ORGANISATION EUROPÉENNE POUR LA RECHERCHE NUCLÉAIRE

CERN EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

The CERN Accelerator School

Plasma Wake Acceleration

Geneva, Switzerland
23–29 November 2014

Proceedings

Editor: B. Holzer

GENEVA
2016

This report should be cited as:

Proceedings of the CAS–CERN Accelerator School: Plasma Wake Acceleration, Geneva, Switzerland, 23–29 November 2014, edited by B. Holzer, CERN-2016-001 (CERN, Geneva, 2016), http://dx.doi.org/10.5170/CERN-2016-001

A contribution in this report should be cited as:

[Author name(s)], in Proceedings of the CAS–CERN Accelerator School: Plasma Wake Acceleration, Geneva, Switzerland, 23–29 November 2014, edited by B. Holzer, CERN-2016-001 (CERN, Geneva, 2016), pp. [first page] – [lastpage], http://dx.doi.org/10.5170/CERN-2016-001. [first page]

Abstract

These proceedings are of interest to staff and students in accelerator laboratories, university departments and companies working in or having an interest in the field of new acceleration techniques. Following introductory lectures on plasma and laser physics, the course covers the different components of a plasma wake accelerator and plasma beam systems. Topical seminars and an overview about alternative new techniques like dielectric accelerators are included. Lectures on the experimental studies and their latest results, on diagnostic tools and state of the art wake acceleration facilities, both present and planned, complement the theoretical part.

iv

Preface

The CERN Accelerator School – CAS – has the mission to preserve and disseminate the knowledge accumulated at CERN and elsewhere on particle accelerators and storage rings. For topics on general accelerator physics this is achieved by a series of lectures presented in the introductory and advanced level schools, which run in alternate years. In addition certain topics, relevant for the field, are treated in special schools to follow the technological developments and present details of advanced subjects.

Over the past decades, and in the true sense of the word since the invention of the very first particle accelerator by Cockcroft and Walton, progress in high–energy physics has always been the driving force for the development of particle accelerators. New scientific findings went hand in hand with the continuous effort to increase the performance of the particle accelerators. We have learned to apply better and better these techniques and include them in our accelerators and storage rings, and the harvest is impressive: the standard model that proves its solidity more than ever, including the recent discovery of the Higgs particle at LHC in 2012.

Meanwhile many interesting questions of high–energy physics lie beyond the standard model. Keywords are dark matter and dark energy and a big step forward in accelerator technology is needed to address these issues. For this purpose at CERN and other places we are studying already now the next generation colliders and it is no surprise that new methods to achieve an efficient and more compact way to gain the required acceleration gradient are being studied in many places: plasma wake based concepts show gradients of GeV/m, an impressive step forward by several orders of magnitude compared to our standard acceleration techniques.

Without doubt these new techniques deserve support and in several laboratories worldwide the plasma wake field as a new acceleration concept is being studied and developed. With this in mind it was only natural to organise a lecture series, dedicated to the special topic of plasma wake acceleration, in order to bring the world wide expertise in this rapidly developing field together and to train young people. In the present proceedings a number of lectures on laser and plasma physics are introducing the field as well as an overview about conventional accelerators and their limitations. They are followed by more advanced courses which cover a large number of aspects in plasma wake acceleration schemes: the creation of the plasma by high–power lasers or particle beams, the description of the plasma creation process in simulations and the characteristics of the accelerated particle beams including results from latest achievements. Lectures about beam diagnostics, applications of plasma accelerated beams and topical seminars complete the program. We are convinced that the present status of the topic is presented in an adequate manner and we hope that these proceedings will help newcomers as well as experienced colleagues to find some useful and interesting information.

We would like to take the opportunity to thank all lecturers for the enormous task of preparing, presenting, and writing up their topics, for the effort that they did, the time that they spent and the motivation to share their knowledge with the younger colleagues.

Very special thanks must go to the enthusiasm and positive feedback of the participants, the students that attended the school in an unexpectedly large number, for their motivation to get involved in this new topic.

As always, the backing of the CERN management, the guidance of the CAS Advisory and Programme Committees, and the attention to detail of the CAS secretary, Barbara Strasser ensured that the school was held under optimum conditions.

Finally, we thank the CERN E-Publishing Service for their dedication and commitment to the production of this document.

<div align="center">
Bernhard Holzer,
Organiser and Chair of the CERN Accelerator School on Plasma Wake Acceleration
</div>

DRAFT PROGRAMME FOR PLASMA WAKE ACCELERATION
23-29 November, 2014, CERN, Geneva, Switzerland

Time	Sunday 23 November	Monday 24 November	Tuesday 25 November	Wednesday 26 November	Thursday 27 November	Friday 28 November	Saturday 29 November
08:30		Welcome & Introduction to the School	Plasma Wake Generation (Linear)	Plasma Injection Schemes	Beam Driven Propagation (Beams)	Particle Beam Diagnostics II	**CERN VISIT**
09:30		R. Bailey	B. Bingham	J. Faure **10:00** Coffee	R. Assmann	C. Welsch	
09:30		Introduction & Historical Overview	Acceleration of e- in a Plasma	**10:30**	Beam Driven Propagation (Lasers)	Radiation Generation	**CERN VISIT**
10:30		V. Malka	B. Bingham	Injection Extraction & Matching	B. Cros	K. Ta Phuoc	
		COFFEE	COFFEE		COFFEE	COFFEE	
11:00	A R R I V A L	Accelerator Physics & Limitations I	Plasma Wake Generation (Non-Linear)		Plasma Sources	Beam Driven Plasma Acceleration (Experiment)	**CERN VISIT**
12:00		B. Holzer	L. Silva	M. Ferrario	S. Hooker	E.Gschwendtner	
12:00		Introduction to Plasma Physics I	Blow Out Regime	Modelling and Simulations I	Plasma Diagnostics	Laser Driven (Experiment)	**CERN VISIT**
13:00		P. Gibbon	L. Silva	A. Pukhov	M. Kaluza	S. Mangles	
		LUNCH	LUNCH	LUNCH	LUNCH	LUNCH	
14:30		Accelerator Physics & Limitations II	Laser Driven Systems	Modelling and Simulations II	Particle Beam Diagnostics I	Case Study	D E P A R T U R E
15:30		B. Holzer	Z. Najmudin	A. Pukhov	C. Welsch		
		TEA	TEA	TEA	TEA	TEA	
16:00		Introduction to Plasma Physics II	Beam Driven Systems	**Seminar** Dielectric Laser Acceleration	**Seminar** Acceleration of Protons and Ions	Applications	
17:00		P. Gibbon	P. Muggli	P.Hommelhoff	M. Roth	S. Karsch	
17:00	*Registration*	Introduction to Laser Physics & High Power Lasers	Case Study	Case Study	**Seminar** Standard Model and Beyond	Case Study Summary	
18:00		C. Wahlstroem			P. Sphicas	J. Osterhoff	
19:00	Dinner	Welcome Drink Dinner	Dinner	Special Dinner	Dinner	Dinner	

Contents

Plasma Wake Accelerators: Introduction and Historical Overview

V. Malka

Laboratoire d'Optique Appliquée, ENSTA-ParisTech, CNRS, Ecole Polytechnique, Université Paris-Saclay, Palaiseau, France

Abstract

Fundamental questions on the nature of matter and energy have found answers thanks to the use of particle accelerators. Societal applications, such as cancer treatment or cancer imaging, illustrate the impact of accelerators in our current life. Today, accelerators use metallic cavities that sustain electric fields with values limited to about 100 MV/m. Because of their ability to support extreme accelerating gradients, the plasma medium has recently been proposed for future cavity-like accelerating structures. This contribution highlights the tremendous evolution of plasma accelerators driven by either laser or particle beams that allow the production of high quality particle beams with a degree of tunability and a set of parameters that make them very pertinent for many applications.

Keywords

Accelerator; laser; plasma; laser plasma accelerator; laser wakefield.

1 Introduction

This article corresponds to the introductory lecture given at the first CAS-CERN Accelerator School on Plasma Wake Acceleration on 21–28 November, 2014. Having this school dedicated to Plasma Accelerators at CERN, where an important part of the story of high energy physics has been written and where the worlds larger accelerators and the brighter and more energetic particle beams are produced, represents in itself the realization of a dream that shows the maturity and the vitality of the field. Having an unexpected level of participation shows also the dynamism of this field of research with an impressive growth of groups in Europe and all over the world. Accelerator physics started almost 130 years ago with the discovery of the cathodic tube. Since then, accelerators have gained in efficiency and in performance delivering energetic particle beams with record energy and luminosity values. During the last century, they have been developed for fundamental research, for example, for producing intense picosecond X-ray pulses in synchrotron machines, or more recently even shorter, few femtosecond X-ray pulses in free electron laser machines. Such short X-ray pulses are crucial for the study of ultra-fast phenomena, for example in biology, to follow the DNA structure evolution, or in material science to follow the evolution of crystals. Higher energy accelerators are crucial to answering important questions regarding the origins of the universe, of dark energy, of the number of space dimensions, etc. The largest one available, the Large Hadron Collider, has for example confirmed two years ago the existence of the Higgs boson. Figure 1 illustrates few of the many fundamental discoveries that have been made this last century and have permitted matter from 10^{-10} to 10^{-20} m spatial resolution to be probed.

Moreover, as illustrated in Fig.2, with an annual market of more than a few US billions of dollars, accelerators are used today in many fields such as cancer therapy, ion implantation, electron cutting and melting, and non-destructive inspection, etc.

The accelerating field in superconducting radio-frequency cavities is, due to electrical breakdown of the metallic cavity, limited to about 100 MV/m. It is for this reason that an increase of the particle energy requires an increase of the acceleration length. In the 1950s, Budker and Veksler [1] proposed using plasma collective fields to accelerate charged particles more compactly. In the pioneering theoretical work performed in 1979, Tajima and Dawson [2] showed how an intense laser pulse can excite
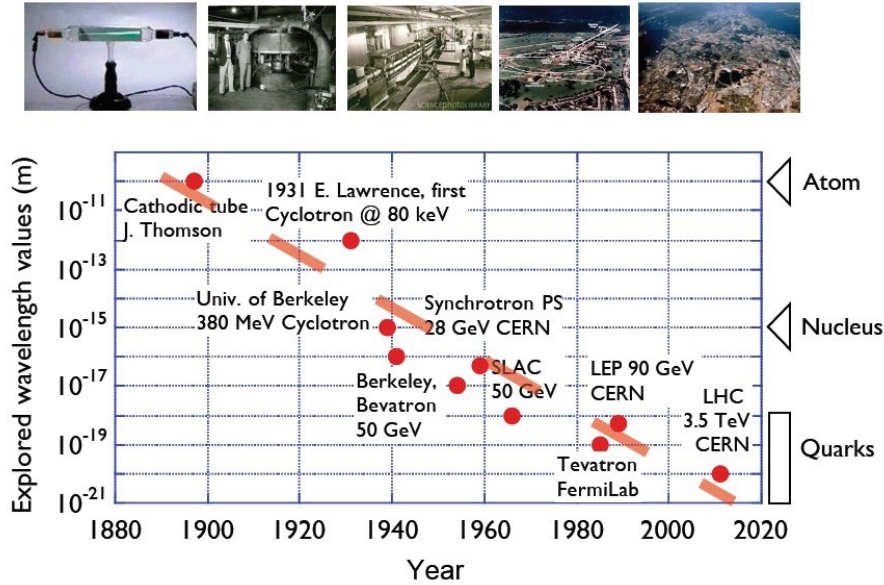
Fig. 1: Evolution of accelerators and their main related discoveries

a wake of plasma oscillations through the non-linear ponderomotive force associated to the laser pulse. In their proposed scheme, relativistic electrons were injected externally and were accelerated in the very high GV/m electric field sustained by relativistic plasma waves. In this early article [2], the authors proposed two schemes: the laser beat wave and the laser wakefield. Several experiments were performed at the beginning of the 1990s following on from their ideas, and injected few MV/m electrons have gained energy in GV/m accelerating gradients using either the beat wave or the laser wakefield scheme. In 1994, at the Rutherford Appleton Laboratory, using the 40 TW powerful Vulcan Laser, hundreds of GV/m accelerating gradients have been generated and used to trap electrons from the plasma itself, and to accelerate them [3] to few tens of MV/m over only 1 mm distance. TV/m accelerating gradients have since been demonstrated in the non-linear regime in the forced laser wakefield scheme. Figure 3 illustrates the compactness of a plasma accelerating cavity.

Application	Total syst. (2007) approx.	System sold/yr	Sales/yr (M$)	System price (M$)
Cancer Therapy	9100	500	1800	2.0 − 5.0
Ion Implantation	9500	500	1400	1.5 − 2.5
Electron cutting and welding	4500	100	150	0.5 − 2.5
Electron beam and X-rays irradiators	2000	75	130	0.2 − 8.0
Radio-isotope production (incl. PET)	550	50	70	1.0 − 30
Non destructive testing (incl. Security)	650	100	70	0.3 − 2.0
Ion beam analysis (incl. AMS)	200	25	30	0.4 − 1.5
Neutron generators (incl. sealed tubes)	1000	50	30	0.1 − 3.0
Total	27500	1400	3680	

Fig. 2: Market of industrial accelerators and their main societal applications

In 1985, Chen and Dawson [4] proposed to use a bunched electron beam to drive plasma wakes with, again, GV/m accelerating gradients. Soon after, the first experiments on Particle WakeField Acceleration (PWFA) were achieved using low energy electron beam drivers. In 1996, T. Katsouleas and

C. Joshi proposed to use an ultra-relativistic electron beam delivered by the SLAC linac to drive GV/m accelerating fields. In 2009, the possibility of driving plasma-wakefield acceleration with a proton bunch was proposed [5], and the authors demonstrated through numerical simulations that TeV energy levels could be reached in a single accelerating stage driven by a TeV proton bunch.
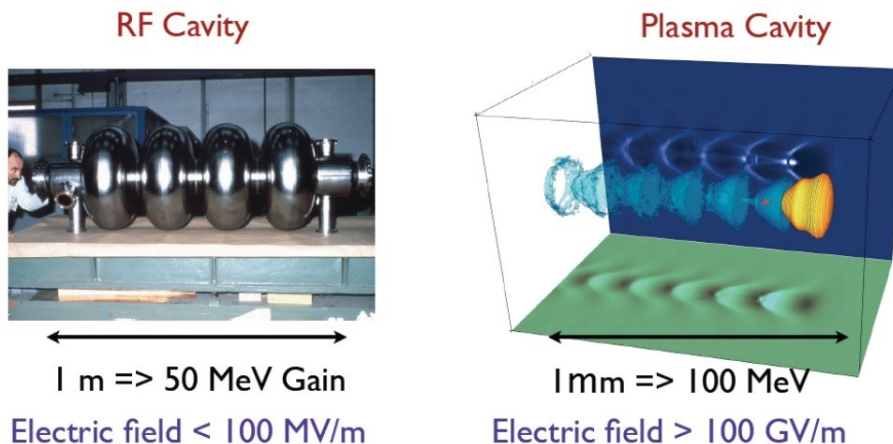
Fig. 3: Compactness of plasma 'cavity'. Left: Radiofrequency cavity. Right: Non-linear laser plasma wakefield. The laser pulse in yellow propagates from left to right, the iso-electronic density is shown in blue and the electron bunch in red.

In both cases, the accelerating gradient results from the rapid electron plasma oscillation that follows the electronic perturbation. If, in an initially uniform and non-collisional plasma a slab of electrons are displaced from their equilibrium position, the restoring force which is applied to this electron slab drives them towards the equilibrium position. For the time scale corresponding to the electrons motion, the motion of the ions can be neglected because of inertia. The typical frequency of electron oscillations around the equilibrium position is called the electron plasma frequency ω_{pe}:

$$\omega_{pe} = \sqrt{\frac{n_e e^2}{m_e \varepsilon_0}}, \tag{1}$$

where n_e is the unperturbed electron density.

If $\omega_{pe} < \omega_0$ (where ω_0 is the laser frequency) then the characteristic time scale of the plasma is longer than the optical period of the incoming radiation. The medium cannot stop the propagation of the electromagnetic wave. The medium is then transparent and it is called 'under-dense'. When $\omega_{pe} > \omega_0$ then the characteristic time scale of the electrons is fast enough to adapt to the incoming wave and to reflect totally or partially the radiation, and the medium is called 'over-dense'.

These two domains are separated at frequency ω_0, which corresponds to the critical density, $n_c = \omega_0^2 m_e \varepsilon_0 / e^2$. For a wavelength $\lambda_0 = 1\,\mu m$, one obtains $n_c = 1.1 \times 10^{21}\,cm^{-3}$. The typical range of electron densities of laser plasma accelerators with current laser technology, is $[10^{17}\,cm^{-3} - 10^{20}\,cm^{-3}]$.

In a uniform ion layer, the density change δn for a periodic sinusoidal perturbation of the electron plasma density is written

$$\delta n = \delta n_e \sin(k_p z - \omega_p t), \tag{2}$$

where ω_p and k_p are the angular frequency and the wave number of the plasma wave.

This density change leads to a perturbation of the electric field $\delta \vec{E}$ via the Poisson equation:

$$\vec{\nabla} \delta \vec{E} = -\frac{\delta n\, e}{\varepsilon_0}. \tag{3}$$

This gives

$$\delta\vec{E}(z,t) = \frac{\delta n_{\mathrm{e}}\, e}{k_{\mathrm{p}}\varepsilon_0}\cos(k_{\mathrm{p}}z - \omega_{\mathrm{p}}t)\vec{e_z}. \tag{4}$$

The electric field associated to the relativistic plasma wave, i.e. with a phase velocity close to the speed of light $v_{\mathrm{p}} = \omega_{\mathrm{p}}/k_{\mathrm{p}} \sim c$ can be described by

$$\delta\vec{E}(z,t) = E_0\frac{\delta n_{\mathrm{e}}}{n_{\mathrm{e}}}\cos(k_{\mathrm{p}}z - \omega_{\mathrm{p}}t)\vec{e_z}, \tag{5}$$

where $E_0 = m_{\mathrm{e}}c\omega_{\mathrm{pe}}/e$.

In the linear case, as shown in Fig. 4, the relative density perturbation is much smaller than one, and the density perturbation with the electric field has a sinusoidal profile. Note that the electric field is dephased by $-\pi/4$ with respect to the electron density. A 1% density perturbation at a plasma density of $10^{19}\,\mathrm{cm}^{-3}$ corresponds to 3 GV/m. In the non-linear case, for a 100% density perturbation at a plasma density of $10^{19}\,\mathrm{cm}^{-3}$ the accelerating field reaches 300 GV/m.
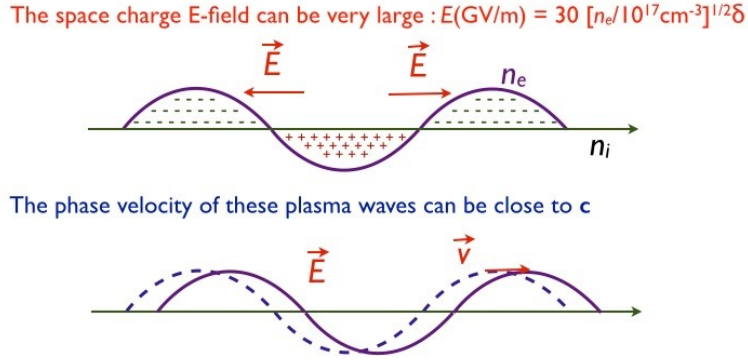
Fig. 4: Density perturbation with the corresponding electric field

We now examine the electron motions in this oscillating electric field in the simplified case of a one-dimensional plasma wave. Figure 5 represents an example of an electron trajectory in a plasma wave. In this phase space, the closed orbits correspond to trapped particles. Open orbits represent untrapped electrons, either because the initial velocity is too low or to high. The curve which separates these two regions is called the separatrix. This separatrix gives the minimum and maximum energies for trapped particles. This is comparable to the hydrodynamic case, where a surfer has to crawl to gain velocity and to catch the wave.

For an electron density much lower than the critical density $n_{\mathrm{e}} \ll n_{\mathrm{c}}$, we find $\gamma_{\mathrm{p}} = \omega_0/\omega_{\mathrm{p}} \gg 1$ and

$$\Delta W_{\max} = 4\gamma_{\mathrm{p}}^2\frac{\delta n_{\mathrm{e}}}{n_{\mathrm{e}}}mc^2. \tag{6}$$

For an electron travelling along the separatrix, the time necessary to reach maximal energy is infinite because there exists a stationary point at energy γ_{p}. for other closed orbits, the electron successively gains and loses energy during its rotation in the phase space. In order to design an experiment, one needs an estimation of the distance an electron travels before reaching maximal energy gain. This length, which is called the dephasing length L_{deph}, corresponds to a $\lambda_{\mathrm{p}}/2$ rotation in the phase space. In order to have a simple analytical estimation, one can assume that the energy gain is small compared to the initial energy of the particle and that the plasma wave is relativistic $\gamma_{\mathrm{p}} \gg 1$, then the dephasing length is written

$$L_{\mathrm{deph}} \sim \gamma_{\mathrm{p}}^2\lambda_{\mathrm{p}}. \tag{7}$$
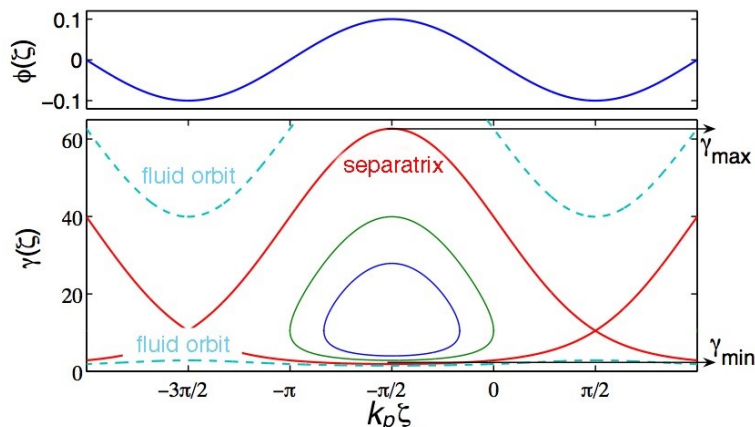
Fig. 5: Upper panel: Potential in phase space. Lower panel: Trajectory of an electron injected ion the potential of the plasma wave in the frame of the wave with the fluid orbit (dashed line), the trapped orbit and in between in red the separatrix.

In these formulas, we have considered a unique test electron, which has no influence on the plasma wave. In reality, a massive trapping of particles modifies the electric fields and distorts the plasma wave. This is called the space-charge or beam loading effect (which results from the Coulomb repulsion force). Finally, this linear theory is difficult to apply to non-linear regimes which are explored experimentally.

2 Laser wakefield accceleration

2.1 Laser wakefield: the linear regime

The ponderomotive force of the laser excites a longitudinal electron plasma wave with a phase equal to the group velocity of the laser close to the speed of light. Two regimes have been proposed to excite a relativistic electron plasma wave.

In the standard laser wakefield acceleration (LWFA) approach, a single short laser pulse excites the relativistic electron plasma wave. As the ponderomotive force associated with the longitudinal gradient of the laser intensity exerts two successive pushes in opposite directions on the plasma electrons, the excitation of the electron plasma wave is maximum when the laser pulse duration is of the order of $1/\omega_{\mathrm{p}}$. For a linearly polarized laser pulse with full width at half maximum (FWHM) $\sqrt{2\ln 2}\, L$ (in intensity), the normalized vector potential, also called the force parameter of the laser beam, is written

$$a(z,t) = a_0 \exp\left[-\left(\frac{k_0 z - \omega_0 t}{\sqrt{2}k_{\mathrm{p}}L}\right)^2\right]. \tag{8}$$

In the linear regime, $a_0 \ll 1$, the electronic response obtained behind a Gaussian laser pulse can be easily calculated [6]. In this case, the longitudinal electric field is given by

$$\vec{E}(z,t) = E_0 \frac{\sqrt{\pi}a_0^2}{4}k_{\mathrm{p}}L \exp(-k_{\mathrm{p}}^2 L^2/4)\cos(k_0 z - \omega_0 t)\vec{e_z}. \tag{9}$$

Equation (9) explicitly shows the dependence of the amplitude of the wave on the length of the exciting pulse. In particular, the maximal value for the amplitude is obtained for a length $L = \sqrt{2}/k_{\mathrm{p}}$ as shown in Fig. 6 for a laser with a normalized vector potential $a_0 = 0.3$. One can note that in the linear regime, the electric field has a sinusoidal shape and reaches maximal values of a few GV/m. For example, for an electron density $n_{\mathrm{e}} = 10^{19}$ cm^{-3}, the optimal pulse duration equals $L = 2.4\ \mu$m

5
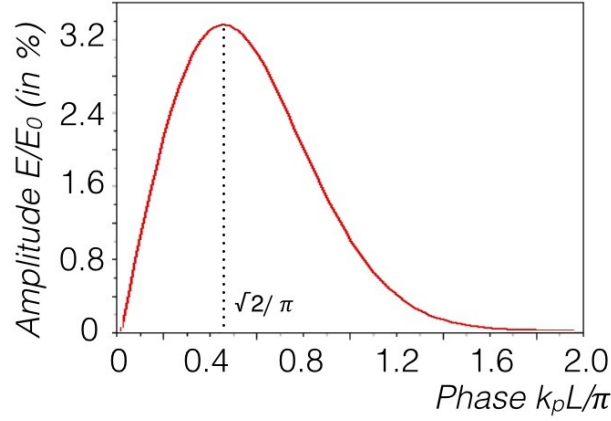
Fig. 6: Amplitude of the electric field as function of the length of a Gaussian laser pulse for a normalized vector potential $a_0 = 0.3$.

(equivalent to a pulse duration $\tau = 8$ fs). For $a_0 = 0.3$, the maximal electric field is in the GV/m range. Figure 7 illustrates the density perturbation and the corresponding longitudinal electric field produced at resonance by a low intensity, $I_{\text{laser}} = 3 \times 10^{17}$ W/cm^2, laser pulse of 30 s duration.
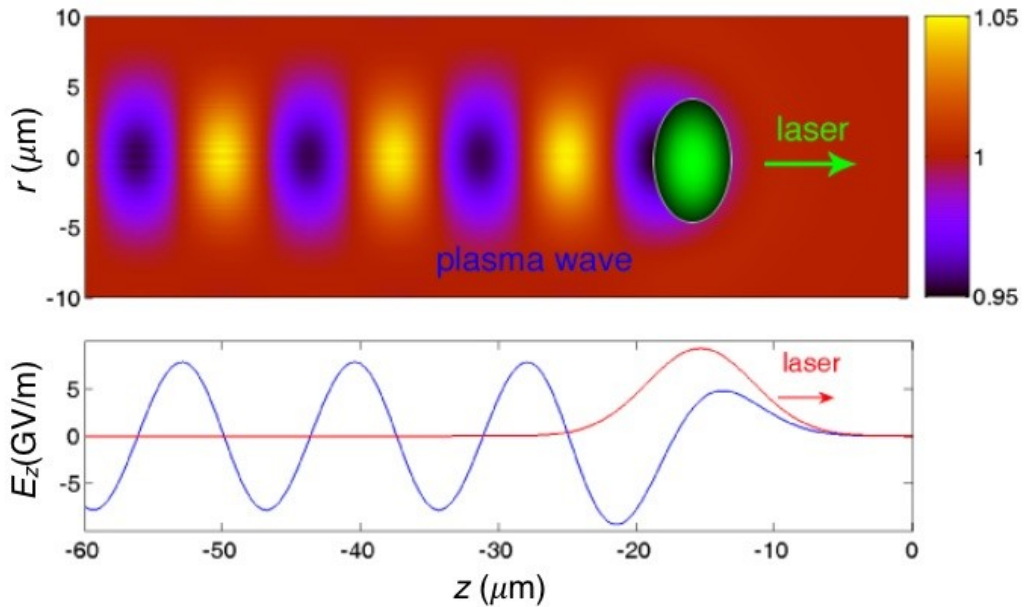
Fig. 7: Density perturbation (top) and electric field (bottom) produced in the linear regime

In experiments carried out at LULI, relativistic plasma waves with 1% amplitude have been demonstrated. As indicated in Fig. 8, 3 MeV electrons have been injected into a relativistic plasma wave driven by a 300 fs laser pulse, some of which were accelerated up to 4.6 MeV [7]. The electron spectra has a broad energy distribution with a Maxwellian like shape, as expected when injecting an electron beam with a duration much longer that the plasma period, and in this case with a duration much longer than the plasma wave live-time.
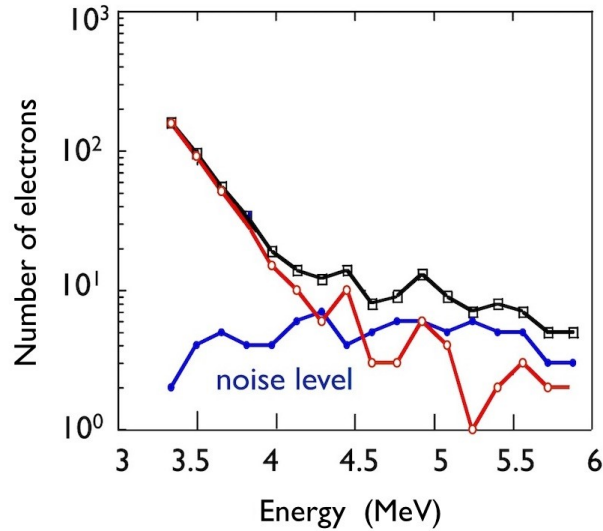
Fig. 8: Electrons spectra obtained at LULI in the laser wakefield scheme

2.2 Laser beatwave

Before the advent of short and intense laser pulses, relativistic plasma waves were driven by the beat-wave of two long laser pulses of a few tens of picoseconds (i.e. with duration much greater than the plasma period). In this case, the plasma frequency ω_p has to satisfy exactly the matching condition, $\omega_p = \omega_1 - \omega_2$, with ω_1 and ω_2 the frequencies of the two laser pulses. The first observation of relativistic plasma waves was performed using the Thomson scattering technique by the group of C. Joshi at UCLA [8]. Acceleration of 2 MeV injected electrons up to 9 MeV [9] and later on, up to 30 MeV [10], were demonstrated by the same group using a CO_2 laser of about 10 μm wavelength. At LULI, 3 MeV electrons were accelerated up to 3.7 MeV in beat wave experiments with Nd:Glass lasers of about 1 μm wavelengths by a longitudinal electric field of 0.6 GV/m [11]. Similar works were also performed in Japan at University of Osaka [12], in the UK at the Rutherford Appleton Laboratory [13], and in Canada at the Chalk River Laboratory [14]. Electron spectra obtained at LULI in the laser beat wave scheme are shown on Fig. 9.

In order to reduce the coupling between electron waves and ion waves which was a limiting factor of previous experiments performed with 100 ps Nd lasers [15], experiments done at the Rutherford Appleton Laboratory with a 3 ps laser pulse have shown excitations of higher amplitude relativistic plasma waves [16].

2.3 Self-modulated laser wakefield

In all of these experiments, because of the duration of the injected electron bunch, which is much longer than the plasma period and even longer than the life time of the plasma, only a very small fraction of injected electrons were accelerated and the output beam had a very poor quality with a Maxwellian-like energy distribution.

Thanks to the development of powerful laser systems with short pulse duration (500 fs), a new regime that allows self-injection of electrons in very intense accelerating gradients with values exceeding 100 GV/m has been discovered. The cumulative effects of the self-focusing and the self-modulation of the laser envelope by the initial perturbation of the electron plasma density generates a train of laser pulses which become resonant with the plasma wave. These effects are described in Fig. 10. The self-modulated laser wakefield regime occurs when the laser pulse duration exceeds the plasma period and when the laser power exceeds the critical power for self-focusing [17–19]. The initial Gaussian laser
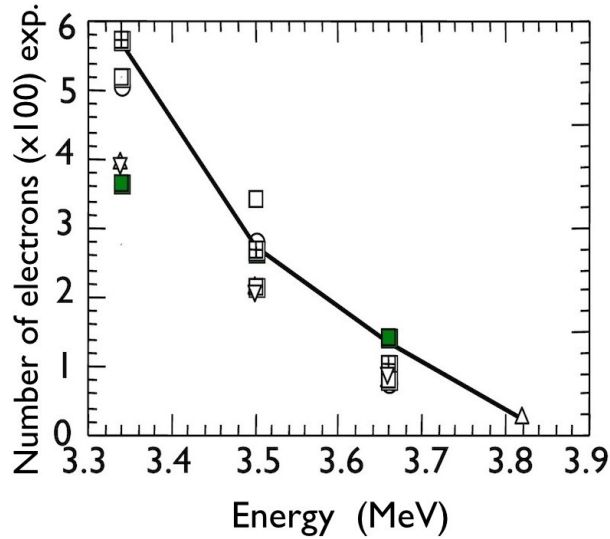
Fig. 9: Electrons spectra obtained at LULI in the laser beat wave scheme

pulse becomes modulated at the plasma wavelength during its propagation. This mechanism, which is close to a Forward Raman Scattering Instability [20], can be described as the decomposition of an electromagnetic wave into a plasma wave at a frequency shifted by the plasma frequency.

In an experiment done at the Rutherford Appleton Laboratory, a relativistic plasma wave was excited by an intense laser ($>5 \times 10^{18}$ W/cm^2), for a short duration (<1 ps), by a 1.054 μm wavelength laser pulse in the self-modulated laser wakefield regime. This is the decay (induced by a noise level plasma wave) of the strong electromagnetic pump wave (ω_0, k_0) into the plasma wave (ω_p, k_p) and two forward propagating electromagnetic cascades at the Stokes ($\omega_0 - n\omega_p$) and anti-Stokes ($\omega_0 + n\omega_p$) frequencies, n being a positive integer, and ω and k being the angular frequency and the wavenumber, respectively, of the indicated waves. The spatial and temporal interference of these sidebands with the laser produces an electromagnetic beat pattern propagating synchronously with the plasma wave. The electromagnetic beat exerts a force on the plasma electrons, reinforcing the original noise level plasma wave which scatters more sidebands, thus closing the feedback loop for the instability.

The solid curve in Fig. 11 shows the electromagnetic frequency spectrum emerging form the plasma with a density of $>5 \times 10^{18}$ cm^{-3}, where the abscissa is the shift in frequency of the forward scattered light from the laser frequency in units of ω_p. The upshifted anti-Stokes and downshifted Stokes signals at $\Delta\omega/\omega_p = \pm 1$ are clearly visible as is the transmitted pump at $\Delta\omega/\omega_p = 0$ and the second and third anti-Stokes sidebands. These signals are sharply peaked, and their widths indicate that the plasma wave which generated these signals must have a coherence time of the order of the laser pulse. The dashed curve shows the spectrum when the density is increased to 1.5×10^{19} cm^{-3}. The most startling feature is the tremendous broadening of the individual anti-Stokes peaks at this higher density. This broadening corresponds to wave-breaking and is mainly caused by the loss of coherence due to severe amplitude and phase modulation as the wave breaks. As wave-breaking evolves, the laser light no longer scatters off a collective mode of the plasma but instead scatters off the trapped electrons which are still periodically deployed in space but have a range of momenta producing, therefore, a range of scatter frequencies.

During experiments carried out in the UK in 1994 [3], the amplitude of the plasma waves reached the wave-breaking limit, where electrons initially belonging to the plasma wave are self-trapped and accelerated to high energies. The fact that the external injection of electrons in the wave is no longer necessary is a major improvement. Electron spectrums extending up to 44 MeV have been measured
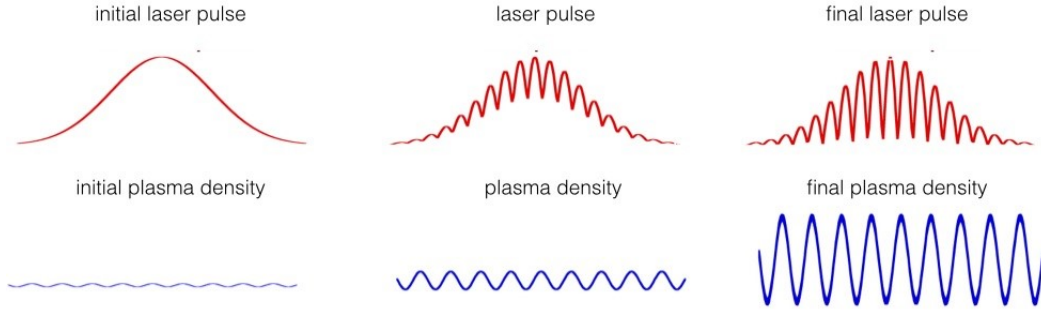
Fig. 10: Evolution of the laser pulse and plasma density in the self-modulated laser wakefield regime

during this first campaign, and up to 104 MeV in the second campaign. This regime has also been reached for instance in the United States at CUOS [21], and at NRL [22]. However, because of the heating of the plasma by these relatively 'long' pulses, the wave-breaking occurred well before reaching the cold wave-breaking limit, which limited the maximum electric field to a few 100 GV/m. The maximum amplitude of the plasma wave has also been measured to be in the range 20–60% [23].
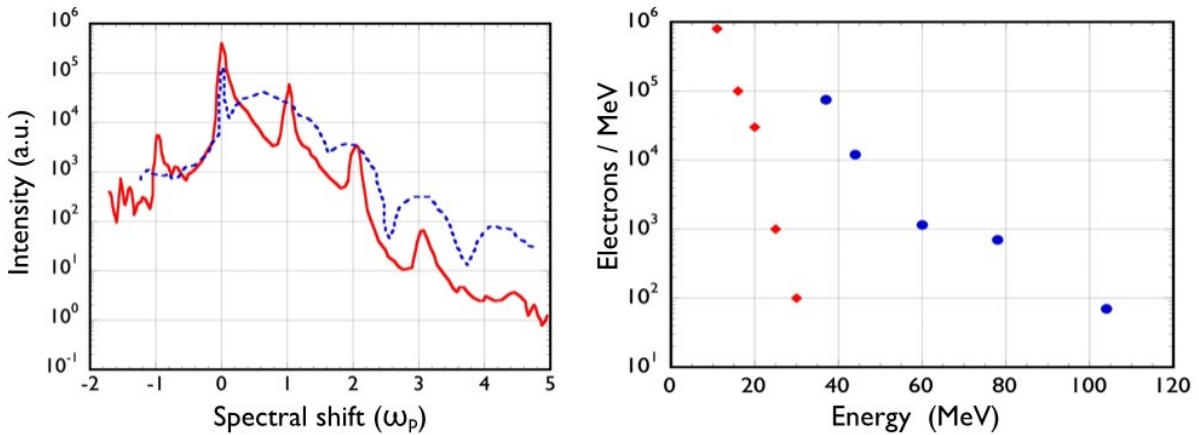
Fig. 11: Frequency and electron spectrum the self-modulated laser wakefield regime for two different electron plasma densities: 0.54×10^{18} cm^{-3} (in red) and 1.5×10^{19} cm^{-3} (in blue).

Experiments performed at LOA since 1999 have shown that an electron beam can also be produced using a compact 10 Hz laser system [24]. Fig. 12 shows two typical electron spectra obtained at 5×10^{19} cm^{-3} and 1.5×10^{20} cm^{-3}. The 0.6 J, 35 fs laser beam was focused tightly to a 6 μm focal spot leading to a peak laser intensity of 2×10^{19} W/cm^2. Electron distributions with electron energy greater than 4 MeV are well fitted by an exponential function, characteristic of an effective temperature for the electron beam. These effective temperatures are 8.1 MeV (2.6 MeV) for electron density of 5×10^{19} cm^{-3} (1.5×10^{20} cm^{-3}), to which correspond typical values of 54 MeV (20 MeV) for the maximum electron energy. This maximum energy is defined by the intersection between the exponential fit and the detection threshold. One can observe an important decrease in the effective temperature and in the maximum electron energy for increasing electron densities.

This point is summarized in Fig. 12 where we present the maximum electron energy as a function of the electron density. It decreases from 70 MeV to 15 MeV when the electron density increases from 1.5×10^{19} cm^{-3} to 5×10^{20} cm^{-3}. Also presented in Fig. 12 is the theoretical value [25]

$$W_{\max} \approx 4\gamma_{\mathrm{p}}^2(E_z/E_0)mc^2 F_{\mathrm{NL}}. \tag{10}$$
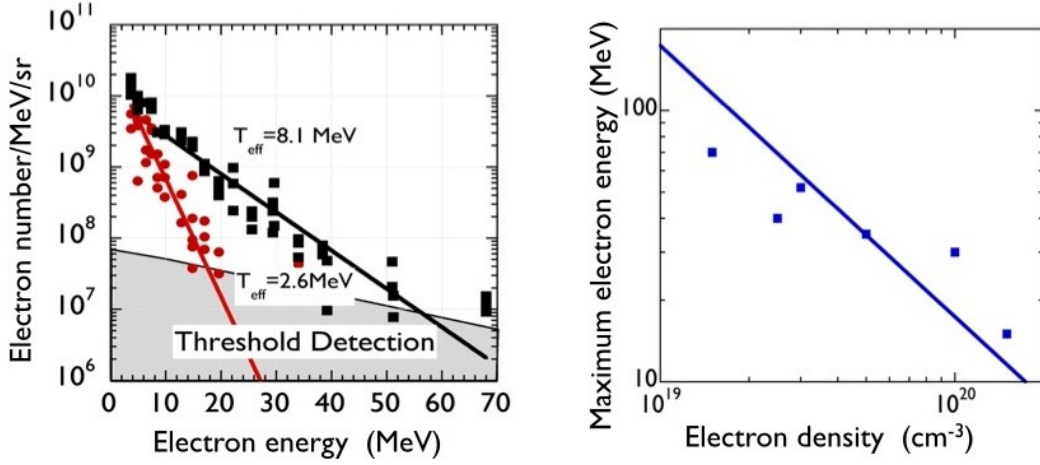
Fig. 12: Left: Typical electron spectra obtained at 5×10^{19} cm^{-3} (squares) and 1.5×10^{20} cm^{-3} (circles). The corresponding effective temperatures are 8.1 MeV (2.6 MeV) for electron density of 5×10^{19} cm^{-3} (5×10^{20} cm^{-3}). Right: Maximum electron energy as a function of the plasma electron density. Experimental data: squares. Theoretical data: line.

Here, the maximum electron energy is greater than the conventional one given by the simple formula $W_{\max} \approx 2\gamma_{\mathrm{p}}^2(E_z/E_0)mc^2$, where γ_{p} is the plasma wave Lorentz factor (which is equal to the critical density to electron density ratio $n_{\mathrm{c}}/n_{\mathrm{e}}$) and E_z/E_0 is the electrostatic field normalized to E_0 ($E_0 = cm\omega_{\mathrm{p}}/e$). The factor of two is due to self-channelling induced by the space-charge field which focuses accelerated electrons for all phases. The correction factor $F_{NL} \approx (\gamma_{\perp 0}n_0/n)^{3/2}$ corresponds to a non-linear correction due to the relativistic pump effect and to self-channelling. In this formula, n_0 is the initial electron density, n the effective one and $\gamma_{\perp 0}$ is the Lorentz factor associated to the laser intensity: $\gamma_{\perp 0} = (1 + a_0^2/2)^{1/2}$. The electron density depression is estimated by balancing the space-charge force and laser ponderomotive force, and evaluated by $\delta n/n = (a_0^2/2\pi^2)(1 + a_0^2/2)^{-1/2}(\lambda_p/w_0)^2$.

In the lower electron density case, the depression correction will introduce an important increase of the maximum energy gain which is multiplied by a factor of 2 at 1.5×10^{19} cm^{-3}. For densities greater than 1.0×10^{20} cm^{-3}, the main contribution is due to the relativistic pump effect, as outlined on the plot in Fig. 12. It is also crucial to note that the fact that the electron maximum energy increases when the electron density decreases demonstrates that electrons are mainly accelerated by relativistic plasma waves. The maximum electron energy calculated at lower density overestimated the experimental ones, indicating that the dephasing length becomes shorter than the Rayleigh length. In order to solve this problem, experiments were performed at LOA using a longer off-axis parabola, more energetic electrons have been measured, with a peak laser intensity ten times smaller than in this first experiment.

Electron beams with Maxwellian spectral distributions, generated by compact high repetition rate ultra-short laser pulses, have been also at this time been produced in many laboratories around the world: at LBNL [26], at NERL [27] and in Germany [28] for instance, and are now currently produced in more than 20 laboratories worldwide.

2.3.1 *Forced laser wakefield*

The forced laser wakefield regime [29] is reached when the laser pulse duration is approximately equal to the plasma period and when the laser waist is about the plasma wavelength. This regime allows a

reduction in heating effects that are produced when the laser pulse interacts with trapped electrons. In this regime, highly non-linear plasma waves can be reached as can be seen in Fig. 13.
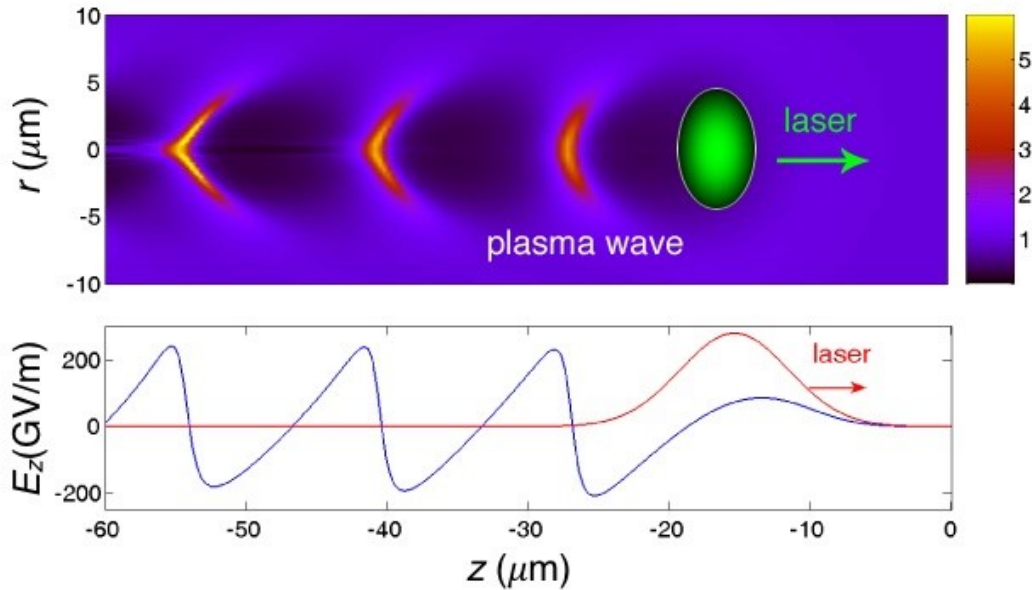
Fig. 13: Density perturbation (top) and electric field (bottom) produced in the non-linear regime

The laser power needs also to be greater that the critical power for relativistic self-focusing in order for the laser beam to shrink in time and in space. Due to self-focusing, pulse erosion can take place, which can allow efficient wake generation. Since the very front of the pulse is not self-focused, the erosion will be more severe. The wake then is mostly formed by this fast rising edge, and the back of the pulse has little interaction with the relativistic longitudinal oscillation of the plasma wave electrons. Indeed, the increase of plasma wavelength due to relativistic effects means that the breaking and accelerating peak of the plasma wave sits behind most, if not all, of the laser pulse. Hence its interaction, and that of the accelerated electrons with the laser pulse, is minimized, thus reducing emittance growth due to direct laser acceleration. Thanks to short laser pulses, plasma heating in the forced laser wakefield regime is significantly lower than in the self-modulated wakefield regime. This allows much higher plasma wave amplitudes to be reached, as well as higher electron energies. Thanks to a limited interaction between the laser and the accelerated electrons, the quality of the electron beam is also improved. Indeed, the normalized transverse emittance measured using the pepper pot technique has given values comparable to those obtained with conventional accelerators with an equivalent energy (normalized r.m.s. emittance $\varepsilon_n = 3\pi$ mm mrad for electrons at 55 ± 2 MeV) [30].

The three-dimensional simulations realized for this experiment showed that the radial plasma wave oscillations interact coherently with the longitudinal field, so enhancing the peak amplitude of the plasma wave. This, coupled with the aforementioned strong self-focusing, are ingredients absent from one-dimensional treatments of this interaction. Even in two-dimensional simulations, it was not possible to observe electrons beyond 200 MeV, as measured in this experiment, since except in three-dimensional simulations, both the radial plasma wave enhancement and self-focusing effects are underestimated. Hence it is only in three-dimensional simulations that $E_{max} \sim E_{wb}$ can be reached. That such large electric fields are generated demonstrates another important difference between FLW and SMWF regimes, since in the latter, plasma heating by instabilities limits the accelerating electric field to an order of magnitude below the cold wave-breaking limit. It should be noted that the peak electric field inferred for these FLW experiments is in excess of 1 TV/m, considerably larger than any other coherent
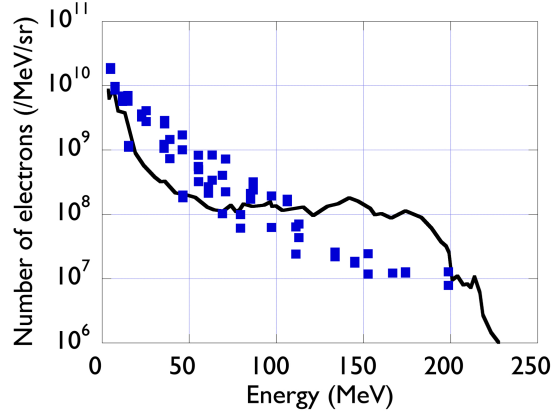
Fig. 14: Typical experimental (blue squares) and calculated (black curve) electron spectrum obtained at $n_e = 2.5 \times 10^{19}$ cm^{-3} with a 1 J, 30 fs laser pulse focused down to a waist of $w_0 = 18\mu$m.

accelerating structure created in the laboratory.

2.3.2 Bubble regime

In 2002, theoretical work based on three-dimensional particle-in-cell (PIC) simulations have shown the existence of a robust acceleration mechanism called the bubble regime [31]. In this regime, the dimensions of the focused laser are shorter than the plasma wavelength in longitudinal and also transverse directions, the laser shape appearing like a ball of light. If the laser energy contained in this spherical volume is large enough, the ponderomotive force of the laser expels radially and efficiently electrons from the plasma, which forms a cavity free from electrons behind the laser surrounded by a dense region of electrons. Behind the bubble, electron trajectories intersect each other. Electrons are injected into the cavity and accelerated along the laser axis, thus creating an electron beam with radial and longitudinal dimensions smaller than those of the laser (see Fig. 15).
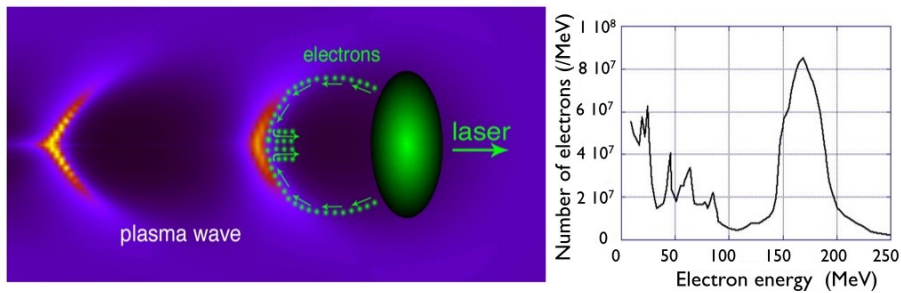
Fig. 15: Left: acceleration principle in the bubble regime. Right: typical quasi-monoenergetic electron spectra measured at LOA.

The signature of this regime is a quasi-monoenergetic electron distribution that results from the localization of injection at the back of the cavity, which gives similar initial properties in the phase space to injected electrons. Since electrons are trapped behind the laser pulse, this reduces or even suppresses interaction with the electric field of the laser. The trapping process stops when the charge contained in the cavity compensates the ionic charge, and the rotation in the phase space also leads to a shortening of the spectral width of the electron beam [32].

Several laboratories have obtained quasi-monoenergetic spectra: in France [33] with a laser pulse shorter than the plasma period, but also with pulses slightly longer than the plasma period in the U.K. [34], in the United States [35], then in Japan [36] and in Germany [37]. The interest in such a beam is a result of its importance for a number of applications: it is now possible to transport and to refocus this beam by magnetic fields. With a Maxwellian-like spectrum, it would have been necessary to select an energy range for the transport, which would have decreased significantly the electron flux. Electrons in the GeV level were also observed in this regime using a uniform plasma [38] or in plasma discharge, i.e. a plasma with a parabolic density profile [39] with a more powerful laser which propagates at high intensity over a longer distance. With the development of PW class lasers, a few GeV electron beam has been reported [40–42].

In all the experiments performed so far, the laser plasma parameters were not sufficient to fully enter the bubble/blowout regimes. Yet, with the increase of laser system power, this regime will be reached, and significant improvement of the reproducibility of the electron beam is expected. Nevertheless, since self-injection occurs through transverse wave-breaking, it is hardly appropriate for a fine tuning and control of the injected electron bunch. Figure 16 shows electron distributions obtained for different densities. It illustrates the transition from a Maxwellian-like spectrum obtained in high density cases in the self-modulated laser wakefield, to the forced laser wakefield regime with an emerging monoenergetic component at moderate density, to a spectrum containing a very well defined monoenergetic component. This transition occurs for densities around $n_e = 1$–3×10^{19} cm^{-3}. The best coupling for obtaining a high charge and a quasi-monoenergetic electron beam is at $n_e = 6 \times 10^{18}$ cm^{-3}. For this density, the image shows a narrow peak around 170 MeV, indicating efficient monoenergetic acceleration with a 24% energy spread corresponding to the spectrometer resolution.
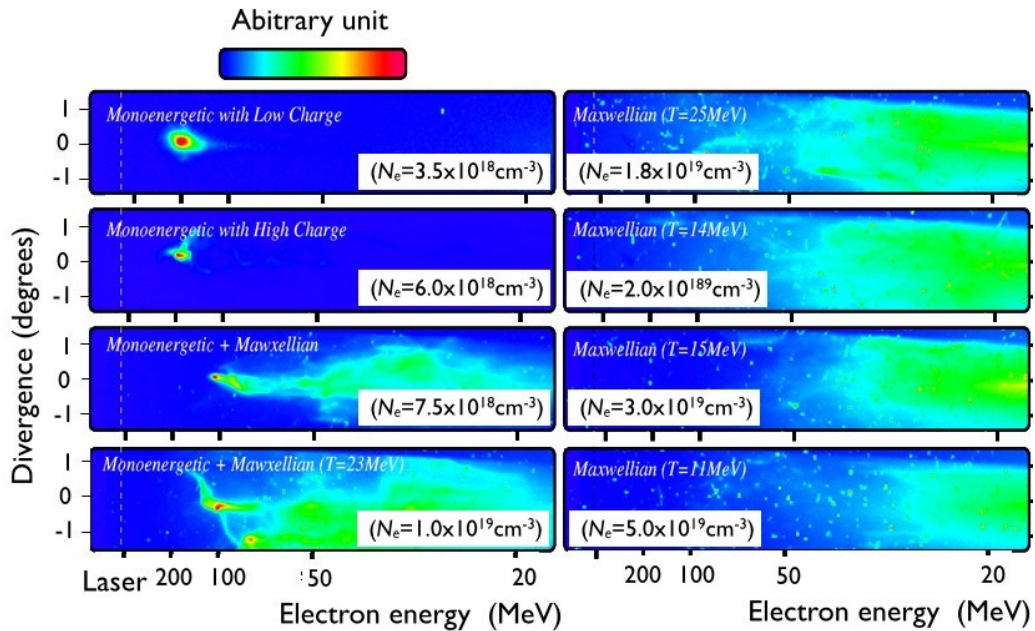
Fig. 16: Electron beam distribution for different plasma densities showing the transition from the self-modulated laser wakefield and the forced laser wakefield to the bubble/blow-out regime. From top to bottom, the plasma density values are 6×10^{18} cm^{-3}, 1×10^{19} cm^{-3}, 2×10^{19} cm^{-3} and 5×10^{19} cm^{-3}.

2.4 Injection in a density gradient

One solution to control electron injection with current laser technology was proposed by S. Bulanov *et al.* [43]. It involves a downward density ramp with a density gradient scale length L_{grad} smaller than the
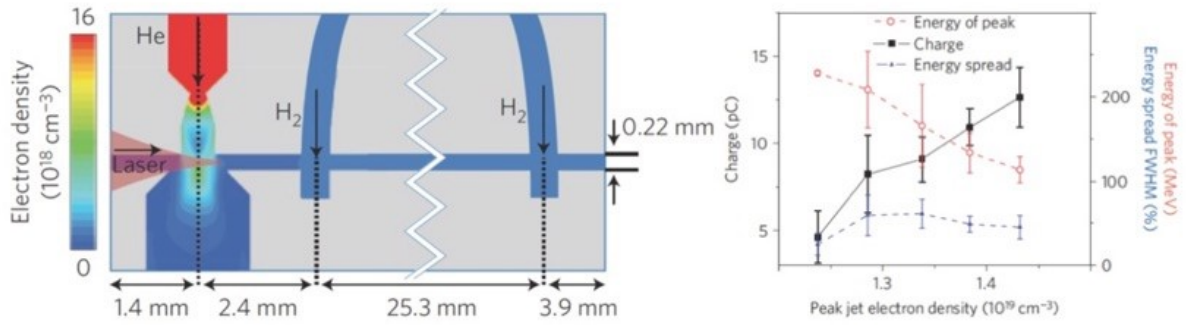
Fig. 17: Top: target schematic representation with embedded supersonic gas jet into a capillary that is filled with hydrogen gas. Bottom: the charge (squares), energy (circles) and energy spread (triangles) as a function of the peak jet density. From A. J. Gonsalves *et al.* [50].

plasma wavelength λ_p. Injection in a downward density ramp relies on the slowing down of the plasma wave velocity at the density ramp. This decrease of the plasma wave phase velocity lowers the threshold for trapping the plasma background electrons and causes wave-breaking of the wakefield in the density ramp. This method can therefore trigger wave-breaking in a localized spatial region of the plasma. Geddes *et al.* [44] showed the injection and acceleration of high charge (>300 pC) and stable quality beams of $\simeq 0.4$ MeV in the downward density ramp at the exit of a gas jet ($L_{grad} \simeq 100\,\mu m \gg \lambda_p$). These results, although very promising, have the disadvantage that the low energy beam blows up very quickly out of the plasma, due to the space-charge effect. To circumvent this issue, one should use a density gradient located early enough along the laser pulse propagation so that electrons can be accelerated to relativistic energies [45]. This can be achieved by using, for instance, a secondary laser pulse to generate a plasma channel transverse to the main pulse propagation axis [46]. In this case, the electron beam energy could be tuned by changing the position of the density gradient. In this pioneering experiment, the electron beam had a large divergence and a Maxwellian energy distribution because of a too low laser energy. However, two-dimensional PIC simulations showed that this method can result in high quality quasi-monoenergetic electron beams [47].

At LOA a density gradient across a laser created plasma channel was used to stabilize the injection [48]. The experiment was performed at an electron density close to the resonant density for the laser wakefield ($c\tau \sim \lambda_p$) to guaranty a post acceleration that delivered high quality electron beams with narrow divergences (4 mrad) and quasi-monoenergetic electron distributions with 50 to 100 pC charge and 10% relative energy spread.

The use of density gradients at the edges of a plasma channel showed an improvement of the beam quality and of the reproducibility with respect to those produced in the bubble/blowout regime with the same laser system and with similar laser parameters. However, the electron energy distribution was still found to fluctuate from shot to shot. The performance of the experiments could be further improved and could potentially lead to more stable and controllable high quality electron beams. In particular, sharper gradients with $L_{grad} \simeq \lambda_p$ coupled with a long plasma can lead to better beam quality [49].

For example, at LBNL, as shown in Fig. 17, electrons at 30 MeV were produced in a density ramp and accelerated up to 400 MeV in a second stage 4 cm parabolic plasma channel formed with a plasma discharge [50]. Here also, the density gradient injection led to an improvement of the stability and of the electron beam quality. The electron energy, divergence, charge and relative energy spread were found to be respectively 400 MeV, 2 mrad, 10 pC and 11%. It was shown that steeper density transitions, with $L_{grad} \ll \lambda_p$, can also cause trapping [51]. Such injection was successfully demonstrated experimentally using the shock-front created by a knife-edge inserted in a gas jet [52,53].
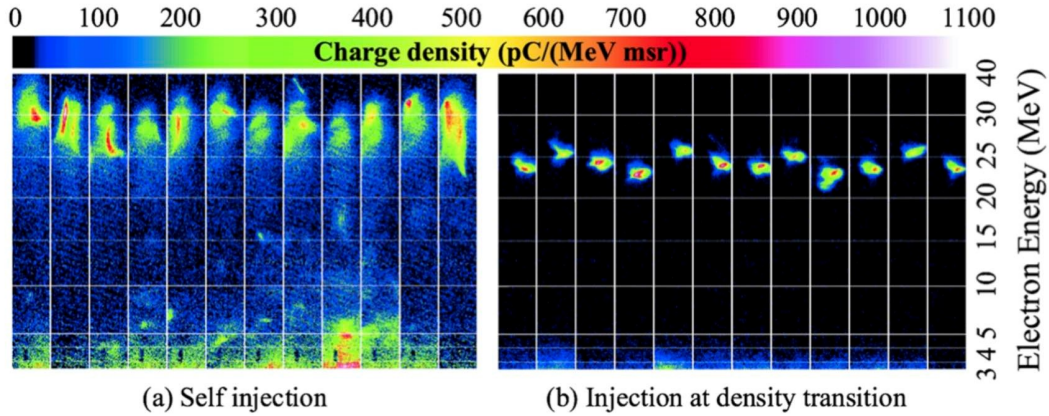
Fig. 18: A few representative shots of the 10% of all the shots with the lowest energy spread for self-injection (top) and injection at a density transition (bottom). The horizontal axis in each image corresponds to the transverse electron beam size; the vertical axis shows electron energy. From K. Schmid *et al.* [53].

Figure 18 illustrates the improvement of injection in a sharp density gradient, with a characteristic length of the order of the plasma wavelength and a peak electron density of about $5 \times 10^{19} \, \text{cm}^{-3}$. The experiment was performed at the Max-Planck-Institut fur Quantenoptik using a multi-TW sub-10 fs laser system that delivered for this experiment pulses with 65 mJ energy on target and with a duration of 8 fs FWHM. The laser pulse was focused down to a spot diameter of $12 \, \mu\text{m}$ FWHM into the gas target yielding a peak intensity of $2.5 \times 10^{18} \, \text{W/cm}^2$. The comparison between the self-injection and density transition injection shows a reduction of the relative energy spread and of the charge of a about a factor of 2.

2.5   Injection with colliding laser pulses

In 2006, stable and tunable quasi-monoenergetic electron beams were measured by using two counter-propagating laser beams in the colliding scheme. The use of two laser beams instead of one offers more flexibility and enables one to separate the injection from the acceleration process [54]. The first laser pulse, the pump pulse, is used to excite the wakefield while the second pulse, the injection pulse, is used to heat electrons during the collision with the pump pulse. After the collision has occurred, electrons are trapped and further accelerated in the wakefield, as shown in Fig. 19.

To trap electrons in a regime where self-trapping does not occur, one has either to inject electrons with energies greater that the trapping energy or dephase electrons with respect to the plasma wave. As mentioned earlier, electrons need to be injected in a very short time ($\ll \lambda_\text{p}/c$) in order to produce a monoenergetic beam. This can be achieved using additional ultra-short laser pulses whose only purpose is to trigger electron injection.

Umstadter *et al.* [21] first proposed to use a second laser pulse propagating perpendicular to the pump laser pulse. The idea was to use the radial ponderomotive kick of the second pulse to inject electrons. Esarey *et al.* [55] proposed a counter-propagating geometry based on the use of three laser pulses. This idea was further developed by considering the use of two laser pulses [56]. In this scheme, a main pulse (pump pulse) creates a high amplitude plasma wave and collides with a secondary pulse of lower intensity. The interference of the two beams creates a beatwave pattern, with a zero phase velocity, that heats some electrons from the plasma background. The force associated with this ponderomotive beatwave is proportional to the laser frequency. It is therefore many times greater than the ponderomotive force associated with the pump laser, that is inversely proportional to the pulse duration at resonance. As a result, the mechanism is still efficient even for modest laser intensities. Upon interacting with this field pattern, some background electrons gain enough momentum to be trapped in the main plasma wave and
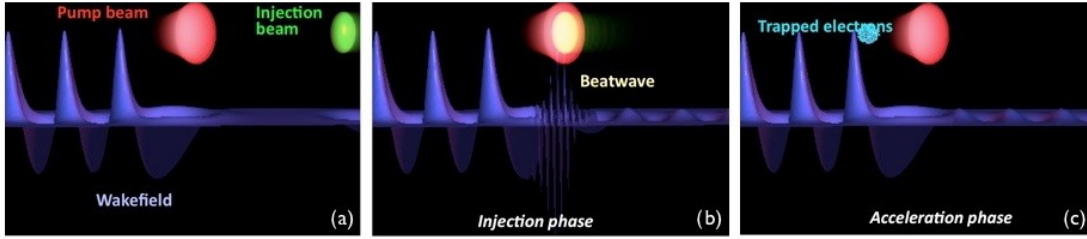
Fig. 19: Scheme of the principle of the injection with colliding laser pulses: (a) the two laser pulses propagate in opposite direction; (b) during the collision, some electrons get enough longitudinal momentum to be trapped by the relativistic plasma wave driven by the pump beam and (c) trapped electrons are then accelerated in the wake of the pump laser pulse.

then accelerated to high energies. As the overlapping of the lasers is short in time, the electrons are injected over a very short distance and can be accelerated to an almost monoenergetic beam.

This concept was validated in an experiment [54], using two counter-propagating pulses. Each pulse had a duration of 30 fs FWHM, with $a_0 = 1.3$ and $a_1 = 0.4$. They were propagated in a plasma with electron density $n_e = 7 \times 10^{18}$ cm^{-3} corresponding to $\gamma_p = k_0/k_p = 15$. It was shown that the collision of the two lasers could lead to the generation of stable quasi-monoenergetic electron beams. The beam energy could be tuned by changing the collision position in the plasma.

PIC simulations in one dimension have been used to model electron injection in the plasma wave at the collision of the two lasers, and their subsequent acceleration. In particular, the PIC simulations were compared to existing fluid models [55] with prescribed electric field. They showed significant differences, such as changes in the behaviour of plasma fields and in the amount of injected charge. The fluid approach fails to describe qualitatively and quantitatively many of the physical mechanisms that occur during and after the laser beams collision [57]. In this approach, the electron beam charge was found to be one order of magnitude greater than in the PIC simulations. For a correct description of injection, one has to describe properly (i) the heating process, e.g. kinetic effects and their consequences on the dynamics of the plasma wave during the beating of the two laser pulses and (ii) the laser pulse evolution which governs the dynamics of the relativistic plasma waves [58]. Unexpectedly, it was shown that efficient stochastic heating can be achieved when the two laser pulses are crossed polarized. The stochastic heating can be explained by the fact that for high laser intensities, the electron motion becomes relativistic which introduces a longitudinal component through the $\mathbf{v} \times \mathbf{B}$ force. This relativistic coupling makes it possible to heat electrons even in the case of crossed polarized laser pulses [59]. Thus the two perpendicular laser fields couple through the relativistic longitudinal motion of electrons. The heating level is modified by tuning the intensity of the injection laser beam or by changing the relative polarization of the two laser pulses [60]. This consequently changes the volume in the phase space of the injected electrons and therefore the charge and the energy spread of the electron beam.

Figure 20 shows, at a given time (42 fs), the longitudinal electric field during and after collision for parallel and crossed polarization. The solid line corresponds to PIC simulation results whereas the dotted line corresponds to fluid calculations. The laser fields are represented by the thin dotted line. When the pulses have the same polarization, electrons are trapped spatially in the beatwave and cannot sustain the collective plasma oscillation, inducing a strong inhibition of the plasma wave which persists after the collision. When the polarizations are crossed, the electron motion is only slightly disturbed compared to their fluid motion, and the plasma wave is almost unaffected during the collision, which tends to facilitate trapping.

Importantly, it was shown that the colliding pulse approach allows control of the electron beam energy which is done simply by changing the delay between the two laser pulses [54]. The robustness
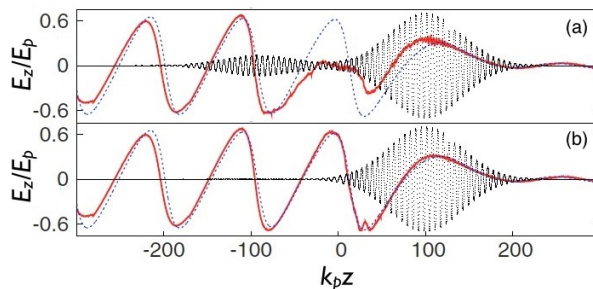
Fig. 20: Longitudinal electric field computed at $t = 43$ fs in one-dimensional PIC simulation (solid red line), and in fluid simulations (dotted blue line). The transverse electric field is also represented (thin dotted line). The laser pulse duration is 30 fs FWHM, the wavelength is 0.8 μm with $a_0 = 2$ and $a_1 = 0.4$. The laser pulses propagate in a plasma with electron density $n_e = 7 \times 10^{18}$ cm^{-3}. In (a) the case of parallel polarization and in (b) the case of crossed polarization.
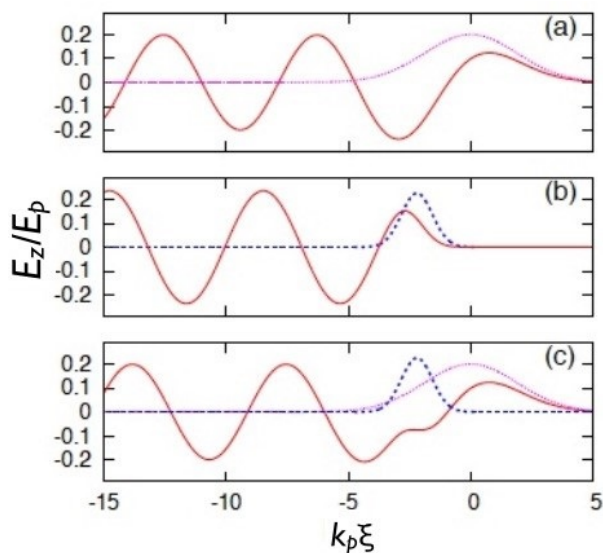
Fig. 21: In red, normalized longitudinal electric field. (a) The laser (in pink) wakefield. (b) The electron bunch (in blue) wakefield. (c) Field resulting from the superposition of the laser and electron beam wakefields. The normalized vector potential is $a_0 = 1$, the laser pulse duration is 30 fs, $n_e = 7 \times 10^{18}$ cm^{-3}, $n_{\text{beam}} = 0.11 \times n_e$, the bunch duration is 10 fs and its diameter is 4 μm. From C. Rechatin, Ph.D. thesis.

of this scheme permitted also very accurate studies of the dynamics of the electric field in presence of a high current electron beam to be carried out. Indeed, in addition to the wakefield produced by the laser pulse, a high current electron beam can also drive its own wakefield as shown in Fig. 21.

The beam loading effect contributes to the reduction of the relative energy spread of the electron beam. It was demonstrated that there is an optimal load which flattened the electric field, leading to the acceleration of all the electrons with the same value of the field, and producing consequently an electron beam with a very small, 1%, relative energy spread [61]. Thanks to the beam loading effect, the most energetic electrons can be slightly slowed down and accelerated to the same energy as the slowest ones. In cases of low charge beam, this effect does not play any role and the energy spread depends mainly on the heated volume. For a very high current, the load is too high and the most energetic electrons slow down so much that they eventually obtain energies smaller than the slowest electrons [61], increasing the relative energy spread. The existence of an optimal load was observed experimentally and supported
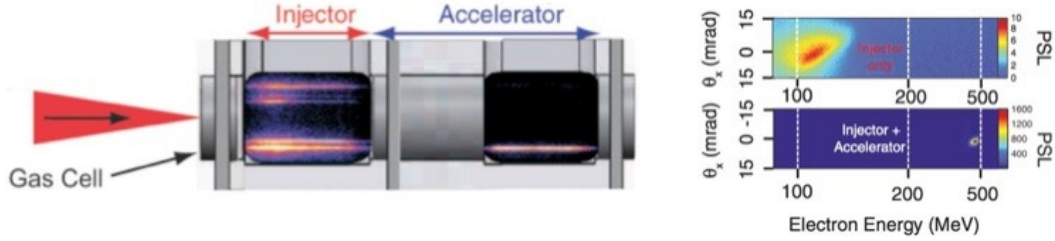
Fig. 22: (a) Schematic of the experimental set-up showing the laser beam, the two-stages gas cell, on the left the injector part and on the right the accelerator part. (b) Magnetically dispersed electron beam images from a 4 mm injector-only gas cell (top) and the 8 mm two-stages cell (bottom). From B. B. Pollock *et al.* [66].

by full three-dimensional PIC simulations. It corresponds to a peak current in the 20–40 kA range. The decelerating electric field due to the electron beam was found to be in the $GV/m/pC$ range.

2.6 Injection triggered by ionization

Another scheme was proposed recently to control the injection by using a high-Z gas and/or a high-Z/low-Z gas mixture. Thanks to the large differences in ionization potentials between successive ionization states of the atoms, a single laser pulse can ionize the low energy level electrons in its leading edge, drive relativistic plasma waves, and inject in the wakefield the inner level electrons which are ionized when the laser intensity is close to its maximum.

Such an ionization trapping mechanism was first demonstrated in electron beam driven plasma wave experiments on the Stanford Linear Collider (SLAC) [62]. Electron trapping from ionization of high-Z ions from capillary walls was also inferred in experiments on laser wakefield acceleration [63]. In the case of a self-guided laser driven wakefield, a mixture of helium and trace amounts of different gases was used [64,65]. In one of these experiments, electrons from the K shell of nitrogen were tunnel ionized near the peak of the laser pulse and were injected into and trapped by the wake created by electrons from majority helium atoms and the L shell of nitrogen. Because of the relativistic self-focusing effect, the laser propagates over a long distance with peak intensity variations that can trigger the injection over a long distance and in an inhomogeneous way, which leads to the production of a high relative energy spread electron beam. Importantly, the energy required to trap electrons is reduced, making this approach of great interest to produce electron beams with a large charge at moderate laser energy. To reduce the distance over which electrons are injected, experiments using two gas cells were performed at LLNL [66], as shown in Fig. 22. By restricting electron injection to a small region, in a first short cell filled with a gas mixture (the injector stage), energetic electron beams (of the order of 100 MeV) with a relatively large energy spread were generated. Some of these electrons were then further accelerated in a second, larger, accelerator stage, consisting of a long cell filled with low-Z gas, which increases their energy up to 0.5 GeV while reducing the relative energy spread to <5% FWHM.

2.7 Longitudinal injection

As has been shown, electron trapping is generally achieved by the wave-breaking of the plasma wake, a process that is by nature uncontrollable and leads generally to poor quality electrons. The presented controlled injection techniques, such as colliding pulse injection, ionization-induced injection and density gradient injection, have been developed to overcome this shortcoming. These methods offer an improved control on the acceleration and lead to better electron features, but they imply generally complex set-ups. For this reason, self-injection remains the most common method for injecting electrons in the plasma wake. Two distinct physical mechanisms can be distinguished: longitudinal and transverse self-injection. In longitudinal self-injection, the trajectory of injected electrons is mainly longitudinal,
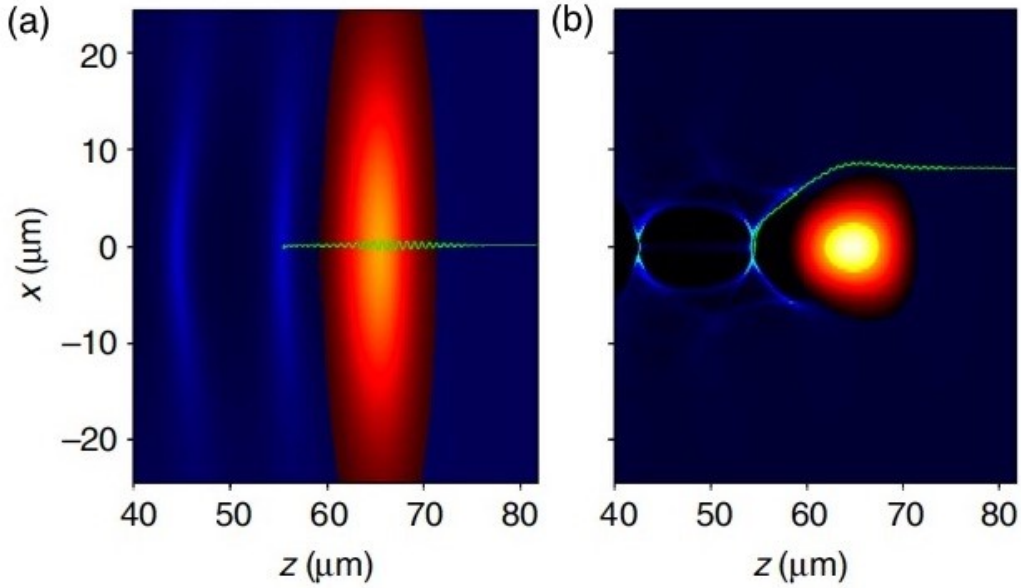
Fig. 23: Schematic for longitudinal and transverse self-injections. (a) Typical trajectory of an injected electron in the longitudinal self-injection mechanism. (b) Typical trajectory of an injected electron in the transverse self-injection mechanism. The blue colour scale represents the electron density. The red to yellow colour scale indicates the laser intensity. The trajectories are given by the green lines.

with a negligible transverse motion. As shown by the schematic in Fig. 23, the injected electrons pass through the laser pulse and gain energy while crossing the plasma wave. When they reach the rear of the first plasma period, their velocity exceeds the wake phase velocity and the electrons are eventually injected. The only electrons that are trapped are those that were initially close to the axis where the laser intensity and the wakefield amplitude are the highest and where the ponderomotive force is small. The longitudinal self-injection mechanism is analogous to one-dimensional longitudinal wave-breaking. In contrast, transverse self-injection occurs in the bubble regime, where the laser ponderomotive force expels electrons from the propagation axis and forms an electron-free cavity in its wake. As shown in Fig. 23 and in Fig. 15, the injected electrons are initially located at approximately one laser waist from the axis. They circulate around the laser pulse and the bubble, and attain a velocity larger than the wake phase velocity when reaching the axis at the rear of the bubble.

During its propagation, the laser pulse evolves, the self-phase-modulation modifies its duration and the relativistic self-focusing modifies its initial transverse shape. As a consequence, the generated wakefield is not uniform along the laser propagation axis and electrons can be self-injected at different positions of the plasma accelerator. Electrons in the second bunch originate from positions close to the laser waist, as expected in the case of trapping by transverse self-injection. In contrast, electrons in the first bunch come from regions close to the axis. When these electrons are injected, the laser spot radius is large and the normalized laser amplitude is still low; hence, the radial ponderomotive force close to the axis is small.

Thus, on-axis electrons are only weakly deviated when crossing the laser pulse, and they remain in the region of largest accelerating field E_z. Moreover, the laser amplitude increases steeply in the region of first injection because of laser self-focusing (see Fig. 24). This reduces the wake phase velocity via the relativistic shift of the plasma wavelength λ_p [67]. The strongly reduced wake phase velocity lowers the threshold for trapping, such that electrons can catch up the plasma wave and be injected despite a low a_0, similarly to density gradient injection.
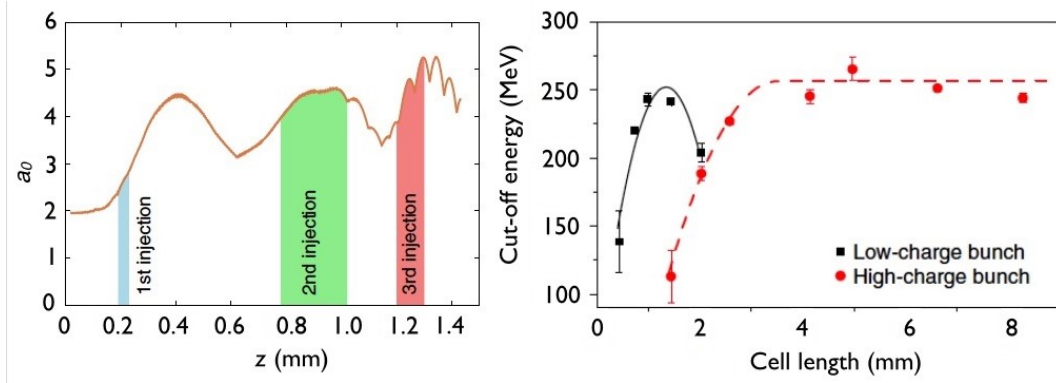
Fig. 24: Left: Evolution of the normalized laser amplitude. The coloured areas indicate the injection regions. Right: Electron spectra for a density of $1.1 \times 10^{19}\,\mathrm{cm}^{-3}$. Mean 5% cut-off energy as a function of the cell length. The solid line is a parabolic fit, the dashed line is a visual guide. The peak accelerating electric field is $E^z_{\max} = 340 \pm 65\,\mathrm{GV/m}$ for the first bunch and $E^z_{\max} = 185 \pm 40\,\mathrm{GV/m}$ for the second bunch.

3 Particle beam wakefield accceleration

3.1 Electron driven plasma wakefield

Wakefields in a plasma can be also driven by an electron bunch that has, at resonance, a length of about half the plasma wavelength. Whereas in the laser wakefield case the radiation pressure, known as the ponderomotive force, pushes away the plasma electrons, here the force is due the space-charge of the electron beam. The plasma electrons are strongly blown out radially, but because of the space-charge attraction of the plasma ions, they are attracted back towards the rear of the beam where they overshoot the beam axis and set up a wakefield oscillation. Here again, charged particles injected in an appropriately phased trailing pulse can then extract energy from the wakefield. Because of the lack of accelerators that deliver suitable electron beams, there are fewer particle-beam-driven plasma acceleration experiments compared with laser accelerator experiments. The first beam-driven plasma wakefield experiments were carried out at the Argonne Wakefield Accelerator Facility in the 1980s [68]. Since then, important experiments done at SLAC by the UCLA/USC/SLAC collaboration have mapped the physics of electron and positron beam-driven wakes and has shown acceleration gradients of $40\,\mathrm{GV/m}$ using electron beams with metre-scale plasmas [69]. In the first important SLAC experiments only one electron bunch was used to excite the wakefield. Since the energy of the drive pulse was 42 GeV, both the electrons and the wake are moving at a velocity close to c, so there is no relative motion between the electrons and the wakefield. Because the electron bunch was also longer that the plasma period, most of the electrons in the drive bunch lose energy in exciting the wake, but some electrons in the back have gained energy from the wakefield as the wakefield changes its sign. Thanks to the high quality, low emittance of the electron bunch, its intensity was so high that the 42 GeV electron beam passed through a column of lithium vapour 85 cm long, the head of the beam created a fully ionized plasma and the remainder of the beam excited a strong wakefield. Figure 25 shows the energy spectrum of the beam measured after the plasma. The electrons in the bulk of the pulse that lost energy in driving the wake are mostly dispersed out of the field of view of the spectrometer camera and so are not seen in the spectrum. However, electrons in the back of the same pulse are accelerated and reach energies up to 85 GeV. The measured spectrum of the accelerated particles was in good agreement with the spectrum obtained from computer simulations of the experiment, as Fig. 25 shows. As said Prof. C. Joshi, 'This is a remarkable result when one realizes that while it takes the full 3 km length of the SLAC linac to accelerate electrons to 42 GeV, some of these electrons can be made to double their energy in less than a metre'.

In this former experiment, a small fraction of electrons of the beam was injected and accelerated. As a consequence, the quality of the accelerated electrons was poor with a long Maxwellian-like tail
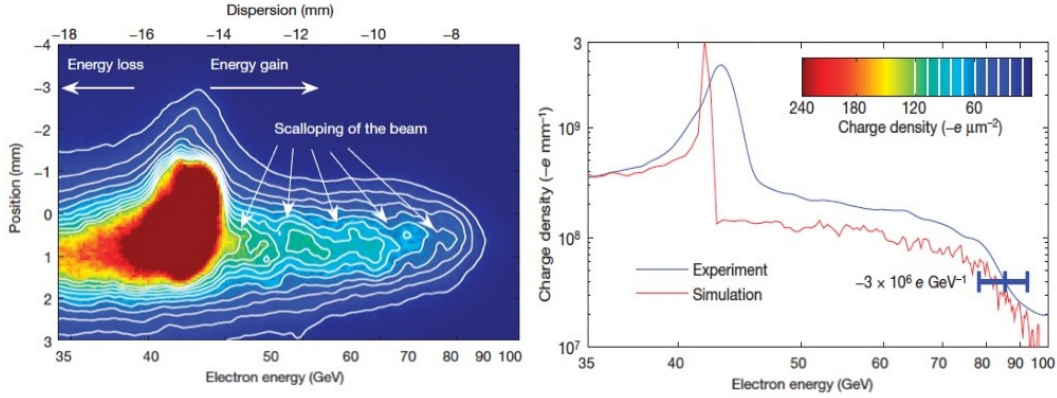
Fig. 25: Energy spectrum of the electrons in the 35–100 GeV range. The dispersion (shown on the top axis) is inversely proportional to the particle energy (shown on the bottom axis). The head of the pulse, which is unaffected by the plasma, is at 43 GeV. The core of the pulse, which has lost energy driving the plasma wake, is dispersed partly out of the field of view of the camera. Particles in the back of the bunch, which have reached energies up to 85 GeV, are visible to the right.

and therefore with also a poor energy transfer efficiency. For future high energy physics purposes, a high efficiency is mandatory to achieve an affordable and compact high-energy collider. To improve this, a second important experiment has been performed at SLAC. In this plasma wakefield acceleration experiment, a charge-density wake with high accelerating fields has been driven by an ultra-relativistic bunch of charged particles (the drive bunch) through a plasma followed by a second bunch of relativistic electrons (the trailing bunch) in the wake of the drive bunch at an appropriate distance that has been efficiently accelerated to higher energy. Whereas in the previous experiment, the total charge of accelerated electrons was insufficient to extract a substantial amount of energy from the wake, here high efficiency acceleration of the trailing bunch of electrons has been demonstrated. Accelerations of approximately 70–80 pC of the trailing bunch have been achieved in an accelerating gradient of about 4.4 GV/m. As presented in Fig. 26, these particles have gained approximately 1.6 GeV of energy per particle, with a final energy spread as low as 1% and an energy-transfer efficiency from the wake to the bunch that exceeded 30%. This acceleration of a distinct bunch of electrons containing a substantial charge and having a small energy spread with both a high accelerating gradient and a high energy-transfer efficiency represents a milestone in the development of plasma wakefield acceleration into a compact and affordable accelerator technology. 6% of the initial electron beam energy (36 J) was transferred to the trailing bunch. This value is comparable to the laser to electron beam energy transfer efficiency from LPAW. The main advantage here being that the driver is more efficient that the laser driver. Accelerators have indeed today a wall-plug efficiency more that 10 times larger than lasers.

3.2 Proton driven plasma wakefield

As it has been shown in all these former experiments, the energy gain was limited by the energy carried by the driver (about 40 J for an e-beam driver and about 100 J for a laser driver) and by the propagation length of the driver in the plasma (few tens of centimetres for the e-driver and few centimetres for the laser driver). The laser pulse and electron bunch driver schemes therefore require the use of many acceleration stages in the tens of GeV each in order to gain TeV energy levels. A 10 GeV stage that delivers an nC of charge corresponds to an energy of 10 J, and it will correspond to 10 kJ for a 10 TeV stage. If one assumes 10% energy transfer efficiency from the driver to the trail bunch, this indicates that the driver energy must contain about 100 kJ. In 2009, for the first time, plasma-wake excitation by a relativistic proton beam has been considered [5]. In this ideal case, the proton driver beam has to be resonant with
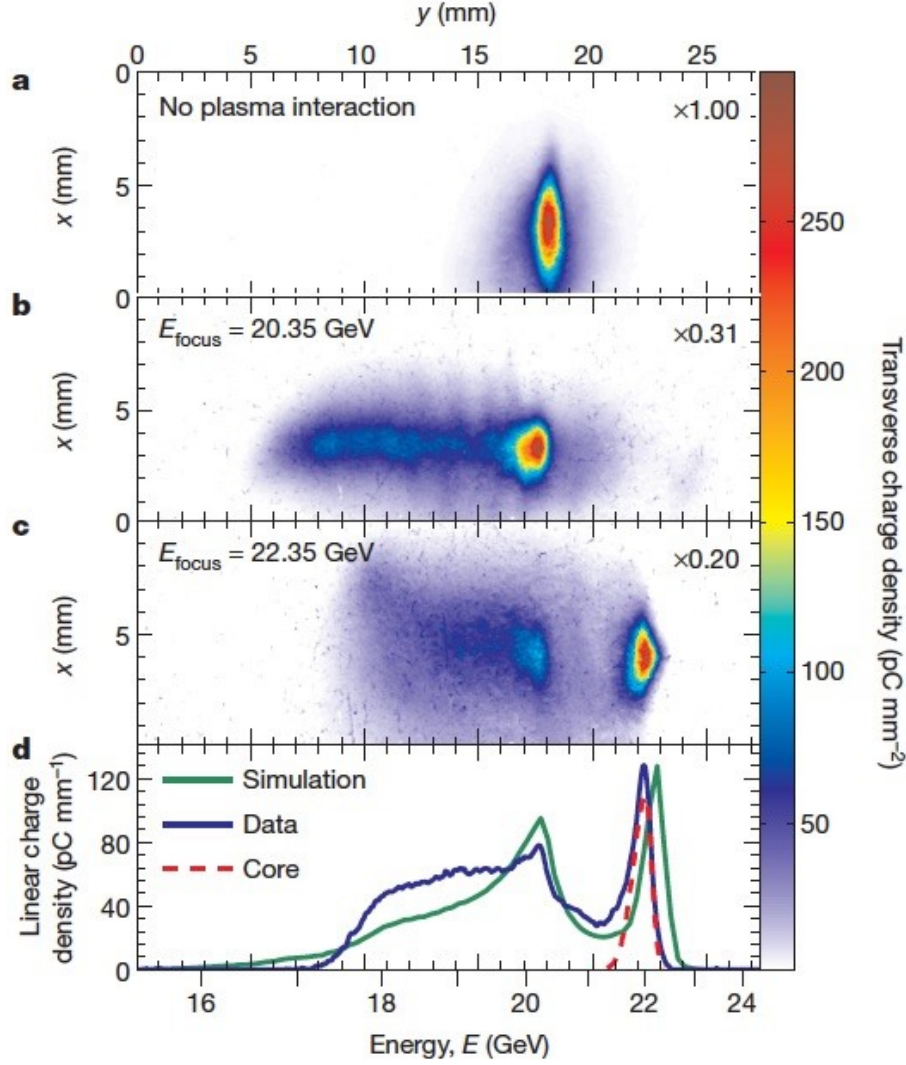
Fig. 26: (a) The dispersed electron beam profile without plasma interaction, where the spectrometer is set to image 22.35 GeV. (b) and (c) The dispersed beam profile after the electron bunches have interacted with the plasma, where the spectrometer is set to image 20.35 GeV and 22.35 GeV. (d) The spatially integrated spectrum (in x) or the linear charge density of the bunches shown in (c) (solid blue line) along with the final spectrum obtained from the simulation (solid green line in (d)). The core of the accelerated trailing beam is shown for the data (dashed red line).

the plasma and it was predicted on the basis of numerical simulations that 10 GeV electrons injected could be accelerated to 0.5 TeV in a 450 m proton wakefield. Unfortunately such a short proton bunch does not exist, and therefore it has been proposed to use the CERN SPS 19 kJ, 400 GeV proton beam that is produced routinely.

Because the length of the driver (about 10 cm) is much longer than the plasma wavelength (about 1 mm) at a density large enough (of 10^{14} cm^{-3}) to reach a GV/m accelerating field, the interaction has to occur in the self-modulated regime. In this regime, also called the self-modulation instability (SMI) [70], the proton bunch is split during its propagation in several micro-bunches that excite resonantly a strong plasma wave. A first experiment, called AWAKE [71], will be performed at CERN in the next two years to demonstrate the possibility of accelerating injected electrons in proton driven plasma wakefield. For this, an extremely uniform long plasma has been developed with a control precision for the density that
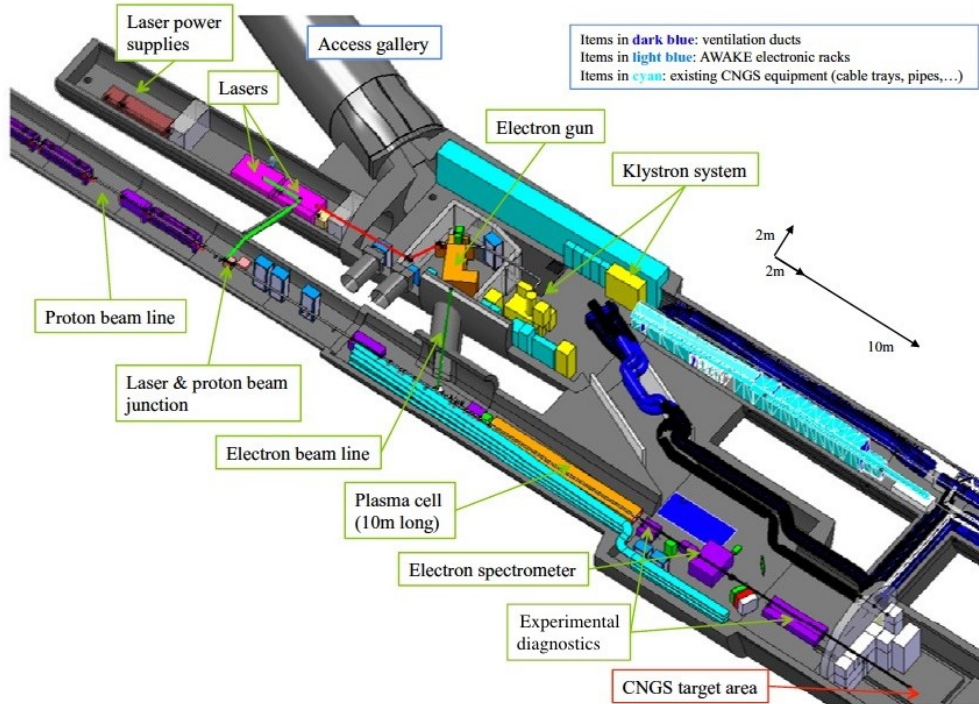
Fig. 27: Design of the layout of the AWAKE experiment

must be within 0.5% over metres. To optimize the coupling efficiency, electrons will be injected at an angle after the self-modulated instability has reached saturation. The conceptual design of the proposed AWAKE experiment is shown in Fig. 27 where the laser and proton bunches are made co-linear. The laser that will ionize the metal vapour is required to seed the self-modulation instability. The self-modulated proton bunch (shown on the left hand side) enters a second plasma section where it drives the plasma wakefield structure (shown on the right side). The electrons are injected in the wakefields and the 2 GeV accelerated electrons will be measured with an electron spectrometer. The AWAKE experiment will be installed in the CERN Neutrinos to Gran Sasso (CNGS) facility. Approximately 5% of electrons are supposed to be trapped and accelerated to the end of the 10 m plasma with accelerating gradients in the few GV/m. In addition to the electron spectrometer, several other diagnostics will be used to characterize the proton beams to better understand the physics of self-modulation. Coherent Transition Radiation (in the visible and in the infra-red) produced when the proton beam passes through a thin foil, will be measured using a streak camera. Additionally, transverse coherent transition radiation will be produced and detected using electro-optical sensors; this will be the first experimental use of this recent concept [72]. First protons to the experiment are expected at the end of 2016 and this will be followed by an initial 3–4 year experimental program of four periods of two weeks of data taking.

4 Future of the laser plasma accelerators

The tremendous progress that has been made in plasma acceleration [73–75], from the first acceleration of externally injected electrons in a GV/m laser wakefield, self-injection in a 100 GV/m laser wakefield with first a 100 MeV broad spectra to the series of experiments with the production of a quasi-monoenergetic electron beam in a laser wakefield with a compact 10 Hz laser system have contributed to boosting this field of research in which tens of laboratories/teams are playing important roles in a competitive and friendly approach. The evolution of short-pulse laser technology with diode pump lasers or fibre lasers, a field in rapid progress, will eventually contribute to the improvement of laser plasma

acceleration and their societal applications, in material science for example for high resolution gamma radiography [76, 77], in medicine for cancer treatment [78, 79], in chemistry [80, 81] and in radiobiology [82–84]. In the near future, the development of compact free electron lasers could open the way to the production of intense X-ray beams, in a compact way, by coupling the electron beam with undulators. Thanks to the very high peak current of a few kiloamperes [85], comparable to the current used at LCLS, the use of laser plasma accelerators for free electron lasers, the so-called fifth generation light source, is clearly identified by the scientific community as a major development. Alternative schemes to produce ultra-short X-ray beams, such as Compton, betatron or Bremsstrahlung X-ray sources, have also been considered. Tremendous progress has been made regarding the study of betatron radiation in a laser plasma accelerator. Since its first observation in 2004 [86] and the first monitoring of electron betatronic motion in 2008 [87], a number of articles have reported in more detail this new source, including measures of a sub ps duration [88] and of a transverse size in the micrometre range [89]. Betatron radiation was used recently to perform with high spatial resolution, of about 10 microns, X-ray contrast phase images in a single shot mode operation [90, 91]. In parallel, similar huge progress has been performed in accelerating electrons and positrons using electron or positron bunches, with here, a gain of a few tens of GeV in a few tens of centimetres accelerating gradient. Wakefields driven by electron beams are good candidates to boost electron energy in a metre long plasma device. The requirement for the driver being very close to the one for FEL purposes, shorter radiation wavelength could be produced by doubling, for example, the electron energy delivered from SLAC or from DESY. Acceleration of electrons and positrons with these drivers are also very relevant for a staging approach for high energy physics purposes. The AWAKE experiment will certainly contribute to defining the roadmap for future larger-scale R&D projects on laser, electron or proton driven plasma wakefield acceleration for future high energy colliders for particle physics. The success of plasma wakefield accelerators will open a pathway towards many exciting societal application, a compact FEL radiation source and a revolutionary plasma-based TeV lepton collider. This revolution could then enable ground-breaking discoveries in many domains, including particle physics.

Acknowledgements

References

[1] G.I. Budker, Proc. CERN Symposium on High Energy Accelerators Vol. 1, p. 68 (1956).

[2] T. Tajima and J.M. Dawson, *Phys. Rev. Lett.* **43**(4) (1979) 267.
http://dx.doi.org/10.1103/PhysRevLett.43.267

[3] A. Modena, A. Dangor, Z. Najmudin, C. Clayton, K. Marsh, C. Joshi, V. Malka, C. Darrow, C. Danson, D. Neely and F. Walsh, *Nature* **377** (1995) 606. http://dx.doi.org/10.1038/377606a0

[4] P. Chen, J.M. Dawson, R.W. Huff and T. Katsouleas, *Phys. Rev. Lett.* **54**(7) (1985) 693.
http://dx.doi.org/10.1103/PhysRevLett.54.693

[5] A. Caldwell, K. Lotov, A. Pukhov and F. Simon, *Nature Phys.* **5** (2009) 363.
http://dx.doi.org/10.1038/nphys1248

[6] L.M. Gorbunov and V.I. Kirsanov, *Sov. Phys. JETP* **66** (1987) 290.

[7] F. Amiranoff *et al.*, *Phys. Rev. Lett.* **81**(5) (1998) 995.
http://dx.doi.org/10.1103/PhysRevLett.81.995

[8] C.E. Clayton, C. Joshi, C. Darrow and D. Umstadter, *Phys. Rev. Lett.* **54**(21) (1985) 2343. http://dx.doi.org/10.1103/PhysRevLett.54.2343

[9] C.E. Clayton, K.A. Marsh, A. Dyson, M. Everett, A. Lal, W.P. Leemans, R. Williams and C. Joshi, *Phys. Rev. Lett.* **70**(1) (1993) 37. http://dx.doi.org/10.1103/PhysRevLett.70.37

[10] M. Everett, A. Lal, D. Gordon, C. Clayton, K. Marsh and C. Joshi, *Nature* **368** (1994) 527. http://dx.doi.org/10.1038/368527a0

[11] F. Amiranoff *et al.*, *Phys. Rev. Lett.* **74**(26) (1995) 5220. http://dx.doi.org/10.1103/PhysRevLett.74.5220

[12] Y. Kitagawa *et al.*, *Phys. Rev. Lett.* **68**(1) (1992) 48. http://dx.doi.org/10.1103/PhysRevLett.68.48

[13] A. Dyson *et al.*, *Plasma Phys. Control. Fusion* **38**(4) (1996) 505. http://dx.doi.org/10.1088/0741-3335/38/4/005

[14] N.A. Ebrahim, *J. Appl. Phys.* **76**(11) (1994) 7645. http://dx.doi.org/10.1063/1.357937

[15] F. Amiranoff *et al.*, *Phys. Rev. Lett.* **68**(25) (1992) 3710. http://dx.doi.org/10.1103/PhysRevLett.68.3710

[16] B. Walton *et al.*, *Opt. Lett.* **27**(24) (2002) 2203. http://dx.doi.org/10.1364/OL.27.002203

[17] P. Sprangle, E. Esarey, J. Krall and G. Joyce, *Phys. Rev. Lett.* **69**(15) (1992) 2200. http://dx.doi.org/10.1103/PhysRevLett.69.2200

[18] T.M. Antonsen, Jr. and P. Mora, *Phys. Rev. Lett.* **69**(15) (1992) 2204. http://dx.doi.org/10.1103/PhysRevLett.69.2204

[19] N.E. Andreev, L.M. Gorbunov, V.I. Kirsanov, A.A. Pogosova and R.R. Ramazashvili, *JETP Lett.* **55** (1992) 571.

[20] C. Joshi, T. Tajima, J.M. Dawson, H.A. Baldis and N.A. Ebrahim, *Phys. Rev. Lett.* **47**(18) (1981) 1285. http://dx.doi.org/10.1103/PhysRevLett.47.1285

[21] D. Umstadter, S.-Y. Chen, A. Maksimchuk, G. Mourou and R. Wagner, *Science* **273**(5274) (1996) 472. http://dx.doi.org/10.1126/science.273.5274.472

[22] C.I. Moore, A. Ting, K. Krushelnick, E. Esarey, R.F. Hubbard, B. Hafizi, H.R. Burris, C. Manka and P. Sprangle, *Phys. Rev. Lett.* **79**(20) (1997) 3909. http://dx.doi.org/10.1103/PhysRevLett.79.3909

[23] C.E. Clayton *et al.*, *Phys. Rev. Lett.* **81**(1) (1998) 100. http://dx.doi.org/10.1103/PhysRevLett.81.100

[24] V. Malka *et al.*, *Phys. Plasmas* **8** (2001) 2605.

[25] E. Esarey, P. Sprangle, J. Krall and A. Ting, *IEEE Trans. Plasma Sci.* **24**(2) (1996) 252. http://dx.doi.org/10.1109/27.509991

[26] W.P. Leemans, J. van Tilborg, J. Faure, C.G.R. Geddes, C.Tóth, C.B. Schroeder, E. Esarey, G. Fubiani and G. Dugan, *Phys. Plasmas* **11**(5) (2004) 2899. http://dx.doi.org/10.1063/1.1652834

[27] T. Hosokai, K. Kinoshita, A. Zhidkov, K. Nakamura, T. Watanabe, T. Ueda, H. Kotaki, M. Kando, K. Nakajima and M. Ueseka. *Phys. Rev. E* **67**(3) (2003) 036407. http://dx.doi.org/10.1103/PhysRevE.67.036407

[28] C. Gahn, G.D. Tsakiris, A. Pukhov, J. Meyer-ter-Vehn, G. Pretzler, P. Thirolf, D. Habs and K.J. Witte. *Phys. Rev. Lett.* **83**(23) (1999) 4772. http://dx.doi.org/10.1103/PhysRevLett.83.4772

[29] V. Malka *et al.*, *Science* **298**(5598) (2002) 1596. http://dx.doi.org/10.1126/science.1076782

[30] S. Fritzler, E. Lefebvre, V. Malka, F. Burgy, A. E. Dangor, K. Krushelnick, S.P.D. Mangles, Z. Najmudin, J.-P. Rousseau and B. Walton, *Phys. Rev. Lett.* **92**(16) (2004) 165006. http://dx.doi.org/10.1103/PhysRevLett.92.165006

[31] A. Pukhov and J. Meyer-ter-Vehn, *Appl. Phys.* B **74**(4-5) (2002) 355. http://dx.doi.org/10.1007/s003400200795

[32] F.S. Tsung, R. Narang, W.B. Mori, C. Joshi, R.A. Fonseca and L. O. Silva, *Phys. Rev. Lett.* **93**(18) (2004) 185002. http://dx.doi.org/10.1103/PhysRevLett.93.185002

[33] J. Faure, Y. Glinec, A. Pukhov, S. Kiselev, S. Gordienko, E. Lefebvre, J.-P. Rousseau, F. Burgy and V. Malka. *Nature* **431** (2004) 541. http://dx.doi.org/10.1038/nature02963

[34] S. Mangles *et al.*, *Nature* **431** (2004) 535. http://dx.doi.org/10.1038/nature02939

[35] C.G.R. Geddes, C. Tóth, J. van Tilborg, E. Esarey, C.B. Schroeder, D. Bruhwiler, C. Nieter, J. Cary and W.P. Leemans, *Nature* **431** (2004) 538. http://dx.doi.org/10.1038/nature02900

[36] E. Miura, K. Koyama, S. Kato, S. Saito, M. Adachi, Y. Kawada, T. Nakamura and M. Tanimoto, *Appl. Phys. Lett.* **86**(25) (2005) 251501. http://dx.doi.org/10.1063/1.1949289

[37] B. Hidding *et al.*, *Phys. Rev. Lett.* **96**(10) (2006) 105004. http://dx.doi.org/10.1103/PhysRevLett.96.105004

[38] N. A. M. Hafz *et al.*, *Nature Photon.* **2** (2008) 571. http://dx.doi.org/10.1038/nphoton.2008.155

[39] W.P. Leemans, B. Nagler, A.J. Gonsalves, Cs. Tòth, K. Nakamura, C.G.R. Geddes, E. Esarey, C.B. Schroeder and S.M. Hooker, *Nature Phys.* **2** (2006) 696. http://dx.doi.org/10.1038/nphys418

[40] W.P. Leemans *et al.*, *Phys. Rev. Lett.* **113**(24) (2014) 245002. http://dx.doi.org/10.1103/PhysRevLett.113.245002

[41] H.T. Kim, K.H. Pae, H.J. Cha, I.J. Kim, T.J. Yu, J.H. Sung, S.K. Lee, T.M. Jeong and J. Lee, *Phys. Rev. Lett.* **111**(16) (2013) 165002. http://dx.doi.org/10.1103/PhysRevLett.111.165002

[42] X. Wang, R. Zgadzaj, N. Fazel, Z. Li, S.A. Yi, X. Zhang, W. Henderson, Y.-Y. Chang,R. Korzekwa, H.-E. Tsai, C.-H. Pai, H. Quevedo,G. Dyer, E. Gaul, M. Martinez, A. C. Bernstein, T. Borger, M. Spinks, M. Donovan, V. Khudik, G. Shvets, T. Ditmire, M. C. Downer, *Nature Commun.* **4** (2013). http://dx.doi.org/10.1038/ncomms2988

[43] S. Bulanov, N. Naumova, F. Pegoraro and J. Sakai, *Phys. Rev.* E **58**(5) (1998) R5257(R). http://dx.doi.org/10.1103/PhysRevE.58.R5257

[44] C.G.R. Geddes, K. Nakamura, G. R. Plateau, Cs. Tóth, E. Cormier-Michel, E. Esarey, C.B. Schroeder, J.R. Cary and W.P. Leemans, *Phys. Rev. Lett.* **100**(21) (2008) 215004. http://dx.doi.org/10.1103/PhysRevLett.100.215004

[45] J.U. Kim, N. Hafz and H. Suk, *Phys. Rev.* E **69**(2) (2004) 026409. http://dx.doi.org/10.1103/PhysRevE.69.026409

[46] T.-Y. Chien, C.-L. Chang, C.-H. Lee, J.-Y. Lin, J. Wang and S.-Y. Chen, *Phys. Rev. Lett.* **94**(11) (2005) 115003. http://dx.doi.org/10.1103/PhysRevLett.94.115003

[47] P. Tomassini, M. Galimberti, A. Giulietti, D. Giulietti, L.A. Gizzi, L. Labate and F. Pegoraro, *Phys. Rev. ST Accel. Beams* **6**(12) (2003) 121301. http://dx.doi.org/10.1103/PhysRevSTAB.6.121301

[48] J. Faure, C. Rechatin, O. Lundh, L. Ammoura and V. Malka, *Phys. Plasmas* **17**(8) (2010) 083107. http://dx.doi.org/10.1063/1.3469581

[49] A.V. Brantov, T.Z. Esirkepov, M. Kando, H. Kotaki, V.Yu. Bychenkov and S.V. Bulanov, *Phys. Plasmas* **15**(7) (2008) 073111. http://dx.doi.org/10.1063/1.2956989

[50] A.J. Gonsalves *et al.*, *Nature Phys.* **7** (2011) 862. http://dx.doi.org/10.1038/nphys2071

[51] H. Suk, N. Barov, J.B. Rosenzweig and E. Esarey, *Phys. Rev. Lett.* **86**(6) (2001) 1011. http://dx.doi.org/10.1103/PhysRevLett.86.1011

[52] K. Koyama, A. Yamazaki, A. Maekawa, M. Uesaka, T. Hosokai, M. Miyashita, S. Masuda and E. Miura, *Nucl. Instrum. Methods Phys. Res.* A **608**(1) (2009) S51. http://dx.doi.org/10.1016/j.nima.2009.05.142

[53] K. Schmid, A. Buck, C.M.S. Sears, J.M. Mikhailova, R. Tautz, D. Herrmann, M. Geissler, F. Krausz and L. Veisz, *Phys. Rev. ST Accel. Beams* **13**(9) (2010) 091301. http://dx.doi.org/10.1103/PhysRevSTAB.13.091301

[54] J. Faure, C. Rechatin, A. Norlin, A. Lifschitz, Y. Glinec and V. Malka, *Nature* **444** (2006) 737. http://dx.doi.org/10.1038/nature05393

[55] E. Esarey, P. Sprangle, J. Krall and A. Ting, *IEEE J. Quant. Electron.* **33**(11) (1997) 1879-1914. http://dx.doi.org/10.1109/3.641305

[56] G. Fubiani, E. Esarey, C. B. Schroeder and W. P. Leemans, *Phys. Rev.* E **70**(1) 016402 (2004). http://dx.doi.org/10.1103/PhysRevE.70.016402

[57] C. Rechatin, J. Faure, A. Lifschitz, V. Malka, and E. Lefebvre, *Phys. Plasmas* **14**(6) (2007) 060702. http://dx.doi.org/10.1063/1.2741387

[58] X. Davoine, E. Lefebvre, J. Faure, C. Rechatin, A. Lifschitz and V. Malka, *Phys. Plasmas* **15**(11) (2008) 113102. http://dx.doi.org/10.1063/1.3008051

[59] V. Malka, J. Faure, C. Rechatin, A. Ben-Ismail, J.K. Lim, X. Davoine and E. Lefebvre, *Phys. Plasmas* **16**(5) (2009) 056703. http://dx.doi.org/10.1063/1.3079486

[60] C. Rechatin, J. Faure, A. Lifschitz, X. Davoine, E. Lefebvre and V. Malka, *New J. Phys.* **11** (2009) 013011. http://dx.doi.org/10.1088/1367-2630/11/1/013011

[61] C. Rechatin, X. Davoine, A. Lifschitz, A. Ben Ismail, J. Lim, E. Lefebvre, J. Faure and V. Malka, *Phys. Rev. Lett.* **103**(19) (2009) 194804. http://dx.doi.org/10.1103/PhysRevLett.103.194804

[62] E. Oz *et al.*, *Phys. Rev. Lett.* **98**(8) (2007) 084801. http://dx.doi.org/10.1103/PhysRevLett.98.084801

[63] T.P. Rowlands-Rees *et al.*, *Phys. Rev. Lett.* **100**(10) (2008) 105005. http://dx.doi.org/10.1103/PhysRevLett.100.105005

[64] A. Pak, K.A. Marsh, S.F. Martins, W. Lu, W.B. Mori and C. Joshi, *Phys. Rev. Lett.* **104**(2) (2010) 025003. http://dx.doi.org/10.1103/PhysRevLett.104.025003

[65] C. McGuffey *et al.*, *Phys. Rev. Lett.* **104**(2) (2010) 025004. http://dx.doi.org/10.1103/PhysRevLett.104.025004

[66] B. B. Pollock *et al.*, *Phys. Rev. Lett.* **107**(4) (2011) 045001. http://dx.doi.org/10.1103/PhysRevLett.107.045001

[67] S. Y. Kalmykov *et al.*, "Numerical modelling of a 10-cm-long multi-GeV laser wakefield accelerator driven by a self-guided petawatt pulse" *New J. Phys.* **12** (2010) 045019. http://dx.doi.org/10.1088/1367-2630/12/4/045019

[68] J.B. Rosenzweig, D.B. Cline, B. Cole, H. Figueroa, W. Gai, R. Konecny, J. Norem, P. Schoessow and J. Simpson, *Phys. Rev. Lett.* **61**(1) (1988) 98. http://dx.doi.org/10.1103/PhysRevLett.61.98

[69] I. Blumenfeld, C. E. Clayton, F. J. Decker, M. J. Hogan, C. Huang, R. Ischebeck, R. Iverson, C. Joshi, T. Katsouleas, N. Kirby, W. Lu, K. A. Marsh, W. .B. Mori, P. Muggli, E. Oz, R. .H. Siemann, D. Walz, and M. Zhou, *Nature* (2007) 741. doi:10.1038/nature05538

[70] N. Kumar, A. Pukhov and K. Lotov, *Phys. Rev. Lett.* **104**(25) (2010) 255003. http://dx.doi.org/10.1103/PhysRevLett.104.255003

[71] R. Assmann *et al.*, *Plasma Phys. Control. Fusion* **56**(8) (2014) 084013. http://dx.doi.org/10.1088/0741-3335/56/8/084013

[72] A. Pukhov and T. Tueckmantel, *Phys. Rev. ST Accel. Beams* **15**(11) (2012) 111301. http://dx.doi.org/10.1103/PhysRevSTAB.15.111301

[73] E. Esarey, C. B. Schroeder and W. P. Leemans, *Rev. Mod. Phys.* **81**(3) (2009) 1229. http://dx.doi.org/10.1103/RevModPhys.81.1229

[74] V. Malka, *Phys. Plasmas* **19**(5) (2012) 055501. http://dx.doi.org/10.1063/1.3695389

[75] C. Joshi, *Phys. Plasmas* **14**(5) (2007) 055501. http://dx.doi.org/10.1063/1.2721965

[76] Y. Glinec *et al.*, *Phys. Rev. Lett.* **94**(2) (2005) 025003. http://dx.doi.org/10.1103/PhysRevLett.94.025003

[77] A. Ben-Ismail, O. Lundh, C. Rechatin, J.K. Lim, J. Faure, S. Corde and V. Malka, *Appl. Phys. Lett.* **98**(26) (2011) 264101. http://dx.doi.org/10.1063/1.3604013

[78] Y. Glinec, J. Faure, V. Malka, T. Fuchs, H. Szymanowski and U. Oelfke, *Med. Phys.* **33**(1) (2006) 155. http://dx.doi.org/10.1118/1.2140115

[79] T. Fuchs, H. Szymanowski, U. Oelfke, Y. Glinec, C. Rechatin, J. Faure and V. Malka, *Phys. Med. Biol.* **54**(11) (2009) 3315. http://dx.doi.org/10.1088/0031-9155/54/11/003

[80] B. Brozek-Pluska, D. Gliger, A. Hallou, V. Malka and Y.A. Gauduel, *Radiation Chem.* **72**(2–3) (2005) 149. http://dx.doi.org/10.1016/j.radphyschem.2004.06.014

[81] Y. Gauduel, Y. Glinec, J.-P. Rousseau, F. Burgy and V. Malka, *Eur. Phys. J.* D **60**(1) (2010) pp 121. http://dx.doi.org/10.1140/epjd/e2010-00152-2

[82] V. Malka, J. Faure and Y. A. Gauduel, *Mutation Research* **704**(1-3) (2010) 142. http://dx.doi.org/10.1016/j.mrrev.2010.01.006

[83] O. Rigaud, N.O. Fortunel, P. Vaigot, E. Cadio, M.T. Martin, O. Lundh, J. Faure, C. Rechatin, V. Malka and Y. A. Gauduel, *Cell Death Disease* **1** (2010) e73. doi: 10.1038/cddis.2010.46

[84] V. Malka, J. Faure, Y. A. Gauduel, E. Lefebvre, A. Rousse and K. Ta Phuoc, *Nature Phys.* **4** (2008) 447. http://dx.doi.org/10.1038/nphys966

[85] O. Lundh *et al.. Nature Phys.* **7** (2011) 219. http://dx.doi.org/10.1038/nphys1872

[86] A. Rousse *et al.*, *Phys. Rev. Lett.* **93** (2004) 135005. http://dx.doi.org/10.1103/PhysRevLett.93.135005

[87] Y. Glinec, J. Faure, A. Lifschitz, J.M. Vieira, R.A. Fonseca, L.O. Silva and V. Malka, *Eur. Phys. Lett.* **81**(6) (2008)64001. http://dx.doi.org/10.1209/0295-5075/81/64001

[88] K.T. Phuoc *et al.*, *Phys. Plasmas* **14**(8) (2007) 080701. http://dx.doi.org/10.1063/1.2754624

[89] K. Ta Phuoc, S. Corde, R. Shah, F. Albert, R. Fitour, J.-P. Rousseau, F. Burgy, B. Mercier and A. Rousse, *Phys. Rev. Lett.* **97**(22) (2006) 225002. http://dx.doi.org/10.1103/PhysRevLett.97.225002

[90] S. Fourmaux *et al.*, *Opt. Lett.* **36**(13) (2011) 2426. http://dx.doi.org/10.1364/OL.36.002426

[91] S. Kneip *et al.*, *Appl. Phys. Lett.* **99**(9) (2011) 093701. http://dx.doi.org/10.1063/1.3627216

Introduction to Particle Accelerators and their Limitations

B.J. Holzer
CERN, Geneva, Switzerland

Abstract

The paper gives an overview of the principles of particle accelerators and their historical development. After introducing the basic concepts, the main emphasis is on sketching the layout of modern storage rings and discussing their limitations in terms of energy and machine performance. Examples of existing machines, among them the Large Hadron Collider (LHC) at CERN, demonstrate the basic principles of and the technical and physical limits that we face in the design and operation of particle colliders. The push for ever higher beam energies motivates the design of future colliders as well as the development of more efficient acceleration techniques.

Keywords

Particle accelerators; history; basic concepts; technologies; technical limits; physical limits.

1   Introduction

The study of matter, from initial theories about the structure of the atom to the discovery of the nucleus and, subsequently, of a variety of particles and their interactions, has been summarized in a scientific picture often called the 'standard model'; along its way it has driven the development of powerful tools to create the particle beams that are needed to analyse the detailed structure of matter. The largest accelerator to date, the proton–proton Large Hadron Collider (LHC) at CERN in Geneva, operates with an energy per beam of 7 TeV, which corresponds to an available centre-of-mass energy of $E_{\mathrm{cm}} = 14$ TeV. The LHC is part of a long tradition of technical and physical progress in creating particle beams, accelerating them, and achieving successful collision with micrometre beam sizes. This article gives a basic introduction to the physics of particle accelerators and discusses some of their limitations. The author has arbitrarily selected ten limitations to focus on, although in fact there are many more that the reader may find in other publications and which could be studied further (and some of them overcome, hopefully).

1.1   Limit I: The geography

For the fun of it, let us start at the end of the line: the largest accelerators and the fact that beyond physical and technical limits there is a serious boundary condition—the landscape. For a given technology, pushing the particle energy of a storage ring to higher and higher values will necessitate larger and larger machines, and we may suddenly encounter the problem that our device no longer fits in the garage at our institute or, as shown in Fig. 1, not even in the entire region surrounding our facility. As regards LHC [1], the largest storage ring at present, and the Geneva region, the space between Lake Geneva and the Jura mountains defined the size of the tunnel for the present LHC. As a consequence, the maximum feasible beam energy available for high-energy physics experiments is determined by the geographical boundary conditions of the Geneva countryside. Certainly, there are extremely high particle energies in cosmic rays, but you will agree that the accelerators driving these are also much much larger!

Before we discuss the high-energy frontier machines, let us take a brief look at the path paved by the ingenious scientific developments dating back to the discovery of the nucleus by Ernest Rutherford. Figure 2 (taken from [2]) shows a comparison of the scattering events, plotted as a function of the scattering angle, predicted by the Thomson model of the atom and the experimental results obtained by

Fig. 1: The proton–proton collider LHC at CERN, Geneva
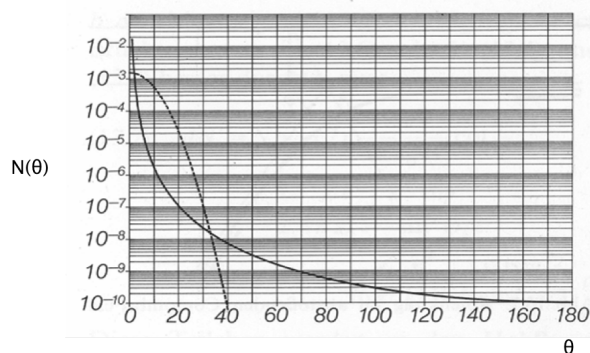
Fig. 2: Comparison of the Thomson model of the atom and the results of Rutherford's experiment; the number of scattering events of alpha particles scattered at the gold foil is plotted as a function of the measured (or predicted) scattering angle.

Rutherford: the discrepancy between the model, which assumes equally distributed charges in matter, and the observed data is evident. As a consequence, the concept of the nucleus was born.

Using alpha particles on the level of MeV is not ideal for precise, triggerable and healthy experiments. So Rutherford discussed with two colleagues, Cockcroft and Walton, the possibility of using artificially accelerated particles. Based on this idea, within only four years Cockcroft and Walton invented the first particle accelerator ever built, and in 1932 they gave the first demonstration of the splitting of a nucleus (lithium) by using a 400 keV proton beam.

Their acceleration mechanism was based on a rectifier or Greinacher circuit, consisting of a number of diodes and capacitors that transformed a relatively small AC voltage to a DC potential which corresponds, depending on the number of diode/capacitor units used, to a multiple of the applied basic potential. The particle source was a standard hydrogen discharge source connected to the high-voltage part of the system, and the particle beam was accelerated to ground potential, hitting the lithium target [3]. A schematic view of the mechanism is shown in Fig. 3 (for details see, for instance, Ref. [4]), and a photograph of such a device which has been used at CERN for many years is presented in Fig. 4. Cockcroft and Walton were awarded the Nobel prize for their invention.

1.2 Limit II: Voltage breakdown in DC accelerators

In parallel to Cockcroft and Walton, but based on a completely different technique, another type of DC accelerator had been invented: Van de Graaff designed a DC accelerator [5] that used a mechanical
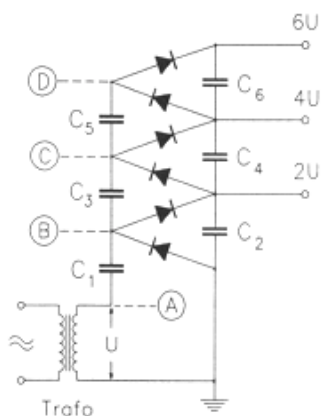
Fig. 3: Schematic layout of the rectifier circuit used by Cockcroft and Walton to generate the high DC voltage needed for their accelerator.

transport system to carry charges, sprayed on a belt or chain, to a high-voltage terminal; see Fig. 5, taken from Ref. [6].

In general these machines can reach higher voltages than the Cockcroft–Walton devices, but they are more limited in terms of particle intensity. Common to all DC accelerators is the limitation on the achievable beam energy due to high-voltage breakdown effects (discharges). Without using an insulating gas (SF_6 in most cases), electric fields will be limited to about $1\,\mathrm{MV\,m^{-1}}$, and even with the most sophisticated devices, like the one in Fig. 6, acceleration voltages on the order of MV cannot be overcome. In fact, the example in Fig. 6 shows an approach that has been applied in a number of situations: injecting a negative ion beam (even H^- is used) and stripping the ions in the middle of the high-voltage terminal allows one to profit from the potential difference twice and thus to make another step of gain in beam energy.

Given the obvious limitations of the DC machines described above, the next step forward is natural. In 1928, Widerøe developed the concept of a AC accelerator. Instead of rectifying the AC voltage, he connected a series of acceleration electrodes in an alternating manner to the output of an AC supply. The schematic layout is shown in Fig. 7 where, for a instant in time, the direction of the electric field is indicated. In principle, this device can produce step by step a multiple of the acceleration voltage, as long as for the negative half-wave of the AC voltage the particles are shielded from the decelerating field.
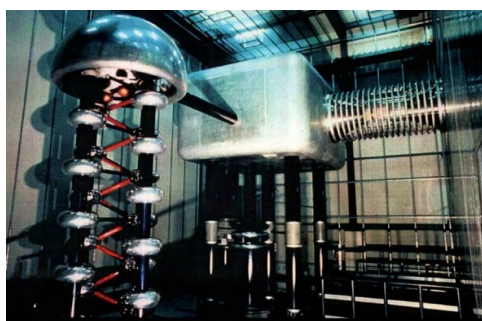
Fig. 4: A Cockcroft–Walton generator that was used at CERN as a pre-accelerator for the proton beams; the device has since been replaced by the more compact and efficient RFQ technique.
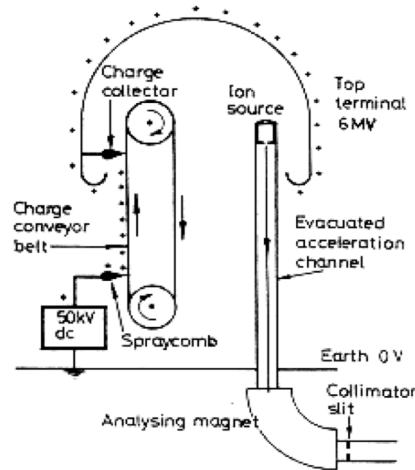
Fig. 5: Schematic design of a Van de Graaff accelerator

Fig. 6: A typical example of a tandem Van de Graaff accelerator; these are very reliable machines for precise measurements in atomic and nuclear physics. (Photo courtesy of the Max Planck Institute for Nuclear Physics, Heidelberg.)

The energy gain after the nth step is therefore

$$E_n = n \cdot q \cdot U_0 \cdot \sin(\psi_s), \tag{1}$$

where n denotes the acceleration step, q the charge of the particle, U_0 the applied voltage per gap, and ψ_s the phase between the particle and the changing AC voltage.

1.3 Limit III: The size of the accelerating structure

A key quantity in such a Wideröe structure is the length of the drift tubes that will protect the particles from the negative half-wave of the sinusoidal AC voltage. For a given frequency of the applied radio-frequency (RF) voltage, the length of the drift tube is defined by the speed of the particle and the duration of the negative half-wave of the sinusoidal voltage, as shown in Fig. 8.

The time-span of the negative half-wave is defined by the applied frequency, $\Delta t = \tau_{\mathrm{rf}}/2$, so for the length of the nth drift tube we get

$$l_n = v_n \cdot \frac{\tau_{\mathrm{rf}}}{2}. \tag{2}$$

Given the kinetic energy of the particle,

$$E_{\mathrm{kin}} = \frac{1}{2}mv^2, \tag{3}$$
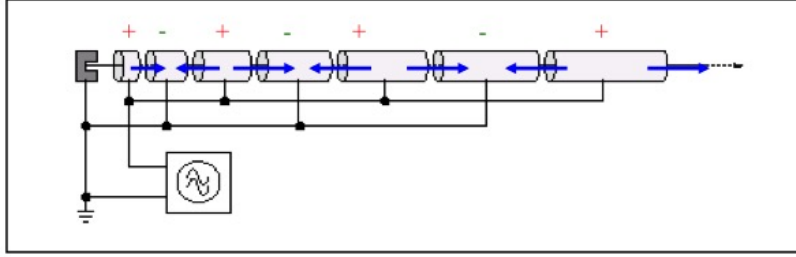
Fig. 7: Schematic view of the Widerøe principle as a fundamental concept for AC (or RF) acceleration
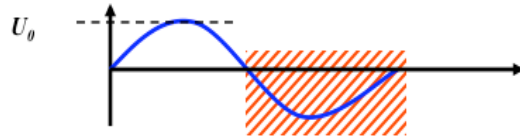
Fig. 8: The frequency (and hence the period) of the RF system and the particle speed determine the length of the drift tubes in the Widerøe structure.
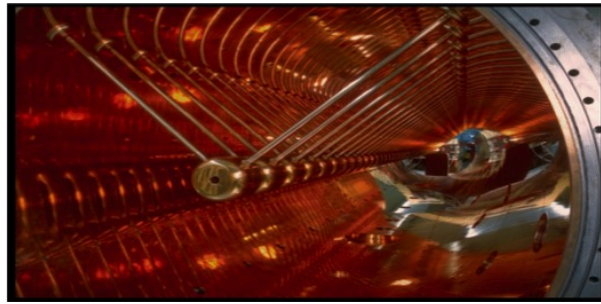
Fig. 9: Unilac at GSI, Darmstadt; clearly visible are the structure of the drift tubes and their increasing length as a function of the particle energy.

we obtain directly that

$$l_n = \frac{1}{\nu_{\mathrm{rf}}} \cdot \sqrt{\frac{nqU_0 \sin \psi_s}{2m}}, \tag{4}$$

which defines the design concept of the machine. Figure 9 shows a photograph of such a device, the Unilac at the Institute for Heavy Ion Research (GSI) in Darmstadt, Germany.

Two remarks should be made in this context.

- The short derivation here is based on the classical approach, and in fact these accelerators are usually optimum for 'low-energetic' proton or heavy-ion beams. Typical beam energies (referring to protons) are on the order of 10 MeV; for example, the present Linac 2 at CERN delivers the protons for LHC operation with an energy of 50 MeV, corresponding to a relativistic β of 0.31.
- For higher energies, even in the case of protons or ions, the speed will at some point approach the speed of light, and the length of the drift tubes and hence the dimension of the whole accelerator will reach a size that may no longer be feasible. More advanced ideas are needed in order to keep the machine within reasonable dimensions, and the next natural step in the historical development was to introduce magnetic fields and bend the particle beam into a circle.

2 Pushing for the highest energies: synchrotrons and storage rings

A significant step forward in achieving high beam energies involves the use of circular structures. In order to apply over and over again the accelerating fields, we try to bend the particles onto a circular path and so bring them back to the RF structure where they will receive the next step-up in energy. To do this, we introduce magnetic (or electric) fields that will deflect the particles and keep them on a well-defined orbit during the complete acceleration process. The Lorentz force that acts on a particle will therefore have to compensate exactly the centrifugal force due to the bent orbit. In general, we can write

$$\mathbf{F} = q \cdot (\mathbf{E} + \mathbf{v} \times \mathbf{B}). \tag{5}$$

For high-energy particle beams, the velocity v is close to the speed of light and so represents a nice amplification factor whenever we apply a magnetic field. As a consequence, it is much more convenient to use magnetic fields for bending and focusing the particles.

Therefore, neglecting electric fields for the moment, we write the Lorentz force and the centrifugal force of the particle on its circular path as

$$F_{\text{Lorentz}} = e \cdot v \cdot B, \tag{6}$$

$$F_{\text{centrifugal}} = \frac{\gamma m_0 v^2}{\rho}. \tag{7}$$

Assuming an idealized homogeneous dipole oriented along the particle orbit, we define the condition for a perfect circular orbit as equality between these two forces; this yields the following condition for the idealized ring:

$$\frac{p}{e} = B \cdot \rho, \tag{8}$$

where we refer to protons and have accordingly set $q = e$. This condition relates the so-called beam rigidity $B\rho$ to the particle momentum that can be carried in the storage ring, and it will ultimately define, for a given magnetic field of the dipoles, the size of the storage ring.

In reality, instead of a continuous dipole field the storage ring will be built out of several dipoles, powered in series to define the geometry of the ring. For a single magnet, the particle trajectory is shown schematically in Fig. 10. In the free space outside the dipole magnet, the particle trajectory follows a straight line. As soon as the particle enters the magnet, it is bent onto a circular path until it leaves the magnet at the other side.
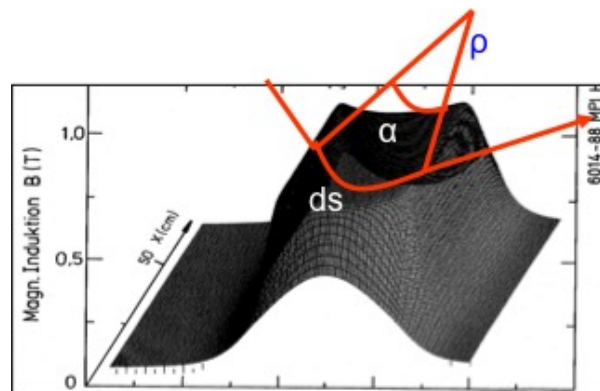
Fig. 10: Dipole field of a storage ring and schematic path of the particles

2.1 Limit IV: The magnetic guide field

The overall effect of the main bending (or 'dipole') magnets in the ring is to define a more or less circular path, which we will call the 'design orbit'. By definition, this design orbit has to be a closed loop, and so the main dipole magnets in the ring have to define a bending angle of exactly 2π overall. If α denotes the bending angle of a single magnet, then

$$\alpha = \frac{\mathrm{d}s}{\rho} = \frac{B\,\mathrm{d}s}{B \cdot \rho}. \tag{9}$$

We therefore require that

$$\frac{\int B\,\mathrm{d}l}{B \cdot \rho} = 2\pi. \tag{10}$$

Thus, a storage ring is not a 'ring' in the true sense of the word but more a polygon, where 'poly' means the discrete number of dipole magnets installed in the 'ring'.

In the case of the LHC, the dipole field has been pushed to the highest achievable values; 1232 superconducting dipole magnets, each of length 15 m, define the geometry of the ring and, via Eq. (10), the maximum momentum for the stored proton beam. Using the equation given above, for a maximum momentum of $p = 7$ TeV/c we obtain a required magnetic field of

$$B = \frac{2\pi \cdot 7000 \cdot 10^9 \text{ eV}}{1232 \cdot 15 \text{ m} \cdot 2.99792 \cdot 10^8 \text{ m s}^{-1}}, \tag{11}$$

or

$$B = 8.33T, \tag{12}$$

to bend the beams. For convenience we have expressed the particle momentum in units of GeV/c. Figure 11 shows a photograph of the LHC dipole magnets, built out of superconducting NbTi filaments, which are operated at a temperature of $T = 1.9$ K.

Fig. 11: Superconducting quadrupole of the LHC storage ring

2.2 Focusing properties

In addition to the main bending magnets that guide the beam onto a closed orbit, focusing fields are needed to keep the particles close together. In modern storage rings and light sources, the particles are kept in the machine for many hours, and a carefully designed focusing structure is needed to maintain the necessary beam size at different locations in the ring.

Following classical mechanics, linear restoring forces are needed, just as in the case of a harmonic pendulum. Quadrupole magnets provide the corresponding property: they create a magnetic field that depends linearly on the particle amplitude, i.e. the distance of the particle from the design orbit:

$$B_x = g \cdot y, \qquad B_y = g \cdot x. \tag{13}$$

The constant g is called the gradient of the magnetic field and characterizes the focusing strength of the magnetic lens in both transverse planes. For convenience it is (like the dipole field) normalized to the particle momentum. The normalized gradient is denoted by k and defined as

$$k = \frac{g}{p/e} = \frac{g}{B\rho}. \tag{14}$$

The technical layout of such a quadrupole is depicted in Fig. 11. As in the case of the LHC dipoles, the quadrupole magnet is built in superconducting technology.

Now that we have defined the basic building blocks of a storage ring, we need to arrange them in a so-called magnet lattice and optimize the field strengths in such a way as to obtain the required beam parameters. An example of what such a magnet lattice looks like is given in Fig. 12. This photograph shows the dipole (orange) and quadrupole (red) magnets in the TSR storage ring in Heidelberg. Eight dipoles are used to bend the beam in a 'circle', and the quadrupole lenses between them provide the focusing to keep the particles within the aperture limits of the vacuum chamber.
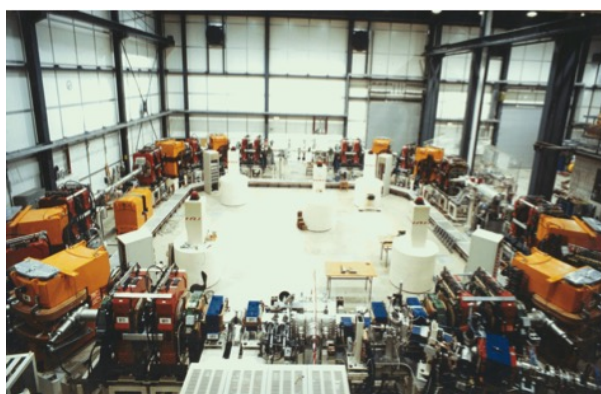
Fig. 12: TSR storage ring, Heidelberg, as a typical example of a separate-function strong focusing storage ring

A general design principle of modern synchrotrons or storage rings should be pointed out here. In general, these machines are built following a so-called separate-function scheme: every magnet is designed and optimized for a certain task, such as bending, focusing, chromatic correction, and so on. We separate the magnets in the design according to the job they are supposed to do; only in rare cases a combined-function scheme is chosen, where different magnet properties are combined in one piece of hardware. To express this mathematically, we use the general Taylor expansion of the magnetic field,

$$\frac{B(x)}{p/e} = \frac{1}{\rho} + k \cdot x + \frac{1}{2!}mx^2 + \frac{1}{3!}nx^3 + \cdots. \tag{15}$$

Following the arguments above, for the moment we take into account only constant (dipole) or linear (quadrupole) terms. The higher-order field contributions will be treated later as (hopefully) small perturbations.

The particles will now follow the 'circular' path defined by the dipole fields, and in addition will undergo harmonic oscillations in both transverse planes. The situation is shown schematically in Fig. 13. An ideal particle will follow the design orbit that is represented by the circle in the diagram. Any other particle will perform transverse oscillations under the influence of the external focusing fields, and the amplitude of these oscillations will ultimately define the beam size.

Unlike a classical harmonic oscillator, however, the equations of motion in the horizontal and vertical planes differ somewhat. Assuming a horizontal focusing magnet, the equation of motion is

$$x'' + x \cdot \left(\frac{1}{\rho^2} + k\right) = 0, \tag{16}$$
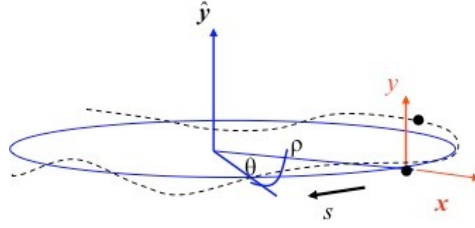
Fig. 13: Coordinate system used in particle beam dynamics; the longitudinal coordinate s moves around the ring with the particle considered.

where k is the normalized gradient introduced above and the $1/\rho^2$ term represents the so-called weak focusing, which is a property of the bending magnets. In the vertical plane, on the other hand, due to the orientation of the field lines and by Maxwell's equations, the forces instead have a defocusing effect; also, the weak focusing term disappears:

$$y'' - y \cdot k = 0. \tag{17}$$

The principal problem arising from the different directions of the Lorentz force in the two transverse planes of a quadrupole field is sketched in Fig. 14. It is the task of the machine designer to find an adequate solution to this problem and to define a magnet pattern that will provide an overall focusing effect in both transverse planes.
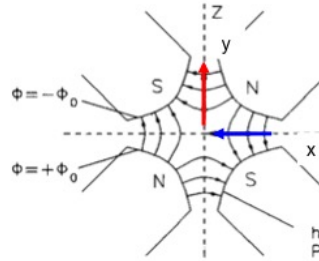
Fig. 14: Field configuration in a quadrupole magnet and the direction of the focusing and defocusing forces in the horizontal and vertical planes.

Following closely the example of the classical harmonic oscillator, we can write down the solutions of the above equations of motion. For simplicity, we focus on the horizontal plane; a 'focusing' magnet is therefore focusing in this horizontal plane and at the same time defocusing in the vertical plane. Starting with initial conditions for the particle amplitude x_0 and angle x_0' in front of the magnet element, we obtain the following relations for the trajectory inside the magnet:

$$x(s) = x_0 \cdot \cos\left(\sqrt{|K|}\,s\right) + x_0' \cdot \frac{1}{\sqrt{|K|}} \sin\left(\sqrt{|K|}\,s\right), \tag{18}$$

$$x'(s) = -x_0 \cdot \sqrt{|K|} \sin\left(\sqrt{|K|}\,s\right) + x_0' \cdot \cos\left(\sqrt{|K|}\,s\right). \tag{19}$$

Here the parameter K combines the quadrupole gradient and the weak focusing effect, $K = k - \frac{1}{\rho^2}$. Usually these two equations are combined into a more elegant and convenient matrix form:

$$\begin{pmatrix} x \\ x' \end{pmatrix}_s = \mathbf{M}_{\text{foc}} \begin{pmatrix} x \\ x' \end{pmatrix}_0, \tag{20}$$

where the matrix $\mathbf{M}_{\text{foc}}$ contains all the relevant information about the magnet element,

$$\mathbf{M}_{\text{foc}} = \begin{pmatrix} \cos(\sqrt{|K|}\,s) & \frac{1}{\sqrt{|K|}}\sin(\sqrt{|K|}\,s) \\ -\sqrt{|K|}\sin(\sqrt{|K|}\,s) & \cos(\sqrt{|K|}\,s) \end{pmatrix}.$$

Schematically, the situation is visualized in Fig. 15.
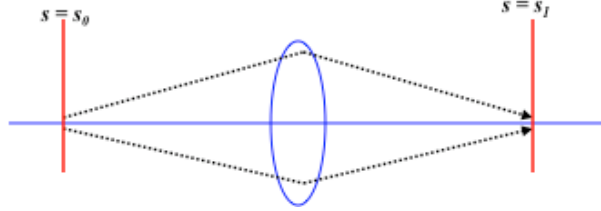
Fig. 15: Schematic principle of the effect of a focusing quadrupole magnet

In the case of a defocusing magnet, we obtain analogously that

$$\begin{pmatrix} x \\ x' \end{pmatrix}_s = \mathbf{M}_{\text{defoc}} \begin{pmatrix} x \\ x' \end{pmatrix}_0, \tag{21}$$

with

$$\mathbf{M}_{\text{defoc}} = \begin{pmatrix} \cosh(\sqrt{|K|}\,s) & \frac{1}{\sqrt{|K|}}\sinh(\sqrt{|K|}\,s) \\ \sqrt{|K|}\sinh(\sqrt{|K|}\,s) & \cosh(\sqrt{|K|}\,s) \end{pmatrix};$$

see Fig. 16.
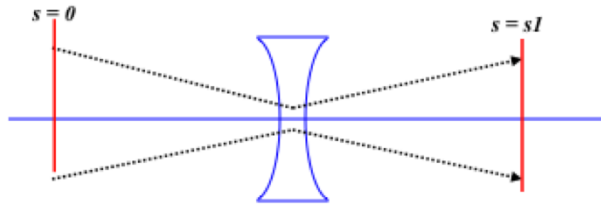
Fig. 16: Schematic principle of the effect of a defocusing quadrupole magnet

For completeness, we also include the case of a field-free drift. With $K = 0$, we obtain

$$\mathbf{M}_{\text{drift}} = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}.$$

This matrix formalism allows us to combine the elements of a storage ring in an elegant way and so it is straightforward to calculate the particle trajectories. As an example, we consider the simple case of an alternating focusing and defocusing lattice, a so-called FODO lattice [4]; see Fig. 17.

As we know the properties of each and every element in the accelerator, we can construct the corresponding matrices and calculate step by step the amplitude and angle of a single particle's trajectory around the ring. Even more conveniently, we can multiply out the different matrices and, given initial conditions x_0 and x'_0, obtain directly the trajectory at any location in the ring:

$$\mathbf{M}_{\text{total}} = \mathbf{M}_{\text{foc}} \cdot \mathbf{M}_{\text{drift}} \cdot \mathbf{M}_{\text{dipole}} \cdot \mathbf{M}_{\text{drift}} \cdot \mathbf{M}_{\text{defoc}} \cdots . \tag{22}$$
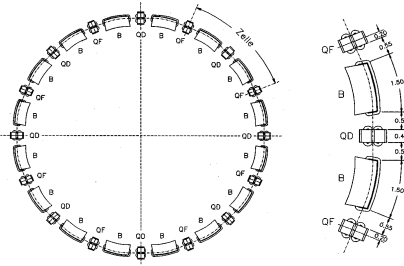
Fig. 17: A simple periodic chain of bending magnets and focusing/defocusing quadrupoles forming the basic structure of a storage ring.
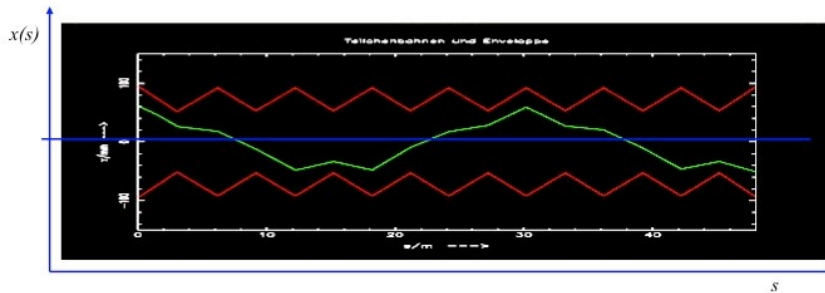
Fig. 18: Calculated particle trajectory in a simple storage ring

The trajectory thus obtained is shown schematically in Fig. 18.

We emphasize the following facts in this context.

– At each moment, or in each lattice element, the trajectory is a part of a harmonic oscillation.
– However, due to the different restoring or defocusing forces, the solution will look different at each location.
– In the linear approximation that we make use of in this context, all particles experience the same external fields, and their trajectories will differ only because of their different initial conditions.
– There seems to be an overall oscillation in both transverse planes while the particle is travelling around the ring. Its amplitude stays well within the boundaries set by the vacuum chamber, and its frequency in the example of Fig. 18 is roughly 1.4 transverse oscillations per revolution, which corresponds to the eigenfrequency of the particle under the influence of the external fields.

Coming closer to a real existing machine, we show in Fig. 19 the orbit measured during one of the first injections into the LHC storage ring. The horizontal oscillations are plotted in the upper half of the figure and the vertical oscillations in the lower half, on a scale of ± 10 mm. Each histogram bar indicates the value recorded by a beam position monitor at a certain location in the ring, and the orbit oscillations are clearly visible. By counting (or, better, fitting) the number of oscillations in both transverse planes, we obtain values of

$$Q_x = 64.31, \qquad Q_y = 59.32. \tag{23}$$

These values, which describe the eigenfrequencies of the particles, are called the horizontal and vertical *tune*, respectively. Knowing the revolution frequency, we can easily calculate the transverse oscillation frequencies, which for this type of machine usually lie in the range of kHz.

As the tune characterizes the particle oscillations under the influence of all external fields, it is one of the most important parameters of the storage ring. Therefore it is usually displayed and controlled at
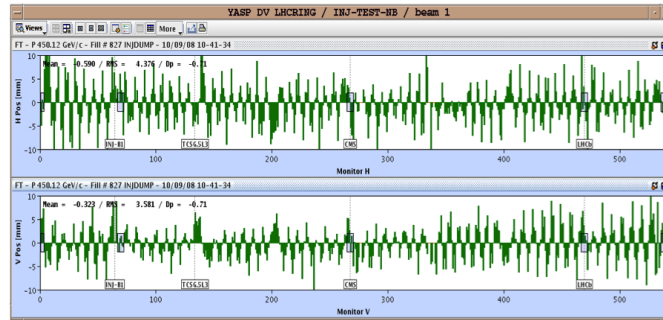
Fig. 19: Measured orbit in LHC during the commissioning of the machine

all times by the control system of such a machine. As an example, Fig. 20 shows the tune diagram of the HERA proton ring [7]; it was obtained via a Fourier analysis of the spectrum measured from the signal of the complete particle ensemble. The peaks indicate the two tunes in the horizontal and vertical planes of the machine, and in a sufficiently linear machine a fairly narrow spectrum is obtained.

Fig. 20: Tune signal of a proton storage ring (HERA-p)

Briefly referring back to Fig. 18, the question is what the trajectory of the particle will look like for the second turn, or the third, or after an arbitrary number of turns. Now, as we are dealing with a circular machine, the amplitude and angle, x and x', at the end of the first turn will be the initial conditions for the second turn, and so on. After many turns the overlapping trajectories begin to form a pattern, such as that in Fig. 21, which indeed looks like a beam having here and there a larger and smaller beam size but still remaining well-defined in its amplitude by the external focusing forces.
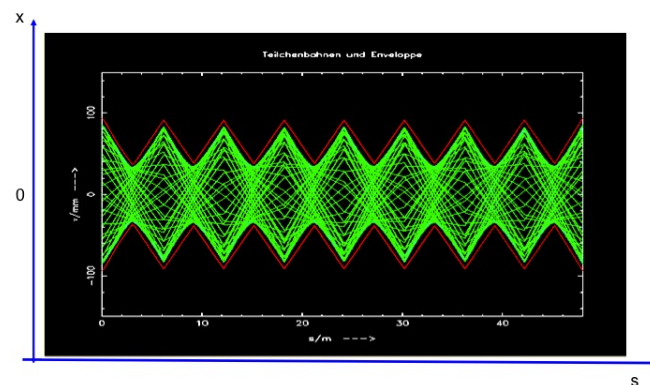
Fig. 21: Many single-particle trajectories together form a pattern that corresponds to the beam size in the ring

To make a long story short [8], a mathematical function, which we call β or amplitude function, can be defined that describes the envelope of the single-particle trajectories. With this new variable, we can rewrite the equation for the amplitude of a particle's transverse oscillations as

$$x(s) = \sqrt{\epsilon}\sqrt{\beta(s)}\cos(\psi(s) + \phi), \tag{24}$$

where ψ is the phase of the oscillation, ϕ is its initial condition, and ϵ is a characteristic parameter of the single particle or, if we are considering a complete beam, of the ensemble of particles. Indeed, ϵ describes the space occupied by the particle in the transverse (here simplified two-dimensional) (x, x') phase space. More specifically, the area in (x, x') space that is covered by the particle is given by

$$A = \pi \cdot \epsilon, \tag{25}$$

and, as long as we consider conservative forces acting on the particle, this area is constant according to Liouville's theorem. Here we take these facts as given, but we point out that, as a direct consequence, the so-called emittance ϵ cannot be influenced by whatever external fields are applied; it is a property of the beam, and we have to take it as given and handle it with care.

To be more precise, and following the usual textbook treatment of accelerators, we can draw in phase space the ellipse of the particle's transverse motion; see, for example, Fig. 22. While the shape and orientation are determined by the optics function β and its derivative, $\alpha = -\frac{1}{2}\beta'$, the area covered is constant.
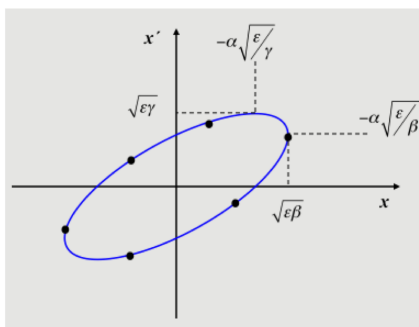
Fig. 22: Ellipse in (x, x') phase space

Let us talk a bit more about the beam as an ensemble of many (typically 10^{11}) particles. Referring to Eq. (24), at a given position in the ring the beam size is defined by the emittance ϵ and the function β. Thus, at a certain moment in time the cosine term in (24) will be 1 and the trajectory amplitude will reach its maximum value. Now, if we consider a particle at one standard deviation (sigma) of the transverse density distribution, then using the emittance of this reference particle we can calculate the size of the complete beam, in the sense that the complete area (within one sigma) of all particles in (x, x') phase space is surrounded (and thus defined) by our one-sigma candidate. Thus the value $\sqrt{\epsilon \cdot \beta(s)}$ will define the one-sigma beam size in the transverse plane. As an example, we use the values for the LHC proton beam: in the periodic pattern of the arc, the beta function is $\beta = 180$ m and the emittance at flat-top energy is roughly $\epsilon = 5 \times 10^{-10}$ rad m. The resulting typical beam size is therefore 0.3 mm. Now, clearly we would not design a vacuum aperture of the machine based on a one-sigma beam size; typically, an aperture requirement corresponding to 12σ is a good rule to guarantee a sufficient aperture, allowing for tolerances from magnet misalignments, optics errors and operational flexibility. In Fig. 23 the LHC vacuum chamber is shown, including the beam screen used to protect the cold bore from synchrotron radiation; it corresponds to a minimum beam size of 18σ.
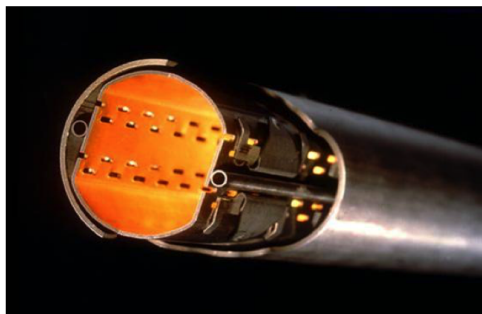
Fig. 23: The LHC vacuum chamber with the beam screen to shield the superconducting magnet bore from synchrotron radiation.

3 Particle colliders

3.1 Limit V: Fixed-target collider

The easiest way to perform physics experiments with particle accelerators is to bang the accelerated beam onto a target and analyse the resulting events. While nowadays in high-energy physics we do not apply this technique any more, it still plays an essential role in the regime of atomic and nuclear physics experiments. The advantage is that it is quite simple once the accelerator has been designed and built, and the particles produced are easily separated due to the kinematics of the reaction. The situation is illustrated in Fig. 24. The particle 'a' that is produced and accelerated in the machine is directed onto the particle 'b', which is at rest in the laboratory frame. The particles produced from this collision are labelled 'c' and 'd' in this example.
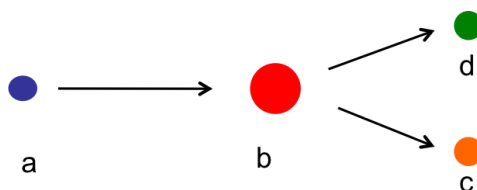
Fig. 24: Schematic diagram of fixed-target collider

While the set-up of such a scheme is quite simple, it is worth taking a closer look at the available energy in the centre-of-mass system. The relativistic overall energy is given by

$$E^2 = p^2 c^2 + m^2 c^4, \tag{26}$$

which holds for a single particle but is equally valid for an ensemble of particles. Most important, the rest energy of the particle ensemble is constant (and is sometimes called the 'invariant mass of the system').

Considering the system of two particles colliding, we can write

$$(E_a^{cm} + E_b^{cm})^2 - (p_a^{cm} + p_b^{cm})^2 c^2 = (E_a^{lab} + E_b^{lab})^2 - (p_a^{lab} + p_b^{lab})^2 c^2. \tag{27}$$

In the frame of the centre-of-mass system we get, by definition,

$$p_a^{cm} + p_b^{cm} = 0, \tag{28}$$

while in the laboratory frame where particle 'b' is at rest we have simply

$$p_b^{lab} = 0. \tag{29}$$

The equation for the invariant mass therefore simplifies to

$$W^2 = (E_a^{cm} + E_b^{cm})^2 = (E_a^{lab} + m_b \cdot c^2)^2 - (p_a^{lab} \cdot c^2). \tag{30}$$

In other words, the energy that is available in the centre-of-mass system depends on the square root of the energy of particle 'a', which is the energy provided by the particle accelerator:

$$W \approx \sqrt{2E_a^{lab} \cdot m_b \cdot c^2} \tag{31}$$

—a quite unsatisfactory situation!

To meet the demand for higher and higher energies in particle collisions, the design of modern high-energy accelerators has naturally concentrated on the development of particle colliders, where two counter-rotating beams are brought into collision at one or several interaction points (Fig. 25).

Fig. 25: Schematic diagram of the collision of two particles with equal energy

If we calculate the available energy in the centre-of-mass system for the case of two colliding beams of identical particles, we get

$$(p_a^{cm} + p_b^{cm})^2 = 0 \tag{32}$$

and, by symmetry, also

$$(p_a^{lab} + p_b^{lab})^2 = 0. \tag{33}$$

So the full energy delivered to the particles in the accelerator is available during the collision process:

$$W = E_a^{lab} + E_b^{lab} = 2\,E_a^{lab}. \tag{34}$$

A 'typical' example of a high-energy physics event in such a collider is shown in Fig. 26.
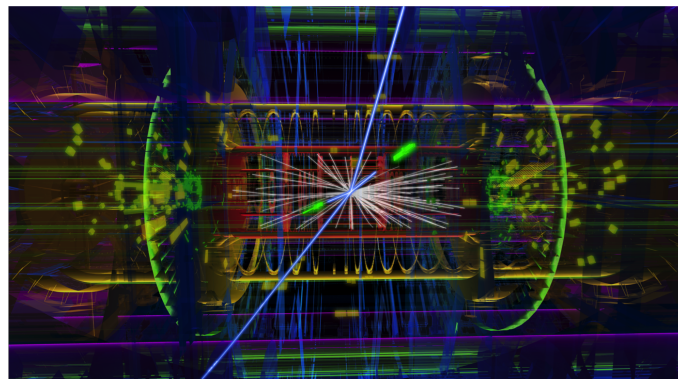
Fig. 26: 'Typical' event observed in a collider ring—a Higgs particle measured in the ATLAS detector

3.2   Limit VI: The unavoidable particle detectors

While it is quite clear that a particle collider ring is a magnificent machine in the quest for higher energies, there is a small problem involved, namely the 'particle detector'. In the arc of the storage ring we can usually find a nice pattern of magnets providing us with a well-defined beam size, expressed as the beta

function. However, special care has to be taken when our colleagues from high-energy physics wish to install a particle detector. Especially when working at the energy frontier, just like for the accelerators, these devices tend to expand considerably in size with the energies required. In Fig. 27, the largest particle detector installed in a storage ring is shown as an impressive example: the ATLAS detector at the LHC.
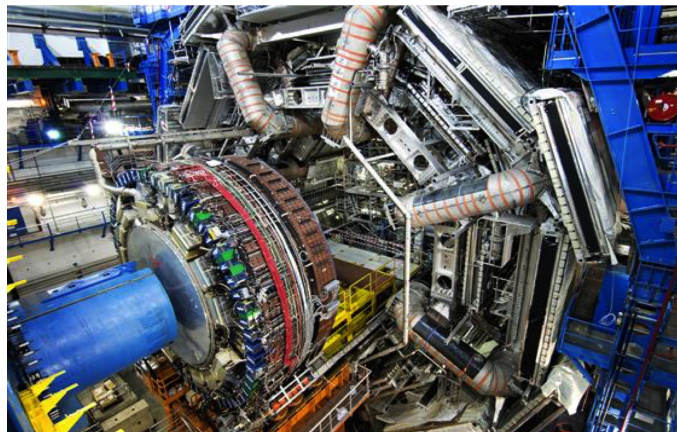
Fig. 27: ATLAS detector at the LHC, which is 46 m in length and has an overall weight of 7000 t

The storage ring has to be designed to provide the space needed for the detector hardware and at the same time create the smallest achievable beam spots at the collision point, which is usually right in the centre of the detector. Unfortunately these requirements are a bit contradictory. The equation for the luminosity of a particle collider depends on the stored beam currents and the transverse spot size of the colliding beams at the interaction point (IP):

$$L = \frac{1}{4\pi e^2 f_0 b} \cdot \frac{I_1 I_2}{\sigma_x^* \sigma_y^*}. \tag{35}$$

At the same time, however, the beta function in a symmetric drift grows quadratically as a function of the distance between the beam waist and the first focusing element, i.e.

$$\beta(s) = \beta^* + \frac{s^2}{\beta^*}. \tag{36}$$

The smaller the beam at the IP, the faster it will grow until we can apply—outside of the detector region— the first quadrupole lenses. As a consequence, this trend sets critical limits on the achievable quadrupole aperture or, for a given aperture, the achievable quadrupole gradient. The focusing lenses right before and after the IP, being placed as closed as possible to the detector, are generally the most critical and most expensive magnets in the machine, and their aperture requirement ultimately determines the luminosity that can be delivered by the storage ring.

For the experts we would like to add that even if the bare aperture requirement can be fulfilled, the resulting chromaticity that is created in the mini-beta insertion and the sextupole strengths that are needed to correct for it usually pose the next limit that we will face.

3.3 Limit VII: The relative rareness of Nobel prize-winning reactions

The rate of events produced in a particle collision process depends not only on the performance of the colliding beams but first and foremost on the probability of creating such an event, the so-called cross-section of the process. In the case of the Higgs particle, which is without doubt the highlight of LHC Run 1, the overall cross-section is displayed in Fig. 28.
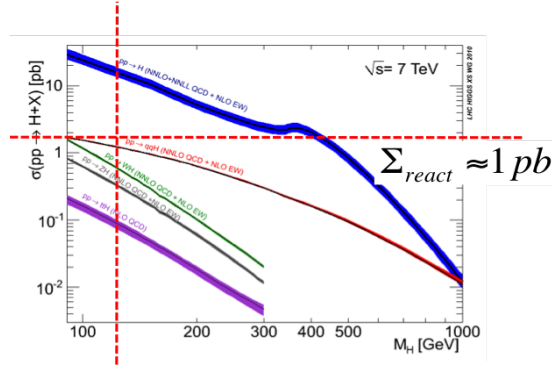
Fig. 28: Cross-section of the Higgs particle for different production processes, courtesy of the CMS collaboration

Without going into details, we can state that the cross-section for Higgs production is on the order of

$$\Sigma_{\text{react}} \simeq 1 \text{ pb.} \tag{37}$$

During the three years of LHC Run 1, i.e. the period 2011–2013, an overall luminosity of

$$\int L \, dt = 25 \text{ fb}^{-1} \tag{38}$$

was accumulated.

Combining these two numbers using the fact that the event rate of a reaction is $R = L \cdot \Sigma_{\text{react}}$, we get a total number of 'some thousand' Higgs particles produced—for a Nobel prize-winning event just at the edge of reliable statistics. Therefore, the particle colliders have to be optimized not only for the highest achievable energies but also for maximum stored beam currents and small spot sizes at the interaction points so as to optimize the luminosity of the machine.

3.4   Limit VIII: The luminosity of a collider ring

Following the arguments above, the design goal here is to prepare, accelerate and store two counter-rotating particle beams in order to profit best from the energy of the two beams during the collision process. Still, there is a price to pay: unlike in fixed-target experiments, where the 'particle' density of the target material is extremely high, in the case of two colliding beams the event rate is basically determined by the transverse particle density that can be achieved at the IP. Assuming Gaussian density distributions in both transverse planes, the performance of such a collider is described by the luminosity

$$L = \frac{1}{4\pi e^2 f_0 b} \cdot \frac{I_1 I_2}{\sigma_x^* \sigma_y^*}. \tag{39}$$

While the revolution frequency f_0 and the bunch number b are ultimately determined by the size of the machine, the stored beam currents I_1 and I_2 and the beam sizes σ_x^* and σ_y^* at the IP have their own limitations.

The most serious limitation comes from the beam–beam interaction itself. During the collision process, individual particles of the counter-rotating bunches feel the space charge of the opposing bunch. In the case of a proton–proton collider, this strong field acts like a defocusing lens, and has a strong impact on the tune of the bunches [9].

In Fig. 29 the situation is shown schematically. Two bunch trains collide at the IP, and during the collision process a direct beam–beam effect is observed. In addition to that, before and after the actual collision, long-range forces exist between the bunches that have a nonlinear component; see Fig. 30.

As a consequence [9], the tune of the beams is not only shifted with respect to the natural tune of the machine but also spread out, as different particles inside the bunches see different contributions from the beam–beam interaction.
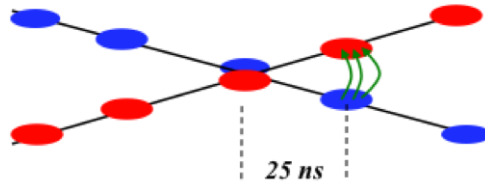
Fig. 29: Schematic view of the beam–beam interaction during the crossing of bunch trains
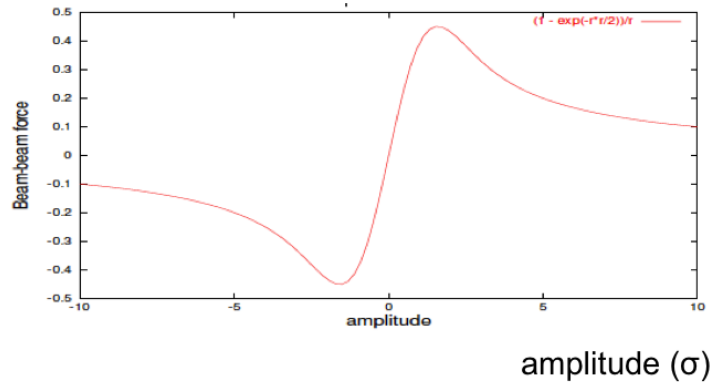
Fig. 30: Beam–beam force as a function of the transverse distance of the particle to the centre of the opposing bunch.

Therefore, in the tune diagram, we no longer obtain a single spot representing the ensemble of particles, but rather a large array that depends in shape, size and orientation on the particle densities, the distance of the bunches at the long-range encounters, and the single-bunch intensities. The effect has been calculated for the LHC and is displayed in Fig. 31.

In a number of cases a useful approximation can be applied, as for distances of about 1–2σ, the beam–beam force in Fig. 30 can be linearized and acts like a quadrupole lens. Accordingly, a tune shift can be calculated to characterize the strength of the beam–beam effect in a collider. Given the parameters described above, and introducing the classical particle radius r_p, the amplitude function β^* at the IP and the Lorentz factor γ, we can express the tune shift due to the linearized beam–beam effect as

$$\Delta Q_y = \frac{\beta_y^* \cdot r_\mathrm{p} \cdot N_\mathrm{p}}{2\pi\,\gamma(\sigma_x + \sigma_y)\sigma_y}. \tag{40}$$

In the case of the LHC, the design value of the beam–beam tune shift is $\Delta Q = 0.0033$, and in daily operation the machine is optimized to run close to this value, which places the ultimate limit on achievable bunch intensities in the collider.

4   Lepton colliders

4.1   Limit IX: Synchrotron light—the drawback of electron storage rings

In proton or heavy-ion storage rings, the design can more or less follow the rules discussed above. But the situation changes drastically as the particles become more and more relativistic. Bent on a circular
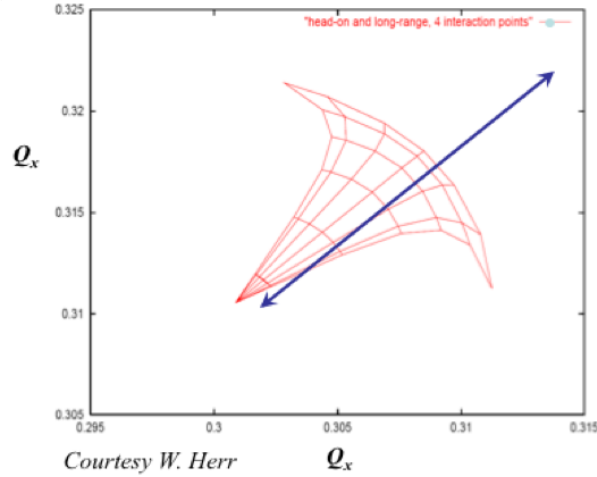
Fig. 31: Calculated tune shift due to beam–beam interaction in the LHC

path, electrons in particular will radiate an intense light, the so-called synchrotron radiation, which will have a strong influence on the beam parameters as well as on the design of the machine.

Summarizing the situation briefly here, the power loss due to synchrotron radiation depends on the bending radius and the energy of the particle beam:

$$P_\text{s} = \frac{2}{3}\alpha\hbar c^2 \frac{\gamma^4}{\rho^2}, \tag{41}$$

where α represents the fine structure constant and ρ the bending radius in the dipole magnets of the ring. As a consequence, the particles will lose energy turn by turn. To compensate for these losses, RF power has to be supplied to the beam at any moment. An example that illustrates the problem nicely is shown in Fig. 32. It plots the horizontal orbit of the former Large Electron–Positron Collider (LEP) storage ring. The electrons, travelling from right to left in the plot, lose a considerable amount of energy in each arc and hence deviate from the ideal orbit towards the inner side of the ring. The effect on the orbit is large: up to 5 mm orbit deviation was observed in the example of Fig. 32. In order to compensate for these losses, four RF stations were installed in the straight sections of the ring to supply the necessary power.
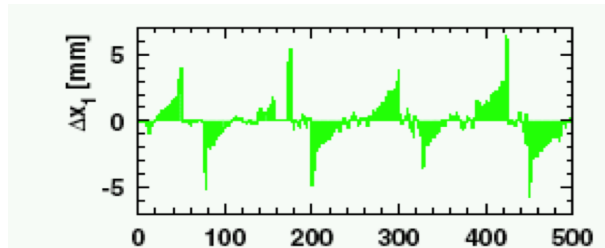
Fig. 32: Measured horizontal orbit of the LEP electron beam; due to synchrotron radiation losses, the particle orbit is shifted towards the inner side of the ring in each arc.

The strong dependence of the synchrotron radiation losses on the relativistic γ factor sets severe limits on the beam energy that can be carried in a storage ring of a given size. The push for ever higher energies means either that storage rings even larger than LEP need to be designed or, to avoid synchrotron radiation, linear accelerating structures should be developed.

Currently, the next generation of particle colliders is being studied [10]. The ring design of the future circular collider (FCC) foresees a 100 km ring to carry electrons (and positrons) of up to 175 GeV energy. The size of this storage ring is far beyond the dimensions of anything that has been designed up to now. A sketch of the machine layout is given in Fig. 33, where the yellow dashed circle delineates the 100 km ring and the white circle represents the little LHC machine.
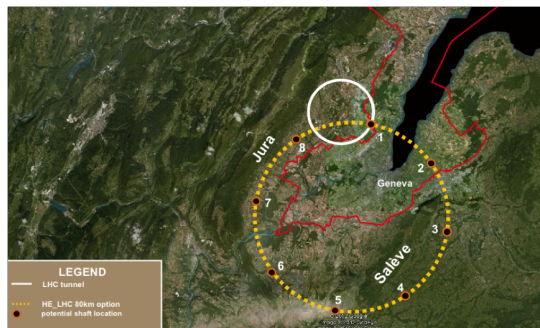
Fig. 33: Schematic view of a possible 100 km FCC design in the Geneva region

For the maximum projected electron energy of $E = 175$ GeV, synchrotron radiation would cause an energy loss of 8.6 GeV, or an overall power of 47 MW of the radiated light at full beam intensity.

4.2 Limit X: Acceleration gradients in linear structures

As far as lepton beams are concerned, ring colliders suffer from the severe limitation caused by synchrotron radiation losses, and at a certain point the construction of such large facilities would not seem reasonable any more. To avoid the problem of synchrotron radiation, linear structures that were discussed earlier and used in the infancy of particle accelerators have become in vogue again. Still, the advantage of circular colliders cannot be completely ignored: even with a modest acceleration gradient in the RF structures, the particles will get turn by turn a certain boost in energy and will at some point reach the desired flat-top energy in the ring.

In a linear accelerator, this kind of repetitive acceleration is by design not possible; within a single pass through the machine, the particles will have to be accelerated to full energy. In order to keep the structure compact, the highest acceleration gradients will therefore be needed. One of the most prominent designs proposed for a possible future collider is the CLIC design [11]. Within one passage through the 25 km long accelerator, the electron beam will get up to 3 TeV, and the same is true for the opposing positron beam. An artist's rendering of this machine is shown in Fig. 34.

Fig. 34: Proposed location of the CLIC linear collider along the Jura mountains in the Geneva region

The main parameters of the CLIC design are listed in Table 1. The accelerating gradient, i.e. the energy gain per meter, is especially to be emphasized; it has been pushed to the maximum value

that is technically feasible, and the limit is ultimately due to the breakdown of the electric field in the accelerating structure.

Table 1: Main parameters of the CLIC study

	500 GeV	**3 TeV**
Site length	13 km	48 km
Loaded acceleration gradient (MV m^{-1})	12	
Beam power per beam (MW)	4.9	14
Bunch charge (10^9 e+/e)	6.8	3.7
Horizontal/vertical normalized emittance ($10^{-6}/10^{-9}$ m)	2.4/25	0.66/20
Beta function (mm)	10/0.07	
Beam size at IP: horizontal/vertical (nm)	45/1	
Luminosity (cm^{-2} s^{-1})	2.3×10^{34}	5.9×10^{34}

A picture of such a CLIC-type structure is shown in Fig. 35. On the right-hand side is an electron microscope photo of the surface after a voltage breakdown. At the spot of the sparking, a little crater can be seen, indicating possible damage to the surface and, as a consequence, deterioration of the achievable gradient which has to be avoided under all circumstances [12]. Although considerably higher than the typical values in circular machines, the gradient $E_{\mathrm{acc}} = 100 \, \mathrm{MV \, m^{-1}}$ in a linear machine still leads to a design of overall length approximately 50 km for a maximum achievable energy of $E_{\mathrm{max}} = 3$ TeV.

5 Conclusion

To summarize, for future lepton ring colliders (or, to be more precise, electron–positron colliders), synchrotron radiation losses set a severe limit on the achievable beam energy; and very soon the size of the machines will become uneconomical. For a given limit in synchrotron radiation power, the dimensions of the machine would have to grow quadratically with the beam energy. Linear colliders are therefore proposed as the preferred way to go. In this case, the maximum achievable acceleration gradient is the key issue. New acceleration methods, namely plasma-based set-ups in which gradients have been observed that are much higher than those seen with conventional techniques, are a most promising concept for the design of future colliders. An impressive example is shown in Fig. 36: within a plasma cell of only a few centimetres in length, electrons are accelerated to several GeV. The gradients achievable are orders of magnitude higher than in any conventional machine (see, e.g., Ref. [13]). Still, there are problems to overcome, such as issues with overall efficiency, beam quality (mainly the energy spread of the beam), and the achievable repetition rate. Nevertheless, we are convinced that this is a promising field worthy of much further study—and this is what the present school is about.
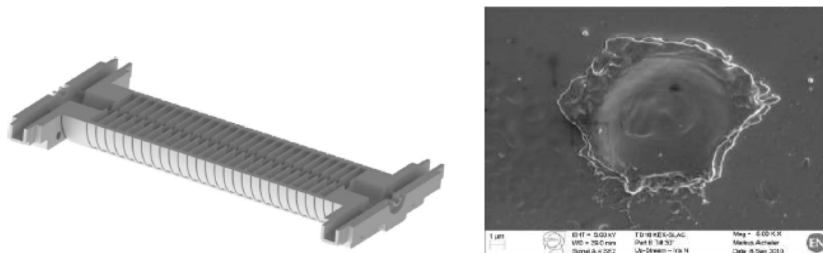
Fig. 35: Accelerating structure of the CLIC test facility CTF3; the electron microscope photo on the right shows the damage to the surface due to discharges in the module.
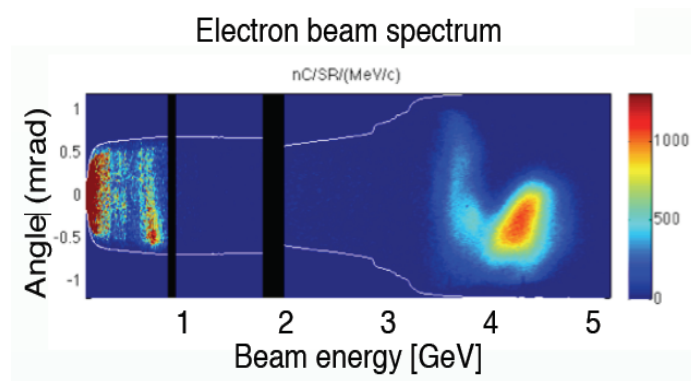
Fig. 36: Electron beam accelerated in the wake potential of a plasma cell; up to 4 GeV is obtained within only a few centimetres of length [13].

References

[1] LHC design report, edited by O.S. Brüning et al., CERN-2004-003 (CERN, Geneva, 2004). http://dx.doi.org/10.5170/CERN-2004-003-V-1, http://dx.doi.org/10.5170/CERN-2004-003-V-2, http://dx.doi.org/10.5170/CERN-2004-003-V-3

[2] C. Gerthsen, O. Kneser and H. Vogel, *Physik* (Springer, Berlin, 1982). http://dx.doi.org/10.1007/978-3-662-09308-5

[3] J.D. Cockcroft and E.T.S. Walton, *Proc. R. Soc. (Lond.)* **A136** (1932) 619, **A137** (1932) 229, **A144** (1934) 704.

[4] K. Wille, *The Physics of Particle Accelerators* (Oxford University Press, Oxford, 2000).

[5] R.J. Van de Graaff, *Phys. Rev.* **38** (1931) 1919.

[6] P.J. Bryant, in Proceedings of the CAS-CERN Accelerator School: 5th General Accelerator Physics Course, edited by Turner S., CERN-1994-001 (CERN, Geneva, 1994). http://dx.doi.org/10.5170/CERN-1994-001.1

[7] HERA Design Team, HERA: a proposal for a large electron proton colliding beam facility at DESY, DESY-HERA-81/10 (1981).

[8] J. Rossbach and P. Schmueser, in Proceedings of teh CAS-CERN Accelerator School: 5th General Accelerator Physics Course, edited by Turner S., CERN-1994-001 (CERN, Geneva, 1994). http://dx.doi.org/10.5170/CERN-1994-001.17

[9] W. Herr, in Proceedings of the CAS-CERN Accelerator School: Intermediate Course on Accelerator Physics, edited by Brandt D., CERN-2006-002 (CERN, Geneva, 2006). http://dx.doi.org/10.5170/CERN-2006-002.379

[10] F. Zimmermann, Challenges for highest energy circular colliders, in Proc. 5th International Particle Accelerator Conf. (IPAC 2014, Dresden), in press.

[11] A multi-TeV linear collider based on CLIC technology: CLIC conceptual design report, edited by M. Aicheler et al., CERN-2012-007 (CERN, Geneva, 2012). http://dx.doi.org/10.5170/CERN-2012-007

[12] A. Palaia, Beam momentum changes due to discharge in high-gradient accelerator structures, Ph.D. thesis, Uppsala University, 2013; ISBN 978-91-554-8802-4.

[13] W.P. Leemans *et al.*, *Phys. Rev. Lett.* **113** (2014) 245002. http://dx.doi.org/10.1103/PhysRevLett.113.245002

[14] A. Cho, *Science* **337** (2012) 141. http://dx.doi.org/10.1126/science.337.6091.141

Introduction to Plasma Physics

P. Gibbon
Forschungszentrum Jülich GmbH, Institute for Advanced Simulation, Jülich Supercomputing Centre, Jülich, Germany

Abstract
These notes are intended to provide a brief primer in plasma physics, introducing common definitions, basic properties, and typical processes found in plasmas. These concepts are inherent in contemporary plasma-based accelerator schemes, and thus provide a foundation for the more advanced expositions that follow in this volume. No prior knowledge of plasma physics is required, but the reader is assumed to be familiar with basic electrodynamics and fluid mechanics.

Keywords
Plasma properties; 2-fluid model; Langmuir waves; electromagnetic wave propagation; dispersion relation; nonlinear waves.

1 Plasma types and definitions

Plasmas are often described as the fourth state of matter, alongside gases, liquids and solids, a definition which does little to illuminate their main physical attributes. In fact, a plasma can exhibit behaviour characteristic of all three of the more familiar states, depending on its density and temperature, so we obviously need to look for other distinguishing features. A simple textbook definition of a plasma [1, 2] would be: a *quasi-neutral* gas of charged particles showing *collective* behaviour. This may seem precise enough, but the rather fuzzy-sounding terms of 'quasi-neutrality' and 'collectivity' require further explanation. The first of these, 'quasi-neutrality', is actually just a mathematical way of saying that even though the particles making up a plasma consist of free electrons and ions, their overall charge densities cancel each other in equilibrium. So if n_e and n_i are, respectively, the number densities of electrons and ions with charge state Z, then these are *locally balanced*, i.e.

$$n_e \simeq Z n_i. \tag{1}$$

The second property, 'collective' behaviour, arises because of the long-range nature of the $1/r$ Coulomb potential, which means that local disturbances in equilibrium can have a strong influence on remote regions of the plasma. In other words, macroscopic fields usually dominate over short-lived microscopic fluctuations, and a net charge imbalance $\rho = e(Z n_i - n_e)$ will immediately give rise to an electrostatic field according to Gauss's law,

$$\nabla \cdot \mathbf{E} = \rho/\varepsilon_0.$$

Likewise, the same set of charges moving with velocities v_e and v_i will give rise to a *current* density $J = e(Z n_i v_i - n_e v_e)$. This in turn induces a magnetic field according to Ampères law,

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}.$$

It is these internally driven electric and magnetic fields that largely determine the dynamics of the plasma, including its response to externally applied fields through particle or laser beams—as, for example, in the case of plasma-based accelerator schemes.

Now that we have established what plasmas are, it is natural to ask where we can find them. In fact they are rather ubiquitous: in the cosmos, 99% of the visible universe—including stars, the interstellar

medium and jets of material from various astrophysical objects—is in a plasma state. Closer to home, the ionosphere, extending from around 50 km (equivalent to 10 Earth radii) to 1000 km, provides vital protection from solar radiation to life on Earth. Terrestrial plasmas can be found in fusion devices (machines designed to confine, ignite and ultimately extract energy from deuterium–tritium fuel), street lighting, industrial plasma torches and etching processes, and lightning discharges. Needless to say, plasmas play a central role in the topic of the present school, supplying the medium to support very large travelling-wave field structures for the purpose of accelerating particles to high energies. Table 1 gives a brief overview of these various plasma types and their properties.

Table 1: Densities and temperatures of various plasma types

Type	Electron density n_e (cm^{-3})	Temperature T_e (eV^a)
Stars	10^{26}	2×10^3
Laser fusion	10^{25}	3×10^3
Magnetic fusion	10^{15}	10^3
Laser-produced	10^{18}–10^{24}	10^2–10^3
Discharges	10^{12}	1–10
Ionosphere	10^6	0.1
Interstellar medium	1	10^{-2}

^{a} 1 eV $\equiv$ 11 600 K.

1.1 Debye shielding

In most types of plasma, quasi-neutrality is not just an ideal equilibrium state; it is a state that the plasma actively tries to achieve by readjusting the local charge distribution in response to a disturbance. Consider a hypothetical experiment in which a positively charged ball is immersed in a plasma; see Fig. 1. After some time, the ions in the ball's vicinity will be repelled and the electrons will be attracted, leading to an altered average charge density in this region. It turns out that we can calculate the potential $\phi(r)$ of this ball after such a readjustment has taken place.
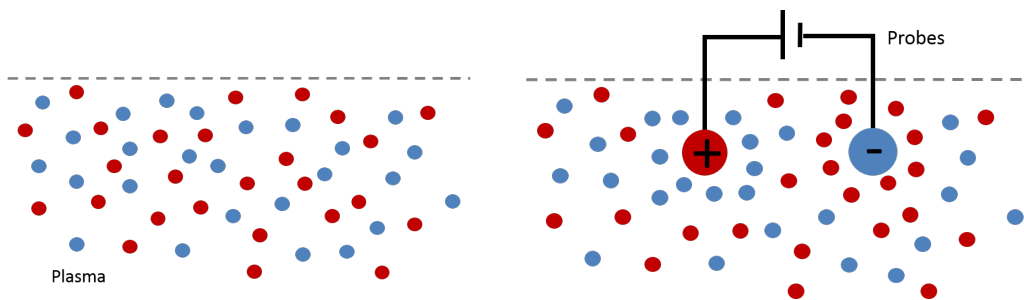
Fig. 1: Debye shielding of charged spheres immersed in a plasma

First of all, we need to know how fast the electrons and ions actually move. For equal ion and electron temperatures ($T_e = T_i$), we have

$$\frac{1}{2}m_e \bar{v}_e^2 = \frac{1}{2}m_i \bar{v}_i^2 = \frac{3}{2}k_B T_e. \tag{2}$$

Therefore, for a hydrogen plasma, where $Z = A = 1$,

$$\frac{\overline{v}_\mathrm{i}}{\overline{v}_\mathrm{e}} = \left(\frac{m_\mathrm{e}}{m_\mathrm{i}}\right)^{1/2} = \left(\frac{m_\mathrm{e}}{A m_\mathrm{p}}\right)^{1/2} = \frac{1}{43}.$$

In other words, the ions are almost stationary on the electron time-scale. To a good approximation, we often write

$$n_\mathrm{i} \simeq n_0, \tag{3}$$

where $n_0 = N_\mathrm{A} \rho_\mathrm{m}/A$ is the material (e.g. gas) number density, with ρ_m being the usual mass density and N_A the Avogadro constant. In thermal equilibrium, the electron density follows a Boltzmann distribution [1],

$$n_\mathrm{e} = n_\mathrm{i} \exp(e\phi/k_\mathrm{B}T_\mathrm{e}), \tag{4}$$

where n_i is the ion density, k_B is the Boltzmann constant, and $\phi(r)$ is the potential created by the external disturbance. From Gauss's law (Poisson's equation), we can also write

$$\nabla^2 \phi = -\frac{\rho}{\varepsilon_0} = -\frac{e}{\varepsilon_0}(n_\mathrm{i} - n_\mathrm{e}). \tag{5}$$

So now we can combine (5) with (4) and (3) in spherical geometry[1] to eliminate n_e and arrive at a physically meaningful solution:

$$\phi_\mathrm{D} = \frac{1}{4\pi\varepsilon_0} \frac{\exp(-r/\lambda_\mathrm{D})}{r}. \tag{6}$$

This condition supposes that $\phi \to 0$ at $r = \infty$. The characteristic length-scale λ_D inside the exponential factor is known as the *Debye length*, and is given by

$$\lambda_\mathrm{D} = \left(\frac{\varepsilon_0 k_\mathrm{B}T_\mathrm{e}}{e^2 n_\mathrm{e}}\right)^{1/2} = 743 \left(\frac{T_\mathrm{e}}{\mathrm{eV}}\right)^{1/2} \left(\frac{n_\mathrm{e}}{\mathrm{cm}^{-3}}\right)^{-1/2} \mathrm{cm}. \tag{7}$$

The Debye length is a fundamental property of nearly all plasmas of interest, and depends equally on the plasma's temperature and density. An *ideal* plasma has many particles per Debye sphere, i.e.

$$N_\mathrm{D} \equiv n_\mathrm{e} \frac{4\pi}{3} \lambda_\mathrm{D}^3 \gg 1, \tag{8}$$

which is a prerequisite for the collective behaviour discussed earlier. An alternative way of expressing this condition is via the so-called *plasma parameter*,

$$g \equiv \frac{1}{n_\mathrm{e}\lambda_\mathrm{D}^3}, \tag{9}$$

which is essentially the reciprocal of N_D. Classical plasma theory is based on the assumption that $g \ll 1$, which implies dominance of collective effects over collisions between particles. Therefore, before we refine our plasma classification, it is worth taking a quick look at the nature of collisions between plasma particles.

1.2 Collisions in plasmas

Where $N_\mathrm{D} \leq 1$, screening effects are reduced and collisions will dominate the particle dynamics. In intermediate regimes, collisionality is usually measured via the *electron–ion collision rate*, given by

$$\nu_\mathrm{ei} = \frac{\pi^{3/2} n_\mathrm{e} Z e^4 \ln\Lambda}{2^{1/2}(4\pi\varepsilon_0)^2 m_\mathrm{e}^2 v_\mathrm{te}^3} \ \mathrm{s}^{-1}, \tag{10}$$

---

[1] $\nabla^2 \to \frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2 \frac{\mathrm{d}\phi}{\mathrm{d}r}\right).$

where $v_{\text{te}} \equiv \sqrt{k_{\text{B}} T_{\text{e}}/m_{\text{e}}}$ is the electron thermal velocity and $\ln \Lambda$ is a slowly varying term, called the Coulomb logarithm, which typically takes a numerical value of order 10–20. The numerical coefficient in expression (10) may vary between different texts depending on the definition used. Our definition is consistent with that in Refs. [5] and [3], which define the collision rate according to the average time taken for a thermal electron to be deflected by $90°$ via multiple scatterings from fixed ions. The collision frequency can also be written as

$$\frac{\nu_{\text{ei}}}{\omega_{\text{p}}} \simeq \frac{Z \ln \Lambda}{10 N_{\text{D}}} \quad \text{with } \ln \Lambda \simeq 9 N_{\text{D}}/Z,$$

where ω_{p} is the electron plasma frequency defined below in Eq. (11).

1.3 Plasma classification

Armed with our definition of plasma ideality, Eq. (8), we can proceed to make a classification of plasma types in density–temperature space. This is illustrated for a few examples in Fig. 2; the 'accelerator' plasmas of interest in the present school are found in the middle of this chart, having densities corresponding to roughly atmospheric pressure and temperatures of a few eV (10^4 K) as a result of field ionization; see Section 1.5.
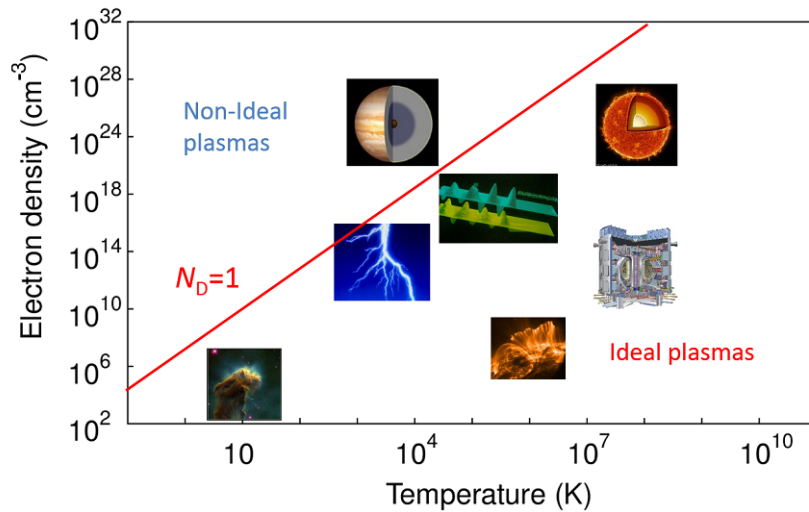
Fig. 2: Examples of plasma types in the density–temperature plane

1.4 Plasma oscillations

So far we have considered characteristics, such as density and temperature, of a plasma in equilibrium. We can also ask how fast the plasma will respond to an external disturbance, which could be due to electromagnetic waves (e.g. a laser pulse) or particle beams. Consider a quasi-neutral plasma slab in which an electron layer is displaced from its initial position by a distance δ, as illustrated in Fig. 3. This creates two 'capacitor' plates with surface charge $\sigma = \pm e n_{\text{e}} \delta$, resulting in an electric field

$$\mathbf{E} = \frac{\sigma}{\varepsilon_0} = \frac{e n_{\text{e}} \delta}{\varepsilon_0}.$$

The electron layer is accelerated back towards the slab by this restoring force according to

$$m_{\text{e}} \frac{\mathrm{d}v}{\mathrm{d}t} = -m_{\text{e}} \frac{\mathrm{d}^2 \delta}{\mathrm{d}t^2} = -eE = \frac{e^2 n_{\text{e}} \delta}{\varepsilon_0},$$
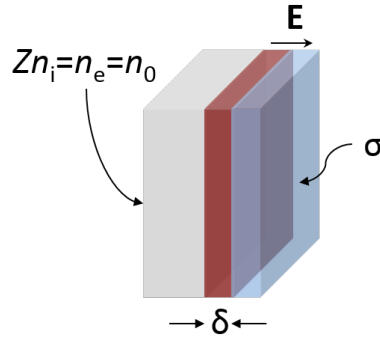
Fig. 3: Slab or capacitor model of an oscillating electron layer

or

$$\frac{\mathrm{d}^2\delta}{\mathrm{d}t^2} + \omega_{\mathrm{p}}^2 \delta = 0$$

where

$$\omega_{\mathrm{p}} \equiv \left(\frac{e^2 n_{\mathrm{e}}}{\varepsilon_0 m_{\mathrm{e}}}\right)^{1/2} \simeq 5.6 \times 10^4 \left(\frac{n_{\mathrm{e}}}{\mathrm{cm}^{-3}}\right)^{1/2} \mathrm{s}^{-1} \tag{11}$$

is the *electron plasma frequency*.

This quantity can be obtained via another route by returning to the Debye sheath problem of Section 1.1 and asking how quickly it would take the plasma to adjust to the insertion of the foreign charge. For a plasma of temperature T_{e}, the response time to recover quasi-neutrality is just the ratio of the Debye length to the thermal velocity $v_{\mathrm{te}} \equiv \sqrt{k_{\mathrm{B}} T_{\mathrm{e}}/m_{\mathrm{e}}}$; that is,

$$t_{\mathrm{D}} \simeq \frac{\lambda_{\mathrm{D}}}{v_{\mathrm{te}}} = \left(\frac{\varepsilon_0 k_{\mathrm{B}} T_{\mathrm{e}}}{e^2 n_{\mathrm{e}}} \cdot \frac{m}{k_{\mathrm{B}} T_{\mathrm{e}}}\right)^{1/2} = \omega_{\mathrm{p}}^{-1}.$$
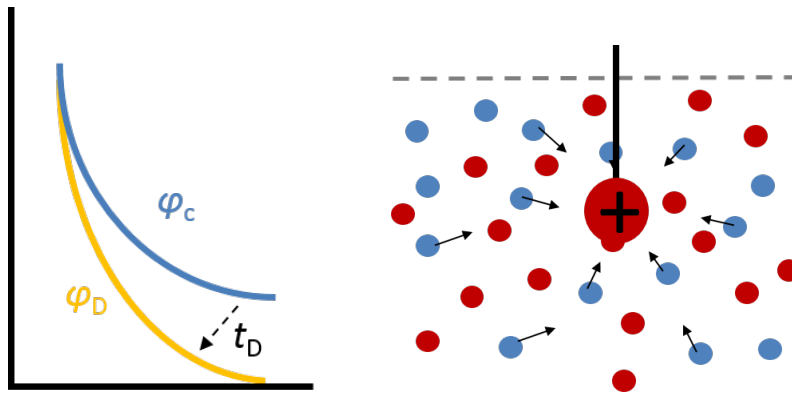
Fig. 4: Response time to form a Debye sheath

If the plasma response time is shorter than the period of a external electromagnetic field (such as a laser), then this radiation will be *shielded out*. To make this statement more quantitative, consider the ratio

$$\frac{\omega_{\mathrm{p}}^2}{\omega^2} = \frac{e^2 n_{\mathrm{e}}}{\varepsilon_0 m_{\mathrm{e}}} \cdot \frac{\lambda^2}{4\pi^2 c^2}.$$

Setting this to unity defines the wavelength λ_μ for which $n_e = n_c$, or

$$n_c \simeq 10^{21} \lambda_\mu^{-2} \text{ cm}^{-3}. \tag{12}$$

Radiation with wavelength $\lambda > \lambda_\mu$ will be reflected. In the pre-satellite/cable era, this property was exploited to good effect in the transmission of long-wave radio signals, which utilizes reflection from the ionosphere to extend the range of reception.

Typical gas jets have $P \sim 1$ bar and $n_e = 10^{18}$–10^{19} cm^{-3}, and the critical density for a glass laser is $n_c(1\mu) = 10^{21}$ cm^{-3}. Gas-jet plasmas are therefore *underdense*, since $\omega^2/\omega_p^2 = n_e/n_c \ll 1$. In this case, *collective effects* are important if $\omega_p \tau_{int} > 1$, where τ_{int} is some characteristic interaction time, such as the duration of a laser pulse or particle beam entering the plasma. For example, if $\tau_{int} = 100$ fs and $n_e = 10^{17}$ cm^{-3}, then $\omega_p \tau_{int} = 1.8$ and we will need to consider the plasma response on the interaction time-scale. Generally this is the situation we seek to exploit in all kinds of plasma applications, including short-wavelength radiation, nonlinear refractive properties, generation of high electric/magnetic fields and, of course, particle acceleration.

(a)　　　　　　　　　　　　　　(b)
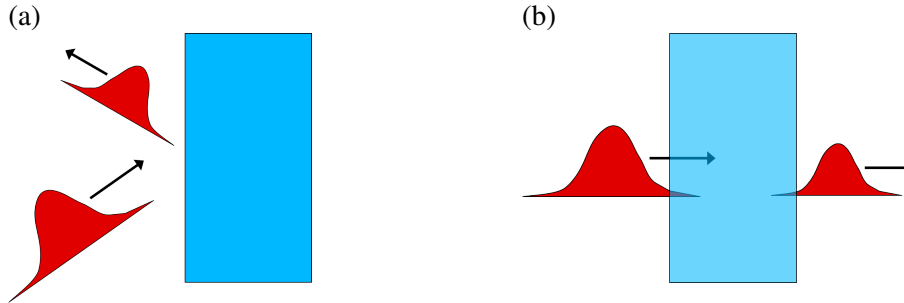
Fig. 5: (a) Overdense plasma, with $\omega < \omega_p$, showing mirror-like behaviour. (b) Underdense plasma, with $\omega > \omega_p$, which behaves like a nonlinear refractive medium.

1.5 Plasma creation

Plasmas are created via ionization, which can occur in several ways: through collisions of fast particles with atoms; through photoionization by electromagnetic radiation; or via electrical breakdown in strong electric fields. The latter two are examples of *field ionization*, which is the mechanism most relevant to the plasma accelerator context. To get some idea of when field ionization occurs, we need to know the typical field strength required to strip electrons away from an atom. At the Bohr radius

$$a_B = \frac{\hbar^2}{me^2} = 5.3 \times 10^{-9} \text{ cm},$$

the electric field strength is

$$E_a = \frac{e}{4\pi\varepsilon_0 a_B^2} \simeq 5.1 \times 10^9 \text{ V m}^{-1}. \tag{13}$$

This threshold can be expressed as the so-called *atomic intensity*,

$$I_a = \frac{\varepsilon_0 c E_a^2}{2} \simeq 3.51 \times 10^{16} \text{ W cm}^{-2}. \tag{14}$$

A laser intensity of $I_L > I_a$ will therefore guarantee ionization for any target material, though in fact ionization can occur well below this threshold (e.g. around 10^{14} W cm^{-2} for hydrogen) due to *multiphoton* effects. Simultaneous field ionization of many atoms produces a plasma with electron density n_e and temperature $T_e \sim 1$–10 eV.

1.6 Relativistic threshold

Before we discuss wave propagation in plasmas, it is useful to have some idea of the strength of the external fields used to excite them. To do this, we consider the classical equation of motion for an electron exposed to a linearly polarized laser field $\mathbf{E} = \hat{y} E_0 \sin \omega t$:

$$\frac{\mathrm{d}v}{\mathrm{d}t} \simeq \frac{-eE_0}{m_\mathrm{e}} \sin \omega t.$$

This implies that the electron will acquire a velocity

$$v = \frac{eE_0}{m_\mathrm{e}\omega} \cos \omega t = v_\mathrm{osc} \cos \omega t, \tag{15}$$

which is usually expressed in terms of a dimensionless oscillation amplitude

$$a_0 \equiv \frac{v_\mathrm{osc}}{c} \equiv \frac{p_\mathrm{osc}}{m_\mathrm{e}c} \equiv \frac{eE_0}{m_\mathrm{e}\omega c}. \tag{16}$$

In many articles and books a_0 is referred to as the 'quiver' velocity or momentum; it can exceed unity, in which case the normalized momentum (third expression) is more appropriate, since the real particle velocity is just pinned to the speed of light. The laser intensity I_L and wavelength λ_L are related to E_0 and ω through

$$I_\mathrm{L} = \frac{1}{2}\varepsilon_0 c E_0^2, \qquad \lambda_\mathrm{L} = \frac{2\pi c}{\omega}.$$

By substituting these into (16) one can show that

$$a_0 \simeq 0.85(I_{18}\lambda_\mu^2)^{1/2}, \tag{17}$$

where

$$I_{18} = \frac{I_\mathrm{L}}{10^{18}\ \mathrm{W\,cm^{-2}}}, \qquad \lambda_\mu = \frac{\lambda_\mathrm{L}}{\mu\mathrm{m}}.$$

From this expression it can be seen that we will have relativistic electron velocities, or $a_0 \sim 1$, for intensities $I_\mathrm{L} \geq 10^{18}\ \mathrm{W\,cm^{-2}}$, at wavelengths $\lambda_\mathrm{L} \simeq 1\ \mu\mathrm{m}$.

2 Wave propagation in plasmas

The theory of wave propagation is an important subject in its own right, and has inspired a vast body of literature and a number of textbooks [4, 5, 8]. There are a great many possible ways in which plasmas can support waves, depending on the local conditions, the presence of external electric and magnetic fields, and so on. Here we will concentrate on two main wave forms: longitudinal oscillations of the kind we have encountered already, and electromagnetic waves. To derive and analyse wave phenomena, there are several possible theoretical approaches, with the suitability of each depending on the length- and time-scales of interest, which in laboratory plasmas can range from nanometres to metres and from femtoseconds to seconds. These approaches are:

   (i) first-principles N-body molecular dynamics;
  (ii) phase-space methods—the Vlasov–Boltzmann equation;
 (iii) two-fluid equations;
 (iv) magnetohydrodynamics (single magnetized fluid).

The first is rather costly and limited to much smaller regions of plasma than usually needed to describe the common types of wave. Indeed, the number of particles needed for first-principles modelling of a tokamak would be around 10^{21}; a laser-heated gas requires 10^{20} particles, still way out of reach of

even the most powerful computers available. Clearly a more tractable model is needed, and in fact many plasma phenomena can be analysed by assuming that each charged particle component of density n_s and velocity $\mathbf{u}_s$ behaves in a fluid-like manner, interacting with other species (s) via the electric and magnetic fields; this is the idea behind approach (iii). The rigorous way to derive the governing equations in this approximation is via *kinetic theory*, starting from method (ii) [2, 5], which is beyond the scope of this paper. Finally, slow wave phenomena on more macroscopic, ion time-scales can be handled with approach (iv) [2].

For the present purposes, we therefore start from the two-fluid equations for a plasma with finite temperature ($T_e > 0$) that is assumed to be collisionless ($\nu_{ie} \simeq 0$) and non-relativistic, so that the fluid velocities are such that $u \ll c$. The equations governing the plasma dynamics under these conditions are

$$\frac{\partial n_s}{\partial t} + \nabla \cdot (n_s \mathbf{u}_s) = 0, \tag{18}$$

$$n_s m_s \frac{d\mathbf{u}_s}{dt} = n_s q_s (\mathbf{E} + \mathbf{u}_s \times \mathbf{B}) - \nabla P_s, \tag{19}$$

$$\frac{d}{dt}(P_s n_s^{-\gamma_s}) = 0, \tag{20}$$

where P_s is the thermal pressure of species s and γ_s the specific heat ratio, or $(2 + N)/N$ with N the number of degrees of freedom.

The continuity equation (18) tells us that (in the absence of ionization or recombination) the number of particles *of each species* is conserved. Noting that the charge and current densities can be written as $\rho_s = q_s n_s$ and $\mathbf{J}_s = q_s n_s \mathbf{u}_s$, respectively, Eq. (18) can be rewritten as

$$\frac{\partial \rho_s}{\partial t} + \nabla \cdot \mathbf{J}_s = 0, \tag{21}$$

which expresses the conservation of *charge*.

Equation (19) governs the motion of a fluid element of species s in the presence of electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$. In the absence of fields, and assuming strict quasi-neutrality ($n_e = Z n_i = n$; $\mathbf{u}_e = \mathbf{u}_i = \mathbf{u}$), we recover the more familiar *Navier–Stokes* equations

$$\begin{aligned}
\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) &= 0, \\
\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} &= \frac{1}{\rho}\nabla P.
\end{aligned} \tag{22}$$

By contrast, in the plasma accelerator context we usually deal with time-scales over which the ions can be assumed to be motionless, i.e. $\mathbf{u}_i = 0$, and also unmagnetized plasmas, so that the momentum equation reads

$$n_e m_e \frac{d\mathbf{u}_e}{dt} = -e\, n_e \mathbf{E} - \nabla P_e. \tag{23}$$

Note that $\mathbf{E}$ can include both external and internal field components (via charge separation).

2.1 Longitudinal (Langmuir) waves

A characteristic property of plasmas is their ability to transfer momentum and energy via collective motion. One of the most important examples of this is the oscillation of electrons against a stationary ion background, or *Langmuir waves*. Returning to the two-fluid model, we can simplify (18)–(20) by setting $\mathbf{u}_i = 0$, restricting the electron motion to one dimension (x) and taking $\frac{\partial}{\partial y} = \frac{\partial}{\partial z} = 0$:

$$\frac{\partial n_e}{\partial t} + \frac{\partial}{\partial x}(n_e u_e) = 0,$$

$$n_{\mathrm{e}}\left(\frac{\partial u_{\mathrm{e}}}{\partial t} + u_{\mathrm{e}}\frac{\partial u_{\mathrm{e}}}{\partial x}\right) = -\frac{e}{m}n_{\mathrm{e}}E - \frac{1}{m}\frac{\partial P_{\mathrm{e}}}{\partial x}, \tag{24}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{P_{\mathrm{e}}}{n_{\mathrm{e}}^{\gamma_{\mathrm{e}}}}\right) = 0.$$

The system (24) has three equations and four unknowns. To close it, we need an expression for the electric field, which, since $\mathbf{B} = 0$, can be found from Gauss's law (Poisson's equation) with $Zn_{\mathrm{i}} = n_0$:

$$\frac{\partial E}{\partial x} = \frac{e}{\varepsilon_0}(n_0 - n_{\mathrm{e}}). \tag{25}$$

The system of equations (24)–(25) is nonlinear and, apart from a few special cases, cannot be solved exactly. A common technique for analysing waves in plasmas is to *linearize* the equations, which involves assuming that the perturbed amplitudes are small compared to the equilibrium values, i.e.

$$n_{\mathrm{e}} = n_0 + n_1,$$
$$u_{\mathrm{e}} = u_1,$$
$$P_{\mathrm{e}} = P_0 + P_1,$$
$$E = E_1,$$

where $n_1 \ll n_0$ and $P_1 \ll P_0$. Upon substituting these expressions into (24)–(25) and neglecting all products of perturbations such as $n_1\partial_t u_1$ and $u_1\partial_x u_1$, we get a set of linear equations for the perturbed quantities:

$$\frac{\partial n_1}{\partial t} + n_0\frac{\partial u_1}{\partial x} = 0,$$
$$n_0\frac{\partial u_1}{\partial t} = -\frac{e}{m}n_0 E_1 - \frac{1}{m}\frac{\partial P_1}{\partial x}, \tag{26}$$
$$\frac{\partial E_1}{\partial x} = -\frac{e}{\varepsilon_0}n_1,$$
$$P_1 = 3k_{\mathrm{B}}T_{\mathrm{e}}\,n_1.$$

The expression for P_1 results from the specific heat ratio $\gamma_{\mathrm{e}} = 3$ and from assuming isothermal background electrons, $P_0 = k_{\mathrm{B}}T_{\mathrm{e}}n_0$ (ideal gas); see Kruer's book [5]. We can now eliminate E_1, P_1 and u_1 from (26) to get

$$\left(\frac{\partial^2}{\partial t^2} - 3v_{\mathrm{te}}^2\frac{\partial^2}{\partial x^2} + \omega_{\mathrm{p}}^2\right)n_1 = 0, \tag{27}$$

with $v_{\mathrm{te}}^2 = k_{\mathrm{B}}T_{\mathrm{e}}/m_{\mathrm{e}}$ and ω_{p} given by (11) as before. Finally, we look for plane-wave solutions of the form $A = A_0\exp\{\mathrm{i}(\omega t - kx)\}$, so that our derivative operators are transformed as follows: $\frac{\partial}{\partial t} \to \mathrm{i}\omega$ and $\frac{\partial}{\partial x} \to -\mathrm{i}k$. Substitution into (27) yields the Bohm–Gross dispersion relation

$$\omega^2 = \omega_{\mathrm{p}}^2 + 3k^2 v_{\mathrm{te}}^2. \tag{28}$$

This and other dispersion relations are often depicted graphically on a chart such as that in Fig. 6, which gives an overview of which propagation modes are permitted for low- and high-wavelength limits.

2.2 Transverse waves

To describe *transverse* electromagnetic (EM) waves, we need two additional Maxwell's equations, Faraday's law and Ampère's law, which we will introduce properly later; see Eqs. (38) and (39). For the time being, it is helpful to simplify things by making use of our previous analysis of small-amplitude longitudinal waves. Therefore, we linearize and again apply the harmonic approximation $\frac{\partial}{\partial t} \to \mathrm{i}\omega$ to get

$$\nabla \times \mathbf{E}_1 = -\mathrm{i}\omega\mathbf{B}_1, \tag{29}$$
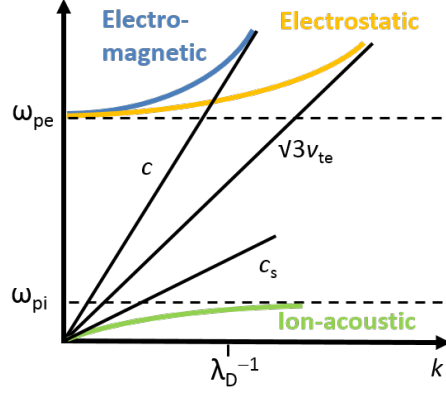
Fig. 6: Schematic illustration of dispersion relations for Langmuir, electromagnetic and ion-acoustic waves

$$\nabla \times \mathbf{B}_1 = \mu_0 \mathbf{J}_1 + i\varepsilon_0 \mu_0 \omega \mathbf{E}_1, \tag{30}$$

where the transverse current density is given by

$$\mathbf{J}_1 = -n_0 e \mathbf{u}_1. \tag{31}$$

This time we look for pure EM plane-wave solutions with $\mathbf{E}_1 \perp \mathbf{k}$ (see Fig. 7) and also assume that the group and phase velocities are large enough, $v_\mathrm{p}, v_g \gg v_\mathrm{te}$, so that we have a *cold* plasma with $P_\mathrm{e} = n_0 k_\mathrm{B} T_\mathrm{e} \simeq 0$.
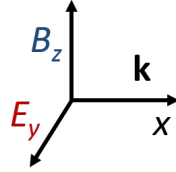
Fig. 7: Geometry for electromagnetic plane-wave analysis

The linearized electron fluid velocity and corresponding current are then

$$\mathbf{u}_1 = -\frac{e}{i\omega m_\mathrm{e}} \mathbf{E}_1,$$

$$\mathbf{J}_1 = \frac{n_0 e^2}{i\omega m_\mathrm{e}} \mathbf{E}_1 \equiv \sigma \mathbf{E}_1, \tag{32}$$

where σ is the *AC electrical conductivity*. By analogy with dielectric media (see, e.g., Ref. [7]), in which Ampère's law is usually written as $\nabla \times \mathbf{B}_1 = \mu_0 \partial_t \mathbf{D}_1$, by substituting (32) into (39) one can show that

$$\mathbf{D}_1 = \varepsilon_0 \varepsilon \mathbf{E}_1,$$

with

$$\varepsilon = 1 + \frac{\sigma}{i\omega\varepsilon_0} = 1 - \frac{\omega_\mathrm{p}^2}{\omega^2}. \tag{33}$$

From (33) it follows immediately that

$$\eta \equiv \sqrt{\varepsilon} = \frac{ck}{\omega} = \left(1 - \frac{\omega_\mathrm{p}^2}{\omega^2}\right)^{1/2}, \tag{34}$$

with

$$\omega^2 = \omega_{\mathrm{p}}^2 + c^2 k^2. \tag{35}$$

The above expression can also be found directly by elimination of $\mathbf{J}_1$ and $\mathbf{B}_1$ from Eqs. (29)–(32). From the dispersion relation (35), also depicted in Fig. 6, a number of important features of EM wave propagation in plasmas can be deduced. For *underdense* plasmas ($n_{\mathrm{e}} \ll n_{\mathrm{c}}$),

$$\text{phase velocity} \quad v_{\mathrm{p}} = \frac{\omega}{k} \simeq c \left(1 + \frac{\omega_{\mathrm{p}}^2}{2\omega^2} \right) > c \,;$$

$$\text{group velocity} \quad v_{\mathrm{g}} = \frac{\partial \omega}{\partial k} \simeq c \left(1 - \frac{\omega_{\mathrm{p}}^2}{2\omega^2} \right) < c \,.$$

In the opposite case of an *overdense* plasma, where $n_{\mathrm{e}} > n_{\mathrm{c}}$, the refractive index η becomes imaginary and the wave can no longer propagate, becoming evanescent instead, with a decay length determined by the *collisionless skin depth* c/ω_{p}; see Fig. 8.
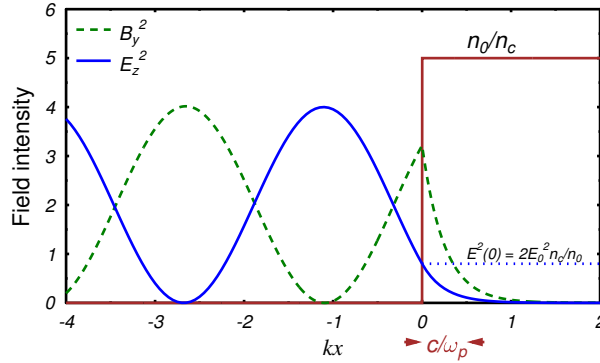
Fig. 8: Electromagnetic fields resulting from reflection of an incoming wave by an overdense plasma slab

2.3 Nonlinear wave propagation

So far we have considered purely longitudinal or transverse waves; linearizing the wave equations ensures that any nonlinearities or coupling between these two modes is excluded. While this is a reasonable approximation for low-amplitude waves, it is inadequate for decscribing strongly driven waves in the relativistic regime of interest in plasma accelerator schemes. The starting point of most analyses of nonlinear wave propagation phenomena is the Lorentz equation of motion for the electrons in a *cold* ($T_{\mathrm{e}} = 0$) unmagnetized plasma, together with Maxwell's equations [5,6]. We make two further assumptions: (i) that the ions are initially singly charged ($Z = 1$) and are treated as an immobile ($v_{\mathrm{i}} = 0$), homogeneous background with $n_0 = Z n_{\mathrm{i}}$; (ii) that thermal motion can be neglected, since the temperature remains low compared to the typical oscillation energy in the laser field ($v_{\mathrm{osc}} \gg v_{\mathrm{te}}$). The starting equations (in SI units) are then as follows:

$$\frac{\partial \mathbf{p}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{p} = -e(\mathbf{E} + \mathbf{v} \times \mathbf{B}), \tag{36}$$

$$\nabla \cdot \mathbf{E} = \frac{e}{\varepsilon_0}(n_0 - n_{\mathrm{e}}), \tag{37}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{38}$$

$$c^2 \nabla \times \mathbf{B} = -\frac{e}{\varepsilon_0} n_{\mathrm{e}} \mathbf{v} + \frac{\partial \mathbf{E}}{\partial t}, \tag{39}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{40}$$

where $\mathbf{p} = \gamma m_{\mathrm{e}} \mathbf{v}$ and $\gamma = (1 + p^2/m_{\mathrm{e}}^2 c^2)^{1/2}$.

To simplify matters, we first assume a plane-wave geometry like that in Fig. 7, with the transverse electromagnetic fields given by $\mathbf{E}_{\mathrm{L}} = (0, E_y, 0)$ and $\mathbf{B}_{\mathrm{L}} = (0, 0, B_z)$. From Eq. (36), the transverse electron momentum is then simply

$$p_y = eA_y, \tag{41}$$

where $E_y = \partial A_y/\partial t$. This relation expresses conservation of canonical momentum. Substituting $\mathbf{E} = -\nabla\phi - \partial\mathbf{A}/\partial t$ and $\mathbf{B} = \nabla \times \mathbf{A}$ into Ampère's equation (39) yields

$$c^2 \nabla \times (\nabla \times \mathbf{A}) + \frac{\partial^2 \mathbf{A}}{\partial t^2} = \frac{\mathbf{J}}{\varepsilon_0} - \nabla \frac{\partial \phi}{\partial t},$$

where the current is given by $\mathbf{J} = -en_{\mathrm{e}} \mathbf{v}$. Now we use a bit of vectorial wizardry, splitting the current into rotational (solenoidal) and irrotational (longitudinal) parts,

$$\mathbf{J} = \mathbf{J}_\perp + \mathbf{J}_\parallel = \nabla \times \mathbf{\Pi} + \nabla\Psi,$$

from which we can deduce (see Jackson's book [7]) that

$$\mathbf{J}_\parallel - \frac{1}{c^2} \nabla \frac{\partial \phi}{\partial t} = 0.$$

Finally, by applying the Coulomb gauge $\nabla \cdot \mathbf{A} = 0$ and $v_y = eA_y/\gamma$ from (41), we obtain

$$\frac{\partial^2 A_y}{\partial t^2} - c^2 \nabla^2 A_y = \mu_0 J_y = -\frac{e^2 n_{\mathrm{e}}}{\varepsilon_0 m_{\mathrm{e}} \gamma} A_y. \tag{42}$$

The nonlinear source term on the right-hand side contains two important bits of physics: $n_{\mathrm{e}} = n_0 + \delta n$, which couples the EM wave to plasma waves, and $\gamma = \sqrt{1 + \mathbf{p}^2/m_{\mathrm{e}}^2 c^2}$, which introduces relativistic effects through the increased electron inertia. Taking the *longitudinal* component of the momentum equation (36) gives

$$\frac{\mathrm{d}}{\mathrm{d}t}(\gamma m_{\mathrm{e}} v_x) = -eE_x - \frac{e^2}{2m_{\mathrm{e}}\gamma} \frac{\partial A_y^2}{\partial x}.$$

We can eliminate v_x using the x component of Ampère's law (39):

$$0 = -\frac{e}{\varepsilon_0} n_{\mathrm{e}} v_x + \frac{\partial E_x}{\partial t}.$$

And the electron density can be determined via Poisson's equation (37):

$$n_{\mathrm{e}} = n_0 - \frac{\varepsilon_0}{e} \frac{\partial E_x}{\partial x}.$$

The above (closed) set of equations can in principle be solved numerically for arbitrary pump strengths. For the moment, we simplify things by linearizing the *plasma* fluid quantities. Let

$$n_{\mathrm{e}} \simeq n_0 + n_1 + \cdots,$$
$$v_x \simeq v_1 + v_2 + \cdots,$$

and neglect products of perturbations such as $n_1 v_1$. This leads to

$$\left(\frac{\partial^2}{\partial t^2} + \frac{\omega_{\mathrm{p}}^2}{\gamma_0} \right) E_x = -\frac{\omega_{\mathrm{p}}^2 e}{2m_{\mathrm{e}} \gamma_0^2} \frac{\partial}{\partial x} A_y^2. \tag{43}$$
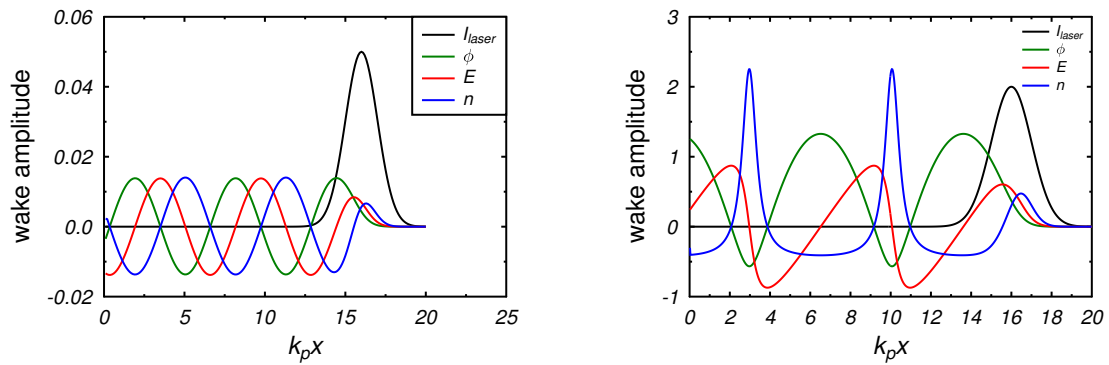
Fig. 9: Wakefield excitation by a short-pulse laser propagating in the positive x direction in the linear regime (left) and nonlinear regime (right).

The driving term on the right-hand side is the *relativistic ponderomotive force*, with $\gamma_0 = (1 + a_0^2/2)^{1/2}$. Some solutions of Eq. (43) are shown in Fig. 9, for low- and high-intensity laser pulses. The properties of the wakes will be discussed in detail in other lectures, but we can already see some obvious qualitative differences between the linear and nonlinear wave forms; the latter are typically characterized by a spiked density profile, a sawtooth electric field, and a longer wavelength.

The coupled fluid equations (42) and (43) and their fully nonlinear counterparts describe a wide range of nonlinear laser–plasma interaction phenomena, many of which are treated in the later lectures of this school, including plasma wake generation, blow-out regime laser self-focusing and channelling, parametric instabilities, and harmonic generation. Plasma-accelerated particle *beams*, on the other hand, cannot be treated using fluid theory and require a more sophisticated kinetic approach, usually assisted by numerical models solved with the aid of powerful supercomputers.

References

[1] F.F. Chen, *Plasma Physics and Controlled Fusion*, 2nd ed. (Springer, New York, 2006).

[2] R.O. Dendy (ed.), *Plasma Physics: An Introductory Course* (Cambridge University Press, Cambridge, 1993).

[3] J.D. Huba, *NRL Plasma Formulary* (Naval Research Laboratory, Washington, DC, 2007), `http://www.nrl.navy.mil/ppd/content/nrl-plasma-formulary`

[4] J. Boyd and J.J. Sanderson, *The Physics of Plasmas* (Cambridge University Press, Cambridge, 2003).

[5] W. Kruer, *The Physics of Laser Plasma Interactions* (Addison-Wesley, Boston, 1988).

[6] P. Gibbon, *Short Pulse Laser Interactions with Matter: An Introduction* (Imperial College Press, London, 2005). http://dx.doi.org/10.1142/p116

[7] J.D. Jackson, *Classical Electrodynamics*, 2nd ed. (Wiley, New York, 1975), 3rd ed. (Wiley, New York, 1998).

[8] J.P. Dougherty, in *Plasma Physics*, Ed. R. Dendy (Cambridge University Press, Cambridge, 1993), Chap. 3.

Appendices

A   Useful constants and formulae

Table A.1: Commonly used physical constants

Name	Symbol	Value (SI)	Value (cgs)
Boltzmann constant	k_B	1.38×10^{-23} J K^{-1}	1.38×10^{-16} erg K^{-1}
Electron charge	e	1.6×10^{-19} C	4.8×10^{-10} statcoul
Electron mass	m_e	9.1×10^{-31} kg	9.1×10^{-28} g
Proton mass	m_p	1.67×10^{-27} kg	1.67×10^{-24} g
Planck constant	h	6.63×10^{-34} J s	6.63×10^{-27} erg-s
Speed of light	c	3×10^8 m s^{-1}	3×10^{10} cm s^{-1}
Dielectric constant	ε_0	8.85×10^{-12} F m^{-1}	—
Permeability constant	μ_0	$4\pi \times 10^{-7}$	—
Proton/electron mass ratio	m_p/m_e	1836	1836
Temperature = 1eV	e/k_B	11 604 K	11 604 K
Avogadro number	N_A	6.02×10^{23} mol^{-1}	6.02×10^{23} mol^{-1}
Atmospheric pressure	1 atm	1.013×10^5 Pa	1.013×10^6 dyne cm^{-2}

Table A.2: Formulae in SI and cgs units

Name	Symbol	Formula (SI)	Formula (cgs)
Debye length	λ_D	$\left(\dfrac{\varepsilon_0 k_B T_e}{e^2 n_e}\right)^{1/2}$ m	$\left(\dfrac{k_B T_e}{4\pi e^2 n_e}\right)^{1/2}$ cm
Particles in Debye sphere	N_D	$\dfrac{4\pi}{3}\lambda_D^3$	$\dfrac{4\pi}{3}\lambda_D^3$
Plasma frequency (electrons)	ω_{pe}	$\left(\dfrac{e^2 n_e}{\varepsilon_0 m_e}\right)^{1/2}$ s^{-1}	$\left(\dfrac{4\pi e^2 n_e}{m_e}\right)^{1/2}$ s^{-1}
Plasma frequency (ions)	ω_{pi}	$\left(\dfrac{Z^2 e^2 n_i}{\varepsilon_0 m_i}\right)^{1/2}$ s^{-1}	$\left(\dfrac{4\pi Z^2 e^2 n_i}{m_i}\right)^{1/2}$ s^{-1}
Thermal velocity	$v_{te} = \omega_{pe}\lambda_D$	$\left(\dfrac{k_B T_e}{m_e}\right)^{1/2}$ m s^{-1}	$\left(\dfrac{k_B T_e}{m_e}\right)^{1/2}$ cm s^{-1}
Electron gyrofrequency	ω_c	eB/m_e s^{-1}	eB/m_e s^{-1}
Electron–ion collision frequency	ν_{ei}	$\dfrac{\pi^{3/2} n_e Z e^4 \ln \Lambda}{2^{1/2}(4\pi\varepsilon_0)^2 m_e^2 v_{te}^3}$ s^{-1}	$\dfrac{4(2\pi)^{1/2} n_e Z e^4 \ln \Lambda}{3 m_e^2 v_{te}^3}$ s^{-1}
Coulomb logarithm	$\ln \Lambda$	$\ln \dfrac{9 N_D}{Z}$	$\ln \dfrac{9 N_D}{Z}$

Table A.3: Useful formulae, with T_e in eV, n_e and n_i in cm^{-3}, and wavelength λ_L in μm

Plasma frequency	$\omega_\mathrm{pe} = 5.64 \times 10^4 n_\mathrm{e}^{1/2} \ \mathrm{s}^{-1}$
Critical density	$n_\mathrm{c} = 10^{21} \lambda_\mathrm{L}^{-2} \ \mathrm{cm}^{-3}$
Debye length	$\lambda_\mathrm{D} = 743 \, T_\mathrm{e}^{1/2} n_\mathrm{e}^{-1/2} \ \mathrm{cm}$
Skin depth	$\delta = c/\omega_\mathrm{p} = 5.31 \times 10^5 n_\mathrm{e}^{-1/2} \ \mathrm{cm}$
Electron–ion collision frequency	$\nu_\mathrm{ei} = 2.9 \times 10^{-6} n_\mathrm{e} T_\mathrm{e}^{-3/2} \ln\Lambda \ \mathrm{s}^{-1}$
Ion–ion collision frequency	$\nu_\mathrm{ii} = 4.8 \times 10^{-8} Z^4 \left(\dfrac{m_\mathrm{p}}{m_\mathrm{i}}\right)^{1/2} n_\mathrm{i} T_\mathrm{i}^{-3/2} \ln\Lambda \ \mathrm{s}^{-1}$
Quiver amplitude	$a_0 \equiv \dfrac{p_\mathrm{osc}}{m_\mathrm{e} c} = \left(\dfrac{I\lambda_\mathrm{L}^2}{1.37 \times 10^{18} \ \mathrm{W\,cm}^{-2}\,\mu\mathrm{m}^2}\right)^{1/2}$
Relativistic focusing threshold	$P_c = 17 \left(\dfrac{n_\mathrm{c}}{n_\mathrm{e}}\right) \ \mathrm{GW}$

Introduction to Plasma Accelerators: the Basics

R.Bingham[1,2] and R. Trines[1]
[1]Central Laser Facility, Rutherford Appleton Laboratory, Chilton, Didcot, Oxfordshire, UK
[2]Physics Department, University of Strathclyde, Glasgow, UK

Abstract
In this article, we concentrate on the basic physics of relativistic plasma wave accelerators. The generation of relativistic plasma waves by intense lasers or electron beams in low-density plasmas is important in the quest for producing ultra-high acceleration gradients for accelerators. A number of methods are being pursued vigorously to achieve ultra-high acceleration gradients using various plasma wave drivers; these include wakefield accelerators driven by photon, electron, and ion beams. We describe the basic equations and show how intense beams can generate a large-amplitude relativistic plasma wave capable of accelerating particles to high energies. We also demonstrate how these same relativistic electron waves can accelerate photons in plasmas.

Keywords
Laser; accelerators; wakefields; nonlinear theory; photon acceleration.

1 Introduction

Particle accelerators have led to remarkable discoveries about the nature of fundamental particles, providing the information that enabled scientists to develop and test the Standard Model of particle physics. The most recent milestone is the discovery of the Higgs boson using the Large Hadron Collider—the 27 km circumference 7 TeV proton accelerator at CERN. On a different scale, accelerators have many applications in science and technology, material science, biology, medicine, including cancer therapy, fusion research, and industry. These machines accelerate electrons or ions to energies in the range of tens of megaelectronvolts to tens of gigaelectronvolts. Electron beams with energies from several gigaelectronvolts to tens of gigaelectronvolts are used to generate intense X-rays in either synchrotrons or free electron lasers, such as the Linear Collider Light Source at Stanford or the European XFEL in Hamburg, for a range of applications. Particle accelerators developed in the last century are approaching the energy frontier. Today, at the terascale, the machines needed are extremely large and costly; even the smaller-scale lower energy accelerators are not small. The size of a conventional accelerator is set by the technology used to accelerate the particle and the final energy required. In conventional accelerators, radio-frequency microwave cavities support the electric fields responsible for accelerating charged particles. In these accelerators, owing to electrical breakdown of the walls, the electric field is limited to about $100 \, \text{MV m}^{-1}$. For more than 30 years, plasma-based particle accelerators driven by either lasers or particle beams have shown great promise, primarily because of the extremely large accelerating electric fields that they can support, about a thousand times greater than conventional accelerators, leading to the possibility of compact structures. These fields are supported by the collective motion of plasma electrons, forming a space charge disturbance moving at a speed slightly below c, the speed of light in a vacuum. This method of particle acceleration is commonly known as plasma wakefield acceleration.

Plasma-based accelerators are the brainchild of the late John Dawson and his colleagues at the University of California, Los Angeles, and are being investigated worldwide with a great deal of success. Will they be a serious competitor and displace the conventional 'dinosaur' variety? The impressive results that have so far been achieved show considerable promise for future plasma accelerators at the energy frontier, as well as providing much smaller 'table-top' ion and electron accelerators. Research on plasma-based accelerators is based on the seminal work by the late John Dawson and his collaborator

Toshi Tajima [1]. The main advantage of a plasma-based accelerator is that it can support accelerating electric fields many orders of magnitude greater than conventional devices that suffer from breakdown of the waveguide structure, since the plasma is already 'broken down'. The collective electric field E supported by the plasma is determined by the electron density $E \propto n^{1/2}$, where n is the electron density, and is known as an electron plasma wave; the collective electric fields are created by a drive beam that may be either a laser or a charged particle beam. These electron plasma waves travel with a phase speed close to the speed of the drive beam. The electric field strength E of the electron plasma wave is approximately determined by the electron density, $E \propto n^{1/2}$, where n is the density in cm^{-3}; for example, a plasma with density 10^{18} cm^{-3} can support a field of about 10^9 V cm^{-1}, a thousand times greater than a radio-frequency accelerator. This translates to a reduction in size of the accelerator and a reduction in cost.

The original plasma accelerator schemes investigated in the 1980s and 1990s were based on a long-pulse laser. Short-pulse lasers did not exist because chirped pulse amplification had not yet been demonstrated in the optical regime, only in the microwave regime. Experiments used the beat-wave mechanism of Tajima and Dawson [1], where two laser beams with a frequency difference equal to the plasma frequency drive a large-amplitude plasma wave. This changed when the process of chirped pulse amplification was ported from microwaves to laser beams by Strickland and Mourou [2, 3]. Suddenly, laser pulses could be produced that were shorter than the plasma wavelength (or skin depth) c/ω_p, where ω_p is the electron plasma frequency. This led to a dramatic change in the shape of the wakefield, from a 'density ripple' with many periods to a one- or two-period 'bubble-shaped' wakefield [4, 5]. The regime where the pulse length of the driving laser or particle beam is of the order of the plasma wavelength is commonly called the 'bubble' or 'blowout' regime. Most laser-driven and particle-driven particle accelerator experiments today are in this regime, and are commonly known as laser wakefield or beam-driven plasma wakefield accelerators. In the laser wakefield accelerator, the radiation pressure of a short, intense laser beam pushes plasma electrons forward and aside, creating a positively charged ion column. As the laser beam passes the displaced electrons snap back, owing to the restoring force of the ions, and overshoot, setting up a plasma density modulation behind the laser pulse. Similar plasma wakefields are set up by relativistic charged particle beams propagating through uniform plasma. A number of reviews on electron acceleration by laser-driven or beam-driven plasma waves have been published [6–10].

Early experiments produced beams with large energy spread, but in 2004 three independent groups in three different countries demonstrated laser wakefield acceleration producing mono-energetic electron beams with good emittance using short-pulse lasers [11–13], a result predicted by Pukhov and Meyer-ter-Vehn [14] and Tsung *et al.* [15]. Many groups worldwide now routinely produce electron beams at gigaelectronvolt energies using this scheme [16–18]. Similar plasma wakefields are set up by relativistic charged particle beams propagating through uniform plasma [19, 20]. In 2007, Joshi's group at the University of California demonstrated acceleration of electrons in metre-long plasma columns using a SLAC charged particle beam as a driver. This resulted in particles near the back of the electron beam doubling their energy from 42 GeV to 85 GeV in a 1 m long lithium plasma [21]. This is a remarkable result, since it takes 3 km of the SLAC linac to accelerate electrons to 42 GeV. The plasma beam-driven wakefield is incorporated into the latest round of experiments at SLAC by a consortium called the Facility for Advanced Accelerator Experimental Tests (FACET) [8]. Both electron and positron beams are accelerated in this facility [22, 23]. Beam-driven plasma wakefields also underpin the proton beam-driven wakefield experiment, AWAKE, which will use the proton beam from CERN's Super Proton Synchrotron and a 10 m long plasma column to produce gigaelectronvolt electrons [20, 24].

Today most experiments are conducted in the bubble regime. This includes experiments at the Berkeley Laboratory Laser Accelerator Center at Lawrence Berkeley Laboratory [18, 25] and the Rutherford Appleton Laboratory's Central Laser Facility [17, 26, 27], as well as many other laser-plasma accelerator experiments around the globe. These experiments have demonstrated mono-energetic electron beams at the gigaelectronvolt scale and planned experiments using lasers will demonstrate acceleration

of electrons to 10 GeV. Recently, FACET experiments demonstrated high efficiency in electron beam production, where the energy transfer from the wakefield to the accelerated bunch exceeded 30% with a low energy spread [22]. Similar impressive results using positrons have also been demonstrated [23].

Despite the successes of these experiments, it is still necessary to improve beam quality, in particular, to produce low energy spread and low emittance, and improve beam focusing. Most of the experiments are guided by plasma simulations that involve high-performance computing clusters. Commonly used simulation codes include the particle-in-cell codes Osiris [28–30], VLPL [31], Vorpal [32], and Epoch [33]. These simulations have already predicted that between 10 and 50 GeV electron beams can be created in one stage of a plasma accelerator.

If plasma accelerators are to take over from conventional machines, a great deal of effort still needs to be put into efficient drivers. Suitable laser efficiencies and pulse rates seem likely with diode-pumped lasers or with fibre lasers, but effort has to be put into these schemes to meet the requirements necessary to drive a wakefield. For beam-driven systems, electron beams at 100 GeV and proton beams with teraelectronvolt energies are required. These exist at the Large Hadron Collider for protons and at FACET for electrons. For an electron–positron system, a key challenge is positron acceleration; some groups are investigating positron acceleration in wakefields. Alternatively, an electron-proton collider or a photon–photon (γ–γ) collider could be built, doing away with the need for positrons, thus saving time and effort.

In the next section, we will discuss the short-pulse laser wakefield accelerator scheme. We will present basic analytical theory that lays the groundwork for all subsequent investigations into laser-driven and beam-driven wakefield acceleration, and provide results from particle-in-cell simulations of three-dimensional laser wakefield acceleration.

2   The laser wakefield accelerator (LWFA)

In the laser wakefield accelerator (LWFA), a short laser pulse, whose frequency is much greater than the plasma frequency, excites a wake of plasma oscillations (at ω_{p}), owing to the ponderomotive force, much like the wake of a motorboat. Since the plasma wave is not resonantly driven, as in the beat-wave, the plasma density does not have to be of a high uniformity to produce large-amplitude waves. We start from an intense laser pulse with electric field amplitude E_0 and frequency ω_0, and define $a_0 \equiv eE_0/(m_{\mathrm{e}}\omega_0 c)$. As this pulse propagates through an underdense plasma, $\omega_0 \gg \omega_{\mathrm{p}}$, the relativistic ponderomotive force associated with the laser envelope, $F_{\mathrm{pond}} \simeq -\frac{1}{2}mc^2(\nabla a_0^2)/\sqrt{1 + a_0^2}$, expels electrons from the region of the laser pulse and excites electron plasma waves. These waves are generated as a result of being displaced by the leading edge of the laser pulse. If the laser pulse length, $c\tau_{\mathrm{L}}$, is long compared with the electron plasma wavelength, the energy in the plasma wave is re-absorbed by the trailing part of the laser pulse. However, if the pulse length is approximately equal to or shorter than the plasma wavelength $c\tau_{\mathrm{L}} \simeq \lambda_{\mathrm{p}}$, the ponderomotive force excites plasma waves or wakefields, with a phase velocity equal to the laser group velocity, which are not re-absorbed. Thus, any pulse with a sharp rise or a sharp fall on a scale of c/ω_{p} will excite a wake. With the development of high-brightness lasers, the laser wakefield concept first put forward by Tajima and Dawson [1] in 1979 has now become a reality. The focal intensities of such lasers are $\geq 10^{19}$ W cm^{-2}, with $a_0 \geq 1$, which is the strong non-linear relativistic regime. Any analysis must, therefore, be in the strong non-linear relativistic regime and a perturbation procedure is invalid.

The maximum wake electric field amplitude generated by a plane-polarized pulse has been given by Sprangle *et al.* [34] in the one-dimensional limit as $E_{\mathrm{max}} = 0.38a_0^2 \left(1 + a_0^2/2\right)^{-1/2}\sqrt{n_0}$ V cm^{-1}. For $a_0 \approx 4$ and $n_0 = 10^{18}$ cm^{-1}, then $E_{\mathrm{max}} \approx 2$ GV cm^{-1}, and the time to reach this amplitude level is of the order of the laser pulse length.

2.1 Model equations describing laser wakefield excitation

To understand the laser wakefield excitation mechanism, it is sufficient to use a model based on one-fluid, cold relativistic hydrodynamics, and Maxwell's equations, together with a 'quasi-static' approximation, a set of two coupled non-linear equations describing the self-consistent evolution in one dimension of the laser pulse vector potential envelope, and the scalar potential of the excited wakefield. Starting from the equation for electron momentum,

$$\frac{\partial \mathbf{p}}{\partial t} + v_z \frac{\partial \mathbf{p}}{\partial z} = -\left(e\mathbf{E} + \frac{1}{c}\mathbf{v} \times \mathbf{B} \right) , \tag{1}$$

where

$$\mathbf{p} = m_0 \gamma \mathbf{v}, \ \gamma = \left(1 + p^2/m_0^2 c^2 \right)^{1/2} ,$$

m_0 and $\mathbf{v}$ being the electron rest mass and velocity.

In Eq. (1), we have assumed that all quantities only depend on z and t, z being the direction of propagation of the (external) pump and

$$\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{A}_\perp}{\partial z} - \hat{z}\frac{\partial \phi}{\partial z} ; \ \mathbf{B} = \nabla \times \mathbf{A}_\perp; \ \mathbf{A}_\perp = \hat{x}A_x + \hat{y}A_y , \tag{2}$$

where $\mathbf{A}_\perp$ is the vector potential of the electromagnetic pulse and ϕ is the ambipolar potential due to charge separation in the plasma.

Using Eqs. (1) and (2), the perpendicular component of the electron momentum is found to be

$$\frac{p_\perp}{m_0 c} = \frac{e}{m_0 c^2}\mathbf{A}_\perp \equiv \mathbf{a}(z,t) , \tag{3}$$

and we can write

$$\gamma = \left[1 + \left(\frac{p_\perp}{m_0 c} \right)^2 + \left(\frac{p_z}{m_0 c} \right)^2 \right]^{1/2} \equiv \gamma_a \gamma_\| , \tag{4}$$

where

$$\gamma_a = \left(1 + \mathbf{a}^2 \right)^{1/2}; \ \gamma_\| = \left(1 - \beta^2 \right)^{-1/2} , \tag{5}$$

and $\beta = v_z/c$.

The equations derived from this model are now the longitudinal component of Eq. (1), the equation of continuity, Poisson's equation, and the wave equation for $\mathbf{a}(z,t)$, which are given by

$$\frac{1}{c}\frac{\partial}{\partial t}\left(\gamma_a \sqrt{\gamma_\|^2 - 1} \right) + \frac{\partial}{\partial z}\left(\gamma_a \gamma_\| \right) = \frac{\partial \phi}{\partial z}; \ \varphi \equiv \frac{e\phi}{m_0 c^2} , \tag{6}$$

$$\frac{1}{c}\frac{\partial n}{\partial t} + \frac{\partial}{\partial z}\left(n\sqrt{1 - 1/\gamma_\|^2} \right) = 0 , \tag{7}$$

$$\frac{\partial^2 \varphi}{\partial z^2} = \frac{\omega_{p0}^2}{c^2}\left(\frac{n}{n_0} - 1 \right) , \tag{8}$$

$$c^2 \frac{\partial^2 \mathbf{a}}{\partial z^2} - \frac{\partial^2 \mathbf{a}}{\partial t^2} = \omega_{p0}^2 \frac{n}{n_0}\frac{\mathbf{a}}{\gamma_a \gamma_\|} . \tag{9}$$

Assuming a driving pulse of the form

$$\mathbf{a}(z,t) = \frac{1}{2}\mathbf{a}_0(\xi,\tau)e^{-i\theta} + \text{c.c.}, \tag{10}$$

where $\theta = \omega_0 t - k_0 z$, ω_0 and k_0 being the central frequency and wavenumber, $\xi = z - v_g t$, $v_g = \partial \omega_0 / \partial k_0$ is the group velocity and τ is a slow time-scale, such that

$$a_0^{-1}\frac{\partial^2 a_0}{\partial \tau^2} \ll \omega_0^2 .$$

Accounting for changes in the pump due to the plasma reaction, the wave equation becomes

$$\left[2\frac{\partial}{\partial\tau}\left(\mathrm{i}\omega_0 a_0 + v_\mathrm{g}\frac{\partial a_0}{\partial\xi}\right) + c^2\left(1 - v_\mathrm{g}^2/c^2\right)\frac{\partial^2 a_0}{\partial\xi^2} + 2\mathrm{i}\omega_0\left(\frac{c^2 k_0}{\omega_0} - v_\mathrm{g}\right)\frac{\partial a_0}{\partial\xi}\right]\mathrm{e}^{-\mathrm{i}\theta} + \mathrm{c.c.}$$

$$= \left[c^2 k_0^2 - \omega_0^2 + \frac{n}{n_0}\omega_\mathrm{p0}^2 > \gamma_a\gamma_\parallel\right]a_0\mathrm{e}^{-\mathrm{i}\theta} + \mathrm{c.c.}\,, \quad (11)$$

where ω_p0 is the plasma frequency of the unperturbed plasma. Equations (6), (7), (8), and (11) form the basic set for this problem in the 'envelope approximation'.

In the weak-pump, weakly relativistic regime, the solution has the structure of a wakefield growing inside the electromagnetic pulse and oscillating behind the pulse, with the maximum amplitude being reached inside the pulse. Using the quasi-static approximation, the time derivative can be neglected in the electron fluid equations, Eqs. (6) and (7), yielding the following constants:

$$\gamma_a\left(\gamma_\parallel - \beta_0\sqrt{\gamma_\parallel^2 - 1}\right) - \varphi = 1\,, \quad (12)$$

$$n\left(\beta_0\gamma_\parallel - \sqrt{\gamma_\parallel^2 - 1}\right) = n_0\beta_0\gamma_\parallel\,, \quad (13)$$

where $\beta_0 = v_\mathrm{g}/c$. The constants of integration have been chosen in such a way that

$$n = n_0,\ \gamma_\parallel = 1,\ \varphi = 0\,,$$

for

$$\gamma_a = 1, \qquad \left(|a_0|^2 = 0\right)\,. \quad (14)$$

Using Eqs. (10) and (13), the general system, Eqs. (6)–(9), can be written as two coupled equations, describing the evolution of the laser pulse envelope $\mathbf{a}_0$ and the scalar potential φ:

$$\frac{\partial^2\varphi}{\partial\xi^2} = \frac{\omega_\mathrm{p0}^2}{c^2}G\,, \quad (15)$$

$$2\mathrm{i}\omega_0\frac{\partial a_0}{\partial\tau} + 2c\beta_0\frac{\partial^2 a_0}{\partial\tau\partial\xi} + \frac{c^2\omega_\mathrm{p0}^2}{\omega_0^2}\frac{\partial^2 a_0}{\partial\xi^2} = -\omega_\mathrm{p0}^2 H a_0\,, \quad (16)$$

where

$$G = \frac{\sqrt{\gamma_\parallel^2 - 1}}{\beta_0\gamma_\parallel - \sqrt{\gamma_\parallel^2 - 1}}\,,\ H = 1 - \frac{\beta_0}{\gamma_a\left(\beta_0\gamma_\parallel - \sqrt{\gamma_\parallel^2 - 1}\right)}\,.$$

This set of non-linear equations, Eqs. (15) and (16), is obtained using a quasi-static approximation, which yields two integrals of the motion, given by Eqs. (12) and (13). The model is valid for electromagnetic pulses of arbitrary polarization and intensities $|a_0|^2 \geq 1$.

Equations (15) and (16) can be solved numerically in the stationary frame of the pulse. Equation (15), Poisson's equation for the wakefield, is solved with the initial conditions $\varphi = 0$, $\partial\varphi/\partial\xi = 0$ by a simple predictor–corrector method. The envelope equation, Eq. (16), describing the evolution of the laser pulse, is written as two coupled equations for the real and imaginary parts of a_0 and solved implicitly.

Numerical solutions of Eqs. (15) and (16) show the evolution of the excited plasma wakefield potential φ and electric field E_w, as well as the envelope of the laser pulse $|a_0|$ [35], all in one spatial dimension. Solutions for a typical laser and plasma configuration are shown in Fig. 1. There is significant distortion of the trailing edge of the laser pulse, resulting in photon spikes. The distortion occurs where the wake potential has a minimum and the density has a maximum. The spike arises as a result of the photons interacting with the plasma density inhomogeneity, with some photons being accelerated (decelerated) as they propagate down (up) the density gradient. This effect was predicted by John Dawson
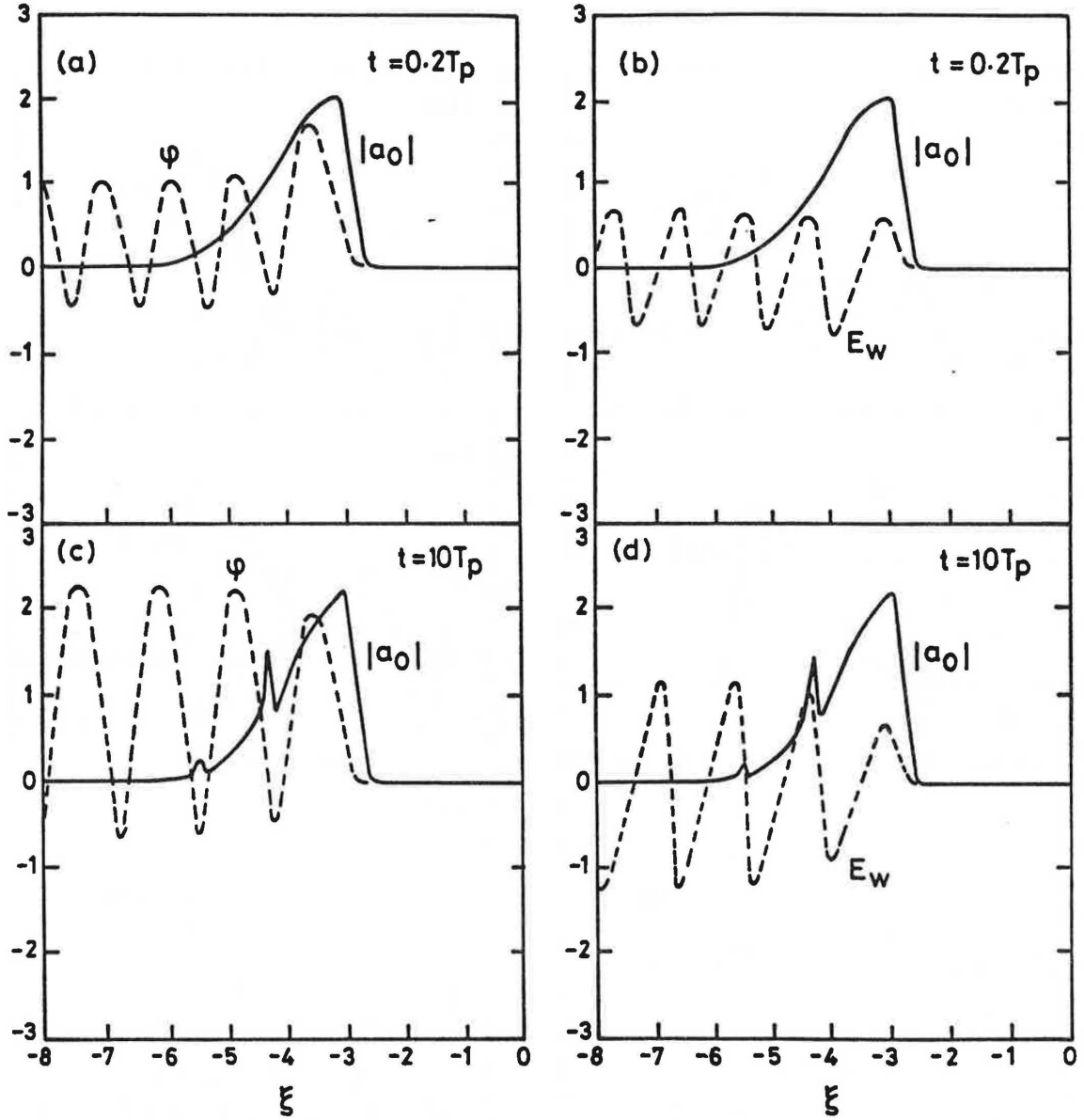
Fig. 1: The values of the magnitude of the normalized vector potential $|a_0|$ (solid curves) and scalar potential φ or wake electric field $E_{\rm w}$ (dashed curves) versus position $\xi = z - v_{\rm g}t$. Here, $|a_0^{\rm in} = 2|$, $\omega_{\rm p0}/\omega_0 = 1$, Gaussian rise $\sigma_{\rm r} = 0.25\lambda_{\rm p}$, Gaussian fall $\sigma_{\rm f} = 1.5\lambda_{\rm p}$. Curves (a) and (b) are at time $t = 0.2T_{\rm p}$; (c) and (d) are at $t = 10T_{\rm p}$.

and his group, and is called the photon accelerator [36]. The distortion of the trailing edge increases with increasing $\omega_{\rm p0}/\omega_0$. The longitudinal potential, $e\phi/(mc^2) > 1$ or $eE_z/(m_{\rm e}\omega_{\rm p0}c) > 1$, is significantly greater than for fields obtained in the plasma beat-wave accelerator. The field amplitude for the beat-wave accelerator is limited by relativistic detuning, while no such saturation exists in the laser wakefield accelerator.

When studying wakefields in more than one spatial dimension, the transverse dimensions of the driving laser pulse or particle beam become important, and the wakefield will assume a characteristic 'bubble' shape, provided the driver is sufficiently short. The bubble regime of plasma-based acceleration has been studied extensively using both analytical theory and full-scale numerical simulations [37–42]. A typical example of a three-dimensional laser-driven bubble-shaped wakefield can be seen in Fig. 2.
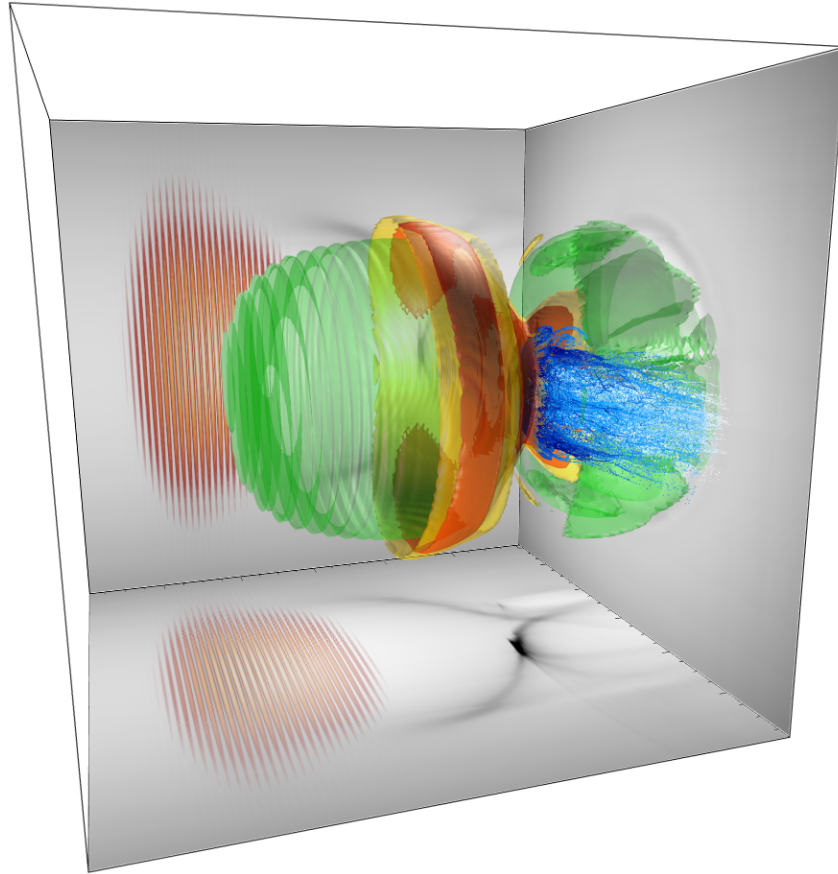
Fig. 2: An example of a three-dimensional wakefield generated by a short, intense laser pulse. Red and yellow, electromagnetic fields; green, electron density; blue, electrons trapped and accelerated by the wakefield. © J. Vieira, IST Lisbon, Portugal.

The electromagnetic fields are coloured red and yellow and the background plasma electron density is coloured green, while the population of electrons trapped and accelerated by the wakefield is coloured blue.

3   Photon acceleration

Photon acceleration is the phenomenon whereby photons interacting with a co-propagating plasma wave can increase or decrease their frequency, and thus their energy. The group speed of photons in plasma is given by $v_\mathrm{g} = c^2 k/\omega = c\sqrt{1 - \omega_\mathrm{p}^2/\omega_0^2} < c$. Thus, when the photon frequency increases, its group speed in plasma also increases, hence the photon accelerates. Photon acceleration [36, 43] is intimately related to short-pulse amplification of plasma waves and is described by a similar set of equations. It has been demonstrated that a relativistic plasma wave can be generated by a series of short electromagnetic pulses [44]. These pulses have to be spaced in a precise manner to give the plasma wave the optimum 'kick'. Conversely, if the second or subsequent photon pulses are put in a different position, 1.5 plasma wavelengths behind the first pulse, the second pulse will produce a wake that is $180°$ out of phase with the wake produced by the first pulse. The superposition of the two wakes behind the second pulse results in a lowering of the amplitude of the plasma wave. In fact, almost complete cancellation can take place. The second laser pulse has absorbed some or all of the energy stored in the wake created by the first pulse. From conservation of photon number density, the increase in energy of the second pulse implies that its frequency has increased. The energy in the pulse is $W = N\hbar\omega$, where N is the total number of

photons in the packet. Wilks *et al.* [36] have shown that the increase in frequency of the 'accelerated' wave packet exactly accounts for the loss in the energy of the accelerating plasma wave, given by

$$\frac{\delta\omega}{\delta x} = \frac{\omega_{\mathrm{pe}}^2 k_{\mathrm{p}}}{2\omega}\frac{\delta n}{n_0} \ . \tag{17}$$

Computer simulations by Wilks *et al.* [36] show that the laser pulse increased in frequency by 10% over a distance $237c/\omega_{\mathrm{p}}$.

Photon acceleration also plays a role in laser modulational instability [45–47]. This instability occurs when the energy of a long laser pulse is bunched longitudinally by a co-propagating plasma wave. The underlying process is periodic acceleration and deceleration of the laser's photons by the density fluctuations of the plasma wave. In turn, the laser's ponderomotive force helps to enhance the plasma wave, creating a positive feedback loop (see also Ref. [48] for the relationship between the ponderomotive force and the photon dispersion in plasma). The laser amplitude modulation due to this instability can be seen in Fig. 1, frames (c) and (d).

Photon acceleration of laser light in laser wakefield interactions was first observed by Murphy *et al.* [49]. The role of photon acceleration in the laser modulational instability was studied by Trines *et al.* [50,51], while aspects of the numerical modelling of wave kinetics have been investigated by Reitsma *et al.* [52,53].

Photon acceleration can be used as a single-shot diagnostic tool to probe laser- or beam-driven wakefields and determine their characteristics [36,49,50]. This is currently pursued within the framework of the AWAKE project [24]. Promising developments have been reported by Kasim *et al.* [54,55].

4 Relativistic self-focusing and optical guiding

In the absence of optical guiding, the interaction length L is limited by diffraction to $L = \pi R$, where R is the Rayleigh length $R = \pi\sigma^2/\lambda_0$, and σ is the focal spot radius. This limits the overall electron energy gain in a single stage to $E_{\mathrm{max}}\pi R$. To increase the maximum electron energy gain per stage, it is necessary to increase the interaction length. Two approaches to keeping high-energy laser beams collimated over a longer region of plasma are being developed. Relativistic self-focusing can overcome diffraction and the laser pulse can be optically guided by tailoring the plasma density profile, forming a plasma channel.

Relativistic self-focusing uses the non-linear interaction of the laser pulse and plasma results in an intensity-dependent refractive index to overcome diffraction. In regions where the laser intensity is highest, the relativistic mass increase is greatest; this results in a reduction of the fundamental frequency of the laser pulse. The reduction is proportional to the laser intensity. Correspondingly, the phase velocity of the laser pulse will decrease in regions of higher intensity. This has the effect of focusing a laser beam with a radial Gaussian profile. This results in the plane wavefront bending and focusing to a smaller spot size. Relativistic self-focusing has a critical laser power threshold, P_{cr}, which must be exceeded; this is given by [56]

$$P \geq P_{\mathrm{cr}} \sim 17\frac{\omega_0^2}{\omega_{\mathrm{p}}^2} \ \mathrm{GW} \ . \tag{18}$$

The laser must also have a pulse length (τ) that is shorter than both a collision period and an ion plasma period, to avoid the competing effects of thermal and ponderomotive self-focusing; this condition is $\tau_{\mathrm{L}} \leq 5 \times 10^9 \sqrt{Z/n_0}$ ps. The shape of the self-focused pulse for intense short pulses is also interesting; at the leading edge the non-linear response is not yet established, there is a finite time for the electrons to respond, and the front of the pulse propagates unchanged. The trailing edge of the pulse compresses radially, owing to the non-linear relativistic self-focusing. It is only the trailing edge of the pulse that is channelled. Another way of forming a guided laser is to perform a density cavity with one short-pulse laser forming a plasma density channel. Alternatively, the channel could be formed by a low-current electron beam, which produces the same effect.

One drawback of the short-pulse laser–plasma interaction is that, at the extreme intensities used, a hole in electron density can be created. This would prevent a wake from being created, or it might result in a very small wake, if the residual density were many orders of magnitude smaller than the original density.

5 Discussion

Plasma acceleration processes continue to be an area of active research. Initial studies of particle acceleration have proved fruitful for current drive schemes and laser accelerators. Particle acceleration in strongly turbulent plasmas is still in its infancy and requires a great deal more research. This area of research is important in astrophysical and space plasma.

Current and future experiments, however, are very far from the parameter range of interest to high-energy physicists, who require something like 10^{11} particles per pulse accelerated to teraelectronvolt energies (for electrons), with a luminosity of 10^{-34} cm^{-2} s^{-1} for acceptable event rates to be achieved. The teraelectronvolt energy range is more than 100 times greater than a single accelerating stage could provide at present; even if the interaction length could be extended by laser channelling, there would still be the requirement of multiple staging, and the need for more energetic lasers.

Researchers, realizing that the next collider will almost certainly be a linear electron–positron collider, are proposing a novel way of building such a device, known as the 'plasma afterburner' concept [57, 58]. Several groups are also developing an entirely new type of electron lens, using focusing by a plasma, to increase the luminosity of future linear colliders [59].

This plays on the fact that relativistic electron beams can be focused by a plasma if the collisionless skin depth c/ω_{pe} is larger than the beam radius. Generally, when a relativistic electron beam enters a plasma, the plasma electrons move to neutralize the charge in the beam on a $1/\omega_{\mathrm{pe}}$ time-scale. However, if the collisionless skin depth is larger than the beam radius, the axial return current flows in the plasma on the outside of the electron beam and the beam current is not fully neutralized, leading to the generation of an azimuthal magnetic field. Consequently, this self-generated magnetic field pinches or focuses the beam in the radial direction. This type of lens exceeds conventional lenses by several orders of magnitude in focusing gradient.

At present, laser-plasma accelerators are being used to accelerate electrons from hundreds of megaelectronvolts to gigaelectronvolts, and to create intense X-ray radiation via betatron motion of the accelerating electrons [26]. A future milestone to be achieved will be the 10 GeV energy level with good beam quality. Electrons in this energy range are ideal as a driver for free electron lasers; at the higher energies predicted, several gigaelectronvolts, it is possible to produce an X-ray free electron laser capable of biological investigations around the water window. Many other applications of plasma-based acceleration are also already being discussed [60].

References

[1] T. Tajima and J.M. Dawson, *Phys. Rev. Lett.* **43**(4) (1979) 267.
    http://dx.doi.org/10.1103/PhysRevLett.43.267

[2] D. Strickland and G. Mourou, *Opt. Comm.* **55**(6) (1985) 447.
    http://dx.doi.org/10.1016/0030-4018(85)90151-8

[3] D. Strickland and G. Mourou, *Opt. Comm.* **56**(3)(1985) 219.
    http://dx.doi.org/10.1016/0030-4018(85)90120-8

[4] W.B. Mori *et al.*, Proc. 1991 Particle Accelerator Conf. (1991), p. 564.

[5] J.B. Rosenzweig *et al.*, *Phys. Rev.* A **44**(10) (1991) R6189(R).
    http://dx.doi.org/10.1103/PhysRevA.44.R6189

[6] E. Esarey *et al.*, *IEEE Trans. Plasma Sci.* **24**(2) (1996) 252. http://dx.doi.org/10.1109/27.509991

[7] R. Bingham *et al.*, *Plasma Phys. Contr. Fusion* **46** (2004) R1.
http://dx.doi.org/10.1088/0741-3335/46/1/R01

[8] C. Joshi, *Phys. Plasmas* **14**(5) (2007) 055501. http://dx.doi.org/10.1063/1.2721965

[9] S.M. Hooker, *Nature Photonics* **7** (2013) 775. http://dx.doi.org/10.1038/nphoton.2013.234

[10] E. Esarey *et al.*, *Rev. Mod. Phys.* **81**(3) (2009) 1229.
http://dx.doi.org/10.1103/RevModPhys.81.1229

[11] S. Mangles *et al.*, *Nature* **431** (2004) 535. http://dx.doi.org/10.1038/nature02939

[12] C. Geddes *et al.*, *Nature* **431** (2004) 538. http://dx.doi.org/10.1038/nature02900

[13] J. Faure *et al.*, *Nature* **431** (2004) 541. http://dx.doi.org/10.1038/nature02963

[14] A. Pukhov and J. Meyer-ter-Vehn, *Appl. Phys.* B **74**(4–5) (2002) 355.
http://dx.doi.org/10.1007/s003400200795

[15] F. Tsung *et al.*, *Phys. Rev. Lett.* **93**(18) (2004) 185002.
http://dx.doi.org/10.1103/PhysRevLett.93.185002

[16] W. Leemans *et al.*, *Nature Phys.* **2** (2006) 696. http://dx.doi.org/10.1038/nphys418

[17] S. Kneip *et al.*, *Phys. Rev. Lett.* **103**(3) (2009) 035002.
http://dx.doi.org/10.1103/PhysRevLett.103.035002

[18] W. Leemans *et al.*, *Phys. Rev. Lett.* **113**(24) (2014) 245002.
http://dx.doi.org/10.1103/PhysRevLett.113.245002

[19] P. Chen *et al.*, *Phys. Rev. Lett.* **54**(7) (1985) 693. http://dx.doi.org/10.1103/PhysRevLett.54.693

[20] A. Caldwell *et al.*, *Nature Phys.* **5** (2009) 363. http://dx.doi.org/10.1038/nphys1248

[21] I. Blumenfeld *et al.*, *Nature* **445** (2007) 741. http://dx.doi.org/10.1038/nature05538

[22] M. Litos *et al.*, *Nature* **515** (2014) 92. http://dx.doi.org/10.1038/nature13882

[23] S. Corde *et al.*, *Nature* **524** (2015) 442. http://dx.doi.org/10.1038/nature14890

[24] R. Assmann *et al.*, *Plasma Phys. Contr. Fusion* **56**(8) (2014) 084013.
http://dx.doi.org/10.1088/0741-3335/56/8/084013

[25] W. Leemans *et al.*, *AIP Conf. Proc.* **1299** (2010) 3. http://dx.doi.org/10.1063/1.3520352

[26] S. Kneip *et al.*, *Nature Phys.* **6** (2010) 980. http://dx.doi.org/10.1038/nphys1789

[27] A. Sävert *et al.*, *Phys. Rev. Lett.* **115**(5) (2015) 055002.
http://dx.doi.org/10.1103/PhysRevLett.115.055002

[28] R.A. Fonseca *et al.*, *Lect. Not. Comp. Sci.* **2331** (2002) 342.
http://dx.doi.org/10.1007/3-540-47789-6_36

[29] R.A. Fonseca *et al.*, *Phys. Plasmas* **10**(5) (2003) 1979. http://dx.doi.org/10.1063/1.1556605

[30] R.A. Fonseca *et al.*, *Plasma Phys. Contr. Fusion* **50**(12) (2008) 124034.
http://dx.doi.org/10.1088/0741-3335/50/12/124034

[31] A. Pukhov, *J. Plasma Phys.* **61**(3) (1999) 425. http://dx.doi.org/10.1017/S0022377899007515

[32] C. Nieter and J.R. Cary, *J. Comp. Phys* **196**(2) (2004) 448.
http://dx.doi.org/10.1016/j.jcp.2003.11.004

[33] T. Arber *et al.*, *Plasma Phys. Contr. Fusion* **57**(11) (2015) 113001.
http://dx.doi.org/10.1088/0741-3335/57/11/113001

[34] P. Sprangle *et al.*, *App. Phys. Lett.* **53**(22) (1988) 2146. http://dx.doi.org/10.1063/1.100300

[35] R. Bingham *et al.*, *Plasma Phys. Contr. Fusion* **34**(4) (1992) 557.
http://dx.doi.org/10.1088/0741-3335/34/4/014

[36] S.C. Wilks *et al.*, *Phys. Rev. Lett.* **62**(22) (1989) 2600.
http://dx.doi.org/10.1103/PhysRevLett.62.2600

[37] S. Gordienko and A. Pukhov, *Phys. Plasmas* **12**(4) (2005) 043109.

http://dx.doi.org/10.1063/1.1884126

[38] W. Lu *et al.*, *Phys. Plasmas* **13**(5) (2006) 056709. http://dx.doi.org/10.1063/1.2203364

[39] W. Lu *et al.*, *Phys. Rev. ST Accel. Beams* **10**(6) (2007) 061301. http://dx.doi.org/10.1103/PhysRevSTAB.10.061301

[40] M. Tzoufras *et al.*, *Phys. Rev. Lett.* **101**(14) (2008) 145002. http://dx.doi.org/10.1103/PhysRevLett.101.145002

[41] M. Tzoufras *et al.*, *Phys. Plasmas* **16**(5) (2009) 056705. http://dx.doi.org/10.1063/1.3118628

[42] S. Martins *et al.*, *Nature Phys.* **6** (2010) 311. http://dx.doi.org/10.1038/nphys1538

[43] R. Bingham, *Phil. Trans. R. Soc.* A **366**(1871) (2008) 1749. http://dx.doi.org/10.1098/rsta.2007.2183

[44] D.A. Johnson *et al.*, *Physica Scripta* **T52** (1994) 77. http://dx.doi.org/10.1088/0031-8949/1994/T52/013

[45] J.F. Drake *et al.*, *Phys. Fluids* **17**(4) (1974) 778. http://dx.doi.org/10.1063/1.1694789

[46] C.E. Max *et al.*, *Phys. Rev. Lett.* **33**(4) (1974) 209. http://dx.doi.org/10.1103/PhysRevLett.33.209

[47] C.J. McKinstrie and R. Bingham, *Phys. Fluids* B **4**(8) (1992) 2626. http://dx.doi.org/10.1063/1.860178

[48] L.O. Silva *et al.*, *Phys. Rev.* E **59**(2) (1999) 2273. http://dx.doi.org/10.1103/PhysRevE.59.2273

[49] C. Murphy *et al.*, *Phys. Plasmas* **13**(3) (2006) 033108. http://dx.doi.org/10.1063/1.2178650

[50] R. M. G. M. Trines *et al.*, *Plasma Phys. Contr. Fusion* **51**(2) (2009) 024008. http://dx.doi.org/10.1088/0741-3335/51/2/024008

[51] R. M. G. M. Trines *et al.*, *Phys. Plasmas* **16**(5) (2009) 055904. http://dx.doi.org/10.1063/1.3125929

[52] A. J. W. Reitsma *et al.*, *Phys. Plasmas* **13**(11) (2006) 113104. http://dx.doi.org/10.1063/1.2366577

[53] A. J. W. Reitsma *et al.*, *Phys. Plasmas* **15**(2) (2008) 023105. http://dx.doi.org/10.1063/1.2834300

[54] M. Kasim *et al.*, *Phys. Rev. ST Accel. Beams* **18**(3) (2015) 032801. http://dx.doi.org/10.1103/PhysRevSTAB.18.032801

[55] M. Kasim *et al.*, *Phys. Rev. ST Accel. Beams* **18**(8) (2015) 081302. http://dx.doi.org/10.1103/PhysRevSTAB.18.081302

[56] P. Sprangle *et al.*, *IEEE Trans. Plasma Sci.* **15**(2) (1987) 145. http://dx.doi.org/10.1109/TPS.1987.4316677

[57] S. Lee *et al.*, *Phys. Rev. ST Accel. Beams* **5**(1) (2002) 011001. http://dx.doi.org/10.1103/PhysRevSTAB.5.011001

[58] D.L. Bruhwiler *et al.*, *Phys. Plasmas* **10**(5) (2003) 2022. http://dx.doi.org/10.1063/1.1566027

[59] G. Hairapetian *et al.*, *Phys. Rev. Lett.* **72**(15) (1994) 2403. http://dx.doi.org/10.1103/PhysRevLett.72.2403

[60] F. Albert *et al.*, *Plasma Phys. Contr. Fusion* **56**(8) (2014) 084015. http://dx.doi.org/10.1088/0741-3335/56/8/084015

Multidimensional Plasma Wake Excitation in the Non-linear Blowout Regime

J. Vieira[1], R.A. Fonseca[1,2] and L.O. Silva[1]

[1]GoLP/Instituto de Plasmas e Fusão Nuclear, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

[2]DCTI/ISCTE Lisbon University Institute, Lisbon, Portugal

Abstract

Plasma accelerators can sustain very high acceleration gradients. They are promising candidates for future generations of particle accelerators for several scientific, medical and technological applications. Current plasma-based acceleration experiments operate in the relativistic regime, where the plasma response is strongly non-linear. We outline some of the key properties of wakefield excitation in these regimes. We outline a multidimensional theory for the excitation of plasma wakefields in connection with current experiments. We then use these results and provide design guidelines for the choice of laser and plasma parameters ensuring a stable laser wakefield accelerator that maximizes the quality of the accelerated electrons. We also mention some of the future challenges associated with this technology.

Keywords

Plasma-based accelerators; Particle-In-Cell simulations; Laser-plasma interactions.

1 Introduction

In their seminal work published more than 30 years ago [1], Toshiki Tajima and John Dawson proposed the concept of the laser wakefield accelerator. Through theoretical calculations and computer simulations they showed that the radiation pressure of an intense laser pulse could drive large-amplitude plasma waves with a phase velocity identical to the group velocity of the laser driver, and characterized by large accelerating electric fields. These electric fields could then accelerate electrons to high energies in very short distances when compared with conventional accelerating devices. This work initiated research efforts that have also motivated the construction of some of the most powerful lasers available today [2].

Laser wakefield accelerators can be thought as energy transformers, converting the energy from the laser driver to the energy of plasma waves, and then converting the energy of the plasma waves to the accelerated particles. The laser acts on the plasma through the ponderomotive force, which expels particles from the regions where the laser fields are more intense to the regions of lower intensities. This force can also be interpreted as the radiation pressure that an intense laser exerts on background plasma electrons [3]. The displacement of background plasma electrons leads to large electrostatic fields due to the space–charge separation between the background plasma electrons and ions, which remain immobile. Typically, these plasma electrostatic fields scale with $E_0[\mathrm{Vcm^{-3}}] \simeq 0.96\sqrt{n_0\ [\mathrm{cm^{-3}}]}$, where n_0 is the background plasma density. Thus, typical plasmas densities of $n_0 \simeq 10^{18}$ cm^{-3} can lead to accelerating electric fields in excess of 1 GV cm^{-1} [4].

In order to excite these very high accelerating fields, current experiments typically use tightly focused, high-intensity and ultra-short laser pulses with transverse spot sizes smaller than $100\ \mu$m, intensities above $I \sim 10^{18}$ W cm^{-2} and pulse durations shorter than 100 fs. Lasers with these properties were not available when the laser wakefield accelerator was first proposed. Recently, however, with the advances in laser technologies, these lasers started to become widely available for experiments. The first

successful experiments capable of generating electron bunches with non-Maxwellian energy distributions were independently reported in 2004 by three experimental groups. These experiments [5–7] used $\sim$1 J lasers, focused to transverse spot sizes of $\sim$10 μm and compressed down to 30 fs. The lasers hit $\sim$1 mm long gas jets producing plasmas with densities on the order of 10^{19} cm^{-3}. After the gas jet, the experiments measured a $\sim$10% energy spread for 50–120 MeV electron bunches at distances ranging between 1 and 3 mm.

These experimental progresses took place even without the advanced theoretical and conceptual plasma-based acceleration framework that is currently available. The first theoretical and computational results for plasma wave excitation and electron acceleration were obtained in the one-dimensional (1D) limit, because the equations can be solved exactly in this limit. The 1D limit, however, presents fundamental limitations that are inherent to the reduced dimensionality. For instance, the motion of free plasma electron oscillations in one dimension is well described by an harmonic oscillator. In multidimensions, however, the plasma single electron trajectories are described by an anharmonic oscillator even for low-amplitude plasma waves [8]. As a result, the period of the oscillations depends on their amplitude of oscillation. Thus, in one dimension, if the background plasma electron flow is laminar during the first plasma oscillation it will remain laminar. On the contrary, the electron flow will inevitably become turbulent in multidimensions [8].

The electromagnetic structure of the plasma wakefield depends mainly on the amplitude of the plasma electron oscillations. Current plasma-based acceleration experiments operate in strongly non-linear regimes, for which the amplitude of the plasma electron radial displacement is much larger than its initial radial position. In this regime, plasma electron trajectories become non-laminar before the end of the first plasma oscillation leading to plasma wave-breaking [9, 10]. Although simplified analytical models exist to describe the structure of the wakefield, the full electromagnetic field structure of the plasma wave can only be captured through numerical simulations. Simulations have then been used to plan and to predict the experimental results, which is also essential to develop and confirm the predictions of the analytical models.

There are several numerical models that can describe plasma acceleration computationally, each employing different approximations. One of the most successful techniques is the particle-in-cell (PIC) method [11, 12]. In PIC codes, space is discretized into a grid that stores electric and magnetic fields. Each grid cell contains simulation macro-particles, each representing an ensemble of real electrons. It is then possible to advance the particles positions and momenta using the Lorentz force equation by interpolating the fields at the particles positions. Current densities, which are defined at the edges of the grid, are used to advance the fields through a set of discretized Maxwell's equations. As a result, PIC modelling retains the kinetic nature of the plasma dynamics, and can be employed to describe the plasma even in turbulent regimes where the flow of the plasma electrons is non-laminar.

There are several categories of PIC codes depending on the physics they retain/neglect. Reduced PIC codes use reduced versions of the Lorentz force (e.g., by neglecting relativistic effects), or reduced sets of Maxwell's equations (e.g., by neglecting magnetic fields). In spite of not retaining the full physics, these codes are typically very computationally effective, and allow for fast turn around simulation times when compared with full PIC codes. Full PIC codes employ almost no physical approximations. Particles advance under the relativistic Lorentz force, and the fields are updated with the full set of Maxwell's equations. Full PIC simulations are well suited to describe plasma accelerators in strongly non-linear regimes, but they are very computationally expensive, requiring the use of large super-computers. In order to optimize available computational infrastructures, and in order to better assist experimental design and interpretation, PIC algorithms have also seen significant advances. The most efficient PIC algorithms can currently achieve very high computing efficiencies from a few thousands up to more than a million cpu-cores [13]. One of the largest laser wakefield acceleration (LWFA) simulations sucessfuly ran in more than 200,000 cores, and enabled the simulation of fine details regarding the acceleration process, while reaching petascale sustainable performance in a production run [13].

Here we will review some of the key advances regarding wake excitation in the linear and the non-linear regimes, in one dimension and in multidimensions, complementing the analytical results with PIC simulation results. In Section 2 we will describe a theory for linear and non-linear plasma wave excitation in one-dimension, and also discuss the physics of beam loading in one dimension. In Section 3 we will derive a phenomenological theory for multidimensional, non-linear plasma waves, capable of describing current experiments. We will then derive an analytical theory for the blowout regime in Section 4 that predicts the full electromagnetic structure of the plasma wave. In Section 5 we will use the theoretical framework derived in Section 4 in order to explore the physics of beam loading in the blowout regime. Some of the outstanding challenges for plasma-based accelerators are described in Section 6 and in Section 7 we present the conclusions.

2 Non-linear plasma waves in the 1D limit

In order to introduce the key physical mechanisms of plasma acceleration, we will first explore the self-consistent generation of plasma waves by a laser pulse in non-weakly relativistic or linear regimes and in the 1D limit.

2.1 Relativistic fluid and Maxwell's equations in weakly relativistic regimes

In order to explore the dynamics in one dimension we consider the limit of wide laser pulses, with transverse spot-sizes much wider than the plasma wavelength. In addition, we will also consider the excitation of linear plasma waves characterized by small density perturbations when compared to the background plasma density. This excitation regime of the plasma waves, usually characterized by sinusoidal plasma density and electric field oscillations, corresponds to the so-called weakly relativistic or linear regime. Our starting point are Maxwell's equations written in the Coulomb gauge for the laser vector potential **A**, which read

$$\frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} + \nabla \times \nabla \times \mathbf{A} = \frac{4\pi}{c}\mathbf{J} - \frac{1}{c}\frac{\partial \nabla \phi}{\partial t}, \tag{1}$$

where t is the time, $\mathbf{J} = n\mathbf{v}$ is the total electric current, ϕ is the scalar (electrostatic) potential and c the speed of light. In addition, n is the local plasma density and $\mathbf{v}$ is the local plasma fluid velocity.

We can rewrite Eq. (1) in the direction of polarisation of a linearly polarized laser pulse. We then assume that the laser is polarized in y and propagates in x. The source term for Eq. (1) is the transverse plasma current driven by the laser fields, J_y. In order to determine J_y, we assume that the plasma ions remain fixed during the interaction. This is a valid approximation because the mass of background plasma ions (m_i) is much higher than the electron mass m_e. Thus, plasma ions barely move during a plasma electron oscillation. The plasma currents are then only due to the motion of plasma electrons. Conservation of canonical momentum for the plasma electrons in the y direction, assumed to be initially at rest, implies that $p_y = eA_y/(m_e c^2)$, where p_y is the electron momentum in the y direction, and where e and m_e are the electron charge and mass, respectively. Since $p_y = v_y \gamma$, the transverse plasma current driven by the laser is simply given by $J_y = nv_y = np_y/\gamma = nA_y/(m_e c\gamma)$. In the former expression, γ is the relativistic factor and v_y is the electron velocity in the y direction. Using this expression for J_y, Eq. (1) then becomes

$$\frac{1}{c^2}\frac{\partial^2 A_y}{\partial t^2} + \frac{\partial^2 A_y}{\partial x^2} = \frac{4\pi}{c}J_y - \frac{1}{c}\frac{\partial \phi}{\partial t} = -\frac{4\pi e^2}{m_e c^2}\frac{n}{\gamma}A_y, \tag{2}$$

where we have used the fact that in one dimension $\partial \phi / \partial y = 0$. We note that the conservation of canonical momentum is strictly valid for the case of plane electromagnetic waves. Thus, although valid exactly in 1D geometries, the conservation of canonical momentum does not hold exactly in multidimensions when the laser pulse has a finite transverse spot-size.

In order to further evaluate Eq. (2), we need to find an expression for the relativistic γ factor involving only known quantities. The relativistic factor γ of plasma electrons is given by

$$\gamma = \sqrt{1 + \frac{p_x^2}{m_e^2 c^2} + \frac{p_y^2}{m_e^2 c^2} + \frac{p_z^2}{m_e^2 c^2}}, \tag{3}$$

where p_x is the longitudinal electron momentum (in the x direction) and where p_z is the transverse electron momentum in z. Although plasma electrons can have transverse velocities in y and in z (as long as the laser has electric field components in y or in z), the motion of plasma electrons is only along x in one-dimension. Since we have assumed that the laser is polarized in y, only $v_y \neq 0$.

As with Eq. (2), we can rewrite the relativistic factor given by Eq. (3) using conservation of canonical momentum in y. Hence, since the laser is polarized in y, $p_y/(m_e c) = eA_y/(m_e c^2)$ and $p_z = 0$. Here it is worth noting that $p_y/(m_e c) = \gamma v_y$ is the proper electron velocity and that the quantity $eA_y/(m_e c^2) = a_y$ is the normalized vector potential. Thus, the conservation of canonical momentum indicates that the normalized laser vector potential, a_y, can be regarded as the momentum associated with the fast quiver motion of electrons on the laser fields. Therefore, the motion of the plasma electrons becomes relativistic when $p_y/(m_e c) = a_y \gtrsim 1$, and remains non- or weakly-relativistic for $a_y \ll 1$. Equation (3) can then be simplified by expanding the relativistic factor for small $p_x/m_e c \ll 1$ and small $a_y \ll 1$, yielding

$$\gamma = \sqrt{1 + \frac{p_x^2}{m_e^2 c^2} + \frac{e^2 A_y^2}{m_e^2 c^4}} \simeq 1 + \frac{1}{2}\frac{p_x^2}{m_e^2 c^2} + \frac{1}{2}\frac{e^2 A_y^2}{m_e^2 c^4}. \tag{4}$$

We can hence write the $1/\gamma$ factor appearing in the wave equation Eq. (2) as

$$\frac{1}{\gamma} \simeq 1 - \frac{1}{2}\frac{e^2 A_y^2}{m_e^2 c^4} - \frac{1}{2}\frac{p_x^2}{m_e^2 c^2}, \tag{5}$$

which after being inserted into the simplied wave Eq. (2) gives

$$\frac{1}{c^2}\frac{\partial^2 A_y}{\partial t^2} + \frac{\partial^2 A_y}{\partial x^2} \simeq -\frac{4\pi e^2}{m_e c^2}n\left(1 - \frac{1}{2}\frac{p_x^2}{m_e^2 c^2} - \frac{1}{2}\frac{e^2 A_y^2}{m_e^2 c^4}\right)A_y. \tag{6}$$

We can further simplify Eq. (6) by using the following ordering to obtain the 1D laser wave equation valid in non/weakly relativistic regimes, which we will confirm *a posteriori*:

$$\mathcal{O}\left(\frac{p_x}{m_e c}\right) \simeq \mathcal{O}\left(\frac{e^2 A_y^2}{m_e^2 c^4}\right) \simeq \mathcal{O}\left(1 - \frac{\delta n}{n_0}\right) \ll 1, \tag{7}$$

where the plasma electron density perturbations are defined according to $n = n_0 + \delta n$, where n_0 is the background plasma density and δn is a small perturbation. As a result, the wave equation Eq. (7) reduces to

$$\frac{1}{c^2}\frac{\partial^2 A_y}{\partial t^2} + \frac{\partial^2 A_y}{\partial x^2} \simeq -\frac{\omega_{p0}^2}{c^2}\left(1 + \frac{\delta n}{n_0} - \frac{1}{2}\frac{e^2 A_y^2}{m_e^2 c^4}\right)A_y. \tag{8}$$

Equation (8) has two unknowns, A_y and δn. In order to close the model we now need an equation for δn. In order to determine an expression for δn, we consider the linearized continuity equation for the plasma electrons and its time derivative, which read

$$\frac{\partial \delta n}{\partial t} + n_0 \nabla \cdot (\delta n \mathbf{v}) = 0 \Rightarrow \frac{\partial^2 \delta n}{\partial t^2} + n_0 \nabla \cdot \frac{\partial \mathbf{v}}{\partial t} = 0, \tag{9}$$

where $\mathbf{v} \ll c$ is given by the Lorentz force equation for non/weakly relativistic regimes

$$\frac{\partial \mathbf{v}}{\partial t} = -\frac{e}{m_e}\mathbf{E} - c^2 \nabla \left(1 + \frac{1}{2}\frac{e^2 A_y^2}{m_e^2 c^4}\right), \tag{10}$$

where $\mathbf{E}$ is the electric field and where the second term on the right-hand side of Eq. (10) is the laser ponderomotive force. Equation (10) neglects the magnetic field force component, a valid assumption because $\mathbf{v} \ll c$. Note that we have assumed that $\gamma = 1$ in Eq. (10). Inclusion of the higher order terms of Eq. (4) for the relativistic factor of plasma electrons in Eq. (10) would also lead to higher order relativistic corrections to the Lorentz force. Taking the divergence of Eq. (10), inserting the resulting expression into the continuity equation and using Gauss's law $\nabla \cdot E \simeq \delta n$ then gives

$$\left(\frac{\partial^2}{\partial t^2} + \omega_p^2\right)\frac{\delta n}{n_0} = \frac{1}{2}\frac{e^2}{m_e^2 c^2}\nabla A_y^2. \tag{11}$$

Equation (11) closes the 1D model for the self-consistent laser-plasma interaction valid in non/weakly relativistic regimes and in one dimension. We note that Eq. (11) confirms the ordering given by Eq. (7) in that $\mathcal{O}\left(\delta n/n_0\right) \simeq \mathcal{O}\left(A_y^2\right)$. Equation (11) is a forced harmonic oscillator excited by the laser radiation pressure (ponderomotive force). The laser expels electrons from regions of maximum laser intensity to regions of lower laser intensity. Background plasma ions provide a restoring force that attracts plasma electrons back to their initial positions. The generation of plasma waves in one dimension then occurs as follows: the laser starts by pushing plasma electrons forward. Since the ions remain fixed, a space charge electrostatic field develops, pushing the plasma electrons back to their original position. When they return to their original longitudinal position, the plasma electrons have a longitudinal backward velocity. They will then continue to move backwards with respect to their initial positions. A space charge force due to the background plasma ions forms again, pulling the plasma electrons forwards. This forms a plasma oscillation. The natural frequency of oscillation is the plasma frequency ω_p.

2.2   1D wakefield excitation by ultra-short lasers in non-linear regimes

After having introduced the excitation of plasma waves in weakly relativistic regimes, we now generalize the 1D model given by Eqs. (8) and (11) to the relativistic and strongly non-linear regimes in the limit of short lasers compared with the plasma period.

It is useful to include dimensionless quantities. Plasma electric fields are then normalized to the cold wave-breaking limit (E_{wb}). The cold wave-breaking limit corresponds to the maximum amplitude that a plasma wave supports in 1D limit and in the non-relativistic regime. Magnetic fields are normalized to the cold wave-breaking limit multiplied by the speed of light c (B_0), scalar and vector potentials normalized to the electron rest energy divided by the elementary charge (ϕ_0), and space and time are normalized to the plasma skin depth (d_0) and inverse plasma frequency t_0, respectively. In addition, momentum is normalized to $m_e c$ and energy to $m_e c^2$. In practical units, dimensionless quantities are then expressed as

$$E_{\mathrm{wb}} = \frac{m_e c \omega_p}{e} \simeq 5.64 \times 10^4 \sqrt{n_0[\mathrm{cm}^{-2}]}\,\mathrm{V/cm}, \tag{12}$$

$$B_0 = \frac{m_e c^2 \omega_p}{e} \simeq 32\sqrt{n_0[\times 10^{16}\,\mathrm{cm}^{-3}]}\,\mathrm{T}, \tag{13}$$

$$\phi_0 \simeq \mathbf{A}_0 = \frac{m_e c^2}{e} \simeq \frac{0.511\,[\mathrm{MeV}]}{e}, \tag{14}$$

$$d_0 = \frac{1}{k_p} \simeq \frac{5.32\,\mu\mathrm{m}}{\sqrt{n_0[10^{18}\,\mathrm{cm}^{-3}]}}, \tag{15}$$

$$t_0 = \frac{1}{\omega_p} \simeq \frac{17\,\mathrm{fs}}{\sqrt{n_0[10^{18}\,\mathrm{cm}^{-3}]}}. \tag{16}$$

Our starting point is the master equation for the momenta of plasma electrons [14, 15]:

$$\frac{\partial^2 \mathbf{p}}{\partial t^2} + c^2 \nabla \times \nabla \times \mathbf{p} = -\left[\omega_p^2 + \frac{1}{m_e}\nabla \cdot \left(\frac{\partial \mathbf{p}}{\partial t} + m_e c^2 \nabla \gamma\right)\right]\frac{\mathbf{p}}{\gamma} - m_e c^2 \frac{\partial \nabla \gamma}{\partial t}. \tag{17}$$

For a linearly polarized laser in y, and for plasma motion in x, the relevant components of Eq. (17) are in *(x,y)* and read

$$\frac{\partial^2 p_x}{\partial t^2} + \left(1 + \frac{\partial^2}{\partial t \partial x}p_x + \frac{\partial^2 \gamma}{\partial x^2}\right)\frac{p_x}{\gamma} + \frac{\partial^2}{\partial t \partial x}\gamma = 0, \tag{18}$$

$$\frac{\partial^2 p_y}{\partial t^2} - \frac{\partial^2 p_y}{\partial x^2} + \left(1 + \frac{\partial^2}{\partial t \partial x}p_x + \frac{\partial^2 \gamma}{\partial x^2}\right)\frac{p_y}{\gamma} = 0. \tag{19}$$

Since we are interested in the laser and plasma dynamics in the region that moves with the laser pulse we will also adopt the speed of light variables. This is a Galilean coordinate transformation to a frame that moves with the laser at c. We then transform time and space according to

$$\psi = ct - x, \tag{20}$$

$$\tau = x/c, \tag{21}$$

where ψ is a measure of the distance to the front of the laser and τ the time (or distance if multiplied by c) travelled by the laser pulse. The speed of light variables allow for the separation of the fast spatial scales associated with the variations in ψ and that occur at the plasma wavelength λ_p, from the slow laser temporal evolution associated with variations in τ and that scale with $\omega_0/\omega_p \gg 1$. The latter scaling can be understood by noting that the Rayleigh length $k_p Z_r$, which defines the typical time (distance) for the laser to evolve, is proportional to ω_0/ω_p. In typical laser wakefield acceleration scenarios, where the plasma is transparent to the laser, $k_p Z_r \propto \omega_0/\omega_p \gg 1$. Hence, $\partial_\tau \simeq \omega_p/\omega_0 \ll \partial_\psi \simeq 1/\lambda_p$. We can therefore neglect ∂_τ in comparison to ∂_ψ. This is also referred to as the Quasi-Static Approximation (QSA) [16]. The QSA is valid as long as the laser pulse envelope does not evolve in the time it takes for an electron to go across the laser pulse. Under the QSA, the master equation components in x and in y given by Eqs. (18) and (19) become

$$\frac{\partial^2 p_x}{\partial \psi^2} + \frac{p_x}{\gamma}\left(1 - \frac{\partial^2 p_x}{\partial \psi^2} + \frac{\partial^2 \gamma}{\partial \psi^2}\right) - \frac{\partial^2 \gamma}{\partial \psi^2} \simeq 0, \tag{22}$$

$$2\frac{\partial^2 p_y}{\partial \psi \partial \tau} + \frac{p_y}{\gamma}\left(1 - \frac{\partial^2 p_x}{\partial \psi^2} + \frac{\partial^2 \gamma}{\partial \psi^2}\right) \simeq 0. \tag{23}$$

Defining $\gamma - p_x = \chi$, we then arrive at the following simplified set of equations describing the coupled laser-plasma evolution and that can be used to describe wakefield excitation even in strongly non-linear regimes in one dimension:

$$\left(\frac{p_x}{\gamma} - 1\right)\frac{\partial^2 \chi}{\partial \psi^2} = -\frac{p_x}{\gamma}, \tag{24}$$

$$2\frac{\partial^2 p_y}{\partial \tau \partial \psi} + \left(1 + \frac{\partial^2}{\partial \psi^2}\right)\frac{p_y}{\gamma} = 0. \tag{25}$$

Equation (24) describes the wakefield evolution and Eq. (25) the evolution of the laser (recall that $p_y = a_y$ due to the conservation of canonical momentum).

Equations (24) and (25) have three unknowns, p_y, p_x and χ. In order to close the model we therefore need an additional relation between these quantities. This relation can be found by integrating Euler's equation. Since plasma waves are electrostatic in one dimension, Euler's equation becomes

$$\frac{\mathrm{d}p_x}{\mathrm{d}t} = -E_x - \frac{\partial \gamma}{\partial x} = \frac{\partial(\phi - \gamma)}{\partial x}, \tag{26}$$

where we have used $E_x = -\partial_x \phi$ in Eq. (26). Recasting Eq. (26) into the speed of light variables gives

$$-\frac{\partial}{\partial \psi} (\gamma - p_x - \phi) = \frac{\partial}{\partial \tau} (\phi - \gamma) \simeq 0, \tag{27}$$

where we have used $\partial_\tau \simeq 0$. As a result, for an electron initially at rest,

$$\chi = \gamma - p_x = 1 + \phi. \tag{28}$$

Using the constant of motion given by Eq. (28), we can now relate p_x, p_y and χ according to the following expressions:

$$p_x = \frac{1 + p_y^2 - \chi^2}{2\chi} \tag{29}$$

$$\gamma = \frac{1 + p_y^2 + \chi^2}{2\chi}. \tag{30}$$

Equations (29) and (30) can be used to close the non-linear model for the excitation of plasma waves given by Eqs. (24) and (25) as in [16]:

$$\frac{\partial^2 \chi}{\partial \psi^2} = -\frac{1}{2} \left(1 - \frac{1 + p_y^2}{\chi^2} \right), \tag{31}$$

$$2\frac{\partial^2 p_y}{\partial \tau \partial \psi} + \frac{p_y}{\chi} = 0, \tag{32}$$

where $p_y = a_y$ is due to the conservation of the canonical momentum. Equations (31) and (32) are a system of non-linear coupled equations that describe wake excitation by ultra-intense and ultra-short laser pulses, valid in one dimension. They can also be referred to as 1D quasi-static equations. Equation (31) describes wakefield excitation and Eq. (32) describes the laser evolution. The source term for the laser evolution is $1/\chi$, and it is possible to show that χ is closely related to the plasma susceptibility. This can be shown with the help of the constant of motion given by Eq. (28) and by considering the continuity equation, which reads:

$$\frac{\partial n}{\partial t} + \frac{\partial (nv_x)}{\partial z} = 0 \Rightarrow \frac{\partial [n(1 - v_x)]}{\partial \psi} = \frac{\partial n}{\partial \tau}. \tag{33}$$

Under the QSA Eq. (33) can be integrated yielding $n(1 - v_x) = n_0$. It is now possible to recover the 1D linear wakefield excitation model given in the previous sub-section by assuming $p_y \ll 1$ and $\chi \sim 1$.

The quasi-static equations (31) and (32) may be integrated analytically for specific laser pulse shapes assuming that the laser pulse remains unchanged. For a square shaped laser profile for instance, it is possible to show that the maximum wake potential ϕ_{max} and maximum longitudinal momentum p_{max} of a plasma electron are respectively given by [17]

$$\phi_{max} \simeq \gamma_\perp^2 - 1, \tag{34}$$

$$p_{max} \simeq \frac{\gamma_\perp^4 - 1}{2\gamma_\perp^2}, \tag{35}$$

and that the maximum electric field of the plasma wave is

$$E_{max} \simeq \frac{\gamma_\perp^2 - 1}{\gamma_\perp}, \tag{36}$$

where $\gamma_\perp = \sqrt{1 + a_y^2}$. Equations (34), (35) and (36) show that higher intensity lasers lead to higher amplitude plasma waves with higher electron energy gain within the plasma wave. These equations

cannot be used to explore particle trapping of background plasma electrons in the plasma wave (self-injection) because the QSA would no longer be valid in this case. Particle trapping occurs when the velocity of the plasma electrons matches the phase velocity of the plasma wave, which corresponds to the laser group velocity. For lasers with sufficiently high intensity, the velocity of the plasma electrons can be larger than the laser group velocity at the back of the first plasma wave. This leads to plasma wave-breaking. In the fluid approach that we have taken, wave-breaking leads to a singularity in the plasma density. For very high electron velocities, $\gamma \simeq p_x$, thus according to Eq. (28), $\phi \to -1$. Since $1/(1+\phi) = 1/\xi = n/\gamma$, then, as $\phi \to -1$, $n \to \infty$. In a kinetic description, wave-breaking corresponds to the crossing of electron trajectories. In relativistic regimes, the wave-breaking electric field limit is $eE_{\mathrm{wb}}/(m_{\mathrm{e}}c\omega_{p0}) = \sqrt{2}\sqrt{\gamma_\phi - 1}$, where v_ϕ is the wake phase velocity and $\gamma_\phi = 1/\sqrt{1 - v_\phi^2}$ is the corresponding relativistic factor.

Key properties of 1D plasma wakefields excited by a short laser pulse in strongly non-linear regimes are shown in Fig. 1. Figure 1 illustrates numerical solutions to the 1D QSA equations given by Eqs. (31) and (32). Figure 1(a) shows the envelope profile of the laser pulse. The electrostatic potential, shown in Fig. 1(b) is strongly non-linear, i.e. ϕ is not well described by sinusoidal oscillations at λ_{p} containing additional high-order harmonics. In addition, the minimum ϕ approaches $\phi \to -1$, as predicted theoretically using the constant of motion. Another distinctive feature of non-linear wake excitation is illustrated in Fig. 1, which shows a typical sawtooth shape for the accelerating electric field. The plasma susceptibility, and thus the plasma density, also reach very high values at the end of the first plasma wave, where the longitudinal electron momentum is at a maximum, which is also consistent with the theoretical arguments provided in the previous paragraph.
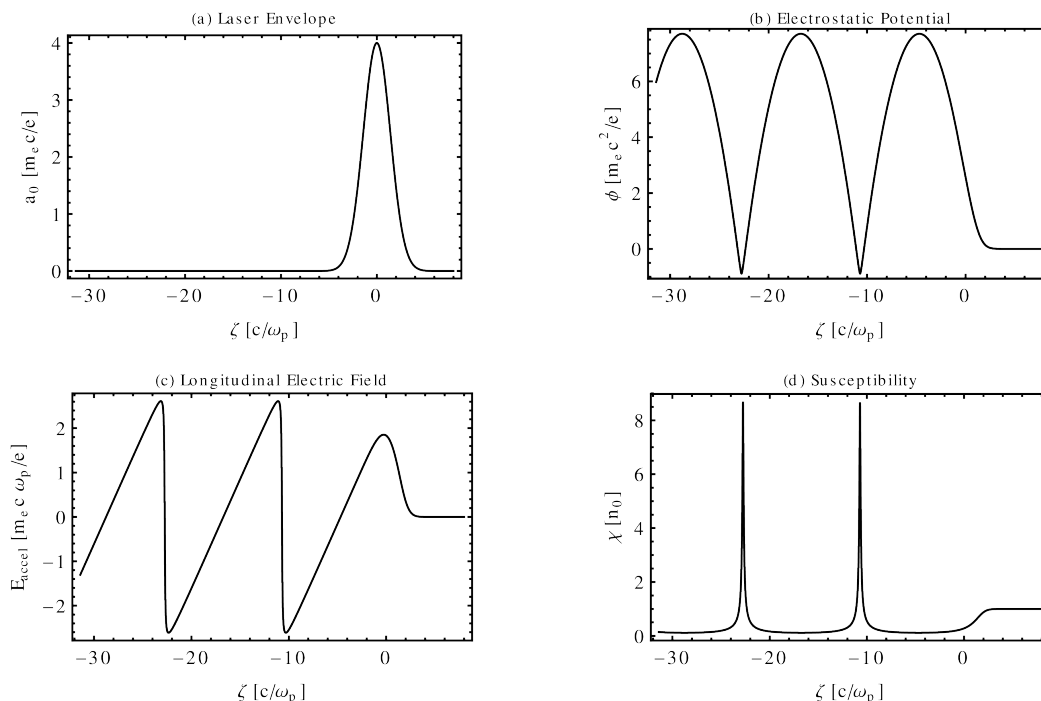
Fig. 1: Wakefield generation by an intense laser beam driver in one dimension: (a) shows the envelope of the laser vector potential; (b) shows the plasma electrostatic potential; (c) shows the longitudinal (accelerating) electric field; and (d) shows the plasma susceptibility.

It is important to note that despite not being able to capture particle trapping in plasma waves, the QSA can be used to explore wakefield excitation and the acceleration of externally injected particle bunches in the plasma wave fields. The computational advantages of the QSA when compared to kinetic

descriptions (such as the PIC method), have led to the development of reduced numerical models. These reduced models are used to predict the output of plasma-based experiments. Examples of such codes are QuickPIC [12], WAKE [18] and HiPACE [19].

2.3 Beam loading in the linear regime

In order to explore the acceleration of particle bunches to maximize the charge that can be accelerated and to minimize the final energy spreads, it is important to investigate how external charged particle bunches influence the wakefield structure. This research topic is typically called beam loading. The study of beam loading in plasma waves is thus important to shape the current profile of the particle bunches for maximizing the quality of the accelerated particles which is critical to potential applications. In this section, we will outline important results on the beam loading of electron bunches in one dimension and in the linear regime.

In Fig. 2 the optimal beam loading conditions in the linear regime and in one dimension are illustrated. Figure 2(a) shows a portion of the initial accelerating field driven by a laser pulse driver in the linear regime. Figure 2(b) shows the wakefield contribution driven by a witness electron bunch with a triangular shape. The resulting wakefield (Fig. 2(c)) shows that the acceleration gradient is constant throughout the entire witness bunch. The witness bunch then decreases the absolute value of the electric field in the beam region, thereby reducing the acceleration gradient. As a result, all beam electrons will accelerate with similar accelerating fields. This configuration then preserves the initial witness bunch energy spread, yielding a scenario of ideal beam loading.

Although a more refined beam loading theory can be derived in the linear regime for witness bunches with arbitrary shapes, we will adopt a simple model assuming a very short, uniform density witness bunch in comparison to the plasma wavelength.

We assume that the bunch density is n_b. An estimate for the maximum number of particles (N_0) that can be loaded into the wakefield in one dimension and in the linear regime can be found by assuming that the total electric field vanishes at the location of the beam. This condition leads to [20]

$$N_0 = 5 \times 10^5 \left(\frac{n_b}{n_0}\right) \sqrt{n_0} A, \tag{37}$$

where A is the area of the transverse bunch section and n_b is the trailing bunch density. The efficiency is 100% if the total number of particles loaded into the wakefield matches Eq. (37). However, this comes at the expense of obtaining 100% energy spread because the front of the beam accelerates at the maximum gradient, while the back of the beam does not accelerate. Since the energy gain $\Delta\gamma \propto E_{accel}$ (where E_{accel} is the accelerating gradient) the final energy spread when the number of particles (N) in the bunch is lower than N_0 at:

$$\frac{\Delta\gamma_{max} - \Delta\gamma_{min}}{\Delta\gamma_{max}} = \frac{E_i - E_f}{E_i} = \frac{N}{N_0}, \tag{38}$$

where $\Delta\gamma_{max/min}$ are the maximum/minimum energy gains by the bunch particles and $E_{i/f}$ are the accelerating electric fields at the front/behind of the accelerating bunch. Equation (38) illustrates the trade-off between number of accelerated particles and final energy spread. Hence, maximizing the accelerated charge also maximizes the energy spread and *vice versa*.

Another key feature related with beam loading is the fraction of the energy absorbed by the trailing particle bunch in the wakefield. Since the energy of the wakefield scales with $\propto E_{accel}^2$, the fraction of energy absorbed by the witness particle bunch is $1 - E_f^2/E_i^2$. Using Eq. (38), $E_f = E_i(1 - N/N_0)$. Thus, the energy conversion efficiency from the wakefield to a trailing bunch of particles is

$$\eta_b = \frac{N}{N_0}\left(2 - \frac{N}{N_0}\right). \tag{39}$$

Equations (38) and (39) also illustrate the trade-off between the beam loading efficiency and the final bunch energy spread. Small energy spreads require smaller total beam charges which leads to lower energy conversion efficiencies. Higher energy conversion efficiencies however lead to higher energy spreads. Although these conclusions are generally valid, it is possible to derive conditions for optimal beam loading using shaped current density profiles, as Fig. 2 shows.
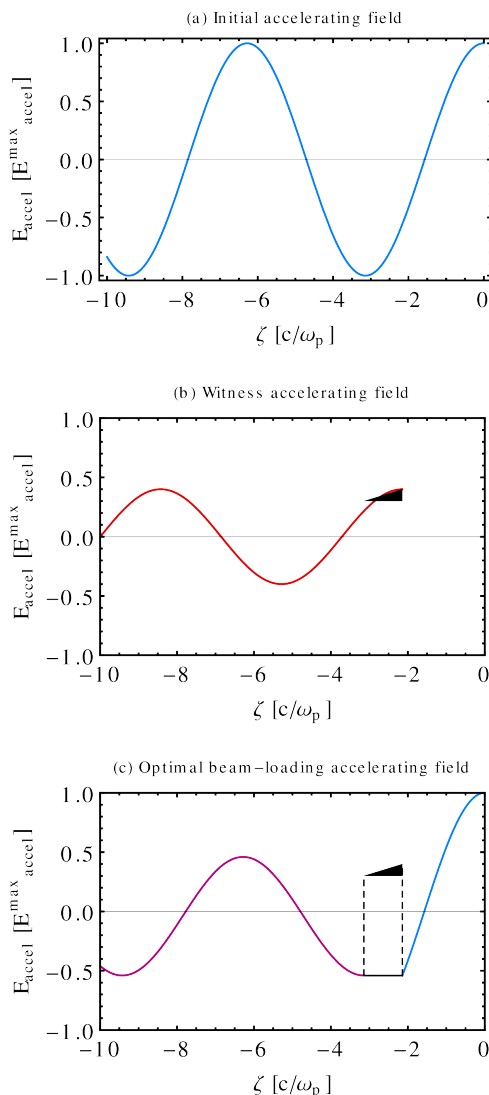
Fig. 2: Illustration of optimal beam loading conditions to minimize the energy spread of an energetic electron beam (dark triangle): (a) shows the accelerating gradient of a plasma wave driven by a laser pulse in the linear regime; (b) shows the accelerating wakefield driven by a witness electron bunch; and (c) shows the accelerating wakefield resulting from the combination (sum) of wakefield in (a) and in (b).

3   Strongly non-linear plasma waves in multidimensions: the blowout regime

The wakefield excitation models derived in the previous Section, valid in one dimension, are useful in describing laser wakefield accelerators qualitatively. However, because in typical experiments the

driver has transverse dimensions smaller or comparable to the plasma wavelength, the 1D theory cannot be employed to describe and predict the experimental outputs quantitatively. In typical experiments, the structure of the wakefields is inherently multidimensional. An accurate understanding of the experiments therefore requires the inclusion of multidimensional effects in both the theory and simulations.

Multidimensional plasma oscillations are intrinsically non-linear and anharmonic even for low-amplitude plasma waves [8]. The period of plasma electron oscillations then depends on their amplitude of oscillation. Although for low-amplitude plasma waves anharmonic effects can be neglected for the first few plasma waves, the cumulative phase shift between close electrons becomes substantial when the number of plasma oscillations increases. As a result, the trajectory of adjacent electrons will always cross in multidimensional plasma waves when close electrons become $\pi/2$ out of phase. Trajectory crossing inevitably leads to wave-breaking, where the initial laminar electron flow becomes turbulent. This is in stark contrast with the 1D theory developed in the previous sections where the flow remains laminar as long as it is laminar during the first plasma oscillations. When trajectory crossing occurs and the electron flow becomes turbulent, the plasma fluid equations used in the previous Section cease to be valid. The properties of the wakefield in these scenarios can only be captured by kinetic descriptions.

When the wakefield amplitude is large, sheath crossing can occur during the first plasma wave. Current experiments operate in this strongly non-linear regime. In this section, we will outline the derivation of a non-linear model for the wakefield excitation in strongly non-linear regimes and in multidimensions which is capable of describing and predicting experimental outputs.

Although we will focus the analysis on laser wakefield acceleration, general conclusions from this section are also valid for non-linear wakefields driven by electron beam drivers because the main properties of the wakefield in strongly non-linear regimes are nearly independent of the nature of the driver. Positron beams can also be used to drive strongly non-linear plasma waves and many of the findings outlined in this section may also be applicable to this case. However, we note that the wakefield generation process can differ significantly from that of lasers or electron bunches. While electron beams or laser pulses repel plasma electrons from the region of the driver, positrons attract plasma electrons towards the axis [24].

Figure 3 shows a particle-in-cell OSIRIS [11, 13] simulation results of a strongly non-linear wakefield excitation in multidimensions driven by an intense electron bunch (Fig. 3(a)), laser pulse (Fig. 3(b)) and positron bunch (Fig. 3(c)). A distinctive signature corresponding to the excitation of strongly non-linear plasma waves in multidimensions is electron cavitation, i.e. the generation of a region void of plasma electrons. This regime is called the blowout [21] or bubble regime [23] when electrons (Fig. 3(a)) or laser pulse (Fig. 3(b)) drivers are employed, and suck-in regime when positron bunch drivers (Fig. 3(c)) are used [24]. Although the wakefields share many similarities, and the main properties of the wakefield are driver independent, Fig. 3 shows that finer details of the blowout region depends on the nature of the driver. Figures 3 also show electron bunches [in the second bucket in Fig. 3(a) and Fig. 3(c) and in the first and second buckets in Fig. 3(b)] in conditions to be accelerated by the plasma.

For the remainder of this report, we define the transverse size of the laser as the laser spot-size W_0. The laser duration is τ_L. The peak vector potential a_0 is related to the laser intensity through $a_0 \simeq 0.8 \left(\lambda_0/1[\mu m]\right) \left[I/\left(10^{18} \mathrm{W\,cm^{-2}}\right)\right]^{1/2}$, where $\lambda_0 = 2\pi c/\omega_0$ is the central laser wavelength and ω_0 is the central laser frequency.

When an intense laser or electron bunch driver propagates through a plasma it radially expels plasma electrons away at its passage (see Fig. 3). If the laser intensity (electron driver density) is sufficiently high, the driver expels nearly all plasma electrons away from the region in which the driver propagates. Electron trajectories cross and accumulate in a thin, high-density electron layer that surrounds an electron void. Ions remain stationary and push plasma electrons back to the axis after the driver has propagated a distance close to $\sim \lambda_p$. At the back of the plasma wave, the shape of the thin electron layer resembles a sphere or a bubble.
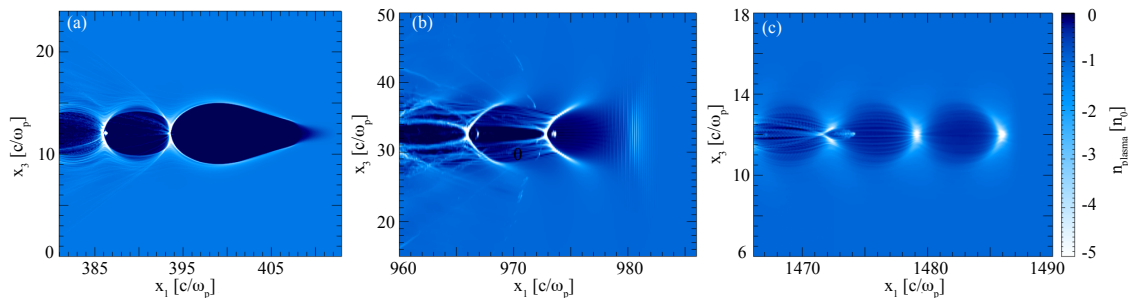
Fig. 3: OSIRIS simulation results illustrating the generation of strongly non-linear plasma waves. The colours are proportional to the electron plasma density. The driver moves from left to right: (a) shows the wakefield driven by an ultra-relativistic particle beam driver; (b) shows the generation of wakefields driven by intense laser beams; and (c) shows the plasma density perturbation associated with a positron bunch driver.

At the back of the sphere, there are strong electron accelerating fields which can be sufficiently strong to capture a fraction of the background plasma electrons into the bubble (self-injection). These electrons can then be focused and accelerate to high energies in the focusing and accelerating fields of the bubble. In the laser wakefield accelerator, reaching these strongly non-linear regimes requires normalized laser vector potentials close to $a_0 \gtrsim 2$ for spot-sizes (W_0) of a few skin-depths. In the plasma (beam driven) wakefield accelerator, the blowout regime requires that particle bunch densities are larger than the background plasma density.

3.1 Phenomenological model for the blowout regime

In order to describe electron acceleration in non-linear laser driven wakefields we first focus on a phenomenological model for the blowout regime driven by laser pulses [25]. Our first goal is first to derive a set of scaling laws for stable laser wakefield acceleration and which are capable of predicting key output beam properties such as energy and charge.

Important wakefield properties are defined by the focusing and accelerating fields. Focusing fields are absent from 1D descriptions but are crucial to the prediction of the outputs of the experiments. The focusing field for a relativistic charged particle traveling at c accelerating in the blowout region is then given by

$$W_\perp = E_r - B_\theta = \frac{r}{2}, \tag{40}$$

where r is the distance to the axis, E_r is the radial electric field and B_θ the azimuthal magnetic field. Equation (40) shows that the field is always focusing for electrons. Linear focusing fields are a feature of strongly non-linear regimes which are absent from wakefield excitation in the linear regime. As the electrons accelerate, they perform transverse harmonic oscillations (also called betatron oscillations). This is important to preserve beam emittance as the acceleration progresses. Achieving small (as small as possible) beam emittances is crucial for high-energy physics and radiation applications. Hence linear focusing fields are crucial for potential applications of plasma accelerators. In the linear regime, the focusing fields are non-linear with high-order contributions in powers of r. Thus, unless the electron bunches are much narrower than the driver spot-size so that they are subject to linear focusing fields near the axis, they will perform anharmonic oscillations, which will lead to higher final bunch emittances.

The accelerating field determines the maximum energy of a witness electron bunch. The accelerating electric field felt by the bunch electrons in the bubble is given by

$$E_{\mathrm{accel}} = \frac{\xi}{2}, \tag{41}$$

where ξ measures the distance to the centre of the bubble which moves nearly at the laser group velocity. The field accelerates electrons for $\xi < 0$ and decelerates them for $\xi > 0$. The linear focusing and accelerating fields are well reproduced in simulations of the laser wakefield accelerator, as shown in Fig. 4.
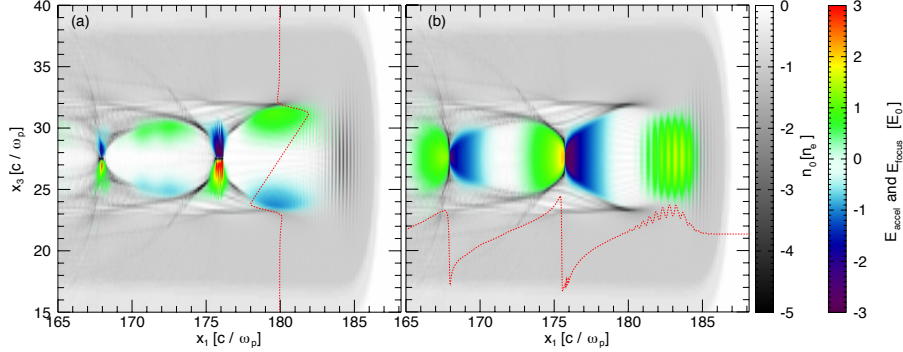
Fig. 4: Three-dimensional (3D) Osiris simulation showing focusing (a) and accelerating fields (b) for a laser wakefield accelerator in the blowout regime. The plasma density is shown in grey colours and fields in blue-red colours. The red dashed lines shown a transverse lineout of the focusing (a) and accelerating (b) fields. Focusing fields are linear within the entire blowout region. Accelerating fields are close to linear except at the back of the bubble.

Equation (41) is valid as long as the witness beam wakefields are negligible. In this case, Eq. (41) shows that the acceleration depends on the initial ξ position of each beam electron. This can increase the final energy spread of the bunch. As with the 1D case, however, it is possible to compensate for this effect by tailoring the shape of the witness bunch current profile in order to preserve the initial energy spread of the witness bunch.

In order to find the average accelerating field in the bubble, which when multiplied by the total acceleration distance determines the final energy energy gain, Eq. (41) needs to be supplemented with an additional expression defining the radius of the bubble. The radius of the bubble, $r_{\rm b}$, can be retrieved by equating the laser ponderomotive (repulsive) force to the ion channel (attractive) force:

$$F_{\rm p} \simeq E_r \Leftrightarrow \frac{a_0}{w_0} = E_r \simeq r_{\rm b} \to r_{\rm b} = \alpha\sqrt{a_0}, \tag{42}$$

where $\alpha = 2$ has been determined through PIC simulations [21, 25] and where we have assumed that $w_0 \simeq r_{\rm b}$. Combining Eq. (41) with Eq. (42) yields an estimate for the average accelerating field given by

$$\langle E_{\rm accel} \rangle \simeq \frac{\sqrt{a_0}}{2}. \tag{43}$$

The maximum acceleration distance corresponds to the smallest distance between pump depletion or dephasing. The pump depletion length, $L_{\rm pd}$, is the length it takes for the laser to exhaust its energy to the plasma through wakefield excitation. For propagation distances larger than $L_{\rm pd}$, the amplitudes of the plasma waves are negligible. Thus, we can assume that the acceleration stops at $L_{\rm pd}$. The dephasing length, $L_{\rm d}$, is the length it takes for a particle to outrun the accelerating phase of the wave, i.e. to go from regions with $\xi < 0$, where $E_{\rm accel} < 0$, to regions with $\xi = 0$ where $E_{\rm accel} = 0$.

Pump depletion in the blowout regime is determined by the rate at which the laser leading edge gives its energy to the plasma. This localized pump depletion process is also called etching. Since the back propagates mostly in vacuum, it does not give energy to the plasma. As the laser propagates, the front of the laser is then locally pump depleted. The pump depletion length is then given by the

product between the laser duration and the velocity at which the laser leading edge etches back, given by $v_{\text{etch}} = c\omega_{\text{p}}^2/\omega_0^2$ [26]

$$L_{\text{pd}} = \frac{\omega_0^2}{\omega_{\text{p}}^2} \left(c\tau_L \right). \tag{44}$$

The maximum dephasing length is given by the length it takes for a particle travelling at c to outrun the accelerating phase of the wakefield traveling with a phase velocity v_ϕ. For an electron initially at $\xi = r_{\text{b}}$, $L_d = cr_{\text{b}}/(c - v_\phi)$. Since the wake phase velocity is $v_\phi = v_g - v_{\text{etch}}$, where v_g is the laser linear group velocity given by

$$v_g = \frac{\partial \omega}{\partial k} = 1 - \frac{1}{2}\frac{\omega_0^2}{\omega_{\text{p}}^2}, \tag{45}$$

the dephasing length is

$$\frac{(c - v_\phi)}{c} L_{\text{d}} = r_{\text{b}} \Leftrightarrow L_{\text{d}} = \frac{2}{3}\frac{\omega_0^2}{\omega_{\text{p}}^2}r_{\text{b}}. \tag{46}$$

Combining Eqs. (44) and (45) yields a criteria for choosing the laser duration for optimal acceleration such that no laser energy is left after the electrons outrun the wave at $L_{\text{pd}} = L_{\text{d}}$:

$$\tau_L = \frac{2}{3}r_{\text{b}}. \tag{47}$$

We can now estimate the maximum energy $\Delta E = q\langle E_{\text{accel}}\rangle L_{\text{accel}}$ gained by an electron in the blowout regime. Denoting the acceleration distance by $L_{\text{accel}} = L_{\text{d}} = L_{\text{pd}}$ then

$$\Delta E = \frac{2}{3}m_{\text{e}}c^2\frac{\omega_0^2}{\omega_{\text{p}}^2}a_0. \tag{48}$$

So far, we have neglected the influence of the transverse laser dynamics on wakefield excitation and electron acceleration. This approximation is valid as long as the laser propagation and wakefields remain stable during L_{accel}. In order to stabilize the transverse laser dynamics, we need to explore how to prevent laser Rayleigh diffraction, one of the key processes that can degrade wakefield excitation and electron acceleration. Theory, simulations and experiments have shown that plasmas can act as optical fibers, guiding the propagation of intense lasers over distances that largely exceed the Rayleigh length. In strongly non-linear regimes, the blowout region refractive index gradients are sufficient to self-guide the body of the driver. Through simulations, it has been found that the optimal condition for stable, self-guided laser propagation occurs when $W_0 = r_{\text{b}} = 2\sqrt{a_0}$ as long as $a_0 > 2$ [21,25]. The laser front, which propagates in a region of nearly undisturbed plasma, may still diffract. This can be avoided if the etching rate exceeds the diffraction rate. This condition is met when

$$a_0 \gtrsim \left(\frac{n_{\text{c}}}{n_{\text{p}}}\right)^{1/5}. \tag{49}$$

Equation (49) is generally valid for $a_0 \gtrsim 4$. For $a_0 \gtrsim 2$, an external parabolic plasma channel needs to be present to externally guide the laser pulse.

In addition to determining maximum accelerating gradients and final energies in the blowout regime, the accelerating and focusing wakefields given by Eqs. (40) and (41) also define important beam loading properties such as the maximum charge that can be accelerated. To estimate the maximum amount of accelerated charge, we assume that a witness electron bunch absorbs all the energy contained in the longitudinal and focusing bubble fields. The electromagnetic energy of the wakefield in the blowout regime is

$$\varepsilon_\parallel \simeq \varepsilon_\perp \simeq \frac{1}{120}\left(k_{\text{p}}r_{\text{b}}\right)^5\left(\frac{m_{\text{e}}^2c^5}{\omega_{\text{p}}e^2}\right), \tag{50}$$

and the energy absorbed by N particles, assuming an average accelerating field gradient $E_{\text{accel}} \simeq r_{\text{b}}/2$, is

$$\varepsilon_{e^-} \simeq m_{\text{e}}c^2 N \left(\frac{k_{\text{p}}r_{\text{b}}}{2}\right)^2 . \tag{51}$$

Matching Eq. (50) to Eq. (51) then gives

$$N \simeq \frac{1}{30} \left(k_{\text{p}}r_{\text{b}}\right)^3 \frac{1}{k_{\text{p}}r_{\text{e}}}, \tag{52}$$

where N is the maximum number of electrons that can be loaded into the wakefield. The acceleration efficiency is the fraction of laser energy that goes into the accelerated electrons. Since the laser energy scales with $r_{\text{b}}^3 a_0^2$ (assuming $W_0 \simeq c\tau_L \simeq r_{\text{b}}$), then the efficiency goes as

$$\Gamma \simeq \frac{1}{a_0}. \tag{53}$$

Equations (48), (52) and (53) illustrate the trade-off between energy gain, maximum number of accelerated particles and efficiency. For a constant laser energy, lower laser a_0s leads to higher efficiencies at the expense of lower accelerated charge and longer accelerating distances that result in final higher energies. Higher laser a_0s lead to lower efficiencies, lower final bunch energies, but to higher charge. In addition, the acceleration distance is also smaller for higher a_0.

The scaling laws derived above can also be rewritten in practical units as

$$\tau[\text{fs}] = 53.22 \left(\frac{\lambda_0 [\mu\text{m}]}{0.8}\right)^{2/3} \left(\frac{\epsilon[\text{J}]}{a_0^2}\right)^{1/3}, \tag{54}$$

$$w_0 = \frac{3}{2}c\tau_L, \tag{55}$$

$$n_0 \left[\text{cm}^{-3}\right] \simeq 3.71 \frac{a_0^3}{P[\text{TW}]} \left(\frac{\lambda_0[\mu\text{m}]}{0.8}\right)^{-2} . \tag{56}$$

The total acceleration distance, final energy and maximum accelerated charge are:

$$L_{\text{accel}}[\text{cm}] \simeq 14.09 \frac{\epsilon[\text{J}]}{a_0^3}, \tag{57}$$

$$\Delta E[\text{GeV}] \simeq 3 \left(\frac{\epsilon[\text{J}]}{a_0^2} \frac{0.8}{\lambda_0[\mu\text{m}]}\right)^{2/3}, \tag{58}$$

$$q[\text{nC}] \simeq 0.17 \left(\frac{\lambda_0[\mu\text{m}]}{0.8}\right)^{2/3} (\epsilon[\text{J}]a_0)^{1/3} . \tag{59}$$

These scalings have been used to guide and predict the output of current laser wakefield acceleration experiments, and to guide the design of future experiments using some of the most powerful lasers soon to become available. They have also been confirmed through numerous 3D PIC simulations performed with different algorithms. For example, PIC simulations performed in relativistic boosted frames have illustrated the acceleration of 12–14 GeV electron bunches with 1–2 nC each ($\sim$4.8 nC in total) using lasers that will soon become available at the Extreme Light Infrastructure (ELI) [27]. Simulations showed that the acceleration could take place over distances smaller than 10 cm. These simulations were performed for strongly non-linear regimes using $a_0 = 4$, where a fraction of the background plasma electrons were trapped and accelerated to these high energies. These results are illustrated in Fig. 5.

Simulations have also been performed at lower laser intensities with $a_0 = 2$, where self-injection is absent, and using laser energies close to 250 J. The acceleration of an externally injected electron bunch
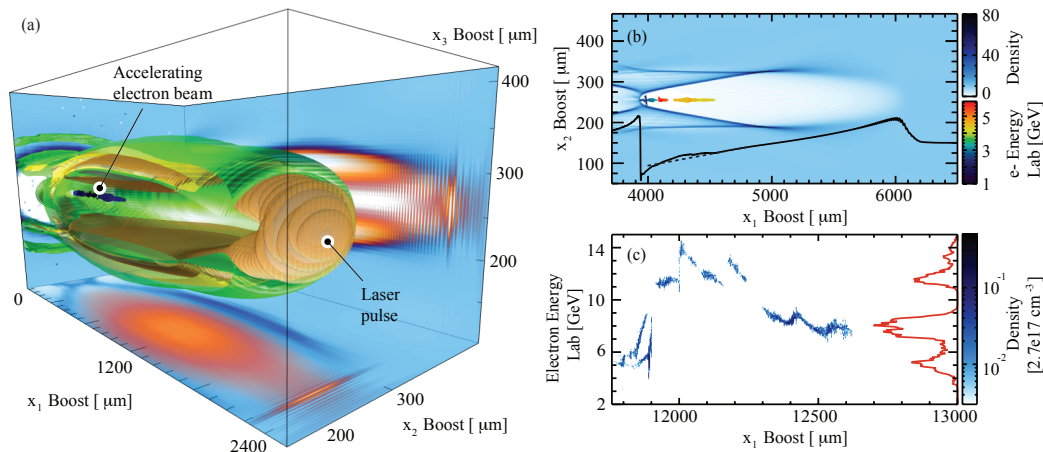
Fig. 5: (Picture taken from Ref. [27]) 3D PIC simulation in a Lorentz boosted frame of a laser wakefield accelerator using a 250 J laser pulse in self-injection, self-guiding scenarios: (a) shows the laser pulse in orange, the accelerated beam in dark blue and plasma density isosurfaces in green and yellow. Blue projections represent the background electron plasma density and orange projections the focusing and laser pulse electric fields; (b) shows a central slice of the simulation box illustrating the plasma electron density (blue) and self-injected electrons coloured according to their energy. The line-out (black) represents the longitudinal electric field (E_{accel}). The dashed line (black) represents the theoretical prediction (Eq. (41)); and (c) represents the phase-space of self-injected electrons. The red line represents the integrated energy spectrum.

was then investigated. Simulations showed the acceleration of a 40 GeV electron beam with 0.3 nC in a preformed parabolic plasma channel 5 m long. These results are shown in Fig. 6.

The scalings presented here are strictly valid for $2 \lesssim a_0 \lesssim 2 \left(\omega_0/\omega_{\mathrm{p}}\right)^{1/4}$. However, electron acceleration can also occur at much higher laser intensities. For instance, using $a_0 = 53$, 3 GeV electron bunches with high charges of around 25 nC could be achieved [28], although at the expense of higher energy spreads. These results are illustrated in Fig. 7.

4 Theory for the blowout regime

In this section, we will outline an analytical derivation for the electromagnetic field structure of the wakefield in the blowout regime [21]. The theory assumes cylindrical symmetry and employs the quasi-static approximation. Thus, as with to the 1D scenario (cf. Section 2), the motion of plasma electrons in multidimensions is also characterized by a constant of motion relating the particle velocity to the wake potential (cf. Eq. (28)). This constant of motion plays a very important role in understanding the trapping process in multidimensions and also in defining the shape of the blowout region. In order to generalize Eq. (28) for the multidimensional scenario, we then start by considering the Hamiltonian of a charged particle in an electromagnetic field:

$$H = \sqrt{m_{\mathrm{e}}^2 c^4 + \left(\mathbf{P} + \frac{e\mathbf{A}}{c}\right)^2} - e\phi, \tag{60}$$

where $\mathbf{P} = \mathbf{p} - e\mathbf{A}/c$ is the canonical momentum and where ϕ is the scalar potential. It is useful to employ the co-moving frame variables that move at the wake phase velocity. In the co-moving frame, variables $\xi = v_\phi t - x$ and $\tau = x$, the Hamiltonian Eq. (60) becomes

$$\mathcal{H} = H - v_\phi \mathbf{P}_\parallel, \tag{61}$$

where $\mathbf{P}_\parallel$ corresponds to the canonical momentum in the longitudinal x direction.
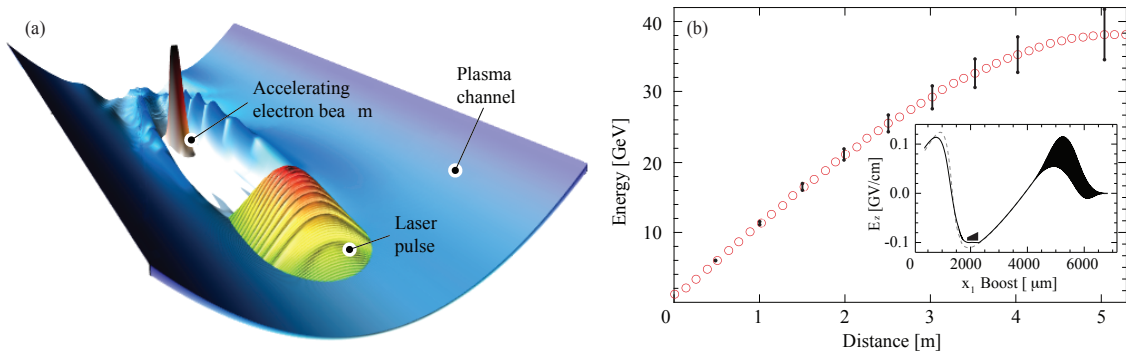
Fig. 6: Picture taken from Ref. [27]) 3D PIC simulation results in a Lorentz boosted frame of a laser wakefield accelerator using a 250 J laser pulse in external injection and external guiding scenarios: (a) illustration of the parabolid plasma channel used to guide the laser pulse (blue). The laser pulse is also shown. An externally injected electron bunch accelerates at the back of the plasma wave; and (b) shows the evolution of the maximum energy as a function of the acceleration distance. The vertical line segments correspond to the energy spread of the bunch. The inset shows the relative position of the externally injected bunch and longitudinal accelerating field. The bunch current profile was designed in order to lead to a flat accelerating field in the region of the bunch to minimize the growth of the energy spread.
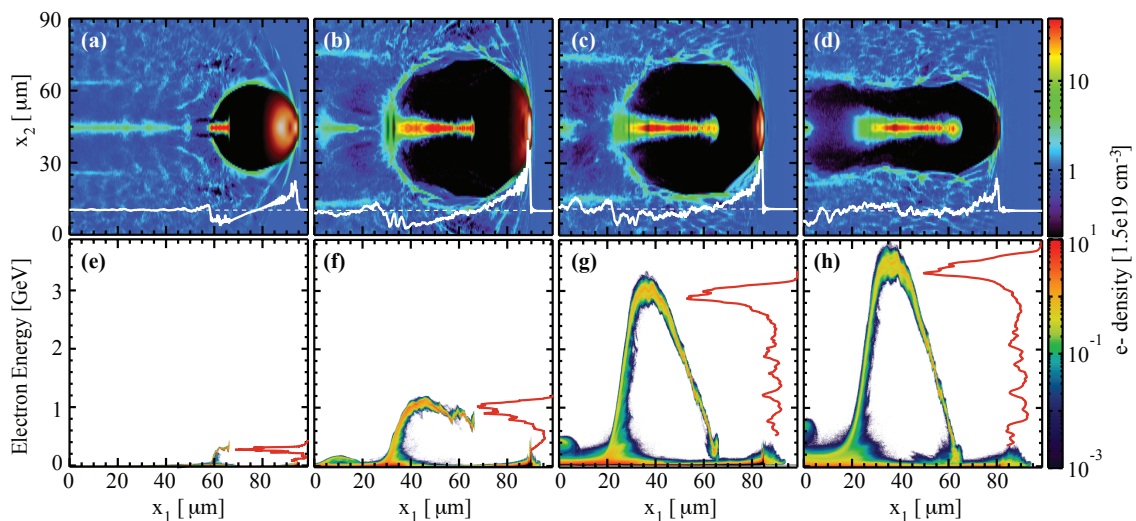
Fig. 7: Picture taken from Ref. [27]). 3D simulation results illustrating a laser wakefield accelerator in the bubble regime using a laser with $a_0 = 53$: (a)–(d) show central slices of the simulation box illustrating the plasma density (colours). On-axis accelerating field lineouts are represented in white; and (e)–(h) show the phase space of the plasma (colours) and integrated energy spectrum (red line).

In order to determine the constant of motion in multidimensions, we integrate Hamilton's equations written in the co-moving frame. Using the chain rule for the co-moving frame variables

$$\frac{\partial}{\partial x} = -\frac{\partial}{\partial \xi}, \tag{62}$$

$$\frac{\partial}{\partial t} = v_\phi \frac{\partial}{\partial \xi} + \frac{\partial}{\partial \tau}, \tag{63}$$

$$\frac{\mathrm{d}\xi}{\mathrm{d}t} = v_\phi - v_\|, \tag{64}$$

where $v_\| = \mathrm{d}x/\mathrm{d}t$ is the longitudinal velocity of an electron. Hamilton's equations can then be written as

$$v_\phi \frac{\mathrm{d}P_\|}{\mathrm{d}t} = -v_\phi \frac{\partial H}{\partial x} = v_\phi \frac{\partial H}{\partial \xi}, \tag{65}$$

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{\partial H}{\partial t} = v_\phi \frac{\partial H}{\partial \xi} + \frac{\partial H}{\partial \tau}. \tag{66}$$

Thus, the temporal evolution for $\mathcal{H}$ becomes

$$\left(v_\phi - v_\|\right)\frac{\mathrm{d}\mathcal{H}}{\mathrm{d}\xi} = \frac{\partial H}{\partial \tau} = \left[\mathbf{v} \cdot \frac{\partial \mathbf{A}}{\partial \tau} - \frac{\partial \phi}{\partial \tau}\right]. \tag{67}$$

We note that $\Delta\mathcal{H}$ depends on the initial and final positions only. Thus [29–32]

$$\Delta\mathcal{H} = \int \frac{\mathrm{d}\mathcal{H}}{\mathrm{d}t}\mathrm{d}t = \int \frac{\mathrm{d}\xi}{v_\phi - v_\|}\frac{\mathrm{d}\mathcal{H}}{\mathrm{d}\xi}, \tag{68}$$

where the integral on the left-hand side is performed over the particles trajectory [29–32]. For a non-evolving wakefield, the right-hand side of Eq. (68) vanishes because $\partial/\partial\tau = 0$. This corresponds to the QSA, where the wakefield does not change during the transit time of a plasma electron. Thus, under the QSA, $\Delta\mathcal{H} = 0$, and hence

$$\Delta H = \Delta\gamma - v_\phi\Delta p_\| - \left(\Delta\phi - v_\phi\Delta A_\|\right) \tag{69}$$

$$= \Delta\gamma - v_\phi\Delta p_\| - \Delta\psi, \tag{70}$$

where $\psi = \phi - v_\phi A_\|$ is called the wake pseudo-potential. Equation (69) generalizes the 1D constant of motion given by Eq. (28) in the presence of magnetic fields (through the presence of the longitudinal vector potential $\Delta A_\|$) for particles born with arbitrary initial velocity in regions with arbitrary electric and magnetic fields.

For a particle initially at rest, and initially in a region of vanishing fields, Eq. (69) reduces to

$$\gamma\left(1 - \beta_\|\right) = 1 + \psi. \tag{71}$$

As with the 1D case, we can relate the particle velocity to the wake potential ψ and determine the onset of self-injection as a function of ψ. When $\beta_\| \to -1$, i.e. when particles move backwards at c, $\psi \to \infty$. When $\beta_\| \to 1$, i.e. when particles move forward at c and are trapped, $\psi \to -1$. We can then state that the limit $\psi \to -1$ is a sufficient, but not necessary, condition for trapping under the QSA.

We will use the constant of motion given by Eq. (71) to derive the equation of motion for plasma electrons that move within the boundary defining the blowout region. The equation of motion is given by the Lorentz force, which can be simplified using Eq. (71). Our calculations now assume $v_\phi = c = 1$. Thus, the required time derivative of the equation of motion is

$$\frac{\mathrm{d}}{\mathrm{d}t} = (1 - v_\phi)\frac{\mathrm{d}}{\mathrm{d}\xi} = \frac{1 + \psi}{\gamma}\frac{\mathrm{d}}{\mathrm{d}\xi}, \tag{72}$$

where we recall that the longitudinal component of the electron velocity $v_\|$ is normalized to c. As a result, $p_\perp = \gamma v_\perp = (1 + \psi)\frac{\mathrm{d}r_\perp}{\mathrm{d}\xi}$ and

$$\frac{\mathrm{d}p_\perp}{\mathrm{d}t} = \frac{1 + \psi}{\gamma}\frac{\mathrm{d}}{\mathrm{d}\xi}\left[(1 + \psi)\frac{\mathrm{d}}{\mathrm{d}\xi}\right]r_\perp, \tag{73}$$

96

where $r_\perp$ and $p_\perp$ are the transverse radius and momentum, respectively. Equation (73) depends on three quantities, namely $p_\perp$, γ and ψ. It is possible to reduce the number of unknowns by using Eq. (71) to write γ as a function of $p_\perp$ and ψ as

$$\gamma = \frac{1 + p_\perp^2 + (1 + \psi)^2}{2(1 + \psi)}. \tag{74}$$

Thus, the Lorentz force equation for the radial motion of a plasma electron becomes

$$\frac{2(1 + \psi)^2}{1 + (1 + \psi)^2 \left(\frac{dr_\perp}{d\xi}\right)^2 + (1 + \psi)^2} \frac{d}{d\xi}\left[(1 + \psi)\frac{dr_\perp}{d\xi}\right] = F_\perp, \tag{75}$$

where $F_\perp = -\left(E_r - v_\parallel B_\theta\right)$ is the radial force acting on a plasma electron assuming cylindrical symmetry, with E_r being the radial electric field and B_θ the azimuthal magnetic field. We note that the left-hand side of Eq. (75) depends on $p_\perp$, which defines the shape of the bubble, and on ψ, which is related to the field structure of the blowout. In order to solve Eq. (75) we need to relate ψ with $p_\perp$. This relation can be established through the field structure of the bubble.

The bubble fields can be written as a function of the scalar and vector potentials in the Coulomb Gauge as

$$E_z = \frac{\partial \psi}{\partial \xi}, \tag{76}$$

$$B_\theta = -\frac{\partial A_z}{\partial r} - \frac{\partial A_r}{\partial \xi}, \tag{77}$$

$$E_r = -\frac{\partial \phi}{\partial r} - \frac{\partial A_r}{\partial \xi}, \tag{78}$$

where scalar and vector potentials can be fully specified once the plasma electron density and currents are known, according to [22]

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial A_r}{\partial r}\right) - \frac{A_r}{r^2} \quad = n_e v_\perp \quad \Leftrightarrow A_r = A_{r0}(\xi)\,r, \tag{79}$$

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial A_\parallel}{\partial r}\right) \quad = n_b + n_e v_\parallel \quad \Leftrightarrow A_\parallel = A_{\parallel 0}(\xi) + \lambda(\xi)\log(r), \tag{80}$$

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial \phi}{\partial r}\right) \quad = n_b + n_e - 1 \quad \Leftrightarrow \phi = \phi_0 - \frac{r^2}{2} + \lambda(\xi)\log(r), \tag{81}$$

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial \psi}{\partial r}\right) \quad = n_e + n_e v_\parallel - 1 \quad \Leftrightarrow \psi = \psi_0(\xi) - \frac{r^2}{4}, \tag{82}$$

$$\frac{1}{r}\frac{\partial}{\partial r}r A_r \quad = -\frac{\partial \psi}{\partial \xi} \quad \Rightarrow A_{r0} = -\frac{1}{2}\frac{d\psi_0}{d\xi}, \tag{83}$$

where we have included the possibility of describing the effect of a particle beam driver, with $\lambda(\xi) = \int_0^\infty r n_b dr$ its current profile. Using Eqs. (79)–(83), the right-hand side of Eq. (75) can be rewritten as

$$F_\perp = -\frac{r}{2} + (1 - v_\parallel)\frac{\lambda(\xi)}{r} + (1 - v_\parallel)\frac{dA_{r0}}{d\xi}r - \frac{1}{\gamma}\nabla_\perp|\frac{a_L}{2}|^2, \tag{84}$$

where the first term (on right-hand side of Eq. (84)) represents the electrostatic field due to the background plasma ions, the second term is the force exerted on the plasma electrons by a charged particle bunch driver, the third term is due to the radial plasma currents, and the last term is the laser ponderomotive force. By expressing $1 - v_\parallel = (1 + \psi)/\gamma$ and replacing γ by Eq. (74), we can rewrite Eq. (84)

as

$$\frac{d}{d\xi}\left[(1+\psi)\frac{dr_b}{d\xi}\right] = r_b\left\{-\frac{1}{4}\left[1 + \frac{1}{(1+\psi)^2} - \left(\frac{dr_b}{d\xi}\right)^2\right]\right\} - \frac{1}{2}\frac{d^2\psi_0}{d\xi^2} + \frac{\lambda(\xi)}{r_b^2} - \frac{1}{\left(\psi_0 - \frac{r_b^2}{4}\right)}\nabla|\frac{a_L}{2}|^2,$$
(85)

which shows that the trajectory of plasma electrons are fully specified by the pseudo-potential ψ. Inversion of Eq. (82) reveals that ψ depends on the radial $n_e(1 - v_\parallel)$ profile:

$$\psi(r,\xi) = \ln r \int_0^r r'\left[n_e(r',\xi)\left(1 - v_\parallel(r',\xi)\right) - 1\right]dr'$$
(86)

$$+ \int_r^\infty r'\ln r'\left[n_e(r',\xi)\left(1 - v_\parallel(r',\xi)\right) - 1\right]dr'.$$
(87)

In order to ensure that the fields vanish away from the blowout region, the profile of $n_e(1 - v_\parallel)$ is also subject to the following boundary condition:

$$\int_0^r r'\left[n_e(r',\xi)\left(1 - v_\parallel(r',\xi)\right) - 1\right]dr' = 0,$$
(88)

which states that the source term for ψ $n_e(1 - v_\parallel)$ is conserved for each transverse slice.

Equations (85), (86) and (88) are general and valid as long as the QSA holds. In order to close our model and to derive the equations that determine the main properties of the blowout regime, we now look for an expression of ψ valid in the blowout regime. In order to determine ψ we then need to find an appropriate model for $n_e(1 - v_\parallel)$ which can describe and reproduce the most important features of the blowout regime.

Figure 8 shows the result of a PIC simulation that illustrates the main properties of the blowout regime that need to be included in a model for $n_e(1 - v_\parallel)$ that can accurately describe the blowout regime. Figure 8 shows that the blowout is characterized by a region empty of plasma electrons, which instead accumulate at the boundary that defines the blowout region. These features are well captured by the simplified model shown by the red line in Fig. 8(b), which depends on the blowout radius r_b, on the value of $n_e(1 - v_\parallel) = n_\Delta$ at $r = r_b$ and on the thickness Δ of the electron sheath that defines the blowout.

The boundary condition given by Eq. (88) yields

$$n_\Delta(\xi) = \frac{r_b^2}{(r_b + \Delta)^2 - r_b^2}.$$
(89)

Using Eq. (86) and Eq. (89) then yields the following expression for ψ as a function of r_b and $\alpha = \Delta/r_b$:

$$\psi[r_b(\xi)] = \frac{r_b^2}{4}\left(\frac{(1+\alpha)^2\ln(1+\alpha)^2}{(1+\alpha)^2} - 1\right) \equiv \frac{r_b^2}{4}\beta.$$
(90)

Equation (90) is a general expression that specifies ψ for given a r_b and α. The plasma response in the blowout regime is always non-linear, but, depending on the maximum value of r_b, it may be relativistic or non-relativistic. When $k_p r_b \ll 1$, the plasma response is non-relativistic. In this limit, the ratio of the blowout radius to the width of the thin electron sheath is much smaller than unity, or $\alpha = \Delta/r_b \gg 1$. In the ultra-relativistic blowout regime, $k_p r_b \gg 1$ and $\alpha = \Delta/r_b \ll 1$, i.e. the width of the electron layer that defines the blowout is much smaller than the blowout radius. We can use Eq. (90) to find ψ in these two limiting scenarios. Hence, in the non-relativistic blowout regime, Eq. (90) reduces to

$$\psi(r,\xi) \simeq \frac{r_b^2(\xi)}{4}\log\left(\frac{1}{r_b}\right) - \frac{r^2}{4},$$
(91)
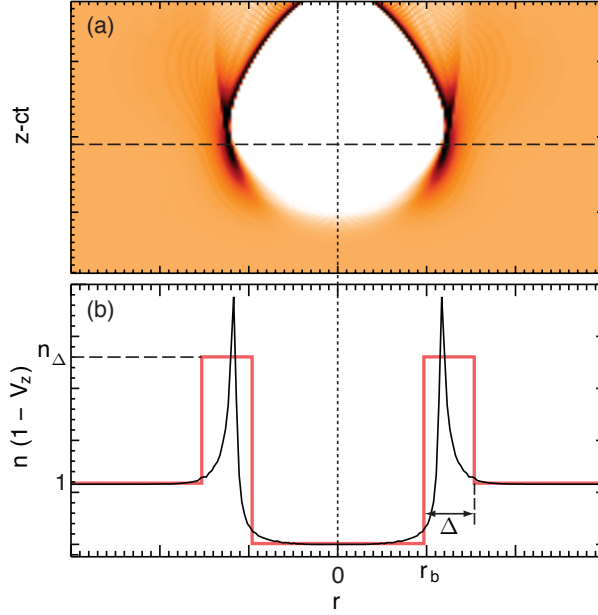
Fig. 8: PIC simulation result illustrating the blowout regime and the model for the $n_e(1-v_\parallel)$ profile in the blowout: (a) shows $n_e(1-v_\parallel)$ for a slice of the simulation box. The laser driver moves to the bottom of the page; and (b) shows a lineout (solid black line) of $n_e(1-v_\parallel)$ taken at the position of the horizontal dashed line in (a). The solid red line shows the simplified model used to describe the blowout regime.

and in the ultra-relativistic blowout regime, Eq. (90) becomes

$$\psi(r,\xi) \simeq (1+\alpha)\frac{r_b^2(\xi)}{4} - \frac{r^2}{4}. \tag{92}$$

In order to obtain a general description of the blowout regime, valid in the relativistic and non-relativistic regimes, we insert the expression for ψ given by Eq. (90) into the Lorentz force equation given by Eq. (85). This results in a non-linear differential equation for the motion of the inner most electron in the blowout $r_b(\xi)$ [21]

$$A(r_b)\frac{d^2r_b}{d\xi^2} + B(r_b)r_b\left(\frac{dr_b}{d\xi}\right)^2 + C(r_b)r_b = \frac{\lambda(\xi)}{r_b} - \frac{1}{4}\frac{d|a|^2}{dr}\frac{1}{\left(1+\beta r_b^2/4\right)^2}, \tag{93}$$

where we have assumed that Δ does not depend on ξ, and where

$$A(r_b) = 1 + \left(\frac{1}{4} + \frac{\beta}{2} + \frac{1}{8}r_b\frac{d\beta}{dr_b}\right)r_b^2, \tag{94}$$

$$B(r_b) = \frac{1}{2} + \frac{3}{4}\beta + \frac{3}{4}r_b\frac{d\beta}{dr_b} + \frac{1}{8}r_b^2\frac{d^2\beta}{dr_b^2}, \tag{95}$$

$$C(r_b) = \frac{1}{4}\left(1 + \frac{1+|a|^2/2}{1+\beta r_b^2/4}\right). \tag{96}$$

In the ultra-relativistic blowout regime ($k_p r_b \gg 1$), the assumption that $\partial_\xi \Delta \simeq 0$ breaks at the back of the bubble where $\Delta \ll r_b$. Numerical solutions to Eqs. (93) and (94) are in very good agreement with full PIC simulations for a wide range of conditions from weakly- to ultra-relativistic blowouts. The agreement is nearly perfect except for the region at the back of the bubble where $\Delta \simeq r_b$. Figure 9(a), which compares Eq. (93) with PIC simulations, shows excellent agreement for almost the entire blowout

region, except at the back of the bubble. Figure 9 considered an electron beam driver to excite the blowout. In the laser driven case, the blowout region is not as well defined. Still, comparisons between theory and simulations are very good when $W_0 \simeq r_m$, where r_m is the maximum blowout radius.
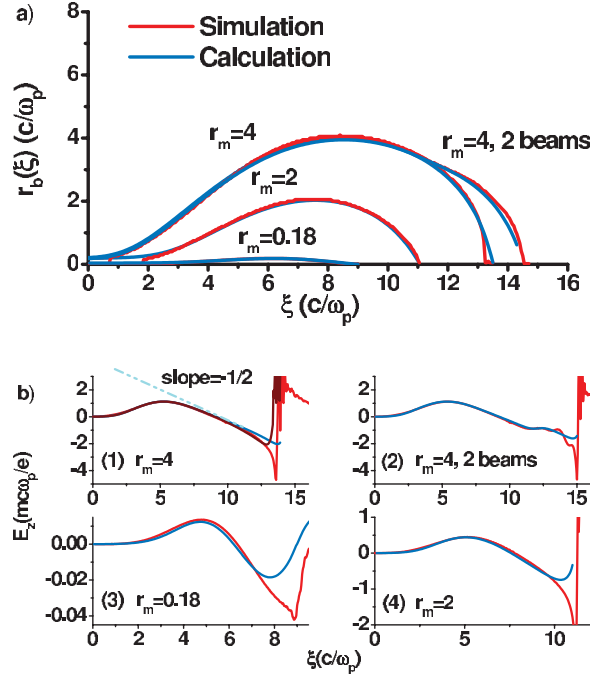
Fig. 9: (Picture taken from Ref. [21]) Comparison between the blowout theory (blue lines) and PIC simulations (red lines): (a) trajectory of the blowout radius $r_b(\xi)$ for different maximum blowout radius r_m, from the non-relativistic blowout regime ($k_p r_m = 0.18$) and strongly relativistic blowout regime ($k_p r_m = 4$); and (b) shows the corresponding on-axis accelerating fields for all the cases in (a). The driver propagates from right to left, in the direction of negating ξ.

It is possible to simplify Eq. (93) in order to predict key wakefield properties in current plasma acceleration experiments, which operate in the ultra-relativistic regime. The shape of the blowout region can be predicted in this extreme regime by assuming $\alpha \ll 1$, for which Eq. (93) simplifies to

$$r_b \frac{d^2 r_b}{d\xi^2} + 2\left(\frac{dr_b}{d\xi}\right)^2 + 1 = \frac{4\lambda(\xi)}{r_b} - \frac{d|a|^2}{dr} \frac{1}{\left(1 + \beta r_b^2/4\right)^2}. \tag{97}$$

At the back of the driver where the right-hand side of Eq. (97) vanishes, Eq. (97) becomes very similar to the equation of a sphere, which is given by

$$r_b \frac{d^2 r_b}{d\xi^2} + \left(\frac{dr_b}{d\xi}\right)^2 + 1 = 0. \tag{98}$$

The main difference between Eqs. (97) and (98) is the multiplication factor in the second term on the left-hand side of both equations. The additional factor of two in Eq. (97) leads to a stronger bending of the blowout radius at the back of the bubble when compared with a sphere. However, for most of the wakefield, the blowout resembles a sphere.

Having determined r_b, and hence ψ, it is now possible to derive the full field structure of the blowout region. The accelerating field is given by Eq. (76), which reads $E_z = \partial\psi/\partial\xi$, and the focusing field acting on a relativistic particle traveling at c is $W_\perp = E_r - B_\theta = -\partial\psi/\partial r$. In ultra-relativistic

regimes, E_z can be evaluated by integrating Eq. (97) near the region where the blowout radius is maximum, and then inserting the resulting expression for $r_{\rm b}$ into Eq. (92). Using Eq. (76) gives

$$E_z = \frac{1}{2}\frac{{\rm d}r_{\rm b}}{{\rm d}\xi} \simeq \frac{\xi}{2}. \tag{99}$$

The maximum accelerating gradient is thus given by $E_z \simeq r_{\rm b}/2$, and we recover the result from the phenomenological theory for the blowout regime. Figure 9(b) compares the theoretical predictions for E_z with simulation results from the non-relativistic to the ultra-relativistic blowout regime, showing very good agreement, except at the back of the bubble where the blowout approaches the axis. We note that, unlike in the linear regime, the accelerating electric field in the bubble is independent of r.

The focusing force can be determined in a similar manner yielding:

$$E_r - B_\theta = \frac{r}{2}, \tag{100}$$

which is also in excellent agreement with simulation results.

5 Beam loading

In the previous section we demonstrated that the blowout regime provides linear focusing and accelerating fields. Linear focusing fields are critical to preserve the emittance of accelerated beams during their acceleration. Matching conditions, relating the initial beam emittance to the linear focusing force in a plasma are well known (see, for instance, [33] and references therein). Linear accelerating fields suggest a way to tailor the currents of accelerated particle bunches ensuring constant accelerating fields for all bunch particles. This is critical for the acceleration of particles with no energy spread growth [34].

In order to investigate beam loading in the strongly non-linear blowout regime we integrate Eq. (97):

$$E_z = \frac{1}{2}r_{\rm b}\frac{{\rm d}r_{\rm b}}{{\rm d}\xi} = -\frac{r_{\rm b}}{2\sqrt{2}}\sqrt{\frac{16\int l(\xi)\xi {\rm d}\xi + C}{r_{\rm b}^4} - 1}, \tag{101}$$

where l is the current density of the electron beam injected into the plasma wave. Equation (101) gives the accelerating electric field in the blowout regime as a function of the position in the bubble ξ and for an arbitrary beam loading current profile. It is interesting to note that the beam loading depends on the integrated current profile of the beam, but not on the particular transverse shape of the bunch. Thus, the accelerating field remains unchanged even if the bunch evolves transversely.

The integral in Eq. (101) can be calculated analytically for the case of a trapezoidal bunch. The optimal beam loading current for a trapezoidal bunch, ensuring constant accelerating electric fields along the bunch is

$$l(\xi) = \sqrt{E_s^4 + \frac{R_{\rm b}^4}{16}} - E_s\left(\xi - \xi_s\right), \tag{102}$$

where the meaning of E_s and ξ_s is shown in Fig. 10. If the beam current profile satisfies Eq. (102), as in the case of Fig. 10, then all bunch particles accelerate with similar acceleration gradients, hence minimizing energy spread variations throughout the propagation. The trailing electron bunch will inevitably be subject to different accelerating fields as it dephases in the plasma wave in laser wakefield accelerators. Thus, growth of energy spread is expected in laser wakefield accelerators. This may be compensated by phase-space rotation near dephasing [35].

The maximum charge that can be loaded into the wakefield can be determined by assuming that the back of the bunch coincides with $r_{\rm b} = 0$. Using Eq. (102), the maximum charge is thus given by

$$Q_s = \frac{\pi}{16}\frac{R_{\rm b}^4}{E_s}, \tag{103}$$
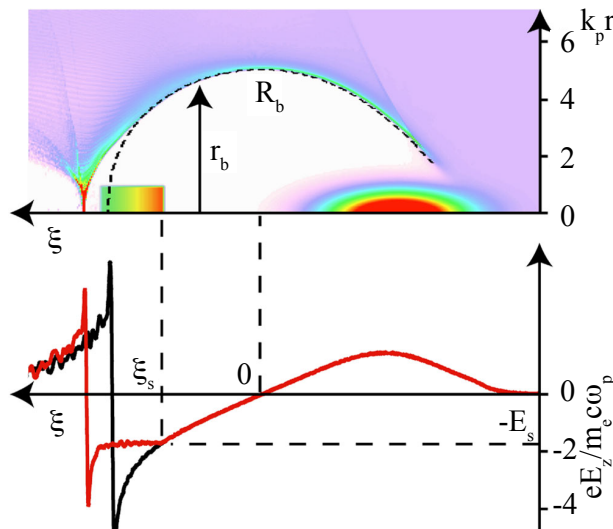
Fig. 10: (Picture taken from Ref. [34]) Top plot shows plasma electron density in the blowout, superimposed with a particle beam driver and a witness beam at the back of the bubble. The driver propagates from left to right. The blowout radius r_b and maximum blowout radius R_b are also indicated. The dashed line in the top plot shows the shape of the bubble in an unloaded scenario where the witness bunch is absent. The initial location of the witness bunch (ξ_s) and corresponding accelerating field E_s is also indicated in the bottom plot. The red (black) lines in the bottom plot show the accelerating fields with (without) the witness electron bunch.

where $R_b = r_b(\xi = 0)$ ($\xi = 0$ is at the centre of the bubble, cf. Fig. 10). Equation (103) shows that smaller E_s leads to higher charges because the beam can be made longer. However, the maximum energy per particle is also lower for smaller E_s. This illustrates the trade-off between the maximum energy gain and the maximum number of accelerated particles. The accelerator efficiency is the ratio between the absorbed energy and the total wakefield energy. This can be expressed as

$$\eta = \frac{\tilde{Q}_s}{Q_s}, \tag{104}$$

where $\tilde{Q}_s$ is the charge of a trapezoidal bunch described by Eq. (102).

6 Challenges

Positron acceleration in plasmas is a major challenge for a future plasma-based linear collider. Although the blowout is ideal for electron acceleration because it has linear focusing and accelerating fields, it cannot be employed to accelerate positrons to high energies because the transverse fields in the blowout regime defocus positrons. Thus, although positrons could accelerate at the front of the bubble, where $E_z > 0$, the acceleration time would be limited to the time it would take for positrons to be defocused away from the blowout region. There are several schemes for positron acceleration in plasmas to overcome this challenge [24, 36, 38, 39, 41]. Positron acceleration could occur in wakes driven by a positron beam driver, for instance. Instead of repelling plasma electrons radially, the positively charged driver sucks plasma electrons inwards. As plasma electrons flow towards the centre of the beam, they provide a positron focusing force. Some of the driver positrons could then accelerate to high energies. Scaling for the onset of the suck-in regime is determined by the time (τ_{col}) it takes for a plasma electron initially

at $r = r_0$ to cross the axis [24], which needs to be smaller than $1/\omega_p$:

$$\tau_{\text{col}} \simeq \sqrt{\pi} \left(\frac{r_0}{\sigma_r} \sqrt{\frac{m_b}{4\pi n_b e^2}} \right) \ll \frac{1}{\omega_p}, \tag{105}$$

where m_b is the positron mass, n_b the positron beam density and σ_r the positron beam transverse size.

Other schemes for positron acceleration have also been explored using hollow plasma channels. A hollow plasma channel contains no plasma for $r < r_c$, where r_c is the radius of the hollow channel. At the centre of the channel, focusing forces are negligible. Thus positrons and electrons may propagate for long distances. The accelerating fields, due to the currents formed at the edges, could then be used to accelerate positrons (and electrons) to high energies [36].

Recently, a new scheme has been proposed for positron acceleration in strongly non-linear plasma waves driven by exotic laser beams with orbital angular momentum and doughnut shaped intensity profiles. Lasers with orbital angular momentum are defined by Laguerre–Gaussian polynomials. These laser beams are interesting from a fundamental point of view because they have spiralling wave fronts that correspond to a well defined orbital angular momentum [37]. They have been used for several applications including super-resolution microscopy, quantum computing and ultra-fast communications. Recently, it has been suggested that these beams could also be used to drive doughnut shaped plasma waves in the blowout regime [38]. The onset of positron focusing and acceleration occurs when the electron sheath that defines the inner doughnut wall merges on-axis. The charge density of the electron sheath on axis can be much larger than the background plasma ion density. Thus, when it merges near the axis, this electron sheath provides a positron-focusing force. At the same time, longitudinal electric fields can accelerate the positrons at the front of the bubble [38]. Similar physics can also be obtained when using higher order Hermite–Gaussian modes [39]. In addition, the use of a ring-shaped electron beam driver could provide similar structures for positron acceleration [40, 41].
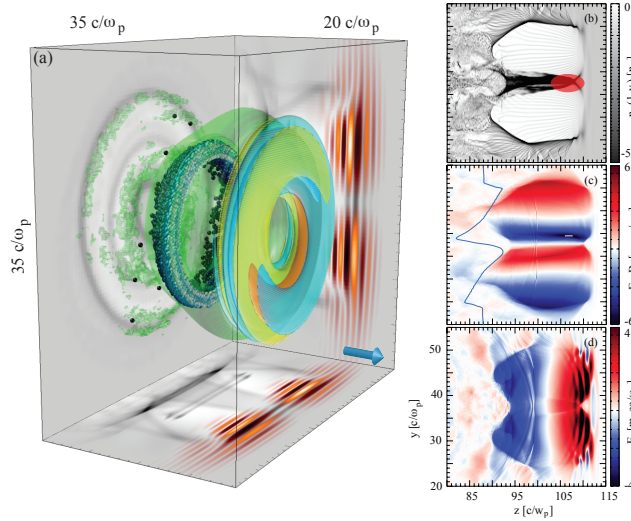
Fig. 11: PIC simulation results showing structures leading to positron acceleration in the strongly non-linear blowout regime using a Laguerre–Gaussian beam driver with orbital angular momentum: (a) shows plasma density isosurfaces (green) illustrating the generation of a doughnut bubble. Plasma density projections are in grey and laser pulse electric field projections in orange. The laser propagates in the direction of the blue arrow. Blue spheres represent a self-injected electron bunch ring; (b) shows the plasma density structures taken from a central slice of the simulation box. The location for positron focusing and accelerating fields is shown by the red shape; and (c)–(d) show the focusing and accelerating field structures taken from a central slice of the simulation box. The blue line in (c) shows a lineout of the focusing wakefields.

The use of proton bunch drivers for plasma accelerators has been recently proposed [42]. The field structure of a proton driven plasma wakefield accelerator is analogous to that of the bubble regime driven by lasers or electron bunch drivers with an additional phase shift associated with the crossing time given by Eq. (105). The beam loading is also identical to the blowout regime driven by intense lasers or particle bunches. Thus, proton driven wakefields can be used to accelerate high-charge electron bunches. It has been proposed that the compressed proton bunches available at CERN could be used to drive intense plasma waves capable of accelerating electrons to nearly 1 TeV in a plasma 600 m long [42].

Proton driven plasma wakefields require proton bunch drivers that are shorter than the plasma skin-depth (cf. Eq. (105)). Proton bunches, such as those at CERN, are, however, much longer and encompass hundreds of plasma wavelengths. The physics behind wakefield excitation, however, differs significantly from the processes described in the previous sections [43–45], but could still be used to drive large-amplitude plasma waves. When a long bunch propagates in a plasma, it may be subject to the so called self-modulation instability. The self-modulation instability leads to the generation of a train of beamlets, each separated by the plasma wavelength. The train of bunches can resonantly excite the plasma wave that grows from the head to the tail of the bunch. Large acceleration gradients can then be produced at the back of the bunch and be used to accelerate particles to high energies. This concept has motivated experiments at CERN (the AWAKE collaboration experiment) [46] and at SLAC [47]. Figure 12 shows an example of a self-modulated long bunch after propagation into a plasma.
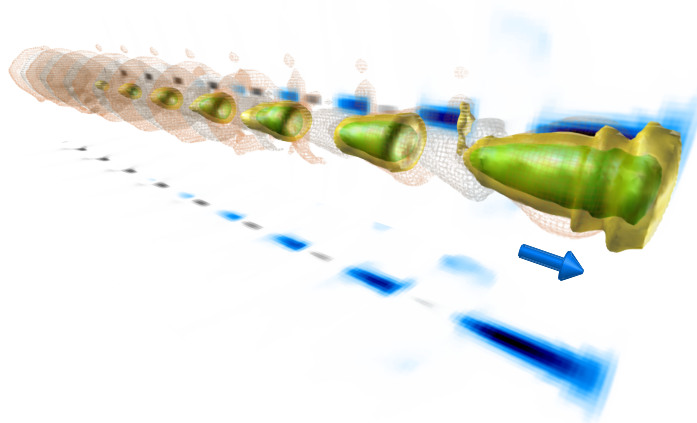
Fig. 12: OSIRIS PIC simulation result illustrating a self-modulated bunch consisting of several beamlets separated by the plasma wavelength. Green and yellow regions show density isosurfaces of the self-modulated particle bunch driver. The orange and grey mesh shows background plasma electron density iso-surfaces.

Another challenge related to higher energy physics applications is the acceleration of spin polarized particle bunches. Neglecting Stern-Gerlach forces, the spin of an ultra-relativistic charged particle propagating in a region with electric and magnetic fields proceeds according to the T-BMT equations:

$$\frac{\mathrm{d}\mathbf{s}}{\mathrm{d}t} = -\left[\left(a + \frac{1}{\gamma}\right)(\mathbf{B} - \mathbf{v} \times \mathbf{E}) - \mathbf{v}\frac{a\gamma}{\gamma+1}\mathbf{v} \cdot \mathbf{B}\right] \times \mathbf{s} = \mathbf{\Omega} \times \mathbf{s}, \tag{106}$$

where a is the anomalous magnetic momentum of an electron. A fully polarized beam has its spins aligned in the same direction, while a fully un-polarized beam has no net spin. For high-energy physics experiments, high spin polarizations above 80% are required. The conditions for keeping beam polarisation unchanged as a particle beam accelerates in plasma wakefields have already been determined [48]. In general, beam depolarisation depends on the elipticity of the particle trajectories as they accelerate. Spin depolarisation will be smaller when the trajectories of the accelerated particles occur in a single

plane and larger when the trajectories are helicoidal. Thus, beam spin polarisation can be maintained as long as the beam emittance is sufficiently small [48].

7 Conclusions

In this report we have outlined wakefield excitation models using various approximations, valid in one- and multidimensional scenarios. Using these models, we have shown how the field structure of plasma waves in the linear and non-linear regimes can be determined, and we have derived a set of scaling laws for the maximum energies and charges that can be achieved in plasma-based accelerators. We have shown that the blowout regime has the potential to accelerate high-charge, high-quality electron bunches to high energies. Laser wakefield acceleration experiments in the blowout regime have demonstrated multi-GeV electron accelerations [9, 10], and future experiments in the field promise to take the technology even closer to applications.

Acknowledgements

References

[1] T. Tajima and J.M. Dawson, *Phys. Rev. Lett.* **43** (4) (1979) 267–270.
http://dx.doi.org/10.1103/PhysRevLett.43.267

[2] ELI beam lines facility, http://www.eli-beams.eu; ELI Nuclear Physics, http://www.eli-np.ro; ELI Atto-second, http://www.eli-hu.hu.

[3] L.O. Silva *et al.*, *Phys. Rev.* E **59**(2) (1999) 2273. http://dx.doi.org/10.1103/PhysRevE.59.2273

[4] E. Esarey *et al.*, *IEEE Trans. Plasma Sci.* **24**(2) (1996) 252. http://dx.doi.org/10.1109/27.509991

[5] C.G.R. Geddes *et al.*, *Nature* **431** (2004) 538. http://dx.doi.org/10.1038/nature02900

[6] J. Faure *et al.*, *Nature* **431** (2004) 541. http://dx.doi.org/10.1038/nature02963

[7] S.P.D. Mangles *et al.*, *Nature* **431** (2004) 535. http://dx.doi.org/10.1038/nature02939

[8] John M. Dawson, *Phys. Rev.* **113**(2) (1959) 383. http://dx.doi.org/10.1103/PhysRev.113.383

[9] W.P. Leemans *et al.*, *Phys. Rev. Lett.* **113**(24) (2014) 245002.
http://dx.doi.org/10.1103/PhysRevLett.113.245002

[10] M. Litos *et al.*, *Nature* **515** (2014) 92. http://dx.doi.org/10.1038/nature13882

[11] R.A. Fonseca *et al.*, OSIRIS: A Three-Dimensional, Fully Relativistic Particle in Cell Code for Modeling Plasma Based Accelerators, *Lecture Notes in Computer Science* Vol. 2331/2002 (Springer, Berlin, Heidelberg 2002), pp. 342–351. http://dx.doi.org/10.1007/3-540-47789-6_36

[12] C. Huang *et al.*, *J. Comp. Phys.* **217**(2) (2006) 658. http://dx.doi.org/10.1016/j.jcp.2006.01.039

[13] R.A. Fonseca *et al.*, *Plasma Phys. Control. Fusion* **55**(12) (2013) 124001.
http://dx.doi.org/10.1088/0741-3335/55/12/124011

[14] P. Gibbon, *Short Pulse Laser Interactions with Matter: An Introduction* (Imperial College Press, Published by World Scientific Publishing Company, London, 2005).
http://dx.doi.org/10.1142/p116

[15] W.B. Mori *et al*, Private communication.

[16] P. Sprangle *et al.*, *Phys. Rev. Lett.* **64**(17) (1990) 2011.
http://dx.doi.org/10.1103/PhysRevLett.64.2011

[17] V.I. Berezhiani and I.G. Murusidze, *Phys. Lett.* A **148**(6–7) (1990) 338.
http://dx.doi.org/10.1016/0375-9601(90)90813-4

[18] Thomas M. Antonsen and P. Mora, *Phys. Plasmas* **4**(1) (1997) 217. http://dx.doi.org/10.1063/1.872134

[19] T Mehrling *et al.*, *Plasma Phys. Control. Fusion* **56**(8) (2014) 084012. http://dx.doi.org/10.1088/0741-3335/56/8/084012

[20] T. Katsouleas *et al*, *Particle Accelerators* **22** (1987) 81.

[21] W. Lu *et al.*, *Phys. Rev. Lett.* **96**(16) (2006) 165002. http://dx.doi.org/10.1103/PhysRevLett.96.165002

[22] W. Lu, Master Thesis, University of California Los Angeles, 2004.

[23] A. Pukhov and J. Meyer-Ter-Vehn, *Appl. Phys.* B **74** (4–5) (2002) 355. http://dx.doi.org/10.1007/s003400200795

[24] S. Lee *et al.*, *Phys. Rev.* E **64**(4) (2001) 045501(R). http://dx.doi.org/10.1103/PhysRevE.64.045501

[25] W. Lu *et al.*, *Phys. Rev.* ST-AB **10**(6) (2007) 061301. http://dx.doi.org/10.1103/PhysRevSTAB.10.061301

[26] C.D. Decker *et al.*, *Phys. Plasmas* **3**(4) (1996) 1360. http://dx.doi.org/10.1063/1.871728.

[27] S.F. Martins *et al.*, *Nat. Physics* **6** (2010) 311. http://dx.doi.org/10.1038/nphys1538

[28] S. Gordienko and A. Pukhov, *Phys. Plasmas* **12**(4) (2005) 043109. http://dx.doi.org/10.1063/1.1884126

[29] A. Pak *et al.*, *Phys. Rev. Lett.* **104**(2) (2010) 025003. http://dx.doi.org/10.1103/PhysRevLett.104.025003

[30] J. Vieira *et al.*, *Plasma Phys. Control. Fusion* **54**(12) (2012) 124044. http://dx.doi.org/10.1088/0741-3335/54/12/124044

[31] J. Vieira *et al.*, *Phys. Rev. Lett.* **106**(22) (2011) 225001. http://dx.doi.org/10.1103/PhysRevLett.106.225001

[32] S. Kalmykov *et al.*, *Phys. Rev. Lett.* **103**(13) (2009) 135004. http://dx.doi.org/10.1103/PhysRevLett.103.135004

[33] C.E. Clayton *et al.*, *Phys. Rev. Lett.* **88**(15) (2002) 154801. http://dx.doi.org/10.1103/PhysRevLett.88.154801

[34] M. Tzoufras *et al.*, *Phys. Rev. Lett.* **101**(14) (2008) 145002. http://dx.doi.org/10.1103/PhysRevLett.101.145002

[35] F.S. Tsung *et al.*, *Phys. Rev. Lett.* **93**(18) (2004) 185002. http://dx.doi.org/10.1103/PhysRevLett.93.185002

[36] C.B. Schroeder *et al.*, *Phys. Plasmas* **20**(8) (2013) 080701. http://dx.doi.org/10.1063/1.4817799

[37] M. Padgett *et al.*, *Phys. Today* **57**(5) (2004) 35. http://dx.doi.org/10.1063/1.1768672

[38] J. Vieira and J.T. Mendonça, *Phys. Rev. Lett.* **112**(21) (2014) 215001. http://dx.doi.org/10.1103/PhysRevLett.112.215001

[39] L.-L. Yu *et al.*, *Phys. Plasmas* **21**(12) (2014) 120702. http://dx.doi.org/10.1063/1.4903536

[40] J. Vieira *et al.*, submitted to the Proc. of the Advanced Accelerator Workshop, 2014.

[41] N. Jain *et al.*, Preprint, arXiv:1410.8762 (2014).

[42] A. Caldwell *et al.*, *Nature Phys.* **5** (2009) 363. http://dx.doi.org/10.1038/nphys1248

[43] N. Kumar *et al.*, *Phys. Rev. Lett.* **104**(25) (2010) 255003. http://dx.doi.org/10.1103/PhysRevLett.104.255003

[44] A. Pukhov *et al.*, *Phys. Rev. Lett.* **107**(14) (2011) 145003. http://dx.doi.org/10.1103/PhysRevLett.107.145003

[45] C.B. Schroeder *et al.*, *Phys. Rev. Lett.* **107**(14) (2011) 145002. http://dx.doi.org/10.1103/PhysRevLett.107.145002

[46] R. Assmann *et al.* and (AWAKE Collaboration), *Plasma Phys. Control. Fusion* **56** (8) (2014) 084013. http://dx.doi.org/10.1088/0741-3335/56/8/084013

[47] J. Vieira *et al.*, *Phys. Plasmas* **19**(6) (2012) 063105. http://dx.doi.org/10.1063/1.4725425

[48] J. Vieira *et al.*, *Phys. Rev.* ST-AB **14** (2011) 071303. http://dx.doi.org/10.1103/PhysRevSTAB.14.099901

Laser Wakefield Accelerators

Z. Najmudin
John Adams Institute for Accelerator Science, The Blackett Laboratory, Imperial College London, UK

Abstract
The one-dimensional wakefield generation equations are solved for increasing levels of non-linearity, to demonstrate how they contribute to the overall behaviour of a non-linear wakefield in a plasma. The effect of laser guiding is also studied as a way to increase the interaction length of a laser wakefield accelerator.

Keywords
Laser; wakefield; accelerator; guiding.

1 Introduction

Large amplitude relativistic plasma waves, driven in the wake of a relativistic driver passing through a plasma, are a potential linear accelerator system [1]. This is commonly called a wakefield accelerator. The plasma, being already ionized, can support accelerating fields many orders of magnitude greater than in conventional accelerators. Indeed, at the time of writing, reports of energy gains of ~ 1 GeV in distances of centimetres are now becoming common [2–5].

Possible drivers of the relativistic plasma wave include light [1], electron [6], positron [7], or proton beams [8]. A generalized one-dimensional (1D) wakefield generation equation describing generation by all of the mentioned drivers is [9]

$$\frac{1}{k_{\mathrm{p}}^2}\frac{\partial^2 \phi}{\partial \zeta^2} = \pm\frac{n_{\mathrm{b}}}{n_0} + \gamma_{\mathrm{p}}^2\left\{\beta_{\mathrm{p}}\left[1 - \frac{(1+a^2)}{\gamma_{\mathrm{p}}^2(1+\phi)^2}\right]^{-1/2} - 1\right\},\tag{1}$$

where the potential ϕ is normalized to mc^2/e and the variations are in the quasi-static frame with $\zeta = z - v_{\mathrm{p}}t$ for a driver of velocity v_{p}. The first term on the right describes the response to a particle beam driver and is $+$ for an electron and $-$ for a positively charged driver. In its absence ($n_{\mathrm{b}} = 0$), and in the limit that the phase velocity of the wave $\gamma_{\mathrm{p}} \gg 1$, then expanding the square brackets and also expanding $\beta_{\mathrm{p}} = (1 - \frac{1}{\gamma_{\mathrm{p}}^2})^{1/2}$ gives a non-linear laser wakefield driver equation:

$$\frac{1}{k_{\mathrm{p}}^2}\frac{\partial^2 \phi}{\partial \zeta^2} = \frac{1}{2}\left[\frac{(1+a^2)}{(1+\phi)^2} - 1\right].\tag{2}$$

In the small amplitude limit, $\phi \ll 1$, this can be further simplified to

$$\left(\frac{\partial^2}{\partial \zeta^2} + k_{\mathrm{p}}^2\right)\phi = \frac{1}{2}k_{\mathrm{p}}^2 a^2,\tag{3}$$

which reveals itself to be a driven oscillator. Below, we obtain Eq. (3) starting from the linear forms of the fluid equations of motion, continuity and Gauss's Law. The electric field and density can then be obtained from $\mathbf{E} = -\nabla\phi$ and Poisson's equation. This allows us to obtain some of the common characteristics of a laser driven wakefield accelerator. By adding further levels of non-linearity, we reveal some of the behaviour of laser wakefield accelerators in the non-linear regime.

2 Laser wakefield

2.1 Basic equations

Beginning with Gauss's Law, and the continuity and fluid motion equations for electrons in a plasma in one dimension (the ions are assumed to be heavy and thus fixed) subject to a laser of peak intensity I and wavelength λ with a normalized vector potential $a = eE/m\omega c \approx 8.9 \times 10^{-6} \left(I\lambda^2\right)^{0.5}$ [SI units]:

$$\epsilon_0 \frac{\partial E}{\partial z} = -e(n_e - n_i), \quad \frac{\partial n_e}{\partial t} + \frac{\partial}{\partial z}(n_e v) = 0, \quad \left[\frac{\partial p}{\partial t} + v\frac{\partial p}{\partial z}\right] = -eE + \frac{1}{\gamma_\perp}mc^2\frac{\partial(a^2)}{\partial z},$$

where $p = \gamma mv$ is the momentum of the fluid element. Note that the last term on the right of the equation of motion is the relativistic ponderomotive force, with the relativistic factor due to the quiver motion usually given by $\gamma_\perp = \sqrt{1 + (a^2/2)}$ for linear polarization. It is this ponderomotive force which mediates the interaction between laser and plasma.

For simplicity, we use normalized units: $\epsilon_0, e, m, c = 1$. We also move to the frame in which the laser driver is stationary (the quasistatic approximation), i.e. $\zeta = z - v_p t \approx z - ct$. So $\frac{\partial}{\partial z} = \frac{\partial}{\partial \zeta}$ and $\frac{\partial}{\partial t} = -c\frac{\partial}{\partial \zeta}$. Also using $n_e = n_0 + n_1$, where n_0 is the initial plasma density and n_1 is therefore the amplitude of the plasma wave, the above equations become

$$\frac{\partial E}{\partial \zeta} = -n_1, \quad n_1 = (n_0 + n_1)\beta, \quad (1 - \beta)\frac{\partial p}{\partial \zeta} = E - \frac{1}{\gamma_\perp}\frac{\partial(a^2)}{\partial \zeta}. \tag{4}$$

Here $\beta = v/c$ and is given by

$$\beta = \frac{p}{(1 + p^2 + a^2)^{1/2}},$$

i.e. the relativistic factor of the electrons depends on not only their longitudinal motions but also their transverse oscillations, which at the intensities used to drive laser wakefield accelerators are also relativistic.

Starting from these base equations, we can now consider different levels of non-linearity and assess their effect on wakefield generation.

2.2 Linear wakes

For small amplitudes, we can assume, $n_1 \ll n_0$, $\beta \ll 1$, $\gamma = 1$ and $p = mc\beta$:

$$\frac{\partial E}{\partial \zeta} = -n_1, \quad n_1 = n_0\beta, \quad \frac{\partial \beta}{\partial \zeta} = E - \frac{\partial(a^2)}{\partial \zeta}.$$
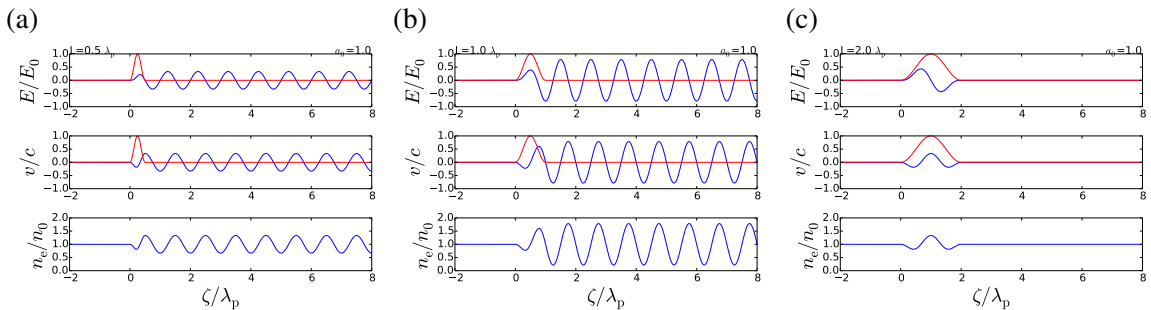
Fig. 1: Linear wakefield generation: plots of E/E_0, v/c, and n_e/n_0 (in blue) for laser pulse $a = a_0 \sin(\pi\zeta/L)$ with pulse length $L = $ (a) $\frac{1}{2}$, (b) 1, (c) 2 λ_p and $a_0 = 0.2$ (in red).

Combining these equations gives us essentially Eq. (2), with (in normalized units) $k_p{}^2 = n_0$. For a normalized laser intensity profile $a^2 = a_0{}^2 \sin^2(\pi\zeta/L)$ for $0 < \zeta < L$, i.e. a half sinusoid, the wake growth is optimized for $L = \lambda_p \equiv 2\pi/k_p$, or alternatively, when the intensity full-width half-max $L_{\text{fwhm}} = \lambda_p/2$. This can be seen in Fig. 1 where these equations have been solved numerically using the scipy.integrate.odeint function in python. One can see that the plasma wave grows at the front end of the laser pulse as the ponderomotive force pushes electrons ahead of it, which then relax to form a plasma oscillation. When resonant, the back of the pulse gives an extra ponderomotive kick to the plasma wave growing it to higher amplitude. At resonance, the wake trailing (behind) the laser pulse has solution

$$n_e = n_0 \left(1 + \frac{\pi}{4}a_0{}^2 \sin k_p\zeta\right), \qquad E_z = E_0 \frac{\pi}{4}a_0{}^2 \cos k_p\zeta, \qquad \phi = -\frac{mc^2}{e}\frac{\pi}{4}a_0{}^2 \sin k_p\zeta,$$

where $E_0 = (mc\omega_p/e)$ is the electric field for sinusoidal oscillations of maximum amplitude, i.e. for $n_1 = n_0$. For yet longer pulse lengths, the wakefield (the plasma wave amplitude following the laser pulse) reduces again, to the point where for $L = 2\lambda_p$, there is no wakefield at all. There is however a driven (or forced) oscillation of the plasma wave whilst the laser pulse is still active (see Fig. 1(c)), which may be useful if only a single bunch needs to be accelerated, rather than a pulse train. There are also smaller amplitude resonances for $L \approx (n + \frac{1}{2})\lambda_p$ for integer $n > 2$.

Clearly there is a problem for linear solutions with $a_0 > 1$, where n_e can become negative. In fact, in the fluid model, electron sheets would cross, so that instead of seeing a returning force, some electrons end up being continually accelerated, a process called 'wavebreaking' [10]. Wavebreaking is interesting because it suggests that plasma electrons can themselves be a source of particles to be accelerated in the accelerator [11]. Thus, this cold linear wavebreaking limit for the electric field $E_0 = (mc\omega_p/e)$ is useful for giving an estimate of the energy output from the accelerator. In the laser driven case, the wakefield travels at the group velocity of the driving pulse, as quantified by $\gamma_{\text{ph}} \approx (\omega_0/\omega_p)$ for a laser of angular frequency ω_0. The maximum energy gain is then given by

$$W_{\text{max}} \approx 2\gamma_{\text{ph}}{}^2 mc^2,$$

and so is dependent on the plasma density. Interestingly, lower density (smaller ω_p, thus higher γ_{ph}) results in greater energy gain. This, however, comes at the expense of the greater laser energy that is required to drive the larger sized wakefield that results from lower density.

2.3 Non-linear wakes

For $n_1 \sim n_0$, the continuity equation gives $n_1 = n_0(1 - \beta)^{-1}$. Since in their fluid motion, the electrons have $-1 < \beta < 1$, one can see that $\frac{1}{2} < n_1/n_0 < \infty$. Because of this non-linearity in the continuity equation, the lowest density in a 1D wake cannot be less than $\frac{1}{2}n_0$. The plasma wave as a result becomes non-linearly steepened, with sharp high-density spikes associated with longer areas of lower electron density plasma. Of course this is only true in one dimension. In three dimensions, at high a_0, transverse motion can lead to $n_e \to 0$, or cavitation. Relativistic effects in the equation of motion add further non-linearity, as in Eq. (4). The effect of these non-linearities can be seen in Fig. 2, for a half sinusoidal laser pulse (as before) with $L = \lambda_p$, and increasing laser strength a_0.

Even for $a_0 = 0.5$ (Fig. 2(a)), the density profiles of the plasma wave have begun to be non-linearly steepened. This is particularly clear for $a_0 = 1.0, 1.5$ (Fig. 2(b,c)), where the plasma wave consists of sharp density spikes either side of shallow long density troughs. As a result, the electric field becomes sawtoothed in nature, with a long region of linearly increasing field followed by a sharp drop due to the density spikes. In the non-linear case too, the fluid approximation breaks down for large amplitudes, as particle sheets cross from neighbouring wave buckets, leading to wavebreaking. Due to the non-linear steepening, the density spikes have much higher density than in the linear case, and the value at which wavebreaking occurs (for a cold relativistic plasma) is given by [12]

$$E_{\text{WB}} = \sqrt{2}(\gamma_{\text{ph}} - 1)^{1/2}E_0.$$
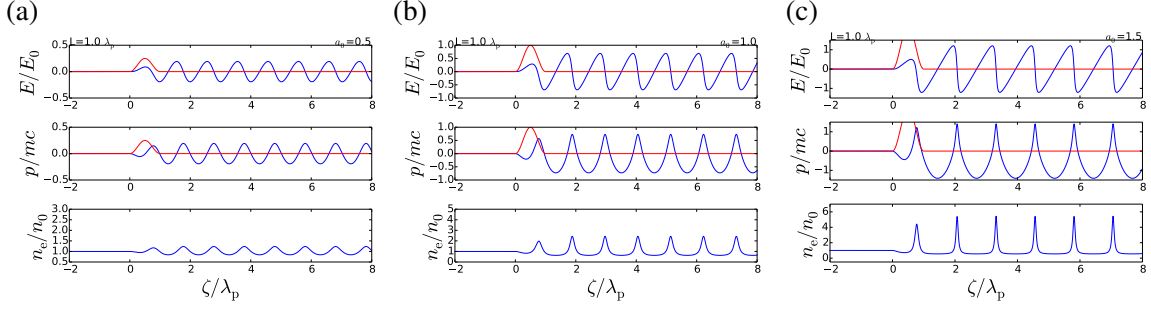
Fig. 2: Non-linear wakefield generation: plots of E/E_0, v/c, and n_e/n_0 (in blue) for laser pulse $a = a_0 \sin(\pi \zeta/L)$ with $a_0 = $ (a) 0.5, (b) 1.0 and (c) 1.5 (in red).

The non-linearities in density also causes lengthening of the plasma wave. This can be seen in Fig. 3(a), where adding relativistic effects results in an increasing lengthening of the laser pulse length at which resonance occurs. This is because the plasma wave wavelength increases as relativistic effects give the plasma electrons greater inertia. Figure 3(b) shows the increase of the maximum electric field with increasing laser strength a_0. At low intensity, the field increases proportional to a_0^2, whereas at high laser strength the dependence becomes closer to a_0. This dependence follows the ponderomotive potential of the laser pulse. Because of the non-linear steepening, the electric field can easily exceed the cold wavebreaking limit E_0 at high a_0. At resonance, the wake amplitude is given by

$$E_\mathrm{max} \approx \frac{a_0^2}{(1 + a_0^2)^{1/2}} E_0.$$
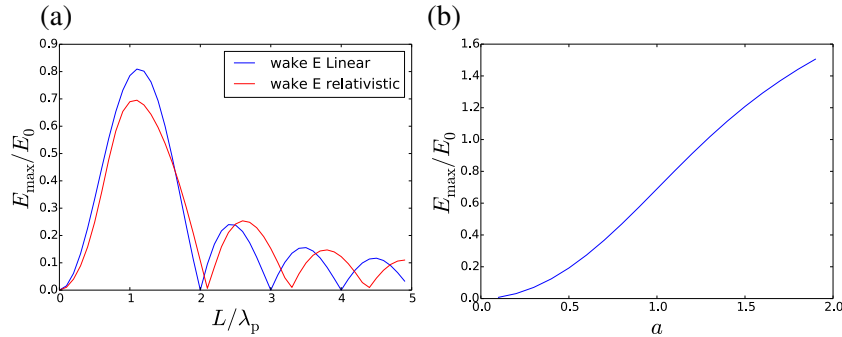
Fig. 3: Wake amplitude for (a) $a_0 = 1$ with varying pulse length for a linear (blue) and non-linear (red) wake, and (b) $L = \lambda_\mathrm{p}$ and varying a_0.

3 Electromagnetic waves

3.1 The wave equation in vacuum

The wave equation in vacuum is

$$\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0.$$

The *paraxial ray approximation* assumes that the phase of the electromagnetic wave varies primarily in z, the direction of propagation, i.e. $\mathbf{E} = E(x, y, z) \exp(\mathrm{i}(kz - \omega t))\hat{\mathbf{x}}$, which is true for gentle focussing

as in wakefield accelerators. So,

$$\nabla^2 \mathbf{E} = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \mathbf{E} + \frac{\partial^2 E}{\partial z^2} \mathrm{e}^{\mathrm{i}(kz-\omega t)} \hat{\mathbf{x}} + \mathrm{i}k\mathrm{e}^{\mathrm{i}(kz-\omega t)} \frac{\partial E}{\partial z} \hat{\mathbf{x}} - k^2 \mathbf{E}.$$

The last term, $k^2\mathbf{E}$, cancels with $\dfrac{1}{c^2}\dfrac{\partial^2 \mathbf{E}}{\partial t^2} = \dfrac{\omega^2}{c^2}\mathbf{E}$ since $\omega \approx ck$, and we can assume that the variation in z is slow, so that $\dfrac{\partial^2 E}{\partial z^2} \to 0$, leading to

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) E - 2\mathrm{i}k \frac{\partial E}{\partial z} = 0. \qquad \text{(cartesian paraxial ray wave equation)}$$

In two dimensions, e.g., taking $\dfrac{\partial^2 E}{\partial y^2} = 0$, the equation forms a time-dependent Schrödinger equation.

3.2 Gaussian optics

For a cylindrically symmetric beam, the paraxial ray equation can be rewritten:

$$\frac{\partial^2 E}{\partial r^2} + \frac{1}{r}\frac{\partial E}{\partial r} - 2\mathrm{i}k \frac{\partial E}{\partial z} = 0. \qquad \text{(paraxial ray wave equation)}$$

This equation has a solution of the form:

$$E(r, z) = \frac{w_0}{w} \exp\left[\frac{-r^2}{w^2} - \frac{\mathrm{i}\pi r^2}{\lambda R} + \mathrm{i}\phi_0 \right],$$

where

$$w = w_0 \sqrt{1 + \left(\frac{z}{z_\mathrm{R}} \right)^2}, \qquad \text{(beam waist)}$$

$$R = \frac{1}{z}\left(z^2 + z_\mathrm{R}{}^2 \right), \qquad \text{(radius of curvature)}$$

$$\tan \phi_0 = \frac{\lambda z}{\pi w_0{}^2}, \qquad \text{(Gouy phase)}$$

where $z_\mathrm{R} = \dfrac{\pi w_0{}^2}{\lambda}$ is the Rayleigh length, and w_0 is the beam waist, and its variation in vacuum can be seen in the Fig. 4. The beam has a minimum at $z = 0$ and like an unconstrained wave packet, will disperse with increasing propagation. R is the curvature of the incoming phase fronts and ϕ_0 is known as the Gouy phase, which flips through zero as a beam passes through focus.

The intensity is then given by $I(r, z)/I_0 = \dfrac{w_0^2}{w^2} \exp\left[\dfrac{-2r^2}{w^2} \right]$, so for $z = \pm z_\mathrm{R}$, the intensity falls by half. As can be seen in Fig. 4, this is also the distance over which the phase fronts are approximately flat. For $z \gg z_\mathrm{R}$, the beam waist expands almost linearly with an angle given by

$$\tan \theta = \lim_{z \gg z_\mathrm{R}} \left[\frac{w}{z} \right] = \frac{w_0}{z_\mathrm{c}} = \frac{\lambda}{\pi w_0}.$$

Or $\theta \sim \dfrac{1}{\pi F}$, where the F-number, $F = \dfrac{w_0}{\lambda} = \dfrac{f}{d}$, which is the ratio of focal length, f, to the beam diameter at the final mirror, d.

For wakefield generation, it would seem advantageous to have as high an intensity as possible, i.e. focusing to small focal spot size. This however implies a short Rayleigh range and so a short acceleration distance. To try to maintain a near constant intensity over the interaction distance ($z_\mathrm{R} > L_\mathrm{deph}$)
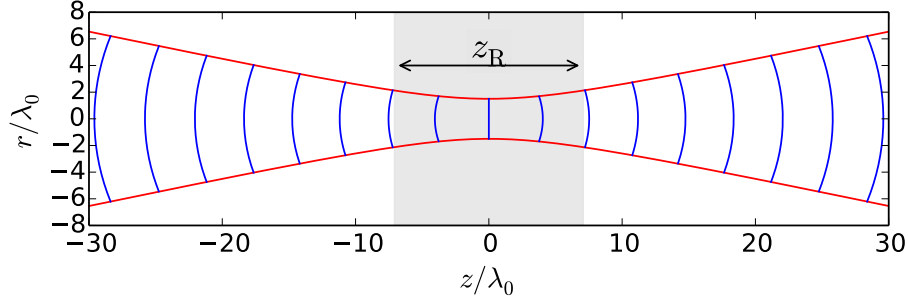
Fig. 4: Evolution of beam waist (red line) of an ideal Gaussian beam, with curvature of phase fronts (blue lines) and Rayleigh range, z_R, (shaded grey) also shown.

would require gentle focussing. This has an implication though for the real size of a plasma accelerator, since the beam diameter is typically constrained by damage thresholds on the final turning mirror ($I_{dam} \sim 10^{12}$ W cm^{-2}), which then has to be further away from final focus for a longer focal length interaction. Luckily, plasma effects can help guide an intense laser beam for distances much longer than the Rayleigh range, alleviating this problem.

3.3 Propagation in plasma

The wave equation including the effect of plasma can be written

$$\nabla^2 \mathbf{E} - \frac{\eta^2}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0,$$

where η, the plasma refractive index, is given by

$$\eta_R \simeq 1 - \frac{\omega_p^2}{2\omega^2} \frac{n(r)}{n_0 \gamma_\perp},$$

for the case of sufficiently underdense plasmas, i.e. $\omega_p^2/\omega^2 \ll 1$. Note that both the density profile $n(r)$ and the relativistic factor $\gamma_\perp(r)$ can now vary radially, in the latter case due to the dependence of $\gamma_\perp$ on intensity, i.e. $\gamma_\perp = \sqrt{1 + (a_0^2/2)}$ for linear polarization.

Either an intensity decrease away from the axis of propagation $(\partial a^2/\partial r^2) < 0$ or a density increase $(\partial n_e^2/\partial r^2 > 0)$ can lead to a higher refractive index on-axis and so, like an optical fibre, causes flow of energy towards the axis or self-focusing. For small variations, the refractive index can be written

$$\eta_R \simeq 1 - \frac{\omega_p^2}{2\omega^2} \left(1 + \frac{\delta n}{n_0} - \frac{a^2}{2}\right),$$

where δn is the density depression, which can be due to a number of effects such as a preformed channel, or ponderomotive expulsion of plasma electrons.

A Gaussian laser pulse with transverse variation $a(r) = a_0 \exp(-r^2/R^2)$, where R is now the beam waist, can be used as a trial solution to the paraxial ray equation. Keeping only dominant terms, the evolution of the beam waist can be approximated by [9]

$$\frac{d^2 R}{dz^2} = \frac{1}{z_R^2 R^3} \left(1 - \frac{P}{P_c}\right).$$

The spot size has been normalized to r_0, the minimum vacuum focal spot size of the laser pulse and the power unit $P_c \simeq 17(\omega^2/\omega_p^2)$ GW, which is dependent on density. Figure 5 demonstrates the effect of
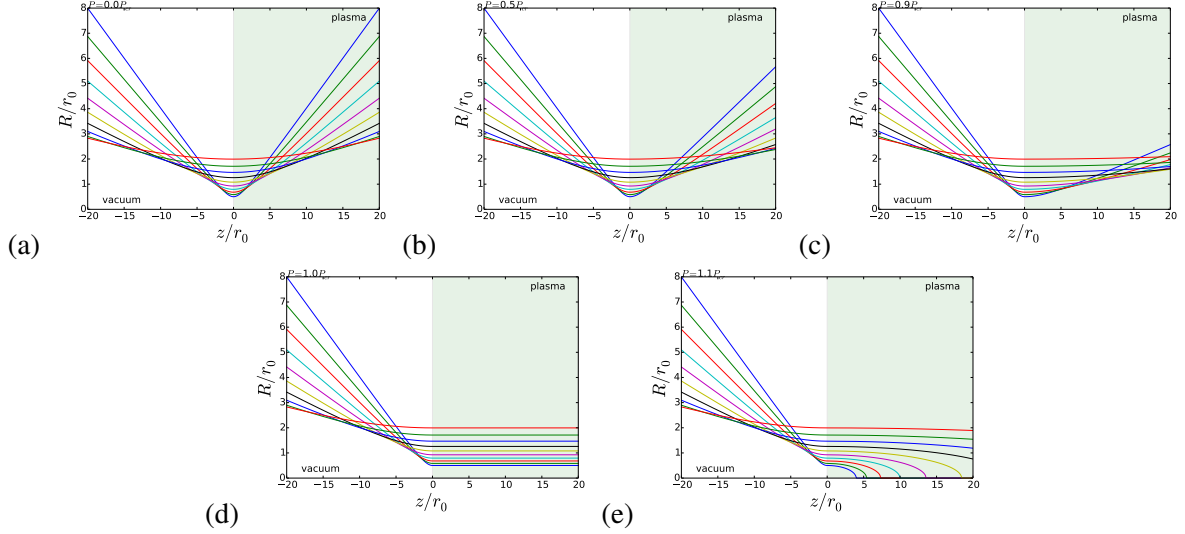
Fig. 5: Behaviour of laser beam waist focussed on boundary of sharp density plasma transition for $P/P_{cr} = $ (a) 0, (b) 0.5, (c) 0.9, (d) 1.0 and (e) 1.1.

P/P_{cr} on the evolution of the beam waist. Increasing P/P_{cr} leads to a greater influence of the second term, which is due to intensity effects, over the first term which causes diffraction. At $P = P_{cr}$, a balance is formed between diffraction and self-focusing and the laser pulse can be guided. Hence this is a *critical* power for self-focusing. For $P > P_{cr}$, the model predicts catastrophic self-focusing. In reality however, including higher order diffraction terms prevents this as strong focussing invariably leads to generation of higher modes and thus filamentation [13].

3.4 Guiding

For wakefield driving short pulses, the leading edge of the pulse interacts with the wrong curvature of density due to the plasma wave (increasing on-axis) and this can cause increased diffraction of the laser pulse. This is true even if the body of the laser pulse is being self-focused. To overcome this, guiding channels are used, where the guiding of the pulse is now determined by

$$\frac{\mathrm{d}^2 R}{\mathrm{d}z^2} = \frac{1}{z_R{}^2 R^3}\left(1 - \frac{\delta n}{\delta n_c}R^4\right).$$

A Gaussian beam with $R = 1$, i.e. $r = r_0$, can be guided, provided there is a density depression with $\delta n = \delta n_c$, where

$$\delta n_c = \frac{1}{\pi r_e r_0{}^2}, \tag{5}$$

and $r_e = e^2/m_e{}^2 c^2$ is the 'classical radius of the electron'. Such channels can be formed, for example, by thermal equilibration of a discharge with its cold walls, or by ponderomotive or thermal expulsion of a plasma by a secondary laser pulse [14]. Channel guiding allows laser energy to be used more efficiently in a laser wakefield and also allows potentially simpler operation in the linear regime.

There is a critical beam or 'matched' spot size at which the beam propagates with minimum channel waist oscillations (see Fig. 6), as can be found by inverting Eq. (5). The goal for laser wakefield operation is to match the Gaussian focussed laser spot to the matched spot of the accelerator, which determines the required focussing parameter. However, as can be seen in the figure, not focussing at the matched spot size will lead to channel oscillations which can lead to a loss of guided energy, as the channels are invariably 'leaky'.
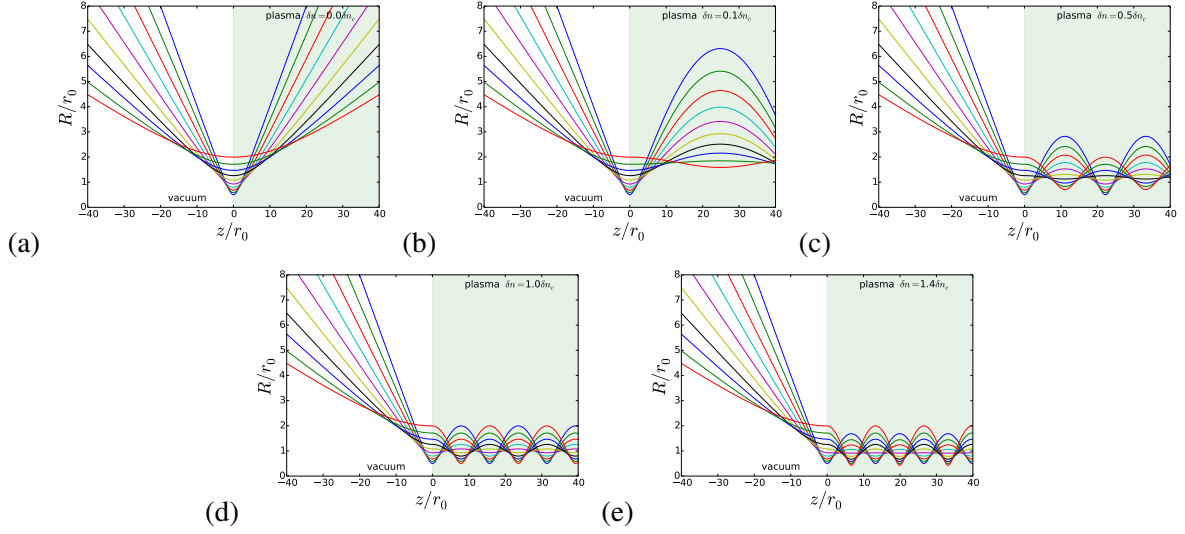
Fig. 6: Behaviour of the waist of the leading edge of a laser pulse focussed on boundary of sharp density transition on a plasma with radial guiding channel with $\delta n / \delta n_{\mathrm{cr}} =$ (a) 0, (b) 0.1, (c) 0.5, (d) 1.0 and (e) 1.4.

3.5   Propagation effects

A laser pulse of course does not stay unchanged as it drives a wakefield, it must lose energy. However this does not necessarily lead to a reduction in intensity, since the pulse can compress as it propagates. In a linear wake, the rate of compression is given by the difference in the group velocity from the first maximum to minimum of the plasma wave. This can be seen in Fig. 7 which shows that due to the density rise, the front of the laser pulse has a slower group velocity than the rear of the pulse, causing the back to 'catch-up' and the pulse to compress. For the case of a pulse with $L \sim \lambda_{\mathrm{p}}$, the variation of pulse length as a function of propagation distance ℓ is [15]

$$\tau = \tau_0 - \frac{n_{e0} l}{2 n_{\mathrm{cr}} c}.$$

The pulse compression has the beneficial effect that the power of the driving pulse can stay relatively constant, even though the energy within the laser pulse is being progressively reduced. The laser pulse also redshifts from the front as those photons that drive the wakefield lose energy. This almost complete loss of energy of the driving photons leads to an etching of the laser pulse. By contrast the trailing photons in the wake are 'squeezed' as they travel in the density depression and so become blue shifted. This process is often called photon acceleration and the amount of blue shifting (energy gain) of the photon is given by [16]

$$\delta \omega = \omega_0 \left(1 - z \frac{\mathrm{d} \beta_{\mathrm{p}}}{\mathrm{d} \zeta} \right) \simeq \omega_0 \left(1 - z \frac{\mathrm{d}}{\mathrm{d} \zeta} \left(\frac{\delta n}{n_0} \right) \right).$$

Photon acceleration can be a useful diagnostic of wakefields in cases where there are no charged particles available to accelerate [17].

4   Conclusion

We have discussed a number of effects which are important for laser wakefield acceleration. We have shown how non-linear effects increase the plasma wave wavelength and also allow larger wakefield amplitudes at high intensities ($a_0 > 1$). This is due to the characteristic peak and trough nature of the non-linear wakes in this regime.
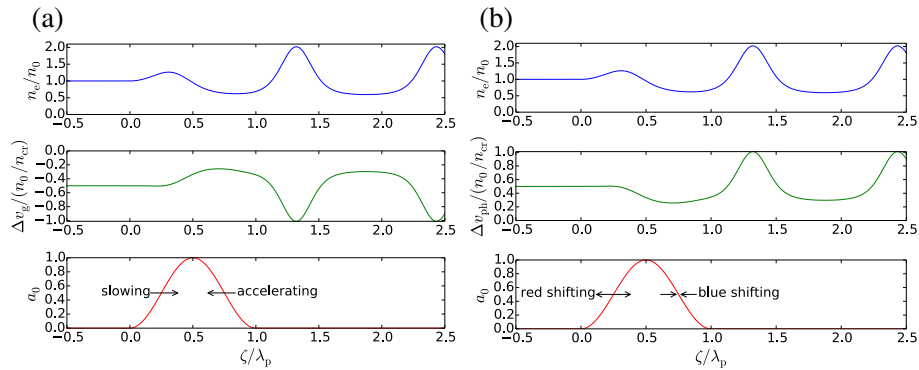
Fig. 7: Compression and photon acceleration of a wakefield driving laser pulse for $a_0 = 1$, $L = \lambda_p$. (a) n_e (top), *group* velocity (middle) and pulse shape (bottom) showing regions of speeding and slowing of the pulse. (b) n_e (top), *phase* velocity (middle) and pulse shape (bottom) showing regions of stretching and compressing of phase.

We have also discussed simple models for the guiding of laser pulses which allow the length of a laser wakefield well beyond the Rayleigh range of the laser pulse. This is important as it allows the full energy gain to be extracted from the laser wakefield. We also discussed how preformed channels can enhance the guiding effect for a laser pulse focussed to the matched spot size of the channel.

References

[1] T. Tajima and J. M. Dawson, *Phys. Rev. Lett.* **43** (1979) 267. http://dx.doi.org/10.1103/PhysRevLett.43.267.

[2] S. Kneip *et al.*, *Phys. Rev. Lett.* **103** (2009) 035002. http://dx.doi.org/10.1103/PhysRevLett.103.035002

[3] X. Wang *et al.*, *Nature Commun.* **4** (2013) 1988. http://dx.doi.org/10.1038/ncomms2988

[4] H.T. Kim, K.H. Pae, H.J. Cha, I.J. Kim, T.J. Yu, J.H. Sung, S.K. Lee, T.M. Jeong and J. Lee, *Phys. Rev. Lett.* **111** (2013) 165002. http://dx.doi.org/10.1103/PhysRevLett.111.165002

[5] W.P. Leemans *et al.*, *Phys. Rev. Lett.* **113** (2014) 245002. http://dx.doi.org/10.1103/PhysRevLett.113.245002

[6] P. Chen, J.M. Dawson, R.W. Huff and T. Katsouleas, *Phys. Rev. Lett.* **54** (1985) 693. http://dx.doi.org/10.1103/PhysRevLett.54.693

[7] B.E. Blue *et al.*, *Phys. Rev. Lett.* **90** (2003) 214801. http://dx.doi.org/10.1103/PhysRevLett.90.214801

[8] A. Caldwell, K. Lotov, A. Pukhov and F. Simon, *Nature Phys.* **5** (2009) 363–367. http://dx.doi.org/10.1038/nphys1248

[9] E. Esarey, P. Sprangle, J. Krall and A. Ting, *IEEE Trans. Plasma Sci.* **24** (1996) 252–288. http://dx.doi.org/10.1109/27.509991

[10] J.M. Dawson, *Phys. Rev.* **113** (1959) 383. http://dx.doi.org/10.1103/PhysRev.113.383

[11] A. Modena *et al.*, *Nature* **377** (1995) 606. http://dx.doi.org/10.1038/377606a0

[12] A.I. Akhiezer and R.V. Polovin, *Sov. Phys. JETP* **3**(5) (1956) 696.

[13] A.G.R. Thomas *et al.*, *Phys. Rev. Lett.* **98** (2007) 2. http://dx.doi.org/10.1103/PhysRevLett.98.054802

[14] E. Esarey, P. Sprangle, J. Krall, and A. Ting, *IEEE J. Quantum Electron.* **33,** (1997) 1879. http://dx.doi.org/10.1109/3.641305

[15] J. Schreiber *et al.*, *Phys. Rev. Lett.* **105** (2010) 235003.

http://dx.doi.org/10.1103/PhysRevLett.105.235003

[16] S.C. Wilks, J.M. Dawson, W.B. Mori, T. Katsouleas and M.E. Jones, *Phys. Rev. Lett.* **62** (1989) 2600. http://dx.doi.org/10.1103/PhysRevLett.62.2600

[17] C.D. Murphy *et al.*, *Phys. Plasmas* **13** (2006) 033108. http://dx.doi.org/10.1063/1.2178650

Bibliography

P. Gibbon, *Short Pulse Laser Interactions with Matter* (World Scientific Publishing Company, New York, 2004).

Beam-driven, Plasma-based Particle Accelerators

P. Muggli
Max Planck Institute for Physics, München, Germany

Abstract

We briefly give some of the characteristics of the beam-driven, plasma-based particle accelerator known as the plasma wakefield accelerator (PWFA). We also mention some of the major results that have been obtained since the birth of the concept. We focus on high-energy particle beams where possible.

Keywords

Plasma; particle accelerator; plasma-based accelerator; plasma wakefield; beam-plasma interaction.

1 Introduction

In plasma-based particle accelerators (PBPAs), particles are accelerated by the wakefields sustained by a periodic plasma density perturbation. It was first realized that wakefields can be driven in plasmas by intense laser pulses [1]. In this case, it is the ponderomotive force associated with the gradient of the laser pulse intensity and not the oscillating electric field itself that acts to displace the plasma electrons and drive the wakefields. Such a PBPA is called a laser wakefield accelerator (LWFA).

Soon after this, it was realized that an intense, relativistic charged-particle bunch can also drive the wakefields [2]. In this case, the (unipolar) space charge field of the bunch acts on the plasma electrons. A PBPA driven by a charged-particle bunch is called a plasma wakefield accelerator (PWFA). This is a particular type of a collinear wakefield accelerator.

The PWFA is one of the advanced accelerator schemes studied as a high-gradient alternative to today's RF technology.

It is, of course, hopeless to pretend to summarize PWFA research in a few pages. However, we attempt to touch on some of the most relevant points, to give an introduction to the field. The many details missing here can be found in the references provided and in other articles. It is left to the reader to do the detailed work, which is the essence of the learning process.

The text is organized as follows. First, we outline a few characteristics of charged particles and of charged-particle bunches relevant for the PWFA. Then we briefly describe how wakefields are driven in the plasma. We next introduce the concept of the transformer ratio. After that, we summarize two-dimensional PWFA linear theory and the relevance of the longitudinal and transverse dimensions of the bunch to the driving process. We also introduce the concepts of the wave-breaking field and beam loading. We then introduce the non-linear regime of the PWFA, including beam focusing, propagation and acceleration, both for electron and positron bunches. The possibility of using a hollow plasma channel for positron bunches is also briefly addressed. We mention the intermediate, quasi-linear or weakly non-linear regime of the PWFA, as well as the self-modulation instability. We end with a few remarks.

2 Charged particles, charged-particle bunches and the PWFA

Charged-particle bunches have a number of characteristics that make them particularly suitable for driving wakefields for PBPAs.

- The electric field of a relativistic charge or bunch is essentially transverse. This can be seen by using the Lorentz transform for the purely radial electric field of a single particle in its rest frame:

$E_{r0} = \frac{q}{4\pi\epsilon_0}\frac{1}{r^2}$. In the laboratory frame where the particle has relativistic factor γ, the fields are $E_{||} = E_{r0}$, $E_\perp = \gamma E_{r0}$, $B_{||} = 0$, $B_\perp = (v_b/c^2)E_{r0}$. Therefore, $E_\perp = \gamma E_{||} \gg E_{||}$ for $\gamma \gg 1$.

– A charged-particle bunch experiences a space charge force that makes it diverge (transversely) and lengthen (longitudinally). However, when the particles are relativistic, their relative distance does not change significantly over distances of interest (a few metres for PBPAs). For particles with energies γ and $\gamma + \Delta\gamma$, with $\Delta\gamma \ll \gamma$, the dephasing ΔL over a propagation distance L is $\Delta L/L \cong (1/\gamma^2)(\Delta\gamma/\gamma)$. This dephasing is usually small compared with the length of interest, the plasma wavelength $\lambda_{pe} = 2\pi c/\omega_{pe}$ over plasma lengths utilized in PBPAs. There is, therefore, no lengthening of the relativistic bunch under the conditions of interest here. In the transverse plane, the bunch is subject to the full Lorenz force, $F_\perp = q_b(E_\perp + v_b \times B_\perp)$. Here, $v_b = (1 - 1/\gamma^2)^{1/2} c$ is the velocity of the bunch particles; in cylindrical coordinates, v_b is along the z-axis and $E_\perp = E_r$, $B_\perp = B_\theta$. To evaluate the fields, we need to assume a cylindrical infinitely long *bunch* or beam, with uniform density n_b. In this case, the radial electric field within the bunch radius is simply given by Gauss' law: $E_r = \frac{1}{2}\frac{q_b n_b}{\epsilon_0}r$. The magnetic field is given by Faraday's law, $B_\theta = \frac{1}{2}(\mu_0 q_b n_b)r$, with $q_b = \pm e$ the charge of a bunch particle. The total force therefore reads, with $c^2 = 1/\mu_0\epsilon_0$,

$$F_\perp = q_b(E_r + v_b \times B_\theta) = q_b\frac{1}{2}\frac{q_b n_b}{\epsilon_0}\left(1 - \frac{v_b^2}{c^2}\right)r = \frac{1}{\gamma^2}q_b E_r . \tag{1}$$

That is, the total transverse force is reduced to $1/\gamma^2$ multiplied by the pure space charge force $F_{sc} = q_b E_r$, i.e., as soon as the particles become relativistic ($\gamma \gg 1$), the transverse dynamics are dominated by the emittance and the external focusing forces.

– A beam with geometric emittance ϵ_g has a beta-function at its waist, defined as $\beta_0 = \sigma_0^2/\epsilon_g$, where its transverse r.m.s. size is σ_0. This parameter depends on the beam emittance, which can, in principle, be made very small. The beta-function of a charged-particle beam is the equivalent of the Rayleigh length of a photon beam focused to a transverse size w_0: $Z_R = \pi\frac{w_0^2}{\lambda_0}$, where λ_0 is the wavelength of the laser pulse. In both cases, the transverse size of the beam increases by a factor of $\sqrt{2}$ over β_0 or Z_R. Since the vast majority of today's laser pulses that are short and intense enough to drive a PBPA have $\lambda_0 \cong 800$ nm, the Rayleigh length is determined only by the focal size, w_0. For example, choosing a typical value $\sigma_0 = w_0 = 10$ μm, we obtain $\beta_0 = 0.8(0.08)$ m (for a typical Stanford Linear Accelerator Center (SLAC) beam geometric emittance, $\epsilon_g = 1.26 \times 10^{-9}(1.26 \times 10^{-10})$ m rad) in the x-(y-)plane and $Z_R = 400$ μm. The value of β_0 and Z_R determines the distance over which the beam remains small and can thus drive wakefields (in the absence of external forces or a guiding mechanism). A future linear collider beam is expected to have an ϵ_g of the order of $(1 - 8) \times 10^{-11}$ m rad from 200 and 500 GeV [3].

– The space charge field of the bunch is (at least) partially cancelled by the plasma (see P. Gibon's lecture). Its current is also cancelled by the plasma return current. Even an initially radially uniform plasma, therefore, focuses the bunch and counters its natural divergence, resulting from its non-zero incoming emittance. In addition to the beam having a relatively long beta-function, plasma focusing can maintain a small beam transverse size and the wakefield driving and acceleration over long distances.

– The velocity of the particle bunch in the plasma is independent of the plasma density. For a bunch with particle energy E_0 and relativistic factor $\gamma = E_0/m_e c^2 - 1 \propto E_0/m_e c^2$ it is simply given by the particle's velocity: $v_b = (1 - 1/\gamma^2)^{1/2} c$. The laser pulse velocity is given by the group velocity of the light in the plasma and around frequency ω_0: $v_g = (1 - \omega_{pe}^2/\omega_0^2)^{1/2} c \le c$. Here $\omega_{pe} = (n_{e0}e^2/\epsilon_0 m_e)^{1/2}$ is the plasma electron (angular) frequency in a plasma of (electron) density n_{e0}.

In the next section, we describe the principles of the PWFA.

3   The PWFA

In the PWFA, the mostly transverse space charge field of the relativistic charged-particle bunch travelling in a neutral plasma displaces the plasma electrons (Fig. 1). The positively or negatively charged bunch driving the wakefields is called the drive bunch. The plasma ions experience the same force as the electrons but, because of their larger mass or inertia, respond on a much longer time scale, of the order of the inverse of the ion plasma frequency, $\omega_{\mathrm{pi}} = \left(n_{\mathrm{i0}} Z^2 e^2 / \epsilon_0 M_i\right)^{1/2}$. In an initially neutral plasma, the ion density is $n_{\mathrm{i0}} = n_{\mathrm{i0}}$. Here, Z is the number of ionized electrons per atom ($Z = 1$ for the protons of a hydrogen plasma or singly ionized atom with more electrons) and M_{i} is the ion mass. Ions are thus usually considered as immobile over the typical $\omega_{\mathrm{pe}}^{-1}$ time scale of a single wakefield period since $\omega_{\mathrm{pe}} \gg \omega_{\mathrm{pi}}$. When the drive bunch density is much larger than the plasma density, the ions can move over the same time scale as the electrons and ion motion must be considered [4–6].
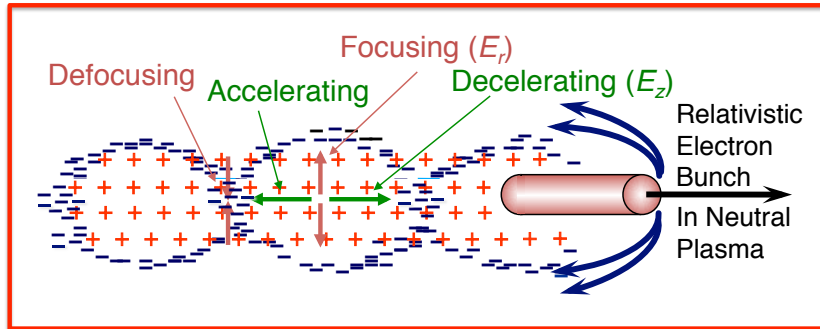
Fig. 1: Cartoon of the PWFA principle. The electron bunch displaces the plasma electrons and forms regions of the plasma that are globally positively and negatively charged. This electron plasma density perturbation sustains the wakefields. The main electric field components of the wakefields are shown and can be schematically determined from the charge distribution. This is the case of a bunch with density larger than the plasma density ($n_{\mathrm{b}} \gg n_{\mathrm{e0}}$, non-linear PWFA regime) and a pure ion column is formed. The arrows showing the backward motion of the plasma electrons suggest this as seen in a window moving with the drive bunch.

In the case of a negatively charged drive bunch (e.g., an electron bunch), once displaced, the plasma electrons feel the restoring force of the plasma ions, are attracted back to the axis, overshoot and oscillate. This oscillation of the plasma electrons with period $\cong 2\pi/\omega_{\mathrm{pe}}$ and phased by the drive bunch moving at approximately the speed of light is the plasma wake. The direction of the fields can be determined form the charge distribution in the second accelerator structure or *bubble* (see Fig. 1). On the axis there is an alternation of positively and negatively charged regions. In the front of the structure, the longitudinal E_z field is in the forward direction (for the bunch moving to the right), corresponding to a decelerating field or force. The drive bunch particles lose energy to the plasma in expelling the plasma electrons (for the case of the negatively charged drive bunch). In the back of the structure, the E_z field direction is reversed, allowing for the possibility for particles in the back of the bunch, or in a trailing witness bunch, to gain energy from the wakefields. In the middle of the structure, the transverse field points outwards, corresponding to a focusing force for this drive bunch or for a witness bunch of the same charge sign. Between the structures, regions of strongly compressed plasma electron density correspond to defocusing regions.

The wakefield can be driven (energy loss) and sampled (energy gain) by the electrons of a single bunch, approximately one plasma wavelength long, as suggested by Fig. 1 and as demonstrated next. However, this leads to a large final energy spread, since all phases of the wakefields are sampled.

Wakefields can be driven by a shorter bunch and sampled by another short, trailing bunch, called a witness bunch. This is shown in Fig. 2. This can lead to a narrow final energy spectrum for the witness bunch, as demonstrated in Ref. [7] and recently in Ref. [8].

The drive bunch can also be positively charged, as suggested in Fig. 2. In this case, the plasma electrons are first attracted towards the beam axis, but then sustain the same kind of wakefield, with just a phase shift with respect to the negatively charged drive bunch.
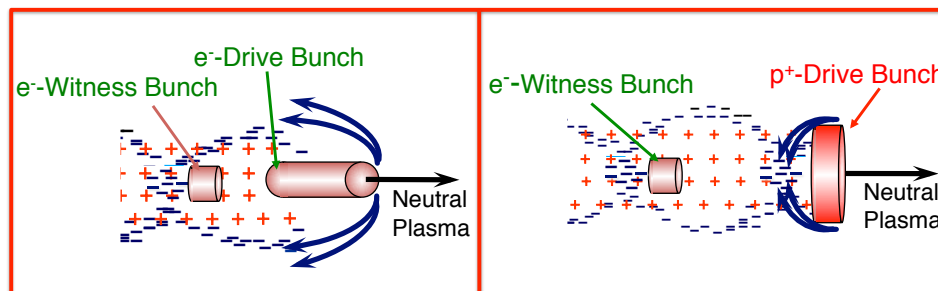
Fig. 2: The left panel shows the wakefields driven by a negatively charged drive bunch. The witness bunch that is accelerated is also negatively charged. The right panel shows a schematic of the wakefields driven by a positively charged drive bunch. The witness bunch is negatively charged.

Driving wakefields with a charged particle bunch was first demonstrated in Ref. [9], interestingly, with a drive-witness electron bunch train. Driving wakefields with a positively charged particle bunch was first demonstrated with positrons in Ref. [10].

The witness bunch can also have a positive charge (e.g., positrons). In this case, it must be placed in the corresponding accelerating and focusing phase (region) of the wakefields.

4 Transformer ratio

The transformer ratio is an important concept for the PWFA. It is the ratio of the accelerating field amplitude behind the drive bunch(es) to the decelerating field amplitude *within* the drive bunch(es). It can be defined from general wakefield characteristics (i.e., also in RF systems).

Consider the wakefield amplitude per unit charge $W(\xi)$ at a position ξ at or behind a particle. Following Ref. [11], the rate of energy change (loss in this case) by an infinitely short bunch with N_1 charges e of energy E_1 (per particle) located at $\xi = 0$ is

$$\frac{\mathrm{d}(N_1 E_1)}{\mathrm{d}z} = -(N_1 e)\, N_1 e W(0) = -N_1^2 e^2 W(0) \,. \tag{2}$$

Similarly, the rate of energy change of the second bunch with charge $N_2 e$ and energy E_2 located at $\xi = \xi_2$ is

$$\frac{\mathrm{d}N_2 E_2)}{\mathrm{d}z} = -N_2^2 e^2 W(0) + (N_2 e)\, N_1 e W(\xi_2) = -N_2^2 e^2 W(0) + N_1 N_2 e^2 W(\xi_2) \,. \tag{3}$$

Note that N_1 and N_2 remain constant along the plasma and that the result is independent of the sign of the charge e; it depends on the sign of $W(\xi_2)$, i.e., on the position (or phase) of the second bunch in the wakefield of the first one. The sum of the energy change by the two bunches must be smaller or equal to zero, thus

$$\frac{\mathrm{d}(N_1 E_1)}{\mathrm{d}z} + \frac{\mathrm{d}(N_2 E_2)}{\mathrm{d}z} = -N_1^2 e^2 W(0) - N_2^2 e^2 W(0) + N_1 N_2 e^2 W(\xi_2) \leq 0 \,, \tag{4}$$

or

$$\left(N_1^2 + N_2^2\right) W(0) + N_1 N_2 W(\xi_2) \geq 0 \,. \tag{5}$$

Since this relation has to be true for all values of N_1 and N_2, we obtain

$$2W(0) \le -W(\xi_2) \ . \tag{6}$$

The accelerating gradient G is, from Eqs. (3) and (6),

$$G = \frac{\mathrm{d}E_2}{\mathrm{d}z} = -N_2 e^2 W(0) + N_1 e^2 W(\xi_2) \le -N_2 e^2 W(0) + 2N_1 e^2 W(0) = e^2 W(0)\left(2N_1 - N_2\right) \ . \tag{7}$$

The distance over which the first bunch loses all its energy is

$$N_1 E_1 = N_1^2 e^2 W(0)L \ , \tag{8}$$

or

$$L = \frac{E_1}{N_1 e^2 W(0)} \ . \tag{9}$$

Thus, the change in energy of the trailing particles is

$$\Delta E_2 = GL \le \frac{e^2 W(0)}{N_1 e^2 W(0)} E_1 \ , \tag{10}$$

and

$$\Delta E_2 \le E_1 \left(2 - \frac{N_2}{N_1}\right) \ . \tag{11}$$

This expression is a maximum for $N_2 = 0$, in which case $\Delta E_2 \le 2E_1$, showing that the energy gain per trailing particle is at most equal to twice the incoming energy of a drive particle. Since both drive and witness bunches travel the same distance in the plasma, the transformer ratio can be defined from the longitudinal wakefield amplitude: $R = E_+/E_-$. Here, E_- is the maximum decelerating field *within* the drive bunch and E_+ is the maximum accelerating field *behind* the drive bunch, where the particles of the witness bunch can be placed. Note that global energy conservation implies that $N_2 \Delta E_2 \le N_1 E_1$, or, more generally, $N_2 \Delta E_2 \le N_1 \Delta E_1$ over a distance smaller than or equal to the full energy depletion length defined in Eq. (9). It is also important to note that these expressions relate the energy change (loss/gain) of the different particles. This means that, for example, a bunch of 20 GeV particles could lose (almost, they must remain relativistic) all their energy (i.e., lose 20 GeV per particle) and transfer 40 GeV to each particle of a witness bunch of 100 GeV. This is what would happen in a staged accelerator, where each plasma section (stage) would be driven by a 20 GeV bunch, while a witness bunch would keep gaining (up to) 40 GeV per stage.

This concept is similar to that for an electrical transformer, $U_2 I_2 \le U_1 I_1$ (index 1 for primary and 2 for secondary; this expression is written for power rather than energy), where the number of particles per bunch divided by the bunch length is the current. The PWFA is a transformer in which the energy of a high-charge, low-energy bunch is extracted and added to that of a lower-charge witness bunch, through the plasma wakefields.

It is interesting to notice that the transformer ratio concept is valid for all collinear accelerators and that the expressions were derived from very simple and general assumptions of energy conservation and linear superposition of the wakefields. In real systems, the bunches can evolve in the wakefields and other considerations that come into play (beam loading, etc.) that can reduce the transformer ratio reached.

Transformer ratios much larger than two can be obtained by tailoring the bunch longitudinal density or current profile, or by using a sequence of bunches [12]. The transformer ratio has been measured in PWFA experiments [13].

5 PWFA linear regime

The linear regime is interesting because in this regime many fundamental aspects of the PWFA can be calculated directly. It is, a priori, not interesting for a collider because (by definition) it does not use the full potential of the plasma for sustaining large fields. These large fields are reached when the plasma perturbation is also large, i.e., in the non-linear regime of the PWFA. In addition, in this regime the wakefields have continuous longitudinal and radial variations and the beam and plasma characteristics are directly proportional to each other (by definition). Therefore, the characteristics of the system generally evolve along the propagation distance, and are not necessarily suitable for a long-lasting acceleration process that aims to produce a high-quality bunch with a small final energy spread and emittance.

5.1 Linear theory

The linear theory of the PWFA can be derived using a cold, non-relativistic fluid model for the plasma. A clear derivation can be found in Ref. [14]. This model uses Newton's equation for a fluid element, the continuity equation and Poisson's equation. The equations are linearized (see examples in Ref. [15] for the linearization process) and a wave equation for the plasma electron density perturbation $n_1 \ll n_{\mathrm{e}0}$ can be derived. It is driven by the bunch density n_{b} and reads

$$\frac{\partial^2 n_1}{\partial \xi^2} + k_{\mathrm{pe}}^2 n_1 = \frac{q_{\mathrm{b}}}{e} k_{\mathrm{pe}}^2 n_{\mathrm{b}} \,, \tag{12}$$

where $n_{\mathrm{b}} \ll n_{\mathrm{e}0}$ is assumed. These equations are written in the coordinate system of the bunch, i.e., $\xi = z - v_{\mathrm{b}}t$, which is often written as $\xi = z - ct$ for relativistic bunches. Equations are written as a function of space coordinates (z and ξ) rather than the time coordinate, a more natural choice for a linear accelerator. In these coordinates, $\frac{\mathrm{d}}{\mathrm{d}z} = \frac{\mathrm{d}}{\mathrm{d}\xi}$ and $\frac{\mathrm{d}}{\mathrm{d}t} = -v_{\mathrm{b}}\frac{\mathrm{d}}{\mathrm{d}\xi} \cong -c\frac{\mathrm{d}}{\mathrm{d}\xi}$. Note the term related to the bunch charge (e.g., $q_{\mathrm{b}}/e = \mp 1$ for an electron or positron bunch), which is the source term driving the (harmonic) oscillator. Note also that a similar equation can be obtained for the case of a laser pulse driving the plasma density perturbation with the ponderomotive force as the source term. This equation can be solved in 1D for a delta function bunch or charge to obtain the Green's function for the wakefield generation (n_1 perturbation) and the longitudinal wakefields using Poisson's equation. This longitudinal component has a $\cos(k_{\mathrm{pe}}\xi)$ dependency. The radial dependency can be obtained from 2D theory for a given transverse density profile (Gaussian, parabolic, etc.) and involves Bessel functions [16] and a $\sin(k_{\mathrm{pe}}\xi)$ dependency. For a bunch with longitudinal and radial Gaussian profiles, the wakefields read

$$W_z(\xi, r) = \frac{e}{\epsilon_0} \int_{-\infty}^{\xi} n_{\mathrm{b}_\parallel}(\xi') \cos\left[k_{\mathrm{pe}}(\xi - \xi')\right] \mathrm{d}\xi' \cdot R(r) \,, \tag{13}$$

$$W_\perp(\xi, r) = \frac{e}{\epsilon_0 k_{\mathrm{pe}}} \int_{-\infty}^{\xi} n_{\mathrm{b}_\parallel}(\xi') \sin\left[k_{\mathrm{pe}}(\xi - \xi')\right] \mathrm{d}\xi' \cdot \frac{\mathrm{d}R(r)}{\mathrm{d}r} \,, \tag{14}$$

where $R(r)$ is the transverse dependency given by

$$R(r) = k_{\mathrm{pe}}^2 \int_0^r r' \, \mathrm{d}r' n_{b\perp}(r') I_0(k_{\mathrm{pe}}r') K_0(k_{\mathrm{pe}}r) + k_{\mathrm{pe}}^2 \int_r^\infty r' \mathrm{d}r' n_{b\perp}(r') I_0(k_{\mathrm{pe}}r) K_0(k_{\mathrm{pe}}r') \,, \tag{15}$$

where I_0 and K_0 are the zeroth-order modified Bessel functions of the first and second kind, respectively. Note that the W notation is typical of wakefields in structures. It is important to understand the meaning of these equations.

On the one hand, in linear theory, W_z is associated only with an E_z component, since v_{b} is along the z-axis and the Lorentz force has no magnetic contribution along z (to first, linear order). On the other hand, $W_\perp$ has two components and is often written as $E_r - v_{\mathrm{b}}B_\theta$ in cylindrical coordinates. The terms have a ξ dependency that is a pure 1D, longitudinal dependency. This dependency is obtained from the Green's function for wakefield excitation. The $R(r)$ and $\mathrm{d}R/\mathrm{d}r$ terms express the deviation

from 1D theory obtained in 2D. The longitudinal (accelerating) wakefield is a maximum on the axis with $R(r = 0)$ (see [17]). For $k_{pe}\sigma_r \gg 1$, i.e., $\sigma_r \gg c/\omega_{pe}$, $R(r = 0) \rightarrow 1$, i.e., one recovers the 1D result, as expected. For $k_{pe}\sigma_r \ll 1$, i.e., $\sigma_r \ll c/\omega_{pe}$, $R(r = 0) \rightarrow k_{pe}^2\sigma_r^2 (0.0597 - \ln(k_{pe}\sigma_r))$. This means that when the plasma density is decreased for a fixed σ_r (or σ_r for a fixed n_{e0}, and in both cases a fixed bunch charge) the radial component contributes to a decrease in W_z, even if n_b/n_{e0} increases. This is because 1D theory applies to an infinite sheet bunch (in the perpendicular direction). In 2D, the plasma responds with a natural transverse extent of (a few) c/ω_{pe}. Therefore, when $k_{pe}\sigma_r \ll 1$, the charge contained in the (few) c/ω_{pe} is less than in the 1D case and so is W_z (for a given charge density).

A few important remarks about these equations:

– They are symmetrical with respect to the bunch charge sign, which means that, in the linear regime, the wakefields are the same for both charge signs, to within a phase factor of $\pi/2$.
– The wakefields are necessarily initially decelerating within the drive bunch because it displaces the plasma electrons and has to lose energy in the process. The wakefields are necessarily initially focusing within the drive bunch because the plasma electrons move to shield the bunch fields, i.e., to neutralize the bunch charge and decrease the space charge fields. The bunch is thus focused by the $v_b \times B_\theta$ term (see Section 1). This is true whatever the bunch charge sign.
– The fields have continuous longitudinal (sin, cos) and radial ($R(r)$) variations. This means that different longitudinal and radial parts of the bunches are accelerated and focused differently by the wakefields. This leads to a broad final energy spread and to emittance growth, respectively. It also means that upon propagation the bunch and the wakefields will evolve self-consistently. This is, in general, not desirable for an accelerator.
– Since the wakefields are driven by the (drive) bunch, they are also tied to the bunch, i.e., they travel with the same velocity as the drive bunch (in the absence of evolution of the bunch or change in plasma density). This (phase) velocity of the wakefields is therefore also close to the speed of light (for $\gamma \gg 1$ and without significant evolution of the bunch). This is possible because in a cold plasma the Langmuir electrostatic wave has zero group velocity [15].

5.2 Bunch size

The cold plasma responds collectively to perturbations, with the fastest time scale given by $\sim 1/\omega_{pe}$ and the smallest spatial scale given by $\sim c/\omega_{pe}$. The bunch transverse and longitudinal sizes can be adjusted to maximize the effect of the bunch on the plasma electrons and thus also on the wakefield's amplitude.

5.2.1 Longitudinal size

For PBPAs, the plasma wave of interest is the Langmuir, electrostatic plasma wave with natural frequency ω_{pe}. It is natural to think that a perturbation (i.e., particle bunches with $n_b \ll n_{e0}$) in time and space shorter rather than longer than the previously described scales is most effective at driving the wave. We note here that a periodic perturbation with period at these time or spatial scales is also very effective in driving the wake. However, most of today's experiments use a single, short particle bunch (or laser pulse). Interest in using periodic excitation was in vogue when sufficiently intense short laser pulses did not exist [18] and is being re-examined, in particular, to take advantage the large amounts of energy stored in long proton bunches (PWFA) [19] and to ease the power requirements on laser systems (LWFA) [20].

Solutions to Eq. (12) are well known (see, for example, Refs. [21, 22]). They indicate that for a finite time or impulse excitation and for weak damping cases, it is the amplitude of the Fourier component of the excitation at the system resonant frequency (here, ω_{pe}) and the damping factor that determine the oscillation amplitude. Most particle bunches have a density or current distribution that is close to Gaussian: $n_b(t) \sim n_{b0} \exp(-t^2/2\sigma_t^2)$ or $n_b(\xi) \sim n_{b0} \exp(-\xi^2/2\sigma_\xi^2)$, where ξ is the position along the bunch (sometimes z is also used). The Fourier transform of such a bunch profile is simply

$\tilde{n}_{\rm b}(k) = \tilde{n}_{\rm b0} \exp(-k_{\rm pe}^2 \sigma_\xi^2/2)$. Therefore, when $k_{\rm pe}\sigma_\xi \leq \sqrt{2}$, significant excitation occurs. For single-bunch experiments, where particles in the front and core of the bunch drive the wakefields and lose energy, while particles in the back gain energy from the wakefields, the bunch length $k_{\rm pe}\sigma_\xi \cong \sqrt{2}$ is optimum [23].

Using linear theory, one can calculate the wakefields driven by bunches with various parameters for a fixed plasma density. However, for a meaningful comparison, one needs to specify which parameters are held constant. For example, as the r.m.s. bunch length is varied, is the number of particles in the bunch or the bunch current kept constant? A way to get around this difficulty is to use the transformer ratio as a figure of merit instead, since wakefield amplitudes can always be obtained by increasing the charge (in linear theory). For example, one can use Eq. (13) to calculate the wakefield's amplitude and then calculate $R = E_+/E_-$. Figure 3 shows the transformer ratio obtained for various bunch lengths W of a square bunch or the r.m.s. length σ_z of a Gaussian bunch, with all other parameters kept constant. For the case of the square bunch, the maximum $R = 2$ is obtained for $W = \lambda_{\rm pe}/2$. In this case, all bunch particles reside in the decelerating phase of the wakefields. For the Gaussian bunch case, the maximum $R = 2$ is reached for $\sigma_z/\lambda_{\rm pe} \cong 0.2$ or $k_{\rm pe}\sigma_z \cong 1.25$. This is close to the predicted $k_{\rm pe}\sigma_z \cong \sqrt{2}$.

This scaling of the wakefield with bunch length and its corresponding increase in amplitude [23] with plasma density has been observed in experiments with electron bunches. A gradient of the order of ~ 100 MeV m^{-1} with bunches with $\sigma_z \sim 600$ μm in a plasma with $n_{\rm e} \sim 2 \times 10^{14}$ cm^{-3} was measured [24]. A gradient in excess of ~ 50 GeV m^{-1} with bunches with $\sigma_z \sim 20$ μm in a plasma with $n_{\rm e} \sim 2.3 \times 10^{17}$ cm^{-3} has also been measured [25, 26]. The scaling has also been confirmed by numerical simulations [27].
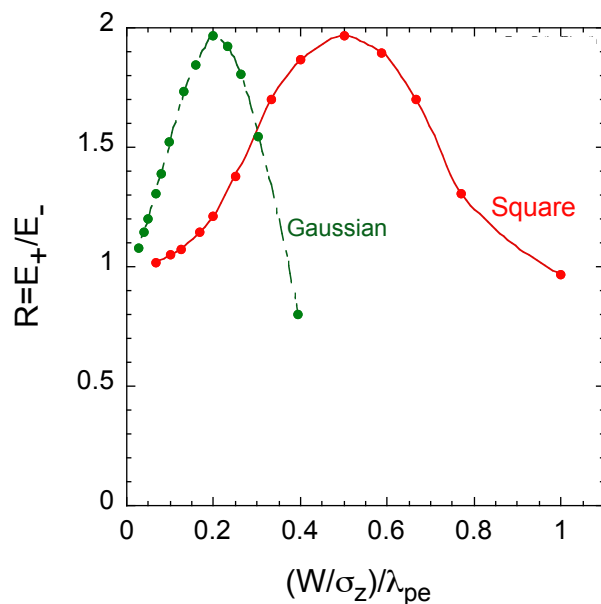
Fig. 3: Transformer ratio $R = E_+/E_-$ as a function of the width W of a square longitudinal profile electron bunch or Gaussian r.m.s. width σ_z, both normalized to the plasma wavelength $\lambda_{\rm pe}$.

5.2.2 Multiple drive bunches

We have seen that Eq. (12) has a natural periodic solution with period $2\pi/\omega_{\rm pe}$. We also saw that the plasma, as an oscillator with natural frequency $\omega_{\rm pe}/2\pi$ selects the frequency content of the drive bunch or train at that frequency. It is therefore also natural to use a (pre-formed) train of bunches with a separation equal to the plasma period to resonantly drive the wakefield. In linear theory, the wakefields

driven by a train are simply the algebraic sum of the wakefields driven by each bunch (with the proper phase). Driving wakefields with multiple bunches has been demonstrated, for example, in Ref. [7]. Figure 4 shows the case of two drive bunches and a witness bunch. The plasma density is chosen so that the plasma wavelength is equal to the bunch spacing ($\lambda_{\mathrm{pe}}(n_{e0}) = \Delta z$), so that both bunches are in the decelerating phase of the total wakefields. The witness bunch follows at a distance $1.5\lambda_{\mathrm{pe}}$ from the second drive bunch, is therefore in the energy gain phase of the wakefields and gains energy from the plasma.

There is little advantage in using a train, unless it is easier to generate a train than a single bunch, or the train is used either to increase the energy extraction efficiency [28] or to increase the transformer ratio [12]. In the former case, the drive and witness bunch are interleaved with a new drive bunch, to replenish the wakefields of the energy extracted by the previous witness bunch. The multi-bunch scheme, relevant for a high-efficiency collider, is not discussed here. In the latter case, the train is a particular form of shaped bunch.

A transformer ratio greater than two with two bunches has been demonstrated in a dielectric based accelerating structure [29], showing the applicability of the wakefield principles to various acceleration schemes.

The case of driving wakefields with a self-generated train of bunches using the self-modulation instability is described in Section 7.
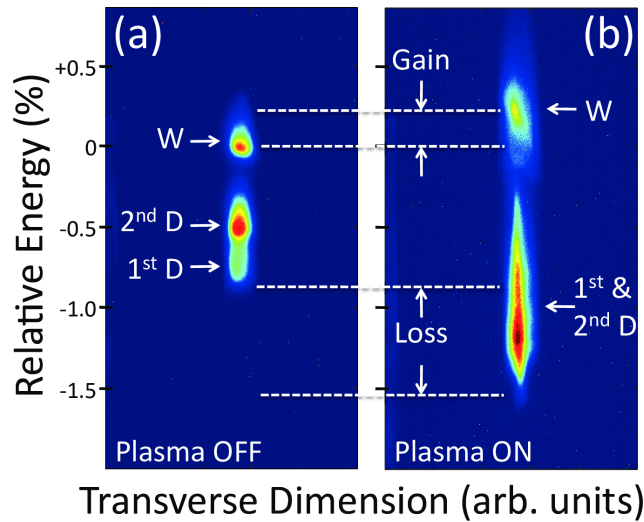
Fig. 4: Energy spectra obtained without (a) and with (b) plasma for the case of two drive bunches (labelled 1st and 2nd D) spaced longitudinally by $\Delta z = \lambda_{\mathrm{pe}}$ (at this plasma density) and a witness bunch (labelled W) following the second drive bunch at $1.5\lambda_{\mathrm{pe}} = 1.5\Delta z$ (from Ref. [7], reproduced with permission from IEEE). In panel (b), the two drive bunches lose energy to the plasma and the witness bunch gains energy from it. The two drive bunches appear merged on panel (a), owing to coherent synchrotron radiation (CSR) effects [30] but are separated in time [31, 32].

5.2.3 Transverse size

The bunch transverse size must be kept smaller than the (cold) plasma skin depth c/ω_{pe} or, equivalently, such that $k_{\mathrm{pe}}\sigma_{r0} < 1$, where $k_{\mathrm{pe}} = \omega_{\mathrm{pe}}/c$. In the opposite case, the bunch is subject to the current filamentation instability (CFI) [33]. The occurrence of CFI results in the breaking up of the bunch current density into filaments of larger current density at the c/ω_{pe} scale in the transverse direction, as observed experimentally [34].

The ability to focus the beam to a small size often limits the maximum density, and thus the maximum accelerating gradient, at which the PBPA can be operated without risk of filamentation. This density is given by

$$n_e \leq \frac{1}{4\pi} \frac{1}{r_0 \sigma_{r0}^2} .$$ (16)

Here, $r_0 = \frac{1}{4\pi\epsilon_0} \frac{e^2}{m_0 c^2} \cong 2.82^{-15}$ m is the classical radius of the electron.

5.3 Maximum accelerating field

The linear PWFA theory is valid for small perturbations of the equilibrium quantities (e.g., $n_1 \ll n_{e0}$). In this case the perturbations are sinusoidal. Although not strictly correct, one can estimate the maximum longitudinal electric field that can be expected by assuming a density perturbation with amplitude equal to the initial density: $n_1 = n_{e0}$. Using the Fourier transform of Poisson's equation with $|\nabla| \to k_{pe}$,

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\epsilon_0} \to k_{pe}E = \frac{\omega_{pe}}{c} E = \frac{-e n_{e0}}{\epsilon_0} = \frac{e^2 n_{e0}}{m_e \epsilon_0} \frac{m_e}{(-e)} \to \frac{\omega_{pe}}{c} |E| = \frac{m_e \omega_{pe}^2}{e} ,$$ (17)

and the corresponding field, known as the cold plasma wave-breaking field, E_{WB} is given by

$$E_{WB} = \frac{m_e c \omega_{pe}}{e} .$$ (18)

(The ions are assumed not to move; their density remains uniform and thus they do not contribute to components at wave number k_{pe}.) It is clear that non-linearities will appear long before the plasma density perturbation reaches n_{e0} and the field reaches E_{WB}. The wakefields become non-sinusoidal (including higher harmonics) and particle trajectories cross [35], invalidating the assumption of fluid behaviour for the plasma. However, this value is a good estimate of the field amplitudes that can be reached in a PBPA, as shown for example by simulations for the PWFA [27]. Results in Ref. [27] show that the wakefield amplitude and scaling generally follow those predicted by linear theory even into the non-linear regime. The expression for E_{WB} is often quoted in an engineering form,

$$E_{WB} \propto 0.96 \text{ GV m}^{-1} \sqrt{n_e[10^{14} \text{ cm}^{-3}]},$$ (19)

and is often used to show the potential of PBPAs for high-gradient acceleration. Plasma electron densities of 10^{14} to 10^{19} cm^{-3} are routinely used in PBPAs with corresponding accelerating fields in the range of a few GV m^{-1} to hundreds of GV m^{-1}.

Crossing of the plasma electron trajectories associated with reaching E_{WB} [35] leads to plasma electron self-injection in the LWFA scheme. We note here that PWFAs usually operate at wakefield phase velocities too large and longitudinal field amplitudes too small to trap plasma electrons. However, ionization injection, now often used in the LWFA, was first observed in the PWFA [36].

5.4 Beam loading

Beam loading is used in many accelerators, in particular to decrease the energy spread resulting from the finite length of bunch when compared with the accelerator wavelength and to maximize the energy extraction or transfer efficiency [37]. It will naturally play an important role in PBPAs because they operate at high frequencies ($\omega_{pe}/2\pi > 100$ GHz) and thus with very short wave periods (<1 mm) and accelerator cavity size (<1 mm^3).

Beam loading can be seen as the simple addition of the bunch wakefields to the fields of the accelerating structure (RF structure or plasma wakefields of the drive bunch). The bunch cannot be placed at the peak of the field, where the accelerating field is constant for a narrow phase range, since (in

the linear regime and in a PBPA) the transverse field becomes defocusing at that point. This is not the case in most RF structures that have no significant transverse field component (TM modes). It has to be placed ahead of the peak for effective beam loading where the field derivative is opposite to its own.

One can illustrate beam loading by using the PWFA linear equations given earlier [17].
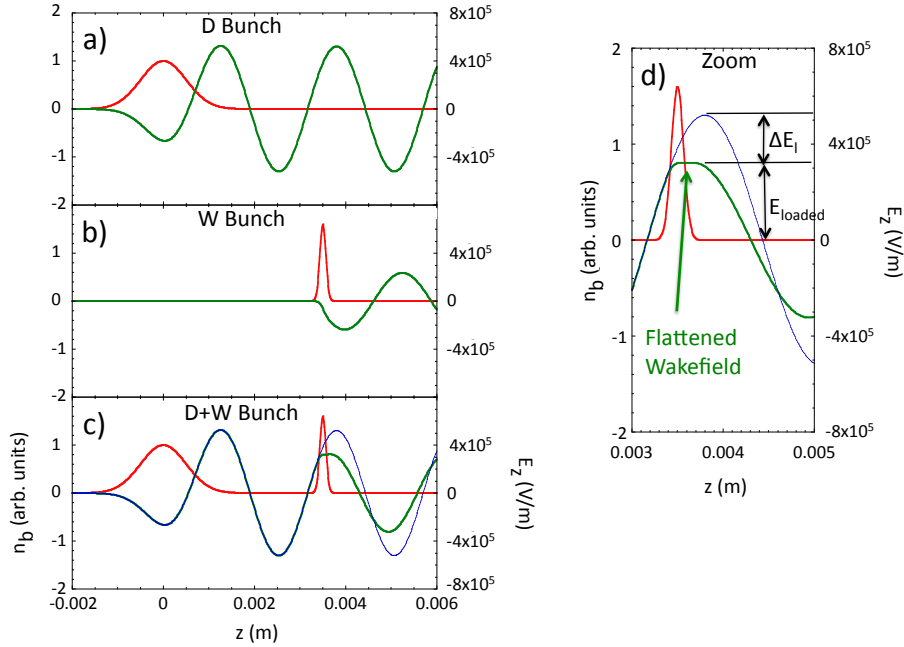
Fig. 5: Linear beam loading example: (a) drive bunch density profile (red line) and longitudinal wakefield E_z (green line), (b) same for the witness bunch, (c) same for the drive and witness bunches together. The field of the drive bunch only is shown as the blue line in panel (c). A zoom around the witness bunch is shown in panel (d). The bunches move to the left.

Figure 5(a) shows a Gaussian bunch and the longitudinal wakefields E_z it is driving in the plasma. Note that this bunch drives wakefields with a transformer ratio very close to two. Figure 5(b) shows the wakefields driven by a shorter, following, witness bunch. This bunch is made shorter than the drive bunch, in order for it to sample a narrow phase range ($\ll \pi/2$), to minimize the energy spread expected for the field variation along the bunch. The total field is shown in Fig. 5(c). The parameters of the witness bunch (charge, length, relative position, etc.) were chosen to illustrate beam loading. The witness bunch is located in a region of increasing field so that the addition of its own field flattens the total field (Fig. 5(d)). This more constant or flat region of accelerating field within the witness bunch reduces its final energy energy spread. However, this more uniform field (E_{loaded}) is also lower than the peak field, more than 25% in this case (ΔE_l), decreasing the effective transformer ratio by the same fraction. The fraction of energy, proportional the field amplitudes squared, extracted from the wakefields by the witness is on the order of 50%. The energy spread ΔE reduces from about 25% to less than 8% in this non-optimized case.

Note that beam loading in the non-linear regime has also been studied [38, 39]. Narrow energy spread and high-energy transfer efficiency can be reached, at the expense of tailoring the witness bunch current profile. Beam-loading effects have been observed in recent PWFA experiments, leading to significant wakefield to witness bunch energy transfer efficiency and to narrow final energy spread [8].

6 Non-linear PWFA regime

In the linear PWFA regime, the wakefields vary along and across the bunch and field structure (see Eqs. (13) and (14)). This leads to significant final energy spread and to emittance growth. In the non-linear PWFA regime, the accelerating field in the pure ion column is independent of radius. It varies longitudinally, but this can be mitigated by using beam loading. The focusing field varies linearly with radius, which preserves the emittance of a bunch with Gaussian position and velocity distributions. It is also independent of the longitudinal position. Therefore, the non-linear PWFA regime has significant advantages over the linear regime, at least for electron beam (or negatively charged particle bunches).

While there is no strict theory for the PWFA in the non-linear regime, an interesting theory based on the parameters of the electron bubble and of the plasma electron sheet that sustains it has been developed [40, 41]. It was used, for example, for the beam loading study in the non-linear regime mentioned previously [38, 39]. However, it is beyond the scope of this manuscript.

6.1 Electron bunch propagation

In a vacuum, a (Gaussian) beam expands because of its emittance, with a characteristic increase in transverse size of $\sqrt{2}$ per β_0 propagation distance. This also means that the bunch density ($\propto 1/\sigma_r^2$) decreases by a factor of two, as does the wakefield amplitude that it can drive (see Eq. (13) for the linear regime case).

The plasma provides (transverse) focusing through the transverse wakefield $W_\perp$ (see Eq. (14) for the linear case), so that the beam can remain transversely small and keep driving large amplitude wakefields.

In the linear regime, the beam continuously evolves in transverse size and distribution, since $W_\perp$ is a function of (r, ξ) and its emittance grows. It can be shown that the r.m.s. transverse position and velocity distributions are preserved by a transverse force that varies linearly with radius: $F_\perp \sim r$. This is the case in the non-linear regime, in which $n_b \geq n_{e0}$ and the plasma structure sustaining the wakefields consists of a 'bubble' that is void of plasma electrons. It thus consists of a pure ion column with uniform density (if there is no ion motion). In the ion column, there is no magnetic field since there is no plasma return current (it flows in a sheet around the bubble). Therefore, the transverse force experienced by the electrons in the column is simply $\overrightarrow{F}_\perp = -e(\overrightarrow{E}_r + \overrightarrow{v}_b \times \overrightarrow{B}_\theta) = -e\overrightarrow{E}_r$. The electric field of the ion column can be estimated using Poisson's equation, assuming that it is infinitely long and cylindrically symmetrical, with $n_{i0} = n_{e0}$:

$$\overrightarrow{\nabla} \cdot \overrightarrow{E} = \frac{\rho_{\text{ion}}}{\epsilon_0} \rightarrow E_r = \frac{1}{2} \frac{n_{e0} e}{m_e} r \; . \tag{20}$$

Thus, $F_\perp = e E_r \sim r$.

Bunch particles have an equation of motion in the pure ion column given by

$$\frac{\mathrm{d}^2 r(z)}{\mathrm{d}z^2} + K^2 r(z) = 0 \; . \tag{21}$$

For the case of the pure ion column, $K^2 = \frac{1}{\gamma m_e c^2} \frac{F_\perp}{r} = \frac{1}{2\gamma m_e c^2} \frac{n_e e^2}{\epsilon_0}$. This is a harmonic oscillator equation with general solution $r(z) = r_0 \exp(\pm i k_\beta z)$, where $k_\beta = \frac{1}{\sqrt{2\gamma}} \left(\frac{n_e e^2}{\epsilon_0 m_e c^2} \right)^{1/2} = \frac{k_{\text{pe}}}{\sqrt{2\gamma}}$ is the betatron wave number.

The evolution of the bunch transverse r.m.s. size is then described by the envelope equation (see Ferrario's lecture), which, neglecting acceleration, reads

$$\frac{\mathrm{d}^2 \sigma_r(z)}{\mathrm{d}z^2} + K^2 \sigma_r(z) = \frac{\epsilon_g^2}{\sigma_r^3(z)} \rightarrow \sigma_r'' + K^2 \sigma_r = \frac{\epsilon_g^2}{\sigma_r^3} \; . \tag{22}$$

This equation is very similar to that for the individual particles (Eq. (21)), with r replaced by σ_r and the addition of the emittance term. It shows that at locations where σ_r is large, the $K^2 \sigma_r$ term dominates

the ϵ_g/σ_r^3 term, $\sigma_r'' < 0$ and the beam is focused by the external force (assuming $K^2 > 0$). As the beam is focused, its size becomes smaller, the ϵ_g/σ_r^3 term eventually dominates the $K^2\sigma_r$ term, and the bunch diverges because of its emittance. Therefore, in general, the bunch transverse size oscillates between a minimum and maximum (i.e., remains positive). The particles also oscillate but, of course, cross the axis. There is a situation in which the bunch size does not oscillate. Envelope oscillation amplitudes can be calculated if one considers the simple and practical case of a bunch focused at the entrance of the plasma (i.e., $\sigma_r'(z = 0) = 0$) with size σ_{r0}. In this case, Eq. (22) can be integrated to obtain the smallest and largest bunch sizes as a function of the matched bunch size σ_{rm}. Rewriting Eq. (22) as

$$\sigma_r'' = \left(\frac{\epsilon_g^2}{\sigma_r^4} - K^2\right)\sigma_r , \tag{23}$$

and first setting the term in parentheses to zero and with the initial conditions specified leads to

$$\sigma_{rm}^4 = \frac{2\epsilon_0\gamma m_e c^2\epsilon_g^2}{n_{e0}e^2} , \tag{24}$$

or the more general condition for the bunch and plasma parameters:

$$\frac{\sigma_{rm}^4 n_e}{\gamma\epsilon_g^2} = \frac{2\epsilon_0 m_e c^2}{e^2} . \tag{25}$$

When the conditions: $\sigma_r = \sigma_{r0}$, $d^2\sigma_r/dz^2 = d\sigma_r/dz = 0$ are satisfied at the plasma entrance ($z = 0$), the beam radius remains constant along the plasma and the beam is said to be matched to (the focusing force of) the plasma.

Note that $\sigma_{rm} \sim \gamma^{1/4}$ is weakly dependent on particle energy, which may allow for adiabatic matching of the bunch size to the plasma focusing upon acceleration. Note also that while the bunch envelope size does not oscillate in the matched case, particles do.

In general, the (unmatched) beam envelope size oscillates in the ion column; it reaches the two sizes $\sigma_1 = \sigma_0$ and $\sigma_2 = \frac{\sigma_{rm}}{\sigma_0}\sigma_{rm}$. These sizes can be calculated by multiplying Eq. (23) by σ_r', integrating it and setting $\sigma_r' = 0$. Therefore, when $\sigma_0 > \sigma_{rm}$, $\sigma_2 < \sigma_{rm}$, and when $\sigma_0 < \sigma_{rm}$, $\sigma_2 > \sigma_{rm}$. This means that the maximum beam size is *always* larger than or equal to the matched size.

The electrons oscillate with the betatron wavelength $\lambda_\beta = \sqrt{2\gamma}\lambda_{pe}$, and the bunch envelope with a periodicity half of λ_β. The beam envelope oscillations were clearly observed with an electron beam by changing the plasma density [42]. Matching of the bunch to the plasma was also observed [24].

The bunch electrons oscillating in the ion column emit synchrotron radiation called betatron radiation (see lecture by K. Ta Phuoc). This radiation was observed for the first time in a PBPA as X-radiation [43]. It was later observed as γ-radiation at larger plasma densities [44]. This is now a main feature of the LWFA, which is a very interesting source of (betatron) radiation. These considerations on the bunch transverse size oscillation indicate that minimum betatron power is emitted when the bunch is matched to the plasma, since the betatron radiation power scales as $P_{\text{betatron}} \sim r_0^2 \sim \sigma_{r,\max}^2 \geq \sigma_{r,m}^2$. This can be used experimentally to find the matching condition by minimizing the amount of betatron radiation emitted by the bunch.

Matching the beam to the pure ion column focusing also has another advantage: it minimizes the sensitivity of the beam angle at the plasma exit as a function of the beam and plasma parameters. When matched, $\sigma_r'' = 0$ or $(\sigma_r')' = 0$, i.e., the exit angle is minimum. Figure 6 shows the beam angle at the plasma exit as a function of the relative variation of the beam size at the plasma entrance and of the plasma density. This figure was obtained by integrating numerically the envelope Eq. (22) in a plasma with constant density (i.e., $K^2 = $ constant). When the bunch is not matched, the angle of the beam at the plasma exit is $\theta \cong r_\beta/(\lambda_\beta/8)$, where r_β is the radial oscillation amplitude of the beam envelope around the matched radius size.
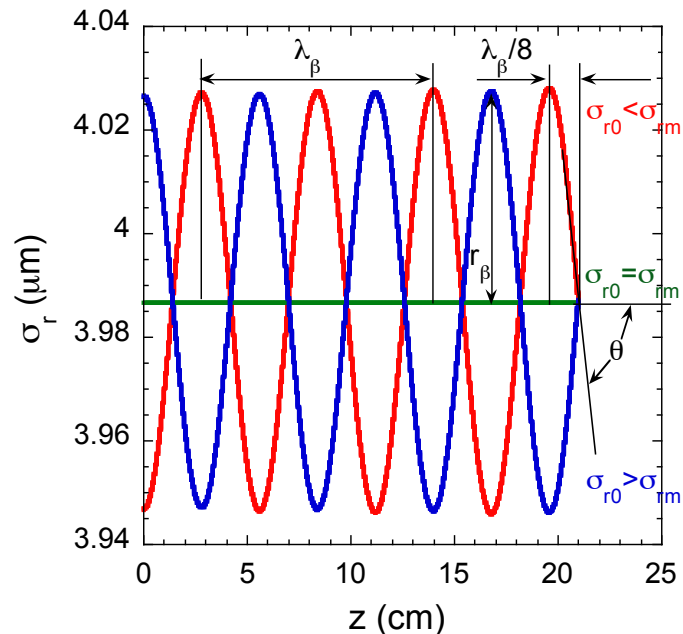
Fig. 6: Bunch envelope size along a plasma with uniform density for the case of a matched bunch (green line) and a slightly mismatched bunch with initial size too large (blue line) or too small (red line). The betatron wavelength and exit angle are shown.

Matching of the beam to the focusing field of non-linear wakefields has been demonstrated experimentally. Figure 7, from Ref. [24], shows the SLAC electron bunch transverse size measured downstream from the plasma, as a function of plasma density. Near the matched density, the variation in transverse size decreases, compared with the unmatched cases (lower density in this case).

Matching was also demonstrated with lower bunch energy [45].

The very strong plasma focusing force implies that the matched beam size is quite small. For example, for a typical SLAC electron beam with $\gamma \cong 56\,000$ ($E_0 \cong 28$ GeV) and a normalized emittance $\epsilon_N = 5 \times 10^{-5}$ mm mrad, Eq. (24) indicates that $\sigma_{rm} \cong 4$ µm. This is a quite difficult size to produce at the plasma entrance. However, a continuous ramp in the plasma density, typical of real plasmas, rather than a step function, can be used to match the beam adiabatically to the plasma focusing force [46]. Adiabatic matching occurs when the variation of the bunch envelope size is small in one betatron period. Numerical integration of Eq. (23) shows that this condition is easily satisfied. Plasma sources used for PWFA experiments, such as metal vapour sources [47, 48], naturally provide a neutral density ramp, which, once ionized by a laser pulse or by the bunch itself [49], becomes a continuous plasma ramp. Figure 8 shows an example of a calculation in which the entrance plasma ramp is used to match the beam to the plasma and the exit ramp, to increase the bunch size and decrease its divergence in vacuum. Increasing the bunch size at the plasma exit decreases the emittance term, owing to the bunch energy spread (see M. Ferarrio's lecture).

Matching into and out of the plasma is a very interesting and important research topic. Indeed, the larger focusing strength of the plasma, compared with that of conventional magnetic optics, leads to much shorter beta-functions in the plasma than outside of it, where the beam is used or conditioned to be injected into the next PBPA. Indeed with the definition of the beam beta, Eq. (23) can be rewritten as

$$\beta_{\text{matched}} = \frac{\sigma_{rm}^2}{\epsilon_g} = \frac{1}{K} \ . \tag{26}$$
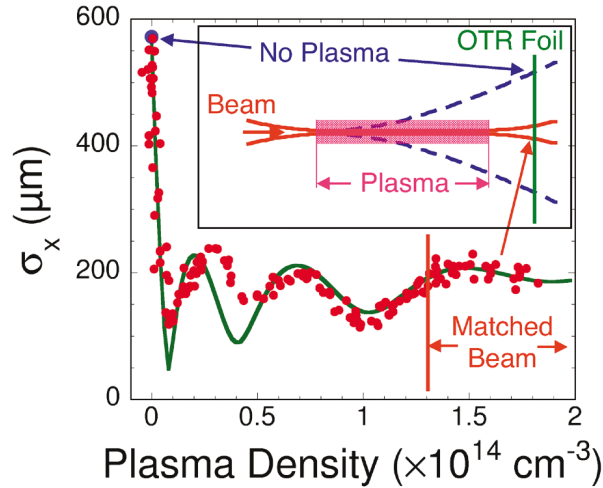
Fig. 7: Transverse bunch size measured as a function of plasma density a distance ≈ 1 m downstream from the plasma exit. As the plasma density and the ion focusing column strength also increases ($\sim n_{e0}$), the beam approaches the matching condition observed in the range $(1.25–2.5)\times10^{14}$ cm^{-3}. See Ref. [24] for experimental parameters (Reproduced with permission from APS).

It is important to note that the pure ion column, so favourable for an electron bunch or more generally for a negatively charged bunch, does not exist for a positively charged bunch (e^+, p^+, etc.). The plasma is asymmetrical for the charge signs of the two bunches since, in any case, only the light plasma electrons move (at the $1/\omega_{pe}$ time scale). This difference has been observed with a positron bunch in conditions similar to those for an electron bunch [10, 50] (see Section 7.3). Focusing of a positron bunch by a short plasma has also been demonstrated [51].

7 Plasma and magnetic focusing

It is interesting to compare the focusing strength of the plasma ion column with that of a conventional quadrupole magnet. In a quadrupole magnet, focusing is achieved by imposing a magnetic field perpendicular to the particles' trajectories. The radius of curvature of the relativistic particle's trajectory in the magnetic field $B_\perp$ is given by the Larmor radius; $r_L = \beta\gamma m_e c/qB_\perp \cong \gamma m_e c/qB_\perp$ ($\beta = (1 - 1/\gamma^2)^{1/2} \cong 1$). In a magnet (and magnetic field) of length L, this results in a deflection angle $\theta \ll 1$, such that $r_L\theta \cong L$. Therefore, $\theta \cong L/r_L = qB_\perp L/\gamma m_e c$. For the magnet to act as a lens, free of geometric aberration, the deflection angle and thus the magnetic strength must increase linearly with radius and reverse at the axis. This is the characteristic of the field with a quadrupole symmetry. A quadrupole magnet with a larger field gradient $B_\perp/r$ corresponds to a stronger, shorter focal length, focusing element. Quadrupoles usually use electromagnets to generate the field, but the largest field gradients are achieved in permanent magnet quadrupoles (PMQs). For example, the field gradient reached $\sim$290 T m^{-1} over 10 cm, with an aperture of 7 mm in the PMQ of Ref. [52].

Using Eq. (20) for the ion column field, and dividing it by r to obtain the gradient and by c to produce the proper units of magnetic field, we obtain $E_r/rc \cong 3\,kT/m$ for $n_e = 10^{14}$ cm^{-3}, as in Refs. [42, 43]. For the higher plasma densities of Refs. [25, 26, 44], $n_e = 10^{17}$ cm^{-3}, one obtains $E_r/rc \cong 3\,MT/m$. These values clearly show the potential for very strong focusing by plasmas in a scheme known as the plasma lens [16, 53]. We note, however, that the transverse extent of the focusing region is limited to $\approx c/\omega_{pe} < 1$ mm.

Since plasma lenses have such strong focusing gradients they can be shorter (and smaller) than magnetic ones. They could be of potential interest for replacing the km-long final focusing system of
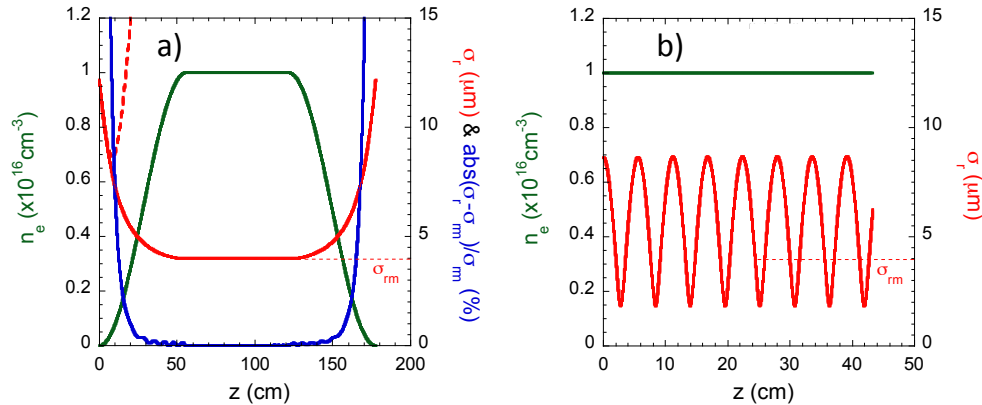
Fig. 8: (a) Beam transverse size in vacuum (dashed red line) and plasma (red line) with density ramps at the entrance and exit (green line). In this case, the beam vacuum focus is $\sigma_{r0} = 8.6$ μm at $z = 8.18$ cm, so that with the initial ramp, the beam reaches the matched radius $\sigma_{rm} = 3.8$ μm in the constant density region ($55.6 < z < 122$ cm). The total plasma length is 1.77 m with a constant density of 10^{16} cm^{-3}, the particle's γ is 56 000, the beam normalized emittance is 5×10^{-5} m rad^{-1} and the betatron wavelength in the constant density region is $\lambda_\beta = 2\pi/K = 11.2$ cm. The blue line represents the absolute value of the relative deviation between the beam size (σ_r) and the matched beam size (see Eq. (24)): $\left| \frac{\sigma_r - \sigma_{rm}}{\sigma_{rm}} \right|$. (b) Transverse beam size over the first 45 cm of plasma with the same density (and beam parameters) as panel (a) in the case where the beam is focused to the same $\sigma_{r0} = 8.6$ μm, but at the entrance of the plasma with a step-function density. The beam is mismatched and the envelope oscillations are indeed between the incoming size and that predicted by $\sigma_2 \leq \sigma_{rm}$ in the previous paragraph, $\cong 1.8$ μm, in this case.

a collider [14]. This is another area where plasmas could contribute to reducing the size and cost of a future collider. We note however that plasma lenses have their own challenges, including the possible need of a *drive bunch* to create the lens out of a neutral plasma.

Quadrupoles are, by definition, focusing in one plane and defocusing in the other. They must therefore be used in pairs (a doublet) or in threes (a triplet) to build an element that focuses in both planes. They are also usually short compared with their focal length and are thin lenses (though they may be long magnets). A lens is said to be thin when its thickness or length is much shorter than its focal length.

By contrast, the plasma ion column (or the plasma in general) is focusing (or defocusing) in both planes at the same time. Because of their large focusing strength, plasma lenses can be thin or thick lenses. Mismatched plasma accelerating sections are very thick focusing elements. The beam has multiple foci along the plasma, since its length can be many betatron wavelengths long (first focus near $\lambda_\beta/4$). For example, in the experiments described in Refs. [24, 42], the beam has up to three foci along the plasma with $\lambda_\beta = 81$ cm, maximum density 1.9×10^{14} cm^{-3}, and length $\cong 1.5$ m $\cong 2\lambda_\beta$. In the case of Ref. [13], the plasma length is about $46\lambda_\beta$.

7.1 Electron acceleration

The accelerating field amplitude is usually not directly measured. The accelerating gradient is obtained by dividing the change in energy (loss or gain) by the (measured or assumed) plasma length. This is a quantity that is integrated over the plasma length, including the density ramp, in experiments. The measured energy gain and loss depend on the actual accelerating field along the plasma and on the fact that there are (or are not) particles sampling the field. The dependency of the longitudinal wakefield along the bunch was measured by time resolving, at the *ps* scale, the energy change of the electrons along a

single bunch [24]. The general longitudinal shape of the field is as expected from simple consideration: decelerating, switching to accelerating (Fig. 9). Acceleration by trailing particles has been observed at a plasma density roughly satisfying $k_{pe}\sigma_z \cong \sqrt{2}$ ($\sigma_z \cong 700$ μm in this case).

The transverse wakefield structure was also inferred [54] using the transverse and longitudinal wakefields $90°$ out of the phase relation predicted by linear theory. Once again, the general agreement between experiment and theory is very good.

The maximum accelerating field (averaged over the plasma length) has been measured by varying the bunch length over a relatively wide range of plasma densities and lengths [55]. The results are shown in Fig. 10 from Ref. [55]. In this case, $\sigma_z \cong 20$ μm and the plasma density is much larger ($\sim$1000 times larger than in [24]) to approach $k_{pe}\sigma_z \cong \sqrt{2}$. The left-hand panel of Fig. 10 shows that, at the lower density, the energy gains (and thus accelerating fields) are low because the bunch is too short $k_{pe}\sigma_z > \sqrt{2}$ for this density. The bunch is also too short to sample the peak field. However, the energy gain increases with decreasing bunch length and increasing plasma length, as expected. At what turned out to be the optimum density (middle panel), the energy gain exhibits a (shallow) maximum for parameters close to $k_{pe}\sigma_z \cong \sqrt{2}$ for the three plasma lengths. Wakefields are again probably larger with shorter bunch lengths, but there may be no particles to sample them. At the highest density (right-hand panel), the energy gain is lower than at the optimum density and increases with shorter bunches because more of the bunch charge participates in driving the wakefields as $k_{pe}\sigma_z$ approaches $\sqrt{2}$ with decreasing σ_z. Since the bunch is long when compared with σ_z, there are probably particles to sample the maximum field amplitude.

In Ref. [24], the measured (average) accelerating gradient is $\sim$200 MeV m^{-1}, lower than expected from $E_{WB} = 1.3$ GV m^{-1} for $n_{e0} = 1.9 \times 10^{14}$ cm^{-3}. In Ref. [26], it is $\sim$52 GeV m^{-1}, comparable to the expected value of $E_{WB} = 46$ GV m^{-1} for $n_{e0} = 2.3 \times 10^{17}$ cm^{-3}. Both experiments were in the $n_b > n_{e0}$ non-linear regime and used a single electron bunch to drive and sample wakefields. These results show the main features expected from the wakefield excitation. The results are in general agreement with the predictions of linear theory.
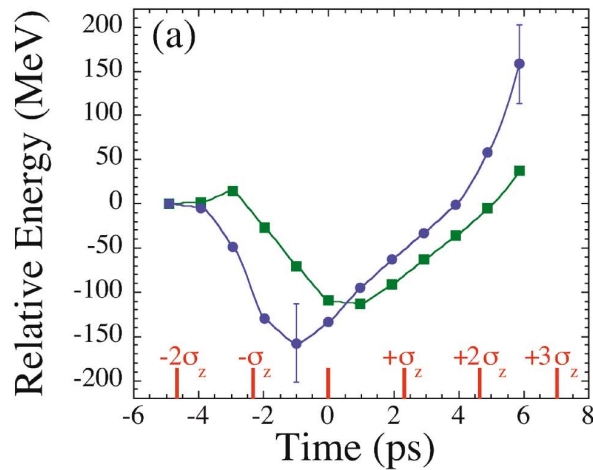
Fig. 9: Energy change, measured along the electron bunch with $\sim$1 ps resolution. The energy change reflects the shape of the wakefields along the bunch (the wakefield amplitude is the change in energy divided by the plasma length). The wakefields have the expected shape, i.e., essentially zero ahead of the bunch ($\sim$−6 ps), with maximum energy loss in the core of the bunch ($\sim$0 ps) and, eventually, energy gain in the back of the single bunch ($>$+4 ps). The green line was obtained at a lower plasma density than the blue one. Energy gain in the back of the bunch (right-hand panel) is observed at a density for which $k_{pe}\sigma_z \cong \sqrt{2}$. For parameters, see Ref. [24] (Reproduced with permission from APS.)
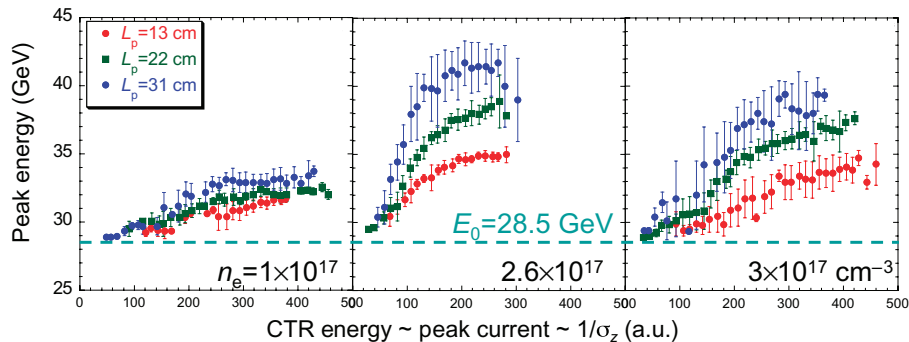
Fig. 10: Maximum energy of the particles (in a single bunch) measured at three plasma densities ($n_e = 1$, 2.6 and 3×10^{17} cm^{-3}) and each at three plasma lengths ($L_p = 13$ (red symbols), 22 (green symbols) and 31 cm (blue symbols)) as the bunch is made shorter, and therefore the coherent transition radiation (CTR) energy emitted by the bunch increases. Higher CTR energy corresponds to higher peak current, shorter bunches. The error bars represent the r.m.s. energy measurement variation in each bin. For parameters, see Ref. [55] (Reproduced with permission from IOP.)

Further experiments using two electron bunches showed a narrow energy spread of the witness bunch as well as good transfer efficiency between the wakefields and the witness bunch, due to a significant beam-loading effect [8].

Measured energy gains are usually in excellent agreement with numerical simulation results with experimental parameters.

7.2 Positron bunch propagation

In the linear PWFA regime, the wakefields are similar for an electron and a positron bunch with a simple phase change (see Figs. 1 and 2). Therefore, the accelerated bunch suffers from similar large final energy spread and emittance growth. In the non-linear regime, the locations of the wakefields that are focusing for a positron bunch are confined to the regions of return of the plasma electrons to the axis (see Figs. 1 and 2). These regions are small and have no strong variations of the fields. The locations that are accelerating are those immediately behind those pinch regions (regions that are decelerating for an electron bunch!). There is, therefore (and a priori), no good regime for accelerating a positron bunch to high energies while maintaining its incoming emittance.

The propagation of positron bunches in plasmas has been studied experimentally. In particular, the focusing or plasma lens effect was observed in a short plasma [51].

Propagation in a long plasmas ($L \gg \lambda_{pe}$) was also studied [50, 56]. The difference for electron bunch propagation and the formation of a beam halo due to the non-linear focusing force was observed directly (Fig. 11). The associated emittance growth was inferred from numerical simulations that described the experiment [50].

7.3 Positron acceleration

The acceleration of positrons in plasmas has been observed [10]. As expected, the measured energy change was lower than with an electron bunch with comparable parameters. The inferred wakefield along the bunch also has a more linear shape, i.e., a greater sinusoidal dependency than that of the electron bunch [24].

Recent results obtained at SLAC suggest that in some cases a situation favourable for large acceleration without significant emittance growth might exist in a uniform plasma.
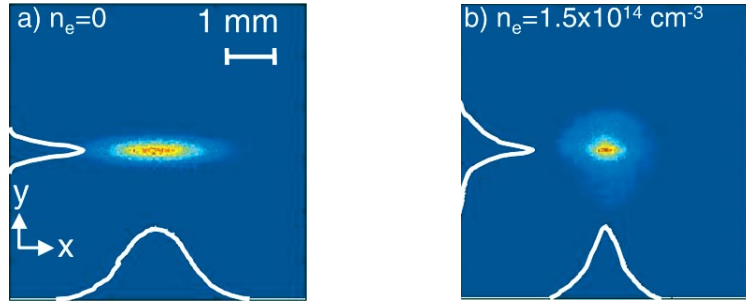
Fig. 11: Images of the positron transverse size obtained (a) without and (b) with plasma, a distance ∼1 m downstream from the plasma exit. The images are obtained using optical transition radiation (OTR). The beam has an elliptical transverse shape without plasma (a) because it is focused to a round spot at the plasma entrance and has different emittances in the x- and y-planes (typical of linear colliders). The image with plasma (b) shows a focused core surrounded by a halo of particles. The halo is the result of the evolving and non-linear character of the plasma focusing force acting on the positron bunch. For parameters, see Ref. [50] (Reproduced with permission from APS.)

7.4 Hollow plasma channel for positron bunches

One possible solution to the emittance preservation during the acceleration of a positron bunch in a plasma is to use a hollow plasma channel. In such a channel, the plasma density is zero from the axis to a radius r_c and a constant density beyond r_c. Simulation results show that r_c of the order of c/ω_{pe} may be optimal in terms of accelerating gradient [57]. A drive bunch (electron or positron) displaces the plasma electrons in much the same way as in a uniform plasma and the plasma electrons converge towards the axis of the beam and channel. They sustain a longitudinal wakefield structure similar to that formed in a uniform plasma that has accelerating and decelerating regions. However, since there are no plasma ions in the channel, there are essentially no transverse forces to defocus a positron witness bunch. In a hollow plasma channel, a positron witness bunch can therefore be accelerated without the emittance growth associated with the unfavourable transverse wakefields. Since there are no focusing fields, the length of the acceleration is limited by the beta-function of the beam, which increases in size transversely and eventually reaches the channel radius. However, as previously noted, the emittance of the beams foreseen for a future accelerator might be low enough to allow for acceleration in metre-long plasma channels.

Hollow plasma channels could be produced, for example, by heating a plasma with a long (∼ ps) laser pulse and letting it expand [58] or by using specially designed phase optics to ionize a gas only in a hollow cylinder region [59].

8 Quasi-linear PWFA regime

An interesting regime, between the linear and non-linear regime can be reached and may prove advantagous for operation with multiple drive bunches. This regime is called the quasi-linear or weakly non-linear regime.

This regime is reached when $n_b \cong n_{e0}$. In this case, blow-out is reached essentially only over the bunch volume, unlike in the non-linear regime, where it is obtained over a much large radius. This regime therefore has the same focusing properties as in the non-linear regime (for an electron bunch). However, the longitudinal wakefields do not attain non-linear values and can therefore be added *linearly* with multiple bunches, possibly allowing for a large transformer ratio or large energy extraction efficiency. The addition of the wakefields and the possibility of reaching a large transformer ratio, comparable with

that predicted by simple linear PWFA theory has been demonstrated in numerical simulations [60–62]. However, the charge and position of each bunch must be studied using simulations, to produce optimum results.

9 Self-modulation instability

Most PWFA experiments are performed with single electron bunches or short preformed trains of electron bunches.

Electron bunches can be compressed to extremely short lengths and focused to small transverse sizes, so that PWFA experiments can be performed at high plasma densities ($k_{\mathrm{pe}}\sigma_z \cong \sqrt{2}$, $k_{\mathrm{pe}}\sigma_r < 1$) and operate at high gradients. However, electron bunches carry relatively small amounts of energy as compared with proton bunches. For example a typical SLAC electron bunch with 25 GeV per particle and 2×10^{10} particles per bunch carries about 80 J. A typical CERN SPS proton bunch with 400 GeV per particle and 3×10^{11} particles per bunch carries about 19 kJ. Once accelerated in the LHC to 7 TeV, it carries 336 kJ.

A future electron/positron linear collider is expected to produce bunches with at least 250 GeV per particle and 2×10^{10} particles per bunch, i.e., carrying 800 J. A PBPA-based collider using an 80 J SLAC-like drive bunch would therefore need to stage at least 10 plasma sections to produce an 800 J bunch. Staging has been envisaged since the beginning of the PWFA. However, staging implies distances between the plasma sections that might be much longer than the plasma sections, leading to gradient dilution: the average accelerating gradient (final energy divided by accelerator length or energy gain per plasma stage divided by distance between stages) might be significantly lower than the peak accelerating gradient (gradient in the plasma sections).

A possible alternative is to use a proton bunch as a drive bunch. This idea was explored numerically [63]. While results showed the possibility of an energy gain as high as ∼600 GeV in a single ∼500 m plasma, the 1 TeV, $\sigma_z = 100$ μm, proton bunch used as a driver does not exist. CERN SPS and LHC bunches are ∼12 cm long. It is interesting to note that while with electrons or more generally negatively charged bunches, a bunch with $\sigma_r \leq \sigma_z \leq c/\omega_{\mathrm{pe}}$ is optimum, a shape with $\sigma_z < \sigma_r$ seems more effective for positively charged bunches.

It was recently proposed to use a transverse instability, the self-modulation instability (SMI) to drive large amplitude wakefields with bunches much longer than the plasma wavelength [19]. The SMI arises from the interplay between the transverse wakefields that can periodically focus (or defocus) the particle bunches and increase (or decrease) the bunch density that drives the wakefields more strongly (or weakly) (see Eq. (14)). The instability modulates the long bunch into a train of bunches with a period approximately equal to the plasma wavelength. The train then resonantly drives wakefields to much larger amplitudes than the long bunch would in a plasma with lower density, such that $k_{\mathrm{pe}}\sigma_z \cong \sqrt{2}$. One can estimate the *gain* in wakefield amplitude when using SMI by taking E_{WB} as the measure of the maximum possible accelerating field and assuming that the SMI produces micro-bunches of length such that $k_{\mathrm{pe}}\sigma_r \cong \sqrt{2}$ separated by λ_{pe}.

It is important to note that the modulation has a longitudinal period as a result of the transverse (and not longitudinal) motion of the bunch particles (since dephasing caused by energy gain or loss is negligible for relativistic particles).

Experiments with electron bunches of one to five λ_{pe} showed that periodic wakefields are indeed driven by bunches with $\lambda_{\mathrm{pe}} > \sigma_z$ [64]. The observation was based on longitudinal wakefields; however, those are always accompanied by transverse wakefields.

Experiments aimed at directly observing the effect of the transverse wakefields on the bunch radius are ongoing at SLAC with electron and positron bunches [65]. Preliminary results show evidence of radial modulation, formation of a beam halo and energy loss to wakefields. Experiments with low-energy electron bunches are planned at DESY Zeuthen [66].

A major experiment, known as AWAKE, is in the design and installation phase at CERN [67]. It will use the 400 GeV SPS proton bunch and a 10 m long plasma with a baseline density $n_{e0} = 7 \times 10^{14}$ cm^{-3} [68].

The first goal of the experiment is to observe and characterize the SMI of the proton bunch. In a second phase, an electron bunch with σ_z of the order of one to two λ_{pe} will be externally injected to sample the wakefields. In a third phase, an electron bunch shorter than λ_{pe}, possibly produced by a LWFA [69], will be injected to narrow the final energy spread through loading of the wakefields and to preserve the accelerated bunch emittance [70]. Details of the AWAKE experiment can be found in the lecture by E. Gschwendtner.

10 Final remarks

The PWFA is one of the advanced accelerator schemes studied as a high-gradient alternative to today's RF technology. The PWFA has its advantages and its limits, and only further research in the field will help determine its relevance to a current or future electron or positron accelerator. The application of the PWFA is to high-energy accelerators, such as those driving X-ray free electron lasers (FELs) or the electron/positron collider that may be chosen to complement the LHC.

It is clear that many subjects have not been addressed in this manuscript: plasma sources, staging of plasma sources, bunch head erosion, scattering on the plasma ions, etc. However, we hope that this manuscript will foster further reading about and understanding of the growing body of theoretical, numerical and experimental PWFA work. Moreover, we hope that curiosity, talent and enthusiasm of many scientists, potentially captured by this CERN accelerator school, will take this concept to its limit.

References

[1] T. Tajima and J. Dawson, *Phys. Rev. Lett.* **43**(4) (1979) 267.
    http://dx.doi.org/10.1103/PhysRevLett.43.267

[2] P. Chen *et al.*, *Phys. Rev. Lett.* **54**(7) (1985) 693. http://dx.doi.org/10.1103/PhysRevLett.54.693

[3] ILC Technical Design Report, vol. 3.ii: Accelerator Baseline Design (2013),
    http://www.linearcollider.org/ILC/Publications/Technical-Design-Report

[4] J.B. Rosenzweig *et al.*, *Phys. Rev. Lett.* **95**(19) (2005) 195002.
    http://dx.doi.org/10.1103/PhysRevLett.95.195002

[5] R. Gholizadeh *et al.*, *Phys. Rev. Lett.* **104**(15) (2010) 155001 (2010).
    http://dx.doi.org/10.1103/PhysRevLett.104.155001

[6] J. Vieira *et al.*, *Phys. Rev. Lett.* **109**(14) (2012) 145005
    http://dx.doi.org/10.1103/PhysRevLett.109.145005

[7] P. Muggli *et al.*, in *Proceedings of the 2011 Particle Accelerator Conference* (IEEE, Piscataway, NJ, 2011), p. 712, http://accelconf.web.cern.ch/AccelConf/PAC2011/papers/tuobn3.pdf

[8] M. Litos *et al.*, *Nature* **515**(7525) (2014) 92. http://dx.doi.org/10.1038/nature13882

[9] J.B. Rosenzweig *et al.*, *Phys. Rev. Lett.* **61** (1988) 98. http://dx.doi.org/10.1103/PhysRevLett.61.98

[10] B.E. Blue *et al.*, *Phys. Rev. Lett.* **90** (2003) 214801.
    http://dx.doi.org/10.1103/PhysRevLett.90.214801

[11] R.D. Ruth *et al.*, SLAC-PUB-3374 (1984).

[12] E. Laziev *et al.*, EPAC IEEE (1988), p. 523.

[13] I. Blumenfeld *et al.*, *Phys. Rev. ST Accel. Beams* **13**(11) (2010) 111301.
    http://dx.doi.org/10.1103/PhysRevSTAB.13.111301

[14] P. Chen *et al.*, *Part. Accel.* **20** (1987) 171.

[15] F.F. Chen, *Introduction to Plasma Physics* (Springer, New York, 1974).
http://dx.doi.org/10.1007/978-1-4757-0459-4

[16] P. Chen *et al.*, *IEEE Trans. Plasma Sci.* **15**(2) (1987) 218.
http://dx.doi.org/10.1109/TPS.1987.4316688

[17] T. Katsouleas *et al.*, *Part. Accel.* **22**, (1987) 81.

[18] A. Modena *et al.*, *Nature* **377** (2002) 606. http://dx.doi.org/10.1038/377606a0

[19] N. Kumar *et al.*, *Phys. Rev. Lett.* **104**(25) (2010) 255003.
http://dx.doi.org/10.1103/PhysRevLett.104.255003

[20] S.M. Hooker *et al.*, *J. Phys. B: At. Mol. Opt. Phys.* **47**(23) (2014) 234003.
http://dx.doi.org/10.1088/0953-4075/47/23/234003

[21] V. Barger and M. Olsson, *Classical Mechanics, A Modern Perspective* (McGraw-Hill, New York, 1973).

[22] J.B. Marion and S.T. Thornton, *Classical Dynamics of Particles and Systems*, 3rd ed. (Harcourt Brace Jovanovich, San Diego, 1988).

[23] S. Lee, *et al.*, *Phys. Rev.* E **61**(6) (2000) 7014. http://dx.doi.org/10.1103/PhysRevE.61.7014

[24] P. Muggli *et al.*, *Phys. Rev. Lett.* **93**(1) (2004) 014802.
http://dx.doi.org/10.1103/PhysRevLett.93.014802

[25] M.J. Hogan *et al.*, *Phys. Rev. Lett.* **95**(5) (2005) 054802.
http://dx.doi.org/10.1103/PhysRevLett.95.054802

[26] I. Blumenfeld *et al.*, *Nature* **445** (2007) 741. http://dx.doi.org/10.1038/nature05538

[27] C. Joshi *et al.*, *Phys. Plasmas* **9**(5) (2002) 1845. http://dx.doi.org/10.1063/1.1455003

[28] R. Maeda *et al.*, *Phys. Rev. ST Accel. Beams* **7**(11) (2004) 111301.
http://dx.doi.org/10.1103/PhysRevSTAB.7.111301

[29] C. Jing, *et al.*, *Phys. Rev. Lett.* **98**(14) (2007) 144801.
http://dx.doi.org/10.1103/PhysRevLett.98.144801

[30] V. Yakimenko *et al.*, *Phys. Rev. Lett.* **109**(16) (2012) 164802.
http://dx.doi.org/10.1103/PhysRevLett.109.164802

[31] P. Muggli *et al.*, *Phys. Rev. Lett.* **101**(5) (2008) 054801.
http://dx.doi.org/10.1103/PhysRevLett.101.054801

[32] P. Muggli *et al.*, *Phys. Rev. ST Accel. Beams* **13**(5) (2010) 052803.
http://dx.doi.org/10.1103/PhysRevSTAB.13.052803

[33] Y. Sentoku *et al.*, *Phys. Rev. Lett.* **90**(15) (2003) 155001.
http://dx.doi.org/10.1103/PhysRevLett.90.155001

[34] B. Allen *et al.*, *Phys. Rev. Lett.* **109**(18) (2012) 185007.
http://dx.doi.org/10.1103/PhysRevLett.109.185007

[35] J.M. Dawson, *Phys. Rev.* **113**(2) (1959) 383. http://dx.doi.org/10.1103/PhysRev.113.383

[36] E. Oz *et al.*, *Phys. Rev. Lett.* **98**(8) (2007) 084801.
http://dx.doi.org/10.1103/PhysRevLett.98.084801

[37] G. Saxon, *Proc. Phys. Soc.* B **67**(9) (1954) 705. http://dx.doi.org/10.1088/0370-1301/67/9/306

[38] M. Tzoufras, *et al.*, *Phys. Rev. Lett.* **101**(14) (2008) 145002.
http://dx.doi.org/10.1103/PhysRevLett.101.145002

[39] M. Tzoufras, *et al.*, *Phys. Plasmas* **16**(5) (2009) 056705. http://dx.doi.org/10.1063/1.3118628

[40] W. Lu *et al.*, *Phys. Plasmas* **12**(6) (2005) 063101. http://dx.doi.org/10.1063/1.1905587

[41] W. Lu *et al.*, *Phys. Plasmas* **13**(5) (2006) 056709. http://dx.doi.org/10.1063/1.2203364

[42] C.E. Clayton *et al.*, *Phys. Rev. Lett.* **88**(15) (2002) 154801.
http://dx.doi.org/10.1103/PhysRevLett.88.154801

[43] S. Wang *et al.*, *Phys. Rev. Lett.* **88**(13) (2002) 135004.
http://dx.doi.org/10.1103/PhysRevLett.88.135004

[44] D.K. Johnson *et al.*, *Phys. Rev. Lett.* **97**(17) (2006) 175003 (2006).
http://dx.doi.org/10.1103/PhysRevLett.97.175003

[45] N. Barov, *et al.*, *Phys. Rev. Lett.* **80**(1) (1998) 81. http://dx.doi.org/10.1103/PhysRevLett.80.81

[46] K.A. Marsh *et al.*, *Proc. 2005 Particle Accelerator Conference* (2005), pp. 2702–2704.
http://dx.doi.org/10.1109/PAC.2005.1591234

[47] P. Muggli *et al.*, *IEEE Trans. Plasma Sci.* **27**(3) (1999) 791. http://dx.doi.org/10.1109/27.774685

[48] S.Z. Green *et al.*, *Plasma Phys. Control. Fusion* **56**(8) (2014) 084011.
http://dx.doi.org/10.1088/0741-3335/56/8/084011

[49] C.L. O'Connell *et al.*, *Phys. Rev. ST Accel. Beams* **9**(10) (2006) 101301.
http://dx.doi.org/10.1103/PhysRevSTAB.9.101301

[50] P. Muggli *et al.*, *Phys. Rev. Lett.* **101**(5) (2008) 055001.
http://dx.doi.org/10.1103/PhysRevLett.101.055001

[51] J.S.T. Ng, *et al.*, *Phys. Rev. Lett.* **87**(24) (2001) 244801.
http://dx.doi.org/10.1103/PhysRevLett.87.244801

[52] Y. Iwashita *et al.*, *IEEE Trans. Appl. Superconduct.* **18**(2) (2008) 957-960.
http://dx.doi.org/10.1109/TASC.2008.921302

[53] R. Lee *et al.*, *Phys. Rev. Lett.* **31**(23) (1973) 1390. http://dx.doi.org/10.1103/PhysRevLett.31.1390

[54] V. Yakimenko *et al.*, *Phys. Rev. Lett.* **91**(1) (2003) 014802.
http://dx.doi.org/10.1103/PhysRevLett.91.014802

[55] P. Muggli *et al.*, *New J. Phys.* **12** (2010) 045022 (2010).
http://dx.doi.org/10.1088/1367-2630/12/4/045022

[56] M.J. Hogan *et al.*, *Phys. Rev. Lett.* **90**(20) (2003) 205002.
http://dx.doi.org/10.1103/PhysRevLett.90.205002

[57] S. Lee, *et al.*, *Phys. Rev.* E **64**(4) (2001) 045501(R) http://dx.doi.org/10.1103/PhysRevE.64.045501

[58] W.D. Kimura *et al.*, *Phys. Rev. ST Accel. Beams* **14**(4) (2011) 041301.
http://dx.doi.org/10.1103/PhysRevSTAB.14.041301

[59] S. Gessner *et al.,* private communication.

[60] Y. Fang *et al.*, in *Proceedings of the 2011 Particle Accelerator Conference* (IEEE, Piscataway, NJ,
2011), p. 391, http://accelconf.web.cern.ch/accelconf/PAC2011/papers/mop158.pdf

[61] J.B. Rosenzweig *et al.*, *AIP Conf. Proc.* **1299** (2010) Annapolis, (Maryland), 13–19 June 2010,
500. http://dx.doi.org/10.1063/1.3520373

[62] J.B. Rosenzweig *et al.*, *AIP Conf. Proc.* **1507** (2012) 612 http://dx.doi.org/10.1063/1.4773767

[63] A. Caldwell *et al.*, *Nature Phys.* **5** (2009) 363. http://dx.doi.org/10.1038/nphys1248

[64] Y. Fang *et al.*, *Phys. Rev. Lett.* **112**(4) (2014) 045001.
http://dx.doi.org/10.1103/PhysRevLett.112.045001

[65] J. Vieira *et al.*, *Phys. Plasmas* **19**(6) (2012) 063105. http://dx.doi.org/10.1063/1.4725425

[66] M. Gross *et al.*, in *Proceedings of the 2014 Particle Accelerator Conference* (IEEE, Piscataway,
NJ, 2014), p. 1525, http://accelconf.web.cern.ch/AccelConf/IPAC2014/papers/tupme074.pdf

[67] R Assmann *et al.*, and AWAKE Collaboration, *Plasma Phys. Control. Fusion* **56**(8) (2014)
084013. http://dx.doi.org/10.1088/0741-3335/56/8/084013

[68] E. ï£¡z *et al.*, *Nucl. Instr. Meth. Phys. Res.* A **740** (2014) 197.
http://dx.doi.org/10.1016/j.nima.2013.10.093

[69] P. Muggli *et al.*, in *Proceedings of the 2014 Particle Accelerator Conference* (IEEE, Piscataway,
NJ, 2014), p. 1470, http://accelconf.web.cern.ch/AccelConf/IPAC2014/papers/tupme048.pdf

[70] V.K. Berglyd Olsen *et al.*, in *Proceedings of the 2015 Particle Accelerator Conference* (IEEE, Piscataway, NJ, 2015), p. 2551,
http://accelconf.web.cern.ch/AccelConf/IPAC2015/papers/wepwa026.pdf

[71] K. V. Lotov *et al.*, *Phys. Plasmas* **21**(12) (2014) 123116. http://dx.doi.org/10.1063/1.4904365

[72] W. Panofsky and W. Wenzel, *Rev. Sci. Instrum.* 27(11) (1956) 967.
http://dx.doi.org/10.1063/1.1715427

Plasma Injection Schemes for Laser–Plasma Accelerators

J. Faure
Laboratoire d'Optique Appliquée, ENSTA-CNRS École Polytechnique, Palaiseau, France

Abstract

Plasma injection schemes are crucial for producing high-quality electron beams in laser–plasma accelerators. This article introduces the general concepts of plasma injection. First, a Hamiltonian model for particle trapping and acceleration in plasma waves is introduced; ionization injection and colliding-pulse injection are described in the framework of this Hamiltonian model. We then proceed to consider injection in plasma density gradients.

Keywords

Laser-plasma accelerators; laser wakefield; electron injection; ionization injection; colliding pulse injection; gradient injection.

1 Introduction

In accelerator physics, the starting point of a high-energy machine is always the injector. The injector can be extremely important; it usually delivers a particle beam at the MeV level, and its characteristics—such as emittance, bunch duration and energy spread—strongly impact the parameters of the final beam. By analogy, it seems only natural that laser–plasma accelerators should also include an injection stage; this would permit (i) decoupling of the injection and acceleration stages, (ii) more control over the beam parameters and the possibility to tune them independently, and (iii) better stability of the beam. However, injecting electrons into the plasma wakefield accelerating structure is not an easy task: the wavelength of the accelerating structure is on the order of $\lambda_p \simeq 10\ \mu$m; here, the plasma wavelength is $\lambda_p = 2\pi c/\omega_p$ (where c is the speed of light and $\omega_p = (n_e e^2/\epsilon_0 m_e)^{1/2}$ is the plasma frequency), so that $\lambda_p = 10$–$30\,\mu$m for electron plasma densities in the 10^{18}–10^{19} cm^{-3} range. The production of monoenergetic electron bunches in such a micrometre structure requires that the injected beam have a duration shorter than $\lambda_p/c = 30$–100 fs and that it be focused to a few microns. In addition to these stringent requirements, the injected particle beam should be synchronized at the femtosecond level with the laser pulse driving the wakefield, which also poses a considerable experimental challenge. The production of such short and well-synchronized bunches is at the edge of conventional radio-frequency (r.f.) accelerator technology, and experiments using external injectors have not yet succeeded in providing high-quality electron beams with narrow energy spreads [1, 2].

In many laser wakefield acceleration experiments, electrons are *self-injected* into the accelerating structure: when the plasma wave amplitude reaches very high levels (close to the wave-breaking threshold), background plasma electrons can be injected into the plasma waves [3–7]. This is somewhat analogous to dark current in a r.f. accelerator: at high field strengths, the cavities start to release electrons. This self-injection mechanism can lead to the production of high charge and, on occasion, narrow energy spreads in the range of a few percent to 10% can be achieved. However, it has proven difficult to obtain high-quality, i.e. narrow, energy spreads in a stable manner using this method. In addition, electron injection results from a succession of nonlinear effects, such as relativistic self-focusing [8, 9] of the laser pulse, spectral broadening and self-steepening [10–13]. This is the reason that self-injection is rather difficult to control precisely and does not allow tuning of the injected beam parameters. In this context, many researchers in the field of laser–plasma accelerators have started to develop injection methods in which the injected electrons originate from the plasma itself (as opposed to external injection requiring an existing electron source).

While several methods have been proposed for injecting plasma electrons into an existing plasma wakefield, they can be summarized by the following general principles.

- Create electrons at the right phase in the wakefield. Even if an electron is created at rest, it can end up being trapped provided that it is born at the appropriate phase. This is the idea behind ionization injection.
- Give an initial kick to electrons so that their initial longitudinal velocity is high enough for trapping. This is analogous to surfers paddling in order to gain momentum and catch the wave. Several methods are based on this idea, including injection by colliding pulses.
- Slow down the wakefield to facilitate trapping. Decreasing the phase velocity of the wakefield locally can be achieved by tailoring the plasma density, for example, and it leads to a controlled injection mechanism. Injection in downward density ramps is based on this idea.

In this article we will review some of these injection mechanisms. In Section 2, a Hamiltonian formalism based on Refs. [14, 15] is derived in order to find the trapping threshold for an electron interacting with an intense laser pulse and a plasma wave. In Section 3, the findings from the Hamiltonian model will be applied to the particular case of ionization injection where high-Z atoms provide the source of injected electrons. In Section 4 we show how electrons can be injected using colliding and counter-propagating laser pulses. Finally, in Section 5, we show how electrons can also be trapped in wakefields with decreasing phase velocities.

2 One-dimensional Hamiltonian model

2.1 Assumptions of the model

The goal of this section is to derive analytically the trajectories of electrons in a laser field and a plasma wave [15]. In particular, some electrons can be trapped and accelerated in the plasma wakefield. The theoretical framework is the following.

- We start with a 1D model: we consider only the motion of electrons along the longitudinal coordinate z, and neglect the role of the radial electric fields. In this case, the wakefield potential ϕ depends only on z and t.
- For simplicity, we assume that the laser pulse does not change during its propagation. Consequently, the plasma wakefield is also stationary along the propagation. This is important because it will allow us to use a conservation-of-energy law.
- The plasma is modelled by an electron fluid. This fluid is described by macroscopic quantities such as its density $n(\mathbf{r}, t)$ and velocity $\mathbf{v}(\mathbf{r}, t)$. Let us note that in such a model, kinetic effects (e.g. trapping, wave-breaking) are not taken into account.
- Plasma ions are assumed to be immobile. This is justified when the typical time for ion motion ($\omega_{\mathrm{pi}}^{-1}$) is large compared to the driver pulse duration (i.e. $\tau \ll \omega_{\mathrm{pi}}^{-1}$).
- The electron fluid is cold. In the case of a laser driver, this is justified when the quiver velocity of electrons in the laser field is orders of magnitude greater than the thermal velocity: $v_{\mathrm{osc}} \simeq eE_{\mathrm{laser}}/(m_{\mathrm{e}}\omega_0) \gg v_{\mathrm{th}} = (k_{\mathrm{B}}T_{\mathrm{e}}/m_{\mathrm{e}})^{1/2}$.

The laser pulse is represented by its normalized vector potential $\mathbf{a} = e\mathbf{A}/m_{\mathrm{e}}c$, where $\mathbf{A}$ is the laser vector potential. For a pulse propagating along the z-axis and polarized along the x-axis,

$$\mathbf{a} = \hat{a}(z, t)\cos(k_0 z - \omega_0 t)\,\mathbf{e}_x, \tag{1}$$

where k_0 and ω_0 are the wavevector and frequency of the laser electromagnetic field, and $\hat{a}$ is an envelope function that represents the longitudinal shape of the pulse. We assume a Gaussian shape for $\hat{a}$ so that

$\hat{a}^2(\zeta) = a_0^2 \exp(-\zeta^2/L_0^2)$, where $\zeta = z - v_{\mathrm{g}}t$ with v_{g} being the laser group velocity, L_0 is the laser pulse length, and a_0 is the laser peak amplitude, $a_0 = 8.6 \times 10^{-10}\sqrt{I\,[\,\mathrm{W\,cm^{-2}}]\,\lambda^2\,[\mu\mathrm{m}]}$, with I and λ being the laser intensity and the central wavelength, respectively. The plasma wakefield is represented by its normalized potential $\phi = e\Phi/m_{\mathrm{e}}c^2$, which is obtained by solving the equation

$$\frac{\partial^2 \phi}{\partial \zeta^2} = k_{\mathrm{p}}^2 \gamma_{\mathrm{p}}^2 \left[\beta_{\mathrm{p}} \left(1 - \frac{1+a^2}{\gamma_{\mathrm{p}}^2(1+\phi)^2} \right)^{-1/2} - 1 \right]. \tag{2}$$

Here, γ_{p} is the Lorentz factor corresponding to the plasma wave phase velocity v_{p}: $\gamma_{\mathrm{p}} = (1 - v_{\mathrm{p}}^2/c^2)^{-1/2}$. It is also assumed that the plasma wave velocity is given by the group velocity of the excitation laser: $v_{\mathrm{p}} \simeq v_{\mathrm{g}}$ and $\gamma_{\mathrm{p}} \simeq \gamma_{\mathrm{g}}$. Finally, the plasma wavevector is $k_{\mathrm{p}} = \omega_{\mathrm{p}}/v_{\mathrm{p}}$. Eq. (2) is the nonlinear plasma wave equation, which is valid even for $a > 1$, i.e. for relativistic laser intensities with $I > 10^{18}\,\mathrm{W\,cm^{-2}}$. Figure 1 shows the result for a 20 fs laser pulse with $a_0 = 2$ and a plasma density of $n_{\mathrm{e}} = 7 \times 10^{18}\,\mathrm{cm^{-3}}$. Note that the laser field was averaged over the fast time-scale at ω_0 for clarity (keeping the fast frequency does not affect our conclusions). Notice the characteristic nonlinear shape of the plasma density perturbation, which in turn causes a nonlinear longitudinal electric field with a sawtooth shape.
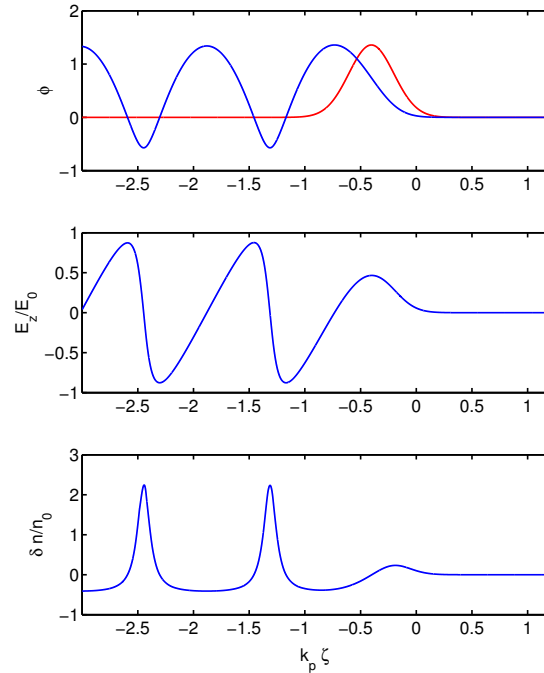
Fig. 1: The top panel plots the laser amplitude a (red) and the wakefield potential ϕ (blue); the middle panel shows the corresponding normalized electric field E_z/E_0 (where $E_0 = m_{\mathrm{e}}c\,\omega_{\mathrm{p}}/e$ is the wave-breaking field). The bottom panel shows the electron density perturbation $n(\zeta)/n_0 - 1$.

2.2 Hamiltonian for an electron interacting with a laser and a plasma wave

Let us consider an electron with Lorentz factor γ in a 1D wakefield, represented by the normalized potential ϕ. Its Hamiltonian reads $H = \gamma - \phi(z - v_{\mathrm{g}}t) = \sqrt{1 + u_\perp^2 + u_z^2} - \phi(z - v_{\mathrm{g}}t)$, where $u_\perp = p_\perp/(m_{\mathrm{e}}c)$ and $u_z = p_z/(m_{\mathrm{e}}c)$ are the transverse and longitudinal normalized linear momenta, respectively. The Hamiltonian depends on z and t in a particular manner: it depends only on the

variable $\zeta = z - v_{\mathrm{g}}t$. We can use this fact to change variables using a canonical transformation $(z, u_z) \rightarrow (\zeta, u_z)$. We use a second-type generating function $F_2(z, u_z) = u_z(z - v_{\mathrm{g}}t)$; the transformation satisfies $\zeta = \partial F_2/\partial u_z = z - v_{\mathrm{g}}t$ and $u_z = \partial F_2/\partial z$, and the new Hamiltonian is given by $H' = H + \frac{1}{c}\frac{\partial F_2}{\partial t}$. In this case, the new Hamiltonian (denoted simply by H from now on) is

$$H = \sqrt{1 + u_\perp^2 + u_z^2} - \phi(\zeta) - \beta_{\mathrm{p}}u_z \tag{3}$$

where $\beta_{\mathrm{p}} = v_{\mathrm{p}}/c$.

This Hamiltonian has several constants of motion. We now introduce the canonical momentum $\mathbf{P} = \mathbf{p} + q\mathbf{A}$. In normalized units, the canonical momentum is denoted by $\mathbf{U}$ and we see that $U_z = u_z$ and $U_\perp = u_\perp - a$ ($q = -e$ for electrons), because in 1D the laser field has only a transverse component. Writing the Hamiltonian with just the transverse canonical momentum gives

$$H = \sqrt{1 + (U_\perp + a)^2 + u_z^2} - \phi(\zeta) - \beta_{\mathrm{p}}u_z.$$

From Hamilton's equations, one finds that in 1D, the transverse canonical momentum is conserved:

$$\dot{U}_\perp = -\frac{\partial H}{\partial r_\perp} = 0 \implies u_\perp(\zeta) - a(\zeta) = \mathrm{const.}$$

In the case of an electron initially at rest in front of the laser pulse, $\zeta_{\mathrm{i}} = +\infty$, this gives the important result that $u_\perp(\zeta) = a(\zeta)$. The other constant of motion is simply the energy: this Hamiltonian does not depend on time but only on ζ; as a consequence, the energy of the system is conserved and H is conserved along an electron trajectory. Hence, for an electron with initial energy H_0, one can solve for its longitudinal momentum u_z to get

$$u_z = \beta_{\mathrm{p}}\gamma_{\mathrm{p}}^2(H_0 + \phi) \pm \gamma_{\mathrm{p}}\sqrt{\gamma_{\mathrm{p}}^2(H_0 + \phi)^2 - \gamma_\perp^2} \tag{4}$$

with $\gamma_\perp^2 = 1 + u_\perp^2$. Once $a(\zeta)$ and $\phi(\zeta)$ are known, this equation gives the electron trajectory in (ζ, u_z) phase space. Figure 2 shows various electron trajectories that were obtained for different initial conditions. For electrons initially at rest in front of the laser pulse, i.e. $\zeta_{\mathrm{i}} = +\infty$ and $u_z(\zeta_{\mathrm{i}}) = u_\perp(\zeta_{\mathrm{i}}) = 0$, the Hamiltonian is $H_0 = 1$. The trajectory of such electrons is referred to as the fluid orbit and corresponds to the trajectory of plasma background electrons that contribute to the formation of the plasma wakefield (black line in lower panel of Fig. 2). These electrons are not trapped and oscillate in the plasma wakefield with low energies. The separatrix is the special trajectory which separates the trapped from the untrapped orbits. It can be found by considering an electron moving along z with $v_z = v_{\mathrm{p}}$ (or $u_z(\zeta_{\mathrm{min}}) = \beta_{\mathrm{p}}\gamma_{\mathrm{p}}$) and located initially at a minimum of the potential $\phi(\zeta_{\mathrm{min}}) = \phi_{\mathrm{min}} < 0$, i.e. at a node of the electric field. Conservation of canonical momentum reads $u_\perp(\zeta_{\mathrm{min}}) = a(\zeta_{\mathrm{min}})$, so that the Hamiltonian on the separatrix is

$$H_{\mathrm{sep}} = \frac{\sqrt{1 + a^2(\zeta_{\mathrm{min}})}}{\gamma_{\mathrm{p}}} - \phi_{\mathrm{min}}. \tag{5}$$

This trajectory is represented by a red line in the lower panel of Fig. 2. The dashed blue lines represent some typical trapped orbits for which electrons gain significant energy in the plasma wave. Trapped orbits are obtained when the Hamiltonian satisfies $H_{\mathrm{i}} \leqslant H_{\mathrm{sep}}$. To prove this, let us consider a trapped electron behind the laser pulse at phase ζ_{i} and moving with velocity $v_z(\zeta_{\mathrm{i}}) = v_{\mathrm{p}}$. At this point in phase space, the plasma wave potential is such that $\phi(\zeta_{\mathrm{i}}) = \phi_{\mathrm{i}} \geqslant \phi_{\mathrm{min}}$. In addition, following Eq. (3), conservation of the Hamiltonian can be written as $H_{\mathrm{i}} = (1 + \beta_{\mathrm{p}}^2\gamma_{\mathrm{p}}^2)^{1/2} - \phi_{\mathrm{i}} - \beta_{\mathrm{p}}^2\gamma_{\mathrm{p}} = 1/\gamma_{\mathrm{p}} - \phi_{\mathrm{i}}$. Since $\phi_{\mathrm{i}} \geqslant \phi_{\mathrm{min}}$, we get $H_{\mathrm{i}} \leqslant H_{\mathrm{sep}}$, which is a necessary and sufficient condition for trapping.
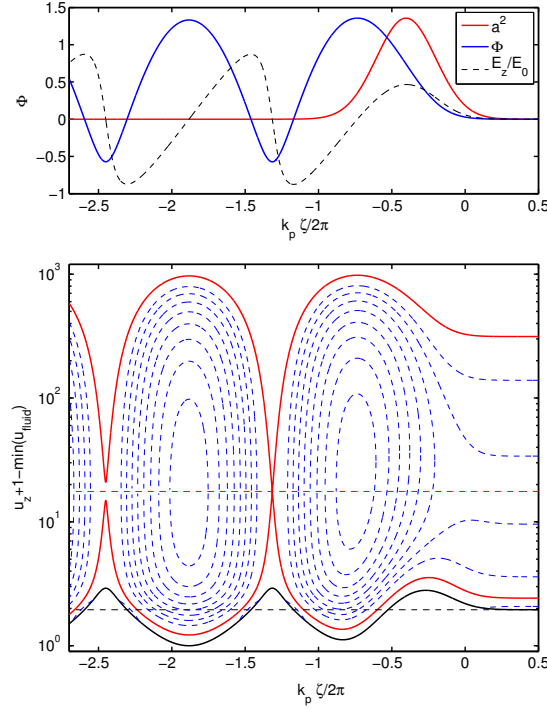
Fig. 2: The top panel plots the laser amplitude a^2 (red), the wakefield potential (blue), and the longitudinal electric field (black dashed); the bottom panel shows the associated electron trajectories in phase space. The black line in the bottom panel represents the fluid orbit, the red line the separatrix, and the dashed blue lines various trapped orbits. The y-axis is in log-scale and, to avoid negative values, we have plotted $u_z + 1 - \min(u_{\text{fluid}})$, where $u_{\text{fluid}}(\zeta)$ is the fluid orbit. The horizontal black dashed line represents $u_z = 0$, while the horizontal red dashed line represents $u_z = \beta_{\text{p}}\gamma_{\text{p}}$. The parameters are $a_0 = 2$, $n_{\text{e}}/n_{\text{c}} = 0.44\%$, $\lambda = 0.8\ \mu\text{m}$ and $\tau = 20$ fs.

2.3 Trapping thresholds

Now that the basic trajectories have been established, it is possible to determine the trapping threshold of plasma electrons in a plasma wave. To do so, we consider that an electron located in front of the laser pulse will be trapped in the wakefield if its longitudinal momentum satisfies $u_z(+\infty) > u_z^{\text{sep}}(+\infty)$, which simply means that the initial momentum has to be larger than the initial momentum on the separatrix, $u_z^{\text{sep}}(+\infty)$. According to Eq. (4), this is simply $u_z^{\text{sep}}(+\infty) = \beta_{\text{p}}\gamma_{\text{p}}^2 H_{\text{sep}} - \gamma_{\text{p}}\sqrt{\gamma_{\text{p}}^2 H_{\text{sep}}^2 - 1}$. Therefore, an electron with initial energy $E > E_{\text{trap}}$ will be trapped and accelerated in the wakefield, where

$$E_{\text{trap}} = m_{\text{e}}c^2\Big(\sqrt{1 + \{u_z^{\text{sep}}(+\infty)\}^2} - 1\Big). \tag{6}$$

Figure 3 shows the variation of the trapping threshold with the plasma wave amplitude ϕ_{min}, as well as its variation with the plasma wave Lorentz factor γ_{p}. Clearly, trapping is easier for high-amplitude plasma waves and/or for small phase velocities (small values of γ_{p}). Note that as $\phi_{\text{min}} \to -1$, the trapping threshold E_{trap} tends toward zero. This is the onset of wave-breaking: the plasma wave amplitude becomes so high that all plasma electrons that were initially at rest get injected into the plasma wave. As $\phi_{\text{min}} \to -1$, the longitudinal electric field reaches the cold wave-breaking limit, $E_{\text{WB}} = \sqrt{2}(\gamma_{\text{p}} - 1)^{1/2}E_0$, where $E_0 = m_{\text{e}}c\omega_{\text{p}}/e$ was defined earlier. Thus, this model is able to give a 1D picture of self-injection, although one has to keep in mind that in experiments self-injection is a 3D process. More complicated models are necessary to fully capture the physics of self-injection in 3D [16].
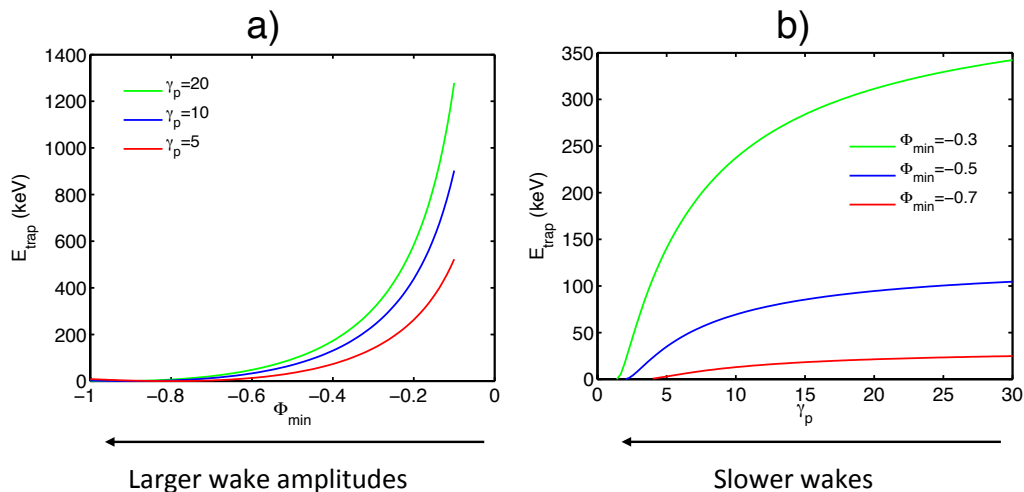
Fig. 3: (a) Trapping thresholds plotted as a function of wake amplitude (where ϕ_{min} represents the minimum of the wake potential), for three different wake phase velocities. (b) Trapping thresholds plotted as a function of γ_p, the Lorentz factor associated with the wakefield velocity, for three different wake amplitudes.

3 Ionization injection

As we have just seen, unless the wakefield reaches wave-breaking amplitudes, it is not possible for plasma electrons to be injected unless a different method is used. Ionization injection is probably the easiest method for injecting electrons in a wakefield [17, 18]. The idea is to use a high-Z gas so that the first levels of ionization occur at low intensity (typically below 10^{16} W cm^{-2}). These electrons are born at rest in a region where the laser intensity is relatively low and provide the electrons which form the plasma wave and follow fluid orbits. On the other hand, ionization from inner shells occurs at higher intensities (typically for $I > 10^{18}$ W cm^{-2}), so these electrons are born at rest in regions of strong fields. It follows that they are born inside the plasma wave, at a totally different phase from fluid electrons. The whole challenge is to put these electrons on trapped orbits so that they will be injected and accelerated. Figure 4 illustrates the different classes of trajectories of electrons born at the front of the pulse and in the middle of the pulse. In this section, we follow the approach of Refs. [17, 19].
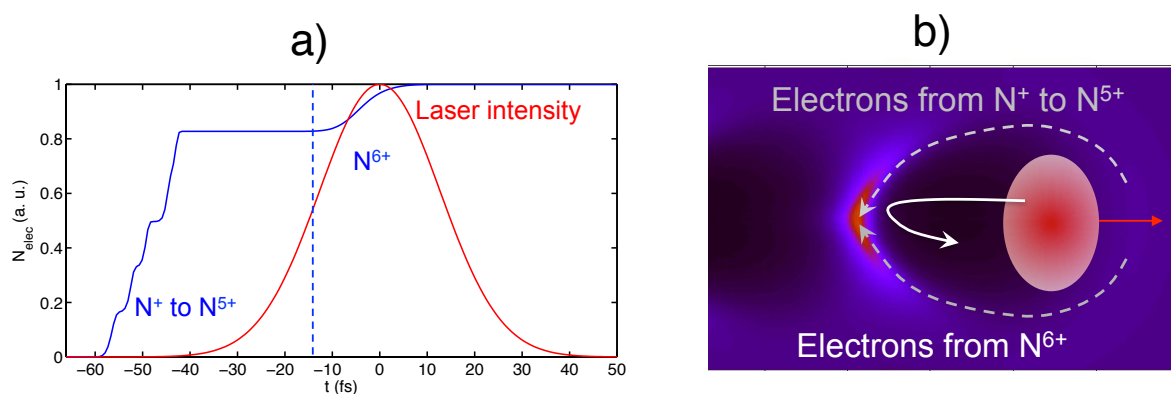
Fig. 4: Principle of ionization injection. (a) Ionization of nitrogen by an intense laser pulse at 10^{19} W cm^{-2}; the first five electrons, from N$^+$ to N^{5+}, appear at the front of the laser pulse, while electrons from N^{6+} appear in the middle of the laser pulse. (b) Schematic showing the different classes of trajectories for electrons born at the front of the pulse (fluid trajectories) or in the middle of the pulse (trapped trajectories).

We now determine the condition(s) under which these electrons will be trapped. We first make a few assumptions concerning ionization. In our model, ionization is assumed to occur via barrier suppression ionization [20]. Considering an ionization level with energy E_i, the threshold intensity required to ionize the electron is given by

$$I_{\text{thresh}} \, [\text{W cm}^{-2}] = 4 \times 10^9 \, E_i^4 \, [\text{eV}]/Z^{*2},$$

where Z^* is the charge of the resulting ion. Furthermore, we assume that electrons are born at rest (which is a good approximation when considering the energy that the electron will gain in the laser and plasma field). Technically, if the electron is born at phase ζ_{ion}, it should witness the corresponding laser amplitude $a(\zeta_{\text{ion}})$. However, in linear polarization, ionization occurs mostly at the peak of the electric field, i.e. $a(\zeta_{\text{ion}}) \simeq 0$. Therefore, in the case of an ionized electron born at rest, conservation of canonical momentum reads $u_\perp(\zeta) = a(\zeta) - a(\zeta_{\text{ion}}) \simeq a(\zeta)$. The initial Hamiltonian for such electrons can be found from Eq. (3) as

$$H_{\text{ion}} = 1 - \phi_{\text{ion}},$$

and the ionized electron trajectory can be computed by inserting H_{ion} into Eq. (4). The conditions for trapping are simply that the intensity should be high enough for ionizing a given electron level, i.e. $a(\zeta_{\text{ion}}) > a_{\text{thresh}}$, and that the electron should be born on a trapped orbit, i.e. $H_{\text{ion}} < H_{\text{sep}}$. These conditions define the region of phase space where ionized electrons are trapped and further accelerated in the wakefield. Typically, these conditions are fulfilled only in the setting of a high-intensity laser pulse ($a > 1$) and a large-amplitude plasma wave. Figure 5 illustrates this discussion: the green area in the upper panel defines the region of phase space where electrons can be injected; the corresponding trajectories are plotted in green in the lower panel (note that the fast laser frequency was considered for this case, as electrons are born at zeros of the vector potential). As discussed previously, electrons are born at rest, oscillate in the laser field, and gain energy as they get on a trapped orbit.

The injection volume in phase space depends on the wakefield amplitude and the ionization threshold, so in principle it can be controlled via the laser intensity. The injection volume is responsible for the injected charge as well as the energy spread. Obviously, large injection volumes lead to large injected charge and larger energy spreads. On the other hand, smaller energy spreads can be obtained by reducing the injection volume at the expense of charge. Experiments have demonstrated the concept of ionization injection using nitrogen and argon [17, 18]. In practice, it is quite difficult to control the injection volume experimentally: ionization injection occurs all along the propagation, as long as the laser intensity exceeds the threshold intensity; as a result, electrons are injected continuously along the propagation and the energy distributions are quite broad [17, 18]. Narrower energy spreads have been obtained by using a two-stage laser–plasma accelerator [21]: in the first stage, a short nitrogen gas jet is used to inject electrons while a second gas jet is used to boost the acceleration. In Ref. [22], by restricting the first stage to a $500 \, \mu\text{m}$ jet, the injection was kept relatively local, resulting in energy spreads at the 10% level.

To summarize our discussion on ionization injection, this method is easy to implement experimentally as it requires simply the use of a high-Z gas. It is a straightforward way to increase the injected charge without reaching wave-breaking amplitudes. It usually results in increased injected beam loads, but controlling the beam quality can be rather difficult, as (i) the injection volume is directly related to the laser intensity, and (ii) injection tends to occur continuously along the propagation.

4 Colliding-pulse injection

We now present the colliding pulse injection scheme, where an auxiliary laser pulse is used to trigger a very localized injection ($\simeq 10 \, \mu\text{m}$), resulting in beams with small energy spread. The idea of using an additional laser pulse for injecting electrons was first proposed by Umstadter *et al.* [23]. It was further developed by Esarey *et al.* [24], who proposed a scheme based on counter-propagating laser pulses. In its simplest form, the scheme uses two counter-propagating ultra-short laser pulses with the same
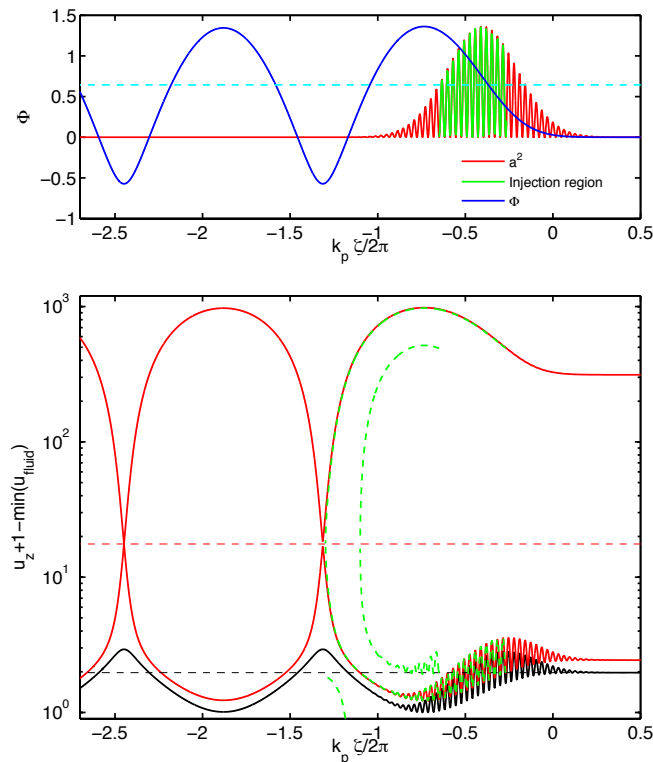
Fig. 5: Ionization injection—a phase space picture. The upper panel plots the laser amplitude a^2 (red) and the wakefield potential (blue); the dashed light-blue horizontal line indicates the ionization threshold $a_{\text{thresh}} = 1.37$, corresponding here to $I_{\text{thresh}} = 4 \times 10^{18} \text{ W cm}^{-2}$, and the region of phase space where the injection conditions are satisfied is shown in green. The lower panel shows the fluid orbit (black), the separatrix (red), and the trajectories followed by ionized electrons at the front and back of the injection region. The parameters are $a_0 = 2$, $n_{\text{e}}/n_{\text{c}} = 0.44\%$, $\lambda = 0.8 \,\mu\text{m}$ and $\tau = 20$ fs.

central wavelength and polarization. The first laser pulse, the 'pump' pulse, creates a wakefield, whereas the second laser, the 'injection' pulse, is used only for injecting electrons. The laser pulses collide in the plasma, and their interference creates a laser beat-wave pattern which can pre-accelerate plasma background electrons. If the laser intensities are high enough, this pre-acceleration permits the injection and trapping of electrons in the wakefield and their further acceleration to relativistic energies. The principle of the method is depicted in Fig. 6. Analytical work [25] and simulations [25, 26] have shown that this two-stage acceleration mechanism can lead to the production of high-quality electron bunches with narrow energy spread, small divergence and ultra-short duration, even when using relatively modest lasers (e.g. with normalized vector potential $a_0 = 1$ for the pump pulse and $a_1 = 0.3$ for the injection pulse).

A theoretical description of electron injection by colliding laser pulses was first given by Esarey *et al.* in Ref. [24]. We will briefly review the principles of their fluid model. First, the pump laser, with normalized vector potential a_0, excites a plasma wave. Before the arrival of the laser pulse, the plasma background electrons are initially at rest. In the absence of wave-breaking and self-trapping, these electrons are not trapped in the plasma wave. As we have seen earlier, they simply oscillate back and forth in the plasma wave potential along the fluid orbit, and their momenta oscillate between $u_{z,\text{min}}^{\text{fluid}}$ and $u_{z,\text{max}}^{\text{fluid}}$, where

$$u_{z,\text{max,min}}^{\text{fluid}} = \beta_{\text{p}} \gamma_{\text{p}}^2 (1 + \phi_{\text{min,max}}) - \gamma_{\text{p}} \sqrt{\gamma_{\text{p}}^2 (1 + \phi_{\text{min,max}})^2 - 1}, \qquad (7)$$

with ϕ_{min} and ϕ_{max} representing the minimum and maximum of the potential, respectively. In order to
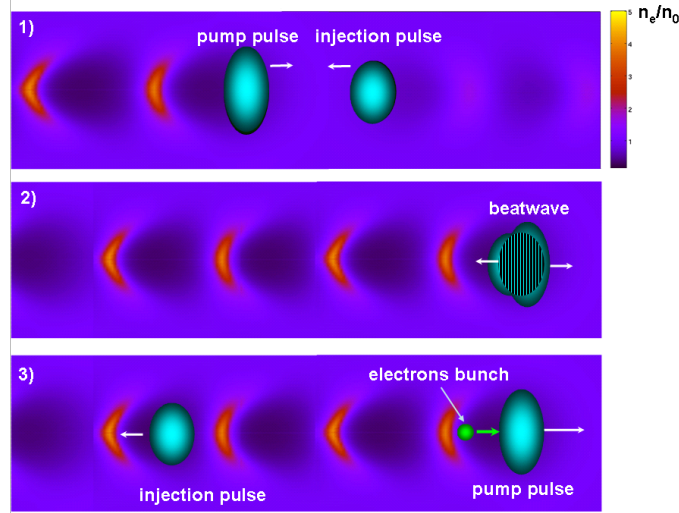
Fig. 6: Principle of colliding-pulse injection

be accelerated in the wakefield, electrons need to move along trapped orbits, inside the separatrix. When electrons move along the separatrix, their momenta oscillate between $u_{z,\mathrm{min}}^{\mathrm{sep}}$ and $u_{z,\mathrm{max}}^{\mathrm{sep}}$, where

$$u_{z,\mathrm{max,min}}^{\mathrm{sep}} = \beta_{\mathrm{p}}\gamma_{\mathrm{p}}(1 + \gamma_{\mathrm{p}}\Delta\phi) \pm \gamma_{\mathrm{p}}\sqrt{(1 + \gamma_{\mathrm{p}}\Delta\phi)^2 - 1} \tag{8}$$

with $\Delta\phi = \phi_{\mathrm{max}} - \phi_{\mathrm{min}}$.

Therefore, one needs to find a way to push plasma background electrons, which follow the fluid trajectories, into trapped trajectories. The interference of two laser pulses generates a beat-wave which is able to heat electrons and provide just such a mechanism. The laser pulses are represented by their normalized vector potentials $\mathbf{a}_{0,1}$, where the subscripts 0 and 1 represent, respectively, the pump pulse and the injection pulse. Assuming that the lasers are counter-propagating along the z direction, one can write the normalized vector potential as

$$\mathbf{a}_{0,1} = \frac{a_{0,1}(\zeta_{0,1})}{\sqrt{1+\sigma}}\left[\cos(k_{0,1}\zeta_{0,1})\mathbf{e}_x + \sigma\sin(k_{0,1}\zeta_{0,1})\mathbf{e}_y\right], \tag{9}$$

where $\sigma = 0$ for linear polarization and $\sigma = 1$ for circular polarization, $k_0 = -k_1$ are the wavevectors, and $\zeta_{0,1} = z \pm v_{\mathrm{g}}t$. Although in general the two waves can have different frequencies, all experiments to date have used identical frequencies, so here we assume that the two laser pulses have the same frequency ω_0.

When the two pulses overlap, they interfere, and the resulting squared electromagnetic field can be written as $\mathbf{a}^2 = (\mathbf{a}_0^2 + \mathbf{a}_1^2 + 2\mathbf{a}_0 \cdot \mathbf{a}_1)/(1 + \sigma)$. The last term is the beat-wave term, and it cancels out for crossed polarizations. In the case where the two polarizations are parallel, it generates a standing wave (i.e. the beat-wave has zero phase velocity) with a spatial scale of $\lambda_0/2$. The ponderomotive force in the beat-wave is very large (because $F_{\mathrm{beat}} \simeq 2k_0 a_0 a_1$), and it can pre-accelerate the plasma electrons. Neglecting the plasma potential, electron trajectories in the beat-wave are governed by the Hamiltonian

$$H_{\mathrm{b}} = \sqrt{1 + u_z^2 + \mathbf{a}^2}. \tag{10}$$

In order to obtain an analytical estimate, we will assume circular polarization ($\sigma = 1$) for the pump and injection beams, so that $\mathbf{a}^2 = (a_0^2 + a_1^2)/2 + a_0 a_1 \cos(2k_0 z)$. The separatrix in this beat-wave pattern is then given by

$$u_z^{\mathrm{beat}} = \pm\sqrt{a_0 a_1(1 - \cos 2k_0 z)}; \tag{11}$$

so on the beat-wave separatrix, electrons oscillate between $u_{z,\text{min}}^{\text{beat}}$ and $u_{z,\text{max}}^{\text{beat}}$ (see Fig. 7), where

$$u_{z,\text{max,min}}^{\text{beat}} = \pm\sqrt{2a_0 a_1}. \tag{12}$$

For instance, an electron trapped in the beat-wave created by two laser pulses with $a_0 = 2$ and $a_1 = 0.3$ can gain about 200 keV, which is sufficient for getting trapped in the wakefield.
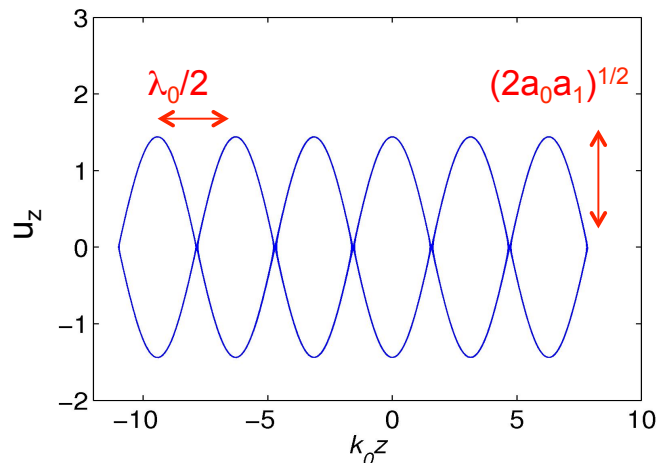
Fig. 7: Trajectories of electrons following the beat-wave separatrix at the collision of the two circularly polarized laser pulses.

One can define an approximate threshold for injection into the wakefield by applying a phase space separatrix overlap condition [24,25]. Specifically, island overlap requires the following conditions:

$$
\begin{aligned}
u_{z,\text{max}}^{\text{beat}} &> u_{z,\text{min}}^{\text{sep}}, \\
u_{z,\text{min}}^{\text{beat}} &< u_{z,\text{min}}^{\text{fluid}}.
\end{aligned}
\tag{13}
$$

These conditions are illustrated in Fig. 8(a). Using criterion (13), we find that the injection threshold can be reached with $a_0 \simeq 1$ and $a_1 \simeq 0.1$ (for this estimation, we have used pulse durations of $\tau = 30$ fs to calculate the plasma wave amplitude ϕ), for plasma density in the range of $n_e = 10^{18}$–10^{19} cm^{-3}; see the results in Fig. 8(b).

When the pulse polarizations are linear, the Hamiltonian in Eq. (10) is time-dependent and no longer integrable. It has been shown that in this case, electron motion is stochastic [27–29] and the simple calculations above do not hold. In fact, stochastic heating turns out to be an efficient mechanism for heating and subsequently injecting electrons [28,30]. Other effects, such as plasma wave inhibition [31], make the physics of colliding pulses more complicated than what this simple model tends to suggest. Full 3D particle-in-cell simulations are necessary to fully capture all the physical ingredients [30], but the concepts presented here still hold.

The colliding-pulse injection method requires some heavy experimental investment, as a second intense laser pulse is needed. The two laser pulses need to be precisely synchronized at the femtosecond level and spatially overlapped at the micron level. Despite its complexity, colliding-pulse injection has intrinsic advantages, some of which are summarized as follows.

– The injection is very localized as it occurs only during the collision of the two laser pulses, i.e. $10\,\mu$m for 30 fs pulses. This helps to reduce the energy spread and to inject very short electron bunches. The injected electron beams were measured to be only a few femtoseconds in duration [32].
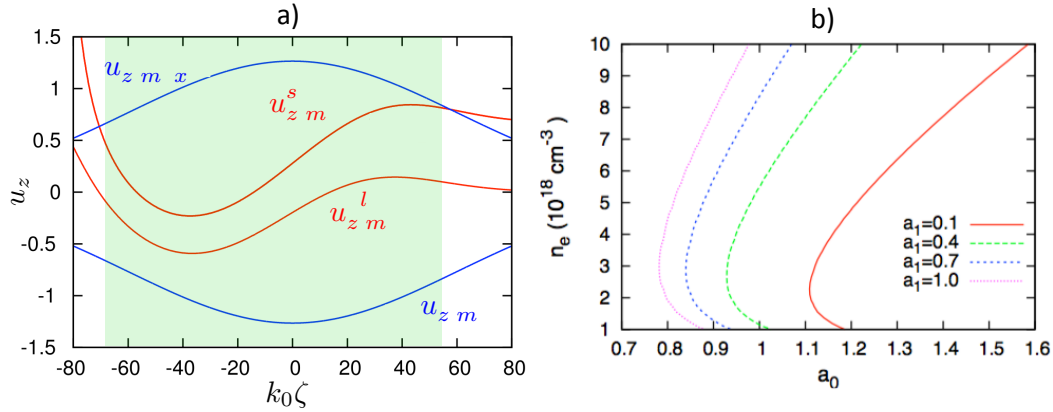
Fig. 8: (a) Phase space representation of the injection conditions in (13) in the case of colliding pulses: the blue lines represent the amplitude of the beat-wave separatrix, while the red lines are the wakefield separatrix and the fluid orbit; the area shaded in green represents in injection volume. (b) Plots of the injection thresholds for various laser and plasma parameters (a 30 fs and 800 nm laser pulse was assumed).

– The physics of injection is largely determined by the heating at the collision, so injection is less sensitive to intensity fluctuations of the main laser pulse (though the intensity of the main pulse still plays a role, as it determines the wakefield amplitude). First experiments have shown that stable beams with narrow energy spreads can be achieved in this way [33, 34].

– The location of the collision and hence the injection can be easily controlled. Therefore it is possible to tune the acceleration length. Experiments have shown that by controlling this parameter, the beam energy can be controlled over a wide range, typically 10–200 MeV [33].

– The injection volume can be controlled by tuning the intensity of the injection laser pulse. For high a_1, the electron heating at the collision is large and results in large injected beam loads as well as larger energy spreads. The energy spread can be tuned and reduced at the 1% level by decreasing the injection laser amplitude, as demonstrated experimentally in Ref. [35].

Finally, based on theoretical studies, a three-pulse scheme has been proposed where the main beam is used only to generate the wakefield [24]; two auxiliary beams are used for injection only. Their polarization is orthogonal to the main beam polarization so that the two injection pulses do not interfere with the main beam. The advantage of this method is that it provides an additional tuning knob: the injection phase can now be chosen by tuning the distance between the main pulse and the collision point. More recently, several variations of the colliding-pulse scheme have been proposed; the basic concepts remain similar but the analysis extends to the 3D case [36, 37].

5 Injection in density gradients

We now focus on another scheme, in which the plasma needs to be engineered: by tailoring the plasma density, one can gain some control over the plasma wave phase velocity. As we have seen earlier, the lower the phase velocity, the lower the trapping threshold. Therefore, by setting up a local decrease in the phase velocity, one can trigger injection in a local manner as well [38–40]. This can be achieved by sending the laser pulse through a downward density ramp, which causes the wakefield to slow down. A full Hamiltonian description of this problem is beyond the scope of this article; in a density transition, the wakefield potential also depends on z, as $\phi(z, \zeta)$, and the constants of motion that we have derived in Section 2 no longer hold in this case. Consequently, we will develop a simple fluid model that will provide some physical intuition of this process.

We start with the fluid equation describing the excitation of the wakefield by an intense laser pulse, restricting ourselves to the weakly relativistic case where $a^2 \ll 1$. We consider a gentle density gradient,

$\frac{1}{n_0}\frac{\partial n_0}{\partial z} \ll k_{\mathrm{p}}$, or $L_{\mathrm{g}}k_{\mathrm{p}} \ll 1$, where L_{g} is the gradient scale length. In this case, the plasma wave equation reads

$$\left(\frac{\partial^2}{\partial t^2} + \omega_{\mathrm{p}}^2(z)\right)\phi = \omega_{\mathrm{p}}^2(z)\frac{\langle a^2 \rangle}{2}. \tag{14}$$

Note the spatial dependence of the plasma frequency in the gradient, $\omega_{\mathrm{p}}(z) \propto \sqrt{n_0(z)}$. We can now perform the usual change of variables to solve the problem in the moving frame: $\zeta = z - v_{\mathrm{g}}t$ and $\tau = t$. By applying the quasi-static approximation, one can then neglect the $\partial/\partial\tau$ terms. The previous equation, written in the co-moving variables, becomes

$$\left(\frac{\partial^2}{\partial\zeta^2} + k_{\mathrm{p}}^2(z)\right)\phi = k_{\mathrm{p}}^2(z)\frac{\langle a^2 \rangle}{2} \tag{15}$$

where $k_{\mathrm{p}}(z) = \omega_{\mathrm{p}}(z)$. This equation can be integrated, and the solution behind the laser pulse has the form

$$\phi(\zeta, z) = \phi_0(z)\sin[k_{\mathrm{p}}(z)(z - v_{\mathrm{g}}t)], \tag{16}$$

where the wakefield amplitude is $\phi_0(z) = -\frac{\sqrt{\pi}}{4}a_0^2(z)k_{\mathrm{p}}(z)L_0\exp\{-k_{\mathrm{p}}(z)^2L_0^2/4\}$ and its phase is $\varphi = k_{\mathrm{p}}(z)(z - v_{\mathrm{g}}t)$, so that one can compute the local oscillation frequency and wavevector:

$$\omega = -\partial\varphi/\partial t = k_{\mathrm{p}}(z)/v_{\mathrm{g}} = \omega_{\mathrm{p}}(z), \tag{17}$$

$$k = \partial\varphi/\partial z = k_{\mathrm{p}}(z) + \partial k_{\mathrm{p}}/\partial z(z - v_{\mathrm{g}}t). \tag{18}$$

It follows that in a downward density gradient ($\partial k_{\mathrm{p}}/\partial z < 0$ and $z - v_{\mathrm{g}}t < 0$ behind the laser pulse), the wavevector increases with time (alternatively, the plasma wavelength decreases with time). In contrast, the plasma frequency does not depend on time: $\omega = \omega_{\mathrm{p}}(z)$. As a result of this time-varying wavevector, the phase velocity $v_{\mathrm{p}}(z,t) = \omega_{\mathrm{p}}(z)/k(z,t)$ is

$$v_{\mathrm{p}}(z,t) = \frac{v_{\mathrm{g}}}{1 + (z - v_{\mathrm{g}}t)\frac{1}{k_{\mathrm{p}}}\frac{\mathrm{d}k_{\mathrm{p}}}{\mathrm{d}z}}. \tag{19}$$

Consequently, as the wavevector increases with time, the phase velocity decreases. Injection occurs behind the laser pulse when the wakefield becomes slow enough to trap plasma background electrons.

These effects are illustrated in Fig. 9. In panel (a), we display the wakefield potential in the case where a laser pulse is tightly focused into a density gradient. The density gradient has a Gaussian shape and $k_{\mathrm{p0}}L_{\mathrm{g}} = 50$, where $k_{\mathrm{p0}} = \omega_{\mathrm{p0}}/v_{\mathrm{g0}}$ is defined using the maximum density at the top of the downward ramp. One can see clearly that the plasma frequency decreases with z as the density decreases, as expected. In addition, while the initial phase velocity is close to c, the bending of the wakefield in the (t, z) plane indicates that the phase velocity decreases with time. In panel b), the phase velocity is plotted for various density scale lengths L_{g}. It can be seen that the phase velocity decreases faster for sharper gradients, indicating that trapping is likely to occur more rapidly in this case.

Numerous experiments have demonstrated that trapping in density gradients is an efficient method for injecting electrons into wakefields [41–44]. This controlled injection scheme has resulted in more stable beams [41, 43] with energy spreads in the 10% range. Some experiments have utilized the density down ramp at the exits of short gas jets, resulting in the production of low-energy beams of a few hundred keV [41, 44]. Other research groups have engineered a sharp density gradient by creating a shock in the gas flow [43, 45] or by using another laser pulse to create a density perturbation [42]. In this case, the injection location can be controlled and the resulting beam energy can be tuned [45, 46].

6 Conclusion

In conclusion, plasma injection schemes are at the forefront of current research in laser–plasma accelerators. Researchers continue to propose new injection schemes to test new ideas and to provide high-quality
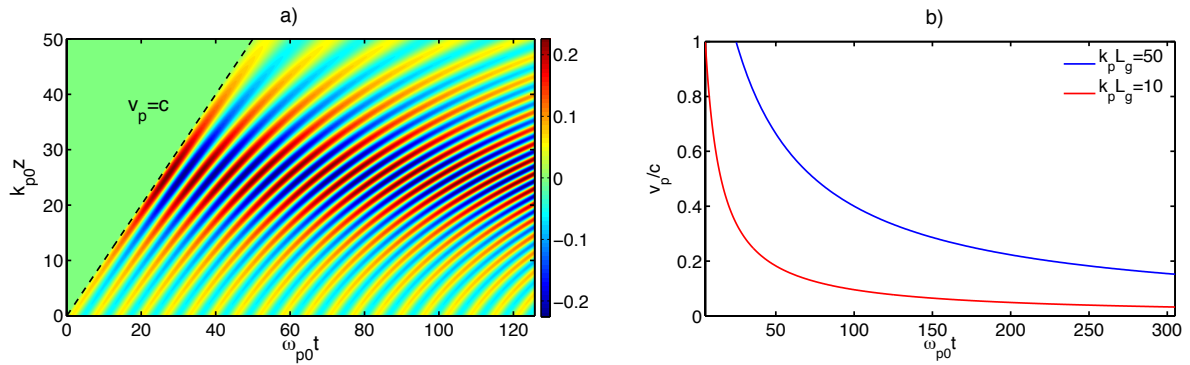
Fig. 9: (a) Wakefield potential map $\phi(z,t)$ in the case where a laser pulse is focused in a density gradient; Eq. (16) was used to compute the potential; the dashed line represents the speed of light and one can see clearly the bending of the wakefield in the (t, z) plane, indicating that the phase velocity is decreasing with time. (b) Illustration of the decrease in the phase velocity with time for two different gradient scale lengths L_g (computed using Eq. (19)).

beams while maintaining a relatively simple experimental set-up [47, 48]. The idea of combining several methods might prove useful in the future; for example, performing colliding-pulse injection in a density gradient could increase the trapped charge while still preserving the beam quality [49]. Similarly, combining ionization injection and density gradient injection has the potential to yield interesting results.

Acknowledgements

References

[1] M. Everett, A. Lal, D. Gordon, C. Clayton, K. Marsh and C. Joshi, *Nature* **368** (1994) 527. http://dx.doi.org/10.1038/368527a0

[2] F. Amiranoff, S. Baton, D. Bernard, B. Cros, D. Descamps, F. Dorchies, F. Jacquet, V. Malka, G. Matthieussent, J.R. Marquès *et al.*, *Phys. Rev. Lett.* **81** (1998) 995. http://dx.doi.org/10.1103/PhysRevLett.81.995

[3] A. Modena, A. Dangor, Z. Najmudin, C. Clayton, K. Marsh, C. Joshi, V. Malka, C. Darrow, D. Neely and F. Walsh, *Nature* **377** (1995) 606. http://dx.doi.org/10.1038/377606a0

[4] D. Umstadter, S.-Y. Chen, A. Maksimchuk, G. Mourou and R. Wagner, *Science* **273** (1996) 472. http://dx.doi.org/10.1126/science.273.5274.472

[5] J. Faure, Y. Glinec, A. Pukhov, S. Kiselev, S. Gordienko, E. Lefebvre, J.-P. Rousseau, F. Burgy and V. Malka, *Nature* **431** (2004) 541. http://dx.doi.org/10.1038/nature02963

[6] C.G.R. Geddes, C.Tóth, J. van Tilborg, E. Esarey, C.B. Schroeder, D. Bruhwiler, C. Nieter, J. Cary and W.P. Leemans, *Nature* **431** (2004) 538. http://dx.doi.org/10.1038/nature02900

[7] S.P.D. Mangles, C.D. Murphy, Z. Najmudin, A.G.R. Thomas, J.L. Collier, A.E. Dangor, E.J. Divall, P.S. Foster, J.G. Gallacher, C.J. Hooker *et al.*, *Nature* **431** (2004) 535. http://dx.doi.org/10.1038/nature02939

[8] G. Sun, E. Ott, Y.C. Lee and P. Guzdar, *Phys. Fluids* **30** (1987) 526. http://dx.doi.org/10.1063/1.866349

[9] A.B. Borisov, A.V. Borovskiy, O.B. Shiryaev, V.V. Korobkin, A.M. Prokhorov, J.C. Solem, T.S. Luk, K. Boyer and C.K. Rhodes, *Phys. Rev.* **A45** (1992) 5830. http://dx.doi.org/10.1103/PhysRevA.45.5830

[10] E. Esarey, C.B. Schroeder, B.A. Shadwick, J.S. Wurtele and W.P. Leemans, *Phys. Rev. Lett.* **84** (2000) 3081. http://dx.doi.org/10.1103/PhysRevLett.84.3081

[11] D.F. Gordon and B. Hafizi, *Phys. Rev. Lett.* **90** (2003) 215001. http://dx.doi.org/10.1103/PhysRevLett.90.215001

[12] J. Faure, Y. Glinec, J.J. Santos, F. Ewald, J.-P. Rousseau, S. Kiselev, A. Pukhov, T. Hosokai and V. Malka, *Phys. Rev. Lett.* **95** (2005) 205003. http://dx.doi.org/10.1103/PhysRevLett.95.205003

[13] J. Schreiber, C. Bellei, S.P.D. Mangles, C. Kamperidis, S. Kneip, S.R. Nagel, C.A.J. Palmer, P.P. Rajeev, M.J.V. Streeter and Z. Najmudin, *Phys. Rev. Lett.* **105** (2010) 235003. http://dx.doi.org/10.1103/PhysRevLett.105.235003

[14] E. Esarey and M. Pilloff, *Phys. Plasmas* **2** (1995) 1432. http://dx.doi.org/10.1063/1.871358

[15] E. Esarey, P. Sprangle, J. Krall and A. Ting, *IEEE Trans. Plasma Sci.* **24** (1996) 252. http://dx.doi.org/10.1109/27.509991

[16] I. Kostyukov, A. Pukhov and S. Kiselev, *Phys. Plasmas* **11** (2004) 5256. http://dx.doi.org/10.1063/1.1799371

[17] A. Pak, K.A. Marsh, S.F. Martins, W. Lu, W.B. Mori and C. Joshi, *Phys. Rev. Lett.* **104** (2010) 025003. http://dx.doi.org/10.1103/PhysRevLett.104.025003

[18] C. McGuffey, A.G.R. Thomas, W. Schumaker, T. Matsuoka, V. Chvykov, F.J. Dollar, G. Kalintchenko, V. Yanovsky, A. Maksimchuk, K. Krushelnick *et al.*, *Phys. Rev. Lett.* **104** (2010) 025004. http://dx.doi.org/10.1103/PhysRevLett.104.025004

[19] M. Chen, E. Esarey, C.B. Schroeder, C.G.R. Geddes and W.P. Leemans, *Phys. Plasmas* **19** (2012) 033101. http://dx.doi.org/10.1063/1.3689922

[20] M.V. Ammosov, N.B. Delone and V.P. Krainov, *Sov. Phys. JETP* **64** (1986) 1191.

[21] B.B. Pollock, C.E. Clayton, J.E. Ralph, F. Albert, A. Davidson, L. Divol, C. Filip, S.H. Glenzer, K. Herpoldt, W. Lu *et al.*, *Phys. Rev. Lett.* **107** (2011) 045001. http://dx.doi.org/10.1103/PhysRevLett.107.045001

[22] G. Golovin, S. Chen, N. Powers, C. Liu, S. Banerjee, J. Zhang, M. Zeng, Z. Sheng and D. Umstadter, *Phys. Rev. ST Accel. Beams* **18** (2015) 011301. http://dx.doi.org/10.1103/PhysRevSTAB.18.011301

[23] D. Umstadter, J.-K. Kim and E. Dodd, *Phys. Rev. Lett.* **76** (1996) 2073. http://dx.doi.org/10.1103/PhysRevLett.76.2073

[24] E. Esarey, R.F. Hubbard, W.P. Leemans, A. Ting and P. Sprangle, *Phys. Rev. Lett.* **79** (1997) 2682. http://dx.doi.org/10.1103/PhysRevLett.79.2682

[25] G. Fubiani, E. Esarey, C. Schroeder and W. Leemans, *Phys. Rev.* **E70** (2004) 016402. http://dx.doi.org/10.1103/PhysRevE.70.016402

[26] H. Kotaki, S. Masuda, M. Kando, J. Koga and K. Nakajima, *Phys. Plasmas* **11** (2004) 3296.

[27] J.T. Mendonça, *Phys. Rev.* **A28** (1983) 3592. http://dx.doi.org/10.1103/PhysRevA.28.3592

[28] Z.-M. Sheng, K. Mima, Y. Sentoku, M.S. Jovanović, T. Taguchi, J. Zhang and J.M. ter Vehn, *Phys. Rev. Lett.* **88** (2002) 055004. http://dx.doi.org/10.1103/PhysRevLett.88.055004

[29] A. Bourdier, D. Patin and E. Lefebvre, *Physica* **D206** (2005) 1. http://dx.doi.org/10.1016/j.physd.2005.04.017

[30] X. Davoine, E. Lefebvre, J. Faure, C. Rechatin, A. Lifschitz and V. Malka, *Phys. Plasmas* **15** (2008) 113102. http://dx.doi.org/10.1063/1.3008051

[31] C. Rechatin, J. Faure, A. Lifschitz, V. Malka and E. Lefebvre, *Phys. Plasmas* **14** (2007) 060702. http://dx.doi.org/10.1063/1.2741387

[32] O. Lundh, J. Lim, C. Rechatin, L. Ammoura, A. Ben-Ismail, X. Davoine, G. Gallot, J.-P. Goddet, E. Lefebvre, V. Malka *et al.*, *Nat. Phys.* **7** (2011) 219. http://dx.doi.org/10.1038/nphys1872

[33] J. Faure, C. Rechatin, A. Norlin, A. Lifschitz, Y. Glinec and V. Malka, *Nature* **444** (2006) 737. http://dx.doi.org/10.1038/nature05393

[34] J. Faure, C. Rechatin, A. Norlin, F. Burgy, A. Tafzi, J.-P. Rousseau and V. Malka, *Plasma Phys. Control. Fusion* **49** (2007) B395. http://dx.doi.org/10.1088/0741-3335/49/12B/S36

[35] C. Rechatin, J. Faure, A. Ben-Ismail, J. Lim, R. Fitour, A. Specka, H. Videau, A. Tafzi, F. Burgy and V. Malka, *Phys. Rev. Lett.* **102** (2009) 164801. http://dx.doi.org/10.1103/PhysRevLett.102.164801

[36] X. Davoine, E. Lefebvre, C. Rechatin, J. Faure and V. Malka, *Phys. Rev. Lett.* **102** (2009) 065001. http://dx.doi.org/10.1103/PhysRevLett.102.065001

[37] R. Lehe, A.F. Lifschitz, X. Davoine, C. Thaury and V. Malka, *Phys. Rev. Lett.* **111** (2013) 085005. http://dx.doi.org/10.1103/PhysRevLett.111.085005

[38] S. Bulanov, N. Naumova, F. Pegoraro and J. Sakai, *Phys. Rev.* **E58** (1998) R5257. http://dx.doi.org/10.1103/PhysRevE.58.R5257

[39] H. Suk, N. Barov, J.B. Rosenzweig and E. Esarey, *Phys. Rev. Lett.* **86** (2001) 1011. http://dx.doi.org/10.1103/PhysRevLett.86.1011

[40] A.V. Brantov, T.Z. Esirkepov, M. Kando, H. Kotaki, V.Y. Bychenkov and S.V. Bulanov, *Phys. Plasmas* **15** (2008) 073111. http://dx.doi.org/10.1063/1.2956989

[41] C.G.R. Geddes, K. Nakamura, G.R. Plateau, C. Tóth, E. Cormier-Michel, E. Esarey, C.B. Schroeder, J.R. Cary and W.P. Leemans, *Phys. Rev. Lett.* **100** (2008) 215004. http://dx.doi.org/10.1103/PhysRevLett.100.215004

[42] J. Faure, C. Rechatin, O. Lundh, L. Ammoura and V. Malka, *Phys. Plasmas* **17** (2010) 083107. http://dx.doi.org/10.1063/1.3469581

[43] K. Schmid, A. Buck, C.M.S. Sears, J.M. Mikhailova, R. Tautz, D. Herrmann, M. Geissler, F. Krausz and L. Veisz, *Phys. Rev. ST Accel. Beams* **13** (2010) 091301. http://dx.doi.org/10.1103/PhysRevSTAB.13.091301

[44] Z.-H. He, B. Hou, J.H. Easter, J. Faure, K. Krushelnick, J.A. Nees and A.G.R. Thomas, *New J. Phys.* **15** (2013) 053016. http://dx.doi.org/10.1088/1367-2630/15/5/053016

[45] A. Buck, J. Wenz, J. Xu, K. Khrennikov, K. Schmid, M. Heigoldt, J.M. Mikhailova, M. Geissler, B. Shen, F. Krausz *et al.*, *Phys. Rev. Lett.* **110** (2013) 185006. http://dx.doi.org/10.1103/PhysRevLett.110.185006

[46] P. Brijesh, C. Thaury, K.T. Phuoc, S. Corde, G. Lambert, V. Malka, S.P.D. Mangles, M. Bloom and S. Kneip, *Phys. Plasmas* **19** (2012) 063104. http://dx.doi.org/10.1063/1.4725421

[47] N. Bourgeois, J. Cowley and S.M. Hooker, *Phys. Rev. Lett.* **111** (2013) 155004. http://dx.doi.org/10.1103/PhysRevLett.111.155004

[48] L.-L. Yu, E. Esarey, C.B. Schroeder, J.-L. Vay, C. Benedetti, C.G.R. Geddes, M. Chen and W.P. Leemans, *Phys. Rev. Lett.* **112** (2014) 125001. http://dx.doi.org/10.1103/PhysRevLett.112.125001

[49] G. Fubiani, E. Esarey, C.B. Schroeder and W.P. Leemans, *Phys. Rev.* **E73** (2006) 026402. http://dx.doi.org/10.1103/PhysRevE.73.026402

Injection, Extraction and Matching

M. Ferrario
Frascati National Laboratory, National Institute for Nuclear Physics (INFN-LNF), Rome, Italy

Abstract

In this lecture we introduce from basic principles the main concepts of beam focusing and transport in modern accelerators using the beam envelope equation as a convenient mathematical tool. Matching conditions suitable for preserving beam quality are derived from the model for significant beam dynamics regimes. An extension of the model to the case of plasma accelerators is introduced. The understanding of similarities and differences with respect to traditional accelerators is also emphasized.

1   Introduction

Light sources based on high-gain free electron lasers or future high-energy linear colliders require the production, acceleration and transport up to the interaction point of low divergence, high-charge density electron bunches [1]. Many effects contribute in general to the degradation of the final beam quality, including chromatic effects, wake fields, emission of coherent radiation, and accelerator misalignments. Space charge effects and mismatch with the focusing and accelerating devices typically contribute to emittance degradation of high-charge density beams [2], hence the control of beam transport and acceleration is the leading edge for high-quality beam production.

In particular, further development of plasma-based accelerators requires careful phase space matching between plasma acceleration stages, and between plasma stages and traditional accelerator components. It represents a very critical issue and a fundamental challenge for high-quality beam production and its applications. Without proper matching, significant emittance growth may occur when the beam is propagating through different stages and components due to the large differences of transverse focusing strength. This unwanted effect is even more serious in the presence of finite energy spread.

In this paper we introduce from basic principles the main concepts of beam focusing and transport in modern accelerators using the beam envelope equation as a convenient mathematical tool. Matching conditions suitable for preserving beam quality are derived from the model for significant beam dynamics regimes. An extension of the model to the case of plasma accelerators is introduced. The understanding of similarities and differences with respect to traditional accelerators is also emphasized. A more detailed discussion of the previous topics can be found in the many classical textbooks on this subject, as listed in Refs. [3–6].

2   Laminar and non-laminar beams

An ideal high-charge particle beam has orbits that flow in layers that never intersect, as occurs in a laminar fluid. Such a beam is often called a laminar beam. More precisely, a laminar beam satisfies the following two conditions [6].

1   All particles at a given position have identical transverse velocities. On the contrary, the orbits of two particles that start at the same position could separate and later cross each other.

2   Assuming the beam propagates along the z axis, the magnitudes of the slopes of the trajectories in the transverse directions x and y, given by $x'(z) = \mathrm{d}x / \mathrm{d}z$ and $y'(z) = \mathrm{d}y / \mathrm{d}z$, are linearly proportional to the displacement from the z axis of beam propagation.

Trajectories of interest in beam physics are always confined to the inside of small, near-axis regions, and the transverse momentum is much smaller than the longitudinal momentum, $p_{x,y} << p_z \approx p$. As a consequence, it is convenient in most cases to use the small angle, or *paraxial*, approximation, which allows us to write the useful approximate expressions $x' = p_x/p_z \approx p_x/p$ and $y' = p_y/p_z \approx p_x/p$.

To help understand the features and the advantages of a laminar beam propagation, the following figures compare the typical behaviour of a laminar and of a non-laminar (or thermal) beam.

Figure 1 illustrates an example of orbit evolution of a laminar mono-energetic beam with half width x_0 along a simple beam line with an ideal focusing element (solenoid, magnetic quadrupoles, or electrostatic transverse fields are usually adopted to this end), represented by a thin lens located at the longitudinal coordinate $z = 0$. In an ideal lens, focusing (defocusing) forces are linearly proportional to the displacement from the symmetry axis z so that the lens maintains the laminar flow of the beam.
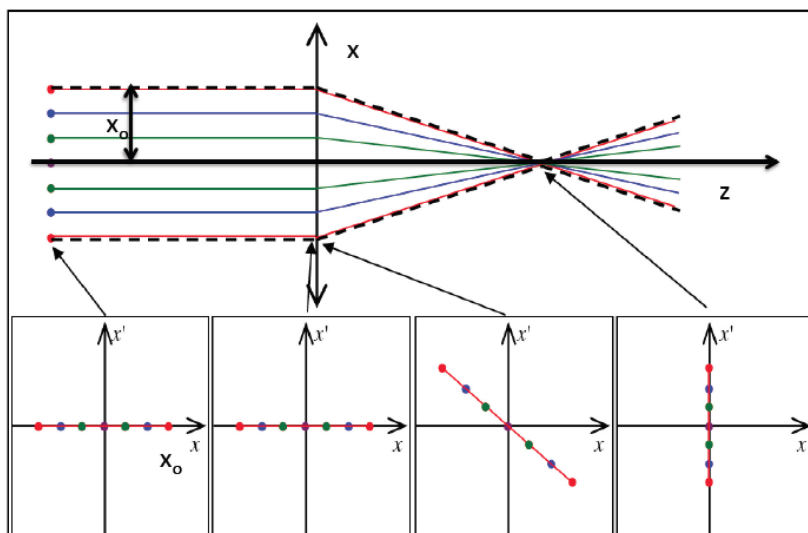
Fig. 1: Particle trajectories and phase space evolution of a laminar beam [7]

The beam shown in Fig. 1 starts propagating completely parallel to the symmetry axis z; in this particular case the particles all have zero transverse velocity. There are no orbits that cross each other in such a beam. Ignoring collisions and inner forces, like coulomb forces, such a parallel beam could propagate an infinite distance with no change in its transverse width. When the beam crosses the ideal lens it is transformed in a converging laminar beam. Because the transverse velocities after the linear lens are proportional to the displacement off-axis, particle orbits define similar triangles that converge to a single point. After passing through the singularity at the focal point, the particles follow diverging orbits. We can always transform a diverging (or converging) beam to a parallel beam by using a lens of the proper focal length, as can be seen by reversing the propagation axis of Fig. 1.

The small boxes in the lower part of the figure depict the particle distributions in the trace space (x,x'), equivalent to the canonical phase space $(x,p_x \approx x'p)$ when p is constant, i.e. without beam acceleration. The phase space area occupied by an ideal laminar beam is a straight segment of zero thickness. As can be easily verified, the condition that the particle distribution has zero thickness proceeds from condition 1; the segment straightness is a consequence of condition 2. The distribution of a laminar beam propagating through a transport system with ideal linear focusing elements is thus a straight segment with variable slope.

Particles in a non-laminar beam have a random distribution of transverse velocities at the same location and a spread in directions, as shown in Fig. 2. Because of the disorder of a non-laminar beam, it is impossible to focus all particles from a location in the beam toward a common point. Lenses can influence only the average motion of particles. Focal spot limitations are a major concern for a wide variety of applications, from electron microscopy to free electron lasers and linear colliders. The phase space plot of a non-laminar beam is no longer a straight line: the beam, as shown in the lower boxes of Fig. 2, occupies a wider area of the phase space.
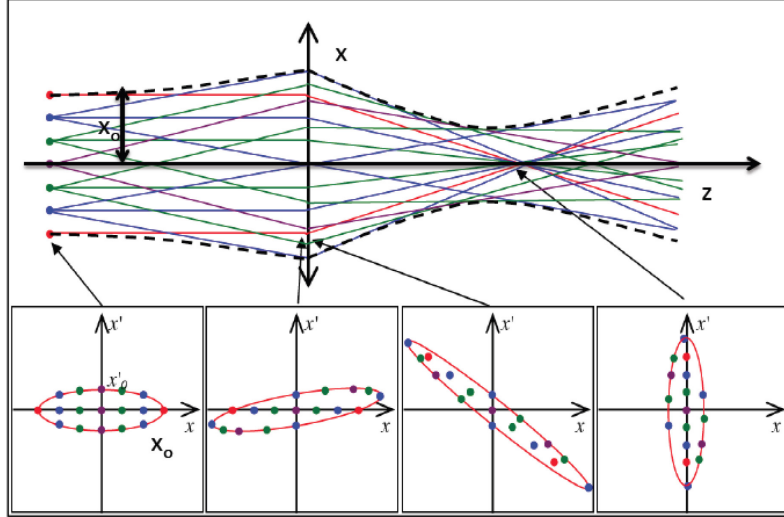
Fig. 2: Particle trajectories and phase space evolution of a non-laminar beam [7]

3   The emittance concept

The phase space surface A occupied by a beam is a convenient figure of merit for designating the quality of a beam. This quantity is the emittance ε_x and is usually represented by an ellipse that contains the whole particle distribution in the phase space (x,x'), such that $A = \pi \varepsilon_x$. An analogous definition holds for the (y,y') and (z,z') planes. The original choice of an elliptical shape comes from the fact that when linear focusing forces are applied to a beam, the trajectory of each particle in phase space lies on an ellipse, which may be called the trajectory ellipse. Being the area of the phase space, the emittance is measured in [mm mrad] or more often in [μm].

The ellipse equation is written as

$$\gamma_x x^2 + 2\alpha_x xx' + \beta_x x'^2 = \varepsilon_x \tag{1}$$

where x and x' are the particle coordinates in the phase space and the coefficients $\alpha_x(z)$, $\beta_x(z)$, $\gamma_x(z)$ are called Twiss parameters, which are related by the geometrical condition:

$$\beta_x \gamma_x - \alpha_x^2 = 1 \tag{2}$$

As shown in Fig. 3 the beam envelope boundary X_{max}, its derivative $(X_{max})'$ and the maximum beam divergency X'_{max}, i.e. the projection on the axis x and x' of the ellipse edges, can be expressed as a function of the ellipse parameters:

$$\begin{cases} X_{max} = \sqrt{\beta_x \varepsilon_x} \\ \left(X_{max}\right)' = -\alpha\sqrt{\dfrac{\varepsilon}{\beta}} \\ X'_{max} = \sqrt{\gamma_x \varepsilon_x} \end{cases} \tag{3}$$
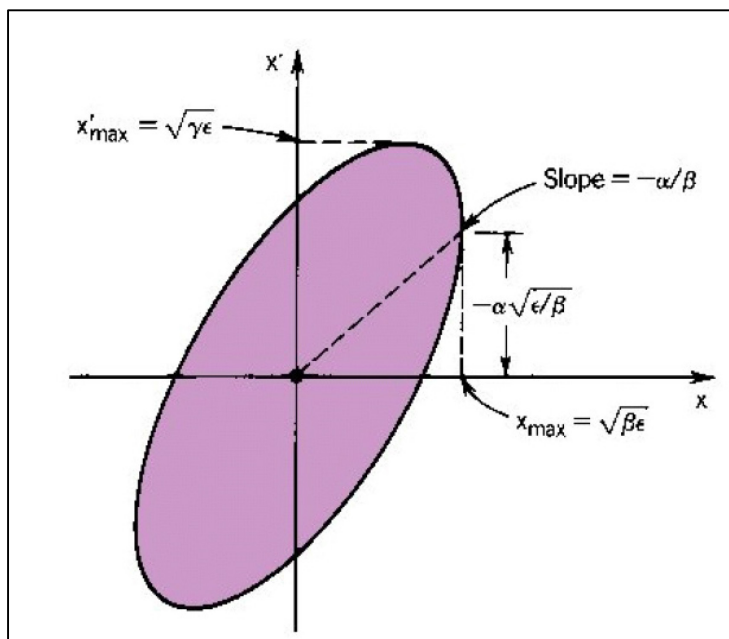
Fig. 3: Phase space distribution in a skewed elliptical boundary showing the relationship of Twiss parameters to the ellipse geometry [6].

According to Liouville's theorem the 6D (x,p_x,y,p_y,z,p_z) phase space volume occupied by a beam is constant, provided that there are no dissipative forces, no particles lost or created, and no coulomb scattering among particles. Moreover, if the forces in the three orthogonal directions are uncoupled, Liouville's theorem also holds for each reduced phase space (x,p_x), (y,p_y), (z,p_z) surfaces and hence emittance also remains constant in each plane [3].

Although the net phase space surface occupied by a beam is constant, nonlinear field components can stretch and distort the particle distribution in the phase space, and the beam will lose its laminar behaviour. A realistic phase space distribution is often very different from a regular ellipse, as shown in Fig. 4.
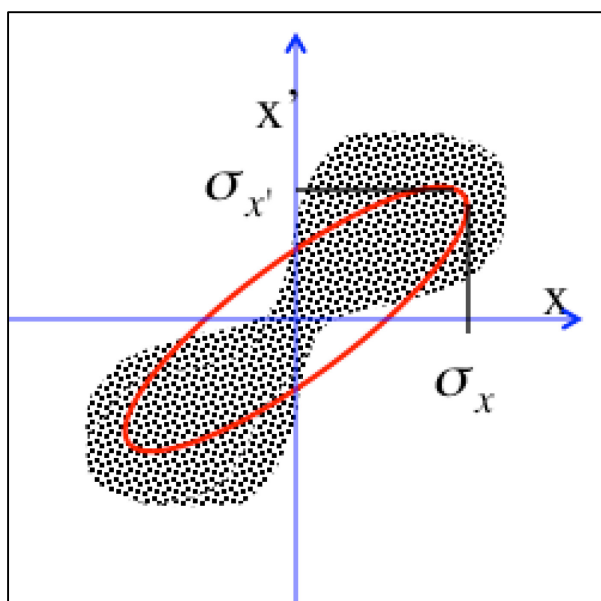
Fig. 4: Typical evolution of phase space distribution (black dots) under the effects of non-linear forces with the equivalent ellipse superimposed (red line).

We introduce, therefore, a definition of emittance that measures the beam quality rather than the phase space area. It is often more convenient to associate to a generic distribution function $f(x,x',z)$ in the phase space a statistical definition of emittance, the so-called *r.m.s. emittance*:

$$\gamma_x x^2 + 2\alpha_x xx' + \beta_x x'^2 = \varepsilon_{x,\mathrm{rms}}$$

(4)

such that the equivalent-ellipse projections on the x and x' axes are equal to the r.m.s. values of the distribution, implying the following conditions:

$$\begin{cases} \sigma_x = \sqrt{\beta_x \varepsilon_{x,\mathrm{rms}}} \\ \sigma_{x'} = \sqrt{\gamma_x \varepsilon_{x,\mathrm{rms}}} \end{cases}$$

(5)

where

$$\begin{cases} \sigma_x^2(z) = \langle x^2 \rangle = \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} x^2 f(x,x',z)\,\mathrm{d}x\mathrm{d}x' \\ \sigma_{x'}^2(z) = \langle x'^2 \rangle = \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} x'^2 f(x,x',z)\,\mathrm{d}x\mathrm{d}x' \end{cases}$$

(6)

are the second moments of the distribution function $f(x,x',z)$. Another important quantity that accounts for the degree of (x,x') correlations is defined as

$$\sigma_{xx'}(z) = \langle xx' \rangle = \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} xx' f(x,x',z)\,\mathrm{d}x\mathrm{d}x' \,.$$

(7)

From Eq. (3) it holds also $\sigma_x' = \dfrac{\sigma_{xx'}}{\sigma_x} = -\alpha_x \sqrt{\dfrac{\varepsilon_{x,\mathrm{rms}}}{\beta_x}}$, see also Eq. (16), which allows us to link the correlation moment Eq. (7) to the Twiss parameter as

$$\sigma_{xx'} = -\alpha_x \varepsilon_{x,\mathrm{rms}}$$

.

(8)

One can easily see from Eq. (3) and Eq. (5) that $\alpha_x = -\dfrac{1}{2}\dfrac{\mathrm{d}\beta_x}{\mathrm{d}z}$ also holds.

By substituting the Twiss parameter defined by Eq. (5) and Eq. (8) into condition 2 we obtain [5]

$$\frac{\sigma_{x'}^2}{\varepsilon_{x,\mathrm{rms}}}\frac{\sigma_x^2}{\varepsilon_{x,\mathrm{rms}}} - \left(\frac{\sigma_{xx'}}{\varepsilon_{x,\mathrm{rms}}}\right) = 1$$

.

(9)

Reordering the terms in Eq. (8) we end up with the definition of *r.m.s. emittance* in terms of the second moments of the distribution:

$$\varepsilon_{\mathrm{rms}} = \sqrt{\sigma_x^2 \sigma_{x'}^2 - \sigma_{xx'}^2} = \sqrt{\left(\langle x^2 \rangle \langle x'^2 \rangle - \langle xx' \rangle^2\right)}$$

(10)

where we omit, from now on, the subscribed x in the emittance notation: $\varepsilon_{\mathrm{rms}} = \varepsilon_{x,\mathrm{rms}}$. Root mean square emittance tells us some important information about phase space distributions under the effect of linear or non-linear forces acting on the beam. Consider, for example, an idealized particle distribution in phase space that lies on some line that passes through the origin as illustrated in Fig. 5.

(a)   (b)
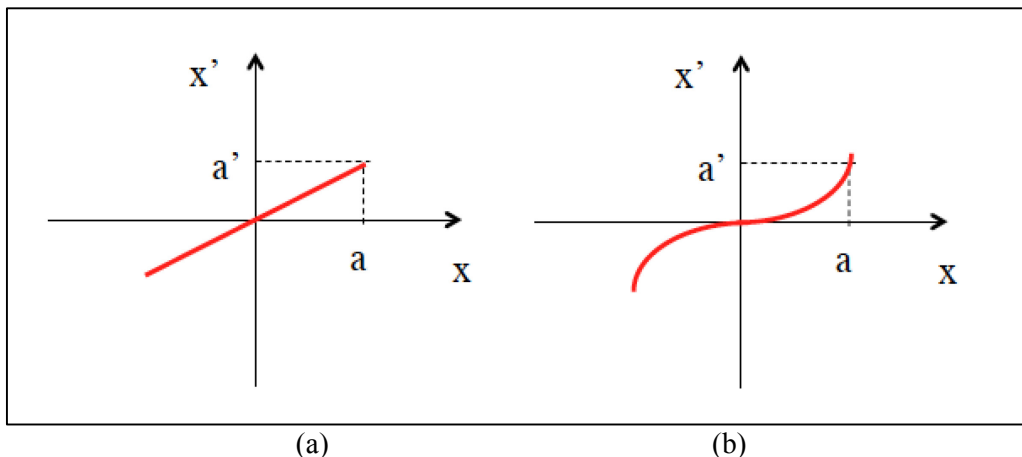
Fig. 5: Phase space distributions under the effect of (a) linear or (b) non-linear forces acting on the beam

Assuming a generic correlation of the type $x' = Cx^n$ and computing the r.m.s. emittance according to Eq. (10) we have

$$\varepsilon_{\text{rms}}^2 = C\sqrt{\langle x^2\rangle\langle x^{2n}\rangle - \langle x^{n+1}\rangle^2} \quad \begin{cases} n=1 \Rightarrow \varepsilon_{\text{rms}} = 0 \\ n>1 \Rightarrow \varepsilon_{\text{rms}} \neq 0 \end{cases}. \tag{11}$$

When $n = 1$ the line is straight and the r.m.s. emittance is $\varepsilon_{\text{rms}} = 0$. When $n > 1$ the relationship is nonlinear, the line in phase space is curved, and the r.m.s. emittance is in general not zero. Both distributions have zero area. Therefore, we conclude that even when the phase space area is zero, if the distribution is lying on a curved line its r.m.s. emittance is not zero. The r.m.s. emittance depends not only on the area occupied by the beam in phase space, but also on distortions produced by non-linear forces.

If the beam is subject to acceleration it is more convenient to use the r.m.s. normalized emittance, for which the transverse momentum $p_x = p_z x' = m_o c \beta \gamma x'$ is used instead of the divergence:

$$\varepsilon_{n,\text{rms}} = \frac{1}{m_0 c}\sqrt{\sigma_x^2 \sigma_{p_x}^2 - \sigma_{xp_x}^2} = \frac{1}{m_0 c}\sqrt{\left(\langle x^2\rangle\langle p_x^2\rangle - \langle xp_x\rangle^2\right)} = \sqrt{\left(\langle x^2\rangle\langle (\beta\gamma x')^2\rangle - \langle x\beta\gamma x'\rangle^2\right)}. \tag{12}$$

The reason for introducing a normalized emittance is that the divergences of the particles $x' = p_x/p$ are reduced during acceleration as p increases. Thus, acceleration reduces the un-normalized emittance, but does not affect the normalized emittance. Assuming a small energy spread within the beam, the normalized and un-normalized emittances can be related by the approximated relation $\langle \beta\gamma\rangle\varepsilon_{\text{rms}}$. This approximation, which is often used in conventional accelerators, may be strongly misleading when adopted for describing beams with significant energy spread, like those presently produced by plasma accelerators. A more careful analysis is reported below [8].

When the correlations between the energy and transverse positions are negligible (as in a drift without collective effects), Eq. (12) can be written as

$$\varepsilon_{n,\text{rms}}^2 = \langle \beta^2\gamma^2\rangle\langle x^2\rangle\langle x'^2\rangle - \langle \beta\gamma\rangle^2\langle xx'\rangle^2 . \tag{13}$$

Consider now the definition of relative energy spread $\sigma_\gamma^2 = \dfrac{\langle \beta^2\gamma^2\rangle - \langle \beta\gamma\rangle^2}{\langle \beta\gamma\rangle^2}$, which can be inserted into Eq. (13) to give

$$\varepsilon_{n,\text{rms}}^2 = \langle \beta^2\gamma^2\rangle\sigma_\gamma^2\langle x^2\rangle\langle x'^2\rangle + \langle \beta\gamma\rangle^2\left(\langle x^2\rangle\langle x'^2\rangle - \langle xx'\rangle^2\right). \tag{14}$$

Assuming relativistic particles ($\beta = 1$) we get

$$\varepsilon_{n,\mathrm{rms}}^2 = \left\langle \gamma^2 \right\rangle \left(\sigma_\gamma^2 \sigma_x^2 \sigma_{x'}^2 + \varepsilon_{\mathrm{rms}}^2 \right).$$

(15)

If the first term in the parentheses is negligible, we find the conventional approximation of normalized emittance, as $\left\langle \gamma \right\rangle \varepsilon_{\mathrm{rms}}$. For a conventional accelerator this might generally be the case. Considering, for example, beam parameters for the SPARC_LAB photoinjector [9], at 5 MeV the ratio between the first and the second term is $\sim 10^{-3}$; while at 150 MeV it is $\sim 10^{-5}$. On the other hand, using typical beam parameters at the plasma–vacuum interface, the first term is of the same order of magnitude as for conventional accelerators at low energies; however, due to the rapid increase of the bunch size outside the plasma ($\sigma_{x'} \sim$ mrad) and the large energy spread ($\sigma_\gamma > 1\%$), it becomes predominant compared to the second term after a drift of a few millimetres. *Therefore the use of approximated formulas when measuring the normalized emittance of plasma accelerated particle beams is very inappropriate* [10].

4   The root mean square envelope equation

We are now interested in following the evolution of particle distribution during beam transport and acceleration. One can take profit of the first collective variable defined in Eq. (6), the second moment of the distribution termed r.m.s. beam envelope, to derive a differential equation suitable for describing r.m.s. beam envelope dynamics [11]. To this end let us compute the first and second derivative of σ_x [4]:

$$\frac{\mathrm{d}\sigma_x}{\mathrm{d}z} = \frac{\mathrm{d}}{\mathrm{d}z}\sqrt{\left\langle x^2 \right\rangle} = \frac{1}{2\sigma_x}\frac{\mathrm{d}}{\mathrm{d}z}\left\langle x^2 \right\rangle = \frac{1}{2\sigma_x}2\left\langle xx' \right\rangle = \frac{\sigma_{xx'}}{\sigma_x}$$

$$\frac{\mathrm{d}^2\sigma_x}{\mathrm{d}z^2} = \frac{\mathrm{d}}{\mathrm{d}z}\frac{\sigma_{xx'}}{\sigma_x} = \frac{1}{\sigma_x}\frac{\mathrm{d}\sigma_{xx'}}{\mathrm{d}z} - \frac{\sigma_{xx'}^2}{\sigma_x^3} = \frac{1}{\sigma_x}\left(\left\langle x'^2 \right\rangle + \left\langle xx'' \right\rangle\right) - \frac{\sigma_{xx'}^2}{\sigma_x^3} = \frac{\sigma_{x'}^2 + \left\langle xx'' \right\rangle}{\sigma_x} - \frac{\sigma_{xx'}^2}{\sigma_x^3}$$

(16)

Rearranging the second derivative Eq. (16) we obtain a second-order nonlinear differential equation for the beam envelope evolution,

$$\sigma_x'' = \frac{\sigma_x^2 \sigma_{x'}^2 - \sigma_{xx'}^2}{\sigma_x^3} + \frac{\left\langle xx'' \right\rangle}{\sigma_x}$$

(17)

or, in a more convenient form using the r.m.s. emittance definition Eq. (10),

$$\sigma_x'' - \frac{1}{\sigma_x}\left\langle xx'' \right\rangle = \frac{\varepsilon_{\mathrm{rms}}^2}{\sigma_x^3} \ .$$

(18)

In Eq. (18) the emittance term can be interpreted physically as an outward pressure on the beam envelope produced by the r.m.s. spread in trajectory angle, which is parameterized by the r.m.s. emittance.

Let's now consider, for example, the simple case with $\left\langle xx'' \right\rangle = 0$, describing a beam drifting in free space. The envelope equation reduces to

$$\sigma_x^3 \sigma_x'' = \varepsilon_{\mathrm{rms}}^2 \ .$$

(19)

With initial conditions σ_0, σ'_0 at z_0, depending on the upstream transport channel, Eq. (19) has a hyperbolic solution:

$$\sigma(z) = \sqrt{\left(\sigma_0 + \sigma'_0\left(z - z_0\right)\right)^2 + \frac{\varepsilon_{\mathrm{rms}}^2}{\sigma_0^2}\left(z - z_0\right)^2} \ .$$

(20)

Considering the case of a beam at waist ($\langle xx' \rangle = 0$) with $\sigma'_0 = 0$, using Eq. (5), the solution Eq. (20) is often written in terms of the β function as

$$\sigma(z) = \sigma_0 \sqrt{1 + \left(\frac{z - z_0}{\beta_{\mathrm{w}}} \right)^2} \tag{21}$$

This relation indicates that without any external focusing element the beam envelope increases from the beam waist by a factor $\sqrt{2}$ with a characteristic length $\beta_w = \frac{\sigma_0^2}{\epsilon_{rms}}$ as shown in Fig. 6.
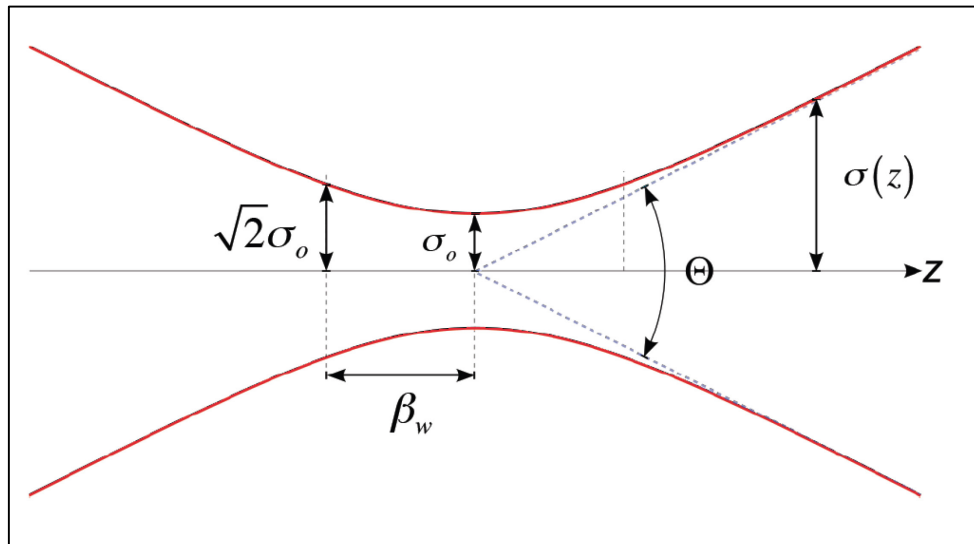
Fig. 6: Schematic representation of the beam envelope behaviour near the beam waist

At waist the relation $\varepsilon^2_{\mathrm{rms}} = \sigma^2_{0,x} \sigma^2_{0,x'}$ also holds, which can be inserted into Eq. (20) to get $\sigma^2_x(z) = \sigma^2_{0x'}(z - z_0)^2$. Under this condition Eq. (15) can be written as follows:

$$\varepsilon^2_{n,\mathrm{rms}}(z) = \langle \gamma^2 \rangle \left(\sigma^2_\gamma \sigma^4_{x'} (z - z_0)^2 + \varepsilon^2_{\mathrm{rms}} \right)$$

showing that beams with large energy spread and divergence undergo a significant normalized emittance growth even in a drift of length $(z - z_0)$ [8, 12].

Notice also that the solution Eq. (21) is exactly analogous to that of a Gaussian light beam for which the beam width $w = 2\sigma_{ph}$ increases away from its minimum value at the waist w_0 with characteristic length $Z_R = \frac{\pi w_0^2}{\lambda}$ (Rayleigh length) [4]. This analogy suggests that we can identify an effective emittance of a photon beam as $\varepsilon_{\mathrm{ph}} = \frac{\lambda}{4\pi}$.

For the effective transport of a beam with finite emittance it is mandatory to make use of some external force providing beam confinement in the transport or accelerating line. The term $\langle xx'' \rangle$ accounts for external forces when we know x'' given by the single particle equation of motion:

$$\frac{\mathrm{d}p_x}{\mathrm{d}t} = F_x. \tag{22}$$

Under the paraxial approximation $p_x \ll p = \beta\gamma mc$ the transverse momentum p_x can be written as $p_x = px' = \beta\gamma m_0 cx'$, so that

$$\frac{\mathrm{d}p_x}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(px'\right) = \beta c \frac{\mathrm{d}}{\mathrm{d}z}\left(px'\right) = F_x \tag{23}$$

and the transverse acceleration results to be:

$$x'' = -\frac{p'}{p}x' + \frac{F_x}{\beta cp}. \tag{24}$$

It follows that

$$\left\langle xx'' \right\rangle = -\frac{p'}{p}\left\langle xx' \right\rangle + \frac{\left\langle xF_x \right\rangle}{\beta cp} = \frac{p'}{p}\sigma_{xx'} + \frac{\left\langle xF_x \right\rangle}{\beta cp}. \tag{25}$$

Inserting Eq. (25) into Eq. (18) and recalling Eq. (16), $\sigma_x' = \dfrac{\sigma_{xx'}}{\sigma_x}$, the complete r.m.s. envelope equation is

$$\sigma_x'' + \frac{p'}{p}\sigma_x' - \frac{1}{\sigma_x}\frac{\left\langle xF_x \right\rangle}{\beta cp} = \frac{\varepsilon_{n,\mathrm{rms}}^2}{\gamma^2 \sigma_x^3} \tag{26}$$

where we have included the normalized emittance $\varepsilon_{n,\mathrm{rms}} = \gamma\varepsilon_{\mathrm{rms}}$. Notice that the effect of longitudinal accelerations appears in the r.m.s. envelope equation as an oscillation damping term, called 'adiabatic damping', proportional to p'/p. The term $\left\langle xF_x \right\rangle$ represents the moment of any external transverse force acting on the beam, such as that produced by a focusing magnetic channel.

5   External forces

Let's now consider the case of an external linear force acting on the beam in the form $F_x = \pm kx$. It can be focusing or defocusing according to the sign. The moment of the force results as

$$\left\langle xF_x \right\rangle = \pm k \left\langle x^2 \right\rangle = \pm k\sigma_x^2 \tag{27}$$

and the envelope equation becomes

$$\sigma_x'' + \frac{\gamma'}{\gamma}\sigma_x' \mp k_{ext}^2 \sigma_x = \frac{\varepsilon_{n,rms}^2}{\gamma^2 \sigma_x^3} \tag{28}$$

where we have explicitly used the momentum definition $p = \gamma mc$ for a relativistic particle with $\beta \approx 1$ and defined the wavenumber $k_{ext}^2 = \dfrac{k}{\gamma m_0 c^2}$.

Typical focusing elements are quadrupoles and solenoids [3]. The magnetic quadrupole field is given in Cartesian coordinates by

$$\begin{cases} B_x = B_0 \dfrac{y}{\mathrm{d}} = B_0' y \\[2mm] B_y = B_0 \dfrac{x}{\mathrm{d}} = B_0' x \end{cases} \tag{29}$$

where d is the pole distance and B'_0 the field gradient. The force acting on the beam is $\vec{F}_\perp = qv_z B'_0 \left(y\hat{j} - x\hat{i} \right)$ that, when B_0 is positive, is focusing in the x direction and defocusing in y. The focusing strength is $k_{\text{quad}} = \dfrac{qB'_0}{\gamma m_0 c} = k_{\text{ext}}^2$.

In a solenoid the focusing strength is given by $k_{\text{sol}} = \left(\dfrac{qB_0}{2\gamma m_0 c} \right)^2 = k_{\text{ext}}^2$. Notice that the solenoid is always focusing in both directions, an important properties when the cylindrical symmetry of the beam must be preserved. On the other hand, being a second-order quantity in γ it is more effective at low energy.

It is interesting to consider the case of a uniform focusing channel without acceleration described by the r.m.s. envelope equation

$$\sigma''_x + k_{\text{ext}}^2 \sigma_x = \frac{\varepsilon_{\text{rms}}^2}{\sigma_x^3} \tag{30}$$

By substituting $\sigma_x = \sqrt{\beta_x \varepsilon_{\text{rms}}}$ in Eq. (30) one obtains an equation for the 'betatron function' $\beta_x(z)$ that is independent of the emittance term:

$$\beta''_x + 2k_{\text{ext}}^2 \beta_x = \frac{2}{\beta_x} + \frac{\beta_x'^2}{2\beta_x}. \tag{31}$$

Equation (31) contains just the transport channel focusing strength and, being independent of the beam parameters, suggests that the meaning of the betatron function is to describe the transport line characteristic by itself. The betatron function reflects exterior forces from focusing magnets, and is highly dependent on the particular arrangement of the quadrupole magnets. The equilibrium, or matched, solution of Eq. (31) is given by $\beta_{\text{eq}} = \dfrac{1}{k_{\text{ext}}} = \dfrac{\lambda_\beta}{2\pi}$, as can be easily verified. This result shows that the matched β_x function is simply the inverse of the focusing wave number, or equivalently is proportional to the 'betatron wavelength' λ_β.

6   Space charge forces

Another important force acting on the beam is the one produced by the beam itself due to the internal coulomb forces. The net effect of the coulomb interaction in a multi-particle system can be classified into two regimes [3]:

–   *collisional regime*, dominated by binary collisions caused by close particle encounters;

–   *collective regime* or *space charge regime*, dominated by the self-field produced by the particles' distribution that varies appreciably only over large distances compared to the average separation of the particles.

A measure for the relative importance of collisional versus collective effects in a beam with particle density n is the relativistic *Debye length*,

$$\lambda_{\text{D}} = \sqrt{\frac{\varepsilon_0 \gamma^2 k_{\text{B}} T_{\text{b}}}{e^2 n}} \tag{32}$$

where the transverse beam temperature T_b is defined as $k_B T_b = \gamma m_0 \langle v_\perp^2 \rangle$, and k_B is the Boltzmann constant. As long as the Debye length remains small compared to the particle bunch transverse size the beam is in the space-charge dominated regime and is not sensitive to binary collisions. Smooth functions for the charge and field distributions can be used in this case, and the space charge force can be treated like an external applied force. The space charge field can be separated into linear and nonlinear terms as a function of displacement from the beam axis. The linear space charge term defocuses the beam and leads to an increase in beam size. The nonlinear space charge terms also increase the r.m.s. emittance by distorting the phase-space distribution. Under the paraxial approximation of particle motion we can consider the linear component alone. We shall see below that the linear component of the space charge field can also induce emittance growth when correlation along the bunch are taken into account.

For a bunched beam of uniform charge distribution in a cylinder of radius R and length L, carrying a current $\hat{I}$ and moving with longitudinal velocity $v_z = \beta c$, the linear component of the longitudinal and transverse space charge field are given approximately by [13]

$$E_z(\zeta) = \frac{\hat{I}L}{2\pi\varepsilon_0 R^2 \beta c} h(\zeta),$$ (33)

$$E_r(r,\zeta) = \frac{\hat{I}r}{2\pi\varepsilon_0 R^2 \beta c} g(\zeta) .$$ (34)

The field form factor is described by the functions

$$h(\zeta) = \sqrt{A + (1-\zeta)^2} - \sqrt{A + \zeta^2 + (2\zeta - 1)} ,$$ (35)

$$g(\zeta) = \frac{(1-\zeta)}{2\sqrt{A^2 + (1-\zeta)^2}} + \frac{\zeta}{2\sqrt{A^2 + \zeta}}$$ (36)

where $\zeta = z/L$ is the normalized longitudinal coordinate along the bunch and $A = R/\gamma L$ is the beam aspect ratio. The field form factors account for the variation of the fields along the bunch. As γ increases, $g(\zeta) \to 1$ and $h(\zeta) \to 0$, thus showing that space charge fields mainly affect transverse beam dynamics. It shows also that an energy increase corresponds to a bunch lengthening in the moving frame $L' = \gamma L$, leading to a vanishing longitudinal field component, as in the case of a continuous beam in the laboratory frame.

To evaluate the force acting on the beam one must also account for the azimuthal magnetic field associated with the beam current that in cylindrical symmetry is given by $B_\vartheta = \dfrac{\beta}{c} E_r$. Thus, the Lorentz force acting on each single particle is given by

$$F_r = e(E_r - \beta c B_\vartheta) = e(1-\beta^2)E_r = \frac{eE_r}{\gamma^2} .$$ (37)

The attractive magnetic force, which becomes significant at high velocities, tends to compensate for the repulsive electric force. Therefore space charge defocusing is primarily a non-relativistic effect and decreases as γ^{-2}.

In order to include space charge forces in the envelope equation let us start writing the space charge forces produced by the previous fields in Cartesian coordinates:

$$F_x = \frac{e\hat{I}x}{2\pi\gamma^2\varepsilon_0\sigma_x^2 \beta c} g(\zeta).$$ (38)

Then, computing the moment of the force we need

$$x'' = \frac{F_x}{\beta cp} = \frac{eIx}{2\pi\varepsilon_0\gamma^3 m_0\beta^3 c^3\sigma_x^2} = \frac{k_{sc}(\zeta)}{(\beta\gamma)^3\sigma_x^2} \tag{39}$$

where we have introduced the generalized beam perveance

$$k_{sc}(\zeta) = \frac{2\hat{I}}{I_A}g(\zeta) \tag{40}$$

normalized to the Alfven current $I_A = 4\pi\varepsilon_0 m_0 c^3/e = 17$ kA for electrons. Notice that in this case the perveance in Eq. (40) explicitly depends on the slice coordinate ζ. We can now calculate the term that enters in the envelope equation for a relativistic beam,

$$\langle xx''\rangle = \frac{k_{sc}}{\gamma^3\sigma_x^2}\langle x^2\rangle = \frac{k_{sc}}{\gamma^3}, \tag{41}$$

leading to the complete envelope equation

$$\sigma_x'' + \frac{\gamma'}{\gamma}\sigma_x' + k_{ext}^2\sigma_x = \frac{\varepsilon_{n,rms}^2}{\gamma^2\sigma_x^3} + \frac{k_{sc}}{\gamma^3\sigma_x}. \tag{42}$$

From the envelope equation Eq. (42) we can identify two regimes of beam propagation: *space-charge dominated* and *emittance dominated*. A beam is space-charge dominated as long as the space charge collective forces are largely dominant over the emittance pressure. In this regime the linear component of the space-charge force produces a quasi-laminar propagation of the beam, as one can see by integrating one time Eq. (39) under the paraxial ray approximation $x' \ll 1$. A measure of the relative importance of space-charge effects versus emittance pressure is given by the *laminarity parameter*, defined as the ratio between the space-charge term and the emittance term:

$$\rho = \frac{\hat{I}}{2I_A\gamma}\frac{\sigma^2}{\varepsilon_n^2}. \tag{43}$$

When ρ greatly exceeds unity, the beam behaves like a laminar flow (all beam particles move on trajectories that do not cross), and transport and acceleration require a careful tuning of focusing and accelerating elements in order to keep laminarity. Correlated emittance growth is typical in this regime, which can be made reversible if proper beam matching conditions are fulfilled, as discussed below. When $\rho < 1$ the beam is emittance dominated (thermal regime) and the space charge effects can be neglected. The transition to the thermal regime occurs when $\rho \approx 1$ corresponding to the transition energy

$$\gamma_{tr} = \frac{\hat{I}}{2I_A}\frac{\sigma^2}{\varepsilon_n^2}. \tag{44}$$

For example a beam with $\hat{I} = 100$ A $\varepsilon_n = 1$ μm and $\sigma = 300$ μm is leaving the space charge dominated regime and is entering the thermal regime at the transition energy of 131 MeV. From this example one may conclude that the space charge dominated regime is typical of low energy beams. Actually, for applications like linac-driven free electron lasers, peak current exceeding kilo amperes are required. Space charge effects may recur if bunch compressors are active at higher energies and a new energy threshold with higher $\hat{I}$ has to be considered.

7   Correlated emittance oscillations

When longitudinal correlations within the bunch are important, like that induced by space charge effects, beam envelope evolution is generally dependent also on the bunch coordinate ζ. In this case the bunch

should be considered as an ensemble of n longitudinal slices of envelope $\sigma_s(z,\zeta)$, whose evolution can be computed from n slice envelope equations equivalent to Eq. (42) provided that the bunch parameters refer to each single slice: γ_s, γ'_s, $k_{sc,s} = k_{sc}g(\zeta)$. Correlations within the bunch may cause emittance oscillations that can be evaluated, once an analytical or numerical solution [13] of the slice envelope equation is known, by using the following correlated emittance definition:

$$\varepsilon_{rms,cor} = \sqrt{\langle\sigma_s^2\rangle\langle\sigma_s'^2\rangle - \langle\sigma_s\sigma_s'\rangle^2}$$

(45)

where the average is performed over the entire slice ensemble. In the simplest case of a two-slices model the previous definition reduces to

$$\varepsilon_{rms,cor} = |\sigma_1\sigma_2' - \sigma_2\sigma_1'| \ ,$$

(46)

which represents a simple and useful formula for an estimation of the emittance scaling [14].

The total normalized r.m.s. emittance is the given by the superposition of the correlated and uncorrelated terms as

$$\varepsilon_{rms,cor} = \langle\gamma\rangle\sqrt{\varepsilon_{rms}^2 + \varepsilon_{rms,cor}^2} \ .$$

(47)

An interesting example to consider here, showing the consequences of non-perfect beam matching, is the propagation of a beam in the space-charge dominated regime nearly matched to an external focusing channel ($k_{ext} = k_{sol}$), as illustrated in Fig. 7. To simplify our computations we can neglect acceleration, as in the case of a simple beam transport line. The envelope equation for each slice, indicated as σ_s, reduces to

$$\sigma_s'' + k_{ext}^2\sigma_s = \frac{k_{sc,s}}{\gamma^3\sigma_s} \ .$$

(48)
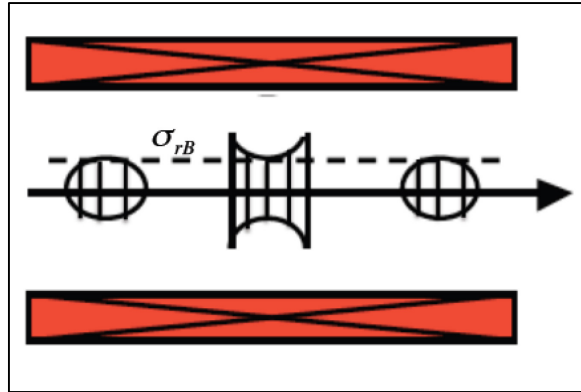
Fig. 7: Schematic representation of a nearly matched beam in a long solenoid. The dashed line represent the reference slice envelope fully matched to the Brillouin flow condition. The other slice envelopes are oscillating around the equilibrium solution.

A stationary solution, called the *Brillouin flow*, is given by

$$\sigma_{s,B} = \frac{1}{k_{ext}^2}\sqrt{\frac{\hat{I}g(\zeta)}{2\gamma^3 I_A}}$$

(49)

where the local dependence of the current $\hat{I}_s = \hat{I}g(\zeta)$ within the bunch has been explicitly indicated. This solution represent the matching conditions for which the external focusing completely balances the internal space-charge force. Unfortunately, since k_{ext} has a slice-independent constant value, the Brillouin matching condition cannot be achieved at the same time for all of the bunch slices. Assuming

that there is a reference slice perfectly matched with an envelope $\sigma_{r,B}$, the matching condition for the other slices can be written as

$$\sigma_{sB} = \sigma_{rB} + \frac{\sigma_{rB}}{2}\left(\frac{\delta I_s}{\hat{I}}\right)$$

(50)

with respect to the reference slice. Considering a small perturbation δ_s from the equilibrium in the form

$$\sigma_s = \sigma_{s,B} + \delta_s$$

(51)

and substituting into Eq. (48) we can obtain a linearized equation for the slice offset

$$\delta_s'' + 2k_{ext}^2\delta_s = 0$$

(52)

which has a solution given by

$$\delta_s = \delta_0 \cos\left(\sqrt{2}k_{ext}z\right)$$

(53)

where $\delta_0 = \sigma - \sigma_{sB}$ is the amplitude of the initial slice mismatch, which we assume for convenience is the same for all slices. Inserting Eq. (53) into Eq. (51) we get the perturbed solution:

$$\sigma_s = \sigma_{s,B} + \delta_0 \cos\left(\sqrt{2}k_{ext}z\right) \ .$$

(54)

Equation (54) shows that slice envelopes oscillate together around the equilibrium solution with the same frequency for all slices ($\sqrt{2}k_{ext}$, often called the plasma frequency) dependent only on the external focusing forces. This solution represents a collective behaviour of the bunch similar to that of the electrons subject to the restoring force of ions in a plasma. Using the two-slices model and Eq. (54) the emittance evolution Eq. (46) results in

$$\varepsilon_{rms,cor} = \frac{1}{4}k_{sol}\sigma_{rB}\left|\frac{\Delta I}{\hat{I}}\delta_0 \sin\left(\sqrt{2}k_{ext}z\right)\right|$$

(55)

where $\Delta I = \hat{I}_1 - \hat{I}_2$. Notice that, in this simple case, envelope oscillations of the mismatched slices induce correlated emittance oscillations that periodically go back to zero, showing the reversible nature of the correlated emittance growth. It is, in fact, the coupling between transverse and longitudinal motion induced by the space-charge fields that allows reversibility. With a proper tuning of the transport line length or of the focusing field one can compensate for the transverse emittance growth at the expenses of the longitudinal emittance.

At first it may seem surprising that a beam with a single charge species can exhibit plasma oscillations, which are characteristic of plasmas composed of two-charge species. But the effect of the external focusing force can play the role of the other charge species, providing the necessary restoring force that is the cause of such collective oscillations, as shown in Fig. 8. The beam can actually be considered as a single component, relativistic, cold plasma.
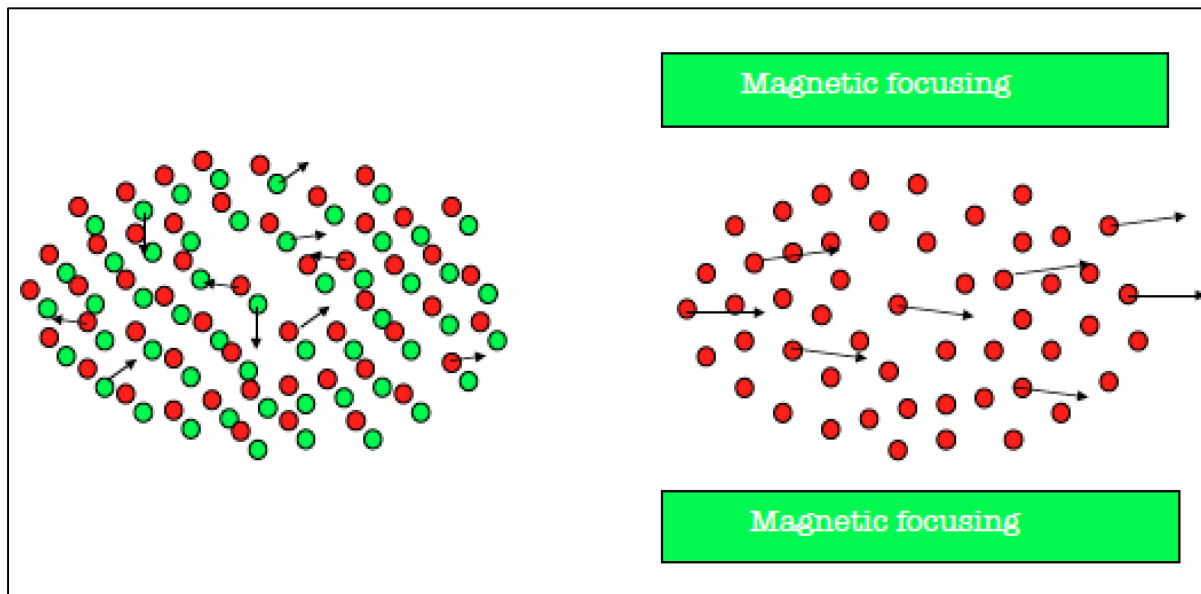
Fig. 8: The restoring force produced by the ions (green dots) in a plasma may cause electron (red dots) oscillations around the equilibrium distribution. In a similar way the restoring force produced by a magnetic field may cause beam envelope oscillations around the matched envelope equilibrium.

It is important to bear in mind that beams in linacs are also different from plasmas in some important respects [5]. One is that beam transit time through a linac is too short for the beam to reach thermal equilibrium. Also, unlike a plasma, the Debye length of the beam may be larger than, or comparable to, the beam radius, so shielding effects may be incomplete.

8 Matching conditions in a plasma accelerator

The concepts developed for the beam transport in the previous sections can be now applied to the case of a plasma accelerator [15], giving important information about the critical topic of beam–plasma matching conditions. To this end we introduce a simplified model for the plasma and for the resulting fields acting on the beam in order to be able to write an envelope equation for the accelerated beam.

In this section we are interested in the case of the external injection of particles in a plasma wave, in the so-called 'bubble' regime, that could be excited by a short, intense laser pulse [15, 16] or by a driving electron beam [17, 18] with beam density n_b larger than the plasma density n_0, $n_b > n_0$. A very simplified model for the plasma behind the driving pulse is illustrated in Fig. 9. We will consider a spherical, uniform ion distribution, as indicated by a dashed circle, with particle density n_0. This model is justified by the fact that, in this regime, the fields are linear in longitudinal and transverse directions, at least in the region of interest for particle acceleration, as that produced by a uniform ion distribution within a sphere of radius $R_{\text{sphere}} \approx \lambda_p / 2$ where $\lambda_p = 2\pi c \sqrt{\varepsilon_0 m / n_0 e^2}$ is the plasma wavelength. A more detailed treatment [19] shows that the correct scaling is $R_{\text{sphere}} = 2\sqrt{(n_b / n_0)}\sigma_r$, where σ_r is the driving beam r.m.s. radius, that for a uniform cylindrical driving bunch gives $R_{\text{sphere}} = \sqrt{\dfrac{4eI}{\pi^3 m c^3 \varepsilon_0}} \dfrac{\lambda_p}{2}$.
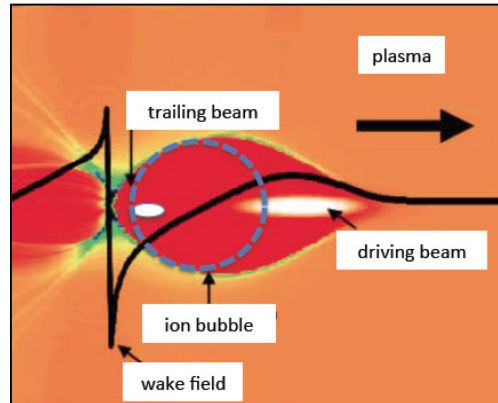
Fig. 9: Schematic representation of the longitudinal wake field (black line) and ion distribution (red area) behind a driving laser or particle beam [16].

The field produced by the ions and experienced by a witness electron beam is purely electrostatic, being the ions at rest in the laboratory frame in the timescale of interest, and is simply given by

$$E_r = \frac{en_0}{3\varepsilon_0} r,$$

(56)

i.e. it has a radial symmetry (other authors, see for example Ref. [17], consider a uniform charged cylindrical ion column producing a transverse field of the form $E_r = \frac{en_0}{2\varepsilon_0} r$). The ion sphere is 'virtually' moving along z with the speed β_d of the driving pulse due to the plasma electron collective oscillation, even if the source of the field remains at rest in the laboratory frame. There are also magnetic fields produced by the plasma electron displacement but, as shown in Ref. [20], the net effect on a relativistic beam is negligible.

The accelerating component of the field is linearly increasing from the moving sphere centre $z_c = \beta_d ct$:

$$E_z(\zeta) = \frac{en_0}{3\varepsilon_0}\zeta$$

(57)

where $\zeta = z - z_c$, and has a maximum on the sphere edge at $\zeta = \lambda_p / 2$. The corresponding energy gained by a witness electron is given by $\gamma = \gamma_0 + \alpha L_{act}$ where L_{acc} is the accelerating length in the plasma and $\alpha(\zeta) = \frac{eE_z(\zeta)}{mc^2} = \frac{1}{3}\left(\frac{2\pi c}{\lambda_p}\right)^2 \zeta$ is the normalized accelerating gradient. The energy spread accumulated by a bunch of finite r.m.s. length σ_z is given by $\frac{\delta\gamma}{\gamma} = \frac{\delta\alpha L_{acc}}{\gamma_0 + \alpha L_{acc}} \approx \frac{\delta\alpha}{\alpha} = \frac{\sigma_z}{\lambda_p}$, showing that ultra-short electron bunches are required to keep energy spread below 1%. In this simplified model, beam loading effects are not considered, nor beam slippage with respect to the driving pulse.

The transverse (focusing) field

$$E_x = \frac{en_0}{3\varepsilon_0} x$$

(58)

at a distance x off the propagation axis is independent of ζ so that correlated emittance growth is not typically induced by the ion focusing field.

In Fig. 10 are shown the plasma wavelength and the longitudinal and transverse fields experienced by a test particle located at $x = 1$ μm and $\zeta = \lambda_p / 4$ versus typical plasma densities, according to Eqs. (54, 55).
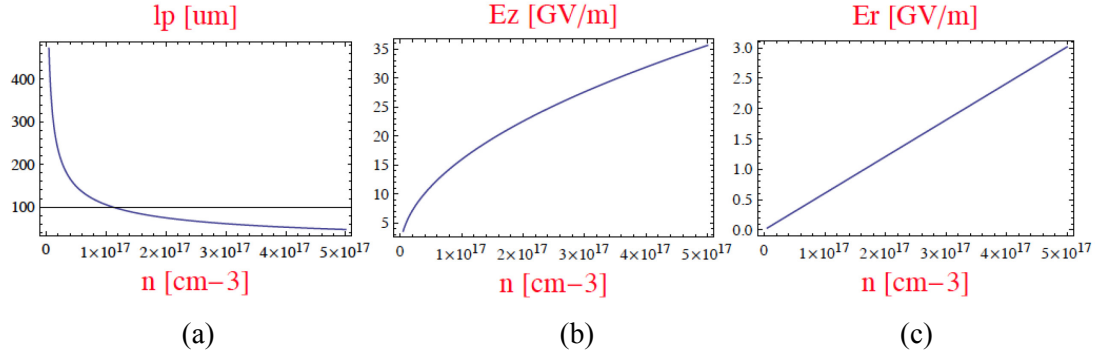
Fig. 10: (a) Plasma wavelength, (b) longitudinal and (c) transverse fields versus typical plasma densities experienced by a test particle located at $x = 1$ μm and $\zeta = \lambda_p / 4$.

As discussed in the previous sections the transverse beam dynamics can be conveniently described by means of a proper envelope equation. To this end let us consider the single particle equation of motion:

$$x'' = \frac{F_x}{\beta cp} = \frac{e^2 n_0}{3 \varepsilon_0 \gamma mc^2} x = \frac{k_p^2}{3\gamma} x$$

(59)

where $k_p = \sqrt{ne^2 / \varepsilon_0 mc^2}$ is the plasma wave number. The moment of the force acting on the beam particles is given by

$$\langle xx'' \rangle = \frac{k_p^2}{3\gamma} \langle x^2 \rangle = \frac{k_p}{3\gamma} \sigma_x^2 .$$

(60)

Inserting into the envelope equation we obtain

$$\sigma_x'' + \frac{\gamma'}{\gamma} \sigma_x' + \frac{k_p^2}{3\gamma} \sigma_x = \frac{\varepsilon_n^2}{\gamma^2 \sigma_x^3} + \frac{k_{sc}^0}{\gamma^3 \sigma_x}.$$

(61)

An equilibrium solution of the previous equation has not yet been found, nevertheless some simplification is still possible and an approximated matching condition exists. As one can see, there are two focusing terms, the adiabatic damping and the ion focusing, and two defocusing terms, the emittance pressure and the space-charge effects. To compare the relative importance of the first two terms it is more convenient to rewrite the previous equation with the new variable $\tilde{\sigma}_x = \sqrt{\gamma} \sigma_x$ leading to the equation

$$\tilde{\sigma}_x'' + \left(\left(\frac{\gamma'}{2\gamma} \right)^2 + \frac{k_p^2}{3\gamma} \right) \tilde{\sigma}_x = \frac{\varepsilon_n^2}{\tilde{\sigma}_x^3} + \frac{k_o^{sc}}{\gamma^2 \tilde{\sigma}_x} .$$

(62)

The beam is space charge dominated, as already discussed in Section 6, when

$$\rho = \frac{k_o^{sc} \tilde{\sigma}_x^2}{\varepsilon_n^2 \gamma^2} = \frac{k_o^{sc} \sigma_x^2}{\varepsilon_n^2 \gamma} \gg 1$$

(63)

and ion focusing dominated when

$$\eta = \frac{4\gamma k_p^2}{3\gamma'^2} >> 1 \ . \tag{64}$$

With the typical beam parameters of a plasma accelerator: 1 kA peak current, 2 μm normalized emittance, injection energy $\gamma_0 = 300$ and spot size about 3 μm, we have $\rho < 1$ and $\eta > 1$. It follows that the envelope equation (61) can be well approximated by the following reduced expression:

$$\sigma_x'' + \frac{k_p^2}{3\gamma}\sigma_x = \frac{\varepsilon_n^2}{\gamma^2\sigma_x^3} \tag{65}$$

with $\gamma(z) = \gamma_0 + \alpha z$. Looking for a particular solution in the form $\sigma_x = \gamma^{-1/4}\sigma_0$ we obtain

$$\left(\frac{5}{16}\gamma'^2 + \frac{1}{3}\gamma k_p^2\right)\sigma_0 = \frac{\gamma\varepsilon_n^2}{\sigma_0^3} \tag{66}$$

that for $\eta > 1$ has a simple solution $\sigma_0 = \sqrt{\dfrac{\sqrt{3}\varepsilon_n}{k_p}}$ giving the matching condition of the beam with the

plasma

$$\sigma_x = \gamma^{-1/4}\sigma_0 = \sqrt[4]{\frac{3}{\gamma}}\sqrt{\frac{\varepsilon_n}{k_p}} \ . \tag{67}$$

In Fig. 11 are shown the matched beam envelope given by Eq. (67) with normalized emittance of 2 μm and injection energy $\gamma = 300$ versus the plasma density. The figure also shows the evolution of the beam envelope in a 10 cm long plasma with density 10^{16} cm^{-3}, corresponding to an accelerating field of 5 GV/m (extraction energy $\gamma = 1300$) and focusing field of 60 MV/m.

(a) (b)
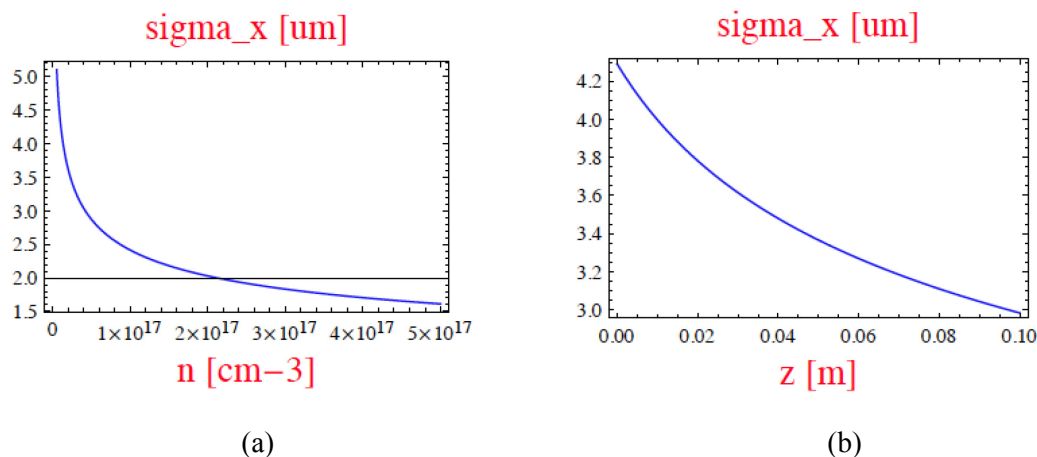
Fig. 11: (a) Matched beam envelope with normalized emittance of 2 μm and injection energy $\gamma = 300$ versus (b) the plasma density and the evolution of the beam envelope in a 10 cm long plasma with density 10^{16} cm^{-3}, corresponding to an accelerating field of 5 GV/m and focusing field of 60 MV/m.

Notice that the beam experiences focusing as γ increases and the beam density increases, leading to a significant perturbation of the plasma fields. A possible solution to overcoming this effect is to taper the plasma density along the channel in order to achieve beam transport with a constant envelope.

It is an interesting exercise to see the effect of a plasma density vanishing as $n(z) = \dfrac{\gamma_0}{\gamma(z)} n_0$,

giving $k_p^2 = \dfrac{e^2 n_0}{\varepsilon_0 m c^2} \dfrac{\gamma_0}{\gamma} = \dfrac{\gamma_0}{\gamma} k_{0,p}^2$. In this case the envelope equation Eq. (61) without space-charge effects becomes

$$\sigma_x'' + \frac{\gamma'}{\gamma}\sigma_x' + \frac{\gamma_0 k_{0,p}^2}{3\gamma^2}\sigma_x = \frac{\varepsilon_n^2}{\gamma^2 \sigma_x^3} \tag{68}$$

which admits a constant equilibrium solution

$$\sigma_x = \sqrt[4]{\frac{3}{\gamma_0}}\sqrt{\frac{\varepsilon_n}{k_{0,p}}}. \tag{69}$$

Figure 12 shows the plasma density along the accelerating section and the resulting equilibrium beam envelope given by Eq. (69), with the same beam parameters as those in Fig. 11

(a)   (b)
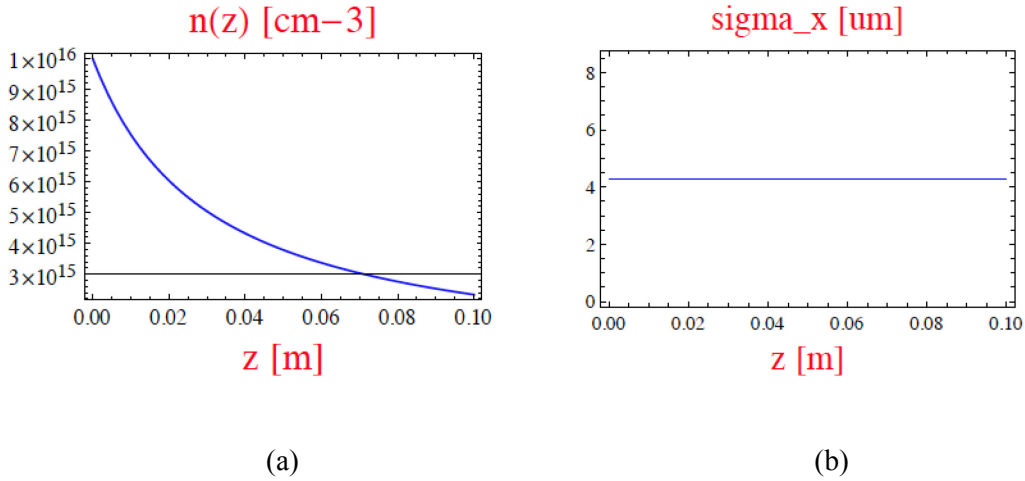
Fig. 12: (a) Plasma density along the accelerating section; (b) the resulting equilibrium beam envelope given by Eq. (69).

On the other hand before injection in the plasma accelerator, the beam has to be focused to the matching spot given by Eq. (67) to prevent envelope oscillations that may cause emittance growth and an enhancement of betatron radiation emission. It has been proposed [21] to shape the plasma density profile in order to gently capture the beam by means of the increasing ion focusing effect. For example, by varying the plasma density as $n(z) = \dfrac{\gamma(z)}{\gamma_0} n_0$ at the entrance of the plasma column, the envelope equation Eq. (61) can be written as

$$\sigma_x'' + \frac{k_{0,p}^2}{3\gamma_0}\sigma_x = \frac{\varepsilon_n^2}{\gamma^2 \sigma_x^3} \tag{70}$$

where $k_p^2 = \dfrac{e^2 n_0}{\varepsilon_0 m c^2}\dfrac{\gamma}{\gamma_0} = \dfrac{\gamma}{\gamma_0} k_{0,p}^2$. This equation has a particular solution assuming that γ'' is negligible,

$$\sigma_x = \sqrt[4]{3\gamma_0}\sqrt{\frac{\varepsilon_n}{\gamma k_{0,p}}}, \tag{71}$$

showing that with a proper choice of the initial plasma density the beam envelope can be gently matched to the accelerating plasma channel.

For additional discussions about injection and extraction beam matching conditions see also some recent papers, Refs. [22–25].

References

[1] T. Shintake, Proc. 22nd Particle Accelerator Conf., Albuquerque, NM, 25–29 June 2007 (IEEE, New York, 2007), p. 89. http://dx.doi.org/10.1109/PAC.2007.4440331

[2] L. Serafini and J.B. Rosenzweig, Physical Review E 55 (1997) p. 7565.

[3] M. Reiser, *Theory and Design of Charged Particle Beams* (Wiley, New York, 1994). http://dx.doi.org/10.1002/9783527617623

[4] J.B. Rosenzweig, *Fundamentals of Beam Physics* (Oxford University Press, New York, 2003). http://dx.doi.org/10.1093/acprof:oso/9780198525547.001.0001

[5] T. Wangler, *Principles of RF Linear Accelerators* (Wiley, New York, 1998). http://dx.doi.org/10.1002/9783527618408

[6] S. Humphries, *Charged Particle Beams* (Wiley, New York, 2002).

[7] N. Pichoff, Beam dynamics basics in RF linacs, Proc. CERN Accelerator School: Small Accelerators, Zeegse, The Netherlands, 2005, CERN-2006-012.
http://dx.doi.org/10.5170/CERN-2006-012

[8] M. Migliorati *et al.*, *Phys. Rev. ST Accel. Beams* **16** (2013) 011302.
http://dx.doi.org/10.1103/PhysRevSTAB.16.011302

[9] M. Ferrario *et al.*, *Nucl. Instrum. Methods. Phys. Res.* B **309** (2013) 183.
http://dx.doi.org/10.1016/j.nimb.2013.03.049

[10] A. Cianchi *et al.*, *Nucl. Instrum. Methods Phys. Res.* A 720 (2013) 153.
http://dx.doi.org/10.1016/j.nima.2012.12.012

[11] F.J. Sacherer, *IEEE Trans. Nucl. Sci.* **NS-18** (1971) 1105.
http://dx.doi.org/10.1109/TNS.1971.4326293

[12] K. Floettmann, *Phys. Rev. ST Accel. Beams* **6** (2003) 034202.
http://dx.doi.org/10.1103/PhysRevSTAB.6.034202

[13] M. Ferrario *et al.*, *Int. J. Mod. Phys.* A **22** (2007) 4214.
http://dx.doi.org/10.1142/S0217751X07037779

[14] J. Buon, Beam phase space and emittance, in Proc. CERN Accelerator School: 5th General Accelerator Physics Course, Jyvaskyla, Finland, 1992, CERN-94-01.
http://dx.doi.org/10.5170/CERN-1994-001

[15] E. Esarey *et al.*, *Rev. Mod. Phys.* **81** (2009) 1229.
http://dx.doi.org/10.1103/RevModPhys.81.1229

[16] C. Joshi and W. B. Mori, *Phil. Trans. R. Soc.* A (2006) **364,** 577. doi:10.1098/rsta.2005.1723

[17] J.B. Rosenzweig *et al.*, *Phys Rev* A **44** (1991) 10. http://dx.doi.org/10.1103/PhysRevA.44.R6189

[18] P. Muggli and M.J. Hogan, *C. R. Physique* **10** (2009) 116.
http://dx.doi.org/10.1016/j.crhy.2009.03.004

[19] W. Lu *et al.*, *Phys. Plasmas* **12** (2005) 063101. http://dx.doi.org/10.1063/1.1905587

[20] W. Lu *et al.*, *Phys. Rev. Lett.* **96** (2006) 165002.
      http://dx.doi.org/10.1103/PhysRevLett.96.165002

[21]   P. Tomassini, Private communication.

[22] R. Lehe *et al.*, *Phys. Rev. ST Accel. Beams* **17** (2014) 121301.
      http://dx.doi.org/10.1103/PhysRevSTAB.17.121301

[23] T. Mehrling *et al.*, *Phys. Rev. ST Accel. Beams* **15** (2012) 111303.
      http://dx.doi.org/10.1103/PhysRevSTAB.15.111303

[24] K. Floettmann, *Phys. Rev. ST Accel. Beams* **17** (2014) 054402.
      http://dx.doi.org/10.1103/PhysRevSTAB.17.054402

[25] I. Dornmair *et al.*, *Phys. Rev. ST Accel. Beams* **18** (2015) 041302.
      http://dx.doi.org/10.1103/PhysRevSTAB.18.041302

Particle-In-Cell Codes for Plasma-based Particle Acceleration

A. Pukhov
University of Düsseldorf, Düsseldorf, Germany

Abstract

In this report we discuss the basic principles of Particle-In-Cell (PIC) codes and their application to plasma-based acceleration. The *ab initio* full electromagnetic relativistic PIC codes provide the most reliable description of plasmas, and their properties are described in detail. However, while they represent the most fundamental model, full PIC codes are computationally expensive. Plasma-based acceleration is a multi-scale problem with very disparate scales. The smallest scale is the laser or plasma wavelength (on the order of one to a hundred microns), and the largest scale is the acceleration distance (which ranges from a few centimetres to metres or even kilometres). The Lorentz boost technique allows the scale disparity to be reduced, at the cost of complicating the simulations and causing unphysical numerical instabilities in the code. Another possibility is to use the quasi-static approximation, whereby the different scales are separated analytically.

Keywords
Particle-In-Cell method; plasma-based acceleration; bubble regime; AWAKE; high-performance computing.

1 Introduction

Plasma-based particle acceleration involves a nonlinear medium, the relativistic plasmas [1]. Modelling this type of medium requires proper numerical simulation tools. Over the past few decades, Particle-In-Cell (PIC) codes have proven to be a very reliable and successful method for kinetic plasma simulations [2–5]. The success of PIC codes relies to a large extent on their analogy with the actual plasma, as suggested by their name. The plasma is in reality an ensemble of many individual particles, electrons and ions, interacting with each other via self-consistently generated fields. PIC codes have a very similar set-up, with the difference being that the number of numerical particles, or macroparticles, that we follow in the code may be significantly smaller than the number of particles in an actual plasma. One could think of one numerical 'macroparticle' as representing a clump, or cloud, of many real particles, which occupy a finite volume in space and all move together with the same velocity. Thus we have a 'numerical plasma' consisting of heavy macroparticles that have the same charge-to-mass ratio as the real plasma electrons and ions, but where each macroparticle substitutes for many real particles.

Depending on the application, different approximations can be used. The most fundamental approximation is provided by the full electromagnetic PIC codes, which solve the Maxwell equations together with the relativistic equations of motion for the numerical particles. These *ab initio* simulations produce the most detailed results, but can be very expensive. In the case of long-scale acceleration, where the driver propagates distances many times greater than its own length, the quasi-static approximation can be exploited. In this case, it is assumed that the driver changes little as it propagates distances that are comparable with its own length. The quasi-static approximation enables separation of fast and slow variables and hence great acceleration of the simulation, but at the cost of omitting radiation: the laser pulse or any emitted radiation cannot be described directly by such codes; rather, an additional module for the laser pulse is required, usually in the envelope approximation.

In this report, we describe the basic principles of PIC methods—both the full electromagnetic codes and the quasi-static approximation—as well as their application to plasma-based acceleration.

2 The basic equations

First, let us formulate the problem we are going to solve. To conduct electromagnetic and kinetic simulations, we have to solve the full set of Maxwell equations [6]:

$$\frac{\partial \mathbf{E}}{\partial t} = c\nabla \times \mathbf{B} - 4\pi\mathbf{j}, \tag{1}$$

$$\frac{\partial \mathbf{B}}{\partial t} = -c\nabla \times \mathbf{E}, \tag{2}$$

$$\nabla \cdot \mathbf{E} = 4\pi\rho, \tag{3}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{4}$$

where we use metric cgs units and c is the speed of light in vacuum.

Let us pause for a moment at this very fundamental system of equations. The electric and magnetic fields, $\mathbf{E}$ and $\mathbf{B}$, evolve according to the time-dependent equations (1) and (2), where the source term is in the form of the current density $\mathbf{j}$. This current is produced by the self-consistent charge motion in our system of particles. It is well known from texts on electrodynamics (see, e.g., Ref. [6]) that the Gauss law (3) together with the curl-free part of Eq. (2) lead to the charge continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0. \tag{5}$$

One can apply the operator $\nabla\cdot$ to the Faraday law (2) and use the Gauss equation (3) for $\nabla \cdot \mathbf{E}$ to obtain (5). The opposite is true as well: if the charge density always satisfies the continuity equation (5), then the Gauss equation (3) is satisfied automatically during the evolution of the system, if it was satisfied initially. By symmetry this is also valid for the magnetic field $\mathbf{B}$. As there is no magnetic charge, Eq. (4) remains valid always if it was valid initially.

This means that we may reduce our problem to solving the two evolution equations (1) and (2), while considering Eqs. (3) and (4) as initial conditions only. This is a very important and fruitful approach; PIC codes using it have a 'local' algorithm, i.e. at each time step information is exchanged between neighbouring grid cells only. No global information exchange is possible because the Maxwell equations have 'absolute future' and 'absolute past' [7]. This property makes the corresponding PIC codes perfectly suitable for parallelization, and in addition the influence of (always unphysical) boundary conditions is strongly reduced.

3 Kinetics and hydrodynamics

Now we define our source term $\mathbf{j}$. In general, to do this, we have to know the distribution function of the plasma particles,

$$F^N(\mathbf{x}_1, \mathbf{p}_1, \ldots, \mathbf{x}_N, \mathbf{p}_N), \tag{6}$$

which defines the probability that an N-particle system takes a particular configuration in the $6N$-dimensional phase space. Here $\mathbf{x}_n$ and $\mathbf{p}_n$ are the coordinates and momentum of the nth particle. The function (6) provides the *exhaustive* description of the system. However, as has been shown in statistical physics (see, e.g., Ref. [8]), knowing the single-particle distribution function for each species of particle may suffice to describe the full system. The sufficient condition is that the inter-particle correlations should be small and can be treated perturbatively. The equation governing the evolution of the single-particle distribution function $f(\mathbf{x}, \mathbf{p})$ is called the Boltzmann–Vlasov equation [9, 10]:

$$\frac{\partial f}{\partial t} + \frac{\mathbf{p}}{m\gamma}\nabla f + \frac{\mathbf{F}}{m}\nabla_p f = \mathrm{St}, \tag{7}$$

where m is the single-particle mass of the species, $\gamma = \sqrt{1 + (p/mc)^2}$ is the relativistic factor, $\mathbf{F}$ is the force, and St is the collisional term (inter-particle correlations).

The kinetic equation (7) for the single-particle distribution function is six-dimensional and still complicated. Solving it either analytically or numerically is a challenge.

However, under appropriate conditions one can make further simplifications. In statistics it has been shown that inter-particle collisions lead to the Maxwellian distribution function (H-theorem; see, e.g., Ref. [8]). In the non-relativistic case the Maxwellian distribution has the form

$$f(\mathbf{x}, \mathbf{v}) = \frac{n(\mathbf{x})}{\sqrt{2\pi T}} \exp\left(-\frac{(\mathbf{v} - \mathbf{V})^2}{2mT}\right),$$ (8)

where $n(\mathbf{x})$ is the local particle density, T is the temperature, and $\mathbf{V}$ is the local streaming velocity. The characteristic time for establishment of the Maxwellian distribution is the inter-particle collision time. Therefore, if the effective collision time in the system is short in comparison with other characteristic times, the distribution function remains Maxwellian. It is sufficient, then, to write evolution equations for the *momenta* of the distribution function, such as the local density

$$n(\mathbf{x}) = \int f(\mathbf{x}, \mathbf{v})\, \mathrm{d}^3 \mathbf{v},$$ (9)

the hydrodynamic velocity

$$\mathbf{V}(\mathbf{x}) = \int \mathbf{v} f(\mathbf{x}, \mathbf{v})\, \mathrm{d}\mathbf{v},$$ (10)

and the temperature

$$T(\mathbf{x}) = \int \frac{(\mathbf{v} - \mathbf{V})^2}{2m} f(\mathbf{x}, \mathbf{v})\, \mathrm{d}^3 \mathbf{v}.$$ (11)

These quantities are said to be fluid-like, or of hydrodynamic type.

4 Vlasov and PIC codes

If, however, the distribution function does deviate, or is expected to deviate, significantly from the Maxwellian distribution, then we have to solve the Boltzmann–Vlasov equation, (7). What would be the appropriate numerical approach here?

At the first glance, it might seem that the most straightforward approach is to solve the partial differential equation (7) using finite differences on the Eulerian grid in phase space. Indeed, this direction has been pursued by several groups [11], and has gained even more popularity with the rapid growth of available computing power. One potential advantage of these 'Vlasov' codes is the possibility of producing 'smooth' results. Indeed, Vlasov codes handle the distribution function, which outputs a smoothly varying real number that gives the probability of finding plasma particles at a particular point in phase space.

Vlasov codes, however, are very computationally expensive, and even one-dimensional problems may require the use of parallel supercomputers. The reason these codes need so much computational power can be seen from Fig. 1(a). It shows schematically a mesh that one would need for a 1d1v Vlasov code. The notation '1d1v' means that the code resolves one spatial coordinate and one coordinate in the momentum (velocity) space. The shaded area represents the region of phase space that is occupied by plasma particles, where the associated two-dimensional distribution function $f(x, p_x)$ is essentially non-zero. The unshaded region is void of particles, and nothing interesting happens there. Nevertheless, one has to maintain such empty regions as parts of the numerical arrays, and process them when solving Eq. (7) on the Eulerian grid. This processing of empty regions leads to an enormous waste of computational power. This drawback becomes even more severe with an increase in the dimensionality of the problem under consideration. The efficiency of Vlasov codes drops exponentially with the number of dimensions, and becomes minuscule in the real 3d3v case, where the computer has to retain in memory and process a six-dimensional mesh, most of it empty of particles.
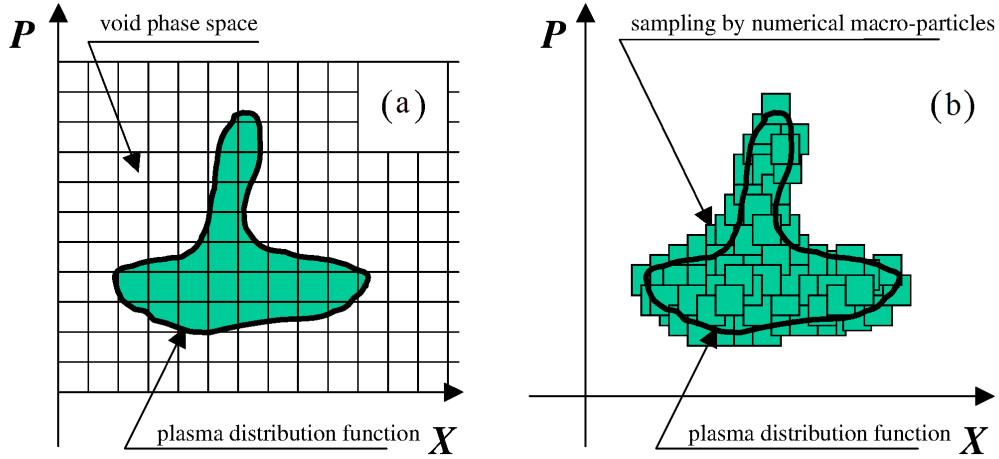
Fig. 1: Kinetic plasma simulations: (a) Vlasov method, using an Eulerian grid in the phase space; (b) PIC method, where numerical macroparticles mark the distribution function.

There is another, currently more computationally effective, method to solve the Boltzmann–Vlasov equation (7), namely the *finite element* method. The principle of this method is illustrated in Fig. 1(b). Again, imagine some distribution function in the phase space (the shaded region). Now, let us approximate, or sample, this distribution function by a set of Finite Phase-Fluid Elements (FPFEs):

$$f(\mathbf{x}, \mathbf{p}) = \sum_n W_n^{\mathrm{ph}}\, S^{\mathrm{ph}}(\mathbf{x} - \mathbf{x}_n, \mathbf{p} - \mathbf{p}_n), \qquad (12)$$

where W_n^{ph} is the 'weight' of the nth FPFE and $S^{\mathrm{ph}}(\mathbf{x}, \mathbf{p})$ is the 'phase shape', or the support function in phase space. The centre of the nth FPFE is positioned at $(\mathbf{x}_n, \mathbf{p}_n)$. We are free to make a choice of the support function. For simplicity, we choose here a six-dimensional hypercube

$$S^{\mathrm{ph}}(\mathbf{x}, \mathbf{p}) = 1, \quad |x_\alpha| < \frac{\Delta x_\alpha}{2}, \quad |p_\alpha| < \frac{\Delta p_\alpha}{2} \quad \text{for } \alpha = x, y, z,$$

where Δx_α is the FPFE size along the j-axis in configuration space, and Δp_α is the FPFE size along the p_α-axis in momentum space.

The 'phase fluid' transports the distribution function along the characteristics of the Boltzmann–Vlasov equation (see, e.g., Ref. [18]). So we have to advance the centres of the FPFEs along the characteristics:

$$\frac{\mathrm{d}\mathbf{x}_n}{\mathrm{d}t} = \frac{\mathbf{p}}{m\gamma}, \qquad (13)$$

$$\frac{\mathrm{d}\mathbf{p}_n}{\mathrm{d}t} = \mathbf{F} + \mathbf{F}_{\mathrm{St}}, \qquad (14)$$

where $\mathbf{F}_{\mathrm{St}}$ denotes the effective 'collisional' force due to the collision term in Eq. (7). The FPFEs follow the evolution of the distribution function in phase space. Of course, Eqs. (13) and (14) are just the relativistic equations of motion of particles! Thus, the FPFE method is equivalent to the PIC method.

A significant advantage of the finite element method over the Vlasov codes is that one does not need to maintain a grid in the full phase space. Instead, the FPFEs sample (or mark) only the interesting regions where particles are present and something important is going on. We still do maintain a grid in the *configuration* space to solve the field equations (1) and (2), but this grid has only three dimensions

(and not six as in the Vlasov case). Thus, PIC codes may be viewed as 'packed' or 'Lagrangian' Vlasov codes. Moreover, the FPFE approach is even more fundamental than the Boltzmann–Vlasov equation itself, because it can easily be generalized to the case where one macroparticle corresponds to just one real particle, and where inter-particle correlations are not small. The corresponding codes are usually called P³M (particle–particle–particle–mesh) codes [5].

As soon as we consider our macroparticles not simply as 'large clumps of real particles' but as finite elements in the phase space, we find that there is no fundamental obstacle to simulation of a cold plasma. Moreover, it is in this setting that the finite element approach really becomes effective computationally and superior to the Vlasov codes. The phase space of a cold plasma is degenerate: the particles occupy a mere three-dimensional hypersurface in the full six-dimensional phase space. Evidently, this hypersurface can be accurately sampled by even a relatively small number of macroparticles (FPFEs). As the system evolves, this surface deforms, stretches and contracts, but it remains degenerate and three-dimensional, unless any heating (i.e. diffusion in the phase space) is present. There is a full stock of interesting physical phenomena associated with relativistic laser–plasma interactions where the physical collisional heating is negligible. Unfortunately, the numerical heating that occurs in the 'standard' PIC codes [4] leads to an unphysical numerical diffusion in the phase space, which spoils the picture. Any code able to successfully simulate initially cold plasma must be energy-conserving.

5   Continuity equation

Historically, the first PIC codes were electrostatic [5], and they have to solve explicitly the Poisson equation

$$\nabla^2 \phi = -4\pi\rho, \tag{15}$$

giving the static electric field

$$\mathbf{E}_\parallel = -\nabla\phi. \tag{16}$$

Generalization to the electromagnetic case seemed quite natural; one would simply add the vector potential $\mathbf{A}$ to get

$$\mathbf{E} = \mathbf{E}_\parallel + \mathbf{E}_\perp = -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}. \tag{17}$$

Yet, there is another way of treating the electromagnetic fields. We established in the previous section that the simultaneous solution of Ampère's law (1) and the continuity equation (5) satisfies the Gauss law (3) automatically. Hence, one can work with the fields $\mathbf{E}$ and $\mathbf{B}$ directly, without introducing the electrostatic potential ϕ and without solving the Poisson equation (15).

It is very advantageous to avoid solving the Poisson equation (15), as it is *nonlocal*. This is an elliptic equation, and its solution essentially depends on the (usually unphysical) boundary conditions. A small perturbation or numerical error at the boundary may give rise to a global perturbation in the full simulation domain. In contrast, the Maxwell equations (1) and (2) are local. Any signal can propagate no faster than the vacuum speed of light, and we refer here to the Minkovski diagram, Fig. 2. As the central event (t_n, x_i, y_j, z_k) we choose some grid cell with indices (n, i, j, k) so that $t_n = n\tau$, $x_i = i\Delta x$, $y_j = j\Delta y$ and $z_k = k\Delta z$. Here $\tau, \Delta x, \Delta y$ and Δz denote the numerical steps along the time, X-, Y- and Z-axes. The light cone separates the full four-dimensional space into the regions of 'absolute past', 'absolute future', and 'absolutely distant' events. Only the events taking place in the 'absolute past' may stay in a casual connection with the central event. Thus, fields at the grid position (t_n, x_i, y_j, z_k) are influenced by the events happening at the instant t_{n-1} at the grid cells located within the circle $c\tau$ around the original cell. If we use an explicit numerical scheme, then the time step is limited through the Courant condition $c\tau < \min(\Delta x, \Delta y, \Delta z)$, and only the immediate neighbouring cells are involved. A numerical scheme that has this physical property is said to be *local*. In this sense, any numerical scheme that involves the solution of an elliptic equation, like the Poisson equation (15), is *nonlocal*.
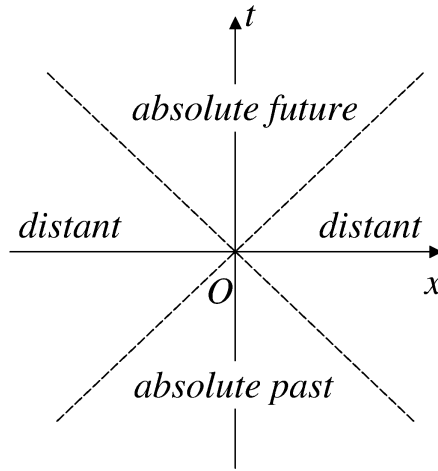
Fig. 2: Minkovski space–time diagram illustrating the causality in special relativity theory

A key issue in developing a local numerical scheme is the method of *current deposition* on the grid during particle motion. Let us consider a cubic FPFE (numerical particle) on a grid. We suppose that the particle and the grid elementary volume (i.e. cell volume) $V_c = \Delta x \Delta y \Delta z$ are *identical*, so that the particle length Δx_α is also the grid step along the α-axis, for $\alpha = x, y, z$. We mark the grid cells with the indices i, j and k along the x-, y- and z-axes

In discussing multi-dimensional PIC codes, we normally use a staggered or Yee lattice (grid), as illustrated in Fig. 3. We define the charge density on the grid at the centres of the cells, $\rho_{i+1/2,\,j+1/2,\,k+1/2}$, as

$$\rho_{i+1/2,\,j+1/2,\,k+1/2} = \sum_n W_n^\rho \, S^\rho(\mathbf{x}_{i+1/2,\,j+1/2,\,k+1/2} - \mathbf{x}_n). \tag{18}$$

The weight and form of the charge density interpolation for the particle are

$$S^\rho(\mathbf{x}) = S_x^\rho(x) S_y^\rho(y) S_z^\rho(z),$$

$$S_j^\rho(\mathbf{x}_j) = 1 - 2\frac{|x_j|}{\Delta_j}, \quad |x_j| < 0.5\Delta_j. \tag{19}$$

The scheme (19) is the 'volume' (or 'area') weighting. It actually assigns the portion of the particle residing in a cell to the cell's centre.
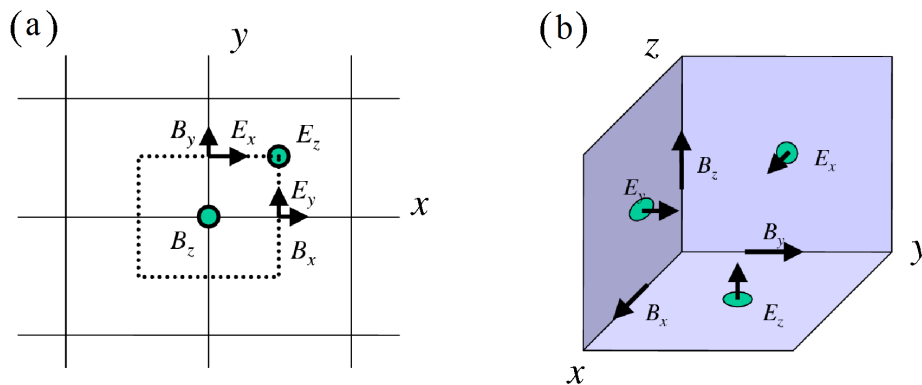
Fig. 3: Yee lattice in (a) 2D and (b) 3D

Now, if the particle moves, it generates current. How should one interpolate this current to the grid cells? One could try to use a straightforward interpolation, say $\mathbf{J} = \sum_n \mathbf{V}_n S_n^\rho$, or others like those discussed in Birdsall and Langdon's book [4]. Generally, such interpolations do not satisfy the continuity equation, i.e. defined in this way the current flux through a cell's boundaries does not represent the actual charge change in the cell. A further consequence is revealed when we integrate in time the Ampère law (1): the electric field obtained does not satisfy the Gauss law (3).

One possible way around this inconsistency is to *correct* the electric field obtained [4]. Suppose that we have advanced the electric field $\mathbf{E}'$ according to Eq. (1) and we run into difficulties with the Gauss law: $\nabla \cdot \mathbf{E}' \neq 4\pi\rho$. We could try to correct the electric field by introducing a potential $\delta\phi$ such that

$$\nabla^2 \delta\phi = -4\pi\rho + \mathbf{E}', \tag{20}$$

and then construct the corrected electric field as

$$\mathbf{E} = \mathbf{E} - \nabla\delta\phi, \tag{21}$$
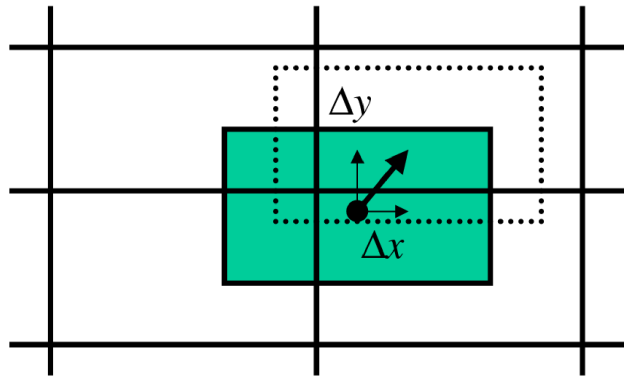
which does satisfy the Gauss law:

$$\nabla\mathbf{E} = 4\pi\rho. \tag{22}$$

Unfortunately, this correction requires us to solve the nonlocal elliptic problem (20).

It turns out, however, that the currents can be defined in a self-consistent way [2]. To do so, one has to follow the particle trajectory in detail, and keep recording of how much charge has passed through each of the cell's boundaries. Fig. 4 illustrates the idea. Let us take a particle with centre located inside the grid elementary volume V_c centred at the grid vertex (i, j, k). The particle's initial position is (x_0, y_0, z_0), and after one time step the particle moves to the new position (x_1, y_1, z_1). To begin with, we suppose that the new position is still inside the elementary volume. We denote the particle displacement by $(\delta x, \delta y, \delta z)$. The particle generates the instantaneous current density $\mathbf{j} = \mathbf{V}W^\rho S^\rho$, and the current fluency $\mathbf{J}$, i.e. the charge that has crossed some surface Ω during the time step τ, is given by

$$\mathbf{J} = \int_\Omega d\mathbf{\Omega} \int_0^\tau \mathbf{V}W^\rho S^\rho \, dt = \int_\Omega d\mathbf{\Omega} \int_{\mathbf{x_0}}^{\mathbf{x_1}} W^\rho S^\rho \, d\mathbf{x}. \tag{23}$$

Fig. 4: Tracing of a particle's trajectory

Thus, we have to integrate along the particle trajectory. If, as usual [4], we are using the second-order finite difference scheme to advance the particle position, then the particle moves along a straight line during one time step. Assuming this, it is easy to calculate that if the particle remained within the elementary volume, it has induced the following currents on the grid:

$$J^x_{i,\, j+1/2,\, k+1/2} = \delta x \, W^\rho (a_y a_z + b_{yz}),$$

$$J^y_{i+1/2,\,j,\,k+1/2} = \delta y\, W^\rho (a_z a_x + b_{zx}),$$

$$J^z_{i+1/2,\,j+1/2,\,k} = \delta z\, W^\rho (a_x a_y + b_{xy}),$$

$$J^x_{i,\,j-1/2,\,k+1/2} = \delta x\, W^\rho [(1 - a_y)a_z - b_{yz}],$$

$$J^x_{i,\,j+1/2,\,k-1/2} = \delta x\, W^\rho [a_y(1 - a_z) - b_{yz}],$$

$$J^x_{i,\,j-1/2,\,k-1/2} = \delta x\, W^\rho [(1 - a_y)(1 - a_z) + b_{yz}],$$

$$J^y_{i+1/2,\,j,\,k-1/2} = \delta y\, W^\rho [(1 - a_z)a_x - b_{zx}],\tag{24}$$

$$J^y_{i-1/2,\,j,\,k+1/2} = \delta y\, W^\rho [a_z(1 - a_x) - b_{zx}],$$

$$J^y_{i-1/2,\,j,\,k-1/2} = \delta y\, W^\rho [(1 - a_z)(1 - a_x) + b_{zx}],$$

$$J^z_{i-1/2,\,j+1/2,\,k} = \delta z\, W^\rho [(1 - a_x)a_y - b_{xy}],$$

$$J^z_{i+1/2,\,j-1/2,\,k} = \delta z\, W^\rho [a_x(1 - a_y) - b_{xy}],$$

$$J^z_{i-1/2,\,j-1/2,\,k} = \delta z\, W^\rho [(1 - a_x)(1 - a_y) + b_{xy}],$$

where

$$a_\alpha = 1 - 2\frac{|x_\alpha + 0.5\Delta_\alpha|}{\Delta_\alpha},$$

$$b_{\alpha\beta} = \frac{1}{12}\delta x_\alpha \delta x_\beta \tag{25}$$

$$\text{for } \alpha, \beta \in \{x, y, z\}.$$

If the particle leaves the elementary volume where it was residing initially, then the full displacement must be split into several 'elementary' motions. During each elementary motion the particle must remain inside an elementary volume surrounding the corresponding vertex of the grid. This 'bookkeeping' of the particle motion does require some programming effort, but it appears to be very important to implement it.

The electric field is then advanced in time according to

$$\mathbf{E}^{n+1} - \mathbf{E}^n = c\tau\hat{\nabla} \times \mathbf{B}^{n+1/2} - 4\pi\mathbf{J}^{n+1/2}, \tag{26}$$

with the particular components of the electric field $\mathbf{E}$ defined at the same positions on the grid as the current $\mathbf{J}$; here $\hat{\nabla}\times$ denotes the finite difference version of the curl operator.

One may think of an alternative approach that is somewhat easier from the programming point of view: why don't we replace the actual, straight motion of the particle during one time step by the *average* of all possible *rectangular* paths along the grid axes which connect the initial and final positions of the particle? This approach was taken by Morse and Nielson [12]; however, this 'fake' integration led to an unacceptably rapid growth of *electromagnetic noise* in their code. The author has also found that even small deviations from the accurate current deposition (24) will immediately result in noise boosting, even if the deviated scheme is still charge-conserving.

Therefore, scheme (24) is *the* method to use to avoid solving elliptic equations and yet still satisfy the Gauss law numerically. In other words, we rigorously enforce the *detailed*—i.e. down to each grid cell—charge conservation and correct continuity equation.

6 Energy conservation

In the previous section we discussed how to develop an electromagnetic PIC code that is rigorously charge-conserving. Another important conservation law we would like to enforce is total energy conservation. Indeed, it is well known that one of the worst plagues of standard PIC codes is the effect of

numerical heating. The numerical 'temperature' (or rather the chaotic energy per numerical particle) is known to grow exponentially until the effective Debye length becomes comparable with the grid size; thereafter the exponential growth transitions to a more moderate linear heating. This is an effect of 'aliasing'—the inconsistent interpolation of the fields defined on the grid to the particle position.

As we hinted in the Introduction, there is no fundamental reason for numerical heating to occur if we adhere to the FPFE paradigm. Now we proceed to design an *energy-conserving electromagnetic code*. We start with the exact analytical equation for the full energy of the system,

$$H = \sum_n m_n c^2 (\gamma - 1) \; + \; \frac{1}{8\pi} \int_V (E^2 + B^2) \, \mathrm{d}V, \tag{27}$$

where m_n is the particle's mass, $\gamma = \sqrt{1 + (p/m_n c)^2}$ is the relativistic γ-factor, and the integration is taken over the full volume V.

Next, we split the electric field into longitudinal $\mathbf{E}_\parallel$ and transverse $\mathbf{E}_\perp$ parts, so that

$$\begin{aligned} \nabla \cdot \mathbf{E}_\perp &= 0, \\ \nabla \times \mathbf{E}_\parallel &= 0. \end{aligned} \tag{28}$$

Then, we introduce a potential ϕ such that $\mathbf{E}_\parallel = -\nabla\phi$. One can show easily that

$$\int_V \mathbf{E}_\perp \mathbf{E}_\parallel \, \mathrm{d}V = - \int_\Omega \phi \mathbf{E}_\perp \, \mathrm{d}\boldsymbol{\Omega} = 0 \tag{29}$$

for an infinite or periodic volume. Here Ω is a surface surrounding the volume. As a consequence, we can write the energy of our system as

$$H = \sum_n m_n c^2 (\gamma - 1) \; + \; \frac{1}{8\pi} \int_V (E_\parallel^2 + E_\perp^2 + B^2) \, \mathrm{d}V = H_{\mathrm{kin}} + H_{\mathrm{S}} + H_{\mathrm{EM}} + H_B, \tag{30}$$

where

$$H_{\mathrm{kin}} = \sum_p m_p c^2 (\gamma - 1) \tag{31}$$

is the kinetic energy of the particles,

$$H_{\mathrm{S}} = \frac{1}{8\pi} \int_V E_\parallel^2 \, \mathrm{d}V \tag{32}$$

is the electrostatic part of the electric field energy,

$$H_{\mathrm{EM}} = \frac{1}{8\pi} \int_V E_\perp^2 \, \mathrm{d}V \tag{33}$$

is the electromagnetic part of the electric field energy, and

$$H_B = \frac{1}{8\pi} \int_V B^2 \, \mathrm{d}V \tag{34}$$

is the magnetic field energy.

The expression (30) for the energy is for continuous fields and individual particles; however, it is straightforward to write an analogue for a finite difference numerical scheme.

We are using the staggered grid (Yee lattice) and have fixed the current interpolation to the grid using the system (24). Now we have to define the *force interpolation* to the actual particle position in such a way that the resulting numerical scheme conserves the Hamiltonian (30).

Most dangerous in terms of numerical heating is the *electrostatic* part of the code. It is the electrostatic plasma waves that are responsible for the Debye shielding. Also, the $\mathbf{v} \times \mathbf{B}$ part of the Lorentz force acting on the particle conserves energy automatically, as does the $\mathbf{B}$-field advance according to the Faraday law (2). Hence, for the time being, we neglect the magnetic field and the magnetic energy part, and we enforce the conservation of $H_\mathrm{E} = H_\mathrm{kin} + H_\mathrm{S} + H_\mathrm{EM}$. The numerical scheme will conserve the energy if it is derived from equations in the canonical form

$$\frac{\mathrm{d}\mathbf{p}_p}{\mathrm{d}t} = -\partial_{\mathbf{x}_p} H_\mathrm{E}, \tag{35}$$

$$\frac{\mathrm{d}\mathbf{x}_p}{\mathrm{d}t} = \partial_{\mathbf{p}_p} H_\mathrm{E}, \tag{36}$$

where the index p runs through all particles.

We rewrite Eqs. (35) and (36) more explicitly as

$$\frac{\mathrm{d}\mathbf{p}_p}{\mathrm{d}t} = -\frac{1}{4\pi} \int_V \mathbf{E} \cdot \partial_{\mathbf{x}_p} \mathbf{E} \, \mathrm{d}V, \tag{37}$$

$$\frac{\mathrm{d}\mathbf{x}_p}{\mathrm{d}t} = \frac{\mathbf{p}_p}{\gamma_p} = \mathbf{V}_p. \tag{38}$$

To deal with Eq. (37), one has to refer to Eq. (26) for the advance in time of the electric field. To get the correct expression for $\partial_{\mathbf{x}_p} \mathbf{E}$, let us displace the particle p by a small distance $\delta\mathbf{x}$. This displacement generates a current $\delta\mathbf{J}$ on the adjacent grid positions according to (24). The resulting change $\delta\mathbf{E}$ in the electric field is

$$\delta\mathbf{E} = -4\pi\delta\mathbf{J} \tag{39}$$

at the same grid positions. Hence, we may rewrite the first canonical equation (37) as

$$\frac{\mathrm{d}\mathbf{p}_p}{\mathrm{d}t} = \int_V \mathbf{E} \cdot \partial_{\mathbf{x}_p} \delta\mathbf{J} \, \mathrm{d}V. \tag{40}$$

The expression (40) has a very simple and clear meaning: to make the PIC code energy-conserving, one has to employ the same scheme for the electric field interpolation to the particle position as for the current deposition. Thus, the energy-conserving interpolation scheme for the $\mathbf{E}$-field is

$$
\begin{aligned}
E_x^p = W^\rho \big[& E_{i,\,j+1/2,\,k+1/2}^x a_y a_z + E_{i,\,j-1/2,\,k+1/2}^x (1-a_y)a_z \\
& + E_{i,\,j+1/2,\,k-1/2}^x a_y (1-a_z) + E_{i,\,j-1/2,\,k-1/2}^x (1-a_y)(1-a_z) \big], \\
E_y^p = W^\rho \big[& E_{i+1/2,\,j,\,k+1/2}^y a_x a_z + E_{i-1/2,\,j,\,k+1/2}^y (1-a_x)a_z \\
& + E_{i+1/2,\,j,\,k-1/2}^y a_x (1-a_z) + E_{i-1/2,\,j,\,k-1/2}^y (1-a_x)(1-a_z) \big], \\
E_z^p = W^\rho \big[& E_{i+1/2,\,j+1/2,\,k}^z a_y a_x + E_{i+1/2,\,j-1/2,\,k}^z (1-a_y)a_x \\
& + E_{i-1/2,\,j+1/2,\,k}^z a_y (1-a_x) + E_{i-1/2,\,j-1/2,\,k}^z (1-a_y)(1-a_x) \big],
\end{aligned}
\tag{41}
$$

It is important that the electric field is taken *at the present particle position* and not averaged along the trajectory. Also, the higher-order corrections b_α which we introduced for the current depositions are absent here. This is because Eq. (40) gives the analytical expression for the infinitesimal particle displacements, i.e. it has to be considered in the limit $|\delta\mathbf{x}| \to 0$.

6.1 Particle push

For advancing the particle in time one could then use the Boris scheme [4], with the electric field $\mathbf{E}$ interpolated according to (41) and the magnetic field $\mathbf{B}$ interpolated using a different scheme, to be discussed later. The Boris scheme is

$$\frac{\mathbf{p}_1 - \mathbf{p}_0}{\tau} = e\left(\mathbf{E} + \frac{1}{c}\frac{\mathbf{p}_1 + \mathbf{p}_0}{2\gamma_{1/2}} \times \mathbf{B} \right), \tag{42}$$

where $\mathbf{p}_0$ and $\mathbf{p}_1$ are the initial and the final particle momenta and $\gamma_{1/2}$ is the γ-factor taken at the middle of the time step. This scheme is time-reversible and *semi*-implicit. It can be analytically resolved for the final momentum $\mathbf{p}_1$ (see Ref. [4]):

$$\mathbf{p}^{n+1/2} = \mathbf{p}^- - e\,\mathbf{E}\,\frac{\tau}{2}, \tag{43}$$

$$\mathbf{p}^{n-1/2} = \mathbf{p}^+ + e\,\mathbf{E}\,\frac{\tau}{2}, \tag{44}$$

$$\frac{\mathbf{p}^+ - \mathbf{p}^-}{\tau} = \frac{q}{2\gamma mc}\,\mathbf{p}^+ + \mathbf{p}^- \times \mathbf{B}, \tag{45}$$

$$\mathbf{p}' = \mathbf{p}^- + \frac{q\tau}{2\gamma}\,\mathbf{p}^- \times \mathbf{B}, \tag{46}$$

$$\mathbf{p}^+ = \mathbf{p}^- + \frac{2}{1 + \left(\dfrac{q\tau B}{2\gamma}\right)^2}\,\mathbf{p}' \times \mathbf{B}. \tag{47}$$

The γ-factor should be calculated after step (43). However, as the magnetic field rotates the particle momentum, the Boris scheme is not exactly symmetric. An alternative scheme has been proposed recently; see Ref. [13]. We derive this alternative scheme below.

The equation of motion is discretized as

$$\frac{\mathbf{p} - \mathbf{p}_0}{\tau} = q\,\mathbf{E} + \frac{q}{2}\left(\frac{\mathbf{p}}{\gamma} + \frac{\mathbf{p}_0}{\gamma_0}\right) \times \mathbf{B}. \tag{48}$$

Here, $\mathbf{p}_0$ is the initial particle momentum and $\mathbf{p}$ is the particle momentum after the push with the corresponding γ-factors. We can rewrite this equation in the form

$$\mathbf{p} = \mathbf{a} + \frac{\mathbf{p}}{\gamma} \times \mathbf{b}, \tag{49}$$

where

$$\mathbf{a} = \mathbf{p}_0 + q\tau\mathbf{E} + \frac{q\tau}{2}\frac{\mathbf{p}_0}{\gamma_0} \times \mathbf{B} \tag{50}$$

and

$$\mathbf{b} = \frac{q\tau}{2}\mathbf{B}. \tag{51}$$

We rewrite (49) as

$$\gamma\mathbf{p} = \gamma\mathbf{a} + \mathbf{p} \times \mathbf{b}. \tag{52}$$

Taking the scalar product of (52) with $\mathbf{p}$ gives

$$p^2 = \mathbf{a} \cdot \mathbf{p}. \tag{53}$$

Taking the scalar product of (52) with $\mathbf{b}$ gives

$$\mathbf{b} \cdot \mathbf{p} = \mathbf{a} \cdot \mathbf{b}. \tag{54}$$

Taking the scalar product of (52) with $\mathbf{a}$ gives

$$\gamma\mathbf{a} \cdot \mathbf{p} - \gamma a^2 = \mathbf{a} \cdot (\mathbf{p} \times \mathbf{b}) = \mathbf{p} \cdot (\mathbf{b} \times \mathbf{a}) = \mathbf{b} \cdot (\mathbf{a} \times \mathbf{p}). \tag{55}$$

Taking the vector product of $\mathbf{a}$ and (52) gives

$$\gamma\mathbf{a} \times \mathbf{p} = \mathbf{p}\,(\mathbf{a} \cdot \mathbf{b}) - \mathbf{b}\,(\mathbf{a} \cdot \mathbf{p}). \tag{56}$$

Combining (53)–(56) gives

$$\gamma\left(p^2 - a^2\right) = \frac{\mathbf{b}}{\gamma}\left[\mathbf{p}(\mathbf{a}\cdot\mathbf{b}) - \mathbf{b}(\mathbf{a}\cdot\mathbf{p})\right] = \frac{\left[(\mathbf{a}\cdot\mathbf{b})^2 - b^2 p^2\right]}{\gamma} \tag{57}$$

or

$$\gamma^2\left(\gamma^2 - 1 - a^2\right) = (\mathbf{a}\cdot\mathbf{b})^2 - b^2\left(\gamma^2 - 1\right). \tag{58}$$

This leads to the quadratic equation

$$\gamma^4 + \gamma^2\left(b^2 - 1 - a^2\right) - b^2 - (\mathbf{a}\cdot\mathbf{b})^2 = 0, \tag{59}$$

which has the solution

$$\gamma^2 = \frac{1 + a^2 - b^2}{2} + \sqrt{\left(\frac{1 + a^2 - b^2}{2}\right)^2 + b^2 + (\mathbf{a}\cdot\mathbf{b})^2}. \tag{60}$$

Now we have to find the particle momentum after the push $\mathbf{p}$. To do this, we take the vector product of $\mathbf{b}$ and (52) to get

$$\gamma\left(\mathbf{b}\times\mathbf{p} - \mathbf{b}\times\mathbf{a}\right) = \mathbf{p}\cdot\mathbf{b} - \mathbf{b}\left(\mathbf{p}\cdot\mathbf{b}\right) = \mathbf{p}\,b^2 - \mathbf{b}\left(\mathbf{a}\cdot\mathbf{b}\right). \tag{61}$$

Using (52), we find that

$$\gamma^2\mathbf{a} - \gamma^2\mathbf{p} - \gamma\mathbf{b}\times\mathbf{a} = \mathbf{p}b^2 - \mathbf{b}\left(\mathbf{a}\cdot\mathbf{b}\right). \tag{62}$$

Solving for $\mathbf{p}$, we obtain

$$\mathbf{p} = \frac{\gamma^2\mathbf{a} + \gamma\mathbf{a}\times\mathbf{b} + \mathbf{b}(\mathbf{a}\cdot\mathbf{b})}{\gamma^2 + b^2}. \tag{63}$$

This pusher is fully implicit and does not require splitting of the Lorentz operator into the electric field push and the magnetic field rotation.

6.2   Energy conservation tests

The interpolation scheme consisting of (24) and (41) conserves the energy *exactly* for time steps which are small enough that the particle does not leave the original grid cell. If, however, motion of the particle becomes highly relativistic, the system will exhibit a slow energy growth. Notwithstanding this small drawback, we have solved one of the major problems with PIC codes. We may now simulate *cold* plasma. As 'cold' usually means non-relativistic 'temperatures', the energy is conserved and numerical heating is absent. If we do have a hot plasma, with temperatures close to relativistic ones, the method of *stochastic sampling* of the phase space becomes valid. Fortunately, the Debye length of such plasmas is many grid cells anyway, and numerical heating is not an issue.

The scheme (41) for electric field interpolation to the particle position is identical to the energy-conserving scheme used in electrostatic codes with the charge–potential (ρ–ϕ) formalism [4,5]. However, as we shall see later, there is a significant difference between these electrostatic codes and the field–current ($\mathbf{E}$–$\mathbf{J}$) formalism used in our electromagnetic simulations.

Figure 5 shows the evolution of total energy in an isolated system of particles for an energy-conserving (EC) algorithm (solid lines) and a 'momentum-conserving' (MC) [4] algorithm (dashed lines) in the following two cases: (a) warm plasma, with Debye length $D = 0.5\Delta x$; (b) cold plasma, with $D = 5 \times 10^{-3}\Delta x$. Although energy conservation for the EC algorithm is not exact, it is much better than in the MC case. The actual energy change is only 2% over 100 plasma oscillations for the EC algorithm; this property makes it possible to simulate a cold plasma.
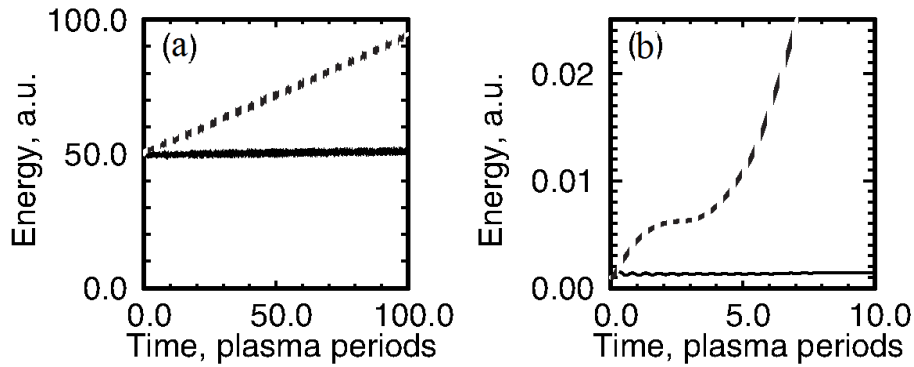
Fig. 5: Energy conservation in Virtual Laser Plasma Lab (VLPL) code based on an energy-conserving (EC) algorithm (solid lines) and numerical heating in a standard momentum-conserving (MC) algorithm (dashed lines), for the cases of (a) warm plasma, with Debye length $D = 0.5\Delta x$, and (b) cold plasma, with $D = 5 \cdot 10^{-3}\Delta x$. The energy change in the VLPL code is within 2% after 100 plasma periods.

7 Momentum and current conservation

It is known that the energy-conserving electrostatic PIC codes do not conserve momentum [4]. Indeed, it is easy to show that, strictly speaking, electric field interpolation to the particle position in the form of (41) does not conserve the total momentum:

$$\frac{d\mathbf{P}_{\text{total}}}{dt} = \sum_p q_p E_p \neq 0, \tag{64}$$

if the particles cross cell boundaries during their motion.

Now, the question is: how detrimental is this lack of momentum conservation to the PIC code?

Momentum non-conservation in electrostatic PIC codes using the ρ–ϕ formalism leads to some notable consequences [4]. If one starts the simulation with electrons drifting with respect to the resting ions, the electrons will experience an average *drag force* from the grid. As a result, after a few plasma periods, an initially regular electron drift becomes chaotic, and the simulation ends up with disordered hot electrons without any net motion with respect to the ions. Although the final energy of the system is preserved, and remains the same as the initial kinetic energy of the drifting electrons, the failure of momentum conservation is spectacular.

However, this spectacular example of momentum non-conservation in the one-dimensional electrostatic PIC code is rather an artefact of the ρ–ϕ formalism. Moreover, even the inital 'equilibrium' of electrons drifting with respect to ions is an artefact itself. Indeed, the original Maxwell equations (1) and (2) simply *do not allow for freely drifting electrons in the one-dimensional geometry*! This drift would correspond to a constant current, which results in a fast build-up of the longitudinal electric field. Consequently, electrons must oscillate around their initial positions at the local plasma frequency. This contradiction with the Maxwell equations has apparently remained unmentioned in regard to the ρ–ϕ formulation of the electrostatic code.

In the more realistic **E**–**J** formalism of electromagnetic codes, the ions have to drift *together* with the electrons, unless the forward electron current is compensated for by some artificial 'return' current, such as the longitudinal part of $\nabla \times \mathbf{B}$, which evidently does not exist in the one-dimensional geometry. A code using the **E**–**J** formalism conserves the net current:

$$\langle \mathbf{J} \rangle = \mathbf{J}_0 = \text{constant}, \tag{65}$$

where the averaging is done in time over the local plasma frequency. Indeed, the only possible deviations in the current are those due to the longitudinal part of the electric field, i.e. the *charge displacement* current.

The current conservation law (65) is a very important property, which may *compensate* for the absence of a detailed momentum conservation. As an example, let us consider the total electron current flowing in the simulation,

$$\mathbf{J}_e = \sum_p e W_p \mathbf{V}_p, \tag{66}$$

where e is the electron charge and W_p is the 'weight' of the numerical particle p. When averaged over the plasma period, the current (66) is conserved by our code. Now we may write the electron momentum in the non-relativistic case as

$$\mathbf{P}_e = \sum_p m_e W_p \mathbf{V}_p = \frac{e}{m_e} \mathbf{J}_e, \tag{67}$$

where m_e is the electron mass.

It follows from (67) that the total momentum is simply proportional to the total current, and so it is conserved on average. This is good news for the energy-conserving PIC code we have designed here: the code does conserve momentum on average in the non-relativistic case. When the particles are moving with relativistic energies, however, the identity (67) breaks down, and the momentum conservation is no longer ideal. Fortunately, the current conservation (66) still imposes a strong enough symmetry to prevent bad consequences such as those discussed in Ref. [4].

8 Maxwell solver: numerical dispersion-free scheme

Earlier, we discussed how to push particles and collect currents on the grid. Now we discuss the finite difference solver for the time-dependent Maxwell equations (1) and (2).

The standard way to propagate the fields on the Yee lattice (Fig. 3) is to use the centred conservative scheme [4]. Let us first consider the two-dimensional geometry for simplicity. In the 2D X–Y geometry, one may distinguish two kinds of polarization: s-polarizations with E_z, B_x and B_y fields, and p-polarizations with E_x, E_y and B_z fields. It can be shown that if the initial condition contains p-polarized fields and J_x and J_y currents only, then the s-polarized fields are not excited at all [4]. Thus, we may take the p-polarization as an example. The standard 2D scheme for the p-polarization is

$$B_{z\,i,j}^{n+1/2} - B_{z\,i,j}^{n-1/2} = \frac{c\tau}{\Delta y}(E_{x\,i,j+1/2}^n - E_{x\,i,j-1/2}^n) - \frac{c\tau}{\Delta x}(E_{y\,i+1/2,j}^n - E_{y\,i-1/2,j}^n), \tag{68}$$

$$E_{x\,i,j+1/2}^{n+1} - E_{x\,i,j+1/2}^n = \frac{c\tau}{\Delta y}(B_{z\,i,j+1}^{n+1/2} - B_{z\,i,j}^{n+1/2}) - 4\pi\tau j_{x\,i,j+1/2}^{n+1/2}, \tag{69}$$

$$E_{y\,i+1/2,j}^{n+1} - E_{y\,i+1/2,j}^n = \frac{c\tau}{\Delta x}(B_{z\,i+1,j}^{n+1/2} - B_{z\,i,j}^{n+1/2}) - 4\pi\tau j_{y\,i+1/2,j}^{n+1/2}. \tag{70}$$

The scheme (68)–(70) uses centred expressions for the finite difference $\nabla\times$ operators, such as

$$(\nabla \times \mathbf{B}_z)_x = \frac{1}{\Delta y}(B_{z\,i,j+1}^{n+1/2} - B_{z\,i,j}^{n+1/2}). \tag{71}$$

The Maxwell equations (1) and (2) and the corresponding scheme (68)–(70) are essentially linear partial differential equations where the only nonlinear source terms are in the form of the currents $\mathbf{j}$. In this case, we decide on the quality of the finite difference scheme (68)–(70) by comparing its dispersion properties with those of the Maxwell equations themselves.

According to the Maxwell equations, all electromagnetic waves in vacuum travel at the speed of light c. There is no dispersion in vacuum; not so for the finite differences. We Fourier-analyse the scheme

(68)–(70) by decomposing the fields in plane waves:

$$\mathbf{E} = \sum_{\mathbf{k}} \mathbf{E_k} \exp(-i\omega_{\mathbf{k}}t + i\mathbf{k}\cdot\mathbf{x}),$$

$$\mathbf{B} = \sum_{\mathbf{k}} \mathbf{B_k} \exp(-i\omega_{\mathbf{k}}t + i\mathbf{k}\cdot\mathbf{x}), \tag{72}$$

where $\mathbf{k}$ is the wavevector, $\omega_{\mathbf{k}}$ is the corresponding frequency, and $\mathbf{E_k}$ and $\mathbf{B_k}$ are the amplitudes of the Fourier harmonics. For the continuum Maxwell equations we have the simple dispersion relation

$$\omega_{\mathbf{k}} = c|\mathbf{k}|, \tag{73}$$

while the discretization in the finite difference scheme (68)–(70) introduces the numerical dispersion

$$\frac{1}{c^2\tau^2} \sin^2 \frac{\omega_{\mathbf{k}}\tau}{2} = \frac{1}{\Delta x^2} \sin^2 \frac{k_x\Delta x}{2} + \frac{1}{\Delta y^2} \sin^2 \frac{k_y\Delta y}{2}, \tag{74}$$

where Δx and Δy are the spatial grid steps and τ is the time step. The time step is limited by the Courant stability condition $c^2\tau^2 \le \Delta x^2\Delta y^2/(\Delta x^2 + \Delta y^2)$. If we run the code close to this limit of stability, then only the waves propagating along the grid diagonals are dispersion-free. The largest numerical dispersion is experienced by the waves running along the grid axes. This is illustrated in Fig. 6(a), where we plot the phase velocity of the numerical modes, $V_{\mathrm{ph}} = \omega_k/k$, for this standard scheme. We mention that we have chosen $\Delta y = 2\Delta x$ here, and this explains the apparent asymmetry of the plot.
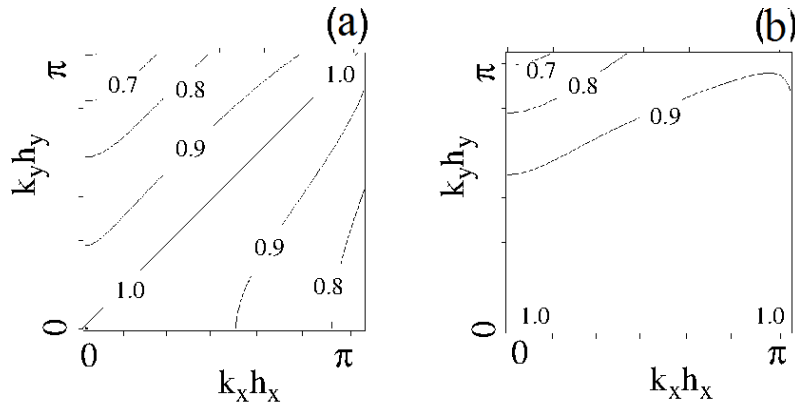
Fig. 6: Plots of the numerical phase velocity v_{ph}/c corresponding to (a) the standard scheme (74) and (b) the NDF scheme (80) used in the VLPL code. The grid cell aspect ratio is $\Delta x/\Delta y = 0.5$. The scheme (74) is dispersionless along the mesh diagonals $k_x = k_y$, while the NDF scheme (80) is dispersionless for waves travelling in the X-direction when $k_y = 0$.

Although numerical dispersion of the standard scheme may not be an issue when one is simulating dense, nearly critical plasma, it can cause trouble in simulations of very underdense plasmas, as it is important for particle acceleration [17]. In the dense plasma case, the plasma dispersion is usually stronger than the numerical dispersion, and so the problem is masked. In low-density plasma, however, the plasma dispersion is small; yet it has to be resolved accurately, as it influences the phase velocity of the laser pulse. As a consequence, one needs to use many grid cells per laser wavelength to obtain physically correct results. Of course, one could send the laser along one of the grid diagonals, but this is extremely inconvenient from the programming point of view.

We now aim at designing a superior numerical scheme which does not have numerical dispersion at all, or removes it to a large extent. Let us return to the finite difference expression for the curl operator,

(71). This centred operator can be written in a different way, using averages from the adjacent cells:

$$(\nabla \times \mathbf{B}_z)_x \rightarrow \frac{1}{2\Delta y}\big(B_{z_{i+1,j+1}}^{n+1/2} + B_{z_{i-1,j+1}}^{n+1/2} - B_{z_{i+1,j}}^{n+1/2} - B_{z_{i-1,j}}^{n+1/2}\big). \tag{75}$$

At first sight it is unclear what we have gained from averaging. It seems evident that the averaged scheme (75) may have only worse dispersion than the simpler scheme (71), but this is only partially true.

It turns out that one can choose a *linear combination* of the two schemes (71) and (75) in such a way that the dispersions of the two schemes *compensate for each other*! The resulting scheme for the p-polarization in the 2D geometry is

$$\begin{aligned}
B_{z_{i,j}}^{n+1/2} - B_{z_{i,j}}^{n-1/2} = \frac{c\tau}{\Delta y}&\Big[b_x\big(E_{x_{i,j+1/2}}^{n} - E_{x_{i,j-1/2}}^{n}\big) \\
&+ a_x\big(E_{x_{i+1,j+1/2}}^{n} - E_{x_{i+1,j-1/2}}^{n} + E_{x_{i-1,j+1/2}}^{n} - E_{x_{i-1,j-1/2}}^{n}\big)\Big] \\
- \frac{c\tau}{\Delta x}&\Big[b_y\big(E_{y_{i+1/2,j}}^{n} - E_{y_{i-1/2,j}}^{n}\big) \\
&+ a_y\big(E_{y_{i+1/2,j+1}}^{n} - E_{y_{i-1/2,j+1}}^{n} + E_{y_{i-1/2,j-1}}^{n} - E_{y_{i-1/2,j-1}}^{n}\big)\Big],
\end{aligned} \tag{76}$$

$$E_{x_{i,j+1/2}}^{n+1} - E_{x_{i,j+1/2}}^{n} = \frac{c\tau}{\Delta y}\big(B_{z_{i,j+1}}^{n+1/2} - B_{z_{i,j}}^{n+1/2}\big) - 4\pi\tau j_{x_{i,j+1/2}}^{n+1/2}, \tag{77}$$

$$E_{y_{i+1/2,j}}^{n+1} - E_{y_{i+1/2,j}}^{n} = \frac{c\tau}{\Delta x}\big(B_{z_{i+1,j}}^{n+1/2} - B_{z_{i,j}}^{n+1/2}\big) - 4\pi\tau j_{y_{i+1/2,j}}^{n+1/2}, \tag{78}$$

where the coefficients of the linear combination of the two schemes are

$$\begin{aligned}
a_x = a_y &= 0.125\frac{\Delta x}{\Delta y}, \\
b_x &= 1 - 2a_x, \\
b_y &= 1 - 2a_y,
\end{aligned} \tag{79}$$

and we have assumed that $\Delta x \leq \Delta y$.

The dispersion relation for numerical scheme (76)–(78) is immediately found to be

$$\begin{aligned}
\frac{1}{c^2\tau^2}\sin^2\frac{\omega\tau}{2} = \frac{1}{\Delta x^2}\sin^2\frac{k_x\Delta x}{2}(b_y + 2a_y\cos k_y\Delta y) \\
+ \frac{1}{\Delta y^2}\sin^2\frac{k_y\Delta y}{2}(b_x + 2a_x\cos k_x\Delta x).
\end{aligned} \tag{80}$$

It follows from (80) that the scheme is stable even at $c\tau = \Delta x$. This is quite a unique property for an explicit multi-dimensional finite difference scheme. In addition, the scheme (80) goes over to the usual Yee scheme in the limit $\Delta_x/\Delta_y \rightarrow 0$.

When used close to the stability limit, the scheme completely removes numerical dispersion along the X-axis (the laser propagation direction). For this reason we call this the NDF scheme, which stands for 'Numerical Dispersion- Free' [14]. The phase velocities of the numerical modes for the NDF scheme are plotted in Fig. 6(b). We mention that the region in which the numerical phase velocities are close to c becomes much wider than for the standard scheme; cf. Fig. 6(a).

The presence of the plasma changes the stability condition slightly, and the maximum τ is limited by the condition

$$1 - \frac{\tau}{\Delta x} > \frac{\omega_{\mathrm{p}}^2\tau^2}{4}, \tag{81}$$

where $\omega_{\mathrm{p}} = \sqrt{4\pi n_e e^2/m_e}$ is the maximum plasma frequency in the simulation domain. For an underdense plasma, however, this is an insignificant change.

The scheme (76)–(78) was written for the p-polarization in the 2D planar geometry. It must be slightly modified before it can be used in the full 3D space. The final version of the 3D NDF scheme is

$$
\begin{aligned}
B_{x_{i+1/2,j,k}}^{n+1/2} - B_{x_{i+1/2,j,k}}^{n-1/2} = & -\frac{c\tau}{\Delta y}\Big[b_z\big(E_{z_{i+1/2,j+1/2,k}}^{n} - E_{z_{i+1/2,j-1/2,k}}^{n}\big) \\
& + a_z\big(E_{z_{i+1/2,j+1/2,k+1}}^{n} - E_{z_{i+1/2,j-1/2,k+1}}^{n} \\
& + E_{z_{i+1/2,j+1/2,k-1}}^{n} - E_{z_{i+1/2,j-1/2,k-1}}^{n}\big)\Big] \\
& + \frac{c\tau}{\Delta z}\Big[b_y\big(E_{y_{i+1/2,j,k+1/2}}^{n} - E_{y_{i+1/2,j,k-1/2}}^{n}\big) \\
& + a_y\big(E_{y_{i+1/2,j+1,k+1/2}}^{n} - E_{y_{i+1/2,j+1,k-1/2}}^{n} \\
& + E_{y_{i+1/2,j-1,k+1/2}}^{n} - E_{y_{i+1/2,j-1,k-1/2}}^{n}\big)\Big],
\end{aligned}
\tag{82}
$$

$$
\begin{aligned}
B_{y_{i,j+1/2,k}}^{n+1/2} - B_{x_{i+1/2,j,k}}^{n-1/2} = & \frac{c\tau}{\Delta x}\Big[b_z\big(E_{z_{i+1/2,j+1/2,k}}^{n} - E_{z_{i-1/2,j+1/2,k}}^{n}\big) \\
& + a_z\big(E_{z_{i+1/2,j+1/2,k+1}}^{n} - E_{z_{i-1/2,j+1/2,k+1}}^{n} \\
& + E_{z_{i+1/2,j+1/2,k-1}}^{n} - E_{z_{i-1/2,j+1/2,k-1}}^{n}\big)\Big] \\
& - \frac{c\tau}{\Delta z}\Big[b_x\big(E_{x_{i,j+1/2,k+1/2}}^{n} - E_{x_{i,j+1/2,k-1/2}}^{n}\big) \\
& + a_x\big(E_{x_{i+1,j+1/2,k+1/2}}^{n} - E_{x_{i+1,j+1/2,k-1/2}}^{n} \\
& + E_{x_{i-1,j+1/2,k+1/2}}^{n} - E_{x_{i-1,j+1/2,k-1/2}}^{n}\big)\Big],
\end{aligned}
\tag{83}
$$

$$
\begin{aligned}
B_{z_{i,j,k+1/2}}^{n+1/2} - B_{z_{i,j,k+1/2}}^{n-1/2} = & \frac{c\tau}{\Delta y}\Big[b_x\big(E_{x_{i,j+1/2,k+1/2}}^{n} - E_{x_{i,j-1/2,k+1/2}}^{n}\big) \\
& + a_x\big(E_{x_{i+1,j+1/2,k+1/2}}^{n} - E_{x_{i+1,j-1/2,k+1/2}}^{n} \\
& + E_{x_{i-1,j+1/2,k+1/2}}^{n} - E_{x_{i-1,j-1/2,k+1/2}}^{n}\big)\Big] \\
& - \frac{c\tau}{\Delta x}\Big[b_y\big(E_{y_{i+1/2,j,k+1/2}}^{n} - E_{y_{i-1/2,j,k+1/2}}^{n}\big) \\
& + a_y\big(E_{y_{i+1/2,j+1,k+1/2}}^{n} - E_{y_{i-1/2,j+1,k+1/2}}^{n} \\
& + E_{y_{i-1/2,j-1,k+1/2}}^{n} - E_{y_{i-1/2,j-1,k+1/2}}^{n}\big)\Big],
\end{aligned}
\tag{84}
$$

$$
\begin{aligned}
E_{x_{i,j+1/2,k+1/2}}^{n+1} - E_{x_{i,j+1/2,k+1/2}}^{n} = & \frac{c\tau}{\Delta y}\Big[b_z\big(B_{z_{i,j+1,k+1/2}}^{n+1/2} - B_{z_{i,j,k+1/2}}^{n+1/2}\big) \\
& + a_z\big(B_{z_{i,j+1,k+3/2}}^{n+1/2} - B_{z_{i,j,k+3/2}}^{n+1/2} \\
& + B_{z_{i,j+1,k-1/2}}^{n+1/2} - B_{z_{i,j,k-1/2}}^{n+1/2}\big)\Big] \\
& - \frac{c\tau}{\Delta z}\Big[b_y\big(B_{y_{i,j+1/2,k+1}}^{n+1/2} - B_{y_{i,j+1/2,k}}^{n+1/2}\big) \\
& + a_y\big(B_{y_{i,j+3/2,k+1}}^{n+1/2} - B_{y_{i,j+3/2,k}}^{n+1/2} \\
& + B_{y_{i,j-1/2,k+1}}^{n+1/2} - B_{y_{i,j-1/2,k}}^{n+1/2}\big)\Big] \\
& - 4\pi\tau j_{x_{i,j+1/2,k+1/2}}^{n+1/2},
\end{aligned}
\tag{85}
$$

$$
\begin{aligned}
E_{y_{i+1/2,j,k+1/2}}^{n+1} - E_{y_{i+1/2,j,k+1/2}}^{n} = & -\frac{c\tau}{\Delta x}\Big[b_z\big(B_{z_{i+1,j,k+1/2}}^{n+1/2} - B_{z_{i,j,k+1/2}}^{n+1/2}\big) \\
& + a_z\big(B_{z_{i+1,j,k+3/2}}^{n+1/2} - B_{z_{i,j,k+3/2}}^{n+1/2} \\
& + B_{z_{i+1,j,k-1/2}}^{n+1/2} - B_{z_{i,j,k-1/2}}^{n+1/2}\big)\Big]
\end{aligned}
$$

$$+ \frac{c\tau}{\Delta z} \Big[b_x \big(B_{x_{i+1/2,j,k+1}}^{n+1/2} - B_{x_{i+1/2,j,k}}^{n+1/2} \big)$$
$$+ a_x \big(B_{x_{i+3/2,j,k+1}}^{n+1/2} - B_{x_{i+3/2,j,k}}^{n+1/2}$$
$$+ B_{x_{i-1/2,j,k+1}}^{n+1/2} - B_{x_{i-1/2,j,k}}^{n+1/2} \big) \Big]$$
$$- 4\pi\tau j_{y_{i+1/2,j,k+1/2}}^{n+1/2}, \tag{86}$$

$$E_{z_{i+1/2,j+1/2,k}}^{n+1} - E_{z_{i+1/2,j+1/2,k}}^{n} = \frac{c\tau}{\Delta x} \Big[b_y \big(B_{y_{i+1,j+1/2,k}}^{n+1/2} - B_{y_{i,j+1/2,k}}^{n+1/2} \big)$$
$$+ a_y \big(B_{y_{i+1,j+3/2,k}}^{n+1/2} - B_{y_{i,j+3/2,k}}^{n+1/2}$$
$$+ B_{y_{i+1,j-1/2,k}}^{n+1/2} - B_{y_{i,j-1/2,k}}^{n+1/2} \big) \Big]$$
$$- \frac{c\tau}{\Delta y} \Big[b_x \big(B_{x_{i+1/2,j+1,k}}^{n+1/2} - B_{x_{i+1/2,j,k}}^{n+1/2} \big)$$
$$+ a_x \big(B_{x_{i+3/2,j+1,k}}^{n+1/2} - B_{x_{i+3/2,j,k}}^{n+1/2}$$
$$+ B_{x_{i-1/2,j+1,k}}^{n+1/2} - B_{x_{i-1/2,j,k}}^{n+1/2} \big) \Big]$$
$$- 4\pi\tau j_{z_{i+1/2,j+1/2,k}}^{n+1/2}, \tag{87}$$

where we are using the following expressions for the free parameters a_α and b_α:

$$a_x = a_y + a_z,$$
$$a_y = 0.125 \frac{\Delta x}{\Delta y},$$
$$a_z = 0.125 \frac{\Delta x}{\Delta z}, \tag{88}$$
$$b_x = 1 - 2a_x,$$
$$b_x = 1 - 2a_y,$$
$$b_x = 1 - 2a_z.$$

Here we have chosen the coefficients (88) in such a way that the scheme is stable provided $\Delta x \leq \Delta y, \Delta z$, and numerical dispersion is removed for waves running along the X-axis.

9 Lorentz boost

Plasma-based particle acceleration is a multi-scale problem, and the scales are very disparate; see Fig. 7. The smallest scale is the laser wavelength λ in the case of laser-driven acceleration, or the plasma wavelength λ_p for beam-driven plasma wakefields. The laser wavelength is on the order of micrometres, while the plasma wavelength can range from tens of micrometres to millimetres. The medium scale is the driver length; it can be comparable to the plasma wavelength in the bubble [16] and blow-out [15] regimes, or be much greater when we are relying on self-modulation in the plasma [19–22]. The largest scale is the acceleration length, which can range from centimetres to hundreds of metres or even kilometres [17]. It is the discrepancy between the driver scale and the acceleration distance that makes the simulations rather expensive.

One possible way to bring the scales together is to change the reference frame from the laboratory frame to a frame that is co-moving with the driver. This is called the Lorentz boost technique [23]. Let us assume that we transform from the laboratory frame L into a frame R moving in the propagation direction of the driver. The relative velocity of the R-frame is $V = \beta c$, and its relativistic factor is $\gamma = 1/\sqrt{1 - \beta^2}$. Then the driver is Lorentz-stretched in the R-frame with factor $\gamma(1 + \beta)$, and the propagation length is compressed by the same factor; see Fig. 7. Thus, potentially, the Lorentz transformation allows us to
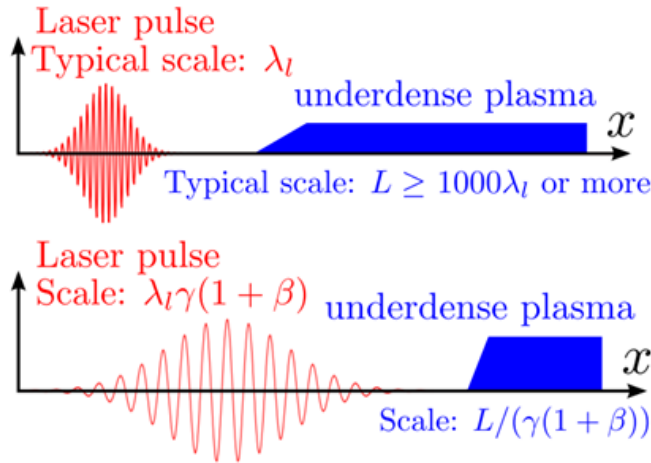
Fig. 7: Scale discrepancy in plasma-based acceleration; transformation to a co-moving frame reduces the discrepancy in scales.

increase the longitudinal grid step and time step—provided it is the grid step that limits the time step—and we have a smaller distance to propagate. The overall enhancement in performance could be huge.

Unfortunately, the background plasma becomes streaming in the R-frame with the same transformation velocity $-V$. The plasma density is higher in the R-frame than in the laboratory frame by a factor of γ. This leads to a source of free energy that can be converted into numerical plasma heating as the plasma particles interact with the numerical spatial grid. The associated numerical instability can have a considerable effect on the simulation quality [24]. The numerical noise generated by the unstable modes can completely mask the regular wake structure, as shown in the example in Fig. 8.
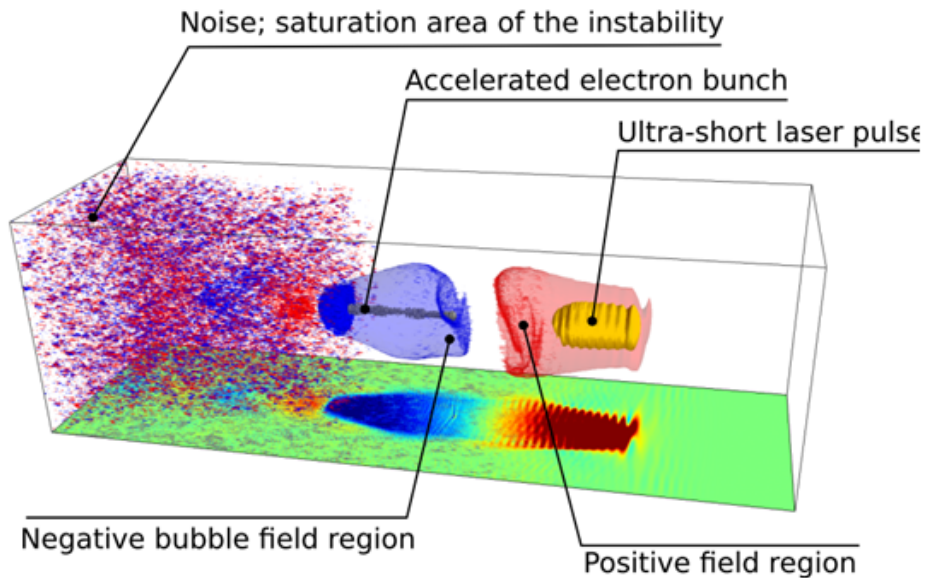
Fig. 8: Three-dimensional view of a numerical instability caused by plasma streaming in the co-moving frame

The main reason for the numerical instability is the Cerenkov resonance between the streaming plasma particles and the numerical electromagnetic modes on the grid. The numerical electromagnetic modes have subluminal phase velocities and can be in particle–wave resonance with the macroparticles

of the background plasma that stream through the grid with the relativistic factor γ. The mechanism of the numerical instability is illustrated in Fig. 9. Plasma fluctuations deflect particles from the straight-line trajectory; this leads to transverse currents. The transverse currents give rise to electromagnetic fields, and some of these electromagnetic modes propagate at exactly the particle velocity and can resonantly exchange energy with the particles. This is the Cerenkov mechanism.
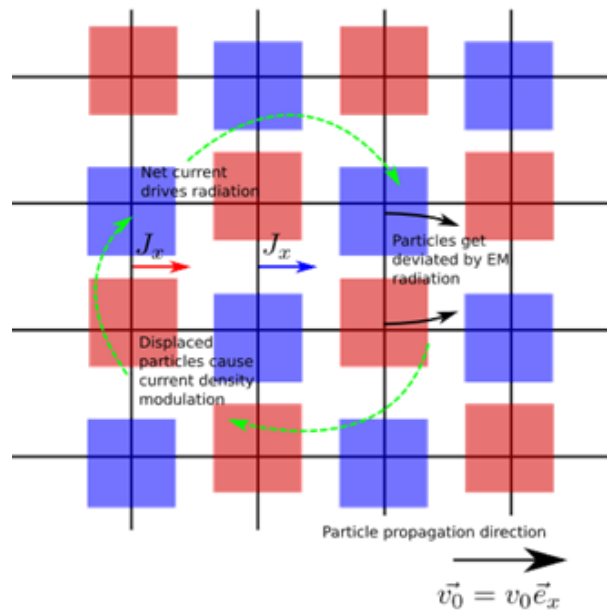
Fig. 9: The Cerenkov resonance mechanism underlying the numerical instability

One can solve the wave–particle dispersion relation for the standard Yee electromagnetic solver and calculate the growth rate of the instability analytically. A comparison of the observed electromagnetic modes in a PIC simulation and the analytical prediction is shown in Fig. 10. This tells us that, indeed, the reason behind the numerical instability is the Cerenkov resonance between relativistically moving plasma particles and numerical electromagnetic modes that have subluminal phase velocities on the Yee grid.
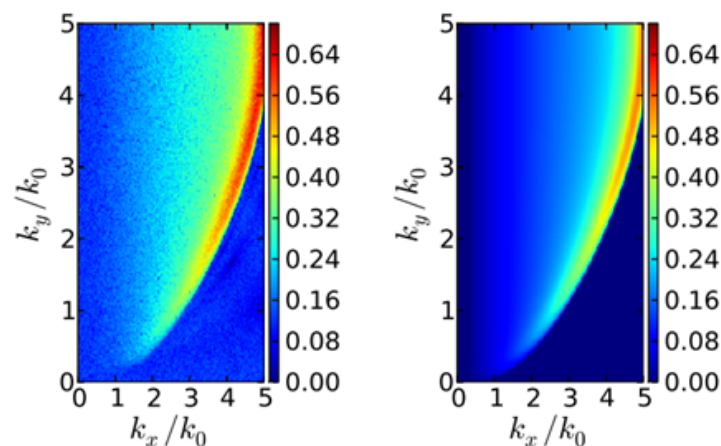
Fig. 10: Comparison of observed growing modes in a PIC simulation (left) and the analytically calculated growth rates of Cerenkov unstable modes (right).

One might hope to remove the instability by choosing a different Maxwell solver. For example, one could use a dispersion-free solver based on Fourier transformation. Indeed, a dispersion-free solver will reduce the instability growth rate, but unfortunately it is not able to eliminate the instability completely. The reason is the spatial and temporal aliasing on the grid; the relativistic particle starts to interact with the numerical modes from the other Brillouin zones. The non-resolved aliased frequencies of the numerical grid can be written as

$$\omega_{\text{eff}} = \pm \left(\sqrt{k_x^2 + k_y^2} - \frac{1}{\Delta t} \right), \tag{89}$$

where Δt is the time step. The resonance condition is then satisfied for electromagnetic waves with wavenumbers

$$k_y(k_x) = \frac{1}{h} \sqrt{h^2 k_x^2 \left(v_0^2 - 1 \right) - 2h k_x v_0^2 + v_0^2}, \tag{90}$$

where h is the grid step. The analytical resonance curve and the growing modes observed in a numerical experiment are shown in Fig. 11.
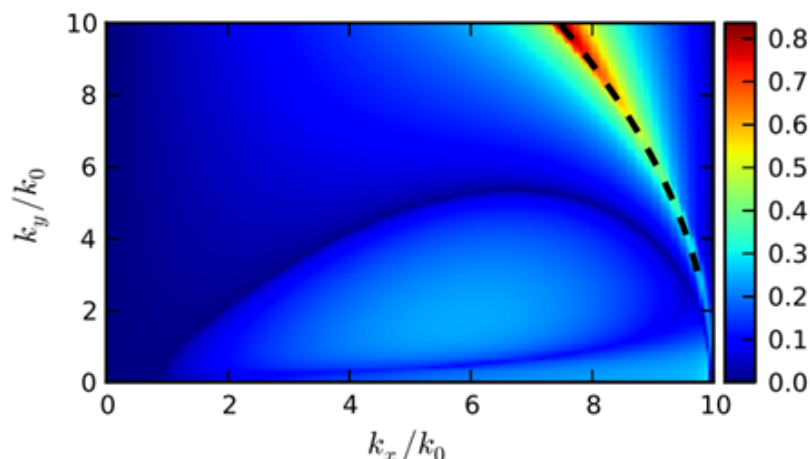
Fig. 11: Observed growing modes in a PIC simulation (colour scale) and the analytical resonance condition due to grid aliasing (dashed line).

The only viable way of taming the numerical instability in Lorentz-boosted simulations is to apply low-pass filters to the deposited currents before the electromagnetic fields are updated at each time step. The filtering reduces the instability to acceptable levels; yet a heavy filtering of the currents can in turn influence the dispersion of numerical modes, so one has to be very careful with this approach [25].

10 Quasi-static codes

Another way to bridge the gap between disparate scales in plasma-based acceleration is by using the quasi-static approximation. It separates explicitly the fast scale of the driver and the slow scale of acceleration [26]. To do so, we introduce new variables

$$\tau = t, \tag{91}$$
$$\zeta = z - ct. \tag{92}$$

We assume that the driver changes slowly as it passes a distance of its own length. Hence, as we are calculating plasma response to the driver, we neglect all derivatives with respect the slow time τ and advance from the front of the driver to the tail to calculate the wakefield configuration at a particular

time τ. After obtaining the fields and the plasma particle distribution, we can advance the driver with a large time step in τ. This procedure enhances the code's performance by many orders of magnitude. Simulations of large-scale plasma-based acceleration that require huge massively parallel computers when using the explicit PIC can be done on a desktop workstation in the quasi-static approximation. Of course, the quasi-static approximation is limited in that it does not describe radiation, only the static electromagnetic fields.

The first quasi-static PIC code, WAKE, was written by Mora and Antonsen [26]. It is a 2D code in cylindrical geometry and uses equations written in terms of the wake potential and the magnetic field. Later, a full 3D code, Quick-PIC, was developed that used the same equations [27]. Another 2D code in cylindrical geometry, called LCODE, has been developed by Lotov and uses equations on fields directly [28]. Here we describe the formalism used by the quasi-static version of the VLPL code. It is a full 3D code in Cartesian geometry. Like LCODE, it uses equations for the fields. Below we derive the quasi-static field equations.

We start again with the Maxwell equations (1)–(4) and write them in terms of the new variables (91) and (92), neglecting derivatives with respect to the slow time τ:

$$c\nabla \times \mathbf{B} = -c\frac{\partial \mathbf{E}}{\partial \zeta} + 4\pi\mathbf{j}, \tag{93}$$

$$c\frac{\partial \mathbf{B}}{\partial \zeta} = c\nabla \times \mathbf{E}, \tag{94}$$

$$\nabla \cdot \mathbf{E} = 4\pi\rho, \tag{95}$$

$$\nabla \cdot \mathbf{B} = 0. \tag{96}$$

First, we take the curl of the Ampère law (93) and the ζ-derivative of the Faraday law (94). Upon combining these two equations, we arrive at the first quasi-static equation on the magnetic field,

$$\nabla_\perp^2 \mathbf{B} = -\frac{4\pi}{c}\nabla \times \mathbf{j}, \tag{97}$$

where the $\nabla_\perp$ operator acts on coordinates transverse to the propagation direction.

Next, we take the gradient of the Poisson law (95). We use a well-known identity from vector analysis, $\nabla(\nabla \cdot \mathbf{E}) = \nabla^2\mathbf{E} + \nabla \times \nabla \times \mathbf{E}$, and obtain for the transverse components of the electric field the equation

$$\nabla_\perp^2 \mathbf{E}_\perp = 4\pi\left(\nabla_\perp\rho - \frac{1}{c}\partial_\zeta\mathbf{j}_\perp\right). \tag{98}$$

For the longitudinal electric field component we obtain

$$\nabla_\perp^2 E_\parallel = 4\pi\frac{\partial}{\partial \zeta}\left(\rho - \frac{1}{c}j_\parallel\right) = \frac{4\pi}{c}\nabla_\perp \cdot \mathbf{j}_\perp, \tag{99}$$

where we have also used the continuity equation

$$\frac{\partial}{\partial \zeta}\rho = \frac{\partial}{\partial \zeta}j_\parallel + \nabla_\perp \cdot \mathbf{j}_\perp. \tag{100}$$

The continuity equation (100) can be used to remove the charge density ρ from the quasi-static equations and so work with the currents only. This can help to reduce the noise in PIC codes.

A typical quasi-static PIC code works as follows (see illustration in Fig. 12). First, the charge density and currents generated by the driver on the numerical grid are gathered. These are the sources that contribute to the basic equations (97)–(99). Then, a layer of numerical macroparticles is seeded at the front boundary (the head of the driver) of the simulation box. These numerical particles advance in the negative ζ direction (towards the tail of the driver) according to Eqs. (97)–(99). As the plasma
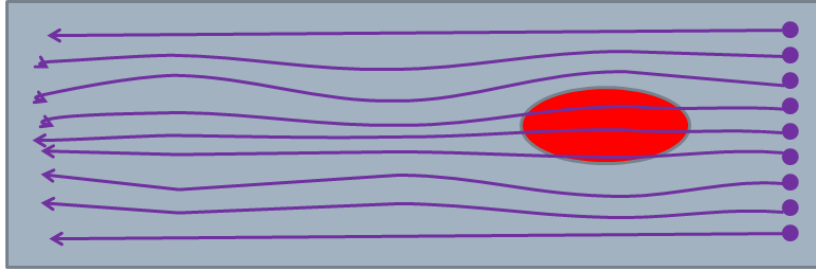
Fig. 12: Illustration of how the quasi-static PIC code is implemented

particles pass the whole simulation domain, the fields and density are defined on the grid and can be used to advance the driver in time τ.

If the driver is a charged particle beam, then we solve the equations of motion for beam particles in the calculated plasma fields. If the driver is a laser pulse, we have to solve an envelope equation on the laser pulse amplitude; this is required because the quasi-static equations do not describe radiation. Thus, an independent analytical model is required for the laser pulse. The laser pulse vector potential is represented as $\mathbf{A}(\tau, \zeta, \mathbf{r}) = \mathrm{Re}[\mathbf{a}(\tau, \zeta, \mathbf{r}) \exp(ik\zeta)]$. The envelope equation for the complex amplitude $\mathbf{a}(\tau, \zeta, \mathbf{r})$ reads [26]

$$\left[\frac{2}{c} \frac{\partial}{\partial \tau} \left(ik_0 + \frac{\partial}{\partial \zeta} \right) + \nabla_\perp^2 \right] \mathbf{a} = \chi(\zeta, \mathbf{r}), \tag{101}$$

where $\chi(\zeta, \mathbf{r}) = \langle 4\pi q^2 n / (\gamma m c^2) \rangle$ is the plasma refraction averaged over all the particles in the cell with charge q, mass m and relativistic factor γ.

The laser pulse acts on the plasma particles via its ponderomotive force

$$F_\mathrm{p} = -\frac{q^2}{\gamma m c^2} \nabla a^2. \tag{102}$$

The ponderomotive force (102) is added to the standard Lorentz force in the particle pusher.

A typical time step of a quasi-static code is shown in Fig. 13. First, there is a cycle that proceeds along the fast variable ζ for the low-frequency (LF) fields. Then, the plasma refraction is calculated and the envelope equation for the high-frequency (HF) fields is updated.
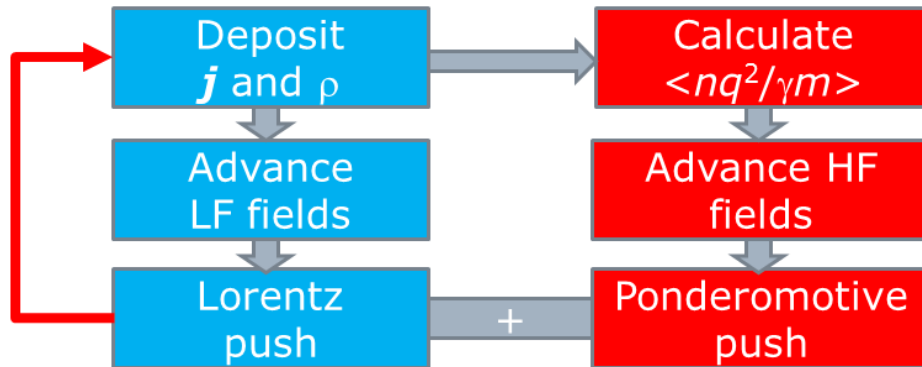
Fig. 13: Typical time step of a quasi-static code; the cycle for low-frequency (LF) plasma fields is followed by the cycle for the envelope equation of the high-frequency (HF) laser driver.

Figure 14 shows a comparison of the quasi-static code (upper half of the simulation frame) and the full PIC code VLPL3D (lower half). We have simulated a blow-out generated by an overdense electron

bunch. The electron bunch density had the profile $n_{\mathrm{b}}(z, \mathbf{r}) = n_{\mathrm{b}0} \exp(-z^2/2\sigma_z^2)\exp(-r^2/2\sigma_r^2)$. The maximum bunch density was two times higher than the background plasma density: $n_{\mathrm{b}0} = 2n_{\mathrm{p}}$. The bunch was spherical with $k_{\mathrm{p}}\sigma_z = k_{\mathrm{p}}\sigma_r = 1$.
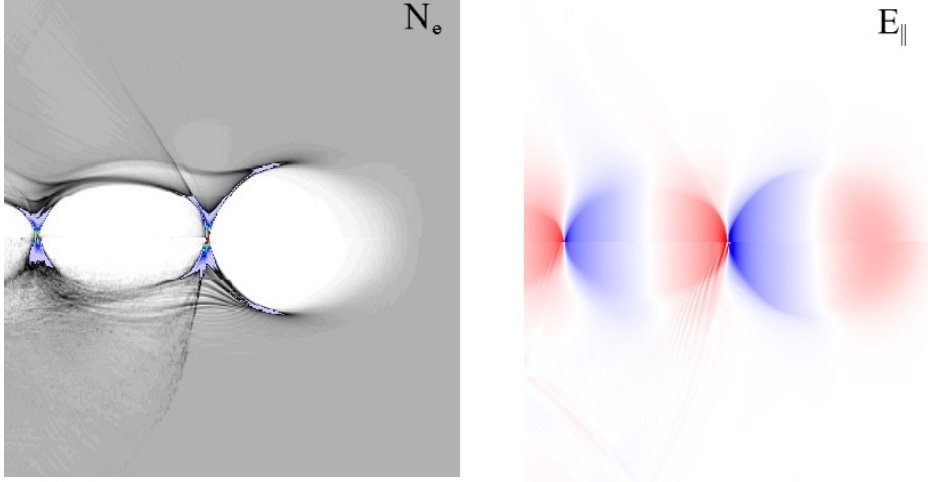
Fig. 14: Blow-out wakefield generated by an overdense electron bunch. The upper half of the frame is taken from a quasi-static simulation, and the lower part from a simulation using a full PIC code, VLPL3D. The left panel shows the plasma density, and the right panel shows the longitudinal electric field component of the wakefield. Cerenkov radiation emitted from the wavebreaking point is seen in the full 3D PIC simulation; this radiation is absent from the quasi-static code.

The two simulations are nearly identical. The only significant difference is the presence of Cerenkov radiation in the full 3D PIC simulation, emitted from the wavebreaking point at the tip of the tail of the first bubble. In the full simulation, we initialized the electron bunch vacuum in front of the plasma layer, so the plasma layer had to have a density ramp. The length of the bubble depends on the plasma density: the higher the density, the shorter the bubble. Consequently, the wavebreaking point moved with a superluminal velocity in the density ramp region and could emit Cerenkov radiation. In the quasi-static code, the Cerenkov radiation cannot be simulated. In addition, no density ramp is needed there: the field distribution is defined only by the local plasma density and by the instantaneous shape of the driver.

11 Computational costs of different codes

It is useful to estimate the number of operations required (called the computational cost) to simulate a particular plasma-based acceleration problem with different codes. The most general code, the full 3D PIC code, uses a 3D spatial grid of size $N_{\|} \times N_{\perp}^2$ cells and N_t time steps. The longitudinal step is limited by the laser wavelength, $h_{\|} \ll \lambda_0$, and the corresponding time step must be such that $\tau \ll \lambda_0/c$. The transverse grid steps are usually limited by the plasma wavelength, $h_{\perp} \ll \lambda_{\mathrm{p}}$. The number of time steps is $N_t = L_{\mathrm{acc}}/c\tau$, where L_{acc} is the acceleration distance. Hence, the number of operations required by the explicit PIC code scales as

$$N_{\mathrm{op}}^{\mathrm{PIC}} \propto \frac{L_{\mathrm{acc}}l_d}{\lambda_0^2} N_{\perp}^2. \tag{103}$$

Here we have assumed that the 'moving window' technique is used, so that the longitudinal size l_d of the simulation box scales with the plasma wavelength λ_{p}.

If one uses the Lorentz boost technique with the transformation relativistic factor γ_{boost}, the number of required operations reduces by a factor of $\gamma_{\mathrm{boost}}^2$ ideally:

$$N_{\mathrm{op}}^{\mathrm{PIC\text{-}LB}} \propto \gamma_{\mathrm{boost}}^{-2} \frac{L_{\mathrm{acc}}l_d}{\lambda_0^2} N_{\perp}^2 = \gamma_{\mathrm{boost}}^{-2} N_{\mathrm{op}}^{\mathrm{PIC}}. \tag{104}$$

The quasi-static approximation relaxes the time step restriction via the Courant condition. In addition, the grid step is no longer limited by the laser wavelength λ_0, but rather by the plasma wavelength λ_p. The time step must resolve the betatron oscillation of the beam particles, $\tau\omega_\beta \ll 1$, where the betatron frequency for a beam particle with mass M and relativistic factor γ_b is $\omega_\beta = \omega_p\sqrt{m_e/2\gamma_b M}$. If the driver is a laser pulse, then the time step must resolve the diffraction length, $c\tau/Z_R \ll 1$, where the characteristic diffraction length of a laser pulse with focal spot R is defined by the Rayleigh length $Z_R = \pi R^2/\lambda_0$. The overall number of operations required by the quasi-static code scales as

$$N_{op}^{QS} \propto \frac{L_{acc}\omega_\beta}{c}\frac{l_d}{\lambda_p}N_\perp^2 = \frac{\lambda_0^2}{\lambda_p\lambda_\beta}N_{op}^{PIC}, \tag{105}$$

where $\lambda_\beta = 2\pi c/\omega_\beta$ is the betatron wavelength. The performance gain, $\lambda_0^2/\lambda_p\lambda_\beta$, of the quasi-static code over the fully explicit PIC code can be huge and easily reach six orders of magnitude.

12 The future of PIC codes

Electromagnetic PIC codes provide a fundamental model for the dynamics of ideal plasma. Particularly in the relativistic regime of short-pulse laser–plasma interactions, these codes are unique in their predictive capabilities. In this regime, the binary collisions of plasma particles are either negligible or can be considered a small perturbation, and thus the electromagnetic PIC codes are the most appropriate tools.

However, the explicit PIC codes do have their limits. As soon as one tries to simulate laser interactions with highly overdense plasmas, these PIC codes become extremely expensive. Indeed, because the scheme is explicit, the code must resolve the plasma frequency and the skin depth. Even for uncompressed solid targets of high-Z materials, the plasma frequency can easily be 30 times higher than the laser frequency. The time and grid steps must be chosen accordingly. For a 3D code, simultaneous refinement of the grid and time steps in all dimensions by a factor of α leads to an increase in computational effort by a factor of α^4. For this reason, the simulation of a highly overdense plasma still poses a challenge for the explicit PIC codes.

A way around this difficulty might be to use implicit PIC codes, such the code LSP [29], or hybrid codes, such as a combination of a hydrodynamic description of the high-density background plasma and a PIC module for the hot electrons and ions [30]. Such codes alleviate the time step limitation, because they suppose the background plasma to be quasi-neutral and thus eliminate the fastest plasma oscillations at the Langmuir frequency. Very large plasma regions of high density can easily be simulated using such codes. Yet, these codes sacrifice a lot of the physics, and for any particular problem it must be checked whether the omitted physics is important or not. One possible way to perform this check is to benchmark the results of implicit codes against the direct PIC simulation on model problems which can be handled by both types of code.

Acknowledgements

References

[1] A. Pukhov, *Rep. Prog. Phys.* **66** (2001) 47. http://dx.doi.org/10.1088/0034-4885/66/1/202

[2] J. Villasenor and O. Buneman, *Comput. Phys. Commun.* **69** (1992) 306.
     http://dx.doi.org/10.1016/0010-4655(92)90169-Y

[3] J. Dawson, *Rev. Mod. Phys.* **55** (1983) 403. http://dx.doi.org/10.1103/RevModPhys.55.403

[4] C.K. Birdsall and A.B. Langdon, *Plasma Physics via Computer Simulations* (Adam Hilger, New York, 1991). http://dx.doi.org/10.1887/0750301171

[5] R.W. Hockney and J.W. Eastwood, *Computer Simulation Using Particles* (McGraw-Hill, London, 1981).

[6] J.D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1975).

[7] L. Landau and E. Lifshitz, *The Classical Theory of Fields*, 2nd ed. (Addison-Wesley, Reading, MA, 1962).

[8] N.A. Krall and A.W. Trivelpiece, *Principles of Plasma Physics* (McGraw-Hill, New York, 1973).

[9] S.I. Braginskii, Transport properties in a plasma, in *Reviews of Plasma Physics*, Ed. M.A. Leontovich (Consultants Bureau, New York, 1965).

[10] A.A. Vlasov, *Many-Particle Theory and Its Application to Plasma* (Gordon and Breach, New York, 1961).

[11] H. Ruhl and P. Mulser, *Phys. Lett.* **A205** (1995) 388; P. Bertrand, A. Ghizzo, T.W. Johnston, M. Shoucri, E. Fijalkov and M. R. Feix, *Phys. Fluids* **B5** (1990) 1028.

[12] R.C. Morse and C.W. Nielsen, *Phys. Fluids* **14** (1971) 830. http://dx.doi.org/10.1063/1.1693518

[13] R.H. Cohen, A. Friedman, D.P. Grote and J.L. Vay, *Nucl. Instrum. Methods Phys. Res.* **A606** (2008) 53. http://dx.doi.org/10.1016/j.nima.2009.03.083

[14] A. Pukhov, *J. Plasma Phys.* **61** (1999) 425. http://dx.doi.org/10.1017/S0022377899007515

[15] K. V. Lotov, *Phys. Rev.* E **69** (2004) 046405. http://dx.doi.org/10.1103/PhysRevE.69.046405

[16] A. Pukhov and J. Meyer-ter-Vehn, *Appl. Phys.* **B74** (2001) 355. http://dx.doi.org/10.1007/s003400200795

[17] E. Esarey, C.B. Schroeder and W.P. Leemans, *Rev. Mod. Phys.* **81** (2009) 1229. http://dx.doi.org/10.1103/RevModPhys.81.1229

[18] F.F. Chen, *Introduction to Plasma Physics and Controlled Fusion* (Plenum Press, New York, 1984). http://dx.doi.org/10.1007/978-1-4757-5595-4

[19] C. Joshi, T. Tajima, J.M. Dawson, H.A. Baldis and N.A. Ebrahim, *Phys. Rev. Lett.* **47** (1981) 1285. http://dx.doi.org/10.1103/PhysRevLett.47.1285

[20] N.E. Andreev, L.M. Gorbunov, V.I. Kirsanov, A.A. Pogosova and R.R. Ramazashvili, *Pis'ma Zh. Ehksp. Teor. Fiz.* **55** (1992) 551.

[21] T.M. Antonsen Jr. and P. Mora, *Phys. Rev. Lett.* **69** (1992) 2204. http://dx.doi.org/10.1103/PhysRevLett.69.2204

[22] A. Pukhov, N. Kumar, T. Tückmantel, A. Upadhyay, K. Lotov and P. Muggli, *Phys. Rev. Lett.* **107** (2011) 145003. http://dx.doi.org/10.1103/PhysRevLett.107.145003

[23] J.-L. Vay, *Phys. Rev. Lett.* **98** (2007) 130405. http://dx.doi.org/10.1103/PhysRevLett.98.130405

[24] J.-L. Vay, C.G.R. Geddes, E. Cormier-Michel and D.P. Grote, *J. Comput. Phys.* **230** (2011) 5908. http://dx.doi.org/10.1016/j.jcp.2011.04.003

[25] B.B. Godfrey and J.L. Vay, arXiv:1502.01387 (2015).

[26] P. Mora and T.M. Antonsen, *Phys. Plasmas* **4** (1997) 217. http://dx.doi.org/10.1063/1.872134

[27] C. Huang, V.K. Decyk, C. Ren, M. Zhou, W. Lu, W.B. Mori, J.H. Cooley, T.M. Antonsen Jr. and T. Katsouleas, *J. Comput. Phys.* **217** (2006) 658. http://dx.doi.org/10.1016/j.jcp.2006.01.039

[28] K.V. Lotov, *Phys. Rev. ST Accel. Beams* **6** (2003) 061301. http://dx.doi.org/10.1103/PhysRevSTAB.6.061301

[29] D.R. Welch, D.V. Rose, B.V. Oliver and R.E. Clark, *Nucl. Instrum. Methods Phys. Res.* **A464** (2001) 134. http://dx.doi.org/10.1016/S0168-9002(01)00024-9

[30] T. Tückmantel and A. Pukhov, *J. Comput. Phys.* **269** (2014) 168. http://dx.doi.org/10.1016/j.jcp.2014.03.019

Laser-driven Plasma Wakefield: Propagation Effects

B. Cros
Laboratoire de Physique des Gaz et des Plasmas, CNRS-Université Paris Sud, Orsay, France

Abstract
In the frame of laser-driven wakefield acceleration, the main characteristics of laser propagation and plasma wave excitation are described, with an emphasis on the role of propagation distance for electron acceleration. To optimize interaction length and maximize energy gain, operation at low plasma density is the most promising regime for achieving ultra-relativistic energies. Among the possible methods of extending propagation length at low plasma density, laser guiding by grazing incidence reflection at the wall of dielectric capillary tubes has several assets. The properties of laser guiding and the measurement of plasma waves over long distances are presented.

Keywords
Laser plasma acceleration; laser guiding; electron acceleration; plasma wave diagnostic.

1 Introduction

Propagation effects play an important role in laser-driven plasma accelerators [1]. The mechanism of plasma wave excitation relies on non-linear effects [2], linked to the presence of background electrons in the medium, driven by short and intense laser pulses [3]. Plasma waves result from the action of the ponderomotive force, which is proportional to the gradient of laser energy. This force expels electrons from the regions of higher intensity. Depending on the laser intensity, the density perturbation, or wakefield, remaining behind the laser pulse may be a periodic plasma wave, e.g., a sinusoidal oscillation [4], or cavities void of electrons, also known as bubbles [5]. Laser and plasma parameters need to be carefully selected to take advantage of non-linear effects.

The high intensity value, typically in the range 10^{17} to 10^{19} W/cm^2, required to drive a wakefield by laser is usually achieved by focusing the laser beam, within a small volume around the focus position, over a length shorter than or of the order of 1 mm. The longitudinal fields associated with plasma waves driven by laser can be as large as 100 GV/m, leading to accelerated electrons with energies of the order of 100 MeV over a typical scale length of 1 mm [6–8]. Acceleration of electrons to ultra-high energies requires that a high acceleration gradient be maintained over a longer distance [9].

Increasing the acceleration distance is one of the current challenges of laser plasma accelerators. Ultra-intense laser beam interaction with matter gives rise to various kinds of non-linear effect [3], which usually increase with propagation distance. Some intrinsic limitations of laser-driven wakefields, such as dephasing of electrons during the acceleration process as they overrun the accelerating phase of the field, or depletion of the laser beam over the propagation distance, can be less severe at low plasma densities. Lowering the plasma density contributes to a reduction in non-linear effects and an increase in acceleration distance, provided the laser intensity can be maintained at the required level over the whole distance. This can be achieved by guiding the laser beam externally, by either a preformed plasma structure or a capillary tube.

The first part of this paper outlines some of the main characteristics of laser-driven plasma wakefields, useful for describing laser propagation and plasma wave excitation. The main parameters governing electron acceleration are also described, with an emphasis on the role of propagation distance. Among the different methods of optimizing interaction length to maximize energy gain, operation at low

plasma density is the most promising regime towards achieving ultra-relativistic energies. Laser guiding by grazing incidence reflection at the wall of dielectric capillary tubes is described in the second part of this paper, from guiding properties to the measurement of plasma waves over long distances.

2 Laser plasma acceleration characteristics

2.1 Laser propagation in a vacuum

The electromagnetic field of a laser is described by Maxwell's equations [10]. The laser pulse can be modelled by Gaussian functions in space and time as a good approximation of experimental profiles. The electric field $\mathbf{E}$ of a bi-Gaussian laser beam, propagating in vacuum along the z axis, is given by:

$$
\begin{aligned}
\mathbf{E}(r, z, t) = & E_{\mathrm{L}} \frac{w_0}{w(z)} \exp\left[-\frac{r^2}{w^2(z)}\right] \exp\left[-2\ln(2)\frac{(z - ct)^2}{c^2 \tau_0^2}\right] \\
& \times \Re\left\{\exp\left[\mathrm{i}\omega_0 t - \mathrm{i}k_0 z - \mathrm{i}k_0 \frac{r^2}{2R(z)} + \mathrm{i}\psi_{\mathrm{g}}(z)\right] \mathbf{e}_\perp\right\} ,
\end{aligned}
\tag{1}
$$

where E_{L} is the amplitude of the electric field, w_0 is the waist or smallest laser transverse size in the focal plane ($z = 0$), c denotes the speed of light in a vacuum, τ_0 is the full width at half maximum of the pulse duration, and $k_0 = 2\pi/\lambda_0$ and $\omega_0 = ck_0$ represent, respectively, the wavenumber and angular frequency of a laser beam with wavelength λ_0. The unit vector $\mathbf{e}_\perp$ indicates the polarization direction of the laser electric field. For a laser field linearly polarized in the x direction, $\mathbf{e}_\perp = \mathbf{e}_x$, while for a circularly polarized one, $\mathbf{e}_\perp = 1/\sqrt{2}(\mathbf{e}_x \pm \mathrm{i}\mathbf{e}_y)$.

The propagation of the Gaussian laser pulse is fully characterized by the beam waist $w(z)$, the radius of curvature of the wavefront $R(z)$, and the Gouy phase shift $\psi_{\mathrm{g}}(z)$. As illustrated in Fig. 1, these parameters evolve along the z axis as

$$
\begin{aligned}
w(z) &= w_0\sqrt{1 + \left(\frac{z}{z_{\mathrm{R}}}\right)^2} , \\
R(z) &= z\left[1 + \left(\frac{z}{z_{\mathrm{R}}}\right)^2\right] , \\
\psi_g(z) &= \arctan\left(\frac{z}{z_{\mathrm{R}}}\right) ,
\end{aligned}
\tag{2}
$$

where $z_{\mathrm{R}} = \pi w_0^2/\lambda_0$ is the Rayleigh length, which represents the position at which the laser beam transverse area is doubled, compared with the one in the focal plane, owing to diffraction. The beam divergence far from the focal plane ($z \gg z_{\mathrm{R}}$) is given approximately by $\theta_{\mathrm{L}} \simeq \lambda_0/\pi w_0$.

The electromagnetic fields $\mathbf{E}$ and $\mathbf{B}$ can be expressed in terms of the scalar potential Φ and the vector potential $\mathbf{A}$,

$$
\begin{aligned}
\mathbf{E} &= -\nabla\Phi - \frac{\partial \mathbf{A}}{\partial t} , \\
\mathbf{B} &= \nabla \times \mathbf{A} ,
\end{aligned}
\tag{3}
$$

with the Coulomb gauge $\nabla \cdot \mathbf{A} = 0$. In a vacuum $\Phi = 0$, and both fields, $\mathbf{E}$ and $\mathbf{B}$, depend only on the vector potential $\mathbf{A}$. It is useful to define a normalized peak vector potential, also called the laser strength, $a_0 = eA_0/(mc) = eE_{\mathrm{L}}/(mc\omega_0)$.

The quantities usually measured during experiments are the laser energy, $\mathcal{E}_{\mathrm{L}}$, and its distribution in space, and the pulse duration, τ_0; they are used to evaluate the laser intensity, which is a key parameter for laser plasma interaction. The laser intensity is defined as

$$
I_{\mathrm{L}} = c^2\varepsilon_0 \langle \mathbf{E} \times \mathbf{B}\rangle = \frac{c\varepsilon_0}{2}|\mathbf{E}|^2 ,
\tag{4}
$$
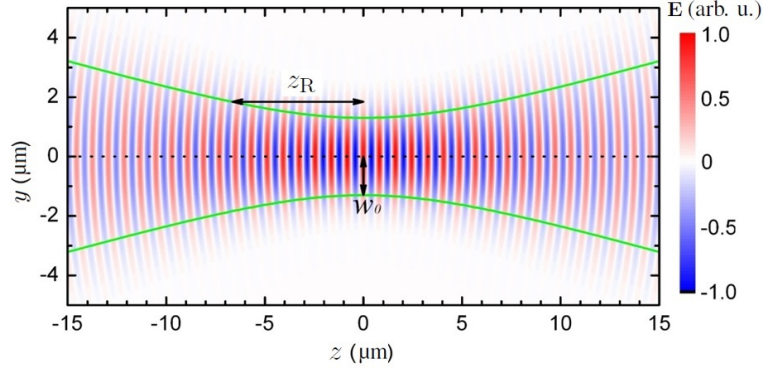
Fig. 1: Electric field of a Gaussian beam in the (y, z) plane; the focal plane is located at $z = 0$

where $\varepsilon_0 = 8.85 \times 10^{-12}$ F/m is the permittivity of free space. The laser power for a Gaussian pulse in time is written as

$$P = 2\sqrt{\frac{\ln(2)}{\pi}}\frac{\mathcal{E}_{\mathrm{L}}}{\tau_0} \simeq \frac{\mathcal{E}_{\mathrm{L}}}{\tau_0} \, , \tag{5}$$

and the corresponding peak laser intensity in the focal plane is

$$I_0 = \frac{2P}{\pi w_0^2} \simeq \frac{2\mathcal{E}_{\mathrm{L}}}{\pi \tau_0 w_0^2} \, . \tag{6}$$

Equation (6) shows that the peak intensity can be calculated from the measurements of energy, duration, and spot size, provided the shape of the laser beam is known. Substituting Eq. (4) into the expression $a_0 = eE_{\mathrm{L}}/(mc\omega_0)$, a_0 can be expressed as a function of the intensity in practical units,

$$a_0 = \sqrt{\frac{e^2}{2\pi^2 \epsilon_0 m_e^2 c^5} \lambda_0^2 I_0} \simeq 0.86 \lambda_0 \, [\, \mu\mathrm{m}] \, \sqrt{I_0 \, [10^{18} \, \mathrm{W/cm^2}]} \, . \tag{7}$$

An example of energy distribution in the focal plane, measured at the Lund Laser Centre during an experiment [11], is shown in Fig. 2. As is often the case, before focusing, the energy delivered by the laser system exhibits a nearly flat-top cylindrically symmetrical distribution in the transverse plane. Therefore, in the focal plane, the corresponding energy distribution is not purely Gaussian, and it can be seen from the right of Fig. 2 that it exhibits a profile close to an Airy pattern. The radial profile is obtained by averaging the energy distribution from the left-hand image over the azimuthal angle. The focal spot shown in Fig. 2 was achieved after optimization of the symmetry of the distribution, by tuning a deformable mirror placed after the compressor, to compensate for aberrations in the laser wavefront. The average radius of the focal spot at first minimum can be determined from the radial profile; in this example, it was measured as (19.7 ± 0.8) μm, which yields an on-axis peak intensity of $(5.4 \pm 0.1) \times 10^{18}$ W/cm^2 and a normalized laser vector potential $a_0 = 1.6$. The energy fraction contained within the grey shaded area in Fig. 2 is estimated to equal 84% of the energy in the focal plane.

2.2 Electron motion in a laser field

For currently used laser systems, the maximum intensity usually exceeds 10^{18} W/cm^2, corresponding to an electric field amplitude larger than 10^{12} V/m. As the laser field is transverse to the direction of propagation, single electrons mainly wiggle in this field, and it cannot be used directly to accelerate electrons in free space.
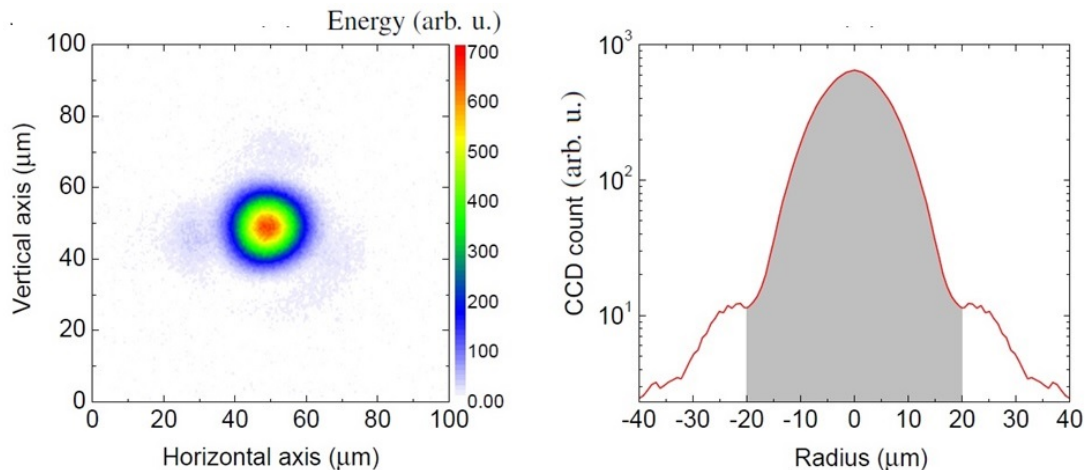
Fig. 2: Left: energy distribution in the focal plane. Right: corresponding averaged radial profile of laser energy in logarithmic scale; the grey shaded area, with a boundary at the first minimum of the focal spot, contains about 84% of laser energy in the focal plane.

The motion of a single electron with charge $-e$ and mass m_e in the laser fields $\mathbf{E}$ and $\mathbf{B}$ is described by the Lorentz equation,

$$\frac{d\mathbf{p}}{dt} = -e(\mathbf{E} + \mathbf{v} \times \mathbf{B}), \tag{8}$$

where $\mathbf{p} = \gamma m_e \mathbf{v}$ is the electron momentum, $\gamma = (1 - \beta^2)^{-1/2}$ is the relativistic factor, and $\beta = v/c$ denotes the normalized velocity. Approximating the laser field as a plane electromagnetic wave polarized along the x axis and propagating along the z axis, $\mathbf{E}(z) = E_L \cos(k_0 z - \omega_0 t)\mathbf{e}_x$, or in terms of vector potential using Eq. (3): $\mathbf{A}(z) = A_0 \sin(k_0 z - \omega_0 t)\mathbf{e}_x$, with $A_0 = E_L/\omega_0$. Taking into account $|\mathbf{B}| = |\mathbf{E}|/c$, the second term in the right-hand side of Eq. (8) can be neglected in the non-relativistic regime, when $\beta \ll 1$. Then Eq. (8) becomes simply

$$\frac{d\mathbf{p}}{dt} = -e\mathbf{E} = e\frac{\partial \mathbf{A}}{\partial t}. \tag{9}$$

Therefore, an electron, initially at rest at $z = 0$, oscillates in the direction of the electric field with a velocity

$$\beta = -\frac{eA_0}{mc} \sin(\omega_0 t) \triangleq -a_0 \sin(\omega_0 t). \tag{10}$$

It follows from Eq. (10) that an electron initially at rest will oscillate in the laser field with no net energy gain.

For $a_0 \gtrsim 1$, the electron oscillation velocity will approach c, and the $\mathbf{v} \times \mathbf{B}$ component in the Lorentz equation must be taken into account. The solution of Eq. (8) in the relativistic regime can be found, for example, in Ref. [12]. In the frame co-moving with the laser pulse, $\mathbf{a}(z) = a_0 \sin(k_0 \xi)\mathbf{e}_x$, where $\xi = z - ct$ is the coordinate in the frame co-moving with the laser. The normalized momentum of the electron can be written as

$$\begin{aligned}
u_x &= \gamma \beta_x = \frac{dx}{d\xi} = a = a_0 \sin(k_0 \xi), \\
u_z &= \gamma \beta_z = \frac{dz}{d\xi} = \frac{a^2}{2} = \frac{a_0^2}{2} \sin^2(k_0 \xi).
\end{aligned} \tag{11}$$
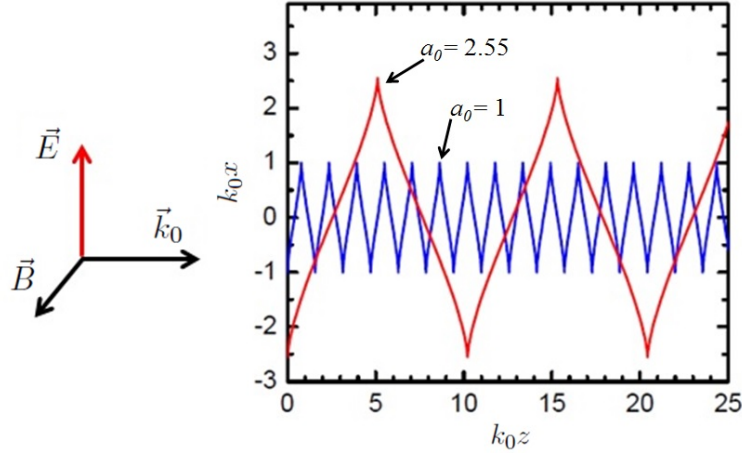
Fig. 3: Trajectory of an electron in the laser transverse field in the laboratory frame for two values of the laser strength.

The electron velocity is always positive in the z direction, so that the $\mathbf{v} \times \mathbf{B}$ force pushes the electron forward. Integrating Eq. (11) gives the electron coordinates along the trajectory:

$$x = -\frac{a_0}{k_0} \cos(k_0\xi) ,$$
$$z = \frac{a_0^2}{8k_0} [2k_0\xi - \sin(2k_0\xi)] . \tag{12}$$

The set of Eq. (12) indicates that the electron not only moves forward but also oscillates at twice the laser frequency in the longitudinal z direction. Figure 3 shows the electron trajectories for two values of a_0. The longitudinal momentum scales with the square of the laser strength as a_0^2, while the transverse one linearly depends on the laser strength by a_0. Hence, for $a_0 \gg 1$, the longitudinal motion of the electron dominates the transverse oscillation. The excursion distances along the x and z axis calculated over one period become equal for $a_0 = 8/\pi \simeq 2.55$. For a laser wavelength of 800 nm, $a_0 = 2.55$ corresponds to $I_0 = 1.4 \times 10^{19}$ W/cm^2.

This simple analysis shows the importance of operation in the relativistic regime, to achieve electron motion along the laser propagation direction. Nevertheless, as shown in Fig. 3, the electron is merely pushed forward during each light cycle but does not gain energy from the laser. The overall net effect is only to transfer the electron to a new longitudinal position. However, in this section, the laser beam was assumed to be a uniform and infinite plane wave. This is generally not the case in experiments, where laser pulses need to be tightly focused to achieve high intensity. The laser intensity is not uniform but Gaussian-like in the transverse plane. As shown next, the ponderomotive force associated with the laser intensity gradient excites a plasma wave, as the plasma plays the role of a transformer, to transfer laser energy to electrons.

2.3 Electromagnetic waves in plasmas

In a plasma, electrons oscillate around an equilibrium position with a characteristic frequency defined as the plasma frequency [2]

$$\omega_{\mathrm{p}} = \sqrt{\frac{n_{\mathrm{e}}e^2}{m_{\mathrm{e}}\varepsilon_0}} . \tag{13}$$

In the non-linear regime, the plasma frequency is modified by relativistic effects to $\omega_{\mathrm{pNL}} = \omega_{\mathrm{p}}/\sqrt{\gamma}$. The dispersion relation of an electromagnetic wave in a plasma can be written as

$$\omega_0^2 = \omega_{\mathrm{p}}^2 + c^2 k^2 \ . \tag{14}$$

For $\omega_0 > \omega_{\mathrm{p}}$, k is real and the wave can propagate in the plasma; k becomes imaginary for $\omega_0 < \omega_{\mathrm{p}}$, and the wave is evanescent. The light is thus either transmitted, damped, or reflected in the plasma, depending on the plasma density. The critical density, n_{c}, is defined as the density for which $\omega_0 = \omega_{\mathrm{p}}$, and can be written in practical units as

$$n_{\mathrm{c}} \left[10^{21} \ \mathrm{cm}^{-3} \right] = \frac{\omega_0^2 m_{\mathrm{e}} \varepsilon_0}{e^2} = \frac{1.12}{\lambda_0^2 \left[\, \mu\mathrm{m} \right]} \ . \tag{15}$$

The critical density corresponding to a laser wavelength of 0.8 μm is 1.75×10^{21} cm^{-3}. For $n_{\mathrm{e}} < n_{\mathrm{c}}$, the plasma is called underdense, and it is called overdense for $n_{\mathrm{e}} > n_{\mathrm{c}}$. The laser-driven wakefield relies on the excitation of a plasma wave in the underdense regime.

The phase and group velocities of the electromagnetic field in the plasma are calculated from Eq. (14):

$$\begin{aligned} v_{\mathrm{ph}} &= \frac{\omega_0}{k} = \sqrt{c^2 + \frac{\omega_{\mathrm{p}}^2}{k^2}} \ , \\ v_{\mathrm{g}} &= \frac{\mathrm{d}\omega_0}{\mathrm{d}k} = \frac{c^2}{v_{\mathrm{ph}}} = \frac{c^2}{\sqrt{c^2 + \omega_{\mathrm{p}}^2/k^2}} \ . \end{aligned} \tag{16}$$

If the evolution of the driving laser pulse in the plasma is not significant during propagation, the phase velocity of the plasma wave is equal to the group velocity of the driving laser [13]. The normalized phase velocity and relativistic factor of the plasma wave are then given by

$$\begin{aligned} \beta_{\mathrm{p}} &= \frac{v_{\mathrm{g}}}{c} = \sqrt{1 - \frac{n_{\mathrm{e}}}{n_{\mathrm{c}}}} \ , \\ \gamma_{\mathrm{p}} &= \frac{1}{\sqrt{1 - \beta_{\mathrm{p}}^2}} = \sqrt{\frac{n_{\mathrm{c}}}{n_{\mathrm{e}}}} \ . \end{aligned} \tag{17}$$

It can be seen that a lower density leads to a higher phase velocity and a larger relativistic factor.

2.4 Plasma wave excitation

In a plasma, the action of the ponderomotive force leads to the excitation of a plasma wave. This force is associated with the second-order electron motion [see Eq. (8)], averaged over a time-scale longer than the laser period. The three-dimensional (3D) ponderomotive force [14] for an electron can be written as

$$\mathbf{F}_{\mathrm{p}} = -m_{\mathrm{e}} c^2 \nabla \left\langle a^2/2 \right\rangle = -m_{\mathrm{e}} c^2 \nabla a_0^2/2 \ . \tag{18}$$

The ponderomotive force can be viewed as the radiation pressure of laser intensity. This force expels charged particles out of the region of high laser intensity, and does not depend on the sign of the charged particle. Furthermore, it is inversely proportional to particle mass $F_{\mathrm{p}} \propto 1/m$. Hence, under the same laser field, the acceleration exerted on a proton is only 10^{-6} times that exerted on an electron, so that ion motion can be neglected for sufficiently low laser strength.

Figure 4 illustrates the excitation of a plasma wave (red solid curve) in the linear (top graph) and non-linear (bottom graph) regimes, where the on-axis density exhibits localized spikes separated by a
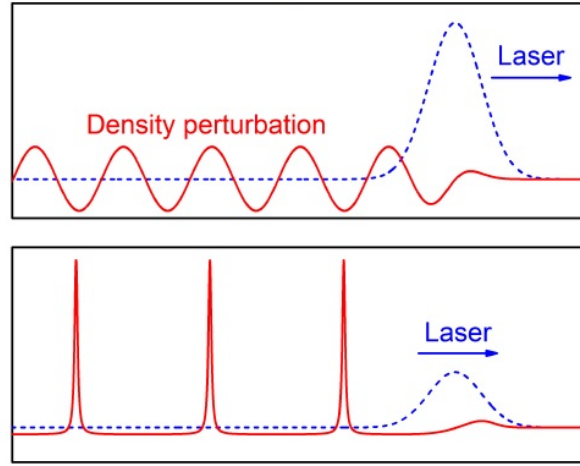
Fig. 4: Laser envelope (blue dashed line) and density perturbation (red solid line) along the axis of propagation in (top) the linear and (bottom) non-linear regimes.

distance longer than the linear wavelength; the laser pulse envelope is indicated by a blue dashed line. The longitudinal electric field associated with a plasma wave is thus a space charge field linked to the periodic distribution of charges oscillating behind the laser pulse. The oscillation length is the plasma wavelength linked to the electron density of the plasma, n_e, by the relation

$$\lambda_p \, [\mu m] \simeq 33 \times (n_e \, [10^{18} \, \text{cm}^{-3}])^{-1/2}. \tag{19}$$

This plasma wave is called a relativistic plasma wave because its phase velocity, given by Eq. (17), is of the order of the laser group velocity in the medium when the plasma frequency ω_p is much smaller than the laser frequency ω_0. The amplitude of the longitudinal accelerating field can be written as:

$$E_p \, [\, \text{GV/m}] \simeq 96(n_e \, [10^{18} \, \text{cm}^{-3}])^{1/2} \frac{\delta n_e}{n_e}. \tag{20}$$

The amplitude of this field is maximum for a density perturbation of 100%, at the upper limit of the linear regime, where the plasma wave breaks.

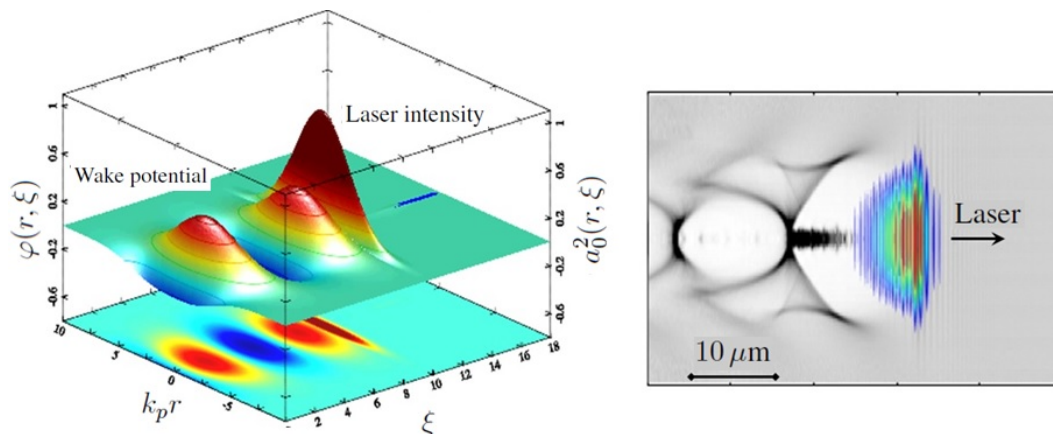### 2.5 Regimes of the laser wakefield

Among the parameters determining the characteristics of accelerated electrons in a plasma are the laser characteristics (amplitude, transverse, and longitudinal sizes) and plasma properties, such as the electron density and distribution. The laser strength a_0 is mainly used to distinguish between linear (or quasi-linear) and non-linear regimes of laser wakefield excitation. Nevertheless, as the non-linear evolution of the laser includes a deformation of the spatial volume delimiting the most intense fraction of the laser pulse, the initial transverse size of the laser, r_L, is also a key parameter. In the quasi-linear regime, $a_0 \simeq 1$ and

$$\frac{k_p^2 r_L^2}{2} > \frac{a_0^2}{\gamma_\perp} , \tag{21}$$

with $\gamma_\perp = (1 + a_0^2/2)^{1/2}$. The bubble or blow-out regime occurs for $a_0 > 1$ and is characterized by

$$k_p r_L \leq 2\sqrt{a_0} . \tag{22}$$

Figure 5 illustrate the main features of these two different regimes with simulation results. The left panel of Fig. 5 shows that the laser wakefield in the linear regime exhibits a regular oscillating behaviour [4].

Fig. 5: Illustration of the two regimes of the laser wakefield. Left: Excitation of a plasma wave in the linear regime; 3D view and projection in the horizontal plane of the normalized laser intensity, a_0, and wake potential, φ. Right: Excitation of a plasma wave in the non-linear regime; map of density in the horizontal plane (grey scale, white is zero) and superposition of laser amplitude (colour scale, red is a maximum).

The ponderomotive force varies as the laser energy gradient and creates a density distribution in the longitudinal and transverse directions. The associated transverse and longitudinal fields can be controlled independently by adjusting the focal spot transverse size and the pulse duration. The accelerating structure is shaped as a sine wave with wavelength λ_p, typically in the range 10–100 μm; its value is adjusted by tuning the plasma electron density. The accelerating field is typically in the range 1–10 GV/m, this value is limited to the wavebreaking field for a non-relativistic cold plasma, $E_0 = m_e c \omega_p / e$, and is of the order of 96 GV/m for a plasma electron density $n_e = 10^{18}$ cm^{-3}. Wavebreaking is characterized by the fact that electron oscillations become so large that the electrons can escape the collective motion. This may be at the origin of electron injection in the non-linear regime. In the linear regime, wavebreaking does not take place, and relativistic electrons need to be produced by an external source and injected into the linear plasma wave to be accelerated.

The right panel of Fig. 5 illustrates the main features of the non-linear regime. As the laser propagates into the plasma, its front edge diffracts while the rest of the pulse self-focuses. Electrons are expelled by the ponderomotive force from the high-intensity volume and a plasma cavity, the white area in Fig. 5, which is void of electrons, is left behind the laser pulse. The expelled electrons travel along the field lines and accumulate at the back of the cavity, where they can be injected and accelerated. The cavity size is typically of the order of 10 μm, and produces accelerated electron bunches with a transverse size of the order of a few micrometres, owing to the focusing field inside the cavity. This regime is also called the blow-out or bubble regime [5, 15].

2.6 Laser modulation in plasma

The optical properties of the plasma are modified during the excitation of a plasma wave, and can, in turn, modify the driving pulse. For the range of intensities of interest for a laser-driven wakefield, there are two important effects, self-focusing and self-compression, corresponding to changes in the transverse size and duration of the laser pulse during its propagation.

2.6.1 Self-focusing

The propagation of a laser pulse in a plasma can be investigated in terms of the spatial refractive index, $\eta(r) = c/v_{ph}$. Recalling the definition of v_{ph} in Eq. (16), the refractive index for an underdense plasma,

with uniform density n_{e0}, and a large amplitude plasma wave, $\omega_{\mathrm{p}}^2(r) = (\omega_{\mathrm{p}0}^2/\gamma)n_{\mathrm{e}}(r)/n_{\mathrm{e}0}$, is

$$\eta(r) = \frac{c}{v_{\mathrm{ph}}} = \left(1 - \frac{\omega_{\mathrm{p}}^2(r)}{\omega_0^2}\right)^{1/2} \simeq 1 - \frac{1}{2}\frac{n_{\mathrm{e}}(r)}{\gamma(r)n_{\mathrm{c}}} \ . \tag{23}$$

Equation (23) shows that the spatial profile of the refractive index $\eta(r)$ can be modified by the relativistic factor $\gamma(r)$ or the density distribution $n_{\mathrm{e}}(r)$. In the weakly relativistic case, Eq. (23) can be expanded as [3, 16]

$$\eta(r) = 1 - \frac{\omega_{\mathrm{p}}^2}{2\omega_0^2}\left(1 - \frac{a^2}{2} + \frac{\Delta n_e}{n_0} + \frac{\delta n_{\mathrm{e}}}{n_0}\right) \ , \tag{24}$$

where $\Delta n_{\mathrm{e}}/n_0$ takes into account the contribution of a preformed plasma channel along the radius; the $\delta n_{\mathrm{e}}/n_0$ term is responsible for plasma wave guiding, self-channelling, and the self-modulation of long pulses; the $a^2/2$ term corresponds to the contribution of relativistic laser guiding.

For a Gaussian pulse, with intensity peaked on-axis, $\partial a^2(r)/\partial r < 0$, which satisfies the condition for refractive guiding, $\partial \eta(r)/\partial r < 0$ or $\partial v_{\mathrm{ph}}(r)/\partial r > 0$. This implies that the on-axis phase velocity is less than the off-axis velocity, making the laser wavefront curved. The plasma thus plays the role of a convex lens, focusing the laser beam towards the propagation axis. This mechanism is known as self-focusing, and is able to balance laser diffraction. As shown in Eq. (24), the modulation of the refractive index leading to beam focusing scales with a^2. Therefore, for a given beam divergence, there is a minimum threshold for laser intensity to balance diffraction. Self-focusing will occur when the laser power P exceeds a critical power P_{c}, which can be written as [17–19]

$$P_{\mathrm{c}} = \frac{8\pi\varepsilon_0 m_{\mathrm{e}}^2 c^5 \omega_0^2}{e^2 \omega_{\mathrm{p}}^2} \simeq 17\frac{\omega_0^2}{\omega_{\mathrm{p}}^2} \ [\,\mathrm{GW}\,] \ . \tag{25}$$

For example, the critical power for a $\lambda = 0.8$ μm laser at plasma density $n_0 = 7 \times 10^{18}$ cm^{-3} is $P_{\mathrm{c}} = 4.25$ TW.

The radial ponderomotive force expels electrons from the axis, thus creating a radial density gradient $\partial(\delta n_{\mathrm{e}}(r))/\partial r > 0$, and a negative transverse gradient of the refractive index $\partial \eta(r)/\partial r < 0$, acting to focus the laser beam, although for $P \ll P_{\mathrm{c}}$, the ponderomotive force is not sufficient to guide the laser [13]. When the laser power approaches the critical power, relativistic self-guiding dominates laser propagation, while ponderomotive channelling enhances self-guiding. The contribution of ponderomotive channelling slightly relaxes the critical power for self-guiding to $P_{\mathrm{c}} = 16.8(\omega_0^2/\omega_p^2)\,[\,\mathrm{GW}\,]$.

2.6.2 Self-compression

In addition to the beam size evolution in the transverse direction, the variation of plasma density along the propagation axis modulates the laser pulse longitudinally. The density variation along the propagation axis $\delta n_{\mathrm{e}}(\xi)$ during the laser pulse makes it experience different local refraction indices, and consequently compress or stretch in the temporal domain.

Using Eq. (24), the local group velocity of the drive laser, or the phase velocity of the plasma wave, can be examined, taking into account the longitudinal dependence of η:

$$v_{\mathrm{g}}/c \simeq \eta \simeq 1 - \frac{\omega_{\mathrm{p}}^2}{2\omega_0^2}\left(1 - \frac{a^2}{2} + \frac{\delta n_{\mathrm{e}}}{n_0}\right) \ . \tag{26}$$

$\delta n_{\mathrm{e}} = 0$ gives the laser group velocity in the background plasma: $v_{\mathrm{g}0}$. For positive density variation $\delta n_{\mathrm{e}} > 0$, the laser group velocity decreases $v_{\mathrm{g}} - v_{\mathrm{g}0} < 0$, whereas $\delta n_{\mathrm{e}} < 0$ corresponds to an increase in laser group velocity: $v_{\mathrm{g}} - v_{\mathrm{g}0} > 0$. As a result of plasma wave excitation (see Fig. 4), the rear of the laser pulse moves faster than its front edge. The laser pulse is temporally compressed; this is also

called pulse shortening. This effect plays an important role in laser-driven wakefields as it contributes to matching the laser pulse duration and plasma wave period.

As a consequence of temporal compression, the spectral bandwidth of the short laser pulse will increase. Most of the laser pulse can be redshifted, while strong blueshifts occur at the back of the pulse for pulses longer than the plasma period [20]. As the spectral modulation of the driving laser is created by the plasma modulation, this provides a good means of diagnosing the excited plasma wave [21].

2.7 Energy gain of an electron in a plasma wave

The energy gain, ΔW, of an electron accelerated in a plasma wave is proportional to the product of the accelerating longitudinal field associated with the plasma wave, E_p, and the length, L_{acc}, over which the electron is submitted to this field,

$$\Delta W = eE_p L_{acc} . \tag{27}$$

The amplitude of the accelerating field and the length of acceleration depend on the regime of the laser wakefield and can be optimized in different ways. Three main mechanisms may limit the acceleration distance: laser diffraction typically limits acceleration to the Rayleigh length; pump depletion defines the length over which half of the laser energy is transferred to the plasma wave; and dephasing, in which accelerated electrons outrun the plasma wave and enter a decelerating phase, defines the dephasing length. Diffraction is a limitation due to laser propagation in the medium and can be overcome in the non-linear regime by self-focusing: in that case, the non-linearities imposed on the medium can shape the transverse density profile to act as a transient focusing length. In the linear regime, where these non-linear effects are negligible, external guiding has to be implemented; the use of capillary tubes for this purpose will be described in Section 3. Electron dephasing and laser depletion are intrinsic to laser plasma acceleration and depend on the electron density. Their dependences are described next.

2.7.1 Dephasing

As the phase of the plasma wave evolves at the group velocity of the laser in the plasma, a relativistic electron accelerated in this wave can explore different phases of the accelerating field. The dephasing length [13, 22], L_d, is defined as the length over which an electron can be accelerated before reaching a decelerating period of the electric field.

In the one-dimensional (1D) linear regime, L_d is evaluated as follows: in the co-moving frame, the maximum length over which the field is accelerating during one period is $\lambda_p/2$. Assuming that the accelerated electron propagates at approximately c, and keeping in mind that the plasma wave phase moves forward with velocity v_g, L_d^{1D} is estimated as

$$L_d^{1D} = \frac{\lambda_p}{2(c - v_g)}c = \frac{\lambda_p}{2(1 - \beta_p)} \simeq \frac{n_c}{n_e}\lambda_p = \frac{\omega_0^2}{\omega_p^2}\lambda_p . \tag{28}$$

This expression can be generalized to two-dimensional (2D), by noting that for a 2D plasma wave, there is only a quarter of the period ($\lambda_p/4$) in which the longitudinal electric field is accelerating and the radial field is focusing. Therefore, the dephasing length becomes

$$L_d^{2D} = L_d^{1D}/2 = \frac{\omega_0^2}{2\omega_p^2}\lambda_p . \tag{29}$$

In the 3D bubble regime, the distance in the co-moving frame for dephasing becomes the bubble radius R_b. The phase velocity of the plasma wave is modified to $\beta_p = 1 - 3\omega_p^2/(2\omega_0^2)$. Accordingly, the dephasing length L_d^{3D} is given by

$$L_d^{3D} = \frac{R_b}{1 - \beta_p} \simeq \frac{2}{3}\frac{\omega_0^2}{\omega_p^2}R_b = \frac{4}{3}\frac{\omega_0^2}{\omega_p^2}\frac{\sqrt{a_0}}{k_p} . \tag{30}$$
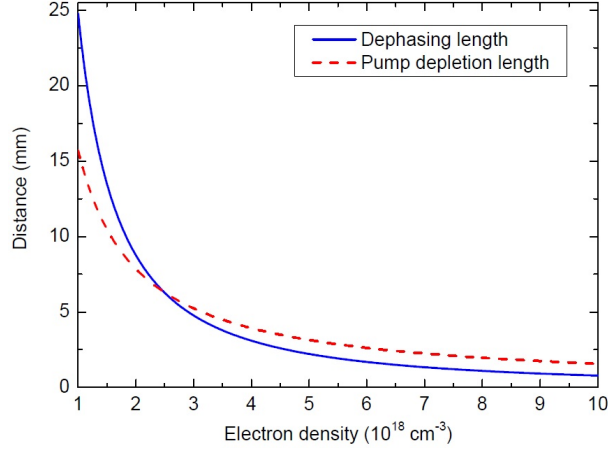
Fig. 6: 3D dephasing length and pump depletion length as functions of plasma electron density. The laser pulse has a duration of 30 fs and a peak intensity corresponding to $a_0 = 4$.

This expression shows that the 3D non-linear dephasing length depends on both the plasma electron density and the laser intensity. A longer $L_{\rm d}^{\rm 3D}$ can be achieved either decreasing the plasma density, as illustrated in Fig. 6, or by increasing the laser intensity.

2.7.2 Pump depletion

Another underlying limit is the so-called pump depletion, associated with the length $L_{\rm pd}$, over which the driving laser becomes depleted. As the laser pulse travels in the plasma, it transfers its energy to the plasma wave. The pump depletion length is defined as the distance over which the energy contained in the plasma wave equals that of the driving laser, $E_{\rm p}^2 L_{\rm pd} = \int E_{\rm L}^2 {\rm d}\xi$. For a Gaussian driving laser pulse, the linear pump depletion length is given by [23]

$$L_{\rm pd}^{\rm L} = \frac{\omega_0^2}{\omega_{\rm p}^2} \frac{c\tau_0}{a_0^2} \ . \tag{31}$$

In the non-linear regime, the pump depletion length is estimated via the etching velocity $v_{\rm etch} \simeq c\omega_{\rm p}^2/\omega_0^2$ [24]; this describes the erosion velocity of the laser front that excites the plasma wave, before the start of self-focusing. The laser will be depleted over the depletion length given by

$$L_{\rm pd}^{\rm NL} = \frac{c}{v_{\rm etch}} c\tau_0 = \frac{\omega_0^2}{\omega_{\rm p}^2} c\tau_0 \ . \tag{32}$$

Lu *et al.* [15] showed that this expression of non-linear pump depletion length $L_{\rm pd}^{\rm NL}$ is valid for 2D and 3D non-linear cases.

Both the 3D dephasing length $L_{\rm d}^{\rm 3D}$ and the pump depletion length $L_{\rm pd}^{\rm NL}$ are plotted as functions of plasma density, for $a_0 = 4$, in Fig. 6. Typically, the two lengths are of the order of a few millimetres in the density range above $n_{\rm e} = 3 \times 10^{18}$ cm^{-3}, where dephasing dominates pump depletion. The most efficient use of laser energy is achieved around the density where $L_{\rm d}^{\rm 3D} \simeq L_{\rm pd}^{\rm NL}$. Below $n_{\rm e} = 2.5 \times 10^{18}$ cm^{-3}, pump depletion occurs over a shorter length than dephasing, and both lengths increase quickly when the density becomes of the order of 1×10^{18} cm^{-3} .

2.7.3 Scaling laws

Scaling laws in the different regimes have been established from phenomenological considerations by Lu *et al.* [15], and are summarized in Table 1. These scaling laws play an important role in understanding

Table 1: Summary of scaling laws for linear and non-linear regimes [15], for the dephasing length L_{d}, the depletion length, L_{pd}, the relativistic factor of the plasma wave γ_{p}, and $\Delta\mathcal{E}$, the energy gain of an electron accelerated over the dephasing length.

	a_0	w_0	L_{d}	L_{pd}	γ_{p}	$\Delta\mathcal{E}/m_{\mathrm{e}}c^2$
Linear	<1	λ_{p}	$\frac{\omega_0^2}{\omega_{\mathrm{p}}^2}\lambda_{\mathrm{p}}$	$\frac{\omega_0^2}{\omega_{\mathrm{p}}^2}\frac{c\tau_0}{a_0^2}$	$\frac{\omega_0}{\omega_{\mathrm{p}}}$	$a_0^2\frac{\omega_0^2}{\omega_{\mathrm{p}}^2}$
1D non-linear	>1	λ_{p}	$4a_0^2\frac{\omega_0^2}{\omega_{\mathrm{p}}^2}\lambda_{\mathrm{p}}$	$\frac{1}{3}\frac{\omega_0^2}{\omega_{\mathrm{p}}^2}c\tau_0$	$\sqrt{a_0}\frac{\omega_0}{\omega_{\mathrm{p}}}$	$4a_0^2\frac{\omega_0^2}{\omega_{\mathrm{p}}^2}$
3D non-linear	>2	$\frac{\sqrt{a_0}}{\pi}\lambda_{\mathrm{p}}$	$\frac{4}{3}\frac{\omega_0^2}{\omega_{\mathrm{p}}^2}\frac{\sqrt{a_0}}{k_{\mathrm{p}}}$	$\frac{\omega_0^2}{\omega_{\mathrm{p}}^2}c\tau_0$	$\frac{1}{\sqrt{3}}\frac{\omega_0}{\omega_{\mathrm{p}}}$	$\frac{2}{3}a_0\frac{\omega_0^2}{\omega_{\mathrm{p}}^2}$

laser plasma acceleration mechanisms, and have been verified in numerous experiments. Scaling laws can be used, for example, to predict the energy gain in the non-linear regime. Using an average value of the accelerating field in the bubble regime, and an average value for acceleration over the dephasing length, the energy gain can be approximately written as

$$\Delta\mathcal{E} \simeq \frac{2}{3}a_0\frac{\omega_0^2}{\omega_{\mathrm{p}}^2}m_{\mathrm{e}}c^2 \simeq m_{\mathrm{e}}c^2\left(\frac{e^2P}{m_{\mathrm{e}}^2c^5}\right)^{1/3}\left(\frac{n_{\mathrm{c}}}{n_{\mathrm{e}}}\right)^{2/3}, \tag{33}$$

which becomes, in practical units,

$$\Delta\mathcal{E}\,[\,\mathrm{GeV}] \simeq 1.7\left(\frac{P\,[\mathrm{TW}]}{100}\right)^{1/3}\left(\frac{0.8}{\lambda_0\,[\mu\mathrm{m}]}\right)^{4/3}\left(\frac{1}{n_{\mathrm{e}}\,[10^{18}\,\mathrm{cm}^{-3}]}\right)^{2/3}. \tag{34}$$

This expression shows that the plasma electron density is a key parameter for tuning the electron bunch energy: a large energy gain over the dephasing length in the 3D non-linear regime can be achieved for low electron density and large laser power. Figure 7 shows the electron energy gain as a function of plasma density and laser power, calculated from Eq. (34). For a laser power of 100 TW, Eq. (34) predicts that electrons can be accelerated up to 1.7 GeV at a plasma density of $n_{\mathrm{e}} = 1 \times 10^{18}$ cm^{-3}. The energy gain displays a stronger dependence on plasma density than on laser power. Hence, to achieve higher electron energy, a lower electron plasma density is desirable. Nevertheless, one must always remember the laser size matching condition, which requires $w_0 = \lambda_{\mathrm{p}}\sqrt{a_0}/\pi$, which is proportional to $1/\sqrt{n_{\mathrm{e}}}$. When n_{e} is decreased, the laser power must be increased accordingly to meet the requirement for producing a larger bubble, as $P = I_0\pi w_0^2/2$, which is proportional to $1/n_{\mathrm{e}}$.

Numerous experimental results in the non-linear regime, where electrons from the plasma are self-injected into the accelerating structure, follow the scaling law for energy gain $\propto 1/n_{\mathrm{e}}$, as illustrated in Fig. 8. Figure 8 shows that when the density is divided by a factor of 10, from $n_{\mathrm{e}} = 8 \times 10^{18}$ cm^{-3} to $n_{\mathrm{e}} = 0.8 \times 10^{18}$ cm^{-3}, the maximum energy of the electron bunch is multiplied by the same factor. It should be noted that parameters other than the plasma density can change significantly between the different cases plotted in this graph: in particular, the laser power has been increased over the years and has contributed to the experimental observation of electrons in the gigaelectronvolt range for petawatt-class laser systems. The length of the plasma was also increased: starting from millimetre-scale gas jets at higher densities, gas cells of length between one and a few centimetres are used in intermediate regimes; the largest energy gain to date was obtained using external guiding by a plasma channel [9].

In conclusion, experimental results are well understood regarding the maximum electron energy dependence with electron plasma density and clearly indicate the way to increase the energy further. Reducing the density significantly increases the dephasing length: the next challenge will be to achieve metre-scale plasma lengths while preserving the laser intensity over the whole distance. Next, we examine some aspects of low-density operation of laser wakefield excitation over large distances.
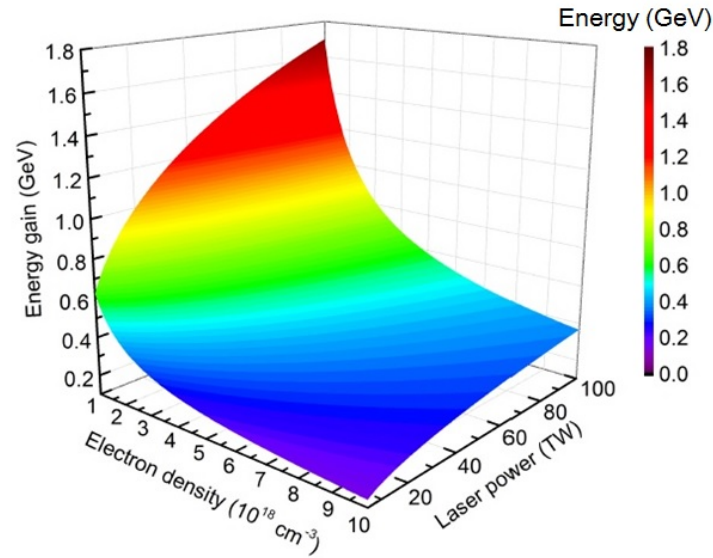
Fig. 7: Energy gain over the dephasing length as a function of plasma density and laser power, calculated using Eq. (34).
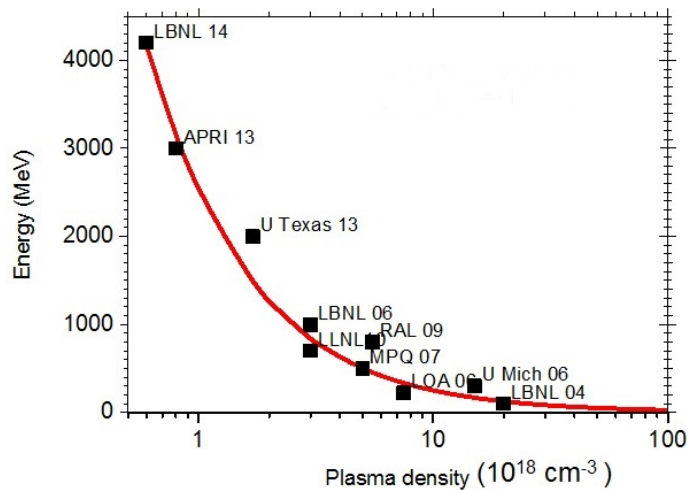

Fig. 8: Energy gain measured in various experiments around the world, plotted with black squares and labels, follow the dependence $\Delta W \propto 1/n$. Labels indicate institute acronyms and year of publication.

3 External guiding with capillary tubes

3.1 Choosing a guiding mechanism

Several methods have been developed to guide the focused laser over a distance longer than the Rayleigh length. Among them, self-focusing is one of the most commonly employed, owing to the simplicity of its implementation. Relying on this scheme, a powerful laser, $P_{\mathrm{L}} > P_{\mathrm{c}}$, can be guided over the pump depletion length, typically several millimetres long (see Fig. 6). To achieve self-focusing, the laser power must be in excess of the critical power, given by $P_{\mathrm{c}} \, \mathrm{GW} = 17\omega_0^2/\omega_{\mathrm{p}}^2$, which increases rapidly when the plasma density is decreased: for a plasma density $n_{\mathrm{e}} = 1 \times 10^{18} \, \mathrm{cm}^{-3}$, P_{c} equals 30 TW.
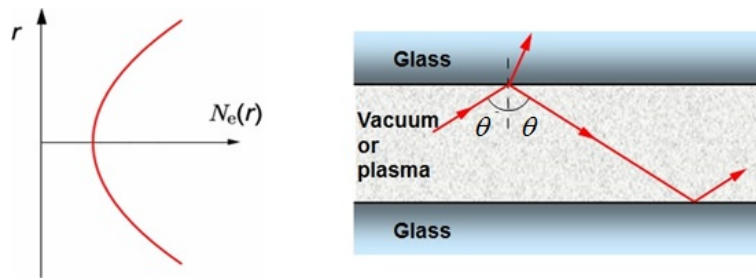
Fig. 9: Guiding methods: left, plasma channel, characterized by varying plasma density along the radius; right, capillary tube with glass wall, and laser reflection from the inner walls.

Self-focusing is the result of a modification of the plasma density by the front of the laser pulse, acting to create a density structure able to compensate diffraction for the rear of the pulse. A density structure acting to compensate diffraction (see Eq. (24)) can be formed prior to the intense laser interaction. A plasma channel can be created with the help of an external electrical discharge [25, 26] or by a heating laser pulse [27, 28]: the hot on-axis plasma electrons move outward, owing to radial hydrodynamic expansion, resulting in a density depletion on-axis and a nearly parabolic transverse density profile, as illustrated in the left panel of Fig. 9. A parabolic plasma channel can guide a Gaussian beam with a constant spot size $w_0 = w_m$, where w_m depends on the curvature of the channel. Under the matching condition, lasers with intensity of the order of 10^{18} W/cm^2 were successfully guided by plasma waveguides over many Rayleigh lengths [27, 29]. Good-quality guiding in plasma channels relies on fine tuning of the incident laser parameters (spot size, duration, energy) with channel parameters. As the guiding mechanism relies on plasma density and its spatial profile, this mechanism cannot be used for applications requiring plasma parameters that are very different from those required for laser guiding in the plasma channel.

When a capillary tube is used, as illustrated in the right panel of Fig. 9, the laser beam is guided by reflections at the inner capillary wall [30], and laser guiding can thus be achieved in vacuum or at low plasma density. This guiding scheme does not rely on laser power, or plasma density, and provides the opportunity to explore a large domain of laser plasma parameters. Laser guiding can, in principle, be achieved inside capillary tubes with total or partial reflection at the inner wall, depending on the material of the tube wall, which may be either a conductor or a dielectric material. Metallic tubes could be used to guide the laser beam without loss at the inner wall. In practice, their surface is usually not optically smooth for the laser wavelength and tube diameters used for laser guiding at high intensities. Dielectric capillaries, such as glass capillaries, are optically smooth and can be manufactured with a good reliability for a large range of inner diameters, wall thicknesses and lengths.

3.2   Eigenmodes of capillary tubes

Solving Maxwell's equations, with boundary conditions for dielectric surfaces at the capillary tube inner wall, gives hybrid mode solutions, with quasi-transverse electromagnetic modes. Solving the wave equation in cylindrical geometry, with boundary conditions describing the continuity of the field components at the boundary between the vacuum inside the tube and the dielectric wall, different families of eigenmodes of capillary tubes are obtained [30]. An incident linearly polarized Gaussian laser beam can be efficiently coupled to the linearly polarized family of hybrid modes, namely the EH_{1m} modes. The transverse electric components of the EH_{1m} modes at zero order can be found in Ref. [30]. For the EH_{1m} modes, the transverse electric field amplitude inside the capillary tube can be written as

$$E_{1m}(r, z, t) = J_0(k_{\perp m} r) \exp(-k_m^l z) \cos(\omega_0 t - k_{zm} z) \, , \qquad (35)$$
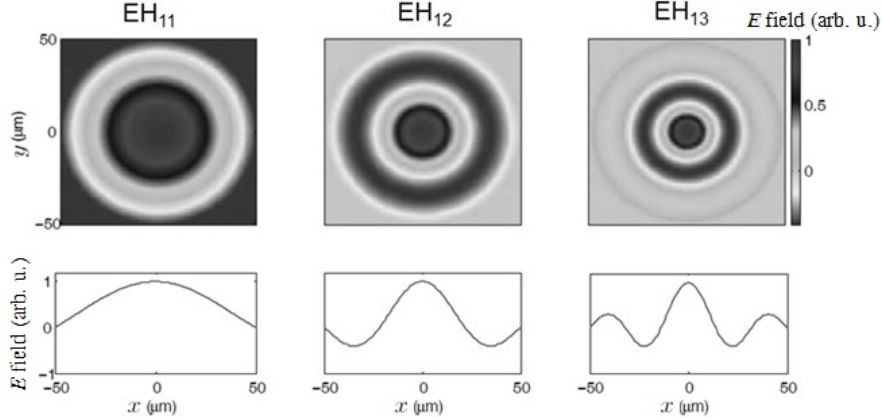
Fig. 10: Field distribution in the transverse plane for the three first hybrid modes EH_{1m}, plotted for $R_{\text{cap}} = 50\,\mu m$; bottom graphs show the corresponding distributions through $y = 0$.

where $k_{\perp m}$ is the transverse wavenumber of the mode with order m, defined as $k_{\perp m} = (k_0^2 - k_{zm}^2)^{1/2}$, and is given by $k_{\perp m} = u_m/R_{\text{cap}}$, R_{cap} is the capillary tube inner radius, u_m is the mth root of $J_0(x) = 0$, and J_0 is the Bessel function of integer order. Table 2 gives the first nine values of u_m. k_{zm} is the longitudinal wavenumber inside the tube, and k_0 is the wavenumber of the laser beam in free space. Figure 10 illustrates the electric field amplitude in the transverse plane for the three first modes. The electric field of the EH_{1m} modes is maximum at the tube axis; for $m = 1$, the field variation is close to the field of a Gaussian beam.

The exponential decay term in Eq. (35) indicates that the electric field is damped along the direction of propagation, z: each reflection at the dielectric capillary wall is associated with a refracted fraction of the laser beam inside the dielectric wall. This refracted fraction is minimum for the smallest perpendicular wavenumber, corresponding to the grazing incidence. The characteristic damping coefficient k_m^l is given by [30]:

$$k_m^l = \frac{u_m^2}{2k_{z0}^2 R_{\text{cap}}^3} \frac{1 + \varepsilon_{\text{r}}}{\sqrt{\varepsilon_{\text{r}} - 1}}\,, \tag{36}$$

where ε_{r} is the dielectric constant of the wall. This shows that k_m^l strongly depends on the capillary radius, R_{cap}, the wavelength of the incident laser beam, and the mode order, through u_m. Laser damping is usually described by the attenuation length L_m^l, defined as

$$L_m^l = \frac{1}{k_m^l} = \frac{2k_{z0}^2 R_{\text{cap}}^3}{u_m^2} \frac{\sqrt{\varepsilon_{\text{r}} - 1}}{1 + \varepsilon_{\text{r}}}. \tag{37}$$

After a propagation distance of L_m^l, the field magnitude decreases by a factor $1/e$ and the beam energy by a factor $1/e^2$, owing to refraction losses. For example, the values of damping length L_m^l for a Ti:sapphire laser ($\lambda_0 = 800$ nm) guided inside a 50 μm radius capillary tube are given in Table 2 for the nine first modes. L_m^l drops rapidly with increasing mode order, which means that higher-order modes are damped over shorter distances. Therefore, the use of the fundamental mode is preferable, to achieve laser guiding over a long propagation distance.

The group velocity is determined from the dispersion relation of an electromagnetic wave in a capillary tube, $k_0^2 = (k_z^2 + k_{\perp m}^2)$, for $k_{\perp m}^2 \ll k_0^2$, as

$$v_{\text{gm}} \simeq c \left(1 - \frac{k_{\perp m}^2}{k_0^2}\right)^{1/2}. \tag{38}$$

Table 2: Values of u_m, L_m^l, and v_{gm} for the first nine modes of a capillary: radius, 50 μm; laser wavelength, $\lambda_0 = 800$ nm.

m	u_m	L_m^l(cm)	v_{gm}/c	$\mathcal{F}_m^{\max}$ (10^{-5})
1	2.404826	91.7	0.99998	2.034
2	5.520078	17.4	0.9999	4.604
3	8.653728	7.1	0.9998	7.2011
4	11.79153	3.8	0.9995	9.8049
5	14.93092	2.4	0.9993	12.4113
6	18.07106	1.6	0.9989	15.0189
7	21.21164	1.2	0.9985	17.6272
8	24.35247	0.9	0.9981	20.2359
9	27.49348	0.7	0.9975	22.8449

The fourth column in Table 2 shows the values of v_{gm}. It can be seen that the group velocity is close to the velocity of light in free space, and decreases as the mode order increases. This again supports the use a lower-order mode with higher group velocity, corresponding to a higher phase velocity for the wakefield.

Another important issue associated with the use of a capillary tube is the threshold of material damage at the inner wall where reflection occurs, which determines the ability of capillary tubes to guide intense lasers. To examine this, we define the normalized flux at the capillary inner wall, $\mathcal{F}_m$, as the ratio of the radial component of the Poynting vector at $r = R_{\text{cap}}$ to the longitudinal component of the on-axis Poynting vector; it is given by [30]

$$\mathcal{F}_m = \frac{k_{\perp m}^2}{k_0^2} J_1^2(k_{\perp m} R_{\text{cap}}) \frac{\cos^2 \theta + \varepsilon_r \sin^2 \theta}{\sqrt{\varepsilon_r - 1}} \ . \tag{39}$$

$\mathcal{F}_m$ depends on the azimuthal angle θ and mode order m. $\mathcal{F}_m$ is minimum for $\theta = 0$, π and maximum for $\theta = \pi/2$, $3\pi/2$. It also depends on the mode eigenvalue and capillary radius through $k_{\perp m}^2$. Its maximum value should be below the threshold of material breakdown, to ensure laser guiding without non-linear interaction at the wall. As shown in Table 2, $\mathcal{F}_m^{\max}$ increases by one order of magnitude from the fundamental mode to the ninth mode. Once again, it emphasizes the advantage of using the fundamental mode. For a glass capillary, the ionization threshold is of the order of 10^{14} W/cm^2 for an 800 nm laser pulse with duration shorter than or of the order of 100 fs [31]. The maximum intensity on-axis guided by a capillary of 50 μm radius on the fundamental EH$_{11}$ mode without wall ionization is thus of the order of 10^{19} W/cm^2.

3.3   Mode coupling

When a laser beam is focused at the entrance of a capillary tube, its energy has to be coupled to the capillary eigenmodes before propagation. In this section, we describe the conditions of coupling for two kinds of laser energy distribution in the focal plane, located at the entrance of the capillary tube: a Gaussian beam and an Airy beam, both of which can be good approximations to describe the beams used in experiments. As shown in the previous section, monomode guiding with the fundamental EH$_{11}$ mode has interesting properties for laser wakefield excitation, so the condition for monomode coupling will be particularly discussed.

3.3.1   Coupling of a Gaussian beam

Laser beams can be described as transverse electromagnetic modes with a transverse Gaussian envelope. Assuming a linearly polarized Gaussian beam focused at the capillary entrance ($z = 0$), the amplitude of
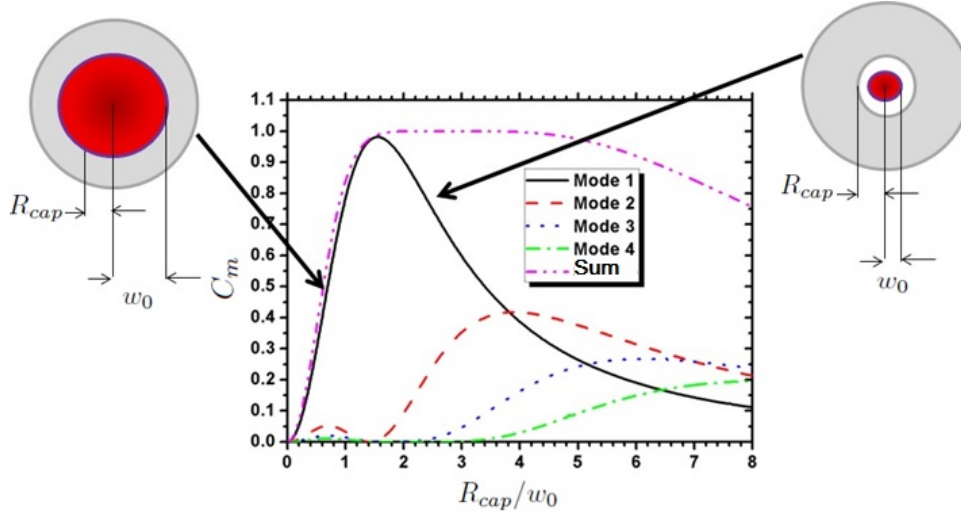
Fig. 11: Coupling coefficient of modes 1 to 4 for a Gaussian beam with waist w_0 incident at the entrance of a capillary with radius R_{cap} as a function of the ratio R_{cap}/w_0.

its electric field is written as $E_{\text{G}}(r) = E_{\text{L}} \exp(-r^2/w_0^2)$, while the electric field inside a capillary tube is a superposition of the EH_{1m} eigenmodes given by Eq. (35). At the entrance ($z = 0$), the continuity of fields reads

$$E_{\text{L}} \exp(-r^2/w_0^2) = \sum_{m=1}^{\infty} A_m E_{1m} = \sum_{m=1}^{\infty} A_m J_0 \left(\frac{u_m r}{R_{\text{cap}}} \right) . \tag{40}$$

The coefficient A_m indicates the amplitude of the EH_{1m} mode, which can be determined using the orthogonality of Bessel functions,

$$A_m = 2E_{\text{L}} \frac{\int_0^1 x \exp(-x^2 R_{\text{cap}}^2/w_0^2) J_0(u_m x)\mathrm{d}x}{J_1^2(u_m)} . \tag{41}$$

The coupling coefficient, C_m, is defined as the fraction of incident energy coupled to the EH_{1m} mode, and can be written as

$$C_m = 8 \left(\frac{R_{\text{cap}}}{w_0} \right)^2 \frac{\left[\int_0^1 x \exp(-x^2 R_{\text{cap}}^2/w_0^2) J_0(u_m x)\mathrm{d}x \right]^2}{J_1^2(u_m)} . \tag{42}$$

C_m thus depends on the mode order m, the capillary radius R_{cap}, and the laser waist w_0. Figure 11 shows the dependence of the coupling coefficient of an incident Gaussian beam on the first four eigenmodes, as a function of the ratio of capillary radius to laser waist, R_{cap}/w_0. It shows that monomode coupling can be achieved when $R_{\text{cap}}/w_0 \simeq 1.55$ (or $w_0/R_{\text{cap}} \simeq 0.65$). In this case, almost 98% of the incident laser energy is coupled to the fundamental EH_{11} mode, and only 1% of laser energy is coupled to higher-order modes. The remaining 1% of laser energy is the energy contained in the wings of the Gaussian function outside the capillary diameter, and it is lost into the material at the front surface of capillary tube.

For $1.2 \leq R_{\text{cap}}/w_0 \leq 2$, more than 90% of the incident laser energy can be coupled to the first mode. When $R_{\text{cap}}/w_0 \to 0$, the laser size increases, and the fundamental mode is predominantly excited, although coupling efficiency decreases quickly. This is because the laser energy is outside the capillary diameter and hits the front surface of the capillary tube, as illustrated by the left-hand sketch in Fig. 11. As R_{cap}/w_0 increases, the laser waist becomes small compared with the optimal waist for quasi-monomode coupling. In this case, the laser energy is nearly 100% coupled into the capillary tube,
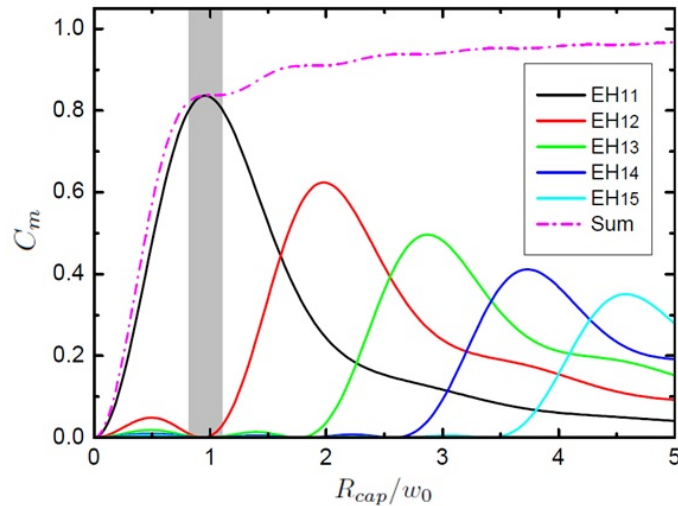
Fig. 12: Coupling coefficient of an Airy beam to the first five eigenmodes as a function R_{cap}/r_0. The grey area indicates the range of R_{cap}/r_0 where more than 80% of the incident laser energy is coupled to the EH$_{11}$ mode.

as indicated by the sum. However, the energy coupled to the fundamental mode decreases at the benefit of the excitation of higher-order modes, resulting in undesired mode beating and severe laser attenuation during the propagation in the capillary tube.

3.3.2 Coupling of an Airy beam

Another case of interest is the coupling of a laser beam with an Airy-like distribution in the focal plane [32]. The electric field amplitude for an Airy beam focused at the capillary entrance ($z = 0$) is written as

$$E_{\mathrm{A}} = E_{\mathrm{L}} \frac{J_1(\nu_1 r/r_0)}{r} \ , \tag{43}$$

where $\nu_1 = 3.8317$ is the first root of the equation $J_1(x) = 0$, and r_0 is the radius corresponding to the first zero. The continuity of the electric field at the capillary entrance gives

$$E_{\mathrm{L}} \frac{J_1(\nu_1 r/r_0)}{r} = \sum_{m=1}^{\infty} A_m E_{1m} = \sum_{m=1}^{\infty} A_m J_0 \left(\frac{u_m r}{R_{\mathrm{cap}}} \right) \ , \tag{44}$$

and the coupling coefficient, C_m, can be calculated from

$$C_m = \frac{4}{J_1^2(u_m)} \left[\int_0^1 J_1 \left(\frac{\nu_1 R_{\mathrm{cap}}}{r_0} x \right) J_0(u_m x) \mathrm{d}x \right]^2 \ . \tag{45}$$

The coupling coefficients for the first five eigenmodes are plotted as functions of R_{cap}/r_0 in Fig. 12. This shows that quasi-monomode guiding of an Airy beam can be achieved for $R_{\mathrm{cap}}/r_0 \simeq 1$. However, even in this case, only 83% of the incident laser energy is coupled to the fundamental mode, while the remaining 17% energy contained by the laser distribution for $r > R_{\mathrm{cap}}$ hits the capillary entrance wall. When R_{cap}/r_0 increases, more laser energy can be coupled inside the capillary tube but essentially to higher-order modes. It is thus less efficient to use an Airy beam than a Gaussian beam in terms of laser coupling and monomode guiding. Another issue is capillary breakdown. The front surface of the capillary tube is exposed to a greater energy for an Airy beam than for one with a Gaussian profile; hence, the peak laser intensity has to be lowered to avoid capillary damage at the entrance.

To summarize, quasi-monomode guiding can be selected by coupling the input laser energy to the fundamental EH$_{11}$ mode: 98% of the energy of a Gaussian beam can be coupled to the fundamental
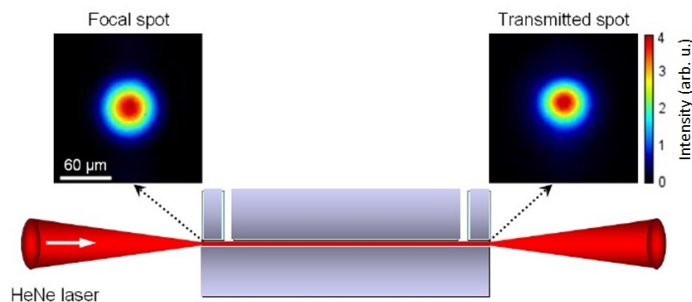
Fig. 13: Example of guiding of a low-intensity laser in the absence of plasma

mode for a waist size $w_0 = 0.645 R_{cap}$. This mode is preferable for laser wakefield acceleration as its group velocity is close to the velocity of light in free space, its damping factor is a minimum for a given capillary radius and wavelength, and the transverse electric field exhibits a variation similar to a Gaussian beam.

3.4 Experimental demonstration of laser guiding

Guiding in capillary tubes at low intensity has been measured and corresponds to theoretical predictions in terms of coupling and transmission. Experimentally, the incident beam should be focused at the entrance plane of the capillary tube, as illustrated in Fig. 13. In this case, a HeNe laser ($\lambda_0 = 632$ nm) was focused at the entrance of a capillary tube, of length 30 mm and inner diameter 127 µm. The focal spot was measured to be Gaussian with a waist of about 43 µm, which corresponds to $w_0 = 0.68 R_{cap}$, close to the matching condition for monomode guiding. Figure 13 also shows the transmitted laser spot at the output of the capillary, which is symmetrical and exhibits the pattern of the fundamental mode.

In this case, a transmission of $90 \pm 2\%$ was obtained experimentally, and can be compared with the theoretical transmission. The transmission is defined as the ratio of incident and transmitted laser energies at the capillary entrance and exit, respectively. The theoretical transmission of energy for the fundamental mode is given by $T_1 = C_1 \exp(-2L_{cap}/L_1^l)$ The theoretical value of the coupling coefficient C_1 of a Gaussian beam for the EH_{11} mode is 97% for $w_0 = 0.68 R_{cap}$. The attenuation length is calculated from Eq. (37) as $L = 230$ cm for $\lambda_0 = 632$ nm and $R_{cap} = 63$ µm. Thus the theoretical transmission is calculated as $T_1 = 94\%$, which is very close to the experimental value, suggesting that excellent alignment and beam quality were achieved in the experiment, leading to quasi-monomode guiding. The sensitivity of the coupling and transmission to misalignments have been studied theoretically: predictions are in excellent agreement with experimental results and are discussed in Ref. [33].

Monomode guiding has been demonstrated experimentally for laser intensities of the order of $I_0 \simeq 10^{16}$ W/cm^2 in a vacuum [34] and the transmission has been measured for different values of the capillary tube radius. In agreement with theoretical predictions, the damping length was found to increase with the cube of capillary radius, as illustrated in Fig. 14. The measured transmissions correspond to the predicted values for quasi-monomode guiding, which was thus measured in a vacuum over a distance of 100 Rayleigh lengths. Analytical predictions for coupling conditions and damping length have been confirmed experimentally for tubes under vacuum.

3.5 Measurement of plasma waves in capillary tubes

The excitation of plasma waves over a length of up to 8 cm was, for the first time [21], demonstrated using laser guiding of intense laser pulses through hydrogen-filled glass capillary tubes. Laser guiding at input intensities up to 10^{18} W/cm^2 was achieved with more than 90% energy transmission in evacuated
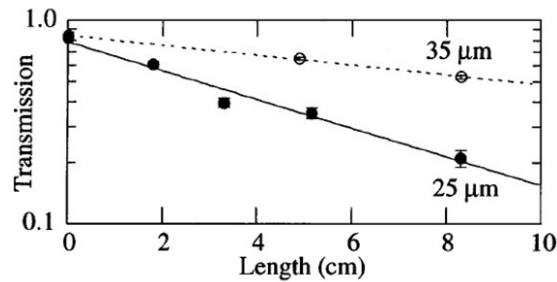
Fig. 14: Measured transmission as a function of propagation distance for two values of the capillary tube radius, $R_{\mathrm{cap}} = 35$ µm (open circles and dotted line), and $R_{\mathrm{cap}} = 25$ µm (filled circles and solid line).
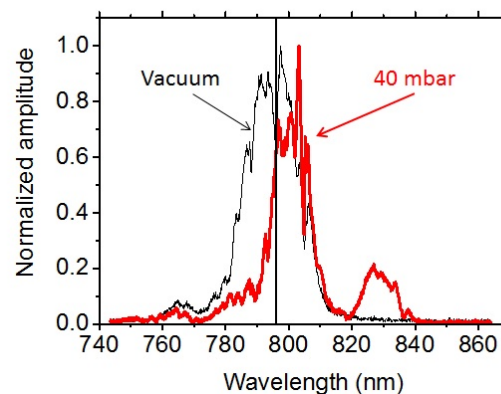
Fig. 15: Spectrum of the laser beam at the output of a capillary tube in vacuum (black solid curve, the vertical line indicates the centre of the spectrum in a vacuum through the capillary) and in 40 mbar of hydrogen (red line).

or hydrogen-filled gas tubes up to 8 cm long, with a radius of $R_{\mathrm{cap}} = 50$ µm. To investigate the linear and moderately non-linear regime, the input intensity was kept below 3×10^{17} W/cm^2, and the electron density was varied in the range 0.05–5×10^{18} cm^{-3}. The plasma wave amplitude was diagnosed by analyzing the spectrum of the transmitted laser radiation. Laser pulses transmitted through gas-filled capillary tubes exhibit broadened spectra. In the range of parameters relevant to this experiment, spectral modifications of the laser pulse driving the plasma wave, after propagating in the plasma over a large distance, are mainly related to changes in the index of refraction of the plasma during the creation of the plasma wave. The front of the laser pulse creates an increase in electron density, leading to a blueshift at the front of the pulse, while the rear of the pulse creates a decrease of electron density with larger amplitude, and thus a redshift of the spectrum.

The signature of plasma wave excitation is a redshift of the laser beam, i.e., a shift of the spectrum towards longer wavelengths, as seen in Fig. 15. The spectrum at the output of a 7 cm long capillary was measured in a vacuum and compared with the spectrum in the presence of 40 mbar of hydrogen filling the tube.

The wavelength shift, $\Delta\lambda/\lambda$ can be shown [35] to be directly related to the energy of the plasma wave with electric field, E_{p}, excited in the plasma volume, V,

$$\frac{\Delta\lambda}{\lambda} \simeq \frac{1}{16\pi\mathcal{E}_{\mathrm{out}}} \int_V E_{\mathrm{p}}^2 \mathrm{d}V \,, \tag{46}$$

where $\mathcal{E}_{\mathrm{out}}$ is the total energy of the transmitted laser pulse. For monomode propagation of a laser pulse
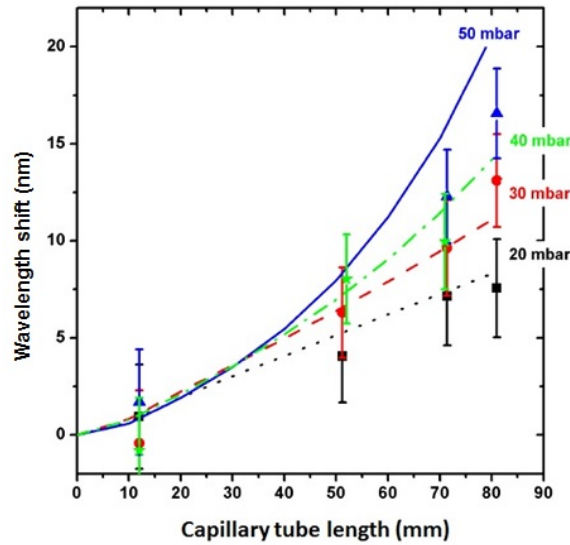
Fig. 16: Wavelength shifts as functions of the length of capillary filled with hydrogen at different pressure values, as determined experimentally (symbols with error bars) and through modelling (curves).

with a Gaussian time envelope, generating a wakefield in the weakly non-linear regime, the wavelength shift can be expressed analytically [21]. For small energy losses at the capillary wall, it is proportional to the peak laser intensity on the capillary axis, and to the length of the capillary, and exhibits a resonant-like dependence on gas pressure directly linked to plasma wave excitation. The dependence of the spectral redshift was measured as a function of filling pressure, capillary tube length, and incident laser energy, and was found to be in excellent agreement with results from modelling, as illustrated in Fig. 16. A linear behaviour of the wavelength shift as a function of length is observed at 20 mbar. The fit of experimental data by simulation results demonstrates that the plasma wave is excited over a length as long as 8 cm. As the pressure is increased, the non-linear laser pulse evolution is amplified, with the propagation length leading to a plasma wave amplitude larger than the linear prediction. The longitudinal accelerating field, inferred from a detailed analysis of laser wakefield dynamics in capillary tubes and numerical simulations [36], is in the range of 1–10 GV/m for an input laser intensity of the order of $I_0 \simeq 10^{17}$ W/cm^2, as shown in Fig. 17. The average product of gradient and length achieved in this experiment was thus of the order of 0.4 GV at a pressure of 50 mbar; it could be increased to several gigavolts by increasing the length and diameter of the capillary tube with higher laser energy.

In conclusion, the outgoing spectra of driving laser pulses measured after propagation in gas-filled capillaries supported by relevant modelling provide detailed information on laser pulse dynamics and on the main characteristics of the accelerating fields excited in the wake of the laser pulses over the long distances necessary for efficient acceleration of electrons to high energies.

4 Conclusion and perspectives

In summary, several features of laser-driven plasma wakefields linked to laser propagation in a plasma were introduced and discussed. Numerous results from previous studies clearly indicate that reducing the plasma density is the best method of increasing the energy of the accelerated electrons. At lower densities, such mechanisms as dephasing between the electron bunch and the plasma wave and the progressive depletion of the driving laser pulse have to be taken into account, as they limit the maximum attainable energy. Increased electron energy of the order of a few gigaelectronvolts is currently achieved at lower plasma densities and higher laser peak powers, in non-linear regimes of laser plasma acceleration, where
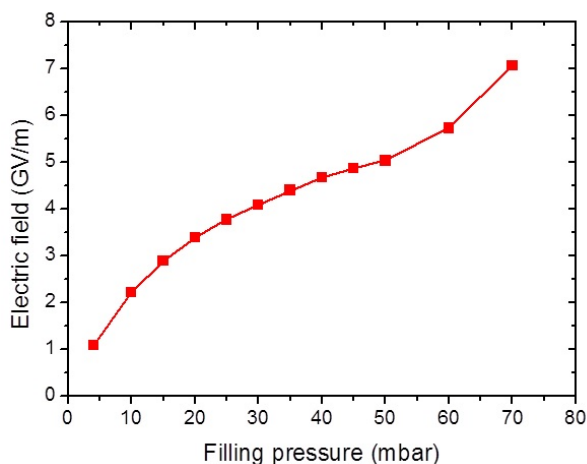
Fig. 17: Average amplitude of the plasma wave electric field measured at the output of a 8 cm long glass capillary tube as a function of filling pressure.

extending the plasma length to the dephasing length is a challenge. Eventually, lowering the plasma density will lead to a regime where self-injection of plasma electrons in the plasma wave does not take place and injection of electrons from external sources has to be implemented.

Ongoing efforts of accelerator development tend to increase the energy of electrons, while improving the electron bunch properties, of which energy spread and transverse emittance are key parameters for beam transport and focusing, as well as subsequent applications. Separating electron injection and acceleration processes, and relying on a number of purely accelerating stages [37], provide solutions to the issues of dephasing and depletion, and make the acceleration process scalable to higher energies, while preserving the bunch quality. Two crucial aspects of multistage laser-driven plasma acceleration are laser guiding over metre-scale distances [38] and control of the properties of the electron bunch for external injection into the plasma wave of the accelerating stage.

Future work will tackle the different issues of multistage laser-driven plasma acceleration. Laser guiding and increased laser energy are expected to produce electron bunches in the 10 GeV range in one stage (see, for example, the BELLA project [39] in the USA or the CILEX APOLLON project [40] in France). Staging is the next milestone for the development of laser-driven plasma accelerators. The APOLLON laser facility [41], under construction in France, will provide two beams in the 1–10 PW range, and a shielded large experimental area built to host multistage experiments [40]. In those experiments, the first stage (i.e., the injector) will be used to generate electrons in the 50–200 MeV range, which will then be transported [42] and focused into the accelerator stage. The quality of the accelerated bunch will thus result from the combined intrinsic properties of the injector and the capability of the transport line to accommodate them.

In the long term, the development of laser-driven accelerators will rely on improvements in the performance of laser systems, in terms of beam quality, reliability, stability, and average power. Plasma stages in the quasi-linear regime provide means to control transverse and longitudinal fields, and consequently the dynamics of accelerated bunches. Electron or positron beams can be accelerated in this regime, and external injection schemes into metre-scale low-density plasma sources need to be developed.

References

[1] T. Tajima and J.M. Dawson, *Phys. Rev. Lett.* **43**(4) (1979) 267. http://dx.doi.org/10.1103/PhysRevLett.43.267

[2] F.F. Chen, *Introduction to Plasma Physics and Controlled Fusion*, 2nd ed. (Plenum, New York, 1983).

[3] E. Esarey *et al.*, *Rev. Mod. Phys.* **81**(3) (2009) 1229. http://dx.doi.org/10.1103/RevModPhys.81.1229

[4] N.E. Andreev *et al.*, *JETP Lett.* **55** (1992) 571.

[5] A. Pukhov and J. Meyer-ter-Vehn, *Appl. Phys.* B **74**(4) (2002) 355. http://dx.doi.org/10.1007/s003400200795

[6] S.P.D. Mangles *et al.*, *Nature* **431** (2004) 535. http://dx.doi.org/10.1038/nature02939

[7] C.G.R. Geddes *et al.*, *Nature* **431** (2004) 538. http://dx.doi.org/10.1038/nature02900

[8] J. Faure *et al.*, *Nature* **431** (2004) 541. http://dx.doi.org/10.1038/nature02963

[9] W.P. Leemans *et al.*, *Phys. Rev. Lett.* **113**(24) (2014) 245002. http://dx.doi.org/10.1103/PhysRevLett.113.245002

[10] B.E.A Saleh and M.C. Teich, *Fundamentals of Photonics* (John Wiley & Sons, Hoboken, NJ, 2007).

[11] J. Ju *et al.*, *Phys. Plasmas* **20**(8) (2013) 083106. http://dx.doi.org/10.1063/1.4817747

[12] P. Gibbon, *Short Pulse Laser Interactions With Matter* (Imperial College Press, London, 2005).

[13] E. Esarey *et al.*, *IEEE Trans. Plasma Science* **24**(2) (1996) 252. http://dx.doi.org/10.1109/27.509991

[14] W.L. Kruer, *The Physics of Laser Plasma Interactions* (Westview Press, Boulder, CO, 2003).

[15] W. Lu *et al.*, *Phys. Rev. ST Accel. Beams* **10**(6) (2007) 061301. http://dx.doi.org/10.1103/PhysRevSTAB.10.061301

[16] W.B. Mori, *IEEE J. Quant. Elec.* **33**(11) (1997) 1942. http://dx.doi.org/10.1109/3.641309

[17] G.-Z. Sun *et al.*, *Phys. Fluids* **30** (1987) 526. http://dx.doi.org/10.1063/1.866349

[18] P. Sprangle *et al.*, *Appl. Phys. Lett.* **53**(22) (1988) 2146. http://dx.doi.org/10.1063/1.100300

[19] E. Esarey *et al.*, *Phys. Fluids* B **5**(7) (1993) 2690. http://dx.doi.org/10.1063/1.860707

[20] J. Schreiber *et al.*, *Phys. Rev. Lett.* **105**(23) (2010) 235003. http://dx.doi.org/10.1103/PhysRevLett.105.235003

[21] F. Wojda *et al.*, *Phys. Rev.* E **80**(6) (2009) 066403. http://dx.doi.org/10.1103/PhysRevE.80.066403

[22] C.G.R. Geddes *et al.*, *Phys. Plasmas* **12**(5) (2005) 056709. http://dx.doi.org/10.1063/1.1882352

[23] B.A. Shadwick *et al.*, *Phys. Plasmas* **16**(5) (2009) 056704. http://dx.doi.org/10.1063/1.3124185

[24] C.D. Decker *et al.*, *Phys. Plasmas* **3**(5) (1996) 2047. http://dx.doi.org/10.1063/1.872001

[25] D.J. Spence and S.M. Hooker, *Phys. Rev.* E **63**(1) (2000) 015401(R). http://dx.doi.org/10.1103/PhysRevE.63.015401

[26] D.J. Spence *et al.*, *J. Opt. Soc. Am.* B **20**(1) (2003) 138. http://dx.doi.org/10.1364/JOSAB.20.000138

[27] C.G.R. Geddes *et al.*, *Phys. Rev. Lett.* **95**(14) (2005) 145002. http://dx.doi.org/10.1103/PhysRevLett.95.145002

[28] W. Leemans *et al.*, *Philos. Trans. R. Soc.* A **364**(1840) (2006) 585. http://dx.doi.org/10.1098/rsta.2005.1724

[29] A. Butler *et al.*, *Phys. Rev. Lett.* **89**(18) (2002) 185003. http://dx.doi.org/10.1103/PhysRevLett.89.185003

[30] B. Cros *et al.*, *Phys. Rev.* E **65**(2) (2002) 026405. http://dx.doi.org/10.1103/PhysRevE.65.026405

[31] D. Du *et al.*, *Appl. Phys. Lett.* **64**(23) (1994) 3071. http://dx.doi.org/10.1063/1.111350

[32] E. Hecht, *Optics*, 4th ed. (Addison Wesley, Reading, MA, 2002).

[33] M. Veysman *et al.*, *J. Opt. Soc. Am.* B **27**(7) (2010) 14008.
http://dx.doi.org/10.1364/JOSAB.27.001400

[34] F. Dorchies *et al.*, *Phys. Rev. Lett.* **82**(23) (1999) 4655.
http://dx.doi.org/10.1103/physrevlett.82.4655

[35] N.E. Andreev and M. Chegotov, *JETP* **101**(1)(2005) 56. http://dx.doi.org/10.1134/1.2010661

[36] N.E. Andreev *et al.*, *New J. Phys.* **12** (2010) 045024.
http://dx.doi.org/10.1088/1367-2630/12/4/045024

[37] W.P. Leemans and E. Esarey, *Physics Today* 62(3) (March 2009) 44.
http://dx.doi.org/10.1063/1.3099645

[38] B.S. Paradkar *et al.*, *Phys. Plasmas* **20**(8) (2013) 083120. http://dx.doi.org/10.1063/1.4819718

[39] BELLA: The Berkeley Lab Laser Accelerator. `http://www.lbl.gov/community/bella/`

[40] B. Cros *et al.*, *Nucl. Instrum. Meth.* A **740**(11) (2014) 27.
http://dx.doi.org/10.1016/j.nima.2013.10.090

[41] J.P. Zou *et al.*, *High Power Laser Sci. Eng.* **3** (2015) e2. http://dx.doi.org/10.1017/hpl.2014.41

[42] A. Chancé *et al.*, *Nucl. Instrum. Meth.* A **740**(11) (2014) 158.
http://dx.doi.org/10.1016/j.nima.2013.10.036

Ion Acceleration—Target Normal Sheath Acceleration*

M. Roth and M. Schollmeier
Institute for Nuclear Physics, Technische Universität Darmstadt, Darmstadt, Germany

Abstract
Energetic ions have been observed since the very first laser-plasma experiments. Their origin was found to be the charge separation of electrons heated by the laser, which transfers energy to the ions accelerated in the field. The advent of ultra-intense lasers with pulse lengths in the femtosecond regime resulted in the discovery of very energetic ions with characteristics quite different from those driven by long-pulse lasers. Discovered in the late 1990s, these ion beams have become the focus of intense research worldwide, because of their unique properties and high particle numbers. Based on their non-isotropic, beam-like behaviour, which is always perpendicular to the emitting surface, the acceleration mechanism is called target normal sheath acceleration (TNSA). We address the physics of the mechanism and its dependence on laser and target parameters. Techniques to explore and diagnose the beams, to make them useful for applications, are also addressed.

Keywords
Laser; ion acceleration; ultraintense lasers; plasma; accelerator.

1 Introduction

Since the first irradiation of a target by a laser, the generation of energetic ions has been well known. The origin of those ions is the electric field generated by the charge separation as a result of the energy transferred from a long-pulse laser to the electrons, and their respective temperature [1]. The ions are then accelerated in the double-layer potential and can reach significant particle energies, expanding isotropically in all directions from the target front surface. Since the advent of ultra-short-pulse lasers with pulse lengths of less than picoseconds, one of the most exciting results obtained in experiments using solid targets is the discovery of very energetic, very intense bursts of ions coming off the rear, non-irradiated surface in a very high quality, beam-like fashion. At the turn of the century, a number of experiments have resulted in proton beams with energies of up to several tens of megaelectronvolts generated behind thin foils irradiated by lasers exceeding hundreds of terawatts [2–4]. Since the first observations, an extraordinary amount of experimental and theoretical work has been devoted to the study of these beams' characteristics and production mechanisms. Particular attention has been devoted to the exceptional accelerator-like spatial quality of the beams, and current research focuses on their optimization for use in a number of groundbreaking applications, addressed in Section 4. But first we will focus on the best understood of all the possible acceleration mechanisms, so-called target normal sheath acceleration (TNSA).

The greater part of this chapter is drawn from Ref. [5]. Review articles about TNSA, the diagnostics of short-pulse laser plasmas, and applications in fast ignition can also be found in Refs. [6–10].

---

2 Target normal sheath acceleration: the mechanism

2.1 Initial conditions

The primary interaction of a high-intensity short laser pulse with a solid target strongly depends on the contrast of the laser pulse, that is, the ratio of unwanted, preceding laser light to the main pulse. At peak intensities exceeding 10^{20} W/cm^2 even a contrast of 10^6 is sufficient to excite a plasma that is expanding towards the incoming main pulse. As a common source of this unwanted laser light, amplified spontaneous emission or prepulses, caused by a limited polarization separation in regenerative amplifiers, have been identified. This ablative plasma sets the stage for a wealth of uncontrolled phenomena at the interaction of the main pulse with the target. The laser beam can undergo self-focusing due to ponderomotive force or relativistic effects, thereby strongly increasing the resulting intensity, or it can break up into multiple filaments, or, finally, it can excite instabilities that ultimately lead to the production of energetic electrons. Moreover, the ablative pressure of blow-off plasma caused by the incident laser energy prior to the main pulse launches a shockwave into the target, which can ultimately destroy the target before the arrival of the main pulse. We address this issue, even though it is not directly related to the TNSA mechanism, because of its influence on the electron spectrum and the thickness of targets that can be used in practice.

2.2 General description

Before going into detail, it is worth taking a step back and qualitatively looking at the general picture of the TNSA ion acceleration mechanism. Let us interpret the process of generating a proton beam by TNSA as a new variation on a familiar theme—acceleration by a sheath electrostatic field generated by the hot-electron component. We assume the interaction of an intense laser pulse well exceeding 10^{18} W/cm^2 with a solid thin foil target as the standard case for TNSA. The interaction of the intense laser pulse with the preformed plasma and the underlying solid target constitutes a source of hot electrons with an energy spectrum related to the laser intensity. This cloud of hot electrons penetrates the foil at, as we shall see, an opening angle of about 30° and escapes into the vacuum behind the target. The target's capacitance, however, allows only a small fraction of the electrons to escape before the target is sufficiently charged that escape is impossible for even megaelectronvolt electrons. Those electrons are then electrostatically confined to the target and circulate back and forth through the target, laterally expanding and forming a charge-separation field on both sides over a Debye length. At the rear surface there is no screening plasma, owing to the short time-scales involved, and the induced electric fields are of the order of several teravolts per metre. Such fields can ionize atoms and rapidly accelerate ions normal to the initially unperturbed surface. The resulting ion trajectories thus depend on the local orientation of the rear surface and the electric field lines driven by the time-dependent electron density distribution. As the ions start from a cold solid surface just driven by quasi-static electric fields, the resulting beam quality is extremely high, as we shall see. This process is illustrated in Fig. 1.

2.3 Electron driver

Current laser systems are not yet capable of accelerating ions directly. Therefore, all existing laser ion acceleration mechanisms rely on the driving electron component and the resulting strong electric field caused by charge separation. Thus the electron driver is extremely important, and will be discussed here in detail. As a rule of thumb, particle-in-cell calculations [11–13] have indicated that the so-called hot-electron component has a logarithmic-slope temperature that is roughly equal to the ponderomotive potential of the laser beam. This is represented by the cycle-averaged kinetic energy of an electron oscillating in the laser electromagnetic field, $T_{\text{hot}} \approx U_{\text{pond}} \approx 1 \text{ MeV} \times (I\lambda^2/10^{19} \text{ W} \cdot \text{µm}^2/\text{cm}^2)^{1/2}$ in the relativistic regime [14]. The relativistic electrons are directed mainly in the forward direction [15]; hence, the particle distribution function can be simplified by a one-dimensional Maxwell–Jüttner distribution, which is close to an ordinary Boltzmann distribution.
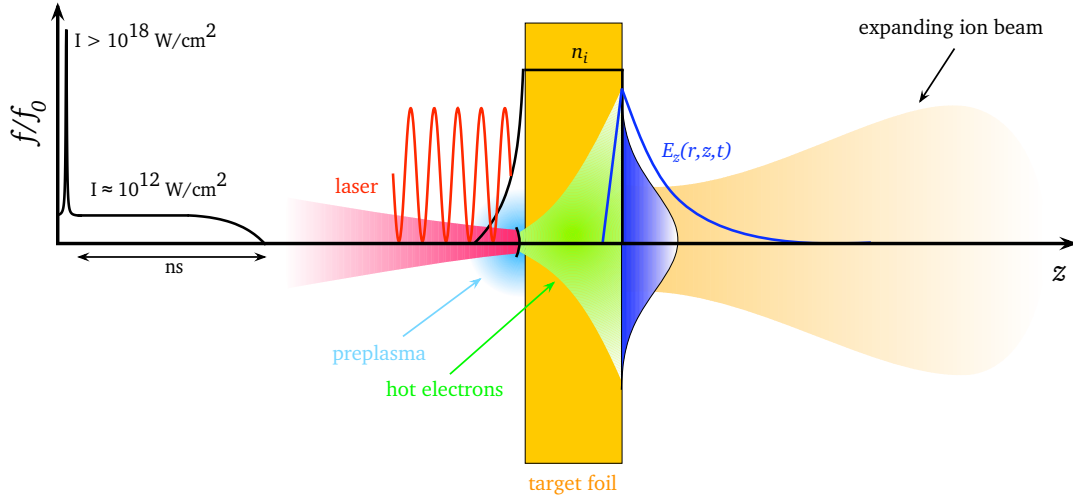
Fig. 1: Target normal sheath acceleration. A thin target foil with thickness $d = 5\text{–}50$ μm is irradiated by an intense laser pulse. The laser prepulse creates a preplasma on the target's front side. The main pulse interacts with the plasma and accelerates megaelectronvolt electrons, mainly in the forward direction. The electrons propagate through the target, where collisions with the background material can increase the divergence of the electron current. The electrons leave the rear side, resulting in a dense sheath. An electric field due to charge separation is created. The field is of the order of the laser electric field ($\sim$ TV/m), and ionizes atoms at the surface. The ions are then accelerated in this sheath field, pointing in the target normal direction.

The conversion efficiency from laser energy to hot electrons is not perfect, and only a fraction η is converted. The total number of electrons is

$$n_0 = \frac{\eta E_{\mathrm{L}}}{c \tau_{\mathrm{L}} \pi r_0^2 k_{\mathrm{B}} T_{\mathrm{hot}}} \; , \tag{1}$$

following a scaling with intensity as

$$\eta = 1.2 \times 10^{-15} I^{0.74} \; , \tag{2}$$

with the intensity in watts per square centimetre reaching up to 50% [16]. For ultra-high intensities, η can reach up to 60% for near-normal incidence and up to 90% for irradiation under 45° [17]. A discussion on which distribution function best fits the experimental data is given in Ref. [18] and, in more detail, in Ref. [19]. However, neither theoretical nor experimental data give a clear answer to the question about the shape of the distribution function.

Given the intensities of modern short-pulse lasers, therefore, copious amounts of energetic electrons are generated and, in contrast with thermal electrons in long-pulse laser plasmas, are pushed into the target. It is fair to estimate that a fraction of $N = \eta E_{\mathrm{L}}/k_{\mathrm{B}} T_{\mathrm{hot}}$ electrons in the megaelectronvolt range are created, where E_{L} is the laser energy. These electrons have typical energies such that their mean free path is much longer than the thickness of the targets typically used in experiments. While the electrons propagate through the target, they constitute a current that exceeds the Alfvén limit by several orders of magnitude. Alfvén found that the main limiting factor on the propagation of an electron beam in a conductor is the self-generated magnetic field, which bends the electrons back towards the source [20]. For parameters of relevance for inertial confinement fusion, a good review is given in Ref. [21]. So as not to exceed the limit of $j_{\mathrm{A}} = m_{\mathrm{e}} c^3 \beta \gamma / e = 17 \beta \gamma$ (kA), the net current must be largely compensated for by return currents, to minimize the resulting magnetic field. The return currents will be driven by the charge
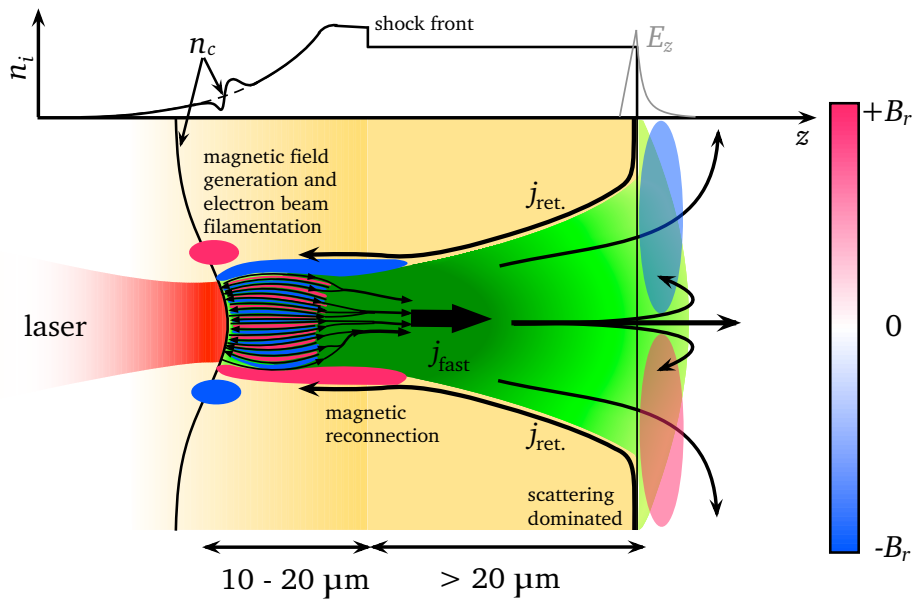
Fig. 2: Schematic of laser-generated fast-electron transport. The laser (shown in red) impinges on a preplasma with exponential density profile from the left side. The light pressure leads to profile steepening, depicted in the graph at the top of the figure. An ablation plasma creates an inward-travelling shockwave that heats, ionizes, and compresses the target. Fast electrons are created by the laser, propagating into the dense plasma towards the target's rear side. The high electron current j_{fast} can lead to filamentation and magnetic field generation (shown by the light red- and blue-coloured areas), as well as driving a return current j_{ret}. The global magnetic field tends to pinch the fast-electron current. Electrons propagating in the dense solid matter interact with the background material by binary collisions. This leads to a spatial broadening of the electron distribution, which becomes the major effect for longer distances. At the rear side, the electrons form a sheath and build up an electrostatic field E_z (grey line in graph). This can lead to refluxing (recirculation) of the electrons, heating the target even further.

separation in the laser–plasma interaction region and strongly depends on the electrical conductivity of the target, as those currents are lower in energy and thereby affected by the material properties. The large counter-streaming currents also give rise to instabilities, which affect the forward motion of the electrons. The influence of limited electrical conductivity on the inhibition of fast-electron propagation has been addressed in Ref. [22], also with respect to space charge separation. Without the return currents, the electric field would stop the electrons in a distance of less than 1 nm [23]. The electric field driving the return current, in turn, can be strong enough to stop the fast electrons. This effect, known as transport inhibition, is prominent in insulators, but almost negligible in conductors [24].

The propagation of electrons through the target is still an active field of research. As depicted in Fig. 2, the laser pushes the critical surface n_c, leading to a steepening of the electron density profile. The motion of the ablated plasma causes a shockwave to be launched into the target, leading to ionization and therefore a modification of the initial electrical conductivity. As soon as the electrons penetrate the cold solid region, binary collisions (multiple small-angle scattering) with the background material are no longer negligible. These tend to broaden the electron distribution, counteracting the magnetic field effect [25]. For long propagation distances ($z \geq 15$ μm), the current density is low enough that broadening due to small-angle scattering becomes the dominating mechanism [26].

The majority of data show a divergent electron transport. The transport full-cone angle of the electron distribution was determined to be dependent on laser energy and intensity, as well as target thickness. For rather thick targets ($d > 40$ μm) this value is around 30° for full width at half maximum (FWHM), whereas for thin targets ($d \leq 10$ μm) published values are of the order of 16° (this figure was indirectly obtained by a fit to proton energy measurements) and are $\approx 150°$ at most [27]. It has been shown that different diagnostics lead to different electron transport cone angles [28], so the nature of the 'true' cone angle dependence with laser and target parameters remains unclear.

When the electrons reach the rear side of the target, they form a dense charge-separation sheath. The outflowing electrons lead to a toroidal magnetic field B_θ, which can spread the electrons over large transverse distances by a purely kinematic $E \times B_\theta$ force [29], sometimes called the fountain effect [30]. The electric field created by the electron sheath is sufficiently strong to deflect electrons back into the target, which then recirculate. Experimental evidence for recirculating electrons is presented in Refs. [17,31,32]. The relevance to proton acceleration was first demonstrated by MacKinnon *et al.* [33], who measured a strong enhancement of the maximum proton energy for foils thinner than 10 μm, compared with thicker foils. With the help of computer simulations, this energy enhancement was attributed to an enhanced sheath density caused by refluxing electrons. Further evidence of refluxing electrons was also found in an experiment discussed in Ref. [34].

Neglecting the complicated interaction for thicknesses below $d \approx 15$ μm, a reasonable estimate for the electron beam divergence is the assumption that the electrons are generated in a region of the size of the laser focus and are purely collisionally transported to the rear side of the target. This is in agreement with most published data. The broadening of the distribution is then due to multiple Coulomb small-angle scattering, given analytically, e.g., by Molière's theory in Bethe's description [35]. To lowest order, the angular broadening, $f(\Theta)$, follows a Gaussian (see Ref. [35], Eq. (27)),

$$f(\theta) = \frac{2e^{-\vartheta^2}}{\chi_c^2 B} \sqrt{\theta / \sin\theta}, \tag{3}$$

where the second term on the right-hand side is a correction for larger angles (from Ref. [35], Eq. (58)). The angle ϑ can be related to θ by $\vartheta = \theta / \chi_c B^{1/2}$. The transcendental equation, $B - \ln B = \ln\left(\chi_c^2 / \chi_{a'}^2\right)$, determines B. The screening angle $\chi_{a'}^2$ is given by $\chi_{a'}^2 = 1.167(1.13 + 3.76\alpha^2)\lambda^2/a^2$, where $\lambda = \hbar/p$ is the de Broglie wavelength of the electron and $a = 0.885a_B Z^{-1/3}$, with the Bohr radius a_B. α is determined by $\alpha = Ze^2/(4\pi\varepsilon_0\hbar\beta c)$ with the nuclear charge Z, electron charge e, and $\beta = v/c$, where ε_0, $\hbar$, and c denote the usual constants. The variable χ_c is given by

$$\chi_c^2 = \frac{e^4}{4\pi\varepsilon_0^2 c^2} \frac{Z(Z+1)Nd}{\beta^2 p^2}, \tag{4}$$

with the electron momentum p and $N = N_A \rho/A$ being the number of scattering atoms, determined by Avogadro's number N_A, material density ρ, and mass number A. χ_c is proportional to the material thickness d and density ρ as $\chi_c \propto (\rho d)^{1/2}$. Since χ_c determines the width of $f(\theta)$, the angular spread of the electron distribution propagating through matter is proportional to its thickness as well as its density. The analytical formula allows us to estimate the broadening of the laser-accelerated electron distribution during the transport through the cold solid target. For a laser intensity $I_L = 10^{19}$ W/cm^2, the mean energy (temperature) is $k_B T_{hot} \approx 1$ MeV. The increase in distribution radius r with target thickness d is shown in Fig. 3. The electrons were propagated in aluminium (red dashed lines) and gold (blue dashed lines). Aluminium does not lead to a strong broadening, owing to its low density and Z; compare this with the broadening observed in gold. The graph shows that in each case the radius at the rear side scales as $r \propto d^2$ (green lines).

The estimate based on an electron distribution broadening determined by small-angle scattering can be used to explain the measured proton beam profiles. It should be noted that although the model
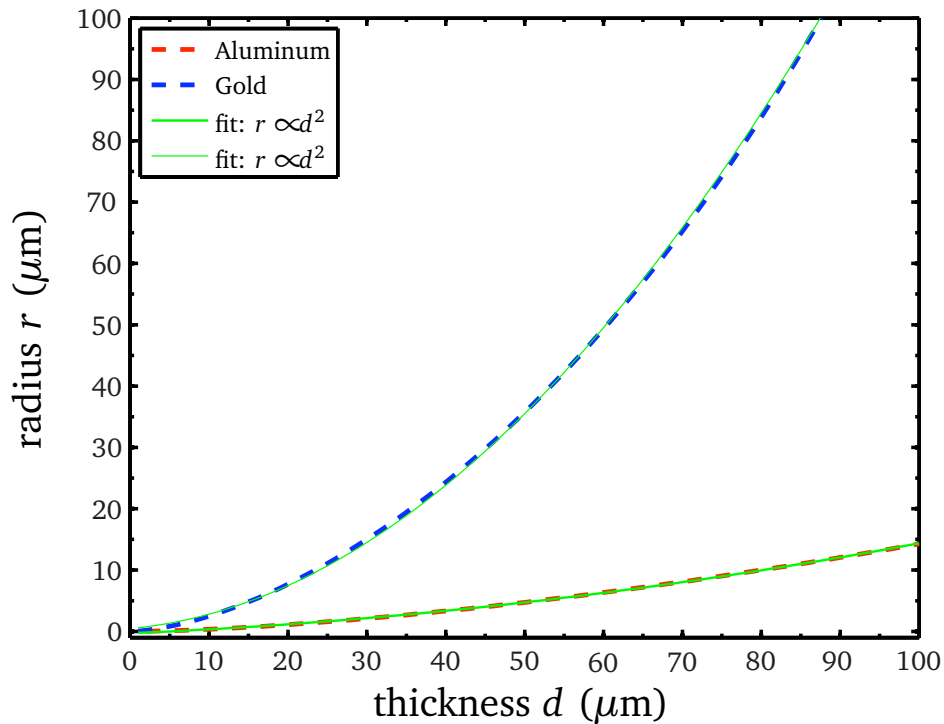
Fig. 3: Increase in radius r of an electron distribution with target thickness d. The calculation utilized Eq. (3), taking an energy of $k_B T_{hot} \approx 1$ MeV, corresponding to a laser intensity $I = 10^{19}$ W/cm^2.

seems to be able to calculate the broadening of the forward-propagating fast-electron distribution generated by intense laser–matter interaction, it could fail to determine the real number of electrons arriving at the rear side of the target. According to Davies [19], the generation of electromagnetic fields, as well as the recirculation of the electrons, must be taken into account, making estimation and even calculation very difficult. Recent experiments by Akli *et al.* [36] have shown that this is true at least for targets thinner than 20 μm, but for thicker foils the assumption of strong recirculation overestimates the number of electrons. Therefore, the question of whether electromagnetic fields and recirculation are essential to determine fast-electron transport from the front to the rear side of the target can still not be satisfactorily answered, although assuming simple collisional broadening gives a a relatively good estimate.

2.4 Target normal sheath acceleration

The electrons are transported through the target to its rear side. The laser creates about 10^{13} electrons, which potentially all propagate through the target. The broadening results in transverse extension, which can be estimated by

$$r_{sheath} = r_0 + d\tan(\theta/2), \qquad (5)$$

where r_0 denotes the laser spot radius, d the target thickness, and θ the broadening angle of the distribution, e.g., calculated using Eq. (3). The electrons exhibit an exponential energy distribution

$$n_{hot}(E) = n_0 \exp\left(-\frac{E}{k_B T_{hot}}\right) \qquad (6)$$

with temperature $k_B T$ and overall density n_0, as given by Eq. (1). The electron density at the rear side (neglecting recirculation), therefore, can be estimated as

$$n_{e,0} = \frac{\eta E_L}{c \tau_L \pi (r_0 + d \tan \theta / 2)^2 k_B T_{hot}} \tag{7}$$

$$\approx 1.5 \times 10^{19} \frac{r_0^2}{(r_0 + d \tan \theta / 2)^2} \frac{I_{18}^{7/4}}{\sqrt{1 + 0.73 I_{18} \lambda_{\mu m}^2} - 1} \; \left[\text{cm}^{-3} \right] . \tag{8}$$

This last approximation was obtained by inserting $E_0 = \sqrt{2I_0 / \varepsilon_0 c} \approx 2.7 \times 10^{12}$ V/m and using Eqs. (1) and (2), together with a practical notation for the electron temperature based on ponderomotive scaling,

$$k_B T_{hot} = m_0 c^2 \left(\sqrt{1 + \frac{I_0 \; [\text{W/cm}^2] \; \lambda_L^2 \; [\mu m^2]}{1.37 \times 10^{18}}} - 1 \right) , \tag{9}$$

in Eq. (7). I_{18} indicates that the intensity has to be taken in units of 10^{18} W/cm^2. The estimate shows that the electron density at the rear side of the target strongly scales with the laser intensity and is inversely proportional to the square of the target thickness. Taking the standard example of a laser pulse with $I = 10^{19}$ W/cm^2, focused to a spot of $r_0 = 10$ μm and assuming a target thickness $d = 20$ μm, the angular broadening according to Eq. (3) is $\theta = 42°$ (FWHM) for electrons with mean energy $k_B T$, as determined by Eq. (9). Hence, the electron density at the target's rear side is $n_{e,0} = 1.4 \times 10^{20}$ cm^{-3}. This is orders of magnitude below the density of the solid and justifies the assumption of a shielded transport through the target.

The electrons arrive at the rear side of the target and escape into the vacuum. The charge separation leads to an electric potential Φ in the vacuum region, according to Poisson's equation. In one dimension, it is given as

$$\varepsilon_0 \frac{\partial^2 \Phi}{\partial z^2} = e n_e . \tag{10}$$

To solve Eq. (10), it is assumed that the solid matter in one half-space ($z \leq 0$) perfectly compensates for the electric potential, whereas for $z \to \infty$ the potential goes to infinity. The derivative $\partial \Phi / \partial z$ vanishes for $z \to \pm\infty$. In the vacuum region ($z > 0$), the field can be obtained analytically [37]. The electron density is taken as

$$n_e = n_{e,0} \exp \left(\frac{e\Phi}{k_B T_{hot}} \right) , \tag{11}$$

where the electron kinetic energy is replaced by the potential energy $-e\Phi$. The initial electron density $n_{e,0}$ is taken from Eq. (8). The solution of the Poisson equation is found by using the ansatz $e\Phi / k_B T_{hot} = -2 \ln(\lambda z + 1)$, where λ is a constant defined by the solution and the '+1' is necessary to fulfil a continuous solution with the condition $\Phi(0) = 0$ at the boundary to the solid matter. The resulting potential is

$$\Phi(z) = -\frac{2 k_B T_{hot}}{e} \ln \left(1 + \frac{z}{\sqrt{2} \lambda_D} \right) \tag{12}$$

and the corresponding electric field reads

$$E(z) = \frac{2 k_B T_{hot}}{e} \frac{1}{z + \sqrt{2} \lambda_D} . \tag{13}$$

In this solution, the electron Debye length

$$\lambda_D = \left(\frac{\varepsilon_0 k_B T_{hot}}{e^2 n_{e,0}} \right)^{1/2} \tag{14}$$

appears; this is defined as the distance over which significant charge separation occurs [38]. Replacing $k_B T_{hot}$ with Eq. (9) and $n_{e,0}$ with Eq. (8) leads to

$$\lambda_D \approx 1.37 \ \mu m \frac{r_0 + d \tan \theta/2}{r_0} \frac{\sqrt{1 + 0.73 I_{18} \lambda^2} - 1}{I_{18}^{7/8}} \ . \tag{15}$$

The Debye length, or longitudinal sheath extension, on the rear side is of the order of a micrometre. It scales quadratically with target thickness (since $d \tan(\theta/2) \propto d^2$) and is inversely proportional to the laser intensity. Thus, a higher laser intensity on the front side of the target leads to a shorter Debye length at the rear side and results in a stronger electric field. The standard example of Eq. (15) leads to $\lambda_D = 0.6 \ \mu m$.

The maximum electric field is obtained at $z = 0$:

$$E_{max}(z = 0) = \frac{\sqrt{2} k_B T_{hot}}{e \lambda_D} \tag{16}$$

$$\approx 5.2 \times 10^{11} \ V/m \ \frac{r_0}{r_0 + d \tan \theta/2} I_{18}^{7/8} \ , \tag{17}$$

$$= 9 \times 10^{10} \ V/m \ \frac{r_0}{r_0 + d \tan \theta/2} E_{12} E_{12}^{3/4} . \tag{18}$$

Hence, the initial field at $z = 0$ is proportional to the laser intensity and depends almost quadratically on the laser's electric field strength. In Eq. (18), the laser's electric field strength is inserted in normalized units of $10^{12} \ V/m$. By inserting the dependence of the broadening with target thickness from Fig. 3, the scaling with the target thickness is obtained as $E_{max}(z = 0) \propto d^{-2}$. The standard example leads to a maximum field strength of $E_{max} \approx 2 \times 10^{12} \ V/m$ just at the surface, that is it is of the order of TV/m or MV/μm. It is only slightly smaller than the laser electric field strength of $E_0 = 8.7 \times 10^{12} \ V/m$. However, at times later than $t = 0$, the field strength is dictated by the dynamics at the rear side of the target, e.g., ionization and ion acceleration.

As just mentioned, the electric field strength instantly leads to ionization of the atoms at the target's rear surface, since it is orders of magnitude above the ionization threshold of the atoms. A simple model to estimate the electric field strength necessary for ionization is the field ionization by barrier suppression (FIBS) model [39]. The external electric field of the laser overlaps the Coulomb potential of the atom and deforms it. As soon as the deformation is below the binding energy of the electron, the electron is instantly freed, hence the atom is ionized. The threshold electric field strength E_{ion} can be obtained with the binding energy U_{bind} as

$$E_{ion} = \frac{\pi \varepsilon_0 U_{bind}^2}{e^3 Z} \ . \tag{19}$$

As the electron sheath at the rear side is relatively dense, the atoms could also be ionized by collisional ionization. However, as discussed by Hegelich [40], the cross-section for field ionization is much higher than the cross-section for collisional ionization for the electron densities and electric fields appearing at the target surface. Taking the ionization energy of a hydrogen atom with $U_{bind} = 13.6$ eV, the field strength necessary for FIBS is $E_{ion} = 10^{10} \ V/m$. This is two orders of magnitude less than the field strength developed by the electron sheath in vacuum calculated earlier. Hence, nearly all atoms (protons, carbon, heavier particles) at the rear side of the target are instantly ionized and, since they are no longer neutral particles, they are then subject to the electric field and are accelerated. The maximum charge state of ions found in an experiment is an estimate of the maximum field strength that appeared. This has been used to estimate the sheath peak electric field value [40], as well as the transverse field extension [41,42].

The strong field ionizes the target and accelerates ions to mega electronvolt energies, if it is applied for a sufficiently long time. The time can easily be calculated by the assumption of a test particle moving

in a static field, generated by the electrons. Free protons were chosen as test particles. The non-linear equation of motion is obtained from Eq. (13). The solution was obtained numerically with MATLAB [43]. It shows that for a proton to obtain a kinetic energy of 5 MeV, the field has to stay for 500 fs in the shape given by Eq. (13). During this time, the proton has travelled 11.3 μm. The electric field will be created as soon as electrons leave the rear side of the target.

Some electrons can escape this field, whereas others with lower energy will be stopped and will be re-accelerated back into the target. Since the electron velocity is close to the speed of light and the distances are of the order of a micrometre, this happens on a time-scale of a few femtoseconds, leading to a situation where electrons are always present outside the rear side of the target. The electric field being created does not oscillate but is quasi-static on the order of the ion acceleration time. Therefore, ultra-short laser pulses, although providing the highest intensities, are not the optimum laser pulses for ion acceleration. The electric field is directed normal to the target's rear surface; hence, the direction of the ion acceleration is normal to the target, giving the process its name, target normal sheath acceleration.

2.5 Expansion models

The laser acceleration of ions from solid targets is a complicated, multidimensional mechanism including relativistic effects, non-linearities, and collective and kinetic effects. Theoretical methods for the various physical mechanisms involved in TNSA range from analytical approaches for simplified scenarios over fluid models up to fully relativistic, collisional three-dimensional computer simulations.

Most of the approaches that describe TNSA neglect the complex laser–matter interaction at the front side of the target as well as the electron transport through the foil. These plasma expansion models start with a hot-electron distribution that drives the expansion of an initially given ion distribution [16, 37, 44–50]. Crucial features such as the maximum ion energy, as well as the particle spectrum, can be obtained analytically, whereas the dynamics have to be obtained numerically. The plasma expansion description dates back to 1954 [51]. Since then, various refinements of the models were obtained, with an increasing activity after the first discovery of TNSA. These calculations resemble the general features of TNSA. Nevertheless, they rely on somewhat idealized initial conditions from simple estimates. In addition to that, the plasma expansion models are one-dimensional, whereas the experiments have clearly shown that TNSA is at least two-dimensional. Hence, these models can only reproduce one-dimensional features, e.g., the particle spectrum of the TNSA process.

Sophisticated three-dimensional computer simulation techniques have been developed for a better understanding of the whole process of short-pulse high-intensity laser–matter interaction, electron transport and subsequent ion acceleration. The simulation methods can be classified as (i) particle-in-cell, (ii) Vlasov, (iii) Vlasov–Fokker–Planck, (iv) hybrid fluid or particle, and (v) gridless particle codes; see the short review in Ref. [23] for a description of each method.

The particle-in-cell method is the most widely used simulation technique. In this method, Maxwell's equations are solved, together with a description of the particle distribution functions. The method more or less resembles a 'numerical experiment' with only a few approximations; hence, a detailed insight into the dynamics can be obtained. The disadvantage is that no specific theory serves as an input parameter and the results must be analyzed like experimental results, i.e., they need to be interpreted and compared with analytical estimates.

2.5.1 *Plasma expansion model*

Plasma expansion is often described as an isothermal rarefaction wave into free space. There is quite a large similarity with the expansion models used to describe TNSA. The isothermal expansion model assumes *quasi-neutrality*, $n_e = Zn_i$, and a constant temperature T_e. Using the two-fluid hydrodynamic model for electrons and ions, the continuity, momentum, and energy conservation equations are used, usually with the assumption of an isothermal expansion (no temperature change in time), no further

source term (no laser), no heat conduction, collisions or external forces, and a pure electrostatic acceleration (no magnetic fields). One can find a self-similar solution [52]:

$$v(z,t) = c_{\mathrm{s}} + \frac{z}{t} , \qquad (20)$$

$$n_{\mathrm{e}}(z,t) = Zn_{\mathrm{i}}(z,t) = n_{\mathrm{e},0} \exp\left(-\frac{z}{c_{\mathrm{s}}t} - 1 \right) , \qquad (21)$$

where v denotes the bulk velocity and $n_{\mathrm{i}}(n_{\mathrm{e}})$ the evolution of the ion (electron) density. The rarefaction wave expands with the sound velocity $c_{\mathrm{s}}^2 = Zk_{\mathrm{B}}T_{\mathrm{e}}/m_{\mathrm{i}}$. By combining these two equations, replacing the velocity with the kinetic energy $v^2 = 2E_{\mathrm{kin}}/m$ and taking the derivative with respect to E_{kin}, the ion energy spectrum $\mathrm{d}N/\mathrm{d}E_{\mathrm{kin}}$ from the quasi-neutral solution per unit surface and per unit energy in dependence of the expansion time t is obtained [46]:

$$\frac{\mathrm{d}N}{\mathrm{d}E_{\mathrm{kin}}} = \frac{n_{\mathrm{e},0}c_{\mathrm{s}}t}{\sqrt{2Zk_{\mathrm{B}}T_{\mathrm{hot}}E_{\mathrm{kin}}}} \exp\left(-\sqrt{\frac{2E_{\mathrm{kin}}}{Zk_{\mathrm{B}}T_{\mathrm{hot}}}} \right) . \qquad (22)$$

The ion number N is obtained from the ion density as $N = n_{\mathrm{e},0}c_{\mathrm{s}}t$. Moreover, the electric field in the plasma is obtained from the electron momentum equation, $n_{\mathrm{e}}eE = -k_{\mathrm{B}}T_e\nabla n_e$, as

$$E = \frac{k_{\mathrm{B}}T_{\mathrm{e}}}{ec_{\mathrm{s}}t} = \frac{E_0}{\omega_{\mathrm{pi}}t}, \qquad (23)$$

with $E_0 = (n_{\mathrm{e},0}k_{\mathrm{B}}T_{\mathrm{e}}/\varepsilon_0)^{1/2}$, where $\omega_{\mathrm{pi}} = (n_{\mathrm{e},0}Ze^2/m_{\mathrm{i}}\varepsilon_0)^{1/2}$ denotes the ion plasma frequency. The electric field is uniform in space (i.e., constant) and decays with time as t^{-1}. The temporal scaling of the velocity is obtained by solving the equation of motion, $\dot{v} = Zq/mE$, with the electric field as before. This yields

$$v(t) = c_{\mathrm{s}}\ln(\omega_{\mathrm{pi}}t) + c_{\mathrm{s}} \qquad (24)$$
$$z(t) = c_{\mathrm{s}}t\left(\ln(\omega_{\mathrm{pi}}t) - 1\right) + c_{\mathrm{s}}\,t . \qquad (25)$$
$$\qquad (26)$$

Both equations satisfy the self-similar solution. The scaling of the ion density is found as $n(t) = n_0/\omega_{\mathrm{pi}}t$.

However, at $t = 0$, the self-similar solution is not defined and has a singularity. Hence, the model of a self-similar expansion is not valid for a description of TNSA at early times and must be modified. Additionally, in TNSA there are more differences. Firstly, the expansion is not driven by an electron distribution, being in equilibrium with the ion distribution, but by the relativistic hot electrons that are able to extend in the vacuum region in front of the ions. There, quasi-neutrality is strongly violated and a strong electric field will build up, modifying the self-similar expansion solution.

Secondly, the initial condition of equal ion and electron densities must be questioned, since the hot-electron density with $n_{\mathrm{e}} \approx 10^{20}$ cm^{-3} is about three orders of magnitude less than the solid density of the rear-side contamination layers. This argument can only be overcome by the assumption of a global quasi-neutrality condition $Zn_{\mathrm{i}} = n_{\mathrm{e}}$.

Thirdly, it might not be reasonable to assume a model of an isothermal plasma expansion. It can be assumed, however, that the expansion is isothermal, since the laser pulse provides 'fresh' electrons from the front side of the target, i.e., the assumption is valid for the laser pulse duration τ_{L}. As will be shown, the main acceleration time period is of the order of the laser pulse duration. This justifies the assumption of an isothermal expansion.

The plasma expansion, including charge separation, was quantitatively described by Mora [46–48] with high accuracy. The main point of this model is a plasma expansion with charge separation at the ion front, in contrast with a conventional, self-similar plasma expansion. The plasma consists of

electrons and protons, with a step-like initial ion distribution and an electron ensemble that is in thermal equilibrium with its potential. The megaelectronvolt electron temperature results in a charge separation being present for long times. It leads to enhanced ion acceleration at the front, compared with the case of a normal plasma expansion. This difference is sometimes named the TNSA effect.

Although only one-dimensional, the model has been successfully applied to experimental data at more than 10 high-intensity short-pulse laser systems worldwide in a recent study [16]. It was separately used to explain measurements taken at the ATLAS-10 at the Max-Planck Institute in Garching, Germany [53] as well as to explain results obtained at the VULCAN PW [54] (with a few modifications). Therefore, it is seen as a reference model, currently used worldwide for an explanation of TNSA. Because of its success in describing TNSA, it will be explained in more detail now.

After the laser acceleration at the foil's front side, the electrons arrive at its rear side and escape into the vacuum. The atoms are assumed to be instantly field ionized, leading to $n_i = n_e/Z$. Charge separation occurs and leads to an electric potential ϕ, according to Poisson's equation:

$$\varepsilon_0 \frac{\partial^2 \phi}{\partial z^2} = e\left(n_e(z) - n_i(z)\right).$$ (27)

The electron density distribution is always assumed to be in local thermal equilibrium with its potential,

$$n_e = n_{e,0} \exp\left(\frac{e\phi}{k_B T_{hot}}\right),$$ (28)

where the electron kinetic energy is replaced by the potential energy $e\phi$. The initial electron density $n_{e,0}$ is taken from Eq. (8). The ions are assumed be of initial constant density $n_i = n_{e,0}$, with a sudden drop to zero at the vacuum interface. The boundary conditions are chosen such that the solid matter in one half-space ($z \leq 0$) perfectly compensates the electric potential for $z \to -\infty$, whereas for $z \to \infty$ the potential goes to infinity. Its derivative $E = -\partial\phi/\partial z$ vanishes for $z \to \pm\infty$. In the vacuum region (initially $z > 0$), the field can be obtained analytically [37]. The resulting potential is

$$\phi(z) = -\frac{2k_B T_{hot}}{e} \ln\left(1 + \frac{z}{\sqrt{2\exp(1)}\lambda_{D,0}}\right) - \frac{k_B T_{hot}}{e},$$ (29)

and the corresponding electric field reads

$$E(z) = \frac{2k_B T_{hot}}{e} \frac{1}{z + \sqrt{2\exp(1)}\lambda_{D,0}}.$$ (30)

The initial electron Debye length is $\lambda_{D,0}^2 = \varepsilon_0 k_B T_{hot}/e^2 n_{e,0}$. The full boundary value problem, including the ion distribution, can only be solved numerically. The result obtained with MATLAB [43] is shown in Fig. 4. The potential ϕ (red curve) is a smooth function and is in perfect agreement with the analytical solution Eq. (29) (grey curve) in the vacuum region. Both are given in units of $k_B T_{hot}/e$. The electron density n_e (green curve), normalized to $n_{e,0}$, follows from Eq. (28). The normalized ion density n_i (in black) is a step function with $n_i(z < 0)/n_{e,0} = 1$ and zero for $z > 0$. The electric field E (blue curve) has a strong peak at the ion front, with $E_{max} = \sqrt{2/\exp(1)}E_0 = 0.86E_0$. The normalization field E_0 is given by $E_0 = k_B T_{hot}/e\lambda_{D,0}$. The z coordinate was normalized with $\lambda_{D,0}$. The subsequent plasma expansion into vacuum is described in the framework of a fluid model, governed by the equation of continuity (left) and the momentum balance (right):

$$\frac{\partial n_i}{\partial t} + \frac{\partial(v_i n_i)}{\partial z}) = 0 \qquad\qquad \frac{\partial v_i}{\partial t} + v_i\frac{\partial v_i}{\partial t} = -\frac{e}{m_p}\frac{\partial \phi}{\partial z}.$$ (31)

The full expansion dynamics can only be obtained numerically. Of particular interest is the temporal evolution of the ion distribution and the evolution of the electric field driving the expansion of
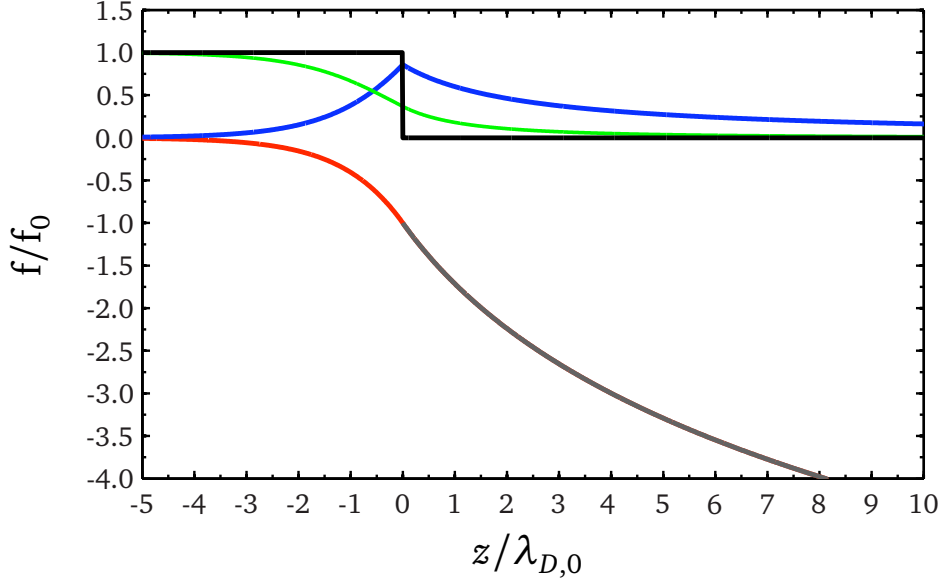
Fig. 4: Solution of Eq. (27). The potential ϕ (red) was obtained numerically. The analytical solution Eq. (29) (grey) is in perfect agreement. Both are given in units of $k_B T_{hot}/e$. The electron density n_e (green), normalized to $n_{e,0}$, follows from Eq. (28). The normalized ion density n_i (black) is a step function with $n_i(z < 0)/n_{e,0} = 1$ and zero for $z > 0$. The electric field E (blue) is given in units of $k_B T_{hot}/e\lambda_{D,0}$. The z coordinate is given in units of $\lambda_{D,0}$.

the bulk. In [5] a Lagrangian code in MATLAB was developed, which solves Eqs. (28), (29), and (31), similar to Ref. [46]. The numerical method is similar to the method described in Ref. [53]; however, the developed code uses MATLAB's built-in bvp4c function for a numerical solution of the boundary value problem in the ion fluid. The initially constant ion distribution is divided into a grid, choosing the left boundary to be $L \gg c_s t$. The boundary value for the potential is $\phi(-L) = 0$. At the right boundary (initially at $z = 0$), the electric field $-\partial\phi_{front}/\partial z = \sqrt{2/e}k_B T_{hot}/e\lambda_{D,front}$ must coincide with the analytical solution of Eq. (30), where the local Debye length must be determined by the potential at the front:

$$\lambda_{D,front} = \lambda_{D,0} \exp\left(\frac{e\phi_{front}}{k_B T_{hot}}\right)^{-1/2}. \tag{32}$$

Initially, the Debye length at the ion front is obtained by inserting Eq. (29) in Eq. (28) to give $\lambda_{D,0,front} = e^{-1}\lambda_{D,0}$. The code divides the fluid region into a regular grid. Each grid element (cell) has position z_j, ion density n_j, and velocity v_j. For each time-step Δt, the individual grid elements are moved according to the following scheme [53]:

$$z_{j'} = z_j + v_j\Delta t + \frac{e}{2m_p}E\Delta t^2, \tag{33}$$

$$v_{j'} = v_j + \frac{e}{m_p}E\Delta t. \tag{34}$$

After that, the density of the cell is changed according to the broadening of the cell due to the movement:

$$n_{j'} = n_j \frac{\Delta x_j}{\Delta x_{j'}}. \tag{35}$$

At the front, the individual cells quickly move forward, resulting in a 'blow-up' of the cells, which dramatically diminishes the resolution. Thus, after each time-step, the calculation grid is mapped onto a

new grid ranging from z_{min} to the ion front position z_{front} with an adapted cell spacing. This method is called rezoning. The new values of v_j and n_j are obtained by third-order spline interpolation, providing very good accuracy.

2.5.2 *Temporal evolution and scaling*

A crucial point in the ion expansion is the evolution of the electric field strength E_{front}, the ion velocity v_{front} and the position z_{front} of the ion front. Expressions given by Mora are [46–48]

$$E_{front} \simeq \left(\frac{2n_{e,0}k_B T_{hot}}{e\varepsilon_0} \frac{1}{1+\tau^2} \right)^{1/2}, \tag{36}$$

$$v_{front} \simeq 2c_s \ln\left(\tau + \sqrt{1+\tau^2} \right), \tag{37}$$

$$z_{front} \simeq 2\sqrt{2e}\lambda_{D,0} \left[\tau \ln\left(\tau + \sqrt{1+\tau^2} \right) - \sqrt{1+\tau^2} + 1 \right], \tag{38}$$

where $e = \exp(1)$ and $\tau = \omega_{pi}t/\sqrt{2e}$. The other variables in these equations are the initial ion density $n_{i,0}$, the ion-acoustic (or sound) velocity $c_s = (Zk_B T_{hot}/m_i)^{1/2}$, T_{hot} is the hot-electron temperature and $\omega_{pi} = (n_{e,0}Ze^2/m_i\varepsilon_0)^{1/2}$ denotes the ion plasma frequency. Owing to the charge separation, the ion front expands more than twice as quickly as the quasi-neutral solution in Eqs. (25) and (26). From Eq. (37), the maximum ion energy is given as

$$E_{max} = 2k_B T_{hot} \ln^2\left(\tau + \sqrt{1+\tau^2} \right). \tag{39}$$

The particle spectrum from Mora's model [46–48] cannot be given in an analytic form, but it is very close to the spectrum of Eq. (22), obtained by the self-similar motion of a fully quasi-neutral plasma expanding into a vacuum. The phrase 'fully quasi-neutral' should indicated that in this solution there is no charge separation at the ion front, hence there is no peak electric field.

A drawback of the model is the infinitely increasing energy and velocity of the ions with time, which results from the assumption of an isothermal expansion. Hence, a stopping condition must be defined. An obvious time duration for the stopping condition is the laser pulse duration τ_L. However, as found by Fuchs *et al.* [16, 45], the model can be successfully applied to measured maximum energies and spectra, as well as to particle-in-cell simulations, if the calculation is stopped at $\tau_{acc} = \alpha(\tau_L + t_{min})$. It was found that, for very short pulse durations. the acceleration time τ_{acc} tends towards a constant value $t_{min} = 60$ fs, which is the minimum time that the energy transfer from the electrons to the ions needs. The variable α takes into account that for lower laser intensities the expansion is slower and the acceleration time must be increased. α varies linearly from 3 at an intensity of $I_L = 2 \times 10^{18}$ W/cm^2 to 1.3 at $I_L = 3 \times 10^{19}$ W/cm^2. For higher intensities, α is constant, at 1.3. Hence, the acceleration time is

$$\tau_{acc} = \left(-6.07 \times 10^{-20} \times (I_L - 2 \times 10^{18}) + 3 \right) \times (\tau_L + t_{min}) \tag{40}$$

for $I_L \in [2 \times 10^{18} - 3 \times 10^{19}]$ W/cm^2, and

$$\tau_{acc} = 1.3 \times (\tau_L + t_{min}) \tag{41}$$

for $I_L \geq 3 \times 10^{19}$ W/cm^2.

Figure 5 shows the temporal evolution of the electric field and the ion velocity at the ion front, respectively. The electric field was normalized to E_0, the ion velocity is divided by the sound velocity. There is a very good agreement between the simulated values (blue circles) and the expressions by Mora [46–48] from Eqs. (36) and (37) (red curve). The maximum deviations from the scaling expressions are 1.6% for the electric field and 0.4% for the velocity. The electric field evolution and the development of the electron and ion density profiles are shown in Fig. 6. The electric field (green curve) peaks sharply
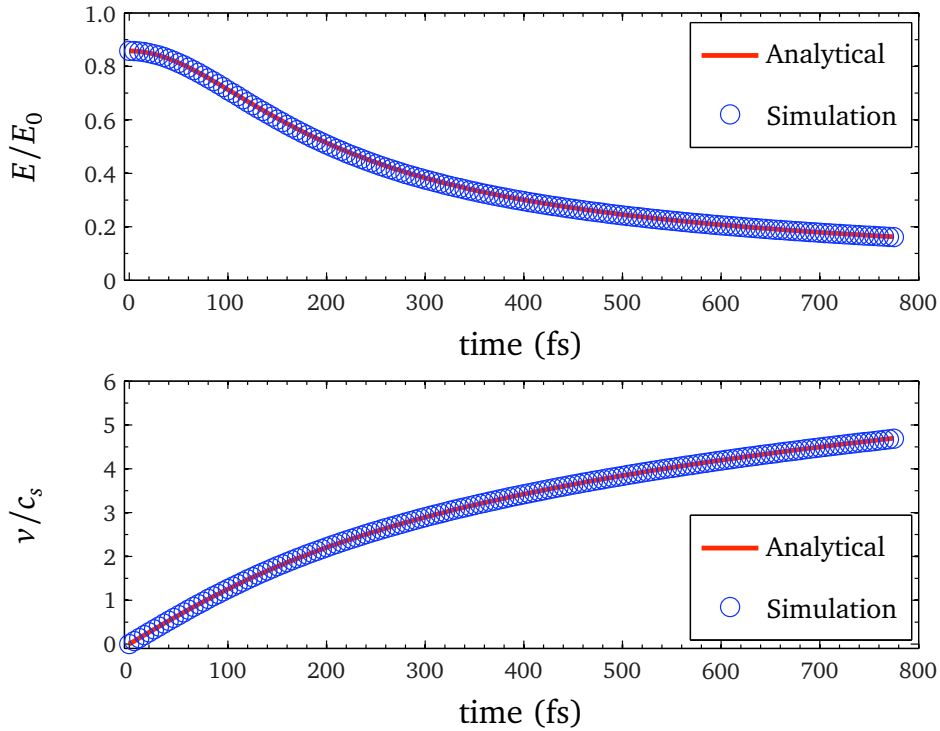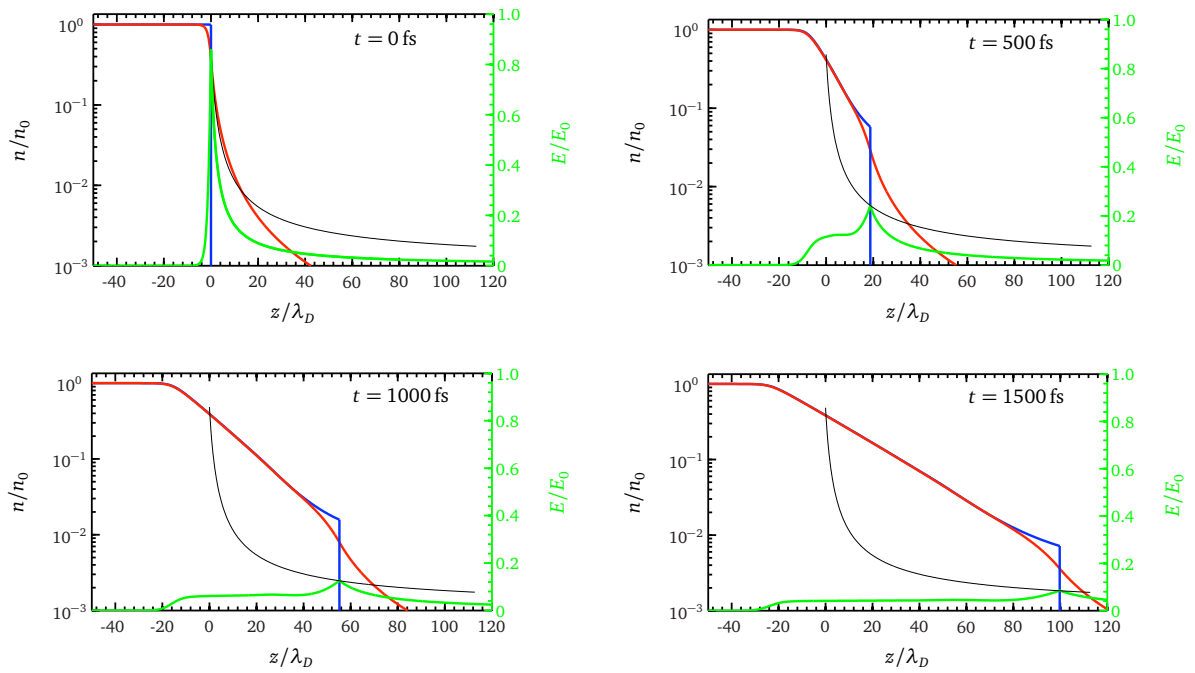
Fig. 5: Temporal evolution of the electric field and the ion velocity at the ion front. There is very good agreement between simulated values (blue circles) and Eqs. (36) and (37) (red curve).

at the ion front for all times. Initially, the ion density (blue curve) is $n_i = n_0$ for $z \leq 0$ and zero for $z > 0$. The electron density (red curve) is infinite and decays proportionally to z^{-2}. Note the different axis scalings for the electric field and the densities; the latter are plotted on a logarithmic scale. For later times, at $t = (500, 1000, 1500)$ fs, the ions are expanded, forming an exponentially decaying profile.

A large part of the expanding plasma is quasi-neutral and can be identified by the constant electric field, as derived in Eq. (23). At the ion front, the charge separation is still present, leading to an enhanced electric field that is a factor of two greater than the electric field in the bulk, in agreement with Ref. [46]. This scaling is maintained for the whole expansion. The scaling of the peak electric field value at the ion front at position z, as given by the analytical expressions in Eqs. (36) and (38), is in perfect agreement with the simulation.

The final proton spectrum is shown in Fig. 7. The numerical solution (blue curve) is close to the analytical solution from the quasi-neutral model of Eq. (22) (red curve). The analytical spectrum is assumed to reach a maximum energy, taken from Eq. (39). The maximum energy in the simulation is $E_{max,num} = 19$ MeV, which is in close agreement to the analytical value of $E_{max,analyt} = 18.5$ MeV. As expected, there is excellent agreement in the spectra for low energies, since in both cases the expansion is quasi-neutral. For high energies, the numerical spectrum deviates from the self-similar model. The numerical spectrum is lower than the self-similar one even though the ion density of the numerical solution increases close to the ion front, as can be seen in Fig. 6 in the deviation of the electron and ion densities close to the front. However, the velocity increase at the front in the simulation is much faster than that in the self-similar solution, owing to the enhanced electric field. Thus, the kinetic energy of the fluid elements close to the ion front is greater than the kinetic energy of fluid elements in a self-similar expansion. The spectrum is obtained by taking the derivative of the ion density with respect to the kinetic energy. It turns out that the kinetic energy increases more strongly than the ion density, hence dN/dE is a little less than the self-similar expansion.

Fig. 6: Temporal evolution of the electric field and the ion and electron density, respectively. The electric field (green curve) peaks sharply at the ion front. The ion density (blue curve) is $n_i = n_0$ for $z \leq 0$ and zero for $z > 0$ for $t = 0$. The electron density (red curve) decays proportionally to z^{-2}. For later times, at $t = (500, 1000, 1500)$ fs, the ions are expanded, forming an exponentially decaying profile.
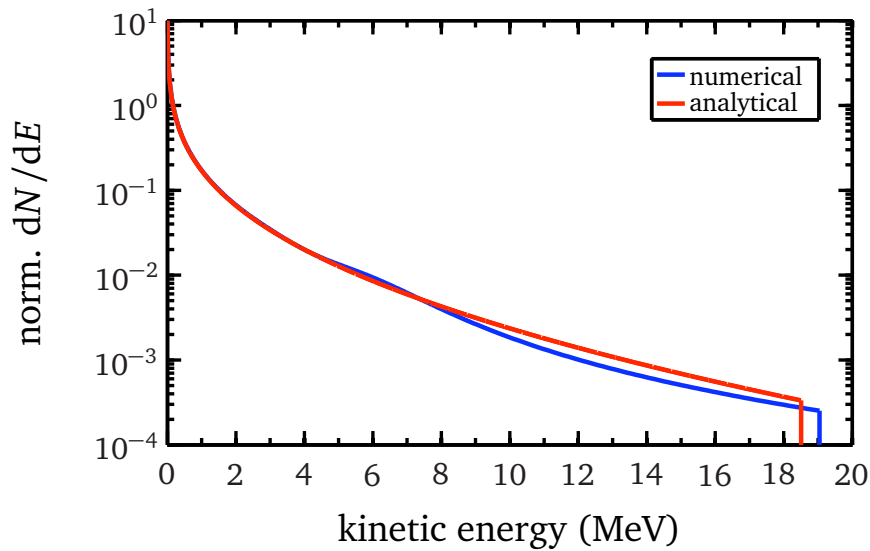
Fig. 7: Energy spectrum dN/dE from the simulation (blue curve) compared to the spectrum of a quasi-neutral plasma expansion (red curve). norm, normalized.

In conclusion, the Lagrangian code and the model developed by Mora show that TNSA-accelerated ions are mainly emitted in the form of a quasi-neutral plasma, with a charge separation at the ion front that leads to an enhanced acceleration compared with the expansion of a completely quasi-neutral plasma. For later times, if $\omega_{pi}t \gg 1$, the analytical expression of the maximum ion energy in Eq. (39) can be used to determine the cut-off energy of TNSA-accelerated ions accurately. The spectral shape of the ions is close to the spectrum of a quasi-neutral, self-similar expansion.

The equations show, that the maximum energy, as well as the spectral shape, strongly scale with the hot-electron temperature. The expression for the initial electric field scales as $E \propto k_B T_{hot} n_e$; hence, a simplistic estimate would assume that both are equally important for the maximum ion energy. In contradiction, however, the investigation has shown that the maximum ion energy only depends weakly on the hot-electron density and is directly proportional to the hot-electron temperature. It is worth noting that this finding is in agreement with results obtained earlier with an electrostatic particle-in-cell code by Brambrink [55]. The hot-electron density—owing to the quasi-neutrality boundary condition— determines the number of generated ions. Both the number of ions and the energy increase with time, again showing that not the shortest and most intense laser pulses are favourable for TNSA, but somewhat longer pulses, of the order of a picosecond. This requires a high laser energy to keep the intensity sufficiently high.

Nevertheless, the model is still very idealized, since it is one-dimensional and isothermal, with the electrons ranging into infinity, and it neglects laser interaction and electron transport. An approach with electrons in a Maxwellian distribution always leads to the same asymptotic behaviour of the ion density [56], hence two-temperature [57] or even tailored [58] electron distributions will lead to different ion distributions. There are many alternative approaches to the one described here, including, e.g., an adiabatic expansion [48]; multitemperature effects [48, 57]; an approach in which an upper integration range is introduced to satisfy the energy conservation for the range of a test electron in the potential [50]; expansion of an initially Gaussian-shaped plasma [47]; and expansion of a plasma with an initial density gradient [59]. Most of these approaches assume an underlying fluid model, where particle collisions are neglected and the fluid elements are not allowed to overtake each other. Hence a possible wave-breaking or accumulation of particles is not included in the models but requires a kinetic description [60, 61]. Furthermore, the transverse distribution of the accelerated ions cannot be determined from a one-dimensional model and requires further modelling. This can be done in the framework of two-dimensional particle-in-cell simulations. Particle-in-cell simulations allow a much more sophisticated description, including relativistic laser–plasma interaction, a kinetic treatment of the particles, and a fully three-dimensional approach.

3 Target normal sheath acceleration: ion beam characteristics

Part of the motivation of the extensive research on laser-accelerated ion beams is based on their exceptional properties (high brightness and high spectral cut-off, high directionality and laminarity, short pulse duration), which distinguish them from those of the lower-energy ions accelerated in earlier experiments at moderate laser intensities. In view of these properties, laser-driven ion beams can be employed in a number of groundbreaking applications in science, technology, and medicine. This section reviews the main beam parameters; Section 4 focuses on established and proposed applications using these unique beam properties.

3.1 Beam parameters

3.1.1 Particle numbers

One of the striking features of TNSA-accelerated ion beams is the fact that the particle number in a forward-directed beam is very high. At present, particle numbers of up to 6×10^{13} protons with energies above 4 MeV have been detected in experiments. This typically leads to a conversion efficiency of laser

to ion beam energy of up to 9%. At these high particle numbers, drawn from a very limited source size, for high-energy short-pulse laser systems, the depletion of the proton contamination layer at the rear surface becomes an issue. This has been addressed by Allen *et al.* [62], who determined that there are 2.24×10^{23} atoms/cm^3 at the rear surface of a gold foil, in a layer 1.2 Å thick. Assuming an area of about 200 µm diameter, the accelerated volume is about $V = 3.8 \times 10^{-11}$ cm^{-3}. Hence, the total number of protons in this area is about $N_{total} = 8.4 \times 10^{12}$, which is close to the integrated number determined in the experiments. Experiments have shown [8] that a rear surface coating of a metal target can provide enough protons up to a thickness of ≈ 100 nm, where the layer thickness causes the onset of instabilities in the electron propagation, owing to its limited electrical conductivity.

3.1.2 Energy spectrum

Based on the acceleration mechanism and the expansion model described earlier, the usual ion energy distribution is an exponential one with a cut-off energy that is dependent on the driving electron temperature. Without special target treatment, and independently of the target material, protons are always accelerated first, as they have the highest charge-to-mass ratio. These protons stem from water vapour and hydrocarbon contamination, which are always present on the target surface, owing to the limited achievable vacuum conditions. Protons from the top-most contamination layer on the target surface are exposed to the highest field gradients and screen the electric field for protons and ions coming from the successive layers. The acceleration of particles from different target depths results in a broad energy distribution, which becomes broader with increasing contamination layer thickness. The inhomogeneous electron distribution in the sheath also leads to an inhomogeneous accelerating field in the transverse direction. The resulting exponential ion energy spectrum constitutes the main disadvantage in laser ion acceleration.

Only three groups so far [63–65] have produced a quasi-monoenergetic ion beam with lasers and an energy spread of 20% or less. Hegelich *et al.* [63] have used 20 µm thick palladium foils that were resistively heated before the acceleration. At temperatures of about 600 K, the targets were completely dehydrogenized, but carbon atoms still remained on the surface. By increasing the target temperature ($T > 1100$ K), the carbon underwent a phase transition and formed a monolayer or graphite (graphene) on the palladium surface, from which C^{5+} ions were accelerated to 3 MeV/u with an energy spread of 17%. An advantage of resistive heating is the complete removal of all hydrogen at once but there are also several disadvantages. The formation of graphene cannot be controlled and the set-up requires a very precise temperature measurement.

Schwoerer *et al.* [64] used 5 µm thick titanium foils coated with 0.5 µm thick hydrogen-rich poly-methylmethacrylate dots of 20 µm × 20 µm on the target's rear surface. This configuration was designed to limit the transverse extension of protons, so that the proton-rich dot would have a smaller diameter than the scale of inhomogeneity of the electron sheath. In this case, all protons experience the same potential. The parasitic proton contamination layer could be reduced by nanosecond-laser ablation; the accelerated protons showed a quasi-monoenergetic energy spectrum peaking at 1.2 MeV.

Ter-Avetisyan *et al.* [65] produced quasi-monoenergetic deuteron bursts by the interaction of a high-intensity, high-contrast ($>10^{-8}$) laser with limited-mass water droplets. The peak position in the spectrum, at 2 MeV, had an energy spread of 20%. This experiment, however, suffered from a low laser–droplet interaction probability.

3.1.3 Opening angle

Figure 8 shows the energy-resolved opening angles for data obtained at Trident (blue circles), LULI-100 TW (green circles), and Z-Petawatt (red circles). The plots have been normalized to the respective maximum energy of each beam. These are 19 MeV for Trident, 16.3 MeV for LULI-100 TW, and 20.3 MeV for Z-Petawatt, respectively.
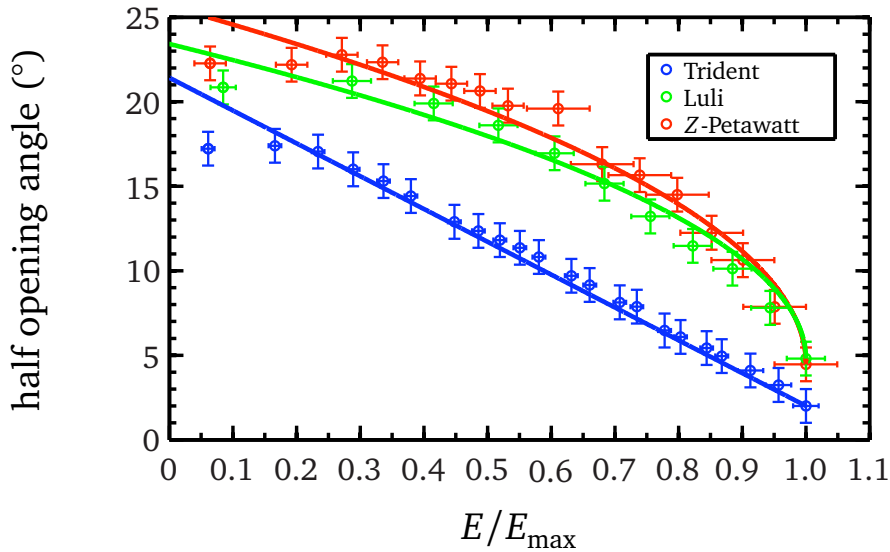
Fig. 8: Energy dependence of the half opening angle. Data obtained at Trident (blue circles), LULI-100 TW (green circles), and Z-Petawatt (red circles). The plots have been normalized to the respective maximum energy of each beam. The opening angle decreases with increasing energy. A parabolic dependency could be fit to the LULI and Z-Petawatt results, the data for Trident produce a linear slope.

Protons with the highest energy are emitted with the smallest opening angle from the source, up to a 5° half angle. Protons with less energy are subsequently emitted in larger opening angles. Below about 30% maximum energy, the opening angle reaches a maximum and stays constant for lower energies. In most cases, the opening angles decrease parabolically with increasing energy, as shown in Fig. 8. In some shots, however, the decrease of opening angle with increasing energy is close to linear, as in the example obtained at Trident. The slope of the opening angle with energy is a result of the initial hot-electron sheath shape at the target surface, as pointed out by Carroll *et al.* [66]. According to Ref. [66], a sheath with Gaussian dependence in the transverse direction results in a strongly curved opening angle–energy distribution, whereas a parabolic hot-electron sheath results in a linear dependency. However, only crude details of the exact modelling of the acceleration process are given in Ref. [66]. In Ref. [5], a more detailed expansion model is described, which is able to explain the experimental results in more detail. It should be noted that the term 'opening angle' is not equivalent to the beam 'divergence'. The divergence of the protons increases slightly with increasing energy, whereas the emitting area (source size) decreases with proton energy [6, 67]. This results in a total decrease of the opening angle measured experimentally.

3.1.4 *Source size*

Figure 9 shows energy-resolved source sizes for the three laser systems Trident (blue circle), LULI-100 TW (green circles), and Z-Petawatt (red circles). As in Section 3.1.3, the energy axis has been normalized to the individual maximum energy of the shot, with the maximum energies as before. Source size decreases with increasing energy. Protons with the highest energies are emitted from sources of about 10 μm diameter and less. For lower energies, the source sizes progressively increase, up to about 200 μm diameter for the lowest energies measurable with radiochromic film imaging spectroscopy, about 1.5 MeV. For even lower energies, the source sizes might be much larger and could reach in excess of 0.5 mm in diameter [42].
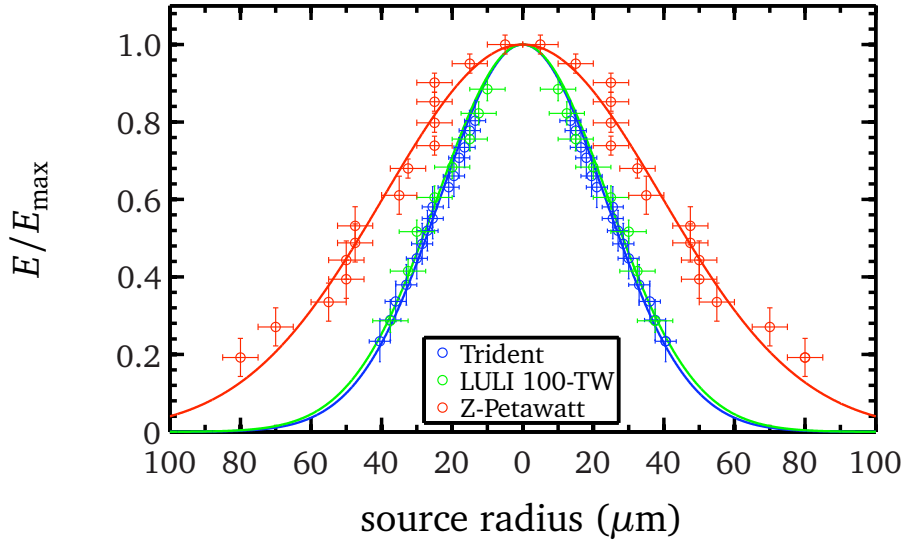
Fig. 9: Energy-resolved source sizes for data from Trident (blue circles), LULI-100 TW (green circles) and Z-Petawatt (red circles). The energy-source-size distribution fits to a Lorentzian with full width at half maximum of 54.8 μm for Trident, 56.5 μm for LULI-100 TW, and 92.8 μm for Z-Petawatt.

The energy dependence of the source size well fits a Gaussian, indicated in Fig. 9. The data could be fit by

$$E = \exp\left(\frac{-(4\ln(2)\text{source size})^2}{\sigma^2}\right) , \qquad (42)$$

where 2σ denotes the FWHM. This fit allows the complete energy-dependent source size to be characterized using only one parameter. The FWHM for Trident with a 10 μm thin gold target is $\sigma = 54.8$ μm. For LULI-100 TW, the source size is $\sigma = 56.5$ μm for a 15 μm thin gold foil. A larger source size has been measured at Z-Petawatt, with $\sigma = 92.8$ μm and a 25 μm thick gold target.

The energy dependence of the source size is directly related to the electric field strength distribution of the accelerating hot-electron sheath at the source. Protons with high energies have been accelerated by a high electric field. Cowan *et al.* [67] relate increasing source size with decreasing energy to the shape of the hot-electron sheath, under the assumption of an isothermal quasi-neutral plasma expansion where the electric field is $E = -(k_B T_e/e)(\delta n_e/n_e)$. A transversely Gaussian electric field distribution would result in a non-analytical expression for the electron density, n_e. Conversely, a realistic assumption of a Gaussian hot-electron distribution would result in a radially linearly increasing electric field, in contradiction to the measured data. Hence, it is concluded that the quasi-neutral plasma expansion, although it is the driving acceleration mechanism for later times, does not explain the observed source sizes. In fact, the source size must develop earlier in the acceleration process, e.g., at very early times, when the electric field is governed by the Poisson equation (Eq. (10)), with $E(z) \propto k_B T_e/\lambda_D \propto \sqrt{k_B T_e n_e}$. With the data from Fig. 9, it can be concluded that there must be a radial dependency of $E(z)$, and hence a Gaussian decay of either the hot-electron temperature or the density, or both.

3.1.5 Emittance

As we have seen, the acceleration of the ions by the TNSA mechanism basically constitutes a quasi-electrostatic field acceleration of initially cold (room-temperature) atoms at rest, which are field ionized and then pulled by the charge-separation field. As the electrons are scattered while being pushed through the target, at least for materials with enough electrical conductivity to provide the return currents, the transport smooths the distribution at the rear surface to result in a Gaussian-like field shape that expands

laterally in time as the electrons start to recirculate. Thus the initial random ion movement in the beam, represented in an extended phase space and often regarded as an effective beam temperature, is very low.

An important parameter in accelerator physics is the transverse emittance of an ion beam. In view of the nature of the ion sources used in conventional accelerators, there is always a spread in kinetic energy and velocity in a particle beam. Each point on the surface of the source emits protons with different initial magnitude and direction of the velocity vector. The emittance ε provides a figure of merit for describing the quality of the beam, i.e., its laminarity [68]. Assume that the beam propagates in the z-direction. Each proton represents a point in the position-momentum space (x, p_x and y, p_y), the phase space. The transverse phase space (e.g., in the x-direction) of the TNSA protons is obtained by mapping the source position (indicated for example by imprinted surface grooves in the radiochromic film) versus the angle of emission p_x/p_z, obtained by the position x of the imprinted line in the radiochromic film and the distance d by $p_x/p_z = x' = \arctan(x/d)$.

In general, the quality of charged-particle beams is characterized by their emittance, which is proportional to the volume of the bounding ellipsoid of the distribution of particles in phase space. By Liouville's theorem, the phase space volume of a particle ensemble is conserved during non-dissipative acceleration and focusing. For the transverse phase space dimensions, the area of the bounding phase space ellipse equals $\pi\varepsilon$, where the emittance ε, at a specific beam energy (or momentum p), is expressed in normalized root-mean-square (r.m.s.) units as

$$\varepsilon_{\mathrm{N,rms}} = (p/mc)\left[<x^2><x'^2> - <xx'>^2\right]^{1/2}. \tag{43}$$

In Eq. (43), m is the ion mass, c is the velocity of light, x is the particle position within the beam envelope, and $x' = p_x/p_z$ is the particle's divergence in the x-direction. At a beam waist, Eq. (43) reduces to $\varepsilon_{\mathrm{N,rms}} = \beta\gamma\sigma_x\sigma_{x'}$, where σ_x and $\sigma_{x'}$ are the r.m.s. values of the beam width and divergence angle. Several effects contribute to the overall emittance of a beam: its intrinsic transverse 'thermal' spread; intrabeam space charge forces; and non-ideal accelerating fields, for example, at apertures in the source or accelerator. For typical proton accelerators (e.g., the CERN SPS or FNAL-Tevatron), the emittance at the proton source is $\approx$0.5 mm·mrad (normalized r.m.s.) and up to 20–80 mm·mrad within the synchrotron. The longitudinal phase space (z–p_z) is characterized by the equivalent energy–time product of the beam envelope; for the CERN SpS, a typical value is $\sim$0.1 eV·s.

The highest quality ion beams have the lowest values of transverse and longitudinal emittance, indicating a low effective ion temperature and a high degree of angle-space and time-energy correlation. In typical TNSA experiments, one may estimate an upper limit for the transverse emittance of the proton beam from Eq. (43), by assuming that initial beamlets were initially focused to a size $\ll$100nm, and that the observed width is entirely due to the initial width of the x' distribution. The upper limit of the emittance for 12 MeV protons is $<$0.002 mm·mrad. This is a factor of $>$100 smaller than typical proton beam sources, and we can attribute this to the fact that during much of the acceleration the proton space charge is neutralized by the co-moving hot electrons, and that the sheath electric field self-consistently evolves with the ions to produce an effectively 'ideal' accelerating structure. The remaining irreducible 'thermal' emittance would imply a proton source temperature of $<$100 eV. The energy spread of the laser-accelerated proton beam is large, ranging from zero to tens of megaelectronvolts; however, owing to the extremely short duration of the accelerating field ($<$10 ps), the longitudinal phase space energy–time product must be less than 10^{-4} eV·s. More details can be found in Ref. [67], from which part of this section was extracted.

3.1.6 Ion species

Since protons are the lightest ions, and have the highest charge-to-mass ratio, they are favoured by the acceleration processes. The protons are present in the target as surface contaminants or in compounds of the target itself or of the target coating. The cloud of accelerated protons then screens the electric field
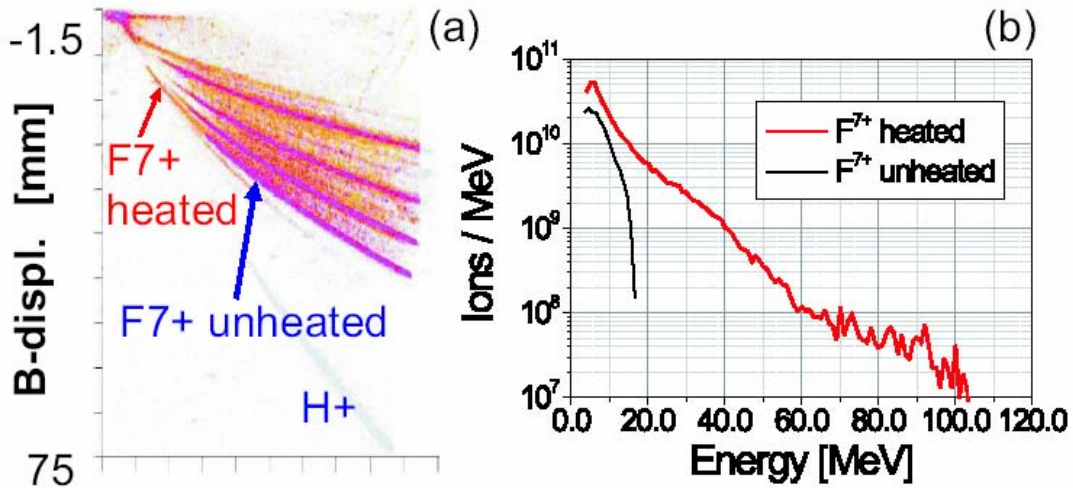
Fig. 10: (a) Overlaid signals of heated and non-heated (blue) W/CaF_2 targets: the proton signals vanish for heated targets, the fluorine signals (especially F^{7+}) go up to much higher energies. (b) Corresponding F^{7+} spectra: more than 5 MeV per nucleon are achieved for F^{7+} ions. From Ref. [40].

generated by the electrons for all the other ion species. The key for the efficient acceleration of heavy ions is the removal of any proton or light ion contaminants. In a few cases, heating of the target was performed prior to the experiments, to eliminate the hydrogen contaminants as much as possible, and to obtain a better, more controllable ion acceleration. In particular, by removing the proton from the targets, or by choosing H-free targets, the acceleration of heavier ions was favoured. For the latter case, recent experiments have demonstrated heavy-ion acceleration of up to more than 5 MeV/u, which corresponds to ion energies that are usually available at the end of a conventional accelerator of hundreds of metres in length [40].

First attempts to remove the hydrogen contaminants used resistively heated Al targets to temperatures of a few hundred degrees Celsius. The partial removal of the hydrogenous contaminants already strongly enhanced the acceleration of carbon ions [8]. A reduction by a factor of ten increased the energy of the carbon ions by a factor of 2.5 and their number by two orders of magnitude. Using tungsten as a thermally stable target resist and coating the rear surface with the material of interest, the target could be heated to more than 1200 degrees Celsius. Such targets show no accelerated protons at all, but instead a strongly increased contribution of heavy ions. The maximum energy could be enhanced by a factor of five compared with aluminium targets and the conversion efficiency again by a factor of ten [40] (see Fig. 10). Where the ohmic heating of target materials of interest is prohibited, owing to a low evaporation point, laser heating has proved an appropriate option for removing the contamination layers. In that case, the intensity of the laser heating the rear surface and evaporating the proton layers and the timing, with respect to the short pulse, have to be matched carefully. The energy spectrum of the heavy ions, together with their charge-state distribution also provides detailed information about the accelerating electric field at the rear surface. It was shown that, in a typical experiment, collisional ionization and recombination in flight is negligible, and so the detected charge states directly image the electric field strength because of the field ionization process. The field strengths that have been obtained match the estimated field strength, also predicted by theory, very well, and range from 10^{11} V/m to $\sim 10^{12}$ V/m.

The accelerating field deduced from ion acceleration is also consistent with observed proton energies with non-heated targets. For example, in typical experiments, fluorine ions were accelerated up to 100 MeV, i.e., more than 5 MeV/nucleon at a maximum charge state of 7^+. This corresponds to an electric field of 2 TV/m, which would have accelerated protons, if present, to energies of up to 25 MeV.

These were exactly the maximum energies found in experiments with non-heated targets under similar experimental conditions. Furthermore, the conversion efficiency is very high. Similar to the results obtained for proton beams, conversion efficiencies of up to 4% from laser to ion beam energy have been measured. Because of the same accelerating mechanism for protons and heavy ions, an excellent beam quality was expected. This section was extracted from [6].

3.2 Target dependence

In the previous section, we have looked extensively at the influence of target thickness and conductivity on the driving electron sheath distribution. To summarize, a highly conducting ultra-thin target is the most favourable for efficient ion acceleration. Moreover, as the electrons can distribute a part of the energy provided by the laser into bremsstrahlung, a low-Z material is preferable. The ultimate thickness of the target is determined by the limited laser contrast, as TNSA requires a sharp density gradient at the rear surface. For effective ion acceleration, an undisturbed back surface of the target is crucial to provide a sharp ion density gradient, since the accelerating field strength is proportional to T_{hot}/el_0, where T_{hot} is the temperature of the hot electrons and l_0 is the larger of either the hot-electron Debye length or the ion scale length of the plasma on the rear surface. The limited contrast of the laser causes a shockwave, launched by the prepulse, that penetrates the target and causes a rarefaction wave that diminishes the density gradient at the back, thereby drastically reducing the accelerating field. The inward moving shockwave also alters the initial conditions of the target material due to shockwave heating and therefore changes, e.g., the target density and conductivity. As a trade-off, one has, however, to note that a certain preplasma at the front side is beneficial to the production of hot electrons, somehow contradicting the need for high-contrast lasers. Also, as the lateral expansion of the electron sheath affects the evolution of the ion acceleration, it has been found that to confine the electron by reducing the transverse dimensions of the targets also enhances the ion particle energy. So the ideal target would resemble an ultra-thin, low-Z, highly conducting target with small lateral dimensions and a large preplasma at the front side.

Meanwhile, high-contrast laser systems are able to irradiate targets as thin as only a few nanometres, and we begin to see the change in the accelerating mechanism to radiation pressure acceleration [70] or laser breakout afterburner type acceleration [71], which is beyond the scope of this paper.

Apart from maximizing the ion beam particle energy, targets can be used to shape the beam for the applications listed next. Ballistic focusing [8, 72–75] and defocusing [8] have been demonstrated by numerous groups, tailoring on a nanoscale using microstructured targets [76] and layered targets to modify the shape of the energy spectrum [64].

3.3 Beam control

Ballistic focusing of laser-accelerated proton beams has been known since 2003 [72] and has been investigated in detail because of the large importance for proton-driven fast ignition [77] and the generation of warm dense matter [78]. Experiments have shown that the real source size of a few hundred micrometres could be collimated down to 30 μm using ballistic focusing from the rear surface of hemispherical targets. However, one has to take into account the real sheath geometry of the driving electrons to interpret the proton beam profile accurately. The driving sheath consists of a superposition of the sheath field normal to the inner surface of the hemisphere and the Gaussian electron density distribution caused by the limited source size of the driving laser focus. Therefore, the experiments in Ref. [72] have indicated that a larger laser focal spot should minimize the second effect and thus result in a better focus quality.

For almost all applications, a precise control of the beam parameter and the possibility of tailoring the beams is of great importance. As we have seen, using the guidance by the shape of the target material, we have large control over the spatial ion beam distribution. However, it might also be instructive and preferable to try to manipulate the ion beam just using optical methods. Results reported so far were obtained with a round laser spot, focused as well as possible to obtain the highest intensities. But, as

found by Fuchs *et al.* [79], the laser focal spot shape is eventually imprinted in the accelerated proton beam. Fuchs *et al.* assumed that the bulk of the hot electrons follow the laser focal spot topology and create a sheath with the same topology at the rear side. The proton beam spatial profile, as detected by a film detector, was simulated with a simple electrostatic model. Fuchs *et al.* took the laser beam profile as an input parameter and assumed the electron transport to be homogeneous, with a characteristic opening angle that needed to be fit to match the measured data. The unknown source size of the protons was fit to best match the experimental results. It was shown, that for their specific target thickness and laser parameters, the fitted broadening angle of the electron sheath at the rear side closely matched the broadening angle expected by multiple Coulomb small-angle scattering. However, Fuchs *et al.* could only fit the most intense part of the measured beam and neglected the lower intense part that originates from rear-side accelerated protons. Additionally, there is no information on the dependence of these findings on target thickness.

Using microgrooved targets, a more detailed understanding has been achieved. The asymmetric laser beam results in asymmetric proton beam profiles. The energy-resolved source size of the protons was deduced by imaging the beam perturbations from the microgrooved surface into a radiochromic film detector stack. It was shown that the protons with the highest energies were emitted from the smallest source. When the laser focus size was increased, the proton source size increased as well. For symmetric as well as asymmetric laser beam profiles, the source-size dependent energy distribution in both cases could be fit to a Gaussian. This leads to the conclusion that the laser beam profile has no significant contribution to the general expansion characteristics of laser-accelerated protons, but can strongly modify the transverse beam profile without changing the angle of the beam spread. For a more detailed analysis of the experimental results, a code for sheath-accelerated beam ray-tracing for ion analysis (SABRINA) was developed, which takes the laser beam parameters as input and calculates the shape of the proton distribution in the detector. The electron transport was modelled to follow the laser beam profile topology closely and a broadening due to small-angle collisions was assumed. It was shown that broadening due to small-angle collisions is the major effect that describes the source size of protons for thick target foils (50 μm). By contrast, thin target foils (13 μm) show much larger sources than expected for small-angle collisions. The physical reason behind this observation remains unclear; it is most likely the result of electron refluxing. Thus, the shape of the sheath at the rear side of the thick targets can be estimated by a simple model of broadening due to multiple small-angle scattering, but the model fails for the description of sheath broadening in thin targets.

The imprint of the laser beam profile affects the intense part of the proton beam profile. This effect must also be present in cases with a round focal spot. Therefore, a focal spot with a sharp peaked laser beam profile will result in a strongly divergent proton beam, as observed in experiments. The findings also explain that in cases where a collimation of the proton beam is required, e.g., proton fast ignition or the injection of the beam into a post-accelerator, not only is a curved target surface necessary, but a large, flat-top laser focal spot is indispensable, to produce a flat proton-accelerating sheath.

4 Applications

Summarizing the beam parameters achievable using the TNSA mechanism one can make a number of conclusions.

The measured particle energies so far extend up to tens of megaelectronvolts (78 MeV protons, 5 MeV/u palladium) and the particles showed complete space charge and current neutralization due to accompanying electrons. In experiments, particle numbers of more than 10^{13} ions per pulse and beam currents in the mega-amp regime were observed. Another outstanding feature is the excellent beam quality, with a transverse emittance of less than 0.004π mm · mrad and a longitudinal emittance of less than 10^{-4} eV · s. Because of these unmatched beam characteristics, a wealth of applications were foreseen immediately. Those applications range from:

– injection of high power ion beams for large-scale basic research facilities;

– new diagnostic techniques for short-pulse phenomena, since the short pulse duration allows for the imaging of transient phenomena;

– modification of material parameters (from applications in material science to warm dense matter research and laboratory astrophysics);

– applications in energy research ('fast igniter' in the inertial fusion energy context);

– medical applications;

– use as a new laser-driven pathway to compact, bright neutron sources.

4.1 Ion source

An important application for TNSA ion beams is for use as next-generation ion sources or accelerators, where the excellent beam quality and strong field gradients can replace more conventional systems. Several collaborations are actively working on that task, spearheaded by the LIBRA collaboration in the UK, the LIGHT collaboration in Germany and a group at JAEA in Japan.

We briefly address a few aspects of current research aimed at applying laser ion beams as a new source.

4.1.1 Collimation and bunch rotation of accelerated protons

One of the main drawbacks of laser-accelerated ions and, in particular, protons are the exponential energy spectrum and the large envelope divergence of the beam. Different techniques have been developed to modify the energy distribution, to produce a more monoenergetic beam. Therefore, special targets were created with thin proton or carbon layers on the rear side, as well as deuterated droplets. In addition, there have been attempts to reduce the initial divergence of the beams by ballistic focusing, with the help of curved hemispherical targets, resulting in a beam focus at a distance of the diameter of the sphere from the laser focus. In a different experiment, a laser-triggered microlens was used to select a small energy interval and to focus the protons with these specific energies to a millimetre spot 70 cm from the target [80].

The total proton yield of typically 10^{13} particles and the extremely high observed phase space density immediately behind the source and prior to any collimator are highly encouraging. As in all cases of sources of secondary particles (antiprotons, muons, rare isotopes, etc.), transmission efficiency and phase space degradation due to the first collimator need to be carefully examined. In particular, higher than first-order focusing properties of the collimator are a serious limitation to the realistically 'usable' fraction of the production energy spectrum, as well as of the production cone divergence. As these same limitations might cause a serious degradation of the transverse emittance of the 'usable' protons, the very small production emittance becomes a relatively irrelevant quantity. Instead, an 'effective' emittance, taking into account transmission loss and blow-up caused by the collimator, should be used. In this context, space charge (non-linear) effects are a further source of emittance degradation—probably not the dominant one—to be carefully examined. Recently, we have shown that the collimation of a laser-accelerated proton beam by a pulsed high-field solenoid is possible and leads to good results in terms of collimation efficiencies. More than 10^{12} particles were caught and transported by the solenoid. The steadiness of the proton beam after collimation could be proven up to a distance of 324 mm from the target position. Inside the solenoid, strong space charge effects occurred, owing to the co-moving electrons, which are forced to circulate around the solenoid's axis at their gyroradius by the strong magnetic field, leading to a proton beam aggregation around the axis (see Fig. 11). Details can be found in Ref. [81].

More detailed calculations of the injection into ion optical structures have been published by Ingo Hofmann [82, 83].
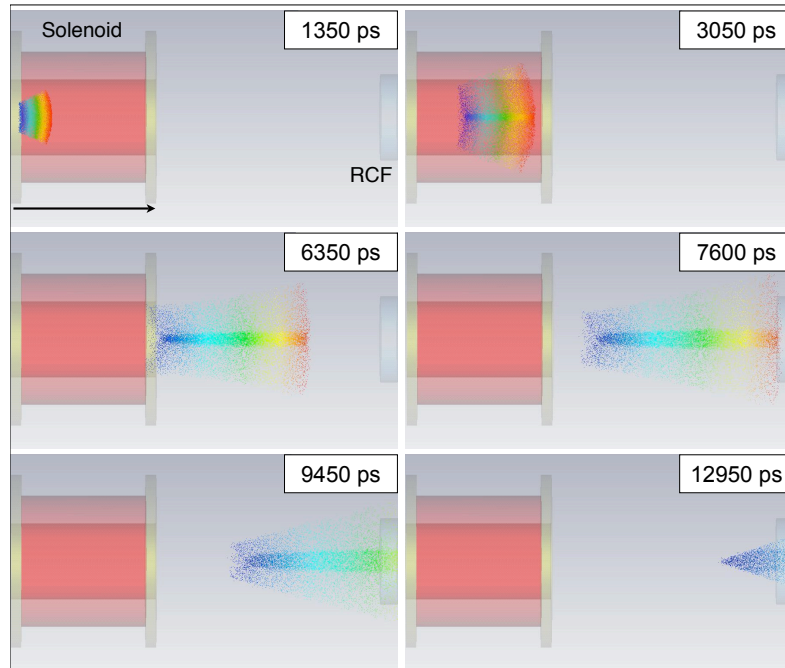
Fig. 11: Simulation of the propagation of target normal sheath acceleration protons through a solenoid magnet up to a detector at six timepoints. For a better view, the accompanying electrons are not shown. A clear aggregation of protons at the axis due to space charge effects is visible. The proton energy in this case ranges from 1 to 5 MeV. RCF, radiochromic film. Published in Ref. [81].

4.1.2 *Chromatic error of solenoid collimation*

In general, pulsed solenoids are a good match to the 'round' production cone of laser-accelerated particles; a quadrupole-based focusing system appears to be disadvantageous in the defocusing plane of the first lens, owing to the relatively large production angles. As an example, we use the short-pulsed solenoid currently under experimental study at the GSI Helmholtz Centre for Heavy-Ion Research. It has a length of 72 mm and a theoretical maximum field strength of 16 T, which is sufficient to parallelize protons at 10 MeV. The distance from the target spot to the solenoid edge is assumed to be 17 mm.

The prevailing higher-order effect of a solenoid is the increase in focal length with particle energy. Owing to the debunching process, different sections along the bunch have different energies and thus focus at different distances. This results in an effective increase in the bunch-averaged emittance to the effect that the tiny initial production emittance should be replaced by a chromatic emittance. To examine the expected behaviour in detailed simulation, we employed the DYNAMION code [84], which includes higher-order effects in amplitudes and energy dependence, as well as space charge effects. The latter are based on particle–particle interaction, which limits the space charge resolution. The solenoid three-dimensional magnetic field was obtained by direct integration using the coil geometry of the experimental solenoid. To quantify the chromatic effect, we consider an ensemble of protons with constant energy spread $\Delta E/E = \pm 0.04$ around a reference energy of 10 MeV. Results for final emittances (ignoring space charge) are shown in Fig. 12 to depend exactly quadratically on the considered production cone opening angle $\delta x'$, which varied up to ± 172 mrad ($\pm 10°$). To test the influence of space charge, we also simulated a case with the number of protons in the bunch, N_b, equal to 3×10^9, which is equivalent to a linac current of $I = 50$ mA (using $I = eN_b f_{RF}$ and assuming that each bucket of a $f_{RF} = 108$ MHz sequence is filled identically). For simplicity, the bunch intensity was chosen to be independent of the opening angle. It is noted that the quadratic law still approximately applies. Since for given $\delta x'$ the dependence on $\Delta E/E$ is found to be practically linear, we can justify the following scaling of the
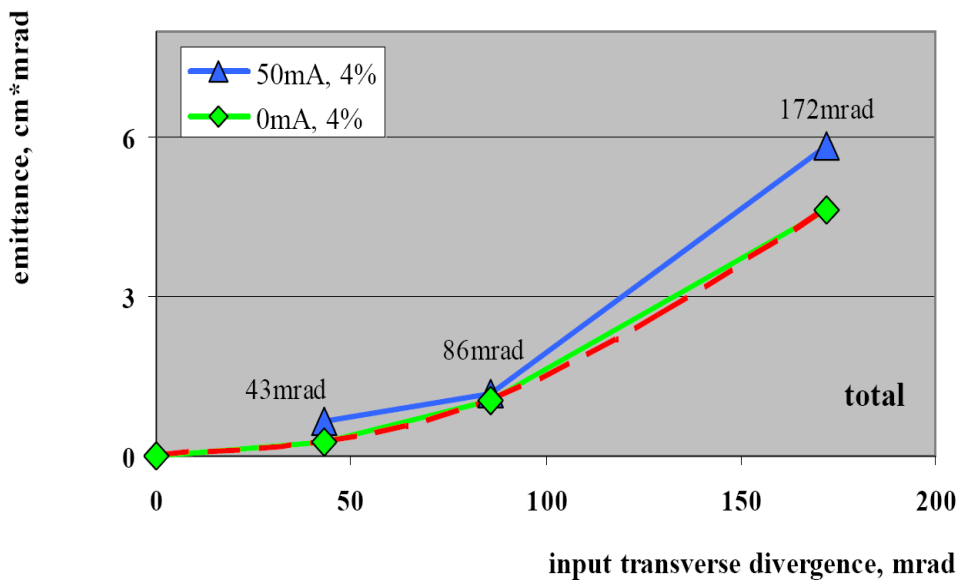
Fig. 12: Dependence of 'chromatic' emittances (here total emittances for 95% of particles) on production cone angle, as obtained by DYNAMION simulations.

chromatic emittances in the absence of space charge:

$$\varepsilon_x = \alpha_c (\delta x')^2 \frac{\Delta E}{E}, \tag{44}$$

with $\alpha_c \approx 0.04$ m/rad for the particular solenoid described here. The law is still approximately true if space charge is included for the assumed bunch intensity. Note that the chromatic emittance is found to be practically independent of the initial spot radius r_{spot}—contrary to the production emittance given by the product $r_{spot}\delta x'$.

4.1.3 Transmission through beam pipe

For planned experiments, it is important to note that the increase of emittance with energy spread will inevitably lead to transmission loss in the finite acceptance of the following beam pipe. To this end, we have assumed a beam pipe of 3 cm radius up to 250 cm distance from the source. We have also assumed a linac current $N_b \approx \Delta E/E$, with $N_b = 2 \times 10^9$ for the lowest value $\Delta E/E = \pm 0.04$. The increase in transmission loss with distance is mostly due to the large spread of focusing angles as a function of the energy spread and is also due, to a lesser extent, to space charge. The surviving energy distribution evaluated at different distances from the source goes down to 35% for the largest initial energy spread case in the previous example of $\Delta E/E = \pm 0.64$ and a correspondingly high current of 560 mA. Obviously, an extended beam pipe serves as an energy filter.

4.1.4 RF bunch rotation

For most applications of laser-accelerated particles, in particular for ion beam therapy, it is desirable to reduce the final energy spread on the target to a fraction of one per cent, to enable focusing on a small target spot. This is achieved by means of a 'bunch rotation' RF cavity applied to the beam after debunching to a length suitable for the RF wavelength. The initial short bunch length increases with debunching proportional to the distance from the source and the considered energy spread. Capture

into the RF bucket of a fraction of a beam within a given transverse emittance defines the ultimate six-dimensional extraction efficiency and the 'usable' part of the total production of protons. As reference value we take an energy spread of $\Delta E/E = \pm 0.04$, which can be reduced to a reasonably small value by using a 500 kV/108 MHz bunch rotation RF cavity approximately 250 cm away from the solenoid. This means that only the central part of the totally transmitted energy distribution—about 20% for the 3 cm aperture limitation—can be captured by the RF bucket. Diagnosing the intensity and six-dimensional emittance of this 'usable' fraction in the presence of the background of the fully transmitted spectrum is a challenge to the diagnostics.

Thus, at the current status, a careful study of the transfer efficiency of these beams into conventional transport and focusing structures is crucial and timely, and will be carried out within the next few years, given the unique prerequisites present among all the international collaborations. The foremost goal of the proposed effort is to determine the properties of the generated proton or ion beams with the prospect of later applications and to examine the possibilities of collimation, transport, debunching and, possibly, post-acceleration in conventional accelerator structures, both theoretically and experimentally.

4.2 Diagnostics

A highly energetic laminar beam of charge particles with a pulse duration of only a few picoseconds constitutes an ideal tool to diagnose transient phenomena. Like a short burst of X-rays, a pulse of laser-driven protons can penetrate a target and reveal important information about its parameters. Laser-driven protons are complementary to X-rays, as the interactions of charged particles are fundamentally different from those of electromagnetic radiation. In the past, ion beams produced by conventional accelerators have already been used to radiograph static and transient samples [85], as well as for the investigation of electric fields in laser-produced plasmas [86]. Several experiments with laser-accelerated beams as probes were performed, to investigate the evolution of highly transient electric fields evolving from charging of laser-irradiated targets [88,92]. These fields change on a picosecond time-scale. Proton beams from ultra-intense laser–matter interactions are accelerated in a few picoseconds, depending on the laser pulse length. Combined with the very low emittance [67], these beams allow for two-dimensional mapping of the primary target with unprecedented spatial and temporal resolution. Using this technique, remnants of relativistic solitons that were generated in a laser plasma have been detected [89], and the accelerating Debye sheath in a TNSA process as well as the ion expansion from the rear side of the target foil could be pictured [90].

Because of the energy spectrum, and owing to the dispersion of the proton pulse at the point of interaction with the target to be probed, different proton energies probe the target at different times. Using the radiochromic film stack technique, the ions can be separated in energy, where the high-energy particles deposit their energy in the rearmost layer, while lower-energy particles are stopped in the front layers. Thus, by unfolding the layers, one can separate the ion energies and therefore the interaction time down to picosecond temporal resolution.

4.2.1 Energy loss

The fundamental contrast mechanism for generating image information is energy loss in the sample, and the consequent shift in the energy distribution of transmitted protons. As one proceeds from the shallow layers to the deepest radiochromic film layers, protons with progressively higher incident energies are preferentially recorded. By examining a portion of the image where the sample contained no intervening material, we can deduce the incident proton energy distribution from the depth dependence of the deposited energy, based on the respective response function.

If the incident protons pass through a thin sample of thickness δx, they lose energy according to their energy loss rate, and the transmitted proton energy, which is incident on the film detector, is

$$E_f \approx E_i - (dE/dx) \cdot \delta x \approx E_i - \Delta E \; . \tag{45}$$

If the sample is thick, so that dE/dx is not constant over the energy range ΔE, then the energy loss is given by the integral along the trajectory,

$$\Delta E = \int_0^{\delta x} dE(dx)dx \; . \tag{46}$$

A limitation of this technique is that energy loss, and therefore thickness information, is encoded as a spectral shift of the proton energy distribution due to energy loss. If the object has a large thickness range, and hence large values of the energy loss, early-time (high incident proton energy) information from thick portions of the sample is recorded in the same final proton energy interval as late-time (low incident proton energy) information from thinner portions of the sample. Deconvolution of the spatial, temporal, and thickness information is complicated, and even self-consistent solutions may not be unique.

In the ideal limit of no transverse scattering, the resolvable thickness variation over a sample is related only to the energy loss and the exponential fall-off of the proton spectrum. This is a strong function of initial proton energy via the energy dependence of dE/dx. For example, for our test case, in which we observe a 2 MeV exponential distribution, a resolvable intensity variation of 1% implies a resolvable energy loss of 20 keV. At a proton energy of 10 MeV, this would correspond to a plastic (CH) thickness of $<5\,\mu m$, and, at 4 MeV, a thickness of $2\,\mu m$.

4.2.2 *Transverse scattering*

In addition to continuously slowing down, the protons also undergo multiple small-angle Coulomb scattering from the nuclei in the sample. In the energy range of interest, proton multiple scattering can be described by a Gaussian fit to the Molière distribution, which is similar to the case of the electrons we have seen earlier. That is, protons are scattered according to a near-Gaussian polar angle distribution,

$$dN/d\Omega \approx \frac{1}{\sqrt{2\pi}\Theta_0} \exp(-\Theta^2/2\Theta_0^2) \; , \tag{47}$$

where Θ_0 is given by

$$\Theta_0 = \frac{13.6\,\mathrm{MeV}}{\beta pc} \sqrt{x/X_0}[1 + 0.038\ln(x/X_0)] \; , \tag{48}$$

with X_0 defined as

$$X_0 = \frac{716.4\,\mathrm{g\,cm^{-2}\,A}}{Z/(Z+1)\ln(287/\sqrt{Z})} \; . \tag{49}$$

Multiple scattering of the protons as they traverse the sample degrades the spatial resolution possible from ideal energy loss imaging, but it can also increase the contrast and hence the thickness resolution. This is because those protons scattered away from the direct line of sight from the source to the film detector are moved from the umbral to penumbral region on the film, thus reducing the flux of protons in the direct shadow. The decrease in proton flux associated with a given film plane being sensitive to higher initial energy protons, owing to the energy loss in the sample, is augmented by a flux reduction from scattering. The very small-angle scattered protons, however, increase the net flux of protons in the penumbral region, which can lead to 'limb brightening' effects, which are usual for image techniques based on scattering (rather than absorption).

4.2.3   *Field deflection*

Probably the most important applications to date of proton probing are related to the unique capability of this technique to detect electrostatic fields in plasmas [91,92]. This has allowed researchers to obtain, for the first time, direct information on electric fields arising through a number of laser–plasma interaction processes [88, 89, 93]. The high temporal resolution is here fundamental in allowing the detection of highly transient fields following short-pulse interaction. When the protons cross a region with a non-zero electric field, they are deflected by the transverse component $E_\perp$ of the field. The proton transverse deflection at the proton detector plane is equal to

$$\Delta r_\perp \approx eL \int_0^b (E_\perp / m_{\mathrm{p}} v_{\mathrm{p}}^2)\mathrm{d}l \; , \tag{50}$$

where $m_{\mathrm{p}} v_{\mathrm{p}}^2/2$ is the proton kinetic energy, e its charge, b the distance over which the field is present, and L the distance from the object to the detector. As a consequence of the deflections, the proton beam cross-section profile undergoes variations showing local modulations in the proton density. Assuming the proton density modulation to be small $\delta n/n_0 \ll 1$, where n_0 and δn are, respectively, the unperturbed proton density and proton density modulation at the detector plane, we obtain $\delta n/n_0 \approx -\mathrm{div}(\Delta r_\perp)/M$, where M is the geometrical magnification. The value of the electric field amplitude and the spatial scale can then be determined if a given functional dependence of E can be inferred a priori, e.g., from theoretical or geometrical considerations. More details can be found in the references cited in this paper and in Ref. [6], from where a part of this section was taken.

4.3   Warm dense matter

The creation of extreme states of matter is important for the understanding of the physics covered in various research fields, such as high-pressure physics, applied material studies, planetary science, inertial fusion energy, and all forms of plasma generation from solids. The primary difficulties in the study of these states of matter are that the time-scales or the changes are rapid (≈ 1 ps), while the matter is very dense and the temperatures are relatively low, $\sim\mathrm{eV}/k_{\mathrm{B}}$. With these parameters, the plasma exhibits long- and short-range orders, which are due to the correlating effects of the ions and electrons. The state of matter is too dense or too cold (or both) to admit standard solutions used in plasma physics. Perturbative approaches using expansions in small parameters for the description of the plasma are no longer valid, proving a tremendous challenge for theoretical models. This region where condensed matter physics and plasma physics converge is the so-called warm dense matter regime [94].

Warm dense matter conditions can be generated in a number of ways, such as by laser-generated shocks [95] or laser-generated X-rays [96, 97], intense ion beams from conventional accelerators [98], or laser-accelerated protons [72–74], to name just a few. Whereas lasers only interact with the surface of a sample, ions can penetrate deeply into the material of interest, thereby generating large samples of homogeneously heated matter. The short pulse duration of intense ion beams furthermore allows for the investigation of the equation of states close to the solid-state density, because of the material's inertia, which prevents the expansion of the sample within the interaction. Moreover, the interaction of ions with matter is dominantly due to collisions and does not include a high-temperature plasma corona as is the case in laser–matter interaction.

The generation of large, homogeneous samples of warm dense matter is accompanied by the challenging task of diagnosing this state of matter, as the usual diagnostic techniques fail under these conditions. The material density results in a huge opacity and the relatively low temperature does not allow traditional spectroscopic methods to be applied. Moreover, the sample size, deposited energy and life-time of the matter state are strongly interrelated and are dominated by the stagnation time of the atoms in the probe. Thus, high spatial and temporal resolution is required to gain quantitative data in those experiments. Owing to the high density of the sample, laser diagnostics cannot be used. The properties of matter could be determined by measuring the expansion after the heating [73] or by measuring the
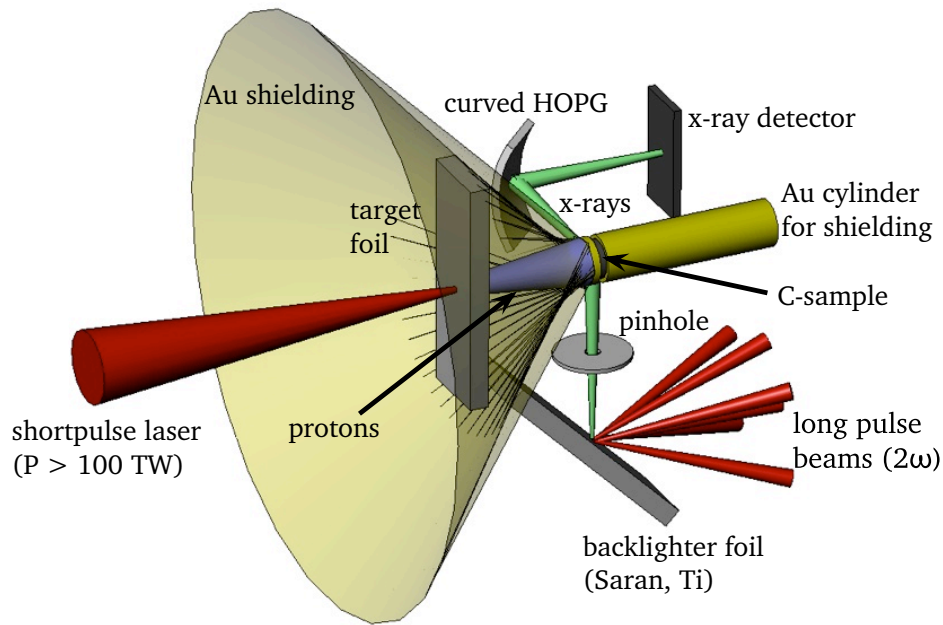
Fig. 13: Experimental scheme to investigate the properties of laser-accelerated proton-heated matter by spectrally resolved X-ray Thomson scattering. HOPG, highly ordered pyrolytic graphite.

thermal radiation emitted by the sample [72]. However, even more interesting are the plasma parameters deep inside the sample, where the ion heating is most effective. An ideally suited diagnostic technique recently developed is X-ray Thomson scattering [96, 99, 100].

Figure 13 shows a typical experimental scheme to investigate the transformation of solid, low-Z material into the warm dense matter state. The experimental scheme requires a high-energy short-pulse laser and one or more long-pulse laser beams in the same experimental vacuum chamber. In recent years, more and more laser facilities have upgraded their laser systems for such types of pump-probe experiment. A chirped pulse amplification laser beam above 100 TW power generates an intense proton beam from a thin target foil. The protons hit a solid-density sample and heat it isochorically, up to a temperature of several eV/k_B. The long-pulse beams are used to drive an intense X-ray source from a titanium or Saran (which contains chlorine) foil. The sample is probed by narrow-band line-radiation from the chlorine or titanium plasma. The scattered radiation is first spectrally dispersed by a highly efficient, highly oriented, pyrolytic graphite crystal spectrometer with von Hamos geometry before it is detected. Extensive gold shielding (partly shown in Fig. 13) is required, to prevent parasitic signals in the scatter spectrometer. From the measured Doppler-broadened, Compton-downshifted signal, the temperature and density can be inferred.

Whereas lasers only interact with the surface of a sample, ions can penetrate deeply into the material of interest, thereby generating large samples of homogeneously heated matter. The short pulse duration of laser-produced ion beams furthermore allows for the investigation of equation of states close to the solid-state density, because the material's inertia prevents expansion of the sample during the interaction. In addition to these unique characteristics, the interaction of ions with matter is dominantly due to collisions and does not include a high-temperature plasma corona, as it is present in laser–matter interaction. The absence of a large radiation background is of importance to the experiment. Large conversion efficiencies have been observed, and significant energy can be transferred from the ultra-intense laser via the ion beam into the sample of interest. Because of the high beam quality, ballistic focusing has been demonstrated, allowing for an increase in local energy deposition, thus leading to
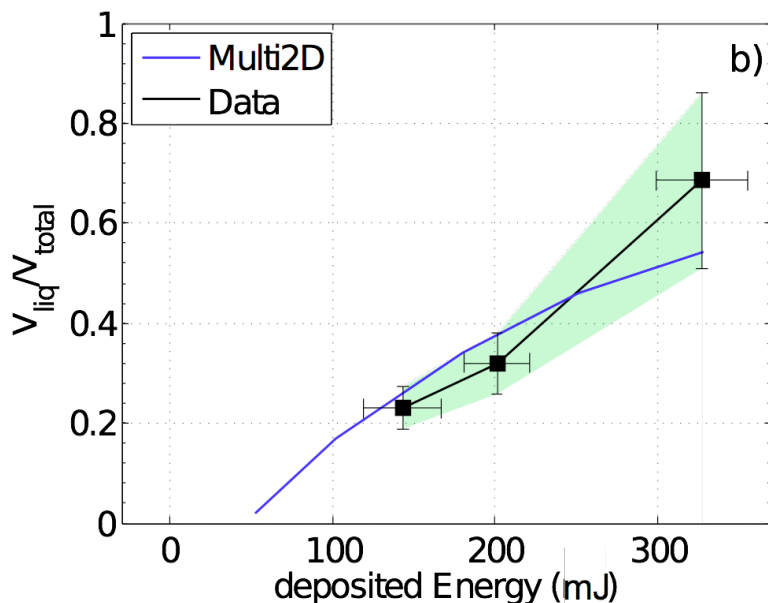
Fig. 14: Molten fraction of a carbon sample versus deposited energy by the proton beam, calculated from the experimental data and simulated with Multi2D [78].

higher temperatures. The use of hemispherical targets, including cone-guided targets to enhance the local proton flux on the material of interest, can even enhance the locally deposited energy.

Using laser pulses in excess of 100 J, the intense proton beams can heat large targets up to several times the melting temperature. In a milestone experiment at TAW last year, the molten fraction in carbon samples heated by intense proton beams was measured [78]. Figure 14 shows some results compared with a radiation hydrodynamics simulation, which uses SESAME as the equation-of-state model. It can be seen that agreement is better at lower temperatures, where ionization is not important, but at higher temperatures, the presence of an ionic component may be important.

To summarize, laser-accelerated proton beams are very well suited to produce macroscopic samples of warm dense matter. Their unique feature, having pulse durations of only a few picoseconds while containing more than 10^{12} protons, cannot be matched by conventional ion accelerators.

4.4 Fast ignition

In conventional inertial fusion, ignition and propagating burn occurs when a sufficient temperature (5–10 keV) is reached within a sufficient mass of deuterium-tritium fuel characterized by a density-radius product greater than the range of an alpha particle, $(\rho r)_\alpha > 0.3$ g cm^{-2}. The necessary conditions for propagating deuterium-tritium burn are achieved by an appropriate balance between the energy gain mechanisms and the energy loss mechanisms. Mechanical work (PdV), alpha particle energy deposition and, to a smaller extent, neutron energy deposition are the principal energy gain mechanisms in deuterium-tritium fuel. Electron heat conduction and radiation are the principal energy loss mechanisms. When the rate of energy gain exceeds the rate of energy loss for a sufficient period of time, ignition occurs. Because of the short burn time and the inertia of the fuel, the contribution of expansion losses is negligible. Fast ignition [101, 102] was proposed as a means to increase the gain, reduce the driver energy and relax the symmetry requirements for compression, primarily in direct-drive inertial confinement fusion. The concept is to precompress the cold fuel and subsequently to ignite it with a separate short-pulse high-intensity laser or particle (electron or ion) pulse. Fast ignition is being extensively studied by many groups worldwide, using short-pulse lasers or temporally compressed heavy-ion

beams. There are several technical challenges for the success of laser-driven fast ignition. Absorption of the igniter pulse generates copious relativistic electrons, but it is not yet known whether these electrons will propagate as a stable beam into the compressed fuel to deposit their energy in a small volume. In principle, heavy-ion beams can have advantages for fast ignition. A focused ion beam may maintain an almost straight trajectory while traversing the coronal plasma and compressed target, and ions have an excellent coupling efficiency to the fuel and deliver their energy in a well-defined volume, owing to the higher energy deposition at the end of their range (Bragg peak) [103]. With the discovery of TNSA ions with excellent beam quality, the idea of using those beams for fast ignition was introduced [77,104]. Protons do have several advantages compared with other ion species [105] and electrons. First, because of their highest charge-to-mass ratio, they are accelerated most efficiently up to the highest energies. They can penetrate a target more deeply to reach the high-density region, where the hot spot is to be formed, because of the quadratic dependence of the stopping power on the charge state. And finally, they do, like all ions, exhibit a characteristic maximum of the energy deposition at the end of their range, which is desirable, to heat a localized volume efficiently.

The basic idea is to use a number of short-pulse lasers to irradiate a thin foil. The protons were accelerated from the rear surface of the foils and, because of the parabolic geometry, are focused into the compressed fuel. One of the requirements for proton fast ignition is the possibility of focusing the proton beams into a small volume. It has recently been demonstrated that proton beam focusing is indeed possible, and spot sizes of about 40 μm have been achieved. This is comparable to what is required by proton fast ignition. Larger irradiated areas on the target front surface, as required for proton fast ignition, would flatten the electron distribution at the rear surface. Not only might this result in a single divergence angle for different energies, but it would also form a much smaller initial divergence angle that could be compensated, in order to reach the required focal spot diameters.

The pulse length at the source is of the right order of magnitude for proton fast ignition, which has already been indicated in first experiments on ion acceleration. The protons are not monochromatic but rather have an exponential energy distribution. This seemed to be a concern at the beginning because of dispersion and pulse lengthening. A close distance to the pellet, however, raises concerns as to whether the thin metallic foil, which is to be the source of the protons, can be kept cold enough that it does not develop a density gradient at the rear surface, which would diminish the accelerating field. A second concern was related to the stopping power. Because of the difference in initial velocity, the energy deposition of protons with different kinetic energies is spread over a larger volume. Slower protons are stopped at a shorter distance and do not contribute to the creation of the igniter spark. Fortunately, further work has relieved those concerns. Simulations by Basko *et al.* (presented at the Fast Ignition Workshop 2002, Tampa, Florida) have shown that the protective shield placed in front of the source can withstand the X-ray flux of the pellet compression and keep the rear surface of the source foil cold enough for acceleration via TNSA. A thickness of a few tens of micrometres, however, provides thermal shielding as well as sufficient mechanical stability. The distance between the source and the ignition spot can be a few millimetres. If this distance is too short for the compression geometry (e.g., if a closely coupled hohlraum is not used), the distance can be adjusted using a similar cone target, as for conventional fast ignition. As for the concerns on the hydrodynamic stability of the proton source foil, the proposed usage of a cone target similar to the one proposed for electron fast ignition [106, 107] has solved most of the problems, by shielding the foil from primary soft X-rays generated in the compression of the capsule (see Fig. 15). Furthermore it was demonstrated that small-scale plasma density gradients at the rear side of the proton source target caused by target preheating have no significant (less than 10%) impact on the TNSA mechanism [45].

A big surprise was the fact that a monochromatic proton beam is actually not the optimum to heat a hot spot in a fusion target. Numerical simulations have shown that one has to take into account the decrease of the stopping power of the nuclear fuel with increasing plasma temperature. Thus, an exponential energy spectrum, like the one generated by this mechanism, is the most favourable. The first

protons with the highest energies penetrate the fuel deeply. By the time the proton number increases and the target temperature rises, the stopping power is reduced, thereby compensating for the lower initial energy of the incoming protons. Thus the majority of the protons deposit their energy within the same volume.

Existing short-pulse lasers have already demonstrated intensities that are sufficient for generating the proton energy spectra required for proton fast ignition. Regardless of the nature of the igniter beam, calculations show a minimum deposited energy required for fast ignition of the order of a few tens of kilojoules. There have been many experiments with different laser systems accelerating proton beams. Interestingly, the laser to ion beam conversion efficiency seems to increase strongly with total laser energy from a fraction of a per cent up to more than 10%. Carefully extrapolating the conversion efficiencies to multikilojoule laser systems, conversion efficiencies of more than 10% can be expected, which would result in the need for a few hundred kilojoules of short-pulse laser energy for proton fast ignition.

The most detailed theoretical analysis so far has been published by Temporal *et al.* [102] for a proton beam with an exponential energy spectrum. Following their assumptions, a total proton energy of about 26 kJ at an effective temperature of 3 MeV is required. This moderate temperature was found to be an optimum between the need for high temperatures to minimize the pulse lengthening caused by the velocity spread and the stopping range for the major part of the spectrum. It is interesting to note that the protons, which effectively heat the hot spot, contain only 10 kJ of the total energy and range from 19 to 10 MeV. If it could be possible to shape the energy spectrum of the laser-accelerated protons, this would strongly influence the required laser beam energy. The total number of protons needed for ignition is close to 10^{16}. Is it conceivable to achieve a consistent scenario for those requirements? A typical proton beam temperature of 3 MeV is commonly obtained in experiments at 5×10^{19} W cm^{-2}. Assuming a pulse length of 4 ps (which would increase the damage threshold of modern dielectric compressor gratings) and a conversion efficiency of 10%, a total laser energy of 260 kJ would be needed.

The use of a cone-guided geometry, as in conventional fast ignition, has been considered to be of great advantage. The source foil can be shielded from the radiation caused by the primary drivers, the source-to-hot spot distance can be tailored precisely, and the pellet can be protected from heat during the injection into the target chamber. A recent experimental campaign to study the influence of the cone walls on the propagation and the transport of TNSA protons has shown that, despite the influence of self-generated electric fields in the cone walls by the recirculating electrons, good focusing may be achievable.

After the initial introduction of laser-accelerated proton beams for fast ignition, theoretical studies have not only quantified the required beam parameters [102, 103], but also recently introduced sophisticated scenarios that have greatly relaxed those parameters. A recent study proposed a combination of two spatially shaped proton beam pulses with a total beam energy that would match laser systems that might be available in the not too distant future [108, 109]. The most recent scenario looks for a ring-shaped proton beam, to impact into the dense fuel and further compress the hot spot, with a subsequent pulse of protons in the centre, to ignite the double-compressed core. This would cause an energy that is further reduced by a factor of two compared with the model described previously. Such laser proton beams have been generated using advanced cone geometries [110].

4.5 Medical applications

Soon after the discovery of TNSA ion beams, the prospects for medical applications have been the focus of research. Besides the possibility of transmuting short-lived isotopes for positron emission tomography, the main interest was in the use as a driver for hadron therapy. Hadron therapy [111–115] is a radiotherapy technique that uses protons, neutrons, or carbon ions to irradiate cancer tumours. The use of protons and carbon ions in radiotherapy has several advantages over the more widely used X-radiotherapy. First, the proton beam scattering on the atomic electrons is weak and thus there is less irradiation of healthy tissues in the vicinity of the tumour. Second, the slowing down length for a proton
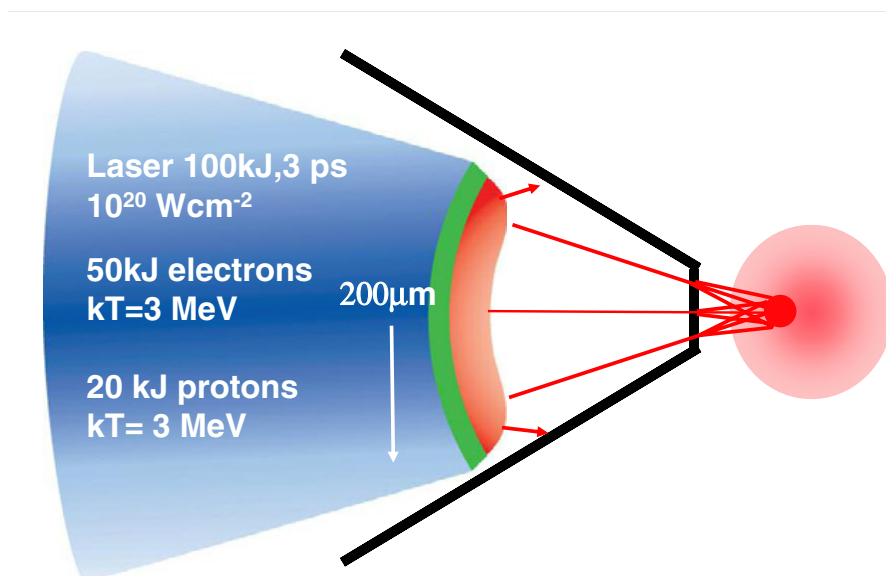
Fig. 15: Proposed concept of using cone-guided proton fast ignition. Courtesy of M. Key (LLNL)

of given energy is fixed, which avoids undesirable irradiation of healthy tissues at the rear side of the tumour. Third, the well localized maximum of proton energy losses in matter (the Bragg peak) leads to a substantial increase of the irradiation dose in the vicinity of the proton stopping point.

The proton energy window of therapeutical interest ranges between 60 and 250 MeV, depending on the location of the tumour. Proton beams with the required parameters are currently obtained using conventional charged-particle accelerators, such as synchrotrons, cyclotrons, and linacs [116]. The use of laser-based accelerators has been proposed as an alternative [117–121], which could lead to advantages in terms of device compactness and cost.

A laser accelerator could be used simply as a high-efficiency ion injector for the proton accelerator, or could replace a conventional proton accelerator altogether. Because of the broad energy and angular spectra of the protons, an energy selection and beam collimation system will be needed [122, 123]. Typically, $\Delta E/E \approx 10^{-2}$ is required for optimal dose delivery over the tumour region. All-optical systems have also been proposed, in which the ion beam acceleration takes place in the treatment room itself and ion beam transport and delivery issues are minimized. In this case, the beam energy spread and divergence would need to be minimized by controlling the beam and target parameters. The required energies for deep-seated tumours (>200 MeV) are still in the future, but appear to be within reach, considering ongoing developments in the field. A demanding requirement to be satisfied is also the system duty factor, i.e., the fraction of time during which the proton beam is available for use, which must not be less than 0.3.

With the recent experimental results of ion beams in the range up to 80 MeV, the lower threshold for medical applications has been achieved. However, for deep-seated tumours it is uncertain whether the TNSA mechanism is still the best option or whether new mechanisms should be explored that not only lead to higher particle energies, but also offer a much smaller energy dispersion to begin with.

4.6   Neutron source

Since the first experiments with ultra-intense lasers, nuclear reactions have been observed and also used to diagnose the hot core part of the laser–plasma interaction [124]. In addition to the generation and detection of radio-isotopes and transmuted nuclei, neutrons are released either as a result of intense

bremsstrahlung or by electron or ion impact. Because of the large conversion efficiency of laser to ion beam energy and the large cross-section for subsequent proton neutron reactions, laser-driven neutron sources based on the TNSA mechanism have become a focus of modern nuclear research.

One has to distinguish between the different neutron generation mechanisms. At proton beam particle energies in the megaelectronvolt range, the interaction and neutron generation relies on the excitation of giant resonances that result in single (p, n) or multiple (p, ×n) neutron emission. The cross-section can be quite large and is energy dependent, peaking at characteristic proton impact energies. In the case of two particles in the exit channel, the neutron spectrum is monoenergetic for a given projectile energy and neutron emission angle. However, since the angle and energy spread of laser-emitted particles is large, only strongly exothermal reactions ($Q \gg E_{\text{proj}}$) will yield roughly monoenergetic neutrons. Which process takes place in a particular instance depends on the combination of target, projectile, and momentum transfer. The cross-sections for these processes are in the range of 100 mb up to 1 b and are therefore quite large.

As the driving ion beam is ultra-short and the release mechanism is prompt, the neutron pulses are very short and originate from a very small region maximizing the net flux on secondary samples. Such a probe exists in the form of fusion neutrons. They are generated by the d(d,n)^{3}He fusion reaction in deuterated targets, and their use as a laser-plasma diagnostic is not fundamentally new. When neutrons are produced from laser-accelerated ions in the bulk of an irradiated $(CD_2)_n$ target, they are emitted within a few picoseconds from a volume of the order of $(10 \, \mu m)^3$. During the neutron pulse, at a distance of several millimetres from the target, fast neutron fluxes of $10^{19}/(\text{cm}^2 \cdot \text{s})$ can be achieved, which is four orders of magnitude higher than current continuous research reactors can deliver.

In the past, the neutron emission caused by (γ, n) and (p, n) reactions from the target have been measured at moderate laser intensities. A typical detector set-up is a silver activation detector attached to a photomultiplier tube. On typical shots, the neutrons are generated by (γ, n) reactions within the target (caused by the bremsstrahlung photons from the relativistic electrons) and by (p, n) reactions of our proton beam impacting on the radiochromic film screen or a dedicated secondary production target. This can be, e.g., a target of deuterated plastic (CD_2), which was irradiated by a beam of TNSA-accelerated deuterons. One can observe the yield of neutrons from (d, d) fusion reactions.

To optimize laser-driven neutron sources, one can perform simulation studies using consolidated findings on particle beam characteristics obtained from laser experiments [125]. The optimization will be according to the absolute neutron yield and angle, as well as the spectral distribution. For neutron generation, we attempt a two-stage target design in which the TNSA ion beam irradiates a secondary sample. The advantage of this design is that we can optimize the proton or deuteron generation using different targets in the first stage. According to the beam properties obtained from the first stage, it will be possible to optimize the neutron production via proton- and deuteron-induced neutron disintegration reactions, respectively, in the second stage. The neutron production target design (second stage) allows adaptation to the desired application. In earlier experimental campaigns at the PHELIX laser facility at the GSI Helmholtz Centre for Heavy-Ion Research, $>10^9$ neutrons per shot from proton-induced reactions in copper have been produced. The integrated number of protons was 10^{12} to 10^{13}. Each neutron yield in these experiments already exceeded that from the accelerator-driven neutron source FRANZ [126] in Frankfurt (Germany) by five orders of magnitude. With the help of the GEANT4 code [127], we can simulate the expected neutron yield for TNSA protons using experimental input spectra. We have simulated the neutron spectra and the production efficiencies using several isotopes within the second-stage target. The thickness of these different targets was 10 mm. The highest production efficiencies were obtained from proton-induced reactions at lithium, beryllium, boron, and vanadium in their natural abundance. As a benchmark, one can compare the simulation results for copper with experimental data, where one finds a good agreement.

The neutron spectra from proton-induced reactions in beryllium and lithium show high particle numbers in the lower energy range and in the range around several megaelectronvolts, respectively. This
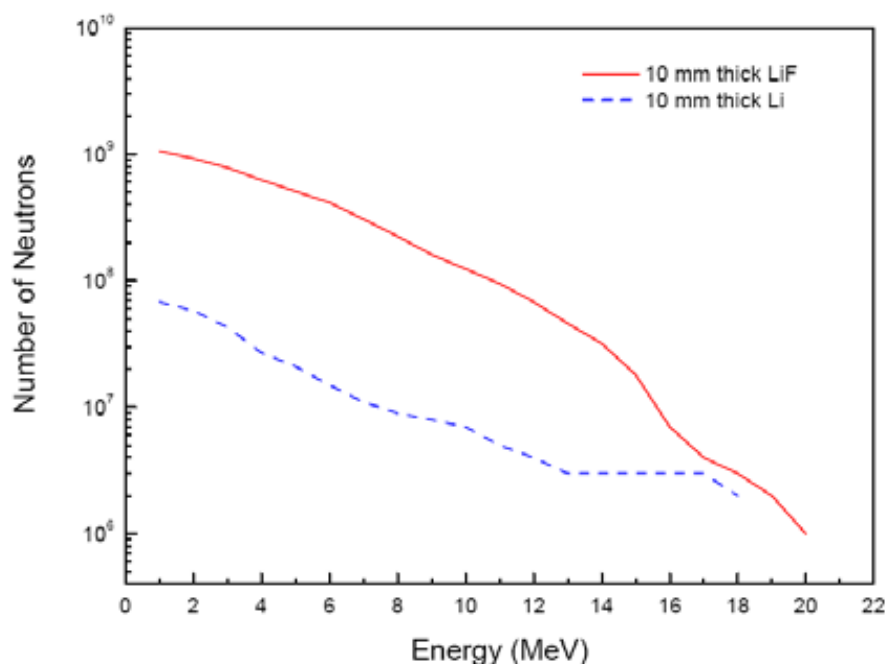
Fig. 16: Simulated neutron spectra for different converter materials

is of interest in transmutation studies and nuclear material science. Figure 16 demonstrates the difference in simulated neutron spectra using lithium in the natural abundance and lithium fluoride. The properties of the initial proton spectrum that was used in the simulation were obtained from experimental results at the PHELIX facility. The initial particle number was 10^{13} and the maximum proton energy was 25 MeV.

In addition, the neutron yield using lithium fluoride is much higher than the neutron yield from proton-induced reactions in lithium. The explanation is that the mass density of lithium fluoride is higher, owing to the interatomic compounds. Lithium fluoride has a mass density of 2.64 g/cm^3 and the mass density of lithium fluoride is only 0.53 g/cm^3. This demonstrates the attraction in using composite targets in future studies of laser-driven neutron sources. In future developments of the optimization of laser-driven neutron sources, we will use more sophisticated composite target designs for adaptation to the desired applications.

References

[1] S.J. Gitomer *et al.*, *Phys. Fluids* **29**(8) (1986) 2679. http://dx.doi.org/10.1063/1.865510

[2] E.L. Clark *et al.*, *Phys. Rev. Lett.* **84**(4) (2000) 670. http://dx.doi.org/10.1103/PhysRevLett.84.670

[3] R.A. Snavely *et al.*, *Phys. Rev. Lett.* **85**(14) (2000) 2945. http://dx.doi.org/10.1103/PhysRevLett.85.2945

[4] A. Maksimchuk *et al.*, *Phys. Rev. Lett.* **84**(18) (2000) 4108. http://dx.doi.org/10.1103/PhysRevLett.84.4108

[5] M. Schollmeier, Ph.D. thesis, Technische Universität Darmstadt, 2008.

[6] M. Borghesi *et al.*, *Fusion Sci. Technol.* **49** (2006) 412.

[7] M. Roth, *J. Instrum* (2011), in press.

[8] M. Roth *et al.*, *Phys. Rev. ST-AB* **5** (2002) 061301.

[9] B.G. Logan *et al.*, *Fusion Sci. Technol.* **49** (2006) 399.

[10] M. Roth, *Plasma Phys. Contr. Fus.* **51**(1) (2009) 014004. http://dx.doi.org/10.1088/0741-3335/51/1/014004

[11] S.C. Wilks *et al.*, *Phys. Rev. Lett.* **69**(9) (1992) 1383.
http://dx.doi.org/10.1103/PhysRevLett.69.1383

[12] S.C. Wilks, *Phys. Fluids B* **5**(7) (1993) 2603. http://dx.doi.org/10.1063/1.860697

[13] B.F. Lasinski *et al.*, *Phys. Plasmas* **6**(5) (1999) 2041. http://dx.doi.org/10.1063/1.873496

[14] S.P. Hatchett *et al.*, *Phys. Plasmas* **7**(5) (2000) 2076. http://dx.doi.org/10.1063/1.874030

[15] P.A. Norreys *et al.*, *Phys. Plasmas* **6**(5) (1999) 2150. http://dx.doi.org/10.1063/1.873466

[16] J. Fuchs *et al.*, *Nature Phys.* **2** (2006) 48. http://dx.doi.org/10.1038/nphys199

[17] Y. Ping *et al.*, *Phys. Rev. Lett.* **100**(8) (2008) 085004.
http://dx.doi.org/10.1103/PhysRevLett.100.085004

[18] M.H. Key *et al.*, *Phys. Plasmas* **5**(5) (1998) 1966. http://dx.doi.org/10.1063/1.872867

[19] J.R. Davies, *Phys. Rev. E* **65**(2) (2002) 026407. http://dx.doi.org/10.1103/PhysRevE.65.026407

[20] H. Alfvén, *Phys. Rev.* **55**(5) (1939) 425. http://dx.doi.org/10.1103/PhysRev.55.425

[21] J.R. Davies, *Phys. Rev. E* **69**(6) (2004) 065401(R).
http://dx.doi.org/10.1103/PhysRevE.69.065402

[22] D. Batani *et al.*, *Phys. Rev. E* **65**(6) (2002) 066409.
http://dx.doi.org/10.1103/PhysRevE.65.066409

[23] A.R. Bell *et al.*, *Plasma Phys. Contr. Fusion* **48**(3) (2006) 37.
http://dx.doi.org/10.1088/0741-3335/48/3/R01

[24] F. Pisani *et al.*, *Phys. Rev. E* **62**(5) (2000) R5927(R).
http://dx.doi.org/10.1103/PhysRevE.62.R5927

[25] J. Honrubia *et al.*, *Laser Part. Beams* **24** (2006) 217.

[26] J.J. Santos *et al.*, *Phys. Plasmas* **14**(10) (2007) 103107. http://dx.doi.org/10.1063/1.2790893

[27] R.B. Stephens *et al.*, *Phys. Rev. E* **69**(6) (2004) 066414.
http://dx.doi.org/10.1103/PhysRevE.69.066414

[28] K.L. Lancaster *et al.*, *Phys. Rev. Lett.* **98**(12) (2007) 125002.
http://dx.doi.org/10.1103/PhysRevLett.98.125002

[29] Y. Sentoku *et al.*, *Phys. Plasma* **14**(12) (2007) 122701. http://dx.doi.org/10.1063/1.2816439

[30] A. Pukhov, *Phys. Rev. Lett.* **86**(16) (2001) 3562. http://dx.doi.org/10.1103/PhysRevLett.86.3562

[31] J.J. Santos *et al.*, *Phys. Rev. Lett.* **89**(2) (2002) 025001.
http://dx.doi.org/10.1103/PhysRevLett.89.025001

[32] H. Nishimura *et al.*, *Plasma Phys. Contr. Fusion* **47**(12B) (2005) 823.
http://dx.doi.org/10.1088/0741-3335/47/12B/S64

[33] A.J. MacKinnon *et al.*, *Phys. Rev. Lett.* **88**(21) (2002) 215006.
http://dx.doi.org/10.1103/PhysRevLett.88.215006

[34] M. Schollmeier *et al.*, *Phys. Plasmas* **15**(5) (2008) 053101. http://dx.doi.org/10.1063/1.2912451

[35] H.A. Bethe, *Phys. Rev.* **89**(6) (1953) 1256. http://dx.doi.org/10.1103/PhysRev.89.1256

[36] K.U. Akli *et al.*, *Phys. Plasmas* **14**(2) (2007) 023102. http://dx.doi.org/10.1063/1.2431632

[37] J.E. Crow *et al.*, *J. Plasma Phys.* **14**(1) (1975) 65. http://dx.doi.org/10.1017/S0022377800025538

[38] N.A. Krall and A.W. Trivelpiece, *Principles of Plasma Physics* (San Francisco Press, San
Francisco, 1986).

[39] S. Augst *et al.*, *Phys. Rev. Lett.* **63**(20) (1989) 2212.
http://dx.doi.org/10.1103/PhysRevLett.63.2212

[40] M. Hegelich *et al.*, *Phys. Rev. Lett* **89**(8) (2002) 085002.
http://dx.doi.org/10.1103/PhysRevLett.89.085002

[41] P. McKenna *et al.*, *Phys. Rev. Lett.* **98**(14) (2007) 145001.

http://dx.doi.org/10.1103/PhysRevLett.98.145001

[42] J. Schreiber *et al.*, *Appl. Phys. B* **79**(8) (2004) 1041 http://dx.doi.org/10.1007/s00340-004-1665-5

[43] http://www.mathworks.com/

[44] J. Schreiber *et al.*, *Phys. Rev.Lett.* **97**(4) (2006) 045005.
http://dx.doi.org/10.1103/PhysRevLett.97.045005

[45] J. Fuchs *et al.*, *Phys. Plasmas* **14**(5) (2007) 053105. http://dx.doi.org/10.1063/1.2720373

[46] P. Mora, *Phys. Rev. Lett.* **90**(18) (2003) 185002. http://dx.doi.org/10.1103/PhysRevLett.90.185002

[47] P. Mora, *Phys. Plasmas* **12**(11) (2005) 112102. http://dx.doi.org/10.1063/1.2134768

[48] P. Mora, *Phys. Rev. E* **72**(5) (2005) 056401. http://dx.doi.org/10.1103/PhysRevE.72.056401

[49] B.J. Albright *et al.*, *Phys. Rev. Lett.* **97**(11) (2006) 115002.
http://dx.doi.org/10.1103/PhysRevLett.97.115002

[50] M. Passoni and M. Lontano, *Laser Part. Beams* **22**(2) (2004) 163.
http://dx.doi.org/10.1017/S026303460422211X

[51] L.D. Landau and E.M. Lifshitz, *Mekhanika sploshnykh sred* (Gostekhizdat, Moscow, 1954).

[52] S. Eliezer, *The Interaction of High Power Lasers with Plasma* (IOP Publishing, Bristol, 2002).

[53] M. Kaluza *et al.*, *Phys. Rev. Lett.* **93**(4) (2004) 045003.
http://dx.doi.org/10.1103/PhysRevLett.93.045003

[54] L. Robson *et al.*, *Nature Phys.* **3** (2007) 58. http://dx.doi.org/10.1038/nphys476

[55] E. Brambrink, Ph.D. thesis, Technische Universität Darmstadt, 2004.

[56] A.V. Gurevich *et al.*, *Sov. Phys. JETP* **22**, 449 (1966)

[57] M. Passoni *et al.*, *Phys. Rev. E* **69**(2) (2004) 026411.
http://dx.doi.org/10.1103/PhysRevE.69.026411

[58] N. Kumar and A. Pukhov, *Phys. Plasmas* **15**(5) (2008) 053103.
http://dx.doi.org/10.1063/1.2913611

[59] T. Grismayer and P. Mora, *Phys. Plasmas* **13**(3) (2006) 032103.
http://dx.doi.org/10.1063/1.2178653

[60] T. Grismayer *et al.*, *Phys. Rev. E* **77**(6) (2008) 066407.
http://dx.doi.org/10.1103/PhysRevE.77.066407

[61] V.Y. Bychenkov *et al.*, *Phys Plasmas* **11**(6) (2004) 3242. http://dx.doi.org/10.1063/1.1738649

[62] M. Allen *et al.*, *Phys. Rev. Lett.* **93**(26) (2004) 265004.
http://dx.doi.org/10.1103/PhysRevLett.93.265004

[63] B.M. Hegelich *et al.*, *Nature* **439** (2006) 441. http://dx.doi.org/10.1038/nature04400

[64] H. Schwoerer *et al.*, *Nature* **439** (2006) 445. http://dx.doi.org/10.1038/nature04492

[65] S. Ter-Avetisyan *et al.*, *Phys. Rev. Lett.* **96**(14) (2006) 145006.
http://dx.doi.org/10.1103/PhysRevLett.96.145006

[66] D.C. Carroll *et al.*, *Phys. Rev. E* **76**(6) (2007) 065401(R).
http://dx.doi.org/10.1103/PhysRevE.76.065401

[67] T.E. Cowan *et al.*, *Phys. Rev. Lett.* **92**(20) (2004) 204801.
http://dx.doi.org/10.1103/PhysRevLett.92.204801

[68] S. Humphries, *Charged Particle Beams* (John Wiley, New York, 2008).
http://www.fieldp.com/cpb.html

[69] F. Nürnberg *et al.*, *Rev. Sci. Instr.* **80**(3) (2009) 033301. http://dx.doi.org/10.1063/1.3086424

[70] A.P.L. Robinson *et al.*, *New J. Phys.* **10** (2008) 013021.
http://dx.doi.org/10.1088/1367-2630/10/1/013021

[71] L. Yin *et al.*, *Phys. Plasmas* **14**(5) (2007) 056706. http://dx.doi.org/10.1063/1.2436857

[72] P.K. Patel *et al.*, *Phys. Rev. Lett.* **91**(12) (2003) 125004. http://dx.doi.org/10.1103/PhysRevLett.91.125004

[73] G.M. Dyer *et al.*, *Phys. Rev. Lett.* **101**(1) (2008) 015002. http://dx.doi.org/10.1103/PhysRevLett.101.015002

[74] P. Antici *et al.*, *J. Phys. IV France* **133** (2006) 1077. http://dx.doi.org/10.1051/jp4:2006133218

[75] R.A. Snavely *et al.*, *Phys. Plasmas* **14**(9) (2007) 092703. http://dx.doi.org/10.1063/1.2774001

[76] E. Brambrink *et al.*, *Phys. Rev. Lett.* **96**(15) (2006) 154801. http://dx.doi.org/10.1103/PhysRevLett.96.154801

[77] M. Roth *et al.*, *Phys. Rev. Lett.* **86**(3) (2001) 436. http://dx.doi.org/10.1103/PhysRevLett.86.436

[78] A. Pelka *et al.*, *Phys. Rev. Lett.* **105**(26) (2010) 265701. http://dx.doi.org/10.1103/PhysRevLett.105.265701

[79] J. Fuchs *et al.*, *Phys. Rev. Lett.* **91**(25) (2003) 255002. http://dx.doi.org/10.1103/PhysRevLett.91.255002

[80] T. Toncian *et al.*, *Science* **312**(5772) (2006) 410. http://dx.doi.org/10.1126/science.1124412

[81] K. Harres *et al.*, *Phys. Plasmas* **17**(2) (2010) 023107. http://dx.doi.org/10.1063/1.3299391

[82] I. Hofman *et al.*, Laser accelerated ions and their potential for therapy accelerators, Proc. HIAT09, Venice, 2009.

[83] I. Hofmann *et al.*, *Phys. Rev. STAB*, **14** (2011) 031304.

[84] S. Yaramishev *et al.*, *Nucl. Instrum. Meth. A* **588** (2006) 1.

[85] D. West and A.C. Sherwood, *Nature* **239** (1972) 157. http://dx.doi.org/10.1038/239157b0

[86] J. Cobble *et al.*, *J. Appl. Phys.* **92**(4) (2002) 1775. http://dx.doi.org/10.1063/1.1494128

[87] M. Borghesi *et al.*, *Appl. Phys. Lett.* **82**(10) (2003) 1529. http://dx.doi.org/10.1063/1.1560554

[88] A. Mackinnon *et al.*, *Rev. Sci. Instrum.* **75**(10) (2004) 3531. http://dx.doi.org/10.1063/1.1788893

[89] M. Borghesi *et al.*, *Phys. Rev. Lett.* **88**(13) (2002) 135002. http://dx.doi.org/10.1103/PhysRevLett.88.135002

[90] M. Borghesi *et al.*, *Laser Part. Beams* **25**(1) (2007) 161. http://dx.doi.org/10.1017/S0263034607070218

[91] M. Borghesi *et al.*, *Phys Plasmas* **9**(5) (2002) 2214. http://dx.doi.org/10.1063/1.1459457

[92] M. Borghesi *et al.*, *Rev. Sci. Instrum.* **74**(3) (2003) 1688. http://dx.doi.org/10.1063/1.1534390

[93] M. Borghesi *et al.*, *Appl. Phys. Lett.* **82**(10) (2003) 1529. http://dx.doi.org/10.1063/1.1560554

[94] R.W. Lee *et al.*, *J. Opt. Am. B* **20**(4) (2003) 770. http://dx.doi.org/10.1364/JOSAB.20.000770

[95] E. Garcia Saiz *et al.*, *Nature Phys.* **4** (2008) 940 093306. http://dx.doi.org/10.1038/nphys1103

[96] S.H. Glenzer *et al.*, *Phys. Rev. Lett.* **90**(17) (2003) 175002. http://dx.doi.org/10.1103/PhysRevLett.90.175002

[97] S.H. Glenzer *et al.*, *Phys. Rev. Lett.* **98**(6) (2007) 065002. http://dx.doi.org/10.1103/PhysRevLett.98.065002

[98] N.A. Tahir *et al.*, *Physica Scripta*, **T132** (2008) 014023. http://dx.doi.org/10.1088/0031-8949/2008/T132/014023

[99] O.L. Landen *et al.*, *J. Quant. Spectrosc. Radiat. Transfer* **71**(2-6) (2001) 465. http://dx.doi.org/10.1016/S0022-4073(01)00090-5

[100] M. Schollmeier *et al.*, *Laser Part. Beams* **24**(3) (2006) 335. http://dx.doi.org/10.1017/S026303460606099X

[101] M. Tabak *et al.*, *Phys. Plasmas* **1**(5) (1994) 1626. http://dx.doi.org/10.1063/1.870664

[102] M. Temporal *et al.*, *Phys Plasmas* **9**(7) (2002) 3098. http://dx.doi.org/10.1063/1.1482375

[103] A. Caruso and V.A. Pais, *Nucl. Fusion* **36**(6) (1996) 745.

http://dx.doi.org/10.1088/0029-5515/36/6/I06

[104] H. Ruhl *et al.*, *Plasma Phys. Rep.* **27**(5) (2001) 363. http://dx.doi.org/10.1134/1.1371596

[105] V.Y. Bychenkov, *Plasma Phys. Rep.* **27**(12) (2001) 1017. http://dx.doi.org/10.1134/1.1426135

[106] P.A. Norreys *et al.*, *Phys. Plasmas* **7**(9) (2000) 3721. http://dx.doi.org/10.1063/1.1287419

[107] R. Kodama *et al.*, *Nature* **412** (2001) 798. http://dx.doi.org/10.1038/35090525

[108] M. Temporal, *Phys. Plasmas* **13**(12) (2006) 122704. http://dx.doi.org/10.1063/1.2400592

[109] M. Temporal *et al.*, *Phys. Plasmas* **15**(5) (2008) 052702. http://dx.doi.org/10.1063/1.2918316

[110] O. Deppert *et al.*, to be published.

[111] W.T. Chu *et al.*, *Rev. Sci. Instrum.* **64**(8) (1993) 2055. http://dx.doi.org/10.1063/1.1143946

[112] V.S. Khoroshkov and K.K. Onosovsky, *Instrum. Exp. Tech.* **38** (1995) 149.

[113] G. Kraft, *Nucl. Instrum. Methods Phys. Res. A* **454**(1) (2000) 1.
       http://dx.doi.org/10.1016/S0168-9002(00)00802-0

[114] R.R. Wilson, *Radiology* **47**(5) (1946) 487. http://dx.doi.org/10.1148/47.5.487

[115] U. Amaldi, *Nucl. Phys. A* **654**(1-2) (1999) C375.
       http://dx.doi.org/10.1016/S0375-9474(99)00264-X

[116] W. Wieszczycka and W.H. Scharf, *Proton Radiotherapy Accelerators*, (World Scientific, River
       Edge, NJ, 2001), p. 323.

[117] S.V. Bulanov *et al.*, *Phys. Lett. A* **299**(2-3) (2002) 240.
       http://dx.doi.org/10.1016/S0375-9601(02)00521-2

[118] S.V. Bulanov *et al.*, *Plasma Phys. Rep.* **28**(5) (2002) 453. http://dx.doi.org/10.1134/1.1478534

[119] E. Fourkal and C. Ma, *Med. Phys.* **30**(7) (2003) 1660. http://dx.doi.org/10.1118/1.1586268

[120] E. Fourkal *et al.*, *Med. Phys.* **31** (2004) 1883.

[121] V. Malke *et al.*, *Med. Phys.* **31**(6) (2004) 1587. http://dx.doi.org/10.1118/1.1747751

[122] E. Fourkal *et al.*, *Med. Phys.* **29**(12) (2002) 2788. http://dx.doi.org/10.1118/1.1521122

[123] E. Fourkal *et al.*, *Med. Phys.* **30**(7) (2003) 1660. http://dx.doi.org/10.1118/1.1586268

[124] M. Guenther *et al.*, *Phys. Plasmas* **18**(8) (2011) 083102. http://dx.doi.org/10.1063/1.3613923

[125] M. Guenther *et al.*, Proc. 2011 ICENES Conference, San Francisco, 2011.

[126] U. Ratzinger *et al.*, Proc. IPAC'10, Kyoto, Japan, 2010.

[127] S. Agostinelli *et al.*, *Nucl. Instrum. Methods A* **506**(3) (2003) 250.
       http://dx.doi.org/10.1016/S0168-9002(03)01368-8

AWAKE, A Particle-driven Plasma Wakefield Acceleration Experiment

E. Gschwendtner
CERN, Geneva, Switzerland

Abstract
The Advanced Proton Driven Plasma Wakefield Acceleration Experiment (AWAKE) aims at studying plasma wakefield generation and electron acceleration driven by proton bunches. It is a proof-of-principle R&D experiment at CERN and the world's first proton driven plasma wakefield acceleration experiment. The AWAKE experiment will be installed in the former CNGS facility and uses the 400 GeV/c proton beam bunches from the SPS. The first experiments will focus on the self-modulation instability of the long (r.m.s ~12 cm) proton bunch in the plasma. These experiments are planned for the end of 2016. Later, in 2017/2018, low energy (~15 MeV) electrons will be externally injected to sample the wakefields and be accelerated beyond 1 GeV.

Keywords
AWAKE; proton-driven plasma acceleration; plasma wakefield acceleration; self-modulation instability.

1 Introduction

The construction of ever larger and costlier accelerator facilities has its limits; accelerating fields of today's RF cavities and microwave technology are limited to about 100 MV/m; hence, several tens of kilometres are required for future linear colliders. New technologies will be needed to push back the energy frontier.

Harnessing wakefields, physicists may be able to produce accelerator gradients hundreds of times higher than those achieved in current radiofrequency cavities. This would allow future colliders to achieve higher energies over shorter distances than is possible today.

With plasma wakefield acceleration, a new energy frontier for particle physics might be realized in an affordable manner; plasma machines could be the vaunted 'table-top' accelerator. Scales of metres rather than kilometres could bring accelerator labs within the reach of any university or industrial laboratory.

For the next generation wakefield accelerator, one could envision so-called 'afterburners', in which existing conventional accelerators, such as the LHC, are used to excite the plasma waves.

Before this new technology can be applied, to fully exploit the high gradients for future accelerators, several key performance parameters need to be benchmarked.

Experimental results have demonstrated the success of the plasma wakefield acceleration and its research: e.g., beam-driven plasma wakefield acceleration experiments [1] performed at Stanford Linear Accelerator Center (SLAC) successfully doubled the energies of some of the electrons in an initial 42 GeV beam in less than 1 m of plasma. Recent results report high-efficiency acceleration of a discrete trailing bunch of electrons containing sufficient charge to extract a substantial amount of energy from the high-gradient, non-linear plasma wakefield accelerator; the core particles gained about 1.6 GeV of energy per particle, with a final energy spread as low as 0.7% and an energy-transfer efficiency from the wake to the bunch that could exceed 30% [2].

2   Beam-driven plasma wakefield acceleration experiments: landscape

Table 1 gives a summary of beam-driven plasma wakefield experiments that are either ongoing or under construction.

2.1   AWAKE

AWAKE is the first plasma wakefield acceleration experiment worldwide to use a proton beam as a drive beam [3]. The first physics is expected late 2016. AWAKE is a proof-of-principle experiment with the aim of providing a design for a particle physics frontier accelerator within the next decade. AWAKE will be explained in detail in the following sections.

2.2   FACET, Stanford Linear Accelerator Center

FACET (Facility for Advanced Accelerator Experimental Tests) is a user facility at the SLAC International Laboratory, providing high energy density electron and positron beams with peak currents of $\approx$20 kA focused down to 30×30 μm^2 transverse spot size at an energy of 20 GeV [4]. The facility studies the acceleration of high-quality and high-efficiency witness bunch beams, as well as the acceleration of positrons. FACET started in 2012 and hosts more than 150 users and 25 experiments. FACET II has been proposed, to start in 2018.

2.3   DESY, Zeuthen

PITZ is the Photo-Injector Test Facility at DESY, Zeuthen. This is a research and development facility [5]; its 20 MeV electron beam is used to study the self-modulation instability (SMI) in a lithium plasma cell.

2.4   DESY FLASH Forward

FLASH Forward [6] aims at the advancement of beam-driven plasma wakefield physics towards applications utilizing the expertise and tools provided by the Free Electron Laser facility at DESY. Its goals are the demonstration of capture and controlled release of externally shaped electron beams, the exploration of novel in-plasma beam-generation techniques, and the assessment of those beams for free electron laser gain.

2.5   Brookhaven ATF

At Brookhaven National Laboratory [7], experiments are performed to study the quasi-non-linear plasma wakefield acceleration (PWA) regime, driven by multiple bunches and visualization with optical techniques.

2.6   SPARC LAB, Frascati

SPARC LAB is a multi-purpose user facility, which includes experiments on laser- and beam-driven plasma wakefield acceleration experiments.

Table 1: Beam-driven plasma wakefield experiments

Facility	Location	Drive beam	Witness beam	Start	End	Goal
AWAKE	CERN, Geneva, Switzerland	400 GeV protons	Externally injected electron beam (PHIN 15 MeV)	2016	2020, or later	Use for future high energy e^-/e^+ collider Study self-modulation instability Accelerate externally injected electrons Demonstrate scalability of acceleration scheme.
SLAC-FACET	SLAC, Stanford, USA	20 GeV electrons and positrons	Two-bunch formed with mask (e^-/e^+ and e^-e^+ bunches)	2012	Sept 2016	Acceleration of high-quality high-efficiency witness bunch Acceleration of positrons FACET II proposal for 2018 operation
DESY, Zeuthen	PITZ, DESY, Zeuthen, Germany	20 MeV electron beam	No witness beam, only drive beam from RF gun	2015	≈2017	Study self-modulation instability
DESY, FLASH Forward	DESY, Hamburg, Germany	X-ray free electron laser type electron beam 1 GeV	Drive and witness in free electron laser bunch, or independent witness bunch (laser wakefield acceleration)	2016	2020+	Application (mostly) for X-ray free electron laser Energy-doubling of FLASH-beam energy Upgrade-stage: use 2 GeV free electron laser drive beam
Brookhaven ATF	Brookhaven National Laboratory, USA	60 MeV electrons	Several bunches, drive and witness formed with mask	Ongoing		Study quasi-non-linear plasma wakefield acceleration regime Study plasma wakefield acceleration driven by multiple bunches Visualization with optical techniques
SPARC LAB	Frascati, Italy	150 MeV electrons	several bunches	Ongoing		multi-purpose user facility: includes laser- and beam-driven plasmas wakefield experiments.

3 AWAKE: components for a particle-driven plasma wakefield acceleration experiment

3.1 Introduction

AWAKE, the advanced proton-driven plasma wakefield acceleration experiment will be installed at CERN in the former CNGS (CERN Neutrinos to Gran Sasso) area and is currently under construction [8]. The first beam for physics experiments is expected by the end of 2016.

AWAKE will use proton bunches, for the first time ever, to drive plasma wakefields [9]. The main physics goals of the experiment are as follows.

— To study the physics of self-modulation of long proton bunches in plasma as a function of beam and plasma parameters. This includes radial modulation and seeding of the instability.

— To probe the longitudinal (accelerating) wakefields with externally injected electrons. This includes measuring their energy spectra for different injection and plasma parameters.

— To study injection dynamics and the production of multigigaelectronvolt electron bunches, either from side injection or from on-axis injection (with two plasma cells). This will include using a plasma density step to maintain the wakefields at the GV/m level over metre distances.

— To develop long, scalable, and uniform plasma cells and develop schemes for the production and acceleration of short bunches of protons for future experiments and accelerators.

In the baseline design of AWAKE at CERN, an LHC-type proton bunch of 400 GeV/c (with an intensity of 3×10^{11} protons/bunch) will be extracted from the CERN Super Proton Synchrotron (SPS) and sent along the 750 m long proton beam line towards a plasma cell, which is installed at the upstream end of the CNGS area (see Fig. 1). The proton beam will be focused to $\sigma_{x,y} = 200$ μm near the entrance of the 10 m long rubidium vapour plasma cell with an adjustable density in the 10^{14} to 10^{15} electrons/cm^3 range.

When the proton bunch, with an r.m.s. bunch length of $\sigma_z = 12$ cm (0.4 ns), enters the plasma cell, it undergoes the SMI, i.e., the development a long bunch of protons into a series of micro-bunches that resonantly drive large wakefields. The effective length and period of the modulated beam is set by the plasma wavelength (for AWAKE, typically $\lambda_{pe} = 1$ mm).

A high power ($\approx$4.5 TW) laser pulse, co-propagating and co-axial with the proton beam, will be used to ionize the neutral gas in the plasma cell and also to generate the seed of the proton bunch self-modulation.
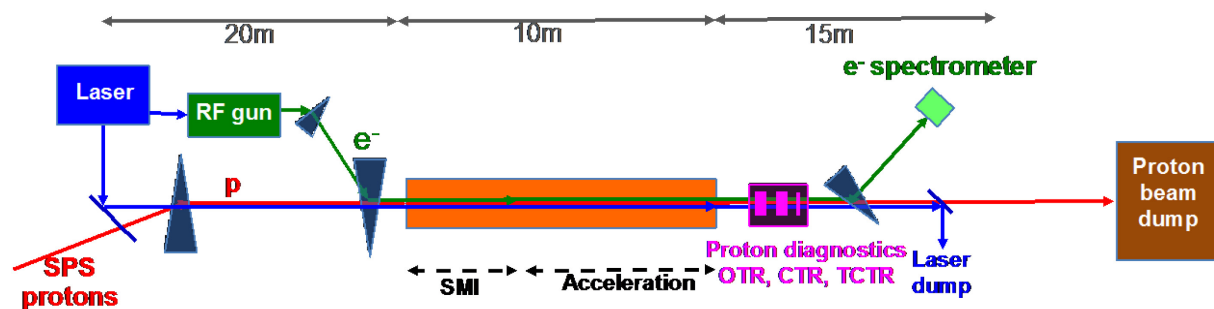
Fig. 1: Baseline design of AWAKE

An electron beam of 1.2×10^9 electrons, which will be injected at 10−20 MeV/c into the plasma cell, serves as a witness beam and will be accelerated in the wake of the proton bunch.

Several diagnostic tools will be installed downstream of the plasma cell to measure the proton bunch self-modulation effects and the accelerated electron bunch properties.
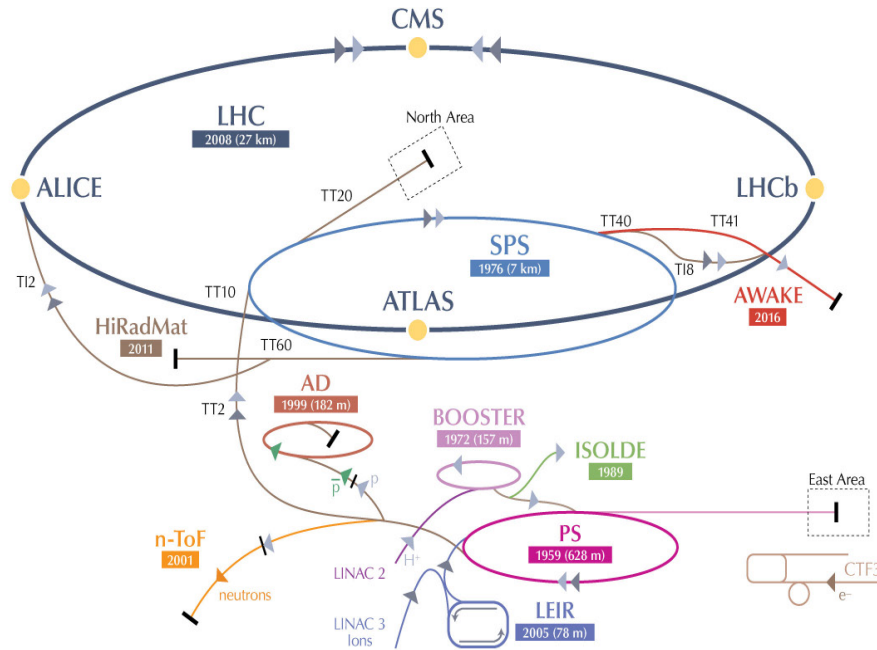
Figure 1 shows the baseline design of the AWAKE experiment. The baseline parameters of the experiment are summarized in Table 2.

In the AWAKE master schedule, the experiment to obtain evidence for the SMI (see Section 3.3) corresponds to Phase 1, and is expected to start by the end of 2016. In Phase 2, AWAKE aims at the first demonstration of proton-driven plasma wakefield acceleration of an electron witness beam; this programme is planned to start by the end of 2017.

Table 2: Baseline parameters of the AWAKE experiment

Parameter	Value	Parameter	Value
Proton beam		**Laser beam to plasma cell**	
Momentum	400 GeV/c	Laser type	Fibre titanium:sapphire
Protons/bunch	3×10^{11}	Pulse wavelength	$L_0 = 780$ nm
Bunch extraction frequency	0.03 Hz (ultimate: 0.14 Hz)	Pulse length	100−120 fs
Bunch length	$\sigma_z = 0.4$ ns (12 cm)	Pulse energy	450 mJ
Bunch size at plasma entrance	$\sigma_{x,y} = 200$ µm	Laser power	4.5 TW
Normal emittance (r.m.s.)	3.5 mm mrad	Focused laser size	$\sigma_{x,y} = 1$ mm
Relative energy spread	$\Delta p/p = 0.45\%$	Energy stability	$\pm 1.5\%$ r.m.s.
Beta function	$\beta^*_x = \beta^*_y = 4.9$ m	Repetition rate	10 Hz
Dispersion	$D^*_x = D^*_y = 0$		
Electron beam		**Plasma source**	
Momentum	16 MeV/c	Plasma type	Laser ionized rubidium vapour
Electrons/bunch	1.2×10^9	Plasma density	7×10^{14} cm^{-3}
Bunch charge	0.2 nC	Length	10 m
Bunch length	$\sigma_z = 4$ ps (1.2 mm)	Plasma radius	>1 mm
Bunch size at focus	$\sigma_{x,y} = 250$ µm	Skin depth	0.2 mm
Normalized emittance (r.m.s.)	2 mm mrad	Wavebreaking field, $E_0 = mc\omega_{cp}/e$	2.54 GV/m
Relative energy spread	$\Delta p/p = 0.5\%$		
Beta function	$\beta^*_x = \beta^*_y = 0.4$ m		
Dispersion	$D^*_x = D^*_y = 0$		

In the following sections, the different components required for a plasma wakefield experiment will be discussed for the example of the AWAKE project and the choices made in the design explained.

**Fig, 2:** Layout of the CERN accelerator chain

3.2 Drive beam

In the baseline design of the AWAKE experiment at CERN, an LHC-type proton bunch of 400 GeV/*c* will be extracted from the CERN SPS (see Fig. 2).

Simulations for optimization studies were performed [10] with the AWAKE baseline parameters as shown in Table 1. Figure 3 shows the simulated wakefield amplitude of a 400 GeV/*c* proton beam along a 10 m long plasma cell. With the chosen emittance, the baseline radius of the proton beam of $r = 0.2$ mm is the optimum; wide beams are not dense enough to drive the wave to the limiting field; narrow beams quickly diverge, owing to transverse emittance.

Figure 4 displays the variation of the driver energy at a constant normalized emittance. The nominal beam energy of the SPS at 400 GeV is well suited for the experiment.
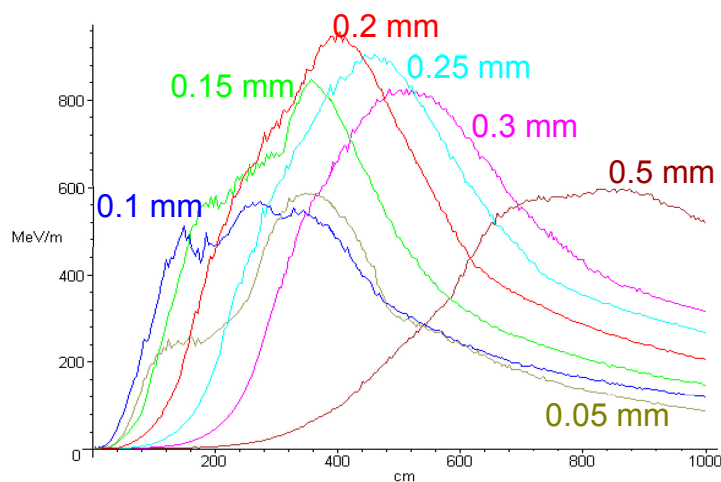
Fig. 3: Wakefield driven by 400 GeV proton beam along 10 m long plasma cell with different transverse beam sizes.
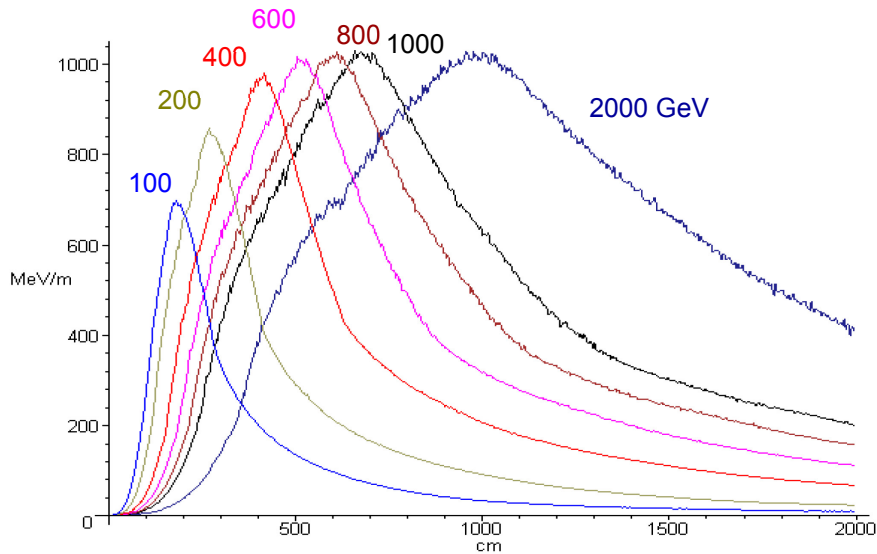
Fig. 4: Wakefield driven by proton beam at constant normalized emittance with different energies along plasma cell.

The AWAKE experiment requires short bunches with very high intensity from the SPS, which have not yet been used in operation by any other experiment. The required bunch intensity is above the instability threshold, so these bunches are unstable. If the bunches are unstable, it is difficult to control their parameters (particle distribution and bunch length) from shot to shot. However, they can be stabilized by increasing their longitudinal emittance (and therefore the bunch length) using controlled emittance blow-up.

To obtain smaller bunch lengths, bunches should be rotated in the longitudinal phase space by a quarter of the synchrotron period just before extraction on the SPS flat top (see Fig. 5). This is achieved by reducing and then sharply increasing the RF voltage. This RF manipulation was not used before in the SPS and will be introduced for AWAKE.

The shortest achievable bunch length is determined by the maximum available RF voltage and the smallest possible longitudinal emittance. The maximum voltage in the two RF systems of the SPS is currently 8 MV at 200 MHz and 600 kV at 800 MHz; this will be increased to 12 MV at 200 MHz (by around 2019) and 1.2 MV at 800 MHz (in 2015).

With this method, the SPS proton beam can be optimized to the following parameters: 3×10^{11} protons/bunch; normalized transverse emittance, 1.7 mm mrad; r.m.s. bunch length, 9 cm (0.3 ns); peak current, 60 A [11]. However, using conservative estimates, an r.m.s. bunch length of 12 cm (0.4 ns) will be used.
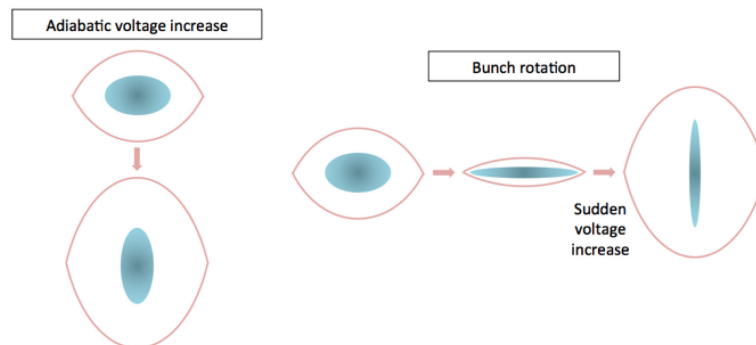
Fig. 5: Bunch rotation of SPS beam in longitudinal phase space instead of adiabatic voltage increase

The longitudinal SPS proton beam size of $\sigma_z = 12$ cm is much longer than the plasma wavelength (typically for AWAKE $\lambda_{\text{pe}} = 1$ mm). However, the long proton beam can be used directly for plasma wakefield acceleration experiments. This initial approach to proton-driven plasma wakefield acceleration will be through self-modulation of a long bunch of protons into a series of micro-bunches that drive resonantly large wakefields [12].

3.3 Self-modulation instability

Plasma wakefields are usually driven by laser pulses or particle bunches approximately one plasma wavelength long: $\sigma_z = \lambda_{\text{pe}}$. Here, $\lambda_{\text{pe}} = 2\pi c/\omega_{\text{pe}}$ is the wavelength of a relativistically moving plasma wave with electron plasma angular frequency $\omega_{\text{pe}} = (n_e e^2/\varepsilon_0 \mu_e)^{1/2}$ in a plasma with density n_e.

The longitudinal amplitude of the wakefield scales as the wave breaking field [11]:
$$E_{\text{WB}} = \mu_e c \omega_{\text{pe}}/e\alpha n_e^{1/2}.$$

This scaling therefore favours short pulses or bunches and large plasma densities to reach large amplitude wakefields. These wakefields, for sufficiently intense drivers, have accelerating or decelerating longitudinal components ($E_z \approx E_{\text{WB}}$) and transverse focusing or defocusing components with comparable amplitudes. In the linear wakefield regime, these fields vary periodically behind the drive bunch and have a $\pi/2$ phase difference.

Unlike recent plasma wakefield accelerator experiments, which employed short bunches ($\sigma_z < \lambda_{\text{pe}}$) to drive intense wakefields, the AWAKE experiment will use much longer proton bunches ($\sigma_z > 100\,\lambda_{\text{pe}}$) to generate plasma wakefields.

The AWAKE experiment will thus operate in the so-called self-modulated plasma wakefield accelerator regime [12]. In this regime, the maximum plasma density for wakefield excitation is given by the condition that the plasma return current must flow outside the drive bunch. This condition is satisfied when the bunch transverse size σ_r is smaller than the cold plasma collisionless electron skin depth c/ω_{pe} or when $k_{\text{pe}}\sigma_r < 1$ ($k_{\text{pe}} = \omega_{\text{pe}}/c$). When this condition is not satisfied, the bunch can be subject to the current filamentation instability. The instability breaks the bunch into transverse current filaments [13] and prevents the efficient excitation of plasma wakefields. For $\sigma_r = 200$ μm, setting $k_{\text{pe}}\sigma_r = 1$ yields $n_e = 7 \times 10^{14}$ cm^3, $\lambda_{\text{pe}} = 1.2$ mm and $E_{\text{WB}} = 3$ GV/m.

With relativistic bunches, energy gain and loss does not lead to significant dephasing between drive bunch particles over metre-scale plasma lengths:
$$\Delta L \approx \frac{1}{\gamma^2} \frac{\Delta\gamma}{\gamma} L \ll \lambda_{\text{pe}}$$

(for particles with energies γ and $\gamma \pm \Delta\gamma$ and plasma length L). This means that there is no longitudinal bunching.

However, the transverse wakefield components can periodically focus and defocus the particles that typically have non-relativistic transverse velocities,
$$\langle v_\perp \rangle \approx \frac{\varepsilon}{\sigma_0} c \ll c,$$

where ε is the beam transverse emittance and σ_0 its waist size [14].

3.3.1 Transverse modulation of a long bunch

When a long and narrow particle bunch travels in a dense plasma; i.e., when $\sigma_z \ll \lambda_{\text{pe}}$, it is subject to a transverse two-stream instability or SMI [13]. The low amplitude transverse wakefields driven by the long bunch modulate its radius with wavelength $\sim\lambda_{\text{pe}}$. This generates a micro-bunch pattern, as seen in Fig. 6 [15].

(z)
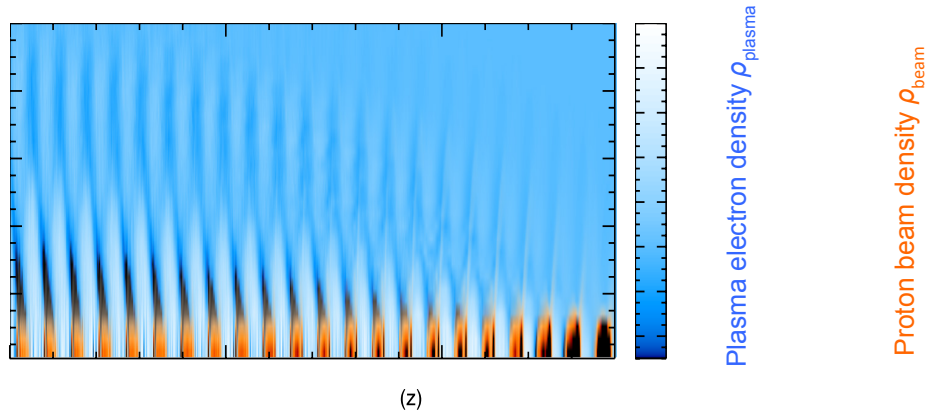
Fig. 6: Self-modulated proton bunch resonantly driving plasma wakefields sustained by plasma density perturbation. The plasma electron density is shown increasing from white to blue and the proton density increasing from yellow to dark red.

This periodic bunch density modulation then resonantly drives wakefields to larger amplitudes, thereby providing a feedback mechanism for the SMI to develop. The SMI is a convective instability that grows both along the bunch and along the plasma. If the drive bunch propagates in a uniform density plasma, then the instability destroys the micro-bunches soon after the maximum field is reached [16]. The reason lies in the slow motion of the defocusing field regions with respect to the bunch. This effect causes a strong decrease in the peak accelerating field, as seen in Fig. 7(a).

As the SMI grows, the interplay between bunch radius and wakefield amplitude leads to an effective wakefield phase velocity slower than that of the drive bunch [17, 18], as seen in Fig. 7(b). The figure shows the location of the accelerating and focusing fields along the bunch ($z - ct$) as a function of propagation distance along the plasma, z. Once the SMI saturates, these two velocities become equal.

It is possible to avoid the destruction of the micro-bunch structure by a proper step-up in the plasma density (Fig. 7(a)), which modifies the instability growth in such a way that the field motion relative to the bunches stops at the optimal moment.
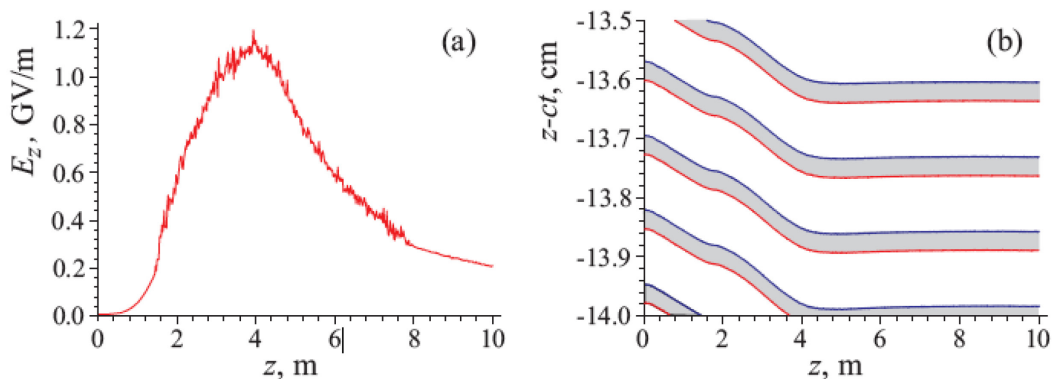
Fig. 7: (a) Maximum amplitude of accelerating field E_z excited along the bunch plotted as a function of position along the plasma. (b) Positions along the bunch ($z - ct$) where the wakefields are both accelerating and focusing for witness electrons are shown in grey as a function of propagation along the plasma. This position varies over the first 4 m of propagation and remains at the same $z - ct$ position after that.

3.3.2   Seeding of the self-modulation instability

The SMI can, in principle, grow from noise in the plasma and the drive bunch. However, seeding of the instability considerably shortens the plasma length needed for the SMI to reach saturation. Calculations [19] and simulations show that the noise level is very low and that the SMI would not grow to a

detectable level over metre-scale plasmas. More importantly, seeding the SMI fixes the phase of the wakefields, a condition necessary to deterministically inject a short witness bunch in the accelerating and focusing phase of the wakefields. The SMI can be seeded by a short laser pulse or particle bunch driving low amplitude wakefields in a preformed plasma ahead of the long drive bunch. A sharp (compared with λ_{pe}) boundary between the drive bunch and the plasma, such as a cut in the bunch current profile or a relativistic ionization front propagating within the drive bunch, also seeds the SMI [20].

The SMI leads to a radially symmetric modulation of the bunch charge density. However, there is an asymmetric competing instability known as the hose instability [21]. This instability is similar to the beam break-up instability in RF accelerators. It grows from noise in the transverse displacement of the bunch-slice centroids and results in a non-axially symmetric displacement of the bunch along its length. The hose instability has a growth rate comparable to that of the SMI, and the two instabilities directly compete. Simulations indicate that seeding helps the SMI dominate over the hose instability [17, 18].

3.4 Plasma cell

The plasma cell for the AWAKE experiment shall fulfil the following requirements.

— Plasma density n_e between 10^{14} and 10^{15} cm^{-3}: Fig. 8 shows the wakefield resulting from a proton beam with AWAKE parameters along a 10 m long plasma cell with different plasma densities [10]. For excessive plasma densities, filamentation and hosing instabilities can occur.

— Allow for seeding.

— Density uniformity $\delta n_e/n_e$ of the order of 0.2 %. The wakefield phase is determined by the plasma density. If the plasma wavelength changes locally, the witness electrons will be defocused. The density must be constant with an accuracy of $\delta_{pe}/4\sigma_z$.

— High-Z gas to avoid background plasma ion motion.

— Gas or vapour easy to ionize.

— Radius R_p larger than ≈3 proton bunch r.m.s. radii.

— Plasma length $L \sim 10$ m.
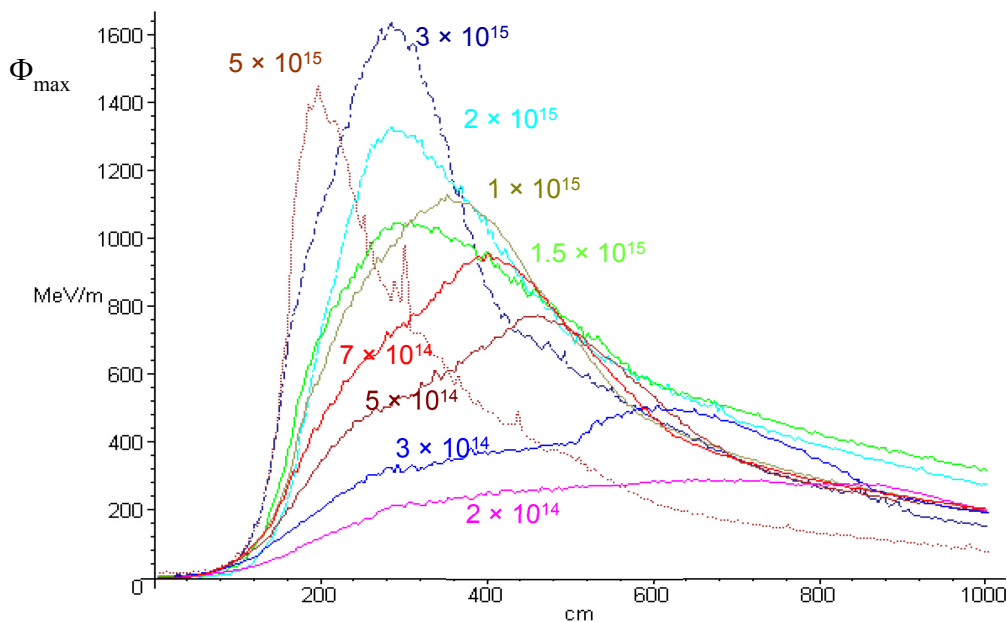
Fig. 8: Wakefield along 10 m long plasma cell with different plasma densities n_e

Several options exist for plasma cells: discharge plasma sources, helicon sources and metal vapour sources.

AWAKE performs developments using the discharge plasma [22] and helicon sources [23] as these are scalable in length.

In the first phases, and with these requirements, AWAKE will use a rubidium vapour source, 10 m long and 4 cm in diameter [24]. The density uniformity is achieved by controlling the tube containing the vapour to ±0.5 °C. This is achieved by circulating synthetic oil inside a thermal insulation around the tube containing the rubidium vapour. The oil temperature can be stabilized to ±0.01 °C. Figure 9 shows the 10 m long prototype of the AWAKE plasma cell installed in a test area at CERN.

Fig. 9: Prototype of 10 m long rubidium vapour plasma cell for AWAKE

A threshold ionization process for the first Rb electron is used to turn the uniform neutral density into a uniform plasma density. The ionization potential is very low, $\Phi_{Rb} = 4.177$ eV; as is the intensity threshold for over the barrier ionization (OBI), $I_{Ioniz} \approx 1.7 \times 10^{12}$ W/cm^2.

3.5 Laser beam

The plasma creation method uses a short laser pulse and has many advantages, since it also serves as ionization front seeding method.

A titanium:sapphire laser pulse with 30−100 fs pulse length, $\lambda_0 = 800$ ns and energy of 20−40 mJ provides the energy necessary to ionize the atoms in the plasma and the intensity to ionize along the 10 m long plasma cell [25].

The laser beam moves co-linearly with the proton beam, creating the ionization front, which acts as if the proton beam is sharply cut and in turn directly seeds the SMI.

Figure 10 shows the distribution of the beams after propagating 4 m in the plasma. Protons are blue, electrons are red and the laser pulse is the pink line at $z−ct = 0$. The laser pulse seeds the SMI for the proton bunch: only the protons that are behind the laser pulse are affected.
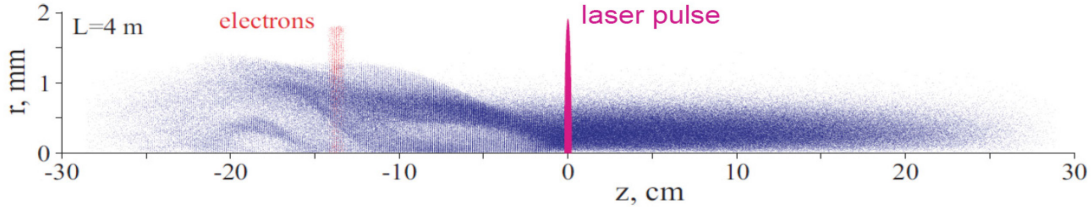
Fig. 10: Beam distribution after propagating 4 m in the plasma. Protons are blue

3.6 Witness beam

The electron beam is externally injected into the plasma wakefield driven by the proton beam. The considerations for the electron beam specifications are that the electrons must be trapped in the grey area shown in Fig. 7b, where the wakefields are both accelerating and focusing. The optimal electron energy is 10−20 MeV, which corresponds to the wakefield phase velocity at the self-modulation stage.

Table 2 shows the electron beam parameters fitting to the requirements that will be used in AWAKE for the first phase of the experiment. Simulations show a trapping efficiency of the order of ~10% when injecting electrons on-axis at the beginning of the plasma cell [26]. With these parameters and after passing through the 10 m long plasma cell, the expected average energy gain of the electrons will be 1.3 GeV.

Initially, the injected electron bunch will be at least one plasma period long, to avoid precise timing of a much shorter electron bunch with respect to the wakefields [27]. However, in this scheme, the trapping efficiency of electrons is very sensitive to the length of a possible density ramp at the plasma entrance, whereas that of protons is not [12].

The numbers are rather conservative, also allowing the use of the provided electron source without major modifications. However, at a later stage, the requirements for the bunch charge (i.e., 1 nC) and bunch length (i.e., 0.3 ps) will be much tighter, to optimize the performance of the electron acceleration.

Many injection parameters must be optimized in the experiment:

Figure 11 shows the comparison between two different electron injection delays with respect to the laser pulse; in Fig. 11(a), the electron beam is well caught in the wakefield and is accelerated to 1.2 GeV after 10 m. In Fig. 11(b), the different injection delay causes a only very weak acceleration.
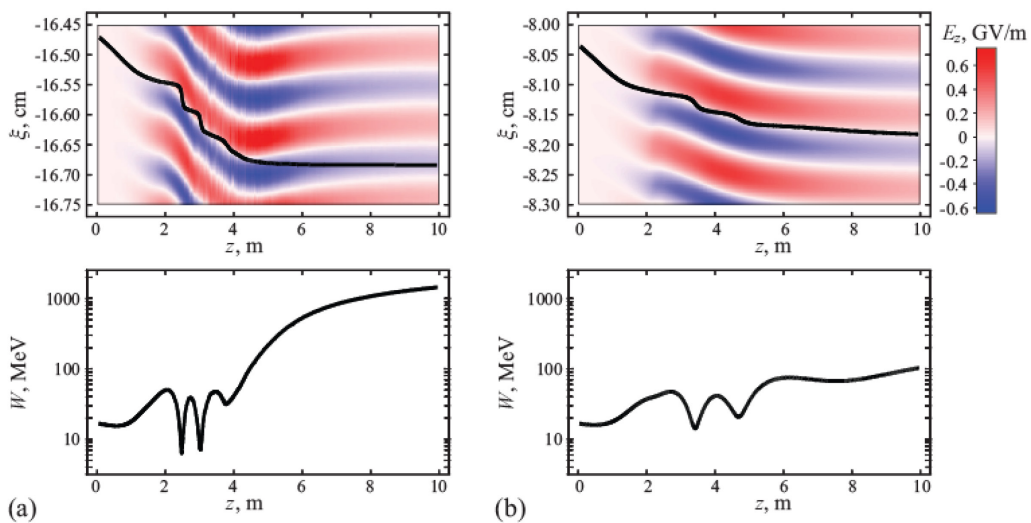
Fig. 11: Co-moving coordinate ξ (top) and energy (bottom) versus propagation distance for two typical test electrons injected with different delays with respect to the laser pulse. The top plot also shows the color map of the on-axis electric field E_z in the vicinity of the electron.

3.6.1   Electron source

The PHIN Photo-Injector built for CTF3 [28], which will have completed its experimental programme for the Compact Linear Collider at the end of 2015, fulfils the baseline electron beam requirements (10−20 MeV/c, 1.2×10^9 electrons/bunch, σ_z =4 ps, $\sigma_{x,y}$ = 250 μm at focus, ε_{norm} = 2 mm mrad) and will be used as electron source for the AWAKE experiment. The electron source consists of a metal photocathode housed in a 2.5-cell normal conducting RF cavity, a load-lock system for cathode exchange, a 3 GHz booster accelerator, to boost the energy to the required 20 MeV/c and electrical and optical diagnostics (beam position monitors, fast current transformer, Faraday cup, and screens for emittance and spectrometer measurements). In addition, the klystron and modulator system will be recuperated from CTF3, providing 3 GHz RF power.

3.7   Diagnostics for the drive beam

The first experiments in AWAKE will be aimed at demonstrating and studying the SMI of the long proton bunch in the dense plasma. Therefore, beam diagnostic tools to measure the radial self-modulation of the bunch density will be developed.

3.7.1   Direct SMI measurements

The direct SMI diagnostic tools [29] are based on transformation of the charge distribution information into a radiation distribution using transition radiation. Transition radiation is emitted when a single charged particle or a collection of charged particles cross a boundary between two media with different dielectric constants (e.g., vacuum—metal interface). The radiation is incoherent at wavelengths much shorter than the characteristic charge distribution scale (length or radius) and in the visible range is called optical transition radiation. It is coherent at wavelengths much longer than the characteristic charge distribution scale and is called coherent transition radiation or transverse coherent transition radiation.

With optical transition radiation, the bunch spatial or temporal characteristics are contained in the light intensity and can be measured using, e.g., a streak camera. The best streak cameras have a single pulse temporal resolution of ~1 ps down to ~200 fs. The period of the modulation (expected to be that of the plasma wave) can, in principle, be measured up to frequencies of ~300 GHz. For the AWAKE plasma densities, the plasma and modulation frequency range is 100−300 GHz.

The modulated proton bunch also emits coherent transition radiation, whose spectrum reflects the bunch longitudinal and transverse structure (radial modulation with longitudinal period approximately that of the plasma wave). Therefore, coherent transition radiation emission at the plasma frequency from the radially modulated proton bunch can, in principle, be detected. The long proton bunch radiation can be filtered out by a section of waveguide in cut-off (high-pass filter) and the high-frequency time evolution of the radiation detected using a fast Schottky diode (~200 ps rise time). In particular, coherent transition radiation emission at high frequencies should be correlated with the ionizing laser position within the proton bunch, i.e., with the SMI seed. The modulation frequency can, in principle, be determined using a heterodyne measurement system, mixing the RF modulation signal of unknown frequency with a known local oscillator frequency in a crystal that generates the difference or intermediate frequency. By choosing appropriate frequencies and a suitable crystal, one can bring the intermediate frequency within the measurement range.

3.7.2   Indirect SMI measurements

Another method is to detect the SMI effect by measuring the angular divergence of the proton beam caused by the SMI in the plasma cell, which is of the order of ~1 mrad (M. Turner and A. Petrenko, private communication). For this purpose, the bunch profile is measured at two different scintillator screens installed downstream of the plasma cell and at a distance of ≈8 m with a resolution of ~0.1 mm.

Figure 12(a) shows the energy deposition of the proton beam in the first scintillator screen downstream of the plasma cell; Fig. 12(b) shows the energy deposition in the second scintillator screen, which is ≈8 m downstream of the first one. With this method the saturation point of the SMI inside the plasma cell can also be measured, at a 2% level.
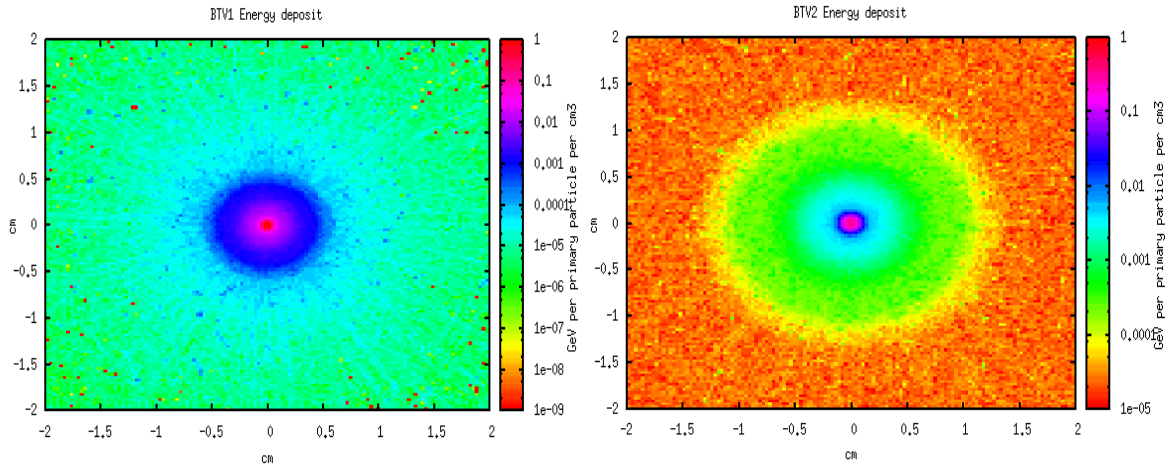
Fig. 12: Energy deposition in GeV/cm^3 per primary proton particle of the self-modulated proton beam downstream of the plasma cell. (a) Scintillator 1 m downstream of the plasma cell. (b) Scintillator 9 m downstream of the plasma cell.

3.7.3   *Witness beam diagnostics*

The purpose of the experiment is to accelerate the externally injected electrons. A magnetic spectrometer was developed with a relatively large energy acceptance, from a few hundred megaelectronvolts to a few gigaelectronvolts [30]. The electron spectrometer system consists of a C-shaped magnet providing a 1.5 T field to separate the electrons from the proton beam and disperse them in energy. The spectrometer system includes a quadrupole doublet in a point-to-point imaging configuration, to focus the beam exiting the plasma onto the spectrometer screen and increase the energy resolution. The electrons impinge on a scintillating screen (the baseline is gadolinium oxysulfide) and an optical line will transport the screen light to a CCD camera. Calculations show that a percentage-level energy resolution can be achieved with a signal-to-noise ratio greater than 1000:1.

4   The AWAKE facility—putting the pieces together

The AWAKE experiment will be integrated in an existing underground area, which previously housed the CNGS facility [31]. The different components described in Section 3 must be put together and connected with beam lines.

Figure 13 shows the integration of the AWAKE experiment in the experimental area.
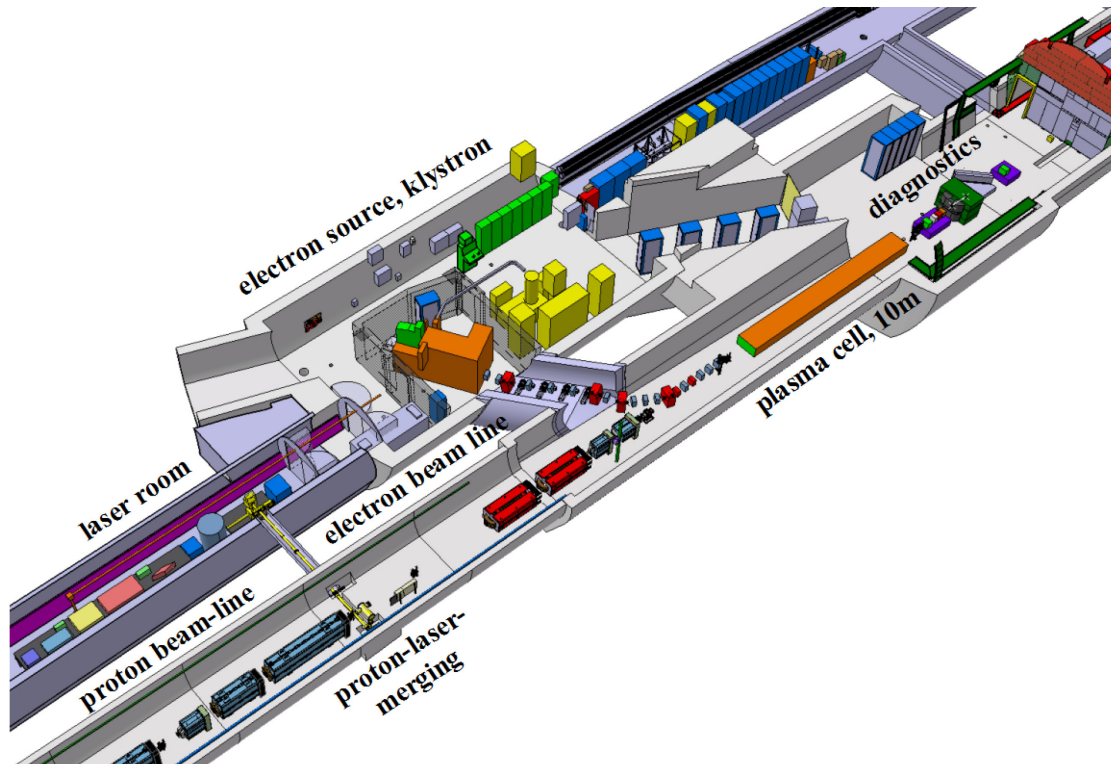
Fig. 13: Integration of the AWAKE experiment in the experimental area

4.1 Proton beam line

The proton beam is extracted from the CERN SPS and sent along the 850 m long proton beam line towards the plasma cell (see Fig. 2). The existing proton beam line is modified only in its last 80 m or so, i.e., in the matching section and in the final focusing part, to comply with the AWAKE requirements. The laser beam delivered from the laser source is merged with the proton beam at a distance of ≈20 m from the plasma cell by adding a proton beam chicane to integrate the laser mirror. An offset of 19.9 mm exists between the proton and the laser beam axis, enough clearance to avoid intercepting protons and inducing losses [32]. The proton and laser beams are made co-axial over the full length of the plasma cell, in particular the 3σ proton beam envelope (≈0.6 mm) must be contained in the 1σ laser spot size (≈1 mm). A pointing precision of 100 µm is required at the cell entrance, resulting in a maximum angular error of 15 µrad for the proton beam line. To achieve this, the ripple in the current of the main dipole power converter is kept below 5×10^{-4}. A 0.2 mm thick aluminium window will be installed in the proton beam line to separate the SPS and the AWAKE vacuums. Modifications and new beam diagnostics are required at the end of the proton beam line; the electronics of the beam position monitors are modified to allow for single bunch measurements. Two BTVs, located about 1 m upstream and downstream of the plasma cell, will perform profile and position measurements and will also be used during set-up to align the proton and laser beams. The synchronization between the two beams will be adjusted with a streak camera connected to a BTV, located ~2 m upstream of the plasma cell.

4.2 Laser beam line

A gallery previously used for CNGS equipment storage is modified to become a dust-free (over-pressurized) and temperature-stabilized area to house the laser system. Owing to the small size of the gallery, the integration of the laser, pulse compressor, laser beam transport optics, laser tables, racks, and clean-room features, such as a special cooling and ventilation system, and a double door access system, is very challenging, but has been successfully achieved [33]. Two laser beams are to be transported from the laser room. The first laser beam is used for plasma production and seeding of the

SMI in the proton bunch. For that aim a 4 m long laser core (50 cm in diameter) was drilled between the laser room and the proton beam line tuning. As the peak power of the 100 fs compressed laser pulse is very high, the laser beam must be injected into the vacuum system of the SPS proton beam before laser pulse compression. Therefore, the vacuum laser beam line connecting the laser with the proton merging area has to fulfil the requirements of the SPS vacuum system (10^{-7} mbar), resulting in a demanding design for the laser line optics and vacuum system.

The second laser beam is required for the generation of the electron beam and has to be transported to the electron source.

4.3   Electron beam line

A new 15 m long electron beam line has been designed to transport the electron beam from the RF gun, across a newly built tunnel (7 m long, 2.5 m wide), towards the proton beam line to be injected into the plasma cell on the same axis as the proton and laser beam. To comply with experimental requirements, the electron beam optics must provide a flexible design so that the focal point can be varied by up to 6 m inside the plasma cell [34]. The proton and the electron beams share the last ≈5 m of the line upstream of the plasma cell. The technical parameters for the new magnets (4 dipoles, 11 quadrupoles, and 11 correctors) have been specified. In addition, specifications were defined for the electron beam line diagnostics; e.g., the beam position monitors should be able simultaneously to measure the position of electrons in the presence of the AWAKE proton beam. A peak-to-peak resolution of 50 μm is required. The electron beam shall be synchronized with the proton and laser beam at a <1 ps level, and will be measured with a streak camera linked to a BTV. To meet the performance criteria of the light transmission system, a straight line between the beam screen and the camera was required, which implied the excavation of a ≈4 m long 15 cm diameter core between the proton line and the electron source area, where the streak camera will be installed.

4.4   Beam synchronization

Synchronization between the laser pulse and the electron beam at the level of a few tens of femtoseconds (a fraction of the plasma period of ≈4 ps) is required for the deterministic injection of the witness bunch into the plasma wakefields. This is achieved by driving the RF gun of the electron source with a laser pulse derived from the same laser system used for plasma ionization and seeding. The synchronization between the proton and the laser beam must be better than 100 ps, i.e., better than the r.m.s. proton bunch length of ≈400 ps. The laser mode-locker ($f_{ML} = 88.17$ MHz) cannot follow the relative revolution frequency change through the SPS acceleration cycle, and therefore the proton bunch in the SPS must be synchronized to the AWAKE reference prior to extraction.

A low phase-noise RF source at 3 GHz is made the master, to which the 34th harmonic of the laser oscillator (at f_{ML}) is locked [35]. Lower frequencies for fast triggers (176.3 MHz), SPS synchronization (8.68 kHz) and laser or electron beam repetition rate (10 Hz) are derived by integer division from the master oscillator. Only the 200 MHz RF source for the synchronization with the SPS requires a fractional divider.

A system is foreseen that allows the exchange of synchronization signals on ≈3 km long fibres between the AWAKE facility and the SPS RF Faraday cage; it has been verified that the jitter and drift of the signal transmission do not exceed the picosecond range, slow variations will be compensated by a feedback loop.

5   Summary

The AWAKE experiment will, for the first time, use a proton bunch to drive wakefields in a plasma. The first experiments, scheduled to begin in late 2016, will aim at studying the development and seeding

of the SMI of the proton bunch in a ≈10 m long, 10^{14} to 10^{15} cm^{-3} electron density plasma. Later experiments, scheduled for 2018, will study the injection of a RF-gun-produced long electron bunch in the wakefields and its acceleration to gigaelectronvolt energies. Future experiments will use two plasma sources and an ultra-short electron bunch to address accelerator related issues [29]. Long-term prospects include using a short proton bunch to drive wakefields without resort to SMI with the possibility of accelerating electrons to very high energies, possibly at the energy frontier.

References

[1] I. Blumenfeld *et al.*, *Nature* 445 (2007) 741. http://dx.doi.org/10.1038/nature05538

[2] M. Litos *et al.*, *Nature* 515 (2014) 92. http://dx.doi.org/10.1038/nature13882

[3] R. Assmann *et al.* (AWAKE Collaboration), *Plasma Phys. Control. Fusion* 56(8) (2014) 084013. http://dx.doi.org/10.1088/0741-3335/56/8/084013

[4] https://www6.slac.stanford.edu/facilities/facet.aspx

[5] M. Gross *et al.*, Preparations for a plasma wakefield acceleration (PWA) experiment at PITZ, *Nucl. Instrum. Methods* A **740** (2014) 74. http://dx.doi.org/10.1016/j.nima.2013.11.042

[6] https://vi-pwfa.desy.de/

[7] https://www.bnl.gov/atf/

[8] A. Caldwell *et al.*, AWAKE design report, CERN-SPSC-2013-013 (2013).

[9] A. Caldwell and K.V. Lotov, *Phys. Plasmas* **18**(10) (2011) 103101. http://dx.doi.org/10.1063/1.3641973

[10] K. Lotov *et al.*, *Phys. Plasmas* **21**(8) (2014) 083107. http://dx.doi.org/10.1063/1.4892183

[11] H. Timko *et al.*, Short high-intensity bunches for the plasma wakefield experiment AWAKE in the CERN SPS, Proc. IPAC2013 (Shanghai, China), p. 1820.

[12] N. Kumar *et al.*, *Phys. Rev. Lett.* **104**(25) (2010) 255003. http://dx.doi.org/10.1103/PhysRevLett.104.255003

[13] B. Allen *et al.*, *Phys. Rev. Lett.* **109**(18) (2012) 185007. http://dx.doi.org/10.1103/PhysRevLett.109.185007

[14] C.E. Clayton *et al.*, *Phys. Rev. Lett.* **88**(15) (2002) 154801. http://dx.doi.org/10.1103/PhysRevLett.88.154801

[15] J. Vieira *et al.*, *Phys. Plasmas* **19**(6) (2012) 063105. http://dx.doi.org/10.1063/1.4725425

[16] K.V. Lotov, *Phys. Plasmas* **18**(2) (2011) 024501. http://dx.doi.org/10.1063/1.3558697

[17] A. Pukhov *et al.*, *Phys. Rev. Lett.* **107**(14) (2011) 145003. http://dx.doi.org/10.1103/PhysRevLett.107.145003

[18] C.B. Schroeder *et al.*, *Phys. Rev. Lett.* **107** (14) (2011) 145002. http://dx.doi.org/10.1103/PhysRevLett.107.145002

[19] K.V. Lotov *et al.*, *Phys. Rev. ST Accel. Beams* **16**(4) 041301(2013). http://dx.doi.org/10.1103/PhysRevSTAB.16.041301

[20] D. Gordon *et al.*, *Phys. Rev.* E **64**(4) (2001) 046404. http://dx.doi.org/10.1103/PhysRevE.64.046404

[21] D.H. Whittum *et al.*, *Phys. Rev. Lett.* **67**(8) (1991) 991. http://dx.doi.org/10.1103/PhysRevLett.67.991

[22] S.N. Bland *et al.*, An argon discharge plasma source for the proton-driven plasma wakefield accelerator experiment at CERN, EDMS 1275630 (2013).

[23] O. Grulke *et al.*, R&D towards a helicon plasma cell, EDMS 1275621 (2013).

[24] E. Oz *et al.*, *Nucl. Instr. Meth. Phys. Res.* A **740** (2014) 197–202.
http://dx.doi.org/10.1016/j.nima.2013.10.093

[25] J. Moody *et al.*, Laser propagation effects during photoionization of meter scale rubidium vapor source, Proc. IPAC'15, Richmond, VA, USA, (2015), paper WEPWA006.

[26] K. Lotov *et al.*, *Phys. Plasmas* **21**(12) (2014) 123116. http://dx.doi.org/10.1063/1.4904365

[27] A. Petrenko *et al.*, Electron injection studies for the AWAKE experiment at CERN, Proc. IPAC2014, Dresden, Germany (2014), p. 1537.

[28] M. Petrarca *et al.*, Performance of the PHIN high charge photo injector, IPAC2010, Kyoto, May 2010, THPEC032. http://www.JACoW.org

[29] P. Muggli, The AWAKE proton-driven plasma wakefield experiment at CERN, WEPWA007, IPAC'15, Richmond, USA (2015).

[30] S. Jolly *et al.*, A spectrometer for proton driven plasma wakefield accelerated electrons at AWAKE, Proc. IPAC2014, Dresden, Germany (2014), p. 1540.

[31] E. Gschwendtner, AWAKE: advanced proton driven plasma wakefield acceleration experiment at CERN, CERN-ACC-2014-0198, LINAC 2014, Geneva, Switzerland.

[32] C. Bracco *et al.*, Beam transfer line design for a plasma wakefield acceleration experiment (AWAKE) at the CERN SPS, Proc. IPAC2013, Shanghai, China (2013), p. 1247.

[33] E. Gschwendtner *et al.*, The AWAKE experimental facility at CERN, Proc. IPAC2014, Dresden, Germany (2014), p. 582.

[34] J.S. Schmidt *et al.*, The AWAKE electron primary beam line, WEPWA039, IPAC'15, Richmond, VA, USA, 2015.

[35] M. Bernardini *et al.*, AWAKE, the proof-of-principle R&D experiment at CERN, in Proc. IPAC'15, Richmond, VA, USA (2015), paper MOAC1.

An Overview of Recent Progress in Laser Wakefield Acceleration Experiments

S.P.D. Mangles
John Adams Institute for Accelerator Science, Blackett Laboratory, Imperial College London, UK

Abstract
The goal of this paper is to examine experimental progress in laser wakefield acceleration over the past decade (2004–2014), and to use trends in the data to understand some of the important physical processes. By examining a set of over 50 experiments, various trends concerning the relationship between plasma density, accelerator length, laser power and the final electron beam energy are revealed. The data suggest that current experiments are limited by dephasing and that current experiments typically require some pulse evolution to reach the trapping threshold.

Keywords
Laser wakefield accelerators; plasma accelerators; laser-plasma acceleration.

1 Introduction

This paper is a summary of a lecture given at the CERN Accelerator School 2015 on plasma-based wakefield accelerators. Its purpose is to provide an overview of recent experimental progress in laser wakefield acceleration, concentrating on the energy frontier and general trends that can be observed in the data produced by the many groups around the world contributing to this growing field. There are now more than 20 active laboratories performing laser wakefield acceleration experiments. We will not detail the key results from each of these laboratories in this paper, although we will examine data from various published experiments and use the trends in the data to try to gain some understanding of the underlying physical processes.

Laser wakefield accelerators were proposed 36 years ago in the seminal 1979 paper by Tajima and Dawson [1] and experiments in laser wakefield acceleration were undertaken as soon as laser pulses of sufficiently short duration and high power became available, thanks to the development of the laser technique called 'chirped pulse amplification' [2]. Some of the early work in what we call laser wakefield acceleration (i.e., where the pulse duration, τ_L is comparable to the plasma period $2\pi/\omega_p$) occurred in the 1990s (e.g., Refs. [3, 4]) using picosecond glass lasers. The field passed a major milestone in the early 2000s as high-power ($\sim$10 TW) femtosecond laser pulses, using titanium sapphire laser systems, became available. One major result from that era occurred when three groups from the UK, USA and France all demonstrated that laser wakefield accelerators could produce electron beams with well-defined energies [5–7]. The electron beams produced by these $\sim$10 TW lasers had energies of $\sim$100 MeV and were produced in plasmas of a plasma density $\sim$10^{19} cm^{-3} that were only $\sim$1 mm long.

One of the key challenges that has driven progress in the field of laser wakefield acceleration is the maximum achievable beam energy, and this paper will concentrate on this challenge. There have, of course, been many other significant areas of experimental progress including: improving beam stability, especially by controlling injection (e.g., Refs. [8, 9]); diagnosis of wakefield accelerators (e.g., Refs. [10, 11]) and the resulting improvements in our understanding of the underlying processes; and the application of laser wakefield accelerators for a range of applications, perhaps most notably their use as novel sources of X-radiation (e.g., Refs. [12–14]). Details of the progress in these areas are outside the scope of this paper.

This paper will examine how the maximum achievable beam energy has progressed over the last decade, and will use a set of published results [5–9,15–60] to try to understand some of the physics behind
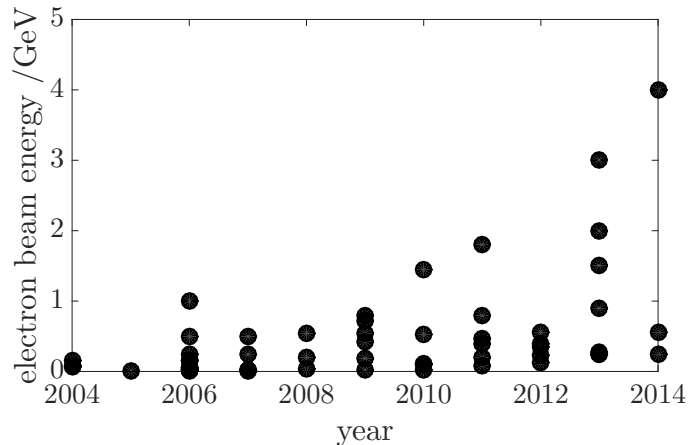
Fig. 1: Reported electron beam energies from laser wakefield experiments at various laboratories over the last decade; data from Refs. [5–9, 15–60].

this progress. The progress in the maximum beam energy in laser wakefield accelerator experiments has been rapid, as shown in Fig. 1, from a maximum beam energy of 0.2 GeV reported in 2002 [61] to the current record of 4 GeV from the group at the Lawrence Berkeley National Laboratory [58] achieved in 2014 – an increase by a factor of 20 in just over a decade. It should be noted that this is by no means an exhaustive list of all published experiments in laser wakefield accelerators (there are just 52 publications in this dataset, whereas a literature search for papers on 'laser wakefield' will find over 1000 papers).

## 2 Overall trends in laser wakefield acceleration experiments

The rapid progress shown in Fig. 1 is impressive. But how has it been achieved? Over the same period of time, short-pulse ($\approx$30 fs) laser systems have become more powerful. Figure 2 shows that there is a clear trend: higher-power lasers are capable of producing higher-energy electron beams. However, these gains were not achieved by simply increasing the laser power, the researchers behind these experiments have often found the optimum conditions for their experiments. Key parameters involved in this optimization include the operating plasma density, the length of the accelerator and the laser intensity. This section will examine the data from various experiments and compare them with predicted trends, to see whether they can confirm those predictions and the underlying physical processes.

2.1 Accelerator length and operating plasma density

Let us first consider the density of the plasma accelerator. The energy gained by a particle of charge q in an accelerating structure is simply proportional to the product of the electric field and the length of the accelerator, d, i.e., $W \simeq qEd$, where E is the average accelerating electric field experienced by the particle. One of the key physical limitations in a laser wakefield accelerator is dephasing. The electrons trapped in a wake are highly relativistic ($\gamma \gg 1$), so they travel at a speed approaching that of light in vacuum ($v_e \to c$), but the phase speed of the wake is determined by the speed of the laser pulse that drives the plasma wave. A simple expression for the speed of a laser pulse in a plasma can be found by using the standard expression for the group velocity of an electromagnetic wave in a plasma,

$$\frac{v_g}{c} = \sqrt{1 - \frac{n_e}{n_c}} \approx 1 - \frac{1}{2}\frac{n_e}{n_c} \, , \tag{1}$$

where n_e is the electron density of the plasma, n_c is the critical density for propagation of the electromagnetic (i.e., when the plasma frequency ω_p equals the frequency of the electromagnetic wave, ω_0) and
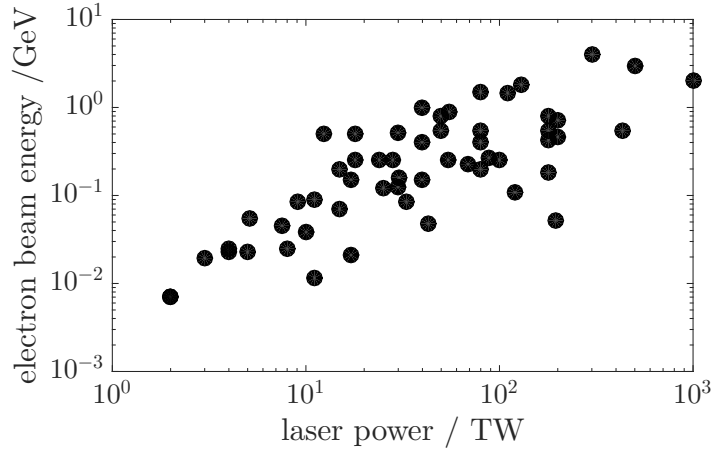
Fig. 2: Variation of reported electron beam energy with laser power from various experiments; data from Refs. [5–9, 15–60].

it is assumed that $n_e \ll n_c$. The wake's phase speed is therefore slightly, but significantly, less than c; crucially, the lower the plasma density, the faster the phase velocity.

Because of this difference between the electron velocity and the wake phase velocity, electrons in a laser-driven wake will outrun the wake[1]. If the electron is injected at the start of the accelerating phase of the plasma wave and then outruns the wave by half a plasma wavelength (i.e., $\lambda_p/2 = \pi c/\omega_p$), it can no longer gain energy from the plasma wave. If the electron has an initial velocity $v_e = \beta_e c$ and the plasma wave has a phase velocity $v_\phi = \beta_\phi c$, then the time it takes for this to occur is $t_d = \lambda_p/(2c(\beta_e - \beta_\phi))$. The dephasing length is then the distance that the electron travels in this time. Since $\beta_e \to 1$ and $\beta_\phi \simeq 1 - \frac{1}{2}(n_e/n_c)$, this reduces to

$$L_{\text{dephasing}} \simeq \frac{n_c}{n_e}\lambda_p \propto n_e^{-\frac{3}{2}} . \tag{2}$$

It is interesting to see how the lengths of the accelerators in the set of experiments vary, and how this compares with what we might expect if dephasing is important. These data are shown in Fig. 3. The top panel shows how the reported electron beam energy varies with the length of the wakefield accelerator. There is a clear correlation – the higher electron energies are achieved with longer accelerators, as we might expect. The bottom panel of Fig. 3 shows how the length of the accelerator and the plasma density at which it was operating are related. The line on this curve is the simple expression for the dephasing length (Eq. (2)).

The maximum electric field that a plasma wave can support increases with plasma density, since it scales as

$$E_{\text{max}} \simeq m_c\omega_p/e \propto \sqrt{n_e} . \tag{3}$$

The maximum energy that can be gained by an electron in a plasma wave as a function of plasma density is therefore expected to be

$$W(n_e) \simeq E_{\text{max}}L_{\text{dephasing}} \propto \frac{1}{n_e} . \tag{4}$$

---

[1]Note that dephasing does not occur in wakefield accelerators driven by highly relativistic charged particle beams as both the accelerated and driver beams are highly relativistic.
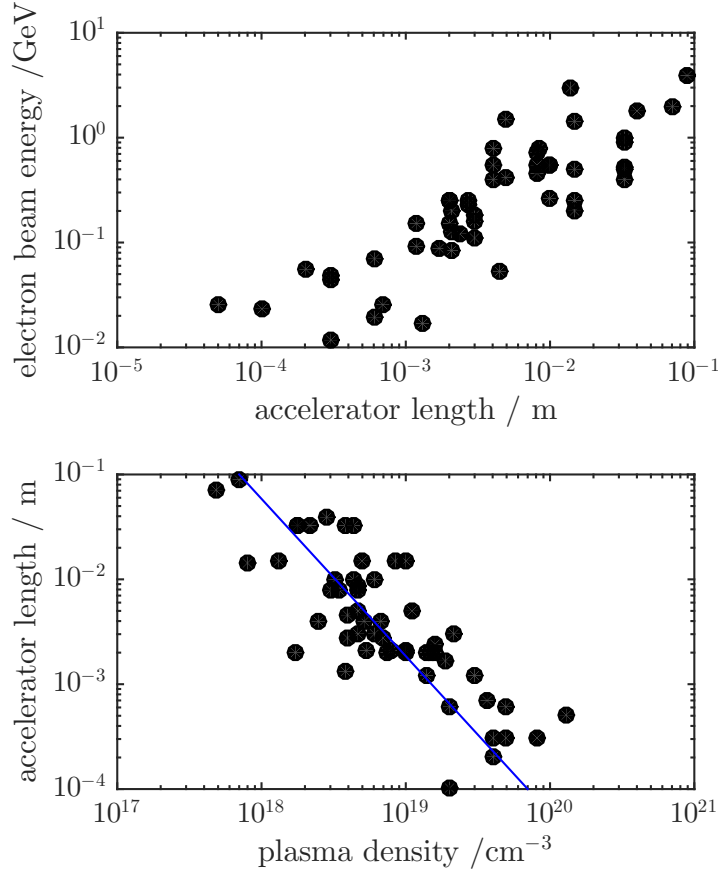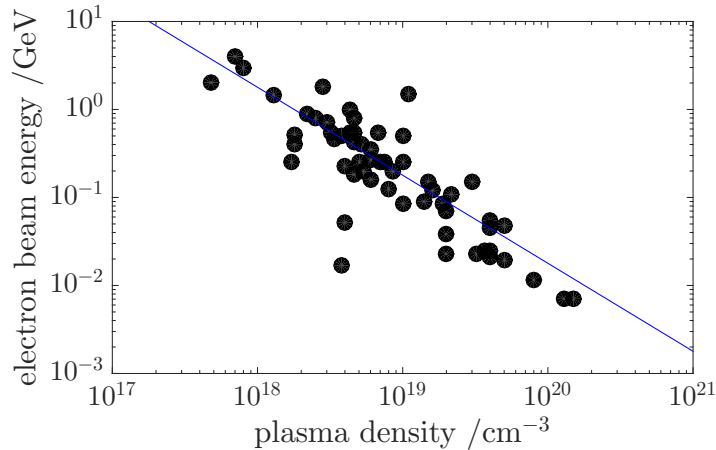
Fig. 3: Top: Variation of reported electron beam energy with accelerator length. Bottom: Relationship between operating plasma density and accelerator length. The line shows the expression for the dephasing length (Eq. (2)). Data from Refs. [5–9, 15–60].

Figure 4 shows how the plasma density and beam energy vary in the set of experiments. The line on Fig. 4 is simply $W/(m_e c^2) = \kappa\, n_c/n_e$ and shows good agreement with the entire dataset for $\kappa = 1$. The scaling laws in Ref. [62], by Wei Lu *et al.*, for the blow-out or 'bubble' regime of wakefield accelerators, suggest that the scaling law should be

$$W(n_e, a_0) \simeq \frac{2}{3} a_0 \frac{n_c}{n_e} m_e c^2 \propto \frac{a_0}{n_e}\,, \tag{5}$$

where $a_0 = eA_0/(m_e c) = eE_0/(m_e \omega_0 c)$ is the normalized peak vector potential (or strength parameter) of the laser pulse. This scaling predicts that the beam energy should not only be proportional to $1/n_e$ but also proportional to the laser strength, a_0. The experiments shown correspond to a wide range of initial laser intensities (corresponding to $a_0 = 0.5$–4.0), yet they do not appear to show a dependence on a_0.

One possible reason for this apparent discrepancy is that the initial value for a_0 is not the value of a_0 that determines the wake amplitude. It is well known that laser pulses can undergo significant evolution once they enter the plasma. The processes of self-focusing, self-compression and photon acceleration [63] can all act to change a_0 as the pulse propagates. Together, these processes can be termed the 'self-evolution' of the laser pulse. One interpretation of the experimental data is, therefore, that the process of self-evolution has occurred until $a_0 \approx 3$ for all of the data shown. Why would this be the case? One reasonable hypothesis is that each point in the dataset corresponds to the maximum

Fig. 4: Variation of reported electron beam energy with the density in the accelerator. The line shows the relation $W/(m_e c^2) = \kappa \, n_c/n_e$ with $\kappa = 1$. Data from Refs. [5–9, 15–60].

energy achieved during a particular experiment and that this will occur at (or at least close to) the lowest density at which a particular experiment can trap and accelerate electrons. Many experiments operate by fixing the laser power and plasma length while varying the plasma density, for reasons of experimental simplicity. When an experiment is conducted in this manner, there will be a minimum density at which electron beams are trapped and accelerated (the trapping threshold). Because self-evolution happens less quickly and severely at lower densities, the maximum a_0 that is reached inside the accelerator will decrease with decreasing plasma density. Therefore, the maximum achieved electron beam energy will correspond to the minimum laser strength required to produce trapping. The fact that the experimental dataset matches the non-linear wakefield scaling but only if $a_0 \approx 3$ suggests that the minimum a_0 required for trapping is $a_0 \approx 3$.

2.2 Laser spot size and matched guiding

The experimental data clearly show that higher-power lasers are required to achieve higher electron beam energies. But what are the physical processes behind this trend? It was argued that the experimental trends are consistent with there being a minimum value of $a_0 \approx 3$, which is needed for trapping, and that this value is reached because of the way the pulse evolves as it propagates.

Consider the relationship between the intensity, I ($\propto a_0^2$), and power, P, of a laser pulse. Since $I = P/A$, where A ($\propto w^2$, the laser spot size) is the focal spot area, we have

$$P \propto a_0^2 w^2 \,. \tag{6}$$

The fact that higher-power lasers are needed to reach the $a_0 \approx 3$ threshold at lower densities therefore implies that the spot size that these laser pulses produce after evolution is larger. Pulse evolution is a result of the feedback between the refractive index gradient associated with the plasma wave and the laser pulse, i.e., it is mediated by the plasma itself. Lower-density plasmas, therefore, have a lesser effect on the laser pulse – resulting in slower evolution – and crucially this affects the properties that the pulse obtains as a result of self-evolution.

One important concept that arises from this is that of the *matched* spot size – i.e., one where the self-focusing caused by the plasma balances the natural diffraction of the laser pulse and stable propagation occurs. In the blow-out regime (where the laser pulse expels practically all the electrons from inside the bubble), the transverse density profile of the bubble is approximately zero and flat inside

the bubble, with very steep walls at the edges. The refractive index of an underdense plasma is $\eta \approx 1 - n_{\text{e}}/(2n_{\text{c}})$, so the refractive index profile of this bubble is similar to that of a single-mode optical fibre, i.e., we have a 'core' of a certain diameter (the bubble diameter) surrounded by 'cladding' (the bubble sheath) of a lower refractive index. The guided mode in such an optical fibre has a transverse size that is approximately equal to the size of the core, so we expect that in a laser wakefield accelerator we will get stable propagation (i.e., no spot size oscillations) when the transverse size of the laser pulse is approximately equal to that of the bubble.

Of course, one main difference between the bubble and an optical fibre is that the size of the bubble is determined by the properties of the laser pulse itself. If the laser spot is too small, then it will drive a bubble that is larger than the laser spot. This over-sized bubble will support a larger mode and the laser pulse will expand, which in turn reduces the bubble size. The radius of the bubble can be found by balancing the ponderomotive force of the laser pulse with the force due to the electric field of the bubble. If the bubble is approximately the same size as the laser pulse then it will be matched. The size of the matched spot can be found by balancing the forces on an electron at the edge of the bubble (i.e., the laser's ponderomotive force and the force due to the electric field inside the bubble). An expression for the matched spot size, w_{m} is

$$w_{\text{m}} \simeq 2\frac{c}{\omega_{\text{p}}}\sqrt{a_0} \, , \tag{7}$$

where the numerical factor of two was found through particle-in-cell simulations [62]. This expression is particularly useful for finding the correct initial parameters of the accelerator, i.e., for a given laser system, one should first determine the spot size at which the threshold $a_0 \simeq 3$ is reached; this expression can then be used to determine the correct operating plasma density. However, when self-focusing plays an important role, we need an expression for the matched spot size that depends not on the initial laser intensity but rather the value of a_0 that it reaches after self-focusing. We can use the fact that the ratio of the laser power, P_{L}, to the critical power for self-focusing, P_{c}, can be written as

$$\frac{P_{\text{L}}}{P_{\text{c}}} = \frac{1}{32}\frac{\omega_p^2}{c^2}\, a_0^2 w^2 \, , \tag{8}$$

and the fact that $a_0^2 w^2$ is constant during focusing (assuming that self-focusing happens more rapidly than any pulse compression) to eliminate a_0 from Eq. (7). This results in the following expression for the matched spot size:

$$w_{\text{m}} \simeq 2\sqrt{2}\frac{c}{\omega_{\text{p}}}\left(\frac{P_{\text{L}}}{P_{\text{c}}}\right)^{\frac{1}{6}} \, . \tag{9}$$

Note that $2w_{\text{m}} \approx \lambda_{\text{p}}$, as long as P_{L} is not many times greater than P_{c}.

It is interesting to examine how the spot size used in experiments compares with this matched spot size. Figure 5 shows the variation in the initial (i.e., vacuum) laser focal spot size with the operating plasma density in the experiments in the dataset. There is a clear overall trend towards larger initial spots at lower densities (and therefore higher electron beam energies), and the spot sizes used are reasonably close to λ_{p} (shown as a solid blue line in Fig. 5). However, given that the evidence suggests that most of these experiments reached $a_0 \approx 3$ as a result of pulse evolution, it is also interesting to compare the initial spot size with the expected matched spot size for $a_0 \approx 3$ (this is shown as a dashed red line in Fig. 5). Most of the experiments are clearly operating at an initial spot size significantly larger than this matched spot size, and none of the selected experiments operate with a spot size below this. This suggests either that most experiments are operating with too large an initial spot size, wasting accelerator length and laser energy while the laser pulse self-focuses, or that there is some experimental advantage in starting at a spot size larger than the matched spot size and letting pulse evolution happen.
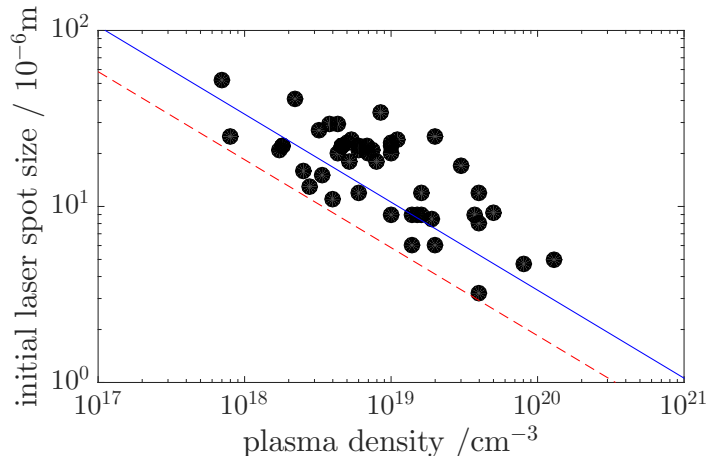
Fig. 5: Variation of initial laser focal spot size with operating plasma density in various laser wakefield acceleration experiments. Solid blue line, plasma wavelength, λ_{p}; red dashed line, matched spot size assuming $a_0 \approx 3$. Data from Refs. [5–9, 15–60].

### 2.3 To guide or not to guide?

It is well known that a tightly focused laser pulse will quickly diffract in a vacuum. Effectively, the pulse can only remain intense over a distance of about a Rayleigh length, $z_{\mathrm{R}} = \pi w_0^2 / \lambda$. The accelerator lengths used in laser wakefield experiments are typically much longer than this and some sort of 'guiding' is therefore needed to keep the laser intensity sufficiently high to drive a wake throughout the structure. There are two principal techniques to achieve this guiding, which both rely on creating a waveguide structure to counteract diffraction. This requires that the transverse plasma density profile has a minimum on-axis. Such a density minimum is naturally created by the laser pulse itself – that is, the bubble itself acts as a waveguide. Alternatively, a preformed density minimum can be formed, for example using a capillary discharge [24]; using a preformed waveguide brings significant complexity to an experiment and restricts diagnostic access (e.g., for wake imaging diagnostics based on ultrashort probes [11]). However, the bubble cannot self-guide the very front slice of the laser pulse, since the density minimum is not formed immediately. Because of this, external channels are expected to be more efficient, i.e., they allow the laser to propagate over greater distances at high a_0. With these points in mind, it is interesting to see whether the experimental evidence supports the use of external waveguides. In Fig. 6, the data are sorted into self-guided and externally guided experiments. The top panel of Fig. 6 appears to show that there is no real advantage to using external guiding structures; the electron beam energy, as a function of plasma density, follows the same trend for both subsets of the data, i.e., they are both limited by dephasing. However, the bottom panel of Fig. 6 reveals the distinct advantage that experiments in externally guided structures have over self-guided ones. The highest electron energy achieved at a given laser power is almost always from an externally guided experiment; the self-guided experiments at the same power tend to produce lower energy electron beams.

   Are these data sufficient to suggest that external guiding structures improve the efficiency of laser wakefield accelerators, owing to reduced energy losses as the pulse propagates? Or is something more subtle occurring? In one experiment with an external waveguide using a $\simeq$20 TW laser [30], differences between the plasma density measured during a high-intensity laser shot and plasma density measurements made offline suggested that, under the conditions in which electron beams were observed, the laser pulse caused additional ionization of the plasma. As the plasma used in that experiment was formed from hydrogen, which is readily ionized, this suggested that high-Z impurities from the walls of the capillary were being ionized by the main pulse and injected into the plasma wave. Since that work, a
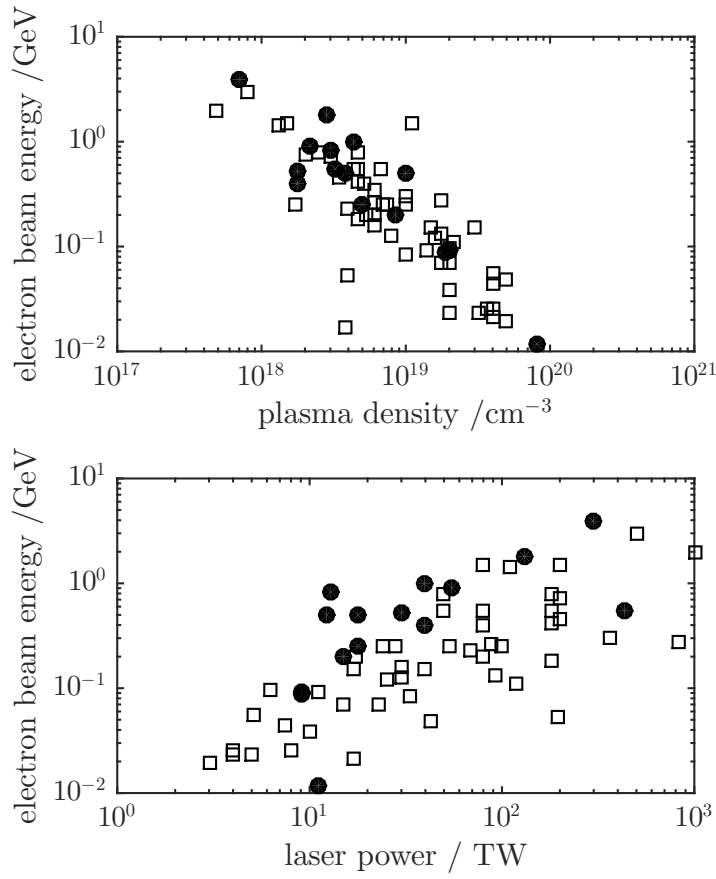
Fig. 6: Variation of reported electron beam energy from various experiments. Top: Variation of beam energy with plasma density. Bottom: Variation of beam energy with laser power. Filled circles, experiments in preformed guiding structures; open squares, experiments without preformed guiding structures. Data from Refs. [5–9, 15–60].

number of groups have proceeded to exploit this ionization injection mechanism to reduce the threshold for injection and increase the beam charge (e.g., Refs. [36, 39]). There are also various other injection techniques, including injection due to propagation in a density gradient [9, 35] and colliding pulse injection [8]. In Fig. 7, the variation of electron beam energy with laser power is plotted again, this time with the data divided into subsets based on injection type. As it is not known whether the majority of capillary-discharge-based experiments rely on self-injection or whether ionization injection plays a role, as in Ref. [30], these experiments have been placed in their own subset. Some of the ionization injection experiments also produce higher electron beam energies for a given laser power than self-injection experiments, and lie on the upper curve of the entire dataset, just as the capillary discharge dataset does. This might suggest that the injection mechanism is more important than the guiding mechanism in determining the maximum energy that can be achieved from a given laser power. This makes sense, as the value that a_0 reaches after self-evolution decreases with decreasing laser power, but alternative injection mechanisms should lower the value that a_0 needs to reach in order for injection to occur. However, injection mechanisms other than ionization injection seem to perform similarly to self-injection in terms of the electron energy that can be obtained for a given laser power. At present, the evidence is still inconclusive; this is clearly a matter that requires further study.
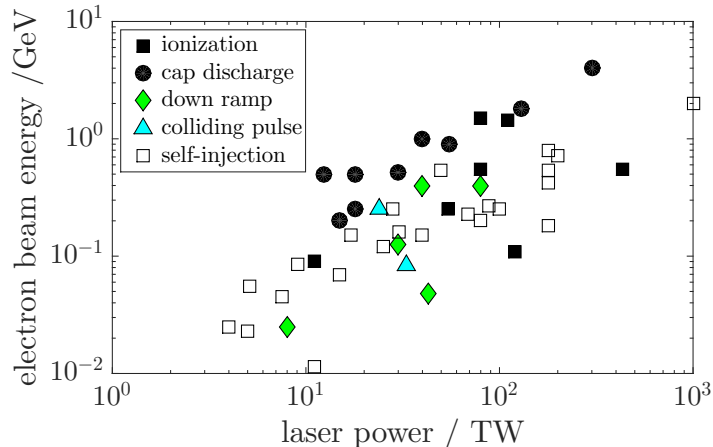
Fig. 7: Variation of reported electron beam energy from various experiments as a function of laser power and for different injection mechanisms. Filled black circles, capillary discharge experiments; black squares, ionization injection; green diamonds, density down-ramp injection; cyan triangles, colliding pulse injection; open black squares, self-injection. Data from Refs. [5–9, 15–60].

3 Future directions

Having examined the trends in experimental data from laser wakefield experiments over the last decade, we find that there is one overriding message. To keep pushing the electron energy achievable from a single stage of a laser wakefield to ever higher values clearly requires operation at lower densities and over longer distances, and such experiments will need more powerful lasers. However, it should be noted that laser power can be increased in two ways, either by increasing the laser energy or by decreasing the pulse duration. Most of the data presented were obtained using laser pulses with a duration of $\sim$30 fs, so the trends in laser power are dominated by the pulse energy. Simulations demonstrate that the pulse length in a laser wakefield accelerator should not be too short. Energy is lost predominantly from the front of the laser pulse (this is the part of the laser that is doing work on the plasma), so the shorter the laser pulse, the more quickly it runs out of energy. As new laser facilities come online around the world, it is worth remembering that the shortest pulse is not necessarily the best for driving laser wakefield accelerators. Pulse durations where $c\tau_L$ is comparable to the plasma wavelength, λ_p, drive efficient wakes and allow the pump depletion length to match the dephasing length [62]. Pulses significantly shorter than $\lambda_p/2$ expend their energy before the maximum energy can be reached.

The experimental trends show that most experiments are attaining electron beam energies that are limited by dephasing. If the community is to be able to keep up the impressive pace it has sustained over the last decade then methods to overcome this limitation will become increasingly important, especially as further increases in laser power become ever more expensive. As a result, techniques to overcome dephasing, such as quasi-phase matching [64], staging [65] and density tapering [66] will all become important areas of research.

The final point that we would like to make is that the laser systems currently used to drive laser wakefield accelerators are woefully inefficient. A titanium sapphire laser has a 'wall plug' efficiency of $\sim$0.1%. They also currently run at very low repetition rates compared with a conventional accelerator. As laser wakefield accelerators push the energy frontier and strive to become workhorses for applications, it will become increasingly important that both repetition rate and efficiency are properly considered and significantly improved. Much more efficient laser architectures are available, including thin-disc [67] and fibre [68] lasers, that can also readily operate at much higher repetition rates. However, these systems do not yet have the capability to produce laser pulses with the high energy and short pulse duration

needed to drive current laser wakefield acceleration experiments. Innovative solutions, including the use of coherent [69] and incoherent [70] combinations of many low-energy laser pulses to drive wakefield experiments, or the use of trains of many low-energy pulses to resonantly drive a high amplitude wakefield [71], may well form the most promising routes to high repetition rate, high-efficiency laser wakefield accelerators suitable for particle physics experiments or light-source based applications.

References

[1] T. Tajima and J.M. Dawson, *Phys. Rev. Lett.* **43**(4) (1979) 267. http://dx.doi.org/10.1103/PhysRevLett.43.267

[2] D. Strickland and G. Mourou, *Opt. Commun.* **55**(6) (1985), 447–449. http://dx.doi.org/10.1016/0030-4018(85)90151-8

[3] K. Nakajima *et al.*, *Phys. Scr.* **T52** (1994) 61. http://dx.doi.org/10.1088/0031-8949/1994/T52/009

[4] F. Amiranoff *et al.*, *Phys. Rev. Lett.* **81**(5) (1998) 995. http://dx.doi.org/10.1103/PhysRevLett.81.995

[5] S.P.D. Mangles *et al.*, *Nature* **431** (2004) 535–538. http://dx.doi.org/10.1038/nature02939

[6] C.G.R. Geddes *et al.*, *Nature* **431** (2004) 538–541. http://dx.doi.org/10.1038/nature02900

[7] J. Faure *et al.*, *Nature* **431** (2004) 541–544. http://dx.doi.org/10.1038/nature02963

[8] J. Faure *et al.*, *Nature* **444** (2006) 737–739. http://dx.doi.org/10.1038/nature05393

[9] A.J. Gonsalves *et al.*, *Nat. Phys.* **7** (2011) 862–866. http://dx.doi.org/10.1038/nphys2071

[10] N.H. Matlis *et al.*, *Nat. Phys.* **2** (2006) 749–753. http://dx.doi.org/10.1038/nphys442

[11] A. Sävert *et al.*, *Phys. Rev. Lett.* **115**(5) (2015) 055002. http://dx.doi.org/10.1103/PhysRevLett.115.055002

[12] A. Rousse *et al.*, *Phys. Rev. Lett.* **93**(13) (2004) 135005. http://dx.doi.org/10.1103/PhysRevLett.93.135005

[13] S. Kneip *et al.*, *Nat. Phys.* **6** (2010) 980–983. http://dx.doi.org/10.1038/nphys1789

[14] M. Fuchs *et al.*, *Nat. Phys.* **5** (2009) 826–829. http://dx.doi.org/10.1038/nphys1404

[15] E. Miura *et al.*, *Appl. Phys. Lett.* **86**(25) (2005) 251501. http://dx.doi.org/10.1063/1.1949289

[16] H. Kotaki *et al.*, *Laser Phys.* **16**(7) (2006) 1107–1110. http://dx.doi.org/10.1134/S1054660X06070140

[17] M. Mori *et al.*, *Phys. Lett.* A **356**(2) (2006) 146–151. http://dx.doi.org/10.1016/j.physleta.2006.06.001

[18] S. Masuda *et al.*, *J. Phys. IV France* **133** (2006) 1127–1129. http://dx.doi.org/10.1051/jp4:2006133229

[19] C.-T. Hsieh *et al.*, *Phys. Rev. Lett.* **96**(9) (2006) 095001. http://dx.doi.org/10.1103/PhysRevLett.96.095001

[20] B. Hidding *et al.*, *Phys. Rev. Lett.* **96**(10) (2006) 105004. http://dx.doi.org/10.1103/PhysRevLett.96.105004

[21] T. Hosokai *et al.*, *Phys. Rev.* E **73**(3) (2006) 036407. http://dx.doi.org/10.1103/PhysRevE.73.036407

[22] S.P.D. Mangles *et al.*, *Phys. Rev. Lett.* **96**(21) (2006) 215001. http://dx.doi.org/10.1103/PhysRevLett.96.215001

[23] S.A. Reed *et al.*, *Appl. Phys. Lett.* **89**(23) (2006) 231107. http://dx.doi.org/10.1063/1.2400400

[24] W.P. Leemans *et al.*, *Nat. Phys.* **2** (2006) 696–699. http://dx.doi.org/10.1038/nphys418

[25] S. Masuda *et al.*, *Phys. Plasmas* **14**(2) (2007) 023103. http://dx.doi.org/10.1063/1.2434248

[26] T. Ohkubo *et al.*, *Phys. Rev. ST Accel. Beams* **10**(3) (2007) 031301. http://dx.doi.org/10.1103/PhysRevSTAB.10.031301

[27] S. Karsch *et al.*, *New J. Phys.* **9** (2007) 415. http://dx.doi.org/10.1088/1367-2630/9/11/415

[28] S.P.D. Mangles *et al.*, *Phys. Plasmas* **14**(5) (2007) 056702. http://dx.doi.org/10.1063/1.2436481

[29] A. Gamucci *et al.*, *IEEE Trans. Plasma Sci.* **36**(4) (2008) 1699–1706. http://dx.doi.org/10.1109/TPS.2008.2000898

[30] T.P. Rowlands-Rees *et al.*, *Phys. Rev. Lett.* **100**(10) (2008) 105005. http://dx.doi.org/10.1103/PhysRevLett.100.105005

[31] N.A.M. Hafz *et al.*, *Nat. Photonics* **2** (2008) 571. http://dx.doi.org/10.1038/nphoton.2008.155

[32] K. Schmid *et al.*, *Phys. Rev. Lett.* **102**(12) (2009) 124801. http://dx.doi.org/10.1103/PhysRevLett.102.124801

[33] S. Kneip *et al.*, *Phys. Rev. Lett.* **103**(3) (2009) 035002. http://dx.doi.org/10.1103/PhysRevLett.103.035002

[34] D.H. Froula *et al.*, *Phys. Rev. Lett.* **103**(21) (2009) 215006. http://dx.doi.org/10.1103/PhysRevLett.103.215006

[35] K. Schmid *et al.*, *Phys. Rev. ST Accel. Beams* **13**(9) (2010) 091301. http://dx.doi.org/10.1103/PhysRevSTAB.13.091301

[36] A.E. Pak *et al.*, *Phys. Rev. Lett.* **104**(2) (2010) 025003. http://dx.doi.org/10.1103/PhysRevLett.104.025003

[37] T.P.A. Ibbotson *et al.*, *New J. Phys.* **12** (2010) 45008. http://dx.doi.org/10.1088/1367-2630/12/4/045008

[38] C.E. Clayton *et al.*, *Phys. Rev. Lett.* **105**(10) (2010) 105003. http://dx.doi.org/10.1103/PhysRevLett.105.105003

[39] C. McGuffey *et al.*, *Phys. Rev. Lett.* **104**(2) (2010) 025004. http://dx.doi.org/10.1103/PhysRevLett.104.025004

[40] H. Lu *et al.*, *Appl. Phys. Lett.* **99**(9) (2011) 091502. http://dx.doi.org/10.1063/1.3626042

[41] S. Fourmaux *et al.*, *New J. Phys.* **13** (2011) 033017. http://dx.doi.org/10.1088/1367-2630/13/3/033017

[42] J.S. Liu *et al.*, *Phys. Rev. Lett.* **107**(3) (2011) 035001. http://dx.doi.org/10.1103/PhysRevLett.107.035001

[43] B.B. Pollock *et al.*, *Phys. Rev. Lett.* **107**(4) (2011) 045001. http://dx.doi.org/10.1103/PhysRevLett.107.045001

[44] O. Lundh *et al.*, *Nat. Phys.* **7** (2011) 219–222. http://dx.doi.org/10.1038/nphys1872

[45] P. Brijesh *et al.*, *Phys. Plasmas* **19**(6) (2012) 063104. http://dx.doi.org/10.1063/1.4725421

[46] M.Z. Mo *et al.*, *Appl. Phys. Lett.* **100**(7) (2012) 074101. http://dx.doi.org/10.1063/1.3685464

[47] S. Kneip *et al.*, *Phys. Rev. ST Accel. Beams* **15**(2) (2012) 021302. http://dx.doi.org/10.1103/PhysRevSTAB.15.021302

[48] R. Weingartner *et al.*, *Phys. Rev. ST Accel. Beams* **15**(11) (2012) 111302. http://dx.doi.org/10.1103/PhysRevSTAB.15.111302

[49] F. Albert *et al.*, *Phys. Rev. Lett.* **111**(23) (2013) 235004. http://dx.doi.org/10.1103/PhysRevLett.111.235004

[50] P.A. Walker *et al.*, *New J. Phys.* **15** (2013) 045024. http://dx.doi.org/10.1088/1367-2630/15/4/045024

[51] M.Z. Mo *et al.*, *Appl. Phys. Lett.* **102**(13) (2013) 134102. http://dx.doi.org/10.1063/1.4799280

[52] S. Corde *et al.*, *Nat. Commun.* **4** (1501) (2013) 1309.6364v1. http://dx.doi.org/10.1038/ncomms2528

[53] S. Chen *et al.*, *Phys. Rev. Lett.* **110**(15) (2013) 155003. http://dx.doi.org/10.1103/PhysRevLett.110.155003

[54] H.T. Kim *et al.*, *Phys. Rev. Lett.* **111**(16) (2013) 165002.
http://dx.doi.org/10.1103/PhysRevLett.111.165002

[55] X. Wang *et al.*, *Nat. Commun.* **4**(1988) (2013). http://dx.doi.org/10.1038/ncomms2988

[56] G. Sarri *et al.*, *Phys. Rev. Lett.* **113** (2014) 224801.
http://dx.doi.org/10.1103/PhysRevLett.113.224801

[57] N.D. Powers *et al.*, *Nat. Photonics* **8** (2014) 28–31. http://dx.doi.org/10.1038/nphoton.2013.314

[58] W.P. Leemans *et al.*, *Phys. Rev. Lett.* **113**(24) (2014) 245002.
http://dx.doi.org/10.1103/PhysRevLett.113.245002

[59] K. Khrennikov *et al.*, *Phys. Rev. Lett.* **114**(19) (2015) 195003.
http://dx.doi.org/10.1103/PhysRevLett.114.195003

[60] M. Schnell *et al.*, *J. Plasma Phys.* **81**(04) (2015) 475810401.
http://dx.doi.org/10.1017/S0022377815000379

[61] V. Malka *et al.*, *Science* **298** (5598) (2002) 1596–1600. http://dx.doi.org/10.1126/science.1076782

[62] W. Lu *et al.*, *Phys. Rev. ST Accel. Beams* **10**(6) (2007) 061301.
http://dx.doi.org/10.1103/PhysRevSTAB.10.061301

[63] W.B. Mori, *IEEE J. Quant. Electron.* **33**(11) (1997) 1942–1953.
http://dx.doi.org/10.1109/3.641309

[64] S.J. Yoon *et al.*, *Phys. Rev. ST Accel. Beams* **15**(8) (2012) 081305.
http://dx.doi.org/10.1103/PhysRevSTAB.15.081305

[65] D. Kaganovich *et al.*, *Phys. Plasmas* **12**(10) (2005) 100702. http://dx.doi.org/10.1063/1.2102727

[66] P. Sprangle *et al.*, *Phys. Rev.* E **63**(5) (2001) 056405.
http://dx.doi.org/10.1103/PhysRevE.63.056405

[67] A. Giesen and J. Speiser, *IEEE J. Sel. Top. Quant. Electron.* **13**(3) (2007) 598.
http://dx.doi.org/10.1109/JSTQE.2007.897180

[68] C. Jauregui, J. Limpert and A. Tünnermann, *Nat. Photonics* **7** (2013) 861–867.
http://dx.doi.org/10.1038/nphoton.2013.273

[69] G. Mourou *et al.*, *Nat. Photonics* **7** (2013) 258–261. http://dx.doi.org/10.1038/nphoton.2013.75

[70] C. Benedetti, *et al.*, *Phys. Plasmas* **21**(5) (2014) 056706. http://dx.doi.org/10.1063/1.4878620

[71] S.M. Hooker *et al.*, *J. Phys. B: At. Mol. Opt. Phys* **47**(23) (2014) 234003.
http://dx.doi.org/10.1088/0953-4075/47/23/234003

Case Studies on Plasma Wakefield Accelerator Design

J. Osterhoff[1], Z. Najmudin[2] and J. Faure[3]
[1] Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany
[2] John Adams Institute for Accelerator Science, Imperial College London, UK
[3] LOA, École Polytechnique-ENSTA-CNRS, Palaiseau, France

Abstract
The field of plasma-based particle accelerators has seen tremendous progress over the past decade and experienced significant growth in the number of activities. During this process, the involved scientific community has expanded from traditional university-based research and is now encompassing many large research laboratories worldwide, such as BNL, CERN, DESY, KEK, LBNL and SLAC. As a consequence, there is a strong demand for a consolidated effort in education at the intersection of accelerator, laser and plasma physics. The CERN Accelerator School on Plasma Wake Acceleration has been organized as a result of this development. In this paper, we describe the interactive component of this one-week school, which consisted of three case studies to be solved in 11 working groups by the participants of the CERN Accelerator School.

Keywords
Plasma wakefield accelerators; applications; case study.

1   Introduction

The CERN Accelerator School on Plasma Wake Acceleration was held in Geneva at CERN from November 23 to 29, with the aim of teaching students and postdoctoral researchers the basic concepts of plasma wakefield acceleration at the intersection of accelerator, laser and plasma physics. An important objective of this school was to include an educational component that deviated in form from the lectures given by experts and fostered interaction between participants.

Therefore, three case studies were developed, pertaining to future applications, which required the design of wakefield accelerators for specific purposes. The students were confronted with problems that left a considerable amount of freedom in approaching their solutions. Intentionally, this demanded self-organization of the students within their working group (there were 11 such groups), a break-up of these cases into subproblems and a distribution of tasks within each group. The students had 4 days, in parallel with the lectures, to work out possible case study solutions and prepare one short presentation per group, which were delivered at the end of the school to the students and lecturers.

This paper summarizes the case study activities. The cases are presented in Section 2. Section 3 discusses the educational objectives of this undertaking, Section 4 describes its organization and Section 5 details the outcome and highlights some specific results.

2   Case study problems

2.1   Case study 1: A plasma-based booster module for the International Linear Collider

2.1.1   Introduction

Today the search for new particles and forces at energies of hundreds or thousands of gigaelectronvolts plays a central role in the field of elementary particle physics. Particle physicists have established a Standard Model for the strong, weak and electromagnetic interactions that passes tests at both low and high energies. The model is extremely successful, and yet it is incomplete in many important respects. New particles and interactions are needed to fill the gaps.

Some of the difficulties of the Standard Model are deep and abstract; their explanations may be found only in the distant future. The Standard Model does not explain how gravity is connected to the other forces of nature. It does not explain why the basic particles of matter are the quarks and leptons, or how many of these there should be.

However, the Standard Model also fails to explain three phenomena that, by rights, should be accounted for at the energies now being probed with particle accelerators, e.g., at the LHC at CERN. Astronomers believe that the dominant form of matter in the universe is a neutral weakly interacting species, called dark matter that cannot be composed of any particle present in the Standard Model. The Standard Model cannot explain why the universe contains atomic matter made of electrons, protons and neutrons but no comparable amount of antimatter.

The problem of the Higgs field is likely to be connected to these questions about the matter content of the universe. Explanatory models of the Higgs field often contain particles with the correct properties to make up the dark matter. There are also strong, independent, arguments that the mass of the dark-matter particle is comparable to the masses—of the order of 100 GeV—of the heaviest particles that receive mass from the Higgs field. The predominance of baryons over antibaryons in the universe could arise from interactions among Higgs fields that violate space-time charge-parity symmetry. More generally, any model of fundamental physics at energies above 100 GeV must contain the Higgs field or some generalization and must account for the place of this field within its structure.

A way to prove the existence of the Higgs field and to study its interactions is to find and study the quantum of this field, the Higgs boson. The International Linear Collider (ILC) was designed to study this particle and other new particles that might be associated with it. It provides an ideal setting for detailed exploration of the origin and nature of the Higgs field. In July 2012, the ATLAS and CMS experiments at the CERN Large Hadron Collider announced the discovery of a new particle with a mass of 125 GeV and many properties of the Higgs boson, as postulated in the Standard Model. The ILC will enable the properties of this Higgs boson to be studied in much greater detail than previously possible. (this section was reproduced, with permission, from the *International Linear Collider Technical Design Report* [1].)

2.1.2   Task

Assume that the ILC (see Fig. 1) is a 250 GeV centre-of-mass high-luminosity linear electron–positron collider built specifically for the study of the Higgs field, based on 1.3 GHz superconducting radio frequency (RF) accelerating technology. The initial programme of the ILC, to discover a 125 GeV Higgs boson, H^0, at this energy, will give the peak cross-section for the reaction $e^+ + e^- \rightarrow Z + H^0$. In this reaction, the identification of a Z boson at the energy appropriate to recoil against the Higgs boson tags the presence of the Higgs boson. At higher energies, other reactions could be studied (see Table 1), which are not accessible in the current design. Wakefield accelerators could offer a cost-efficient path to boost the energy of the ILC and therefore allow for detailed studies of these additional processes.
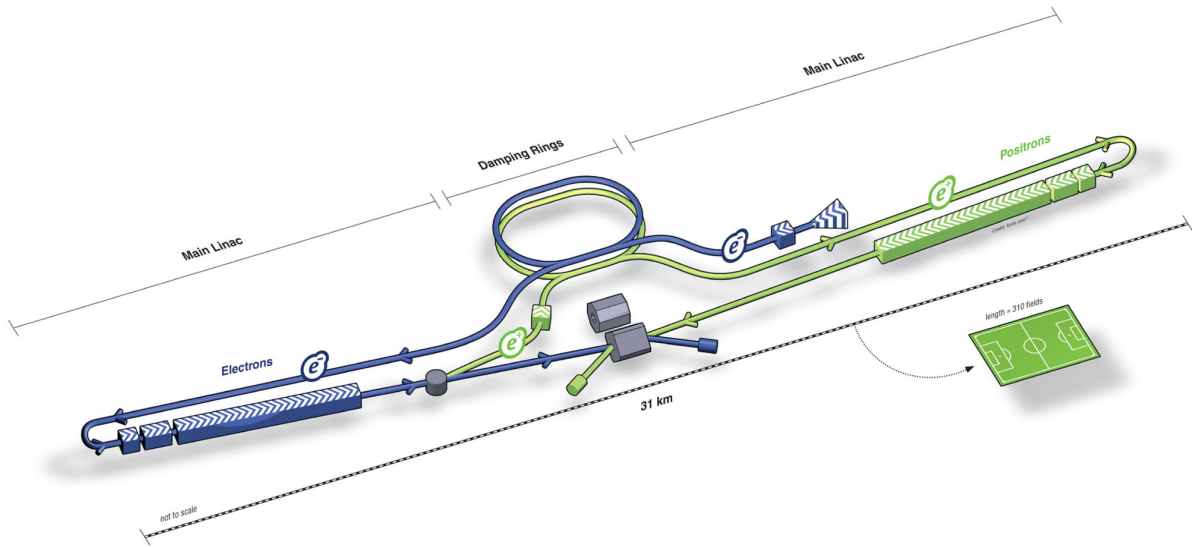
Fig. 1: Basic design of the International Linear Collider (taken from Ref. [1])

Table 1: Reactions of the Higgs field and other events at the energy scale of interest (cf. Ref. [1])

Energy	Reaction	Physics Goal
91 GeV	$e^+e^- \to Z$	ultra-precision electroweak
160 GeV	$e^+e^- \to WW$	ultra-precision W mass
250 GeV	$e^+e^- \to Zh$	precision Higgs couplings
350–400 GeV	$e^+e^- \to t\bar{t}$	top quark mass and couplings
	$e^+e^- \to WW$	precision W couplings
	$e^+e^- \to \nu\bar{\nu}h$	precision Higgs couplings
500 GeV	$e^+e^- \to f\bar{f}$	precision search for Z'
	$e^+e^- \to t\bar{t}h$	Higgs coupling to top
	$e^+e^- \to Zhh$	Higgs self-coupling
	$e^+e^- \to \tilde{\chi}\tilde{\chi}$	search for supersymmetry
	$e^+e^- \to AH, H^+H^-$	search for extended Higgs states
700–1000 GeV	$e^+e^- \to \nu\bar{\nu}hh$	Higgs self-coupling
	$e^+e^- \to \nu\bar{\nu}VV$	composite Higgs sector
	$e^+e^- \to \nu\bar{\nu}t\bar{t}$	composite Higgs and top
	$e^+e^- \to \tilde{t}\tilde{t}^*$	search for supersymmetry

Your task is to design a plasma-based energy-booster section for the ILC. The required beam power currently excludes lasers as wakefield drivers. To achieve a realistic design, consider and motivate a beam-driven approach for both the electron and the positron arms of the ILC. Please be careful to conserve the beam properties required in the ILC accelerator section to sustain a suitable event rate (see Table 2).

Table 2: Beam parameters of the ILC accelerator (cf. [1])

			1st Stage
Centre-of-mass energy	E_{CM}	GeV	250
Collision rate	f_{rep}	Hz	5
Electron linac rate	f_{linac}	Hz	10
Number of bunches	n_{b}		1312
Bunch population	N	$\times 10^{10}$	2.0
Bunch separation	Δt_{b}	ns	554
Pulse current	I_{beam}	mA	5.8
Main linac average gradient	G_{a}	MV m^{-1}	31.5
Average total beam power	P_{beam}	MW	5.9
Estimated AC power	P_{AC}	MW	129
RMS bunch length	σ_{z}	mm	0.3
Electron RMS energy spread	$\Delta p/p$	%	0.190
Positron RMS energy spread	$\Delta p/p$	%	0.152
Electron polarisation	P_{-}	%	80
Positron polarisation	P_{+}	%	30
Horizontal emittance	$\gamma \epsilon_{\mathrm{x}}$	μm	10
Vertical emittance	$\gamma \epsilon_{\mathrm{y}}$	nm	35
IP horizontal beta function	β_{x}^{*}	mm	13.0
IP vertical beta function	β_{y}^{*}	mm	0.41
IP RMS horizontal beam size	σ_{x}^{*}	nm	729
IP RMS veritcal beam size	σ_{y}^{*}	nm	7.7
Luminosity	L	$\times 10^{34}$ cm^{-2}s^{-1}	0.75
Fraction of luminosity in top 1%	$L_{0.01}/L$		87.1%
Average energy loss	δ_{BS}		0.97%
Number of pairs per bunch crossing	N_{pairs}	$\times 10^{3}$	62.4
Total pair energy per bunch crossing	E_{pairs}	TeV	46.5

2.2 Case study 2: An X-ray free-electron laser based on laser wakefield acceleration

2.2.1 Introduction

At gigaelectronvolt energies, the radiation losses of electrons travelling in strong (usually magnetic) fields becomes appreciable. The radiation can easily extend to the X-ray range, as well as being naturally collimated and ultrashort (of the order of femtoseconds). By employing an undulator, a periodically poled magnetic field configuration, this synchrotron radiation can be reinforced in the direction of travel of the electron beam. If the radiated energy is sufficiently bright, it can cause bunching of the electrons, causing them to emit radiation coherently and thus even more strongly. This feedback causes exponential growth of the radiated signal in a device called a free-electron laser (FEL) [2].

To date, two FELs operating at X-ray wavelengths have been demonstrated, the Linac Coherent Light Source at SLAC [3] and SACLA at RIKEN [4], and they are soon to be joined by a handful of others, including the European XFEL at DESY. The enormous increase in the brightness of X-ray sources at these facilities has created dramatic new scientific possibilities, such as the possibility of determining the structure of complex molecules in a single shot, and the ability to generate unique far-from-equilibrium states of ionized matter. However, these facilities are few, and will remain so,

primarily because of the exacting requirements on the electron beam to seed an X-ray FEL. The electron beam must have, simultaneously, high energy, high bunch charge, low energy spread and low emittance. This can be seen by the electron beam properties of the existing FEL labs given in Table 3.

Table 3: Main parameters of operational and planned X-ray FEL facilities

Name	Where	E_{elec} (GeV)	$\Delta E/E_{elec}$	Charge (pC)	ε_n (µm)	E_{phot} (keV)
LCLS	SLAC (USA)	14	$\ll 1\%$	250	1	8
SACLA	J-PARC (Japan)	8.5	$\ll 1\%$	250	0.8	10
XFEL	DESY (Germany)	17.5	$\ll 1\%$	250	1.4	10

Obviously wakefield accelerators could produce high charge beams in a short distance. They are an obvious candidate for generating an electron beam with sufficient energy to drive an X-ray FEL at reduced cost. However, the other parameters required may be harder to produce, and so will require some consideration.

2.2.2   Task

In this case study, we will try to design an accelerator to drive an X-ray FEL with a laser wakefield scheme. The first stage will be to determine the plasma characteristics required to achieve the energies needed for the wakefield accelerator.

Once this has been achieved, it is necessary to consider whether this accelerator will be able to provide the necessary charge, with sufficient beam quality. The requirements of the laser and plasma to be able to create this source should be determined. A survey of the literature would allow us to determine how close we are with present experiments and laser systems. Could other wakefield techniques be used to achieve the required characteristics? Are there advances in FEL research that could help us?

2.3   Case study 3: A high repetition-rate laser-plasma accelerator for electron diffraction

2.3.1   Introduction

One of the current challenges of science is to unravel the dynamical structure of nature on the atomic time-scale; for example, understanding atomic motion during a chemical reaction, or following electron dynamics in a protein after photoexcitation. In the past decade, pulsed sources of X-rays have become available and provide a probe for atomic motion in complex matter at the femtosecond level. For example, the Linac Coherent Light Source at Stanford has reached a high level of performance with the production of unprecedented bright pulses of ultrashort coherent X-rays. In spite of their remarkable capabilities, these sources, based on the use of large electron accelerators, are restricted to a few large facilities with limited beam time access.

In parallel, a less expensive tabletop alternative has been developed, which uses ultrashort electron bunches to probe atomic motion and structural dynamics, either through diffraction [5] or microscopy. In addition to their small size, these electron sources provide another significant advantage in that they offer an elastic scattering cross-section, which is more than five orders of magnitude higher than for X-rays, providing much higher diffraction efficiency. Ultrashort electron bunches have been used successfully for the study of structural dynamics in condensed matter, chemistry and biology. However, the time resolution currently does not exceed 200 fs, preventing the observation of faster phenomena.

Typical electron guns for electron diffraction are d.c. (static acceleration field) photoguns (based on the use of a UV-driven photocathode as the source of electrons) and operate at around 100 keV. Recently, RF guns are being developed, to decrease the pulse duration to below 100 fs. These RF guns

operate at a few megaelectronvolts but they are subject to a timing jitter between the laser and the electron pulses. Consequently, reaching <100 fs or even <10 fs is very difficult, even though interesting dynamics occur on these time-scales. To summarize, the current limitations in ultrafast electron diffraction are:

- space charge of the beam, limiting the bunch duration to >100 fs and the charge to femtocoulomb levels (for static acceleration fields);

- synchronization (jitter) between the excitation laser and the electron bunch (for RF fields), limiting the resolution to >100 fs.

Laser-plasma accelerators offer an opportunity to provide even shorter electron pulses for two main reasons: (i) the huge accelerating fields (10–100 GV/m) prevent space charge and (ii) in principle, the electron bunches are jitter-free because they originate from a laser-driven accelerating structure. In this case study, we will examine the possibility of designing a laser-plasma accelerator with the right parameters (see Table 4) and offering <10 fs electron bunches.

Table 4: Main parameters of typical electron guns used for electron diffraction. The bottom row lists the parameters that the design should obtain.

Accelerator	E_{elec} (MeV)	$\Delta E/E_{elec}$	Charge (fC)	ε_n (µm)	Bunch duration	Rep. rate
DC photogun [5]	0.1	<1%	1	5×10^{-2}	>200 fs	1 kHz
RF photogun [6]	3.5	<1%	100	7×10^{-2}	>100 fs	10 Hz
Laser-plasma	0.1-10	<1%	1-1000	<0.1	<10 fs	>0.1 kHz

2.3.2   Task

In this case study, we will try to design an accelerator that could be used for electron diffraction applications. The target parameters of the design are quite exotic for a laser-plasma accelerator: note, for example, that for the energy should be quite low (<10 MeV) and that it is preferable that the repetition rate be large (>100 Hz; a kilohertz repetition rate is preferable for good statistics in electron diffraction experiments). The beam quality should be very good, for making clear and usable diffraction patterns; in particular, the transverse emittance needs to be low.

A first step in designing this accelerator would be to determine the necessary plasma parameters, followed by the necessary laser parameters. A survey of the literature would allow us to determine how close we are with present experiments and laser systems (note that first experiments at high repetition rate have begun to produce some results [7]). We should then think about what injection techniques could be used to produce the electron beam. We should also discuss what kind of beam quality we could obtain and how close it is from the target parameters, and suggest ways to improve things. Considerations on bunch duration and elongation upon propagation would also be helpful.

3   Educational objectives

The main educational purpose of these case studies was to familiarize the students with advantages, problems and principles of plasma wakefield acceleration. Deliberately, the chosen scenarios were focused on actual visions for applications of plasma-based accelerators, which are currently pursued and studied by a number of research groups worldwide, but remain unsolved in the sense that they have yet to be experimentally realized. Thus, the students were working on cutting-edge problems.

It was intended that these problems should be approached using the knowledge that the students obtained during the lectures at this CERN Accelerator School. Since a number of solution strategies

were applicable to each case, each problem was assigned to three or four different groups of students, who produced a variety of outcomes and approaches.

A significant hurdle in the solution process was the initial strategic discussion and distribution of tasks within each working group. This was intentional. The participants had to interact and self-organize as peers to come up with a successful strategy, strengthening their networking, social interaction and group performance skills. The collective work on a specific problem turned out to be highly motivating for many students. Many of them spent hours following different solution paths, experimenting with different parameters and optimizing their base designs, using every free minute they had between lectures and after dinner, and often working until late at night.

4   Case study organization

On the first day, the three case studies were distributed to three (case study 1) or four different groups (case studies 2 and 3), resulting in 11 working groups altogether. Each group consisted of up to 10 students with backgrounds in accelerator, laser or plasma physics, who worked jointly on the assigned tasks.

The case study work itself was supposed to be done after the end of the lectures in the evenings and during two tutorial sessions (of at least an hour each) on days two and three of the school. For this purpose, three separate seminar rooms were booked with two tutors each. The tutors were selected from the list of lecturers. In these tutorials, students were given the chance to discuss the problems and their solutions with experts.

On the final day of lectures, the students were given an extra hour to prepare their concluding presentations, to be delivered at the end of that day. These presentations were 5 minute breakdowns of the achieved results and suggested designs and were presented to students and lecturers in the audience and a jury, who selected winners for the categories 'best overall solution', 'most innovative solution' and 'most entertaining presentation'.

5   Outcome, highlights and conclusions

The extremely active, productive and creative atmosphere during the case study work that was prevalent during the week resulted in a number of highly interesting solutions and innovative schemes suggested by the different working groups, making the case study programme a great success. In the following, a number of these proposals and highlights are discussed for each case study.

5.1   Case study 1: A plasma-based booster module for the International Linear Collider

The three groups working on this case proposed various design solutions at different plasma densities from 10^{15} cm^{-3} to 10^{16} cm^{-3}. All the groups suggested reshaping the ILC bunch train structure, such that odd and even numbered bunches from the train could be utilized as drivers and witnesses, respectively, in a beam-driven plasma wakefield accelerator set-up. The main issue, which was correctly identified, is the conservation of beam properties in the plasma booster stage, in particular in the positron arm of the ILC. The three proposals differed significantly, proposing the use of hollow-core plasma channels [8–10], plasma acceleration in the linear regime and even the use of plasmas consisting of antihydrogen.

5.2   Case study 2: An X-ray free-electron laser based on laser wakefield acceleration

This case study was considered by four groups, who came up with different design ideas, in particular with respect to the utilized undulator technology. The proposed methods covered cryogenic undulators [11], transverse gradient undulators [12], synchronized seeding from solid-target high harmonic sources [13] and travelling wave Thomson scattering [14]. Beam transport was also taken into account, with an

emphasis on beam emittance preservation and the inclusion of a plasma-based beam dechirper. Proposed electron injection techniques within the plasma target ranged from ionization injection [15, 16] to counter-propagating pulse injection [17].

5.3 Case study 3: A high repetition-rate laser-plasma accelerator for electron diffraction

The four working groups considered the creation of suitable low-emittance and low-energy-spread beams from a plasma to be challenging and therefore presented solutions using various controlled injection methods and a sophisticated electron beam transport system. Electron beams were generated by colliding few-cycle laser pulses or by shock-front injection [18]. In some cases, the transport system also included a post-plasma energy filter for the generation of well-defined narrow-bandwidth spectra.

References

[1] https://www.linearcollider.org/ILC/Publications/Technical-Design-Report

[2] B.W.J. McNeil and N.R. Thompson, *Nat. Photonics* **4** (2010) 814. http://dx.doi.org/10.1038/nphoton.2010.239

[3] P. Emma *et al.*, *Nat. Photonics* **4** (2010) 641–647. http://dx.doi.org/10.1038/nphoton.2010.176

[4] T. Ishikawa *et al.*, *Nat. Photonics* **6** (2012) 540–544. http://dx.doi.org/10.1038/nphoton.2012.141

[5] G. Sciaini and D. Miller, *Rep. Prog. Phys.* **74**(9) (2011) 096101. http://dx.doi.org/10.1088/0034-4885/74/9/096101

[6] P. Musumeci *et al.*, *Appl. Phys. Lett.* **97**(6) (2010) 063502. http://dx.doi.org/10.1063/1.3478005

[7] Z. He *et al.*, *Appl. Phys. Lett.* **102**(6) (2013) 064104. http://dx.doi.org/10.1063/1.4792057

[8] T.C. Chiou *et al.*, *Phys. Plasmas* **2**(1) (1995) 310. http://dx.doi.org/10.1063/1.871107

[9] G. Shvets *et al.*, *IEEE Trans. Plasma Sci.* **24**(2) (1996) 351. http://dx.doi.org/10.1109/27.509999

[10] G. Shvets, Ph.D. thesis, Massachusetts Institute of Technology, 1995.

[11] T. Hara *et al.*, *Phys. Rev. ST Accel. Beams*, **7** (2004) 050702. http://dx.doi.org/10.1103/PhysRevSTAB.7.050702

[12] Z. Huang *et al.*, *Phys. Rev. Lett.* **109**(20) (2012) 204801. http://dx.doi.org/10.1103/PhysRevLett.109.204801

[13] D. von der Linde *et al.*, *Phys. Rev.* A **52**(1) (1995) R25(R). http://dx.doi.org/10.1103/PhysRevA.52.R25

[14] Klaus Steiniger *et al.*, *J. Phys.* B **47**(23) (2014) 234011. http://dx.doi.org/10.1088/0953-4075/47/23/234011

[15] C. Mcguffey *et al.*, *Phys. Rev. Lett.* **104**(2) (2010) 025004. http://dx.doi.org/10.1103/PhysRevLett.104.025004

[16] A. Pak *et al.*, *Phys. Rev. Lett.* **104**(2) (2010) 025003. http://dx.doi.org/10.1103/PhysRevLett.104.025003

[17] J. Faure *et al.*, *Nature* **444** (2006) 737. http://dx.doi.org/10.1038/nature05393

[18] A. Buck *et al.*, *Phys. Rev. Lett.* **110**(18) (2013) 185006. http://dx.doi.org/10.1103/PhysRevLett.110.185006

Participants

ADLI, E.	University of Oslo, Oslo, NO
ALEMANY FERNANDEZ, R.	CERN, Geneva, CH
AMORIM, L.D.	GoLP/IPFN-IST, Lisbon, PT
ATAY, M.	John Adams Institute for Accelerator Science, Oxford, GB
BATSCH, F.	Max Planck Institut fur Physik, Munich, DE
BEAUREPAIRE, B.	LOA-ENSTA, Palaiseau, FR
BELLODI, G.	CERN, Geneva, CH
BERA, R.K.	Institute for Plasma Research, Gandhinagar, IN
BERGLYD OLSEN, V.	University of Oslo, Oslo, NO
BERNARDINI, M.	CERN, Geneva, CH
BETTONI, S.	Paul Scherrer Institut, Villigen, CH
BISESTO, F.G.	INFN-LNF, Frascati, IT
BISKUP, B.	CERN, Geneva, CH
BODENDORFER, M.	CERN, Geneva, CH
BOELLA, E.	GoLP/IPFN-IST, Lisbon, PT
BOROZ, K.	Leibniz Universitat Hannover, Hannover, DE
BRACCO, C.	CERN, Geneva, CH
BRYANT, J.	Imperial College London, London, GB
CALATRONI, S.	CERN, Geneva, CH
CARDELLI, F.	Rome University 'La Sapienza', Rome, IT
CASTILLO RAMIREZ , A.J.	Peoples' Friendship University of Russia, Moscow, RU
CHANCE, A.	CEA-Saclay, Gif-Sur-Yvette, FR
COLE, J.	Imperial College London, London, GB
DALE, J.	DESY, Hamburg, DE
DALENA, B.	CEA-Saclay, Gif-Sur-Yvette, FR
DEACON, L.	University College London, London, GB
DING, H.	LMU, Garching Bei Munchen, DE
DOEBERT, S.	CERN, Geneva, CH
DOEPP, A.	Centro de Laseres Pulsados, Villamayor, ES
DORDA, U.	DESY, Hamburg, DE
FEDOSSEEV, V.	CERN, Geneva, CH
FILIPPI, F.	Rome University 'La Sapienza', Rome, IT
GALLARDO GONZALEZ, I.	Lund University, Lund, SE
GAMBA, D.	CERN, Geneva, CH
GESSNER, S.	SLAC, Menlo Park, US
GIRIBONO, A.	Rome University 'La Sapienza', Rome, IT
GIUBEGA, G.	National Institute for Laser, Plasma and Radiation Physics, Bucharest, RO
GOLDBERG, L.	DESY, Hamburg, DE
GORGISYAN, I.	Paul Scherrer Institut, Villigen, CH
GREBENYUK, J.	DESY, Hamburg, DE
GUILLAUME, E.	Laboratoire d'Optique Appliquee, Palaiseau, FR
HANAHOE, K.	University of Manchester, Manchester, GB
HESSLER, C.	CERN, Geneva, CH
HUEBL, A.	Helmholtz-Zentrum Dresden - Rossendorf, Dresden, DE
ISRAELI, Y.	Max Planck Institut fur Physik, Munich, DE
JONES, J.	STFC Daresbury Laboratory, Warrington, UK
KNIAZIEV, R.	V. N. Karazin Kharkiv National University, Kharkiv, UA
KONONENKO, O.	DESY, Hamburg, DE

KOSCHIEZKI, C.	Max-Born Institut, Berlin, DE
KOURKAFAS, D.	DESY Zeuthen, Zeuthen, DE
KUMAR, P.	University of Lucknow, Lucknow, IN
KUSCHEL, S.	Helmholtz-Insitut Jena, Jena, DE
LAZZARINO, L.L.	University of Hamburg, Hamburg, DE
LEE, S.	Gwangi Institute of Science and Technology (GIST), Gwangju, KR
LIBOV, V.	DESY, Hamburg, DE
LISHILIN, O.	DESY Zeuthen, Zeuthen, DE
LUIS MARTINS, J.	GoLP-IPFN, Lisbon, PT
MAITRALLAIN, A.	CEA Saclay, Gif-Sur-Yvette, FR
MANDRY, S.	University College London, London, GB
MARTINEZ DE LA OSSA, A.	DESY, Hamburg, DE
MARTYANOV, M.	CERN, Geneva, CH
MASSIMO, F.	INFN, Rome, IT
MATTEI, S.	EPFL, Lausanne, CH
MINAKOV, V.	Novosibirsk State University, Novosibirsk, RU
MIRANDA, F.	Schlumberger, Houston, US
MOODY, J.	Max Planck Institut fur Physik, Munich, DE
MUKHERJEE, A.	Institute For Plasma Research, Gandhinagar, IN
OGINOV, A.	Lebedev Physical Institute of RAS, Moscow, RU
OLGUN, H.	DESY, Hamburg, DE
PALMER, C.	DESY, Hamburg, DE
PARDONS, A.	CERN, Geneva, CH
PETRENKO, A.	CERN, Geneva, CH
PLASSARD, F.	CERN, Geneva, CH
POCSAI, M.	Wigner Research Centre for Physics, Budapest, HU
PODER, K.	Imperial College London, London, GB
POMPILI, R.	INFN, Frascati, IT
POURMOUSSAVI, P.	DESY, Hamburg, DE
RIEGER, K.	Max Planck Institut fur Physik, Munich, DE
ROMEO, S.	Rome University 'La Sapienza' and INFN LNF, Rome, IT
SADYKOVA, S.	Juelich Research Centre, Juelich, DE
SALVETER, F.	CERN, Geneva, CH
SCHLUETER, F.	Peter Gruenberg Institut (PGI-6), Juelich, DE
SCHMIDT, J.	CERN, Geneva, CH
SCHMITT, N.	Technische Universitat Darmstadt, Darmstadt, DE
SCHWAB, M.	Institute for Optics and Quantum Electronics, Jena, DE
SCISCIO, M.	Rome University 'La Sapienza', Rome, IT
SHAFIR, G.	Technion, Haifa, IL
SHARMA, A.	ELI-ALPS Hungary, Szeged, HU
SHPAKOV, V.	INFN-LNF, Frascati, IT
SHUKLA, N.	GOLP/IPFN-IST, Lisbon, PT
SINGH, D.	Indian Institute of Technology, Delhi, IN
SINIGARDI, S.	Universita degli Studi di Bologna, Bologna, IT
SNEDDEN, E.	STFC Daresbury Laboratory, Warrington, GB
SPITSYN, R.	Novosibirsk State University, Novosibirsk, RU
TECKER, F.	CERN, Geneva, CH
TUDOR, P.	John Adams Institute, Oxford, GB
TUEV, P.	Novosibirsk State University, Novosibirsk, RU
TURNER, M.	CERN, Geneva, CH
ULLMANN, D.	Helmholtzzentrum f. Schwerionenforschung Jena, Jena, DE

VARGA-UMBRICH, K.	Wigner Research Centre for Physics, Budapest, HU
VERNIER, A.	Laboratoire d'Optique Appliquee, Palaiseau, FR
VRANIC, M.	GoLP/IPFN, Lisbon, PT
WACKER, V.	DESY, Hamburg, DE
WEGNER, R.	CERN, Geneva, CH
WEI, Y.	University of Liverpool, Liverpool, GB
WOOD, J.	Imperial College London, London, GB
ZEMELLA, J.	DESY, Hamburg, DE