Towards Very High Energy Accelerators

Jonathan S. Wurtele
Department of Physics
University of California, Berkeley

1 Introduction

The development of a new generation of colliders beyond the LHC is central to the future of high-energy physics. As was clearly pointed out by Lykken [1], the prospects for high-energy physics may be limited by our ability to build them. This paper describes possibilities for these colliders and some of the challenges we face in building them.

None of the fundamental ideas that underly the advanced accelerator work reported herein are new; they have been suggested, and studied in various levels of detail, for the last two decades. What is new, and encouraging, are the technological advances, the level of understanding and detailed modeling, and the initial experimentation. These changes are affecting a wide variety of concepts. For example, W-band structures are being tested at SLAC, a large collaboration, numbering more than 100 individuals, is working on a neutrino factory and muon collider, and laser-driven plasma accelerators have matured to the point where numerous groups worldwide can accelerate nC of charge to MeV energies.

Such progress may lead one to wonder why none of these options offers an unimpeded path to a high-energy collider. A simplified answer is that the desired luminosity (even extrapolating our ability to produce, accelerate, and focus high quality beams) demands high average power in the colliding beams. Overall system efficiency is thus vital. Preserving the beam quality needed for focusing to very small spot sizes requires that great care be paid to critical engineering issues such as the fabrication of structures, focusing fields, and jitter. Accommodating these numerous and severe constraints requires a mature technology. Many of the beam interactions with plasma are complicated and relevant experimental studies are in an early stage.

The expression for the luminosity illustrates the previous comments, and also makes clear the different approaches that the e^+e^- and $\mu^+\mu^-$ colliders take. We can write the luminosity

$$\mathcal{L} = \frac{N^2 f_c}{4\pi\sigma_x\sigma_y} \propto \frac{P_{av} N n_c}{4\pi\sqrt{\beta_x\beta_y\varepsilon_x\varepsilon_y}}, \tag{1}$$

	Single Beam power (MW)	Normalized emittance $\varepsilon_x/\varepsilon_y$ (10^{-8} m-rad)	β^* (mm)
Muons 3 TeV	14	5000/5000	3
CLIC 3 TeV	11	60/0.1	8/0.1
Matrix Linac 5 TeV	0.7	10/10	0.15/0.15

Table 1: Parameters for Lepton Colliders

where N is the number of particles in a bunch, f_c the frequency of bunch collisions, $\sigma_{x,y} = \sqrt{\beta_{x,y}\varepsilon_{x,y}}$ the (x,y) transverse bunch size, P_{av} the average beam power, and n_c is the number of times a given pair of bunches collides. The luminosity of a multi-TeV collider is of order 10^{35} cm^{-2} sec^{-1}. In an e^+e^- collider, the beams collide once ($n_c = 1$) and must, therefore, be focused to a small spot size, of order $10^{-6}\mu^2$. In a muon collider (see [2] and references therein), the beam emittance is much larger, as is the spot size ($\sim 10\mu^2$). The beams are stored in a ring, with roughly 800 collisions in a storage time (determined by the muon decay). The increase in the number of collisions does not overcome the larger spot size, and the muon collider must also have a larger N, by about a factor of 100. Rough parameters for an e^+e^- collider based on the CLIC scheme [3] and the $\mu^+\mu^-$ collider are given in Table 1. In both cases, the average single beam power exceeds $10MW$.

A second constraint is size. At TeV energies, the machine length of a linear collider quickly becomes very long unless gradients are high. In a structure, the stored energy $U \sim L_{Acc}E_{Acc}^2/\omega_{rf}^2$; at fixed total beam energy $U \sim E_{Acc}/\omega_{rf}^2$. Thus, lower gradients are better from an efficiency standpoint—but worse from a real estate standpoint. Operation at higher frequency is desirable because the stored energy requirements are reduced and the breakdown fields increase. On the other hand, preserving beam quality is easier at lower frequency and less charge. Lower frequency structures are farther from the beam, and thereby have weaker transverse coupling to it. This wakefield coupling is of course a limiting factor on peak single bunch charge in the NLC. So, a shorter linac, with higher frequency and gradients, creates other problems that must be overcome. The present thinking is that if an NLC structure were scaled to W-band, the limiting factor on the gradient obtainable under collider operating conditions would be structure heating, rather than breakdown. Detailed studies of W-band structures are now underway at SLAC [4].

An example of how new ideas might be able to improve some of the overall system parameters is shown in the final row of Table 1. In particular, the average power is now below 1MW. This is accomplished by envisioning an accelerator (the Matrix Linac) and focusing system that is a result of a departure from current thinking [5, 6]. Other high-frequency NLC-like schemes have numbers similar to

the CLIC parameters. The research behind the numbers in Table 1 varies. The CLIC scheme has been an active CERN project for many years, the muon collider has only been an organized collaboration for 3 years, the Matrix Linac has been studied only by a single group at SLAC. I did not list parameters for a plasma-based collider because, in my view, they are less mature than the other concepts. Parameter sets for high-energy colliders not only define hypothetical machines, but also indicate where innovation and research can have a large payoff. With the costs of the NLC now bordering on the prohibitively high, it seems prudent to maintain an effort in alternate concepts for structure design and final focus systems.

We have heard the summary of the NLC research in the talk by Voss. He discussed the NLC concept of the X-band driven linac as well as the CLIC scheme at higher frequency. The development of new technologies, one of which will hopefully be successful enough to reach well-beyond the energies of the X-band driven NLC, has a few primary thrusts:

1. The development of high-gradient systems for linear electron-positron accelerators. This work comprises everything from W-band to plasma-based schemes [7], and the direct laser-driven accelerator structures [8].

2. Muon colliders [2], in which semi-conventional components are employed to produce, cool, and collide muon bunches. Researchers on muon colliders, both in the United States and abroad, have recently shifted their efforts to R&D studies of muon storage rings for neutrino sources. The neutrino-source storage ring [9] is substantially less complex than a muon collider, and would serve as an important development path for this technology.

3. Very large hadron colliders, whose proponents argue that their cost, though they may be hundreds of kilometers in circumference, can be brought down to affordable levels. Little will be said here about the VLHC, because this author wishes to stress machine physics rather than cost issues.

In the following sections, we examine these concepts for high-energy lepton colliders in greater detail.

2 High-frequency RF acceleration: 100 GHz and beyond

The X-band technology of the NLC, with its gradients of order 80 Mev/m, when applied to high-energy colliders ($\geq$ 3 TeV), leads to extremely large machines and average powers. Higher gradients will, of course, lead to smaller machines. The

maximum gradient a structure can hold without breakdown increases with frequency; roughly an order of magnitude in gradient may be expected in switching from X-band to W-band. These compact linear colliders, with gradients of order 0.5–1GeV/m, have a slew of problems that must be overcome. Among them are

1. The development of new structures with tolerable wakefields, thermal properties that allow for rapid pulsing, high breakdown fields, and good shunt impedance.

2. the control of jitter, alignment of components, and emittance preservation,

3. The development of new power sources with sufficient efficiency, peak power, and pulse length and the development of schemes for coupling power from the source into the structure.

4. The development of a new final focus system. Considerations include overall length, the beam-beam interaction (pinch effect, beamstrahlung) and, of course, luminosity and backgrounds.

5. The development of detectors appropriate for operation in this environment, with backgrounds that are likely to be larger than those of the NLC.

6. The development of energy storage systems. The modulators used as storage for the Klystrons are very effective, but not for pulses that are very short. Because fill-time is reduced at higher frequency, a new method of power storage is required. In a two-beam accelerator this is accomplished by using conventional rf or induction linacs to accelerate the low energy beam. Another approach is to use so-called active structures, in which energy is released into the accelerating structure using a photosensitive material and a laser pulse.

The engineering of a scaled NLC structure and source at high frequency becomes very difficult. The wakefield problem, in which the bunch couples to itself (or to subsequent bunches) by exciting fields in the structure, is significantly worse at higher frequencies, where the structures are smaller, and therefore closer to the beam. Advanced structure concepts include the use of advanced materials, active structures, composite structures, disposable structures (plasmas), and a wide variety of possible geometries, such as two-beam accelerators like CLIC.

There has been some preliminary experimentation on W-band structures. The first hurdle in the way of experimental progress was the lack of a coherent high-power source to excite the structure and study its breakdown and thermal properties. One solution has been to operate at the 8th harmonic of X-band and use

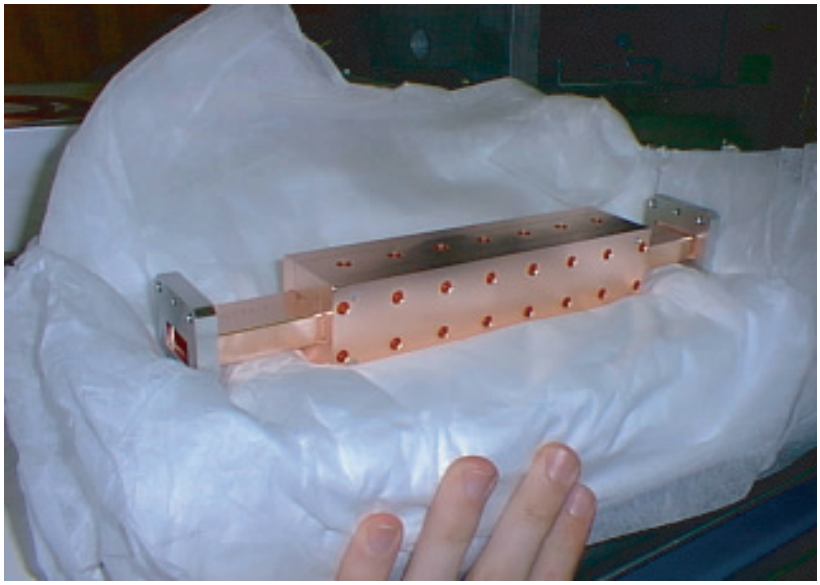

Figure 1: A W-band accelerator structure in which power is supplied in a direction orthogonal to the beam propagation. Bunches propagate through the single set of holes (top side) and rf moves through the rectangular waveguide. Picture courtesy of David Whittum.

the 100 ns bunched beam from the NLCTA. The beam has enough harmonic content to significantly excite the structure. The setup of Whittum and collaborators is sufficiently flexible so that anyone with a W-band structure can consider testing it at the NLCTA. Indeed, the NLCTA can serve to power any structure with a resonant frequency at a low harmonic of its bunch spacing. Unlike a conventional slow wave structure with irises, this smooth structure has a dielectric coating (diamond) on each surface and planar geometry.

An example of a W-band structure is shown in Fig. 1.

A high-energy e^+e^- collider might employ in its final focus neutralized beam collisions, in which the beam-beam interaction is mitigated by the collision of four beams, with a pair of e^+e^- beams colliding with another pair of beams. Assuming identical beam properties, the resultant collision is current and space charge neutral. It will not be stable, so that small perturbations resulting from, say, focusing errors, may grow during the interaction. This growth limits the mitigation of the beam-beam interaction that can be achieved [10]. Scaling laws suggest [11] that operation at higher background levels will be required, and this may require detector development. Another idea for avoiding the undesirable radiative phenomena of the beam-beam interaction is to Compton scatter the e^+e^- beams with high-power lasers upstream from the IP; the machine then becomes

a $\gamma - \gamma$ collider. This idea is also interesting because it provides a new channel for physics studies.

New focusing schemes that can accommodate multi-beam collisions and beam combining and do so in a reasonable length, are required. One novel idea is dynamic focusing [12]. In dynamic focusing, secondary bunches, at about 1% of the linac energy, serve as (moving) lenses for incoming bunches.

Higher frequencies may have an advantage other than gradient. Each pulse of the accelerator will, for reasons of efficiency and bunch stability, consist of a series of n_b microbunches. If these bunches could be combined before they collide, the luminosity will increase by the factor n_b^2. Thus, a train of bunches whose total length is many rf periods is reduced to a single bunch whose length is less than the focal length at the IP. One possibility [6] for accomplishing this is to accelerate each bunch to slightly different energy, and then to use a dispersive section to recombine them. As the frequency is increased, the separation between microbunches decreases, and is therefore easier to compensate for in the dispersive section.

3 The neutrino factory and muon collider

The advantages of accelerating a muon are:

1. There is no significant synchrotron radiation at multi-TeV energies, because the radiation scales as $\propto m^{-4}$. Then, unlike e^+e^-, the particles can be accelerated and collided in circular machines.

2. There are no constituent particles, so that the full energy is available in the center-of-mass.

3. The muon can directly produce the Higgs particle, with a cross section enhanced by m_μ^2/m_e^2 compared to an electron.

There are, of course, many complications associated with developing a collider based on a particle that is produced very dilutely and is unstable:

1. Muons must be produced efficiently.

2. A substantial fraction of the initial 6D phase space must be captured and subsequently cooled by a factor of 100 for the neutrino factory and 10^6 for the collider.

3. Muon decay requires that the cooling and acceleration be done quickly, and that machine components be appropriately shielded from the decay products.

4. At energies beyond 3 TeV (CM), the neutrino flux is sufficiently high that it poses a radiation problem at the site boundary [13]. The radiation hazard results from neutrino interactions with the surroundings; the dose from direct interaction within the body is only ~ 0.1% of the total.

5. The required bunch current is large and collective effects may be important.

The (newly renamed) Neutrino Factory and Muon Collider Collaboration was formed in 1996 and now has more than 135 members. Parameters have been worked out for both a high-energy 3 TeV collider and a low energy "Higgs Factory" machine at 250 GeV. Circular machines are advantageous in that a 3 TeV collider would fit comfortably on the FNAL site.

A reasonably funded R&D program, one in which money is not a primary limiting factor, is anticipated to take of order 5–10 years to determine if such a machine is worth pursuing. A possible first step is to build a neutrino factory. While the neutrino factory is not a high-energy machine, it is a logical first step towards an eventual muon collider. It is a testbed for some of the technology required to build a muon collider, but by no means all. A muon collider is a substantially more difficult machine; the technological obstacles to a neutrino factory will be much easier to overcome than will those impeding a muon collider.

As shown in Fig. 2, a neutrino factory consists of a proton driver, a target, perhaps an initial low frequency cavity for preliminary phase rotation, an induction linac for phase rotation, a cooling section, a linac, one or more recirculating linacs, and a storage ring. Parameters for a possible neutrino factory are given in Table 2. Two observations can be made. First, additional proton power linearly increases the neutrino flux. An important consideration in this regard is the balance between detectors and proton driver power. The divergence angle is such that most neutrinos miss the detector altogether. Increasing the event rate by building larger, or simply more, detectors may become an attractive alternative to increasing the proton driver power. The linear power scaling for the neutrino flux is more appealing than that for the luminosity in a muon collider, which depends quadratically on the number of muons per bunch, and, therefore, also quadratically on the proton driver power and bunch number.

3.1 The individual components of the neutrino factory

The target

Typical targets under consideration are composed of solid carbon or liquid mercury, the former being easier technologically, and the latter having greater technical risk as well as higher yield. A mercury target schematic is shown in Fig. 3. The 20 T solenoid captures the initial pion beam. The pions quickly decay to muons

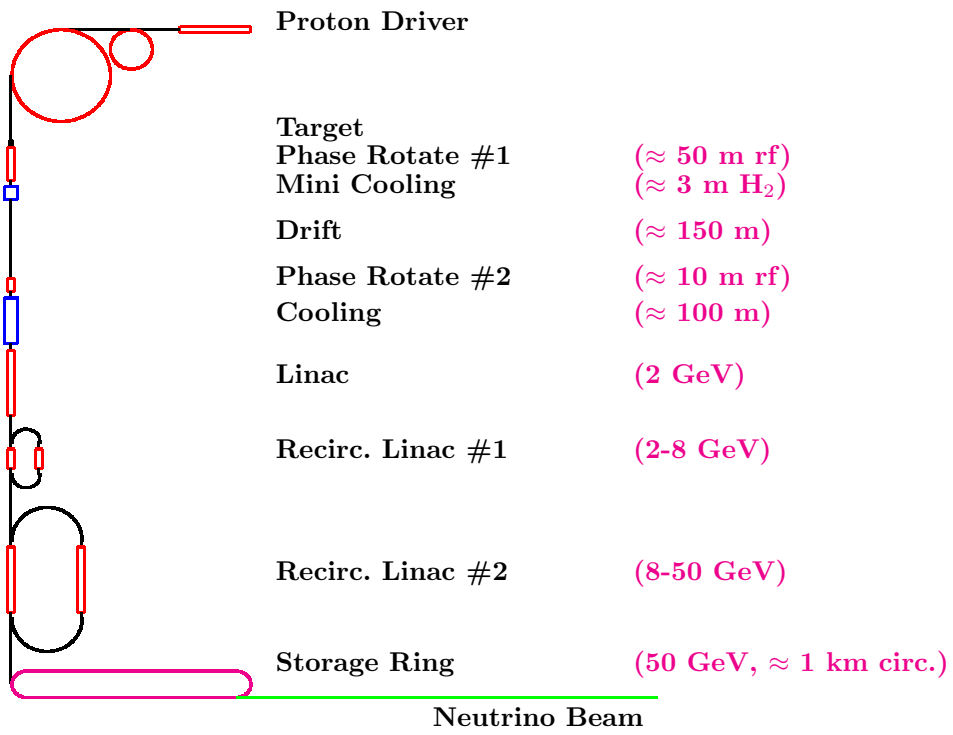

Figure 2: The essential components of a neutrino source (see [9, 14]). The energy depends on the physics goals and may be less than 50 Gev.

with an energy spread of order unity. The preponderance of the muons are low energy, and it is these that must be captured. Critical to the target performance is the pion yield at low energy. Unfortunately, there is a paucity of experimental data in the regime of interest. An important goal of the neutrino factory R&D program is the study of pion production rates and target survivability. Experiments are being discussed to study production with a variety of proton beam energies (from 2 GeV to 24 GeV) and different target materials and geometries. This work will allow for benchmarking of the simulation codes. The target is slanted, so as to reduce pion losses from absorption, and its length is of order an interaction length, so that most of the protons interact. The target typically absorbs about 10% of the proton beam energy; this heating will set a lower bound on its total volume. A discussion of targets can be found in [2].

The pions are produced in a bunch with a 7.5 cm radius, an rms angular spread of roughly 15 degrees, a bunch length of order the proton bunch length (a few ns), and an energy spread of order unity. The maximum yield is for longitudinal momenta of order 200 MeV/c; this is also the average transverse momentum. In order to confine the pions transversely, the target is immersed in

Energy	20–50	GeV
Decay ratio towards detector	40	%
Normalized emittance	0.0032	m-rad
Number of muons/pulse	6	10^{12}
Typical decay angle ($1/\gamma$)	2.0	mrad
Lifetime	3×10^5	$c\gamma\tau_\mu$
Protons on target/pulse	20	10^{12}
Pulses/s	15	Hz
Proton energy	16	GeV
Proton driver power	1	MW
Number of μs/yr	2×10^{20} (towards detector)	

Table 2: Possible parameters for a neutrino factory [15].

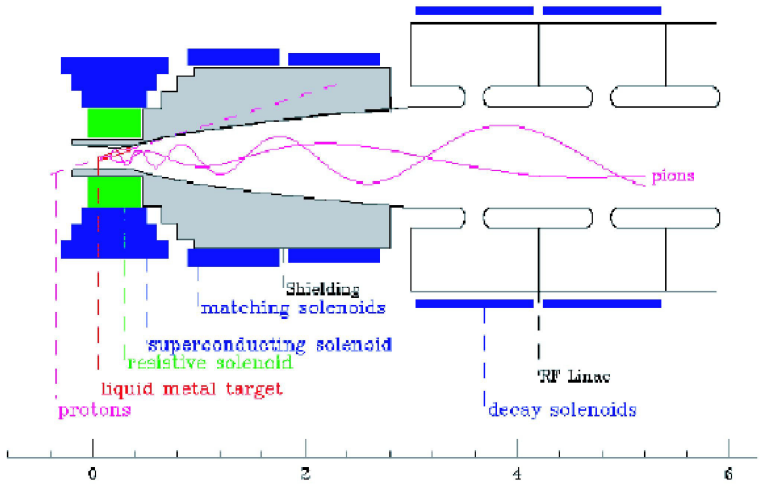

Figure 3: A proton target.

a 20 T magnet. This high field cannot be maintained and is tapered adiabatically by a factor of 16; this reduction causes the beam radius to grow, and the angles to shrink, by a factor of 4. The pions decay into muons and the beam begins to lengthen due to its large parallel velocity spread (of order 0.2 c). A low frequency rf system just beyond the target can be used to give an initial phase rotation to the beam. Such a low frequency cavity is both large and expensive,

and there are questions as to its survivability in a high radiation environment. It is required to preserve the short bunch in the muon collider, but may not be needed in the neutrino factory (although it likely will increase the muon yield at the end of the cooling channel). The energy spread is sufficient that the beam will lengthen in a drift section of order 100 m, with the higher energy particles arriving at the head of the pulse (now 40 ns long). This long continuous beam has an energy variation that is correlated with position in the pulse. This correlation or “chirp" is eliminated by accelerating the tail and decelerating the head of the beam. This process, known as “phase rotation,” may be accomplished with an induction linac. The energy spread after phase rotation is about 15%, and the beam is ready to be bunched into rf buckets and cooled. In the collider, the beam cannot be allowed to lengthen beyond a few nanoseconds, and the phase rotation must take place near the target.

Beam scraping sets aperture limits, and consequently an upper bound on the rf frequency (because the cavity dimension scales inversely with frequency). This upper bound, of order 200 MHz, determines the frequency used in the bunching and cooling sections. The frequency of the accelerator can be increased at higher muon energy. Cooling must occur quickly on the timescale of a muon lifetime, which corresponds to about 1 km of beamline at the muon energies of around 200 MeV. The method of cooling that seems the most promising is ionization cooling. The principle of ionization cooling is straightforward. The muons ionize and lose energy along their direction of motion. The rf accelerates them in the longitudinal direction, thereby reducing the transverse emittance. There are two problems that limit this:

1. Longitudinal emittance growth is caused by the unfavorable dependence of energy loss (δE) on energy. (Landau straggling also contributes in principle but is not important.) At the muon energy in the cooling channel, $d\delta E/dE < 0$, so lower energy muons have higher ionization losses, and the energy spread grows. This is an important feature of the cooling channel, because muons can slip out of rf buckets and be lost. The alternative of operating at a higher average energy, where $d\delta E/dE > 0$ and the energy spread is reduced, would require a much longer cooling channel. The explanation for this points out another engineering constraint on the cooling channel. One e-fold of cooling requires that the muons must lose, and have replenished by the rf, roughly their full energy. If the rf gradients could be increased, and the absorber thickened, then operating at a higher mean energy might be more attractive. Unfortunately, the designs already assume ambitious but not unrealistic gradients (15 MV/m at 200 MHz, for example). Higher energies imply longer channels, and thus higher costs. Muon losses to decay are not affected to lowest order, because the decay length is also dependent on energy. However, for muon cooling, $\beta\gamma \sim 1.8$ and lowest

order approximations are not so accurate.

2. Multiple scattering in the absorber increases the emittance and completes against the cooling. The design of a cooling channel requires that the effects of multiple scattering be mitigated by making the beam spot size small in the absorber. The smaller the beam is, the more its natural spread in transverse velocities minimizes any increase in the spread due to the scattering. As the beam cools, its transverse velocity spread decreases and the spread from the scattering cancels the cooling. The cooling saturates. Because focusing strength is related to magnetic field strength, high fields are desirable. Engineering constraints, such as peak magnetic fields at superconductors, limit the field strength. Scattering from the aluminum windows that hold the LiH absorber and possible beryllium windows on the rf all contribute to scattering and limit the performance.

The longitudinal emittance growth results in particle loss. It is hoped that for the neutrino factory, this growth will be tolerable. For the muon collider, the longitudinal phase space must be reduced by about two orders of magnitude in the cooling channel. Emittance exchange, in which the longitudinal phase space is decreased and transverse phase space increased, is required. There are a few concepts for emittance exchange, based on using dispersion to create a bunch with a correlation in energy and transverse position, and then sending the bunch through a tapered absorber so that higher energy particles lose more energy. The resulting bunch now has a higher transverse emittance, because the correlation is eliminated and dispersion cannot be used to reduce the transverse size. Emittance exchange is one of the critical challenges for the collider. It has yet to be successfully demonstrated in a realistic simulation.

A reasonably complete description of the transverse dynamics of rms beam properties in the cooling channel has recently been derived by Penn [16]. This description should allow for faster evaluation of various magnetic geometries. Until such average models are extended to include longitudinal effects, the modeling of the cooling channel for the collider will still require tracking simulations such as ICOOL and DPGEANT. A schematic of a section of a cooling channel can be seen in Fig. 4. The beam is focused by a series of alternating solenoids. The rf has beryllium windows to allow for a higher gradient (at fixed power). The aluminum windows at each side of the absorber are not shown.

The results of a full front end simulation can be seen in Figs. 5–6. The plots start 100 m downstream from the target. These results were obtained by the late Charles Kim and co-workers at LBNL. Two interesting figures of merit are the total number of muons per proton and the number of muons per proton in the acceptance of the downstream part of the system (*i.e.*, the linac and storage ring). Clearly seen in the initial steep rise in the latter number are the effects of

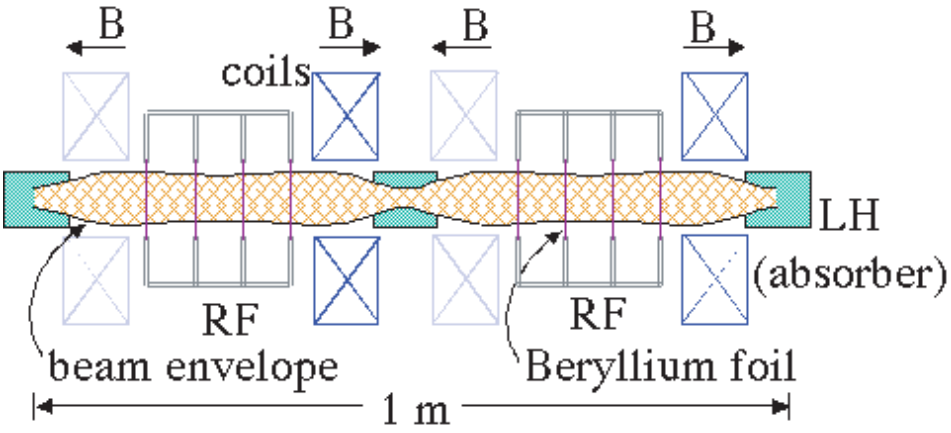

Figure 4: Schematic of a Muon Cooling Channel.

the phase rotation, where the bunch with a large correlated head-to-tail energy spread propagates through the induction linac. The head is decelerated and the tail accelerated so that most of the bunch is within the longitudinal acceptance. The slower rise occurs in the cooling channel, where the total emittance is reduced, as seen in Fig. 6. The losses in the total number of muons come from scraping and longitudinal emittance growth in the cooling channel. Emittance exchange, if feasible, may well be desirable for the neutrino factory. For example, the cooling channel is limited by the large energy acceptance. With emittance exchange after some initial cooling, the energy spread could be reduced and a channel with stronger focusing and better cooling performance employed. This is the conceptual path to the factor of 10^{-6} cooling required in the collider.

3.2 The acceleration and storage

The beam beyond the cooling channel enters a linac to boost its energy to a few GeV. From there, it enters a recirculating linac, which will bring it to about 20 GeV. If the final energy of the neutrino factory is higher than around 20 GeV, a second recirculating linac will be needed. The muons are stored in a ring, with, for example two straight sections, one pointing to a near detector the second to a long baseline detector. The ring will have to be built at a slant. For example, an FNAL to SLAC arm would have to be pointing down at an angle of 17 degrees. Very long baseline arms, with twice this slope, are also under consideration. While a slanted ring is a complication from an engineering standpoint, this aspect of the neutrino factory has obvious political benefits. It is naturally international, and may be of significant benefit to more than one domestic laboratory.

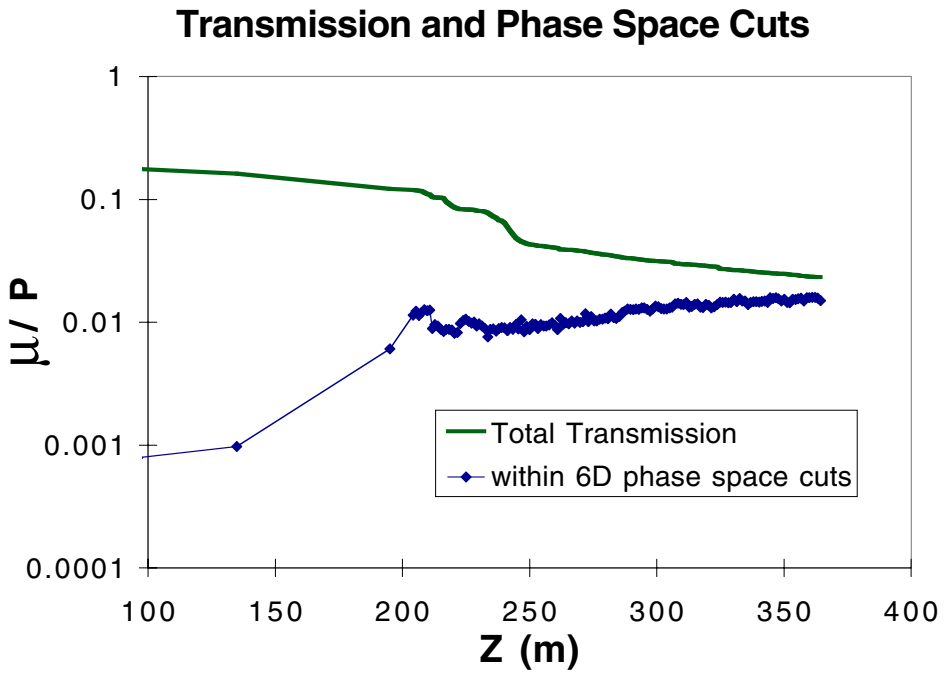

Figure 5: Simulation showing the total number of muons per incident proton, and number of muons per incident proton within the six-dimensional phase space that the accelerator can accept, as a function of propagation distance downstream from the target. Graph courtesy of G. Penn.

3.3 The muon collider ring

At high energy, shielding, about 6 cm thick, must be used to protect the superconducting magnets from heat generated by the decay products. This large aperture limits the on-axis field strength of the magnets. After the IP focus, the beam expands and suffers chromatic aberrations that must be corrected. The bunch length cannot exceed the small focal length at the IP ($\beta^* \sim 3$mm) without incurring "hour glass" reductions in the luminosity. This short bunch could, in principle, be maintained by a large rf system. Such a system would be expensive, and would introduce undesirable increases in the transverse and longitudinal impedances. The plan to have a nearly isochronous ring is preferable, but ambitious. The muon collider ring has aspects that are similar to both hadron and lepton rings. The tune shift is comparable to that in an electron collider ring but, as in the case of a proton ring, there is no radiation damping. Unlike either of these rings, the storage time is short, about 1000 turns. In this respect, especially with the low momentum compaction, the ring has dynamics similar to a

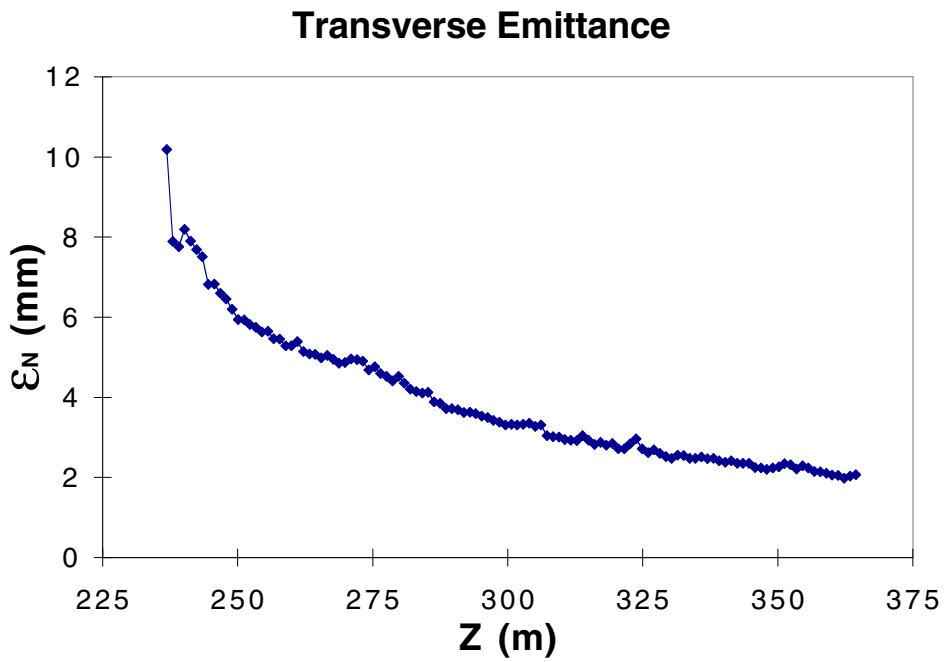

Figure 6: Simulation showing the total number of muons per incident proton, and number of muons per incident proton with the six-dimensional phase space of that the accelerator can accept, as a function of propagation distance downstream from the target. Graph courtesy of G. Penn.

linac (one whose length is thousands of kilometers). Indeed, something similar to BNS damping, a technique used at SLAC to control transverse instabilities, is required for the control of instabilities in the collider ring.

4 The role of plasmas in advanced accelerators

Looked at from a conventional accelerator perspective, a plasma is a disposable structure that must be excited and used on a very short timescale. The advantage of a plasma is that it can maintain large gradients, of order $E_{z,max}$ (GeV/m)$\sim \sqrt{n_0}/10^{15}$, where the plasma density n_0 is measured in cm^{-3}. Plasma has been proposed not only as an accelerating medium, but also as a focusing medium, and as a current neutralizing medium. To achieve this, the plasma must be manipulated over time scales and spacial scales governed by the plasma frequency $\omega_p = e^2 n_0/4\pi\varepsilon_0$, and skin depth, c/ω_p. Here, e is the electron charge, c the speed of light, and ε_0 the permittivity of free space. These fields in the plasma are excited by lasers or by charged particle beams.

The plasma is not confined by external magnetic fields; the timescales are sufficiently short that ion motion is not significant. Over longer, hydrodynamic timescales, the plasma structure disappears; it must be recreated (for example, by ionizing fresh gas from a gas jet) at each pulse of the accelerator. This manipulation of the plasma might well be called femtosecond engineering; many groups are demonstrating some skill at it.

4.1 Laser-driven plasma accelerators

If one wanted to make an accelerator with an rf wavelength of, say, 100 μ, one would quickly find that there are no high-power coherent sources. The laser-driven plasma accelerators operate with "rf wavelengths" in this range. The plasma, an active medium, couples power from the primarily transverse laser light into a primarily longitudinal plasma wave at a longer wavelength. That this could be done was first pointed out in the original paper of Tajima and Dawson [17], where it was shown that an intense short pulse or a beat of two co-propagating pulses will excite a plasma wave. The plasma wave has a phase velocity equal to the group velocity of the laser pulse. The idea has been the subject of much investigation, and is reviewed in [7].

Over the last 20 years, technological advances in short pulse lasers have propelled many researchers into laser-driven plasma-based accelerator research. The scale of these experiments is relatively small, so that they can be done by university as well as national laboratory groups. To gain a sense of this work, look at the large number of experimental papers listed in Table 3. Plasma-based accelerators, with multi-GeV/m gradients, have accelerated particles (but only over mm distances!) to MeV energies. Some points worth noting: The experiments can be distinguished by those that *inject* a beam and those that observe what comes out spontaneously from the plasma. The former experiments use particle beams much longer than the plasma wavelength, hence only a small fraction of the beam experiences maximum acceleration. The final energy spread is of order unity. These experiments use either a beat-wave, with relatively low peak powers, or an intense short pulse. The peak gradients are $\sim$ 1.5 GeV/m. A second class of experiments, (SM-LWFA) uses an intense laser with pulse length significantly greater than a plasma wavelength. The laser pulse undergoes what is known as the forward Raman instability, essentially down-converting to a lower frequency electromagnetic wave and a plasma wave. The instability is sufficiently violent that plasma electrons trapped in the wave are accelerated to relativistic velocities and ejected. A similar rf-structure experiment would be to observe "dark current" when powering a structure. This dark current is highly undesirable in the collider, where it can create backgrounds and must be collimated from the beam. Indeed, rf structures that give off nC of charge in dark current are con-

	I [W/cm^2]	τ_L [ps]	λ [μm]	n_0 [cm^3]	ΔE [MeV]	E_z [GV/m]
LWFA:						
KEK (Japan) [22]	10^{17}	1.0	1.05	10^{15}	5	0.7
LULI (France) [23]	4×10^{17}	0.4	1.05	2×10^{16}	1.6	1.5
PBWA:						
ILE (Japan) [24]	10^{13}	1000	9.6, 10.6	10^{17}	10	1.5
UCLA (USA) [25]	10^{14}	300	10.3, 10.6	10^{16}	28	2.8
LULI (France) [26]	10^{17}	90	1.05, 1.06	10^{17}	1.4	0.6
SM-LWFA:						
LLNL (USA) [27]	10^{18}	0.6	1.05	10^{19}	2	-
KEK (Japan) [28]	10^{17}	1.0	1.05	10^{19}	17	30
RAL (UK) [29]	10^{19}	0.8	1.05	10^{19}	44	100
CUOS (USA) [30]	4×10^{18}	0.4	1.05	3.6×10^{19}	>1	-
NRL (USA) [31]	5×10^{18}	0.4	1.05	1.4×10^{19}	>1	-

Table 3: Parameters and results for laser-driven plasma-based accelerator experiments. The laser intensity I, laser pulse duration τ_L, laser wavelength λ, plasma density n_0, energy gain of the accelerated particles ΔE, and accelerating gradient E_z are listed for each experiment. The experiments listed as LWFA (laser wakefield accelerator) use a laser pulse short compared to the c/ω_p. The PBWA (plasma beatwave accelerator) experiments use two long pulse lasers with a frequency difference $\delta\omega = \omega_p$. SM-LWFA experiments rely on the self-modulation instability to break a long pulse into a series of pulses, thereby generating a plasma wave.

ditioned or operated at lower gradients. These self-modulated experiments are setting upper bounds on the gradient the plasma can hold and still be useful as a component of a linac.

Energies in laser-based schemes are primarily limited by the diffraction of the laser pulse. Diffraction can be mitigated by increasing the spot size, but at a cost of lower peak intensity to drive the wake. Furthermore, lasers with transverse spot size much larger than the plasma skin depth, c/ω_p tend to have instabilities. The natural transverse size of the plasma structure scales with the wavelength, just as in a metallic structure. A better alternative to enlarging the laser pulse is to use the index of refraction of the plasma to guide the pulse. A rarified plasma channel can guide a pulse. This is not surprising, because the plasma index of refraction is less than unity. A fluid simulation of a plasma channel is shown in Fig. 7. The simulation shows the excitation of an electromagnetic wake near the

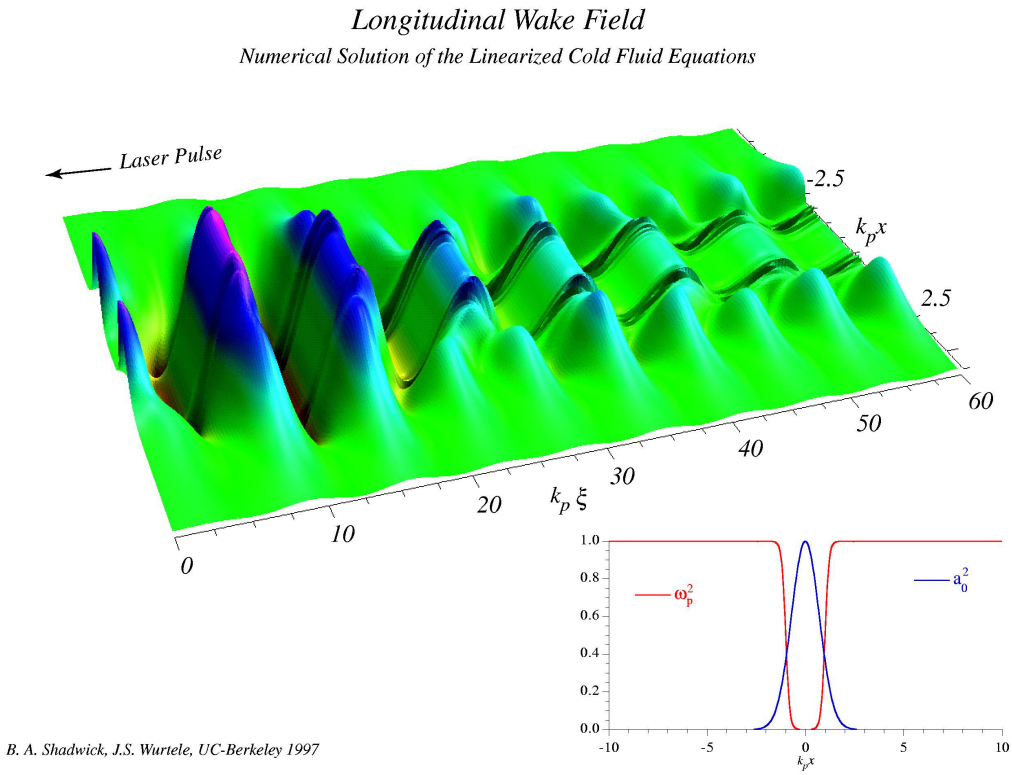

Figure 7: Excited Wake in a Plasma Channel

channel axis and of an electrostatic wake in the bulk plasma. These modes have slightly different frequencies, as predicted by theory. Resonant absorption of the electromagnetic wake leads to finite Q, as seen in the figure. The effective Q of the mode is important for collider applications, where multiple bunches may be needed. For numerical purposes, the bulk $\omega_p = 1$, and the laser power is expressed in terms of the dimensionless vector potential, $a_0 = eE_L/mc\omega_L$. The ideal case of a hollow channel with infinitely steep walls has the further virtue that the accelerating field is independent of transverse position and the focusing fields are linear.

Guiding experiments are shown in Table 4. Progress is steady, and now power levels of 10^{18}W/cm^2 can be propagated over many diffraction lengths. The next

	I [W/cm^2]	τ_L [fs]	λ [μm]	n_0 [cm^3]	$r_{\rm ch}$ [μm]	$L_{\rm guided}$ [cm]
UMCP (USA) [18]	5×10^{15}	500	0.565	7×10^{18}	30	2 (45 Z_R)
LBNL (USA) [19]	5×10^{17}	75	0.8	7×10^{18}	5	0.1 (8 Z_R)
NRL (USA) [20]	10^{17}	400	1.05	10^{18}	500	2 (22 Z_R)

Table 4: Parameters and results for plasma channel laser guiding experiments. The laser intensity I, the laser pulse duration τ_L, the laser wavelength λ, the plasma density n_0, the characteristic radius of the plasma channel $r_{\rm ch}$, and the propagation length of the guided laser pulse ($L_{\rm guided}$ measured in centimeters and number of Rayleigh lengths) are listed for each experiment. This table, and the next two, are courtesy of Carl Schroeder. For a more detailed discussion of these results and theoretical work on plasma accelerator structures, see his thesis [21].)

major challenge is to put all the pieces together and accelerate a bunch with low energy spread and emittance to a substantial fraction of GeV.

The bunches from conventional accelerators are too long to be usefully injected into a plasma accelerator (or direct laser accelerator), where the accelerating wavelength is $1-100\,\mu$ m. One clever way to generate such bunches is by optical injection [36]. This works as follows: A plasma is excited below the self-trapping threshold by an intense short pulse; a second ultra-short pulse, or pair of pulses, then perturbs the plasma locally. Electrons located where the second pulse(s) interacts with the wake are trapped and accelerated. A particular realization of this idea is the colliding pulse injection scheme [37]. As the two short pulses overlap, they form a ponderomotive beat wave, with period of half their wavelength. Plasma electrons are accelerated into the ponderomotive bucket by the beat.

4.2 Beam-driven plasma accelerators

It is clear from earlier discussion that the availability of high power short pulse lasers has led to a significant amount of experimentation. It is unclear, however, whether these laser-driven systems, their potential for short bunch generation not withstanding, will be able to meet the efficiency criteria of high-energy colliders. For example, the laser pulse must be depleted by the generation of the wake, yet not suffer any instabilities in the plasma that would destroy the coherence of the accelerating wake. No experiments have approached such a regime; some results are given in Table 5.

A particle beam driver, which can be ultra-relativistic, is subject to less violent

	γ_{drive} [MeV]	q_{drive} [nC]	$c\tau_b$ [mm]	n_0 [cm^{-3}]	ΔE [MeV]	E_z [MV/m]
ANL (USA) [32]	21	4.0	2.1	10^{12}	0.2	5.0
KEK (Japan) [33]	500	10	3.0	10^{12}	30	30
KhFTI (Ukraine) [34]	2	0.4	17	10^{11}	0.5	0.25
SLAC (USA) [35]	3×10^4	3.2	0.6	10^{15}	-	-

Table 5: Parameters and results for beam-driven plasma-based accelerator experiments. The drive beam energy γ_{drive}, drive beam charge q_{drive}, drive beam length $c\tau_b$, plasma density n_0, energy gain of the accelerated particles ΔE, and accelerating field E_z are listed for each experiment.

Length of each module	10 m
Number of modules	100
Number of electrons in the HEB	5×10^9
Plasma density	10^{15}cm^{-3}
Driver energy	10 GeV
Number of electrons in each macro-bunch	10^{11}
Number of micro-bunches per macro-bunch	7
Driver radius ($0.8c/\omega_p$)	0.14 mm

Table 6: Novosibirsk parameters for a plasma wakefield module of an accelerator. An accelerator of 100 modules would yield a 1 TeV HEB (high-energy bunch) with an energy spread of 3%. From [38].

instabilities than a laser driver, and may be much better suited for the high-energy collider application. Some experimentation has been carried out with relatively low beam energies and plasma gradients by Rosenzweig and collaborators [32]. A serious attempt to think though the critical problems associated with such a system was made by a Novosibirsk group. They proposed a 1 TeV accelerator scheme with 100 10 GeV modules, with each module having the parameters of Table 6.

Two high-energy beam-plasma experiments are underway at SLAC, E150 (plasma lens) and E157 (plasma wakefield). These experiments, which have yet to release major results from their initial runs, will yield important lessons in developing diagnostics and understanding of beam dynamics in plasmas. Should they be successful, we may come to the point where the plasma is "just another" beamline element. A schematic of E157 is shown in Fig. 8. This experiment uses plasmas of order 1 m and is the first to study transport over a few betatron oscillations

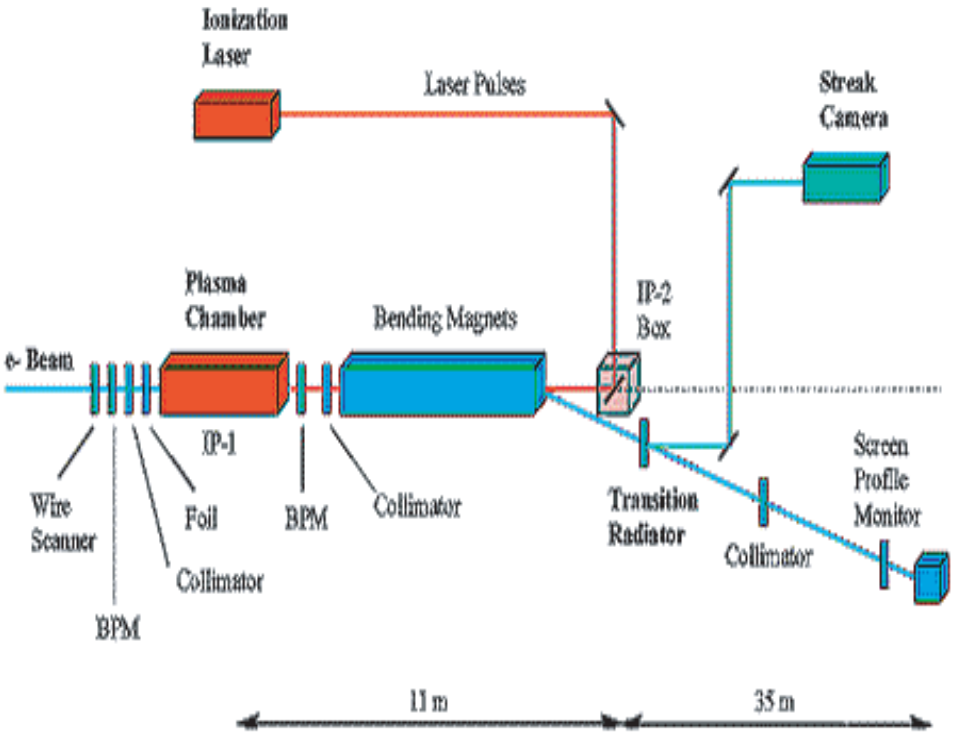

Figure 8: Schematic of the E157 experiment [40].

at such high beam energies.

Beams are subject to hosing instabilities in the plasma. This should not be surprising because the plasma fields are strong enough to induce betatron oscillations in the beam. These transverse displacements induced by the head of the beam kicking the tail can, depending on the experimental situation, look like displacements in an energy-resolving magnet due to energy changes [39].

It is interesting that when the plasma response is linear, the beam is self-focused by its own magnetic field (which may be reduced by plasma return current if the beam radius is $\geq c/\omega_p$). This is a nonlinear focusing force, and undesirable rms emittance growth results. Estimating this growth is not simple. For example, the beam dynamics, in the long-pulse relativistic limit, with no return current, is equivalent to the non-trivial problem of a self-gravitating collection of equal masses in two dimensions.

5 Laser acceleration

Acceleration with laser pulses will allow for the production of short bunches, with the size less than 0.1 μ. These bunches may be of interest to the other users. For example, they could be used to create short pulses of X-rays by Compton backscattering off a laser pulse. They could be made to radiate coherently, and thus copiously, at wavelengths of order their bunch length. For the collider application, the ultrashort bunchlength is not required, and not even necessarily desirable.

6 Conclusions

Concomitant to what is required for understanding the underlying beam physics of new accelerator concepts is research on ways to implement these ideas practically, which will involve getting around the numerous difficulties that can be expected. This will likely lead, as well, to further appreciation for the first possible applications of these ideas. The neutrino factory is a necessary but not sufficient first step to proving the feasibility of the muon collider, and should be of significant physics interest. Laser-driven accelerators may produce ultrashort bunches with diverse applications long before it is clear if they can contribute to the collider problem. The path towards high-energy machines, beyond the LHC and the NLC, is a long one. Active work is going on over a wide range of technologies. With a concerted effort of us all, the next generation of colliders may not be the last.

This work benefited from useful discussions with David Whittum. It was supported by the US Department of Energy, Division of High-Energy Physics.

References

[1] J. Lykken, these proceedings.

[2] A recent status report on the muon collider is Charles M. Ankenbrandt *etal.*, Phys. Rev. ST Accel. Beams, **2**, 081001 (1999). This paper has many references to the early literature.

[3] G. Voss, these proceedings.

[4] See `http://www.slac.stanford.edu/grp/arb/tn`.

[5] D. H. Whittum and S. G. Tantawi, SLAC-PUB-7845 (1998).

[6] F. Zimmermann, D. H. Whittum, and M. E. Hill, SLAC-PUB-7856, Talk given at 6th European Particle Accelerator Conference (EPAC 98), Stockholm, Sweden, 22-26 Jun 1998.

[7] A detailed review can be found in E. Esarey, P. Sprangle, J. Krall, and A. Ting, IEEE Trans. Plasma Sci. **PS-24**, 252 (1996).

[8] T. Plettner, R. L. Byer, T. I. Smith, J. E. Spencer, R. H. Siemann, and Y. C. Huang, in Proceedings Advanced Accelerator Concepts Workshop (Baltimore, Maryland, 6-11 Jul 1998), 118-127, 1999.

[9] A wealth of current thinking on Neutrino Factories can be found in Proceedings of the NuFact99 Workshop, B. Autin, editor, to appear as a special issues of NIMA, 2000. A discussion of history of the neutrino factory can be found at `http://www.fnal.gov/projects/muon_collider/nu/history.html`.

[10] D. H. Whittum and R. H. Siemann, in Proceedings of PAC97 (1997).

[11] S. Chattopadhyay, J. Wurtele, and D. Whittum "Advanced Accelerator Technologies: A Snowmass '96 subgroup summary," Proceedings Snowmass96 Workshop.

[12] J. Irwin, "Dynamic focusing for linear colliders using a stored beam as the lens," in Proceedings Advanced Accelerator Concepts Workshop, Baltimore, Maryland, 6-11 Jul 1998.

[13] B. J. King, in Proceedings of PAC99 (1999).

[14] See for example, S. Geer, Phys. Rev. **D57**, 6989, 1998; and R. B. Palmer, C. Johnson, and E. Keil, CERN-SL-99-070-AP, and in Proceedings of the NuFact99 Workshop, to appear as a special issue of NIMA, 2000; see also "Prospective Study of Muon Storage Rings at CERN," edited by B. Autin, A. Blondel, and J. Ellis (CERN 99-02).

[15] N. Holtkamp, private communication. These are the parameters for the FNAL machine study of a neutrino factory. The physics study considered a range of energies.

[16] G. Penn and J. S. Wurtele, "Envelope Equations for a Solenoidal Focusing System," submitted for publication, and G. Penn, Muon Collider Note 71, available at `http://www-mucool.fnal.gov/notes/notes.html`.

[17] T. Tajima and J. Dawson, Phys. Rev. Lett. **43**, 267 (1979).

[18] C. G. Durfee III, J. Lynch, and H. M . Milchberg, Phys. Rev. **E 51**, 2368 (1995).

[19] P. Volfbeyn, E. Esarey, and W. P. Leemans. Phys. Plasmas **6**, 2269 (1999).

[20] D. Kaganovich, A. Ting, C. I. Moore, A. Zigler, H. R. Burns, Y. Ehrlich, R. Hubbard, and P. Sprangl, Phys. Rev. **E59**, 4769 (1999).

[21] C. B. Schroeder, "Plasma-based Accelerator Structures," PhD thesis, University of California, Berkeley, 1999 and LBNL-44779 (1999).

[22] K. Nakajima, M. Kando, H. Ahn, H. Kotaki, T. Watanabe, T. Ueda, M. Uesaka, H. Nakanishi, A. Ogata, T. Kawakubo, and K. Tani. "Recent Results of Laser Wakefield Acceleration in KEK/U. Tokyo/JAERI," In S. Chattopadhyay, J. McCullough, and P. Dahl, editors, Advanced Accelerator Concepts: Seventh Workshop, pages 83–95, 1996, AIP.

[23] F. Amiranoff, S. Baton, D. Bernard, B. Cros, D. Descamps, F. Dorchies, F. Jacquet, V. Malka, J. R. Marqués, G. Matthieussent, P. Miné, A. Modena, P. Mora, J. Morillo, and Z. Najmudin, Phys. Rev. Lett. **81**, 995 (1998).

[24] Y. Kitagawa, T. Matsumoto, T. Minamihata, K. Sawai, K. Matsuo, K. Mima, K. Nishihara, H. Azechi, K. A. Tanaka, H. Takabe, and S. Nakai, Phys. Rev. Lett. **68**, 48 (1992).

[25] C. E. Clayton, M. J. Everett, A. Lal, D. Gordon, K. A. Marsh, and C. Joshi, Phys. Plasmas **1**, 1753 (1994).

[26] F. Amiranoff, D. Bernard, B. Cros, F. Jacquet, G. Matthieussent, P. Miné, P. Mora, J. Morillo, F. Moulin, A. E. Specka, and C. Stenz, Phys. Rev. Lett. **74**, 5220 (1995).

[27] C. Coverdale, C. B. Darrow, C. D. Decker, W. B. Mori, K.-C. Tzeng, K. A. Marsh, C. E. Clayton, and C. Joshi, Phys. Rev. Lett. **74**, 4659 (1995).

[28] K. Nakajima, D. Fisher, T. Kawakubo, H. Nakanishi, A. Ogata, Y. Kato, Y. Kitagawa, R. Kodama, K. Mima, H. Shiraga, K. Suzuki, K. Yamakawa, T. Zhang, Y. Sakawa, T. Shoji, Y. Nishida, N. Yugami, M. Downer, and T. Tajima , Phys. Rev. Lett. **74**, 4428 (1995).

[29] A. Modena, Z. Najmudin, A. E. Dangor, C. E. Clayton, K. A. Marsh, C. Joshi, V. Malka, C. B. Darrow, and C. Danson, IEEE Trans. Plasma Sci., **PS-24**, 289 (1996).

[30] R. Wagner, S. Y. Chen, A. Maksimchuk, and D. Umstadter, Phys. Rev. Lett. **78**, 3125 (1997).

[31] A. Ting, C. I. Moore, K. Krushelnick, C. Manka, E. Esarey, P. Sprangle, R. Hubbard, H. R. Burris, R. Fischer, and M. Baine, Phys. Plas. **4**, 1889 (1997).

[32] J. B. Rosenzweig, P. Schoessow, B. Cole, W. Gai, R. Konecny, J. Norem, and J. Simpson, Phys. Rev. **A39**, 1586 (1989).

[33] H. Nakanishi, A. Enomoto, A. Ogata, K. Nakajima, D. Whittum, Y. Yoshida, T. Ueda, T. Kobayashi, H. Shibata, S. Tagawa, N. Yugami, and Y. Nishida, NIM **A328**, 596 (1993).

[34] A. K. Berezin, Ya. B. Fainberg, V. A. Kiselev, A. F. Linnik, V. V. Uskov, V. A. Balakirev, I. N. Onishchenko, G. L. Sidel'nikov, and G. V. Sotnikov, Plasma Phys. Rep. **20**, 569 (1994).

[35] R. Assmann, P. Chen, F.-J. Decker, R. Iverson, M. J. Hogan, S. Rokni, R. H. Siemann, D. Walz, D. H. Whittum, P. Catravas, S. Chattopadhyay, E. Esarey, W. P. Leemans, P. Volfbeyn, C. Clayton, R. Hemker, C. Joshi, K. Marsh, W. B. Mori, S. Wang, T. Katsouleas, S. Lee, and P. Muggli, "Progress toward E-157: A 1 GeV plasma wakefield accelerator," in Proceedings of the Particle Accelerator Conference, 330–332, New York, 1999.

[36] D. Umstadter, J. K. Kim, and E. Dodd, Phys. Rev. Lett. **76**, 2073 (1996).

[37] C. B. Schroeder, P. B. Lee, J. S. Wurtele, E. Esarey, and W. P. Leemans. Phys. Rev. **E59**, 6037 (1999).

[38] A. M. Kudryavtsev, K. V. Lotov, and A. N. Skrinsky, NIM **A410**, 388 (1998).

[39] A. Garacia and D. H. Whittum, private communication.

[40] From `http://www.slac.stanford.edu/grp/arb/e157/virtual.html`.