THE CASE FOR AN S-BAND LINEAR COLLIDER

Gustav-Adolf Voss
Deutsches Elektronen-Synchrotron DESY
Notkestraße 85
2000 Hamburg 52, Germany

Abstract

Linear colliders are the only machines which can possibly extend the energy range of electron-positron collisions beyond 200 GeV c.m.s., the maximum energy foreseen for the storage ring LEP. At present the only realistic design of a linear collider seems to be two low emittance linacs for electrons and positrons, aiming their beams at each other and focusing them to extremely small spot sizes at the interaction point. There are many important aspects to the parameter choice of these linacs: Maximum obtainable luminosity, power efficiency, costs, machine and physics background, tolerances and possible beam instabilities, required R & D work, extension of energy and many more. We believe, that cost optimization will lead to a design of a 500 GeV c.m.s. collider with only modest accelerating gradient (about 20 MeV/m) and that S-band linacs operated in a multibunch mode (about 200 bunches per 2 μsec beam pulse) are the best choice in almost every respect. A test facility now under construction, consisting of 4 six meter long accelerating guides powered by two 150 MW klystrons, will facilitate important tests on higher order mode excitation and at the same time be used to try out inexpensive but reliable components suitable for a large collider.

It is a remarkable fact, that, though we love to think of our science of accelerator building as a rational science, optimized solutions to a particular problem, when performed by experts in the field, can still be vastly different. Nowhere is this so blatantly apparent as in the field of linear colliders. Here the frequencies of the proposed linacs vary from 1.3 to 30 GHz, the gradients from 17 to more than 100 MVm^{-1}, and the technology to be used is also very different: Obviously there is quite different technical judgement involved, based perhaps on different experiences (or just enthusiasm). To better understand the present situation it may be useful to recall the historical development of linear collider ideas:

Once it had been realized that, with LEP, electron-positron storage rings had reached their economical limit and that any increase of center of mass energy could, if at all, only be done with colliding linac beams, people were shocked and dismayed by the implication: Such linacs would have enormous size and be very expensive. This shock and dismay triggered the search for new acceleration methods. The eighties were full of workshops on new acceleration methods, using lasers one way or the other, wake fields of tightly bunched driving beams and other untested methods. All of these methods were characterized by extremely high gradients to keep the size of the collider small. But during those years it also became clear that there are functional relationships between luminosity, power consumption, beam strahlung and other important entities which are quite independent of the particular way by which particles are to be accelerated. The importance of the conversion efficiency, with which beam power is produced from utility power was recognized. These considerations led people back to linear accelerator technology, but still with unusually high gradients of 50 to 100 MVm^{-1}. The wish for high gradients persisted although it is clear that such solutions are generally far from an economical optimum: In an optimized linac all costs which scale linearly with length like accelerating structure costs, tunnel costs, cabling costs etc., equal those which scale with the amount of total rf-power, like modulators, klystrons and integrated power consumption over a certain amount of time. Such cost optimizations lead in general to much more modest gradients of 10 to 20 MVm^{-1}. High gradients can only be justified, if a collider has to fit on a given piece of land, if there are large political problems in building tunnels under privately owned land or if construction and tunneling costs are exorbitantly high.

At the beginning, linear colliders were seen as single bunch machines: After the accelerating structure of the linacs was filled with rf energy, only one single bunch of particles was to be accelerated in each of the two opposing linacs. In this mode power efficiency is very poor, because after acceleration of this single bunch in each of the two linacs the whole left-over rf energy stored in the accelerating wave guides is dumped. To keep this energy small, people chose very high linac frequencies, thereby reducing the stored rf energy.

The second reason why high frequencies were favored is the shunt impedance per unit length, which increases with the square root of the frequency in scaled accelerator structures, thereby reducing the power necessary to maintain a certain gradient.

A third reason, why higher frequencies might be adventageous, is the breakdown field strength in linac structures which seems to increase with frequency. Also dark currents, as they might be produced by field emission, require higher gradients at high frequency to be trapped, although the actual amount of dark current depends in a complicated way on field strength and surface finishing.

All of these arguments have some validity, especially if the only objective were to accelerate some small current to very high energies. But if the objective is to produce luminosity at high energies while at the same time keeping the beam strahlung background and wallplug power at acceptable levels, high frequencies may quickly loose some of their attraction and actually might show some severe drawbacks.

The important equations governing linear colliders are simple and their implications clear:

The luminosity L is given by $L = \frac{f\ N^2}{4\pi\sigma_x\sigma_y}$ (1)

N is the number of particles per bunch, f the number of bunch collisions per second. (Neglected here are effects from conceivable bunch-bunch-pinching and small beam-beam crossing angles.) σ_x and σ_y are the transverse beam dimensions (st. d.) at the interaction point.

The beam power is given by $P_B = E\ f\ N$ (2)
where E is the center of mass energy of the collider measured in eV.
The beam power P_B is related to the utility power $P_{ut.}$ by

$$P_B = \eta\ P_{ut.} \quad (3)$$

where η describes the total power conversion efficiency.
Combining equations (1) and (2) one can express the luminosity as

$$L \sim \frac{P_B\ N}{E\ \sigma_x\sigma_y} \quad (4)$$

From equation (4) several simple conclusions can be drawn:

1. The luminosity is directly proportional to the beam power. For a given utility power it is important to have high beam-utility power conversion efficiency.
2. It is desirable to have as many particles per bunch as possible. But N may be limited by higher order mode (HOM) excitation in the accelerating guide. The longitudinal effect of HOM excitation causes an energy spread within a bunch and in a bunch train from bunch to bunch. For scaled accelerating structures this spread is proportional to the square of the frequency. Such energy spread must be limited because of momentum acceptance of the final focus system. A low frequency rf-system is clearly advantageous.
   HOM effects in the transverse direction can cause a distortion of the bunch and thereby a larger effective bunch size. These transverse effects increase with the third power of the accelerating frequency. Higher frequencies require much tighter tolerances in beam-waveguide alignment. Therefore low accelerating frequencies are also clearly to be preferred here, particularly, if N is to be large.
3. For high luminosity the beam size $\sigma_x \cdot \sigma_y$ at the interaction point should be as small as possible.

From the point of higher order mode excitation low frequencies are clearly favored. As a matter of fact, these effects are so strong at X-band frequencies, that no study assuming X-band frequencies uses scaled S-band accelerating structures.The center hole of X-band structures is always considerably larger than that of a scaled S-band structure. This in turn reduces the shunt impedance per unit length almost to that of an S-band accelerating wave guide. The second argument given above in favor of high frequencies - that of higher shunt impedance at higher frequencies - is therefore not relevant. The third reason given - that of the higher breakdown field strength at higher frequencies - is meaningless: Field strengths in a cost optimized linear accelerator are in general much lower than those, at which breakdown occurs. For similar reasons field emission may also be a non-issue. What is left is the first argument of utility power to beam power efficiency, which originally was the strongest reason for the use of very high frequencies.

Here it is the relatively late realization, that high utility power/ beam power conversion efficiency can only be reached in a multibunch operation. In this mode of operation the accelerating waveguides are first filled with rf energy. Then many particle bunches are accelerated in an equally spaced bunch train. During that long beam pulse the rf power is used to maintain the gradient by compensating losses from beam loading and resistive wall heating. If the beam pulse is long compared to the filling time the energy stored in the waveguides becomes small compared to the energy transferred to the beam or going into wall losses. The energy going into wall losses is determined by the shunt impedance per unit length which, as we have seen above, is in an S-band structure comparable to that of a practical X-band structure. It is evident, that the utility power/ beam power efficiency for an S-band machine with long beam pulse is probably as good as that of any higher frequency machine and certainly better than that of machines with short beam pulses. The only machine with potentially even higher efficiency seems to be the superconducting colliding linac. These machines, too, have to be pulsed, albeit at a lower repetition rate. In the desirable but not feasible cw mode the rf power losses would require excessive cryogenic power. Superconducting linacs are therefore subject to the same types of power inefficiencies as normal conducting linacs: Loss of stored energy, as in all pulsed machines, and resistive wall losses, which are much smaller than those in normal conducting machines but, because they occur at 2° K, require substantial utility power for the cryogenic plant.

In a comparison between typical power efficiencies of S-band, X-band and superconducting linacs (1) numbers like 15 %, smaller than 11.4 % and 16.8 % were found for representative studies. All arguments which have been made so far are independent of the energy of the colliding linacs. There is no particular reason, why S-band technology should be less suitable for colliding linacs of very high energy. But there are arguments, why the advantage of the S-band machines is less pronounced for colliding linacs in the TeV-region:

The biggest problems for e^+-e^--machines in the TeV region are beam strahlung and luminosity. Beam strahlung can be understood as synchrotron radiation of particles at the interaction point produced by the extremely strong electromagnetic fields of the opposing bunch. This radiation causes energy losses, which make the center of mass energy between electrons and positrons less well defined. Furthermore, the hard photons created by such beam strahlung can create through further collisions with the opposing bunch severe background problems from pair production and hadronic mini-jet production. In the classical low energy case, where the characteristic photon energy from beam strahlung is small compared with the particle energy, the relative energy loss σ_E is given by

$$\sigma_E \sim \frac{N^2 \; E}{\sigma_{x^2}\,\sigma_z} \qquad (5)$$

or

$$\sigma_E \sim L\,\frac{E}{f\;R\;\sigma_z} \qquad (6)$$

where R is the aspect ratio σ_x/σ_y at the interaction point and σ_z is the bunch length. Since we must try to achieve a luminosity which is proportional to the square of the energy to keep the rate of interesting physics events constant for constant relative energy loss from beam strahlung, the product of $f{\cdot}R{\cdot}\sigma_z$ has to increase with the third power of the energy. Obviously one will try to make R as large as possible in order to keep beam strahlung small. But R may be limited to values smaller than perhaps 300 by imperfections in the linacs (betatron coupling, spurious vertical dispersion and filamentation) and optics limitations in the final focus. The desire for a long bunch length has to be balanced against the wish for a very strong focusing at the interaction point (depth of focus !) The bunch length should also be short in comparison to the rf wave length so as not to introduce additional energy spread in the accelerated beam. To keep the energy spread from beam strahlung small, the bunch collision frequency f will therefore have to increase with the third power of the energy. It can be shown, that the beam power too will have to increase with the third power of the collision energy. (This shows the immense problems for linear colliders in the TeV region.) The number of particles per bunch on the other hand will be inversely proportional to the beam energy, making problems of higher order mode excitation less severe at very high energies.

We have seen, that an S-band linear collider does not seem to have significant disadvantages compared with a higher frequency installation, but that on the other hand problems of higher order mode excitation clearly favor the lower S-band frequency, particularly at lower collision energies.

Another big advantage of the S-band technology of course is that it exists while the higher frequency X-band technology still requires a great deal of development work. This is particularly true for the high power klystrons, which do not yet exist at the higher X-band frequency with an acceptable power rating and efficiency. But the strongest reason for choosing S-band frequencies for the next linear collider is that of the existence of a prototype, which is not available for the higher frequencies:

In the history of high energy synchrotrons and storage rings particle energy for any new machine has almost never increased by more than a factor of 10. The extrapolation of known and optimized accelerator technology to the next generation machine has rarely been more than one order of magnitude in energy and new and detrimental machine physics effects could often be studied in existing accelerators before they became a threat to a new project. Linear colliders are now in a very unique situation: The first project to be built to produce new physics, beyond that which is accessible to storage rings, is already a huge commitment in money and man power. It would be irresponsible to propose such a project without the knowledge and input from a representative pilot project, which must be at least have 10 % of the size of the new proposed project. There are already a number of potential problems known today, which require experimental verification in such a pilot project, before one can propose the construction of a say 500 GeV cms linear collider:

1. The question of emittance preservation in the linacs is most important if one wants to be sure that the final spot size of the beams at the interaction point has the required and expected smallness to ensure the desired luminosity.The invariant emittance, in some studies assumed to be smaller than 3 orders of magnitude compared to numbers reached in the SLC, can easily grow by large factors through imperfections of the linacs (betatron coupling,spurious dispersion and filamentation, multibunch and single bunch instabilities and optics mismatches).
2. The background at the detector, produced by long Gaussian tails of the particle distribution in the bunches, by beam losses in the long linac or by dark currents produced by field emission in very high gradient machines, may be one of the key problems for doing good experiments at linear colliders.
3. Multibunch instabilities may be one of the most critical machine problems. Its suppression and avoidance through precision beam waveguide alignment, through higher order mode absorbers and through frequency variation between acceleration waveguides is most crucial for the next linear collider.

For an S-band collider such a 20 % pilot project exists and can be used to investigate those problems: It is the SLAC linac as used for the SLC project. But the extrapolation of those findings and results to the expected performance of a machine with 4 times the frequency and gradient is not possible. Here a new 10 % pilot machine is necessary, in itself a project comparable with the construction of the SLAC linac. Not only does a pilot project cost extra money, even if under fortuitous circumstances it could, with modifications, be part of the final machine, but its construction and evaluation will also add many years to the schedule of a 500 GeV project.

The question must be asked of course, whether the SLAC linac as operated in the SLC mode is a valid prototype for a 500 GeV collider. Surely this machine was not built to the much tighter specifications necessary for a colliding linac scheme. But two of the most important open questions mentioned above surely can be investigated without problems: That of emittance preservation and that of background. The third question, that of multibunch instability, is harder to investigate because SLAC has not been built with the required alignment tolerances nor with the frequency variation between transverse modes of the accelerating waveguides, which are deemed necessary to combat multibunch instabilities. When SLAC was built this type of instability had not been anticipated. Through deliberate but limited detuning of the already installed waveguides and through additional quadrupole focusing it was possible to suppress this instability in SLAC to a level sufficient for normal linac operation. Meanwhile, the problems of multibunch instabilities are sufficiently well understood and treated by computer simulations, such that a comparison between expected and observed beam behaviour in the SLAC linac should be possible. Although it will not be possible to modify the SLAC linac to the point, where it could be considered to be a fully fledged prototype of a larger colliding linac scheme, it would allow the computer programs and subsequent predictions, which form the basis for a 500 GeV cms design, to be checked.

TABLE 1
Table of S-band linear collider parameters

General Parameters				24 July 1992
energy	GeV	300	500	1000
luminosity	$(cm^2 sec)^{-1}$	$1.1 \cdot 10^{33}$	$4.0 \cdot 10^{33}$	$2.8 \cdot 10^{33}$
active length	m	17640	29412	29412
repetition rate	Hz	50	50	50
number of particles per bunch		$2.1 \cdot 10^{10}$	$2.1 \cdot 10^{10}$	$2.8 \cdot 10^{10}$
Main Linac				
wave length	m	0.10	0.10	0.10
average shunt impedance	MΩ/m	53.6	53.6	53.6
structure length	m	6	6	6
klystron power	MW	150	150	150
number of klystrons		1470	2450	4900
average power	MW	60	110	220
average pulse current	mA	300	300	400
current pulse length	μsec	2	2	0.6
number of bunches per pulse		172	172	50
bunch length (rms)	mm	0.5	0.5	0.5
maximum energy width (peak to peak)	%	0.3	0.3	0.2
Final Focus and Interaction				
β-function at IP $\beta^*_{x,y}$	mm	50, 0.8	16, 1	40, 0.5
beam dimension at IP $\sigma^*_{x,y}$	nm	914, 37	400,32	447, 16
total crossing angle	mrad	2	2	2
disruption parameter D_x, D_y		0.23 / 5.6	0.69 / 8.6	0.38 / 10.6
luminosity enhancement		1.4	1.8	1.4
energy spread $\frac{\Delta E}{E}$ cm	%	0.36	3.3	6.7
Efficiencies				
rf ⟶ beam	%	43	43	22
wall-plug ⟶ beam	%	14	14	8

Tab. I shows a parameter list for a 500 GeV cms linear collider at S-band frequency, which was worked out by the DESY/ Darmstadt collaboration (2). It is characterized by the relatively large beam power (2 x 7.5 MW), which allows the luminosity of $4 \cdot 10^{33}$ cm $^{-2}$ s^{-1} to be reached with a fairly large beam size at the interaction point (400 x 32 nm^2), values close to those aimed for in the Final Focus Test Beam FFTB at SLAC. The large beam size can be produced with relatively large invariant emittances ($5 \cdot 10^{-6}$ m and $5 \cdot 10^{-7}$ m). Actually, the assumed beam size of this collider study is so close to the values of the FFTB that a positive outcome of the FFTB test virtually guarantees that the spot sizes of the S-band study can be reached. The value of the vertical beta function at the interaction point of 1 mm allows the bunch length to be as large as .5 mm. This relatively large number makes the bunch compression between damping rings and linac much less demanding and at the same time reduces the beam strahlung background to very small values. The large beam size at the interaction point also makes the vibration and stabilization tolerances considerably less critical than those in colliding linac studies with much smaller beam sizes.

Also shown are parameter lists for center of mass energies of 300 GeV and 1000 GeV. The lower energy parameter list is interesting for work on the top quark, assumed to be in that energy range. One can see, that the total energy spread as given by machine energy spread from excitation of longitudinal higher modes and the additional spread from beam strahlung is less than .4 %. This number may need some interpretation: About half of the beam has an energy spread smaller than one tenth of that value. The other half has rather large energy tails, such that the calculated rms value becomes .36 %.

The parameter list for the 1000 GeV cms operation assumes a doubling of the number of klystrons and an increase of their effective peak power output by a factor of 2 through some SLED scheme. In such an arrangement it is not possible to increase the luminosity by the desired factor of 4 as compared to the 500 GeV operation. This factor could only be gained by a decrease of the beam spot size and/or an increase in beam power.

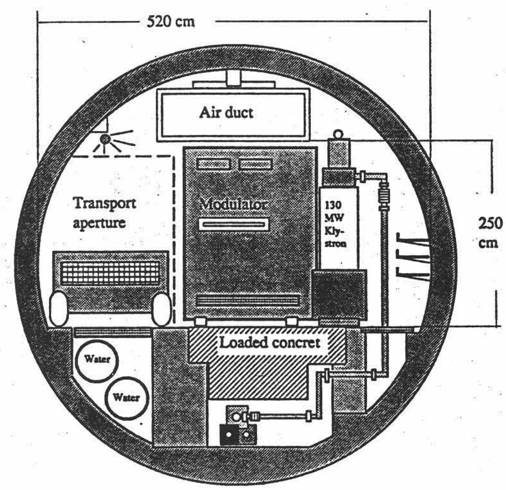

Fig. 1 Conceivable Tunnel cross section for a 500 GeV_{cms} Linear Collider

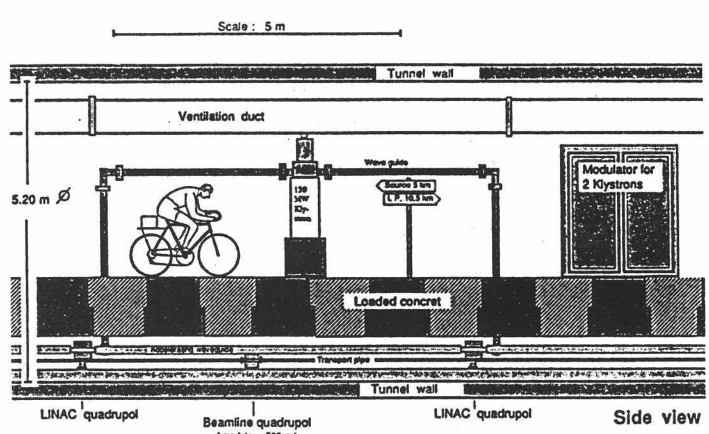

Fig. 2 Sideview of a 500 GeV_{cms} Collider Tunnel

Fig. 1 and Fig. 2 show you a possible schematic layout of the facility and a tunnel cross section. It is believed, that the whole 30 km long installation could be housed in a drilled tunnel. Shielding in the tunnel will be good enough to allow servicing of modulators, klystrons and auxiliary equipment while the machine is in operation. Access to the installation from the surface is only necessary at both ends and at the central collision point.

The most critical aspect of such an installation are the above mentioned multibunch instabilities. Computer studies show, that with a beam-waveguide alignment of better than 0.02 mm, one single transverse higher mode absorber near the front end of each of the 6 m long waveguides and a frequency variation from waveguide to waveguide of up to 10 MHz for the transverse modes, any emittance increase from transverse higher order mode excitation will be negligible. Construction of waveguides with such straightness and monitors to allow a beam-based alignment of such accuracy certainly are a technical challenge. Construction of a pilot project with 4 six meter long constant gradient waveguides powered by two 150 MW klystrons will therefore be an important step on the way to a proposal for a 500 GeV machine. Such a pilot project, presently under construction at DESY, will also involve vibration damping of waveguides and focusing quadrupoles and in particular will test components especially developed for low mass production costs.

We believe, that the technical problems of an S-band linear collider are understood to a large extent and that after a successful conclusion of the FFTB tests and other measurements at the SLC and the pilot project at DESY, a responsible proposal for a 500 GeV machine can be prepared. But it will also be essential that the costs of such a project be as low as possible and a good portion of the R & D effort between now and a proposal should address that question.

References

[1] Proceedings of the Linear Collider Workshop Saariselkä 1991, to be published

[2] Status Report of a 500 GeV S-Band Linear Collider Study, K. Balewski et. al. DESY 91-153, Dec. 1991