphi, z)$, and f is a given function. The prime ($'$) denotes total differentiation with respect to z .

Note that if \mathbf{A} does not explicitly depend on t , p_t is a constant of motion, and the particle trajectories can be calculated without solving the equations for t and p_t .

As derived in Section 7.2.3, the components of \mathbf{A} and their partial derivatives all take the form

$$\sum_i \lambda_i g_m^k(r, z - z_i) \times \begin{cases} \cos(m\varphi) \\ \sin(m\varphi) \end{cases}.$$

The λ_i originate from measurements at the points $z = w_i$. The z_i were defined by $z_i = \frac{1}{2}(w_i + w_{i+1})$. This provides good approximations for \mathbf{A} and its derivatives in the region $0 \leq r < 1$, $-\pi < \varphi < \pi$, $w_0 < z < w_n$.

Now we proceed to solving the system (7.20) by means of a finite difference method. The discrete version of (7.20) is given by

$$\begin{aligned} (\underline{q}, \underline{p})(w_{i+1}) &= (\underline{q}, \underline{p})(w_{i-1}) \\ &+ (w_{i+1} - w_{i-1}) \underline{f}(\underline{q}(w_i), \underline{p}(w_i), \mathbf{A}(w_i, \underline{q}(w_i)), \frac{\partial \mathbf{A}}{\partial \mathbf{x}}(w_i, \underline{q}(w_i))). \end{aligned} \quad (7.21)$$

We apply the initial condition $(\underline{q}, \underline{p})(w_0) = (\underline{q}_0, \underline{p}_0)$, and calculate $(\underline{q}, \underline{p})(w_1)$ from

$$(\underline{q}, \underline{p})(w_1) = (\underline{q}, \underline{p})(w_0) + (w_1 - w_0) \underline{f}(\underline{q}_0, \underline{p}_0, \mathbf{A}(w_0, \underline{q}_0), \frac{\partial \mathbf{A}}{\partial \mathbf{x}}(w_0, \underline{q}_0)).$$

We then find $(\underline{q}_f, \underline{p}_f) = (\underline{q}, \underline{p})(w_n)$ by repeated application of (7.21). Through repeating this procedure for a number of initial locations, and interpolating between the corresponding final locations, the complete transfer function can be obtained.

It should be noted that the description for \mathbf{A} used in the differential equations, is accurate for all $r \leq 1$, and not only for small r , as in the case of a Taylor expansion. Therefore, a transfer function obtained using the above method is accurate for all $r \leq 1$, and its use is not limited to the paraxial region.

Since the steps of the finite difference method are located precisely at the points $z = w_i$, where the field measurements were performed, we find that there are no interpolation errors in the values for \mathbf{A} used in the calculations. By optimizing the interpolation of the boundary values at $r = 1$, such that not only the boundary values, but also their z -derivatives are matched at the points $z = w_i$, we can also remove interpolation errors from the values of the partial derivatives of \mathbf{A} for better results.

A comparison between the method presented in this section, and common perturbative methods reveals the following. For small to moderate values of r , a perturbative method produces an accurate result with less effort than is needed when using our method. It is in situations where non-linear terms dominate the field description where the present method is most useful, e.g. for large values of r , or close to the edge of a multipole device, or in a cluster of devices, where there may be non-linear coupling of overlapping fields. In such cases, the number of terms needed in a perturbative method, and hence its complexity, increases considerably, while our method can be applied just as easily as to a paraxial case. Also, our method does not have any restrictions on the size of the radius r , where a perturbative method of a given order is usually only applicable at radii smaller than a well-defined maximum.

However, even in cases where a perturbative method is the most logical choice, the field description presented in this chapter can be used to obtain the coefficients of the power series terms. This will be demonstrated in the next section.

7.4.3 Incorporating the field description into existing results

As mentioned in Section 7.1, many methods have been developed to obtain analytical expressions for the transfer function of a given magneto-optical device. They usually start from the Hamiltonian for the motion for a charged particle in the magnetic field of the device; then both the magnetic vector potential and the Hamiltonian are approximated by a truncated power series. The truncated Hamiltonian is then used to derive approximated, but still Hamiltonian, equations of motion, for which solutions are obtained using successive substitution. These solutions are mostly given in the form of power series approximations with respect to the transverse coordinates x and y , and the coefficients of the various terms, the so-called aberration coefficients, are expressed in terms of the generalized gradients of the vector potential.

These generalized gradients are in fact the on-axis radial derivatives of \mathbf{A} , which can hardly be obtained from direct measurements, but can, on the other hand, easily be obtained by expanding the basic functions in the description of \mathbf{A} into

powers of r (see also Ref. [19]):

$$g_m^k(r, z) = \sum_{l=0}^{\infty} g_{ml}^k(z) r^{k+2l},$$

$$g_{ml}^k(z) = \begin{cases} \frac{(-1)^n}{2^{k+2l} l! (k+l)! 2\pi} \int_{-\infty}^{\infty} \frac{\omega^{k+2l}}{I_m(\omega)} \cos(\omega z) d\omega, & k - m = 2n, \\ \frac{(-1)^{n+1}}{2^{k+2l} l! (k+l)! 2\pi} \int_{-\infty}^{\infty} \frac{\omega^{k+2l}}{I_m(\omega)} \sin(\omega z) d\omega, & k - m = 2n + 1. \end{cases}$$

The expansion of g_m^k into powers of r is then inserted into (7.10), in order to obtain expansions of $\mathcal{J}_k A_m$, and finally of \mathbf{A} , into powers of r . From (7.16), we find that the z -dependent coefficients in this expansion, i.e. the generalized gradients, are obtained in the same way as the coefficients $J_k A_m$. In other words, we can easily derive accurate approximations for the desired generalized gradients, in term of measurements far from the axis, up to any desired order. This provides a very convenient way to incorporate the field descriptions outlined in this chapter into existing analytical work.

7.5 Conclusions

In this chapter, we have developed a new way of describing the magnetic fields of beam guiding elements. The new description provides a way to describe these fields, and thus the action of beam guiding elements on charged particles, more accurately than can be reached by means of a common power series expansion. Eventually, this will result in a more accurate control over the quality of a bunch of charged particles after acceleration.

The magnetic field inside a magneto-optical device, and its harmonic scalar and vector potentials, have been explored in the area $0 \leq r < R$, $-\infty < z < \infty$. The various multipole contributions to these quantities have been fitted using field measurements at the boundary $r = R$ and shifted basic functions. The same set of measurements and shifts can be used to fit various field-related quantities. The developed procedure is independent of the exact form of the boundary conditions and can be used to fit the field of one device or a cluster of various consecutive devices.

All the approximations of the field and its scalar and vector potentials satisfy Maxwell's equations; the approximated potentials are harmonic, which allows one to apply harmonic potential theory whenever necessary.

As the definitions of the basic functions corresponding to the quantities to be fitted do not contain any physical parameters, these functions can be calculated in advance to the desired accuracy, which greatly simplifies the fitting procedure.

The procedure works for any order multipole contribution, but will be the most useful for lower order contributions, since higher order contributions are more difficult to obtain from measurements, while their effect on particle trajectories will often be small.

Once the vector potential \mathbf{A} for the field of a multipole device has been calculated, it can be inserted in the Hamiltonian equations of motion for a charged particle passing through the device. The description of \mathbf{A} in terms of z -shiftings of basic functions can conveniently be combined with the numerical integration of these equations. Furthermore, by expanding \mathbf{A} into powers of r , descriptions of the z -dependent generalized gradients in terms of z -shifts of basic functions are obtained. These descriptions are much more accurate than descriptions in terms of on-axis field derivative measurements, and are an excellent way of combining the methods presented in this chapter with a lot of existing analytical work on particle optics.

The methods developed in this chapter can also be used for the calculation of the fields and transfer functions of static electro-optical devices, since the electric field of such a device can be written in the same form that was used for the field of a magneto-optical device, and the calculation of the transfer function is similar to the magneto-optical case.

Acknowledgments

The authors wish to thank J. Boersma for his valuable contributions to the theory presented in this chapter.

References

- [1] A. El-Kareh and J. El-Kareh, *Electron beams, lenses, and optics* (Academic Press, New York and London, 1970), Vol. 1.
- [2] D.L. Smith, Nucl. Instrum. Meth. **79**, 144 (1970).
- [3] G.E. Lee-Whiting, Nucl. Instrum. Meth. **99**, 609 (1972).
- [4] G.E. Lee-Whiting, Nucl. Instrum. Meth. **83**, 232 (1970).
- [5] H. Matsuda and H. Wollnik, Nucl. Instrum. Meth. **103**, 117 (1972).
- [6] H. Nakabushi and T. Matsuo, Nucl. Instrum. Meth. **198**, 207 (1982).
- [7] H.L. Hagedoorn, J.I.M. Botman, P.H.A. Mutsaers, and M.J.A. de Voigt, Nucl. Instrum. Meth. B **68**, 23 (1992).

-
- [8] R.W. de Leeuw, P.H.A. Mutsaers, A.V.G. Mangnus, and H.L. Hagedoorn, Nucl. Instrum. Meth. B **83**, 284 (1993).
 - [9] G.W. Grime *et al.*, Nucl. Instrum. Meth. **197**, 97 (1982).
 - [10] G.W. Grime and F. Watt, *Beam optics of Quadrupole Probe-Forming Systems* (Adam Hilger Ltd., Bristol, 1984).
 - [11] K.L. Brown, F. Rothacker, D.C. Carey, and Ch. Iselin, *TRANSPORT, a computer program for designing charged particle beam transport systems*, CERN Report No. CERN 80-04, 1980 (<http://weblib.cern.ch/>).
 - [12] D.C. Carey, K.L. Brown, and Ch. Iselin, *Decay TURTLE, A computer program for simulating charged particle beam transport systems, including decay calculations*, SLAC Report No. SLAC-R-246, 1982.
 - [13] M. Berz, in *Proceedings of the 13th Particle Accelerator Conference, 1989, Chicago* (IEEE, Piscataway, NJ, 1989), Vol. 3, p. 1419.
 - [14] M. Berz, *COSY INFINITY, Version 8 reference manual*, Technical Report MSUCL-1088, National Superconducting Cyclotron Laboratory, Michigan State University, East Lansing, MI, 1997 (<http://cosy.nsl.msu.edu/>).
 - [15] A.J. Dragt, in *Physics of High Energy Particle Accelerators*, edited by R. A. Carrigan, F. R. Huson, and M. Month, AIP Conf. Proc. No. 87 (A.I.P., New York, 1982), pp. 147-313; A. Dragt, *Lie Methods for Nonlinear Dynamics with Applications to Accelerator Physics*, University of Maryland Physics Department Report, 1998 (unpublished).
 - [16] A. Dragt *et al.*, Ann. Rev. Nucl. Part. Sci **38**, 455 (1988).
 - [17] A. Dragt *et al.*, *MARYLIE 3.0, a program for charged particle beam transport based on Lie algebraic methods*, University of Maryland Physics Department Report, 1995 (unpublished).
 - [18] M. Szilagy, *Electron and Ion optics* (Plenum Press, New York, 1988).
 - [19] M. Venturini and A.J. Dragt, Nucl. Instrum. Meth. A **427**, 387 (1999).
 - [20] S. Caspi, IEEE Trans. Magn. **30**, 2419 (1994).
 - [21] R. Trines, Master of Science Thesis, Eindhoven University of Technology internal report, VDF/NK 98-04, 1998 (unpublished).
 - [22] R. Trines *et al.*, in *Proceedings of the 6th European Particle Accelerator Conference, 1998, Stockholm, Sweden*, edited by S. Myers *et al.* (Institute of Physics, Bristol, UK, 1998), 1969.

-
- [23] R. Trines et al., in *Proceedings of the 18th Particle Accelerator Conference, 1999, New York City*, edited by A. U. Luccio and William W. MacKay (IEEE, Piscataway, NJ, 1999), 1539.
 - [24] P.R. Garabedian, *Partial differential equations* (Wiley, New York, 1964).
 - [25] L. Hörmander, *The Analysis of Linear Partial Differential Operators III*, Grundlehren des mathematischen Wissenschaften 274 (Springer-Verlag, Berlin, Heidelberg, 1985).
 - [26] I. Rubinstein and L. Rubinstein, *Partial differential equations in classical mathematical physics* (Cambridge University Press, 1993).
 - [27] G. J. L. M. Brooijmans, Eindhoven University of Technology internal report, VDF/NK 88-15, 1988 (unpublished).
 - [28] H. L. Hagedoorn, J. I. M. Botman, and W. J. G. M. Kleeven, *Hamilton theory as a tool for accelerator physics*, CERN Accelerator School sept. 1991, Noordwijkerhout, CERN Report No. CERN 92-01, 1992 (<http://weblib.cern.ch/>).

Appendix A

The code XOOPIC

For the simulations in Chapter 6, the 1-D version of the particle-in-cell (PIC) code XOOPIC has been used. This code is a combination of the XGrafix graphical package by V. Vahedi and J.P. Verboncoeur [1] and the PIC code OOPIC by J.P. Verboncoeur, A.B. Langdon, and N.T. Gladd [2]. The code OOPIC is originally a 2-D code, from which a 1-D version has been derived by J. Wurtele and H.-J. Lee. The copyright resides with the Regents of the University of California. The code can be freely used and modified for scientific, non-commercial purposes, provided Ref. [2] is cited in every resulting publication. In this appendix, a short description of the algorithms in OOPIC will be given, as well as an overview of the benchmarks we have performed on it. A recent version can be obtained from either Tech-X Corporation (<http://www.techxhome.com>), which has taken over most of the development of the code, or the Plasma Theory and Simulation Group of UC Berkeley (<http://ptsg.eecs.berkeley.edu>).

A.1 Description of the code

A particle-in-cell code is a type of code for simulation of plasma motion, in which the plasma is represented by a large number of so-called macroparticles. Each macroparticle in turn represents a very large number of actual plasma particles (electrons or ions). The electromagnetic fields are known on the points of a (usually rectangular) mesh only, whereas the particles are not confined to the mesh points, but can move freely with respect to the mesh. A space enclosed by neighbouring mesh points is called a *cell*. The charge and current densities, needed to calculate the fields, are obtained by distributing the charge q and current $q\mathbf{v}$ of each particle over the mesh points that define the cell in which the particle resides, hence the name particle-in-cell.

XOOPIC is an object-oriented code, written in C++. The choice of this language over Fortran has been motivated by the much greater flexibility of C++. A major

drawback of C++ compared to Fortran is that automatic optimization of Fortran code produces better results, i.e. a Fortran program will often run faster than a similar program written in C++, if no special measures are taken. To overcome this problem, XOOPIC has been hand-optimized for the most part while it was being written.

XOOPIC is a fully relativistic electromagnetic particle-in-cell code that solves the time-dependent Maxwell equations using Yee's algorithm [3], a second-order leap-frog algorithm. Particle motion is governed by the Lorentz force. It supports a number of boundary conditions for incoming and outgoing waves, dielectric surfaces that also absorb particles, and particle emitters. It also contains routines for Monte Carlo collisions (MCC) and Monte Carlo tunneling ionization (MCTI) in the presence of a neutral gas. A moving simulation window can be employed e.g. to follow a laser pulse, which often saves a lot of time.

In this section, the inner workings of XOOPIC will be described, with the exception of the MCC and MCTI routines, since they have not been used in the simulations presented in this thesis.

A.1.1 General outline

The code consists of two stages: the initialization stage and the simulation stage. In the initialization stage, an input file containing the details of the simulation setup is read and parsed. Then the simulation is initialized in the following steps:

1. Build a list of all species of particles (electrons, ions) specified in the input file,
2. Create a rectangular mesh for the simulation box. The x -coordinate is the longitudinal coordinate,
3. Create the objects that will impose the boundary conditions (BCs) during the simulation,
4. Initialize the fields to zero,
5. Initialize the MCC and MCTI routines, if necessary,
6. Load the particles into the cells,
7. Initialize the charge and current densities, including contributions from dielectric boundaries,
8. If applicable, add initial values for the fields to the current fields.

After the initialization has been completed, the simulation loop is started. This loop is repeated as often as necessary, and consists of the following steps:

1. Advance the fields one timestep dt and apply the appropriate field BCs,
2. Advance the particles one timestep and apply the appropriate particle BCs,
3. Project charge and current densities on the mesh,
4. Emit and/or absorb particles at the boundaries,
5. Invoke MCC and MCTI routines, if necessary,
6. In case of a moving window simulation, shift the contents of the simulation box (fields and particles) one cell to the left, initialize the fields at the rightmost mesh node, and load a fresh batch of particles into the rightmost cell.

In the following subsections, the advance of fields and particles, and the application of boundary conditions, will be described in greater detail.

A.1.2 Fields

XOOPIC calculates the electromagnetic fields from the charge and current densities from the time-dependent Maxwell equations:

$$\frac{d}{dt} \int \mathbf{D} \cdot d\mathbf{S} = \oint \mathbf{H} \cdot d\mathbf{l} - \int \mathbf{J} \cdot d\mathbf{S}, \quad (\text{A.1})$$

$$\frac{d}{dt} \int \mathbf{B} \cdot d\mathbf{S} = - \oint \mathbf{E} \cdot d\mathbf{l}, \quad (\text{A.2})$$

where the line integrals are along cell sides, and the surface integrals are across cell faces. Here, $\mathbf{E} = \mathbf{C}^{-1} \cdot \mathbf{D}$ and $\mathbf{H} = \mathbf{L}^{-1} \cdot \mathbf{B}$, where \mathbf{C}^{-1} and \mathbf{L}^{-1} are diagonal matrices with the dimensionality of capacitance and inductance, respectively. Projection of charges and currents on the grid are done in such a way that the continuity equation is always satisfied:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0. \quad (\text{A.3})$$

The following two equations are then satisfied by default:

$$\frac{\partial}{\partial t}(\nabla \cdot \mathbf{B}) = 0, \quad \frac{\partial}{\partial t}(\nabla \cdot \mathbf{D} - \rho) = 0.$$

So if the conditions $\nabla \cdot \mathbf{B} = 0$ and $\nabla \cdot \mathbf{E} = \rho$ are satisfied from the start, they will be satisfied during the whole course of the simulation automatically. For

the simulations performed for this thesis, the conditions on \mathbf{E} and \mathbf{B} have been satisfied by setting $\mathbf{B} = \mathbf{E} = \mathbf{0}$ and $\rho = 0$ at the start of the simulation.

The mesh on which the fields and charge and current densities are known is a Yee mesh [3], meaning that the transverse electric field and current density, the longitudinal magnetic field, and the charge density are known on integer grid points, while the longitudinal electric field and current density and the transverse magnetic field are known on half-integer grid points.

The fields are advanced in time as follows:

1. Advance \mathbf{B} over $dt/2$, using (A.2) and an Euler forwards algorithm,
2. Advance \mathbf{E} over dt , using (A.1) and an Euler forwards algorithm,
3. Advance \mathbf{B} for another time step $dt/2$.

The resulting leap-frog scheme ensures that the calculation of the fields is second order in dt . The advance of \mathbf{B} takes two stages to ensure that both \mathbf{E} and \mathbf{B} are known at the same time, and not at alternating times.

A.1.3 Particles

Particle motion is calculated using the equations

$$\frac{d\mathbf{x}}{dt} = \mathbf{v}, \quad (\text{A.4})$$

$$\frac{d\mathbf{u}}{dt} = \frac{q}{m}(E + \mathbf{v} \times \mathbf{B}), \quad (\text{A.5})$$

where $\mathbf{u} = \mathbf{v}/\sqrt{1-v^2}$. In the numerical implementation, the position x is known at time t at which the fields are also known, while the “momentum” \mathbf{u} is known at time $t - dt/2$. This way, integration of (A.4)-(A.5) is done through a second order leap-frog algorithm.

The integration of (A.5) works as follows:

1. First, the particle is accelerated over $dt/2$ by the electric field only, $\mathbf{u}_t = \mathbf{u}_{t-1/2} + qdt/(2m) \cdot \mathbf{E}(x_t)$,
2. Then the particle’s velocity is rotated by the magnetic field through Boris’ algorithm [4]:

$$\mathbf{w} = \frac{qdt}{2m\sqrt{1+(u_t/c)^2}}\mathbf{B}(x_t), \quad \mathbf{s} = \frac{2\mathbf{w}}{1+w^2}, \quad \mathbf{u}'_t = \mathbf{u}_t + (\mathbf{u}_t + \mathbf{u}_t \times \mathbf{w}) \times \mathbf{s},$$

3. Finally, the particle is again accelerated over $dt/2$ by the electric field only:
 $\mathbf{u}_{t+1/2} = \mathbf{u}'_t + qdt/(2m) \cdot \mathbf{E}(x_t)$.

This integration scheme is based on the notion that the force exerted by the magnetic field does not change the magnitude of \mathbf{u} while at the same time, it does depend on that magnitude. Therefore, the rotation of \mathbf{u} by \mathbf{B} needs to happen halfway the acceleration by \mathbf{E} to ensure that integration is of second order in dt .

Since the fields are only known on (integer or half-integer) mesh points while the position x_t can be anything, the fields need to be interpolated to obtain $\mathbf{E}(x_t)$ etc. In XOOPIC, a linear interpolation of the fields between the mesh points nearest to x_t has been employed. This interpolation is also called *bilinear weighting*.

Integration of (A.4) is then quite straightforward:

$$\mathbf{x}_{t+1} = \mathbf{x}_t + \frac{dt}{1 + (u_{t+1/2}/c)^2} \mathbf{u}_{t+1/2}.$$

Following the update of particle positions and momenta, the charge and current of each particle need to be projected onto the mesh in order to obtain the charge and current density. The projection of the charge density uses bilinear weighting, as does the current projection in a direction perpendicular to \mathbf{u} . In the direction of \mathbf{u} , the current by that particle is allotted entirely to the nearest grid point. The reason for this is, that the projection of charge and current density needs to satisfy the continuity equation (A.3), and a completely bilinear projection would fail to do so.

When a particle moves across a cell edge, its motion is split up, to ensure that each cell receives an appropriate fraction of the charge and current of that particle.

A.1.4 Boundary conditions

The boundary conditions (BCs) in XOOPIC can roughly be divided into two groups: *dielectric* BCs and *port* BCs. A dielectric BC corresponds to a wall of dielectric or conducting material, while a port BC corresponds to an open end. A dielectric BC can also be used to emit particles into the simulation box, while a port BC can be used to introduce EM waves, e.g. a laser pulse, into the simulation.

A particle that collides into a boundary is always removed from the simulation. In case of a port BC, which represents an open end, the particle is merely deleted. In case of a dielectric BC, the particle's charge will be collected on the dielectric's surface, and the particle may even be reflected back into the simulation box.

An EM wave hitting a dielectric BC is reflected back into the box. A port BC on the other hand absorbs any incident EM waves. This is done by separating

the EM waves one cell away from the boundary into ingoing and outgoing parts. The outgoing part is then stored for two time steps, and subsequently used to remove the reflected outgoing (now ingoing) fields from the EM fields directly at the boundary.

XOOPIC contains a number of port BCs designed to launch EM waves. The BC that is used to introduce laser pulses into the simulation is one of them. They work by imposing a time-dependent value for the electric field only at the location of the boundary. It is then left to the routine that solves the EM field equations to pull the wave farther into the simulation box. The polarization is always linear, on the simple grounds that circular polarization has not been implemented to date.

A.2 Benchmarks

Several benchmarks have been performed on XOOPIC to determine the reliability of its results. The focus has been on the behaviour of Raman instabilities. The following areas of interest are discerned: (i) dispersion relations, (ii) RFS seeding and beat-wave excitation, and (iii) growth rates.

For all simulations, the plasma density is either homogeneous or has the flat-top profile as described in Chapter 6. The peak density is $3.8 \cdot 10^{19} \text{ cm}^{-3}$ and the number of particles per cell is 32 at peak density. We found that for a lower number of particles per cell the (anti-)Stokes peaks in both the frequency and the wavenumber spectrum tend to be too close to the main laser peak. This is possibly caused by the rather crude way in which XOOPIC projects the particles on the grid to obtain charge and current densities. On the other hand, using more particles per cell together with a homogeneous plasma density reduces the noise to an extent that the RBS instability does not have enough to grow on. This does not affect inhomogeneous density profiles though.

The laser pulse has a central wavelength of 800 nm and a Gaussian profile with a FWHM of 50-100 fs. It has a (scaled) peak amplitude a_0 of 0.1-0.5, and is linearly polarized.

A.2.1 Dispersion relations

An EM wave in a homogeneous plasma satisfies the dispersion relation $\omega^2 = \omega_p^2 + c^2 k^2$ where ω and k are the frequency and wave number of the EM wave and ω_p denotes the plasma frequency. This relation holds for the main laser wave (ω_0, k_0) as well as for the (anti-)Stokes satellite waves (ω_{\pm}, k_{\pm}) where $\omega_{\pm} = \omega_0 \pm \omega_p$ and $k_{\pm} = k_0 \pm \omega_p/c$.

From the given values for the laser wavelength and the plasma density it can be inferred that the laser frequency and wavelength satisfy $ck_0 \approx 0.99\omega_0$. To verify this, a simulation has been performed in which the laser pulse (50 fs, $a_0 = 0.5$) traverses a plasma of constant density. During the simulation, the transverse electric field has been recorded versus time at a fixed position, and versus position at a fixed time. From these, a frequency and a wavenumber spectrum have been determined, and the predicted relation between k_0 and ω_0 has indeed been recovered.

Similarly, the backward-going RBS EM wave (ω_-, k_-) needs to satisfy $ck_- \approx 0.985\omega_-$ and $|\omega_- - \omega_0| \approx 0.9\omega_p$, where the factor 0.9 stems from relativistic detuning of the plasma frequency due to the transverse quiver motion of the plasma electrons. From the spectra, we found that $ck_- \approx \omega_-$ and that $|\omega_- - \omega_0| \approx (0.8 \pm 0.1)\omega_p$. More accurate results could not be obtained due to the amount of noise in the RBS fields and spectra.

The Langmuir wave associated with RBS should theoretically have frequency $0.9\omega_p$ and wavenumber $k_L \approx 2k_0 - 0.9\omega_p/c$. From the phase space plots of the plasma electrons, we found that $\tau_L \approx 20 \pm 2$ fs and $\lambda_L \approx (430 \pm 5)$ nm. Together with $\omega_p = 3.48 \cdot 10^{14} \text{ s}^{-1}$ and $\lambda_0 = 808$ nm (in plasma), we find that $\omega_L \approx (0.9 \pm 0.1)\omega_p$ and $|k_L - 2k_0| \approx (0.9 \pm 0.1)\omega_p/c$. These values are in agreement with theoretical predictions, but it should be noted that the accuracy is low due to lack of a regular structure in the Langmuir wave (much noise, wave breaking at every wave tip, etc.).

In general, it looks like the frequencies and wave numbers of the EM and Langmuir waves we encounter in a simulation in the weakly nonlinear regime satisfy the dispersion relations for Raman backscattering. However, the value for the actual plasma frequency we found in the simulations tends to be below the theoretically predicted value of $0.9\omega_p$, an effect that decreases with increasing number of particles per cell. This supports the conjecture that the code could do with better routines for the projection of the charge and current densities on the grid and the Lorentz force on the particles.

A.2.2 RFS seeding and beat-wave excitation

Beat wave excitation, i.e. using a dual-frequency laser pulse with frequency difference ω_p and equal amplitudes for both frequencies, has been shown to increase the wakefield amplitude through enhancement of RFS already by Joshi et al. [5]. RFS seeding, i.e. adding a small amount of Stokes wave to the main laser pulse, has been shown to boost RFS by Fisher and Tajima [6]. Suppression of RBS by RFS through mode coupling has been proposed by Everett et al. [7], who demonstrated this through PIC-simulations and a beat-wave experiment.

All these results have been obtained in the linear regime and their reproduction

is viewed to be a good benchmark for the code.

To reproduce the results on seeding, we performed simulations using a 100 fs laser pulse having $a_0 = 0.2$ and $\lambda_0 = 800$ nm to which a 2% or 5% seed (frequency downshifted by ω_p) has been added. Note that Fisher and Tajima used a 3% seed. For the plasma, a non-constant flat-top profile has been used, to ensure that there would be sufficient RFS growth. In line with the results of Fisher and Tajima, we found that seeding led to an increased RFS growth, a larger and more regular wakefield, and pulse envelope modulation on the scale of λ_p , as compared to the unseeded case. At later simulation times, we also found that (anti-)Stokes peaks emerge that are shifted by non-integer multiples of ω_p with respect to the main laser frequency. Their presence leads to depletion of the laser pulse and eventual degradation of the wakefield, an effect also shown in Ref. [6]. We also found that all the effects of seeding increased with increasing seed amplitude.

Side note: Our simulations also showed that seeding leads to a decrease in RBS, already in the linear regime considered here. This can probably be explained by the mode coupling mechanism of Everett et al. [7].

To reproduce the results of the beat-wave experiment of Everett et al, a simulation has been performed using two laser pulses both having $a_0 = 0.1$, one at 800 nm and one downshifted by ω_p . The plasma slab was identical to that of the seeded simulations. The effects observed were similar to the effects of seeding, only stronger: a very quick growth of RFS and of the wakefield, rapid suppression of any RBS, severe pulse envelope modulation, and soon after that the onset of non-integer (anti-)Stokes peaks, resulting in the complete degradation of both pulse and wakefield. These phenomena are similar to what has been observed experimentally by Everett et al.

Side note: A very small fraction of the plasma electrons has been found to get trapped in and accelerated to about 7.5 MeV by the wakefield, an effect not seen in normal or seeded simulations at this low laser intensity.

A.2.3 Growth rates

A proper investigation of the growth rates of SRS in XOOPIC simulations turned out to be beyond the scope of this thesis. To date, the only result obtained so far is the agreement between analytical and numerical values for the seed amplitude necessary to suppress RBS in an RFS-seeded simulation (cf. Section 6.2). However, there are strong indications that a quantitative reproduction of theoretical growth rates will be quite difficult, since these are usually derived under circumstances (long pump pulses, long plasmas) that are hard to reproduce in simulations because of sheer simulation time. What can and should be verified in the future is how the RFS and RBS growth rates scale with quantities such as the laser intensity and the plasma density. This should be done for single pulse

cases first, and subsequently extended to seeded pulse and beat-wave cases. The amount of RFS necessary to suppress RBS should also be determined for these cases.

We also feel that recovering how the growth rates scale with laser and plasma parameters rather than trying to obtain a quantitative agreement between theory and simulations, is in character with the results presented in Chapter 6, which do not provide any “hard numbers” but are mostly about qualitative behaviour anyway.

References

- [1] V. Vahedi and J.P. Verboncoeur, *XGraftx: An X-Windows Environment for Real-Time Interactive Simulations*, in: Proc. 14th Int. Conf. Num. Sim. Plasmas, Annapolis, MD (1991).
- [2] J. P. Verboncoeur, A. B. Langdon, and N. T. Gladd, Comput. Phys. Commun. **87**, 199 (1995).
- [3] K. S. Yee, IEEE Trans. Antennas Propag. **14**, 302 (1966).
- [4] C.K. Birdsall and A.B. Langdon, *Plasma Physics Via Computer Simulation* (Adam Hilger, Bristol, 1991).
- [5] C. Joshi, W. B. Mori, T. Katsouleas, and J. M. Dawson, Nature **311**, 525 (1984).
- [6] D. L. Fisher and T. Tajima, Phys. Rev. E. **53**, 1844 (1996).
- [7] M. J. Everett *et al.*, Phys. Rev. Lett. **74**, 2236 (1995); Phys. Plasmas **3**, 2041 (1996).

Samenvatting

Het onderwerp van dit proefschrift is het versnellen van elektronen in een plasmagolf, en dan vooral de productie van snelle elektronen door een laserpuls van hoge intensiteit op een plasma af te vuren. Voor een beter begrip van dit onderwerp is een inleiding tot plasma's en plasmaversnelling op zijn plaats.

Hoog-energetische geladen deeltjes hebben zeer veel toepassingen. Ze worden bijvoorbeeld gebruikt in Röntgenbuizen, lineaire versnellers voor radiotherapie, voor de productie van radio-isotopen voor medische diagnostiek, enzovoorts. In fundamenteel onderzoek naar de structuur van alle materie om ons heen laat men sterk versnelde deeltjes op elkaar botsen, om uit de brokstukken gegevens over die structuur af te leiden.

Conventionele deeltjesversnellers werken met elektromagnetische velden in vacuüm. Een probleem hierbij is dat er in vacuüm vrij snel doorslag optreedt als de velden te sterk worden. Om dan toch tot zeer hoge energieën te komen moeten de deeltjes over afstanden tot tientallen kilometers versneld worden. Dat het bouwen van een kilometerslange structuur om deeltjes te versnellen kapitalen kost behoeft geen betoog. Als gevolg hiervan zijn er slechts enkele deeltjesversnellers in de wereld die werkelijk aan het front van de versnelfysica kunnen opereren. In de nabije toekomst zullen er niet meer van dergelijke versnellers bij komen, laat staan versnellers met een nog groter energiebereik. Hierdoor treedt er ernstige stagnatie op in het onderzoek in de ultrahoge-energiefysica.

Om dit probleem te omzeilen hebben onderzoekers gezocht naar een medium dat sterkere elektromagnetische velden kan tolereren dan vacuüm. Versnelling in zo'n medium kan dan over veel kortere afstanden plaatsvinden. Al snel kwam men terecht bij een *plasma*, niet te verwarren met bloedplasma, als medium. Een plasma is een gas dat wordt verhit tot enkele tienduizenden graden Celsius. Bij zulke temperaturen gaan de buitenste elektronen van de gasatomen loslaten. Er ontstaat dan een mengsel van losse elektronen (negatief geladen) en atomen die een elektron missen (ionen, positief geladen): het plasma. De negatieve en positieve ladingen in het plasma trekken elkaar zeer sterk aan, wat het plasma bij elkaar houdt en zorgt dat het lokaal ongeladen blijft. Het is gebleken dat de velden die in een plasma kunnen worden opgewekt duizend tot tienduizend keer zo sterk kunnen zijn als in vacuüm voordat er doorslag optreedt, waardoor het

een veelbelovend medium voor de nieuwe generatie deeltjesversnellers is.

Om dergelijke sterke velden op te wekken in het plasma moeten lokaal de positieve en negatieve ladingen van elkaar gescheiden worden. Dit is niet eenvoudig omdat deze ladingen elkaar zo sterk aantrekken. Het is echter mogelijk gebleken om de ladingen uit elkaar te drijven met behulp van een bijzonder sterke en korte laserpuls. Als een dergelijke puls zich door een plasma beweegt, worden de relatief lichte elektronen door de laserpuls opzij geduwd, terwijl de veel zwaardere ionen nauwelijks reageren. Hierdoor ontstaat er vlak achter de puls een gebiedje met een netto positieve lading. Na passage van de laserpuls proberen de elektronen zo snel mogelijk dit “gat” weer op te vullen, en klotsen nog een paar keer na alvorens tot rust te komen. Hierdoor ontstaat er achter de puls een hekgolf als van een speedboot, waarin geladen deeltjes kunnen worden versneld. De frequentie van deze golf is de eigenfrequentie van het plasma, de zogenaamde *plasmafrequentie*.

Het versnellen van een kluitje geladen deeltjes, meestal elektronen, in de hekgolf van een laserpuls gaat een beetje als surfen op de golven bij Hawaii. Nadat de hekgolf is opgewekt wordt het kluitje elektronen in de golf ingeschoten. Als de snelheid van de elektronen te laag is zal de golf langs het kluitje spoelen zonder het mee te nemen, en als de snelheid te hoog is zal de golf niet veel kracht op de elektronen uit kunnen oefenen. Als het kluitje precies de goede snelheid heeft wordt het door de golf meegesleurd en kan het tot zeer hoge energieën worden versneld. Ook het moment van inschieten is van belang: als de elektronen op het verkeerde moment worden ingeschoten zal de golf ze afremmen in plaats van versnellen. Verder moet het kluitje elektronen bij voorkeur korter zijn dan de golflengte van de hekgolf, omdat het anders in stukken breekt, wat meestal niet de bedoeling is.

In dit proefschrift zijn een aantal aspecten van de hier beschreven laser-plasma-versnelling onderzocht. Het onderzoek is theoretisch van aard: voor de dynamica van zowel de plasma-elektronen als de te versnellen elektronen zijn vergelijkingen opgesteld die deels analytisch, deels numeriek zijn opgelost. Het zwaartepunt ligt bij de injectie van te versnellen elektronen in de hekgolf van de laserpuls, maar er zijn ook aspecten van het gedrag van het kluitje tijdens de versnelling bekeken. Ook is onderzocht hoe het kluitje na versnelling zonder al teveel kwaliteitsverlies bij zijn bestemming kan worden afgeleverd.

Het proefschrift is als volgt opgebouwd. Na een meer uitgebreide inleiding in hoofdstuk 1 en de behandeling van de basistheorie in hoofdstuk 2 is in hoofdstuk 3 onderzocht in hoeverre de energiespreiding van de elektronen in het kluitje onder controle kan worden gehouden tijdens het versnelproces. Omdat het kluitje in het algemeen toch nog vrij groot zal zijn ten opzichte van de golflengte van de hekgolf, en omdat de elektronen elkaar beïnvloeden tijdens het versnelproces, kan de energiewinst per elektron nogal variëren. Door middel van simulaties is bekeken hoe de energiespreiding na versnelling afhangt van parameters als het

moment van inschieten, de lengte van het kluitje elektronen en de hoeveelheid lading in het kluitje. Hierbij is geprobeerd de energiewinst van de elektronen zo groot mogelijk te krijgen onder de voorwaarde dat de energiespreiding minimaal bleef. Het blijkt dat de parameters van de kluit zo te kiezen zijn dat een hoge energiewinst (1 GeV) te behalen is bij een lage energiespreiding (1%). Helaas bleek ook dat de energiespreiding erg gevoelig is voor afwijkingen in met name het moment van inschieten, de zogenaamde *inschietfase*.

In het algemeen wordt de kluit te versnellen elektronen buiten het plasma gemaakt en daarna pas in het plasma ingebracht. Bij deze methode is het amper mogelijk de inschietfase met voldoende nauwkeurigheid in te stellen. Daarom is er gekeken naar een alternatieve methode van injectie. Het is namelijk ook mogelijk de te versnellen elektronen uit het plasma zelf te halen d.m.v. breking van de hekgolf. Dit proces, *interne injectie* genaamd, lijkt wel wat op de breking van golven aan het strand. Het voordeel hierbij is dat op deze wijze ingevangen elektronen vanzelf de juiste fase t.o.v. de plasmagolf hebben. Deze methode is bestudeerd in twee afzonderlijke regimes. In het ene regime is de laserpuls korter dan een enkele golflengte van de hekgolf, in het andere is de laserpuls enkele “hekgolflengten” lang.

In hoofdstuk 4 zijn golfbreking en interne injectie bestudeerd voor een laserpuls korter dan een hekgolflengte. Eerst is er een overzicht gegeven van bestaande theorieën over golfbreking in een plasma. Het blijkt dat eenieder zo zijn eigen definitie van golfbreking hanteert, wat niet tot een beter begrip van de materie leidt. Vervolgens is er een model ontwikkeld voor golfbreking in een koud plasma, waaruit kan worden voorspeld onder welke omstandigheden de hekgolf breekt en hoeveel plasma-elektronen hierbij in de hekgolf worden geïnjecteerd. Het blijkt dat het golfbrekiingsproces zich achter de laserpuls afspeelt. Hierdoor hebben vorm en intensiteit van de puls wel invloed op de amplitude van de hekgolf en daardoor indirect ook op de hoeveelheid ingevangen elektronen, maar niet op het golfbrekiingsproces zelf. De kluit aldus ingevangen elektronen heeft dan globaal steeds dezelfde eigenschappen: een zeer goede inschietfase voor versnelling, erg kort ten opzichte van de hekgolflengte, maar ook een grote energiespreiding, tot 80% van zijn gemiddelde energie.

In hoofdstuk 6 zijn golfbreking en interne injectie bestudeerd voor laserpulsen die enkele hekgolflengten lang zijn. In dit regime blijkt de wisselwerking tussen puls en plasma geheel te worden gedomineerd door de zogenoemde Ramanverstrooiing, beschreven in hoofdstuk 5. Ramanverstrooiing is een gevolg van het feit dat de frequentie van het gebruikte laserlicht meestal nogal verschilt van de plasmafrequentie. Hierdoor ontstaan er zwevingen tussen deze beide frequenties. Het gevolg is dat invallend laserlicht verstrooid wordt aan het plasma, waarbij het verstrooide licht een andere frequentie heeft dan het invallende. Deze zwevingen komen alleen goed tot uiting als de laserpuls langer is dan één golflengte van de hekgolf. Hierdoor treden ze in het regime beschreven in hoofdstuk 4 niet

op. Afhankelijk van de richting van het verstrooide licht ten opzichte van het invallende licht spreken we van terugwaartse, zijwaartse en voorwaartse Raman-verstrooiing.

In het door Ramanverstrooiing gedomineerde regime is de invloed van diverse laser- en plasmaparameters zoals vorm en amplitude van de laserpuls, dichtheidsprofiel van het plasma en “chirp” (frequentieverloop binnen de laserpuls) op het golfbrekingsproces en de elektronenopbrengst numeriek onderzocht, deels in samenwerking met het Lawrence Berkeley Laboratory in Berkeley, Californië. Het blijkt dat de rol die deze parameters spelen grotendeels kan worden verklaard uit hun invloed op beide typen verstrooiing. Dit hangt samen met het feit dat zowel de Ramanverstrooiing als de golfbreking en elektronenvangst zich afspelen binnen de laserpuls zelf, waardoor deze processen elkaar sterk kunnen beïnvloeden. Bij onderzoek naar de rol van terugwaartse Ramanverstrooiing is een nieuw parameter-regime gevonden waarin dit type verstrooiing de productie van hoog-energetisch elektronen vermindert. Dit in tegenstelling tot de gangbare opvatting dat terugwaartse Ramanverstrooiing juist leidt tot meer hoog-energetische elektronen. Verder is gevonden dat voorwaartse Ramanverstrooiing kan worden ingezet om terugwaartse Raman verstrooiing te onderdrukken en zo de elektronenopbrengst te vergroten.

Nadat het kluitje elektronen de versneller heeft verlaten moet het op de plaats van bestemming worden afgeleverd door middel van een zogenoemd bundelgeleidingssysteem. Hoewel dit stadium in principe losstaat van het versnelstadium verdient het toch de nodige aandacht. Tijdens het versnellen probeert men namelijk het kluitje niet alleen een hoge energie maar ook een goede kwaliteit (d.w.z. beperkte spreiding in alle richtingen) mee te geven. Een slordig geconstrueerd bundelgeleidingssysteem kan deze kwaliteit sterk verminderen en dus alle inspanningen teniet doen.

Op grond van deze overwegingen is besloten het laatste hoofdstuk van dit proefschrift te wijden aan een studie van de magnetische lenzen die worden gebruikt in een bundelgeleidingssysteem. Hierbij is vooral aandacht besteed aan een goede beschrijving van de magnetische velden van zulke lenzen. De gangbare aanpak, waarbij deze velden worden ontwikkeld in een machtreeks, heeft als nadeel dat het aantal termen in de veldbeschrijving explosief toeneemt en de resulterende uitdrukkingen nauwelijks meer hanteerbaar zijn. Daarom is gezocht naar een alternatieve beschrijving die geen gebruik maakt van machtreeksen. Dit heeft geleid tot een algemene analytische beschrijving voor het veld van een magnetische lens. Deze beschrijving stelt ons in staat om een scala aan veldgrootheden uit te drukken in uit metingen afkomstige randwaarden van het magnetisch veld, zonder eerst het magnetisch veld uit te hoeven rekenen. Ook het effect van een multipool op elektronenbanen kan op deze manier worden berekend. Dit alles is mogelijk zonder gebruik te maken van machtreeksen, waardoor de uitdrukkingen hanteerbaar blijven.

Curriculum Vitæ

Raoul Trines werd op 10 september 1974 geboren in Eindhoven. In 1992 behaalde hij zijn diploma VWO aan het Sint-Joriscollege aldaar. In hetzelfde jaar startte hij met de gecombineerde propædeuse wiskunde en natuurkunde aan de Technische Universiteit Eindhoven. Na het behalen van het propædeutisch diploma besloot hij van beide richtingen ook het doctoraal gedeelte te doen. Het afstudeeronderzoek werd gelijktijdig aan de faculteiten Wiskunde & Informatica en Technische Natuurkunde verricht, onder toezicht van dr. S.J.L. van Eijndhoven en dr. J.I.M. Botman. Het onderzoek had betrekking op de beschrijving van de velden van magnetische lenzen die worden gebruikt in de deeltjesoptica. Het doctoraal examen in zowel wiskunde als natuurkunde werd behaald in 1998, beide met lof toegekend. Sinds die tijd heeft hij zich bezig gehouden met zijn promotieonderzoek aan de Technische Universiteit Eindhoven, met prof.dr.ir. T.J. Schep en prof.dr. F.W. Sluijter als promotoren. Voor aanvullend onderzoek heeft hij in 2001 vier maanden doorgebracht in de groep van dr. W.P. Leemans aan het Lawrence Berkeley National Laboratory in Berkeley, Californië. Het promotieonderzoek heeft geresulteerd in dit proefschrift.

Raoul Trines was born on September 10, 1974, in Eindhoven. In 1992 he graduated from high school at the local Sint-Joriscollege. The same year he started his college education at the Technische Universiteit Eindhoven, with both mathematics and physics as majors. After having completed the first year, he decided to pursue the M.Sc. degree for both. His M.Sc. research project was conducted simultaneously at the Department of Mathematics & Computer Science and the Department of Applied Physics, under supervision of Dr. S.J.L. van Eijndhoven and Dr. J.I.M. Botman. The topic of his research was the description of the fields of magnetic lenses as used in particle optics. He received his M.Sc. in both mathematics and physics in 1998, both cum laude. Since then he has spent his time on his Ph.D. research at the Technische Universiteit Eindhoven, with Dr. T.J. Schep and Dr. F.W. Sluijter as his thesis advisors. In 2001, he spent four months with the group of Dr. W.P. Leemans at the Lawrence Berkeley National Laboratory, Berkeley, California. His Ph.D. research has resulted in this thesis.

Dankwoord

Het promotieonderzoek waarvan dit proefschrift het verslag is had ik natuurlijk nooit alleen kunnen uitvoeren. Ik ben dan ook behoorlijk wat mensen mijn dank verschuldigd. Op het gevaar af iemand over te slaan wil ik toch een aantal mensen met name noemen.

Allereerst wil ik mijn promotoren Theo Schep en Frans Sluiter bedanken voor het richting geven aan mijn onderzoek. Theo, jouw steun en kritisch oog waren onmisbaar bij het schrijven, zowel van dit proefschrift als van de diverse artikelen. Frans, het samen doorlezen van het proefschrift als geheel was een goed middel om al die kleine foutjes eruit te halen die anders gewoon waren blijven zitten. En nog bedankt dat je het mogelijk maakte dat ik de Lasers 2000 conferentie in Albuquerque kon bezoeken.

Mijn copromotor Leon Kamp wil ik graag bedanken voor de dagelijkse begeleiding en de steun tijdens moeilijke momenten. Mijn FOM-collega Vladimir Goloviznin bedank ik voor zijn deskundige begeleiding tijdens de eerste tweeëneenhalf jaar van mijn promotie.

Collega en vriend Albert Reitsma, bedankt voor de gezellige jaren. Conferenties waren gewoon leuker met jou erbij, en ook de periode dat we samen in Berkeley zaten zal ik niet vergeten. Mijn dank gaat ook uit naar je vrouw Anna: samen zorgden jullie ervoor dat ik me altijd welkom voelde bij jullie.

Marnix van de Wiel, bedankt voor de belangstelling voor mijn onderzoek en voor het doorlezen van mijn proefschrift. Ook bedankt voor het beschikbaar stellen van CPS-gelden voor mijn bezoeken aan Berkeley. Prof. K.-J. Boller wil ik bedanken voor het doorlezen van mijn proefschrift als lid van de kerncommissie.

Jan Botman en Stef van Eijndhoven, zonder jullie hulp was hoofdstuk 7 er nooit gekomen. Bedankt.

Wim Leemans, Eric Esarey, Carl Schroeder, and all the other people of l'OASIS group, I will never forget the good times I had in your group at Berkeley Lab. I also wish to thank you for a fruitful collaboration, the results of which now form an important part of my thesis. And if I ever take part in the LBNL Runaround again, I will definitely prepare myself better!

De medewerkers van het werkverband Theoretische Natuurkunde wil ik bedanken

voor de goede sfeer. Helmi, Inge, Wim, Jan, Thijs, Peter, Paul, Alexej, Boudewijn, Henk, Karsten en vooral Ria: het was gezellig en ik zal jullie missen. Ook de AIO's en studenten mag ik zeker niet vergeten. Altijd waren ze bereid om mijn verhalen aan te horen of een computerprobleem te regelen als ik afleiding nodig had. Stop jullie bugs nou maar in het zuur, want ik ben weg.

De onderzoeksschool CPS ben ik dank verschuldigd voor de financiële ondersteuning van mijn promotieonderzoek.

Muziek maken en alles daar omheen is altijd een goede manier gebleken om me met iets anders bezig te houden dan alleen maar werken. De leden en oud-leden van het Eindhovens Jongeren Ensemble en Le Silence Curieux wil ik hartelijk bedanken voor alle mooie muzikale en niet-muzikale momenten van de afgelopen jaren. Ronald en Han-Wen, succes met jullie eigen promoties.

Mijn broer en zussen wil ik bedanken voor de belangstelling die ze altijd hebben getoond voor mijn onderzoek. Nee Serge, ik heb nog steeds geen nieuw deeltje ontdekt, maar je weet nooit.

Tenslotte wil ik mijn ouders bedanken, omdat ze er altijd voor me waren en zijn. Niets was ze ooit teveel als ik het moeilijk had. Pap en Mam, bedankt voor alles.