REVIEW OF LINEAR COLLIDERS

H. Henke
Technical University, EN-2
1000 Berlin 10, Germany

Introduction

The necessity of new acceleration methods with very high gradients for TeV linear colliders was recognised about ten years ago. A broad exploration of new techniques was then started whose status is summarized in many papers (see e.g. [1]).

During the last years it became then more and more apparent that new methods will need a very long development time, interesting physics can be done in the sub-TeV range and the main problem is not the generation of very high gradients but of an adequate luminosity. Today research and development work is concentrated on final focus design, beam dynamics, beam generation and radio-frequency (RF) acceleration. Although RF power generation is essentially classical the choice of parameters requires new designs and technologies. At SLAC (ref.2, Vlieks and Wilson), KEK (ref.2, Takeda) and Protviro[3,4] new power tubes are being developped in the 7 to 14 GHz and 100 MW range together with pulse compression units. Lawrence Berkeley (LBL) together with Lawrence Livermore (LLNL) Laboratory and CERN have based their work on the two-beam accelerator (TBA) concept to avoid many thousands of power tubes. The LBL/LLNL group is studying a multi-stage free-electron-laser(FEL) at 17 GHz driven by an induction linac, while CERN design foresees a drive linac with 30 GHz RF "transfer structures" and superconducting reaccelerating cavities. Finally, DESY (ref.2, Voss) started to consider a multibunch machine at standard frequency (3 GHz) and gradient (17 MV/m).

Parameter Constraints

The basic difficulty is not the generation of the total voltage, being "only" a question of money and space available, but of the required luminosity. Since electroweak cross sections scale with the inverse square of the particle energy the luminosity should increase with the energy squared and exceed $10^{33}\,cm^{-2}\,s^{-1}$ at 1 TeV per beam. Following ref. [6] the luminosity in a head-on collison is

$$L = \frac{N^2 H_x H_y f}{4\pi\sigma_y^2 R} \qquad (1)$$

where N is the bunch population, f the repetition rate, $R = \sigma_x/\sigma_y$ the beams aspect ratio and H_x, H_y are pinch enhancement factors. With the fractional energy loss δ due to beam-beam radiation one can express the repetition rate as

$$f = 1.4\cdot 10^{-29}\ cm^2 \left(\frac{H_T T}{\delta}\right)^2 \frac{4RH_xH_y}{(H_x+RH_y)^2}\, L \qquad (2)$$

where H_T is the reduction factor for quantum effects depending on T only. T is the ratio of average critical photon energy over particle energy. The product $H_T T$ has a very flat maximum, around 0.2, for all reasonable values . Therefore, with the maximum tolerable value $\delta \approx 0.3$, one gets $\delta/H_T T \approx 1.6$ and the repetition rate, equ.(2), becomes

$$f = 5\cdot 10^3\,s^{-1} \frac{4RH_xH_y}{(H_x+RH_y)^2} \qquad (3)$$

for $L = 10^{33}\,cm^2\,s^{-1}$. For round beams the repetition rate has to be about 5kHz to provide the required luminosity. For flat beams, $R \gg 1$, the radiation loss is decoupled from the luminosity per bunch and the repetition rate can be reduced according to $f \sim R^{-1}$. Aspect ratios of about 100 ($H_x \approx 1$, $H_y \approx 2$) are needed to bring f down to the order of 100 Hz.

The above arguments imply, that in high rep-rate machines beams with small R are preferable, and that in low rep-rate machines ($f \approx 100$ Hz) R and N/σ_y have to be about two orders of magnitude higher. The resulting beam heights are in the order of nm at collision.

The second difficult problem is the emittance preservation during acceleration and transport. The sub-μm beam dimensions prescribe very small emittances. The β-values at collision can not reduced under

$$\beta \gtrsim \frac{\Delta p}{p}\ell \qquad (4)$$

with the ℓ free space between the last quadrupole and the intersection point and $\Delta p/p$ the particle momentum spread. Additionally, in the case of very strong focusing, the final beam spot is radiation limited[7] to

$$\sigma_y \approx 3\cdot 10^{-4}\ \varepsilon_{ny}^{5/7} \qquad (5)$$

where ε_{ny} is the normalized emittance. As a result the emittances are in the range $5\cdot 10^{-7}$ to 10^{-8} m for the different concepts. This is several orders of magnitude smaller than the value achieved in SLC and requires extreme care in the transport channel.

Apart from concerns with emittance preservation, people were troubled about the generation of such low emittances, but since recently progress with damping rings, $\varepsilon \approx 3\cdot 10^{-7}$m has been achieved, confidence that less than 10^{-7} m is not an unsurmountable obstacle has risen. Today, all study groups have paper solutions for an adequate damping ring.

The third big problem is the generation of the necessary RF peak power. Although it is not primarily a basic problem, it is nevertheless very severe in terms of money, reliability and space available. Following the different aims, 250 GeV or 1 TeV per beam, and different philosophies, RF power tubes or distributed power sources (TBA devices), the envisaged solutions cover a large area of R + D work. The choices made are a result of many interconnected and partially conflicting requirements as well as feelings born from a certain well established spirit and experience.

Until recently, belief in high gradients, $E_{acc} \approx 80$ MV/m, and consequently into RF frequencies higher than 3 GHz was common since the RF peak power scales with the square of the gradient but with the square root of the wavelength for a fixed gradient. This belief required the development of new RF power sources. A large effort was out therefore into the development of power tubes together with pulse compression devices, aiming at a tube in the 9 to 12 GHz range and a power output of a few 100 MW. Much work was also invested into the construction of an induction linac driven FEL. 1.8 GW of radiation at 35 GHz have been obtained[8] but phase stability was poor. Another approach studies 30 GHz radiation from a bunched beam passing a travelling wave structure. None of these efforts have yet produced a satisfying result.

Recently DESY (ref.2,Voss) in a joint effort with the University of Darmstadt Germany, started to think of a conventional low gradient low frequency RF system.Assuming a somewhat lower energy, 250 GeV per beam, a gradient of less than 20 MV/m was identified as the cost minimum when trading-off RF equipment and civil engineering costs. The handicap is a long machine with many bunches. This machine is nearly identical to the proposed superconducting (sc) linacs apart from the gradient which is lower by a factor2. Since sc-machines have fill-factors of 0.5, the effectiv gradient is about the same with the advantage of a cheaper and simpler technology.

The boundary conditions for the study groups differ in terms of the desired machine energy, the site available and the planned time schedule. They influence, together with the traditions and experience of the laboratory, the envisaged technical solutions and therefore the machine parameters. Table 1 shows the specific parameters of the different studies.

Tabl 1. Tentative parameter list of linear collider studies

		DESY	SLAC ILC	KEK JLC	Protvino VLEPP	CERN CLIC
c.m.energy	TeV	0.5	0.5	1	1	2
luminosity	$10^{33} cm^{-2} s^{-1}$	1.9	1.7	6	1	1
acc.gradient	MV/m	17	93	100	83	80
acc.frequency	GHz	3	11.4	11.4	14.3	30
repetition rate	Hz	50	360	200	100	1700
no. of bunches/pulse		172	10	10	1	1
bunch population	10^9	7	10	10	>100	5
bunch length	mm	0.2	0.05	0.08	0.7	0.2
nom.vert.emittance	10 μm	6	2	3	6	50
beam height (collision)	nm	7.8	3	1.4	10	12
aspect ratio (collision)		25	100	140	100	5

Accelerating Structure

The most efficient accelerating structure is still the familiar disc-loaded waveguide. It is a travelling-wave structure with the advantage of a common, matched feedpoint for many cells. The power is provided reflection-free in form of short pulses whose duration is given by the length of the structure and the filling speed, controlled by the disc holes. Length and aperture are calculated by trading off dissipation during filling time and effective gradient for a given peak power.

Higher gradients can be achieved at shorter wavelength,since the power is essentially focused into a smaller volume. The scaling for peak- and average power is

$$\hat{P}_{RF} \sim \lambda^{1/2}, \quad \bar{P}_{RF} \sim \lambda^2 \qquad (6)$$

for a fixed gradient. However, beam induced fields (wakefields) per unit length increase with the second and third power of the frequency for longitudinal and transverse component respectively. That means the energy spread increases (bad for final focus) and the transverse stability rapidly deteriorates. Since the transverse wakefields also depend on the aperture with roughly the third power, structures with frequencies higher than 3GHz will have two times larger ratio of aperture to wavelength in order to compensate, at least partially,the wakefield's f-scaling.

The study of different fabrication methods has yielded, that the best is still vacuum brazing of a stack of machined copper cups. The required tolerances and surface finish do not exceed the possibilities of modern machine tools. This applies even for the 30 GHz CERN structure, which is machined to 1 µm tolerances and Nl surface finish on a special diamond-tool lathe. Microwave measurements on a test-stack showed a dispersion diagram in agreement with computer values to within a few ‰ and cell-to-cell phase errors below 1°.

The high accelerating gradients considered for most machines may cause surface deterioration by arcing or by thermal stresses due to the pulsed operation. Very serious problems may also arise by field emission currents. They can destroy surface, the affect field symmetry and, thus generate random deflections which blow-up the emittance. Therefore, the effects related to high gradients are studied at SLAC, KEK and LAL-Orsay.

Single bunch linear colliders are very inefficent. The mains to RF conversion efficiency is typically 20 %. Only a few % of the energy stored in the accelerating structure can be extracted by the bunch, the limit being the tolerable energy spread, resulting in an overall efficiency of about 0.5%. Nearly one order of magnitude increase might be obtained by multibunch operation, since,the extracted energy is multiplied by the number of bunches, while the dissipation is increased by the ratio of beam pulse duration to filling time only. The problem is avoiding bunch-to-bunch deflection (beam break-up). With higher frequency machines this means heavily de-queueing of higher modes by radial slots in every disc, Fig.1. The slots are coupled to waveguides which direct any higher mode power into loads.

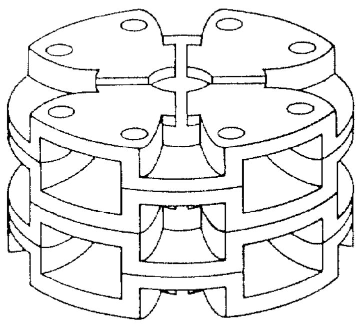

Fig.1 Conceptual design of a stack of four accelerating cells with radial damping lines.

For lower frequency machines damped cavities may not be necessary. Instead, bunches could be placed at the field minimum of two beating transverse modes resulting from two slightly different structures.

RF Power Generation

Apart from the DESY design all machines use frequencies higher than 3 GHz and high gradients. Power requirements vary between 100 and several 100 MW/m. These sources are not available. Three directions of development are being pursued: Discrete sources, distributed sources and the TBA.

Discrete Sources

With discrete sources we mean individual power converters, typically an electron-beam tube, together, eventually, with a pulse compression device. The "workhorse" in accelerator technology is still the old klystron. Essentially, it consists of a gap formed by a cathode and an anode to generate and accelerate the beam, an input cavity to impose a velocity modulation, a drift space to allow for bunching, an output cavity where the bunched beam radiates coherent microwaves and a collector as a beam dump. The beam energy is limited by voltage breakdown in the gun area, the beam current by space charge and cathode area which tends to scale with λ^2 . The peak power output, the product of current times voltage induced in the output cavity, scales, therefore, roughly with λ and it seems unlikely that conventional klystrons could supply high peak power at short wavelength. Several ideas try to overcome the cathode area limitation, e.g. by cathode clusters, ribbon beams or by using a somewhat different principle like a gyro- klystron or a crossed-field generator. Each solution has its proper difficulties such as unsufficient beam stability or RF phase stability.

Probably, the best what could be expected in the near future is a working X-band klystron with a power rating between 50 and 100 MW. This power will have to be increased by a pulse compression device. The device used routinely today is the SLAC energy doubler (SLED), based on the observation that power radiated from a heavily overcoupled cavity is several times the incident power immediately after having switched off the incident wave. A further increase by two can be gained if incident radiation is reversed in phase rather than simply switched off. In practice, the device reaches a power multiplication of about 3. Another technique, called binary pulse compression[10] (BPC), can theoretically give arbitrary gain in steps of a factor two. Two RF sources, Fig. 2a, feed long pulses into a hybrid coupler. The sources' phases are such that they combine at the first coupler output which is connected to a low-loss delay line. Halfway through the pulse, the phases are changed and the combined power exits the second output of the coupler. Both prompt and delayed pulses have half-length with twice power and can be directed to a subsequent stage. The device has passed low power tests successfully and high power test will be done. Recently, SLAC (P.B. Wilson, ref.2) and Protvino (VLEPP) have picked up an idea published elsewhere[11]. In this scheme, called SLED-II, the SLED storage cavities are replaced by short-circuited delay lines, Fig.2 b. The advantage is a stair-case-like output pulse with a flat top.

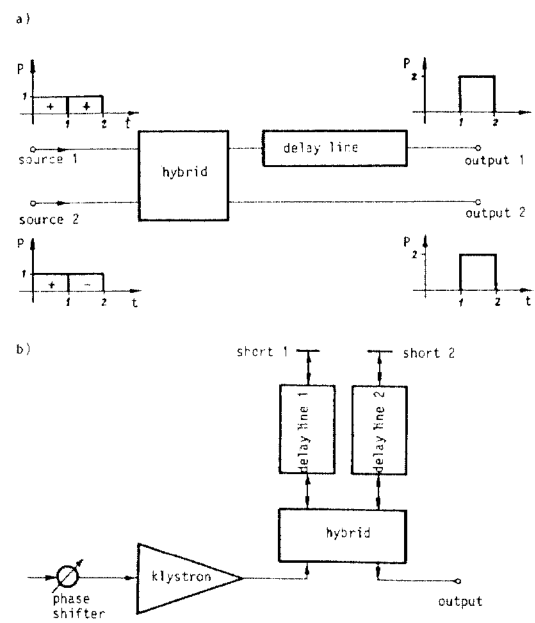

Fig. 2 Diagrams of RF pulse compressors: a) single-stage binary pulse compressor (BPC), b) SLED-II compressor.

Distributed Sources

To avoid an excessive number of discrete sources, reusing the high intensity source beam in a periodic multi-stage device running parallel to the main linac, has been proposed, Fig. 3.

The LBL/LLNL group[12] studies an alternating arrangement of FEL and induction linac units. The RF is generated by FEL radiation, of which a large fraction is extracted and fed to the main linac. The left-over RF serves as phase reference in the next FEL unit. Between the FEL units, the beam is reaccelerated by induction units fed via magnetic pulse compressors. Among other reasons, the relation between RF wave length and beam energy, limits the number of subsequent stages. The electron energy has to be of the order of 10 MeV where phase slip effects and transverse beam stability do not allow more than a handful of stages. The main problem areas to be studied are sideband instabilities, RF phase control, RF extraction and stable parameter dependence.

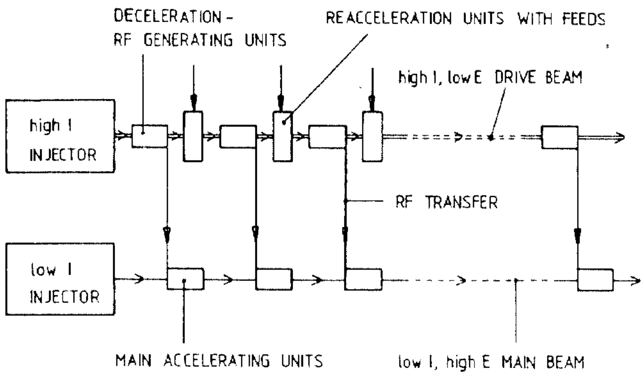

Fig. 3 Schematic drawing of a two-beam accelerator.

Two-Beam Accelerators

In a true TBA, as proposed by A. Sessler[13], the drive-beam has to be highly relativistic, so that it can stretch all along the main linac. Deceleration and reacceleration of the drive-beam are done adiabatically so the distributed RF source is in a steady state.

In a solution, which is being studied at CERN[5], Fig. 4, the RF is generated by a train of bunches passing through 30 GHz travelling-wave (TW) structures. Four trains separated by the structure group delay (the time to empty the structure), make up for one RF pulse. The bunch trains are reaccelerated in sc-cavities, ringing at 350 MHz, which are fed by cw, low peak-power, high-efficiency klystrons.

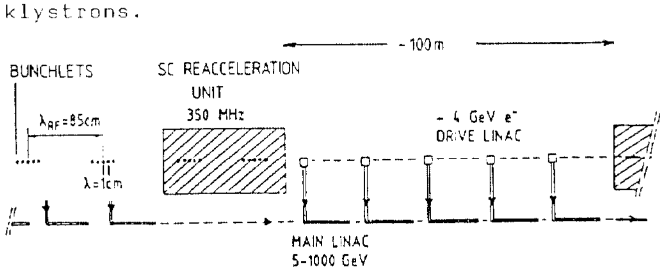

Fig. 4 Two-beam accelerator with superconducting cw drive linac and a microwave main linac.

Problem areas are the wakefield effects induced by the intense drive beam in the surrounding walls and generation and preacceleration of the multibunch beam containing about 10^{12} electrons in mm-long bunches. To study these problems a test facility is being built at CERN, consisting of a laser-driven photocathode in an RF gun and magnetic bunch compressors.

Wakefields and Transverse Alignment[14]

As mentioned above, a main concern is the preservation of the small emittances during acceleration. The reasons for spoiling the emittance are manifold. Jitter motion, i.e. uncorrectable errors from pulse to pulse, will occur at injection and for different machine components. As a result, the beam ellipse jumps around in phase space. The emittance becomes diluted. Since jitter motion are uncorrectable they have to stay below certain values which are directly connected to the beam dimensions. On the other hand, static alignment errors can, at least in principle, be measured and corrected. Tolerances on these errors follow from the allowable chromatic variation of the particles trajectory and are typically more than an order of magnitude looser than for jitter motion. Another reason for emittance dilution is the inevitable energy spread in the bunch generated by longitudinal wakefields. Although this effect can be minimized by making the bunch short and the physical aperture large, it is counter-productive for high frequency machines. These machines need an appreciable energy spread in order to control the transverse wakefield effects. Any bunch with an accidental transverse offset induces transverse fields which tend to deflect the bunch core and tail even further away. The field's strength scales with f^3 and is therefore particularly cumbersome at high frequencies. In that case, the wakefields have to be reduced by choosing a longer wavelength to aperture ratio and have to be controlled by BNS damping[15], a method where defocusing wakefields are counter-balanced by a focusing increasing from bunch head to tail due to a negative energy spread. Instead of providing BNS damping via energy spread, a more powerful method was proposed based on RF quadrupoles. Such a focusing system acts on the bunch scale and can introduce a larger spread in focusing than via energy spread.

Computer simulations and simple two-particle model calculations have shown that transverse jitter motion of quadrupoles has to be in the order of 0.1 µm r.m.s. value Slowly varying errors with larger amplitude have to be corrected on-line. Fast servo-controlled beam steering and servo-controlled alignment remain a necessity even though BNS damping is applied. The required tolerances differ very much between higher and lower frequency machines and the necessary spread in focusing in order to provide optimal damping. Two damping situations, weak and strong, can be distinguished. At lower frequency only weak damping, and therefore a small energy spread, is required. The bunch performs relatively large amplitude betatron motions over many wavelengths. Transverse alignment tolerances for quadrupoles are typically in the 10 µm range and can be as high as 50 µm provided the machine is corrected chromatically. At higher frequencies strong damping is necessary within a few betatron wavelengths. Head and tail particles follow essentially incoherent trajectories while the bunch core oscillates coherently within the same limits. In this case misalignment errors have to be kept below the beam height everywhere, about a few microns.

While multibunching is an obvious way to increase the efficency it has also a series of drawbacks. The bunch-to-bunch energy spread has to be smaller than in a single bunch with BNS damping, a difficult task under beam loading and non-flat RF pulses. The transverse dynamics are dominated by the beam induced first transverse mode in the RF structure. It has to be either damped down to Q-values in the order of 10 and/or tuned so that bunches pass at zero field crossings. Another idea[17] is to build two different structures, identical for the fundamental accelerating mode but slightly different for the first transverse mode, so the transverse fields are beating. By tuning the beat frequency properly, destructive interference can be created at bunch positions.

Final Focus System and Test Facility

At the collison point the beams must be focused to tiny spots in the order of 1 nm or 10 nm depending on the different designs. There are several restrictions for the minimum achievable β-function at collision

$$\beta = \frac{\sigma}{\sigma_0} \ell \qquad (7)$$

with σ_0 and σ the beam dimensions before the last lense and at collision and ℓ the focal length, i.e. the free space on either side of the interaction point. At first, β must be much larger than the bunch length σ_z in order to have the whole bunch within the waist. This limitation is usually fulfilled. Secondly, in a system with no chromatic correction, particles with a different momentum have different focal points, or in other words, the spot size at the collision point is "fuzzed out". The resulting β-function is larger than the relative momentum spread times the focal length, see equ. (4). With a chromatic correction, a factor of 8 may be gained in a design similar to the SLC, so $\beta/\ell > 0.12\ \Delta p/p$ and demagnification factors of the order of 10^{-3} should be possible for $\Delta p/p < 1\%$. Thirdly, in case of excessively strong focusing, the final spot size is blown-up due to quantized radiation, equ. (5).

The required demagnification is one over a few hundred. Following the SLC scheme, this is done with a telescope. The chromaticity is corrected, for horizontal and vertical planes separately, by sextupoles placed in regions of finite dispersion. The dispersion is created by weak and very long bending sections to limit emittance growth by synchrotron radiation.

The final quadrupole can in principle be of conventional design, probably permanent magnets in conjunction with soft steel poles. To achieve the required gradients with poletip fields of little more than 1 Tesla, the aperture has to be very small, half-millimetre radius or less. Alignment precision and vibration amplitudes have to be a fraction of the beam spot size at collision. The corresponding techniques have yet to be developed. Beam-beam deflection and beamstrahlung may be the appropriate diagnostic tools, and micromovers and passive or even active vibration isolations would steer the magnet and keep it in place.

There is a project under way at SLAC (ref.2, Burke) to build a scaled down version of the next linear collider (NLC) final focus, called final focus test facility (FFTF). This experiment's goal is to produce a focal point of 1×0.06 μm with a bunch of 10^{10} electrons. It will be necessary to develop new alignment techniques and beam diagnostic instrumentation to achieve this goal. The experiment is a joint effort between SLAC, KEK, Protvino, LAL-Orsay and DESY.

References

[1] A.M. Sessler, Physics Today, Jan. 1988, pp. 2-10

[2] S. Takeda; G. Voss; A.E. Vlieks; P.B. Wilson ; presented at the 2nd Internat. Workshop on Next Generation Linear Colliders, KEK, Japan, march 28 - april 5, 1990.

[3] L.N. Arapov et.al., VLEPP-Note-05/1990, Protvino, Russia, 1990.

[4] V.E. Balakin and I.V. Syrachev, VLEPP-Note-06/1990, Protvino, Russia, 1990.

[5] W. Schnell, CERN/LEP-RF/88-48, Geneva, Switzerland, 1988.

[6] W. Schnell, SLAC/AP-61, Stanford, USA, nov. 1987.

[7] K. Oide, Phys. Rev. Letters 61 (1988), pp. 1713.

[8] T. Orzechowski et al., Phys. Rev. Lett. 58 (1986), pp. 2172.

[9] R.B. Palmer, SLAC-Pub-4542, Stanford, USA, 1988.

[10] Z.D. Farkas, IEEE Trans. MTT-34 Oct. 1986, pp. 1036.

[11] F. Fiebig and C. Schieblich, Proc. European Particle Accelerator Conf., Rome, Italy, 1988, pp.1075.

[12] A.M. Sessler, D.H. Whittum and J.S. Wurtele, LBL-27765, Lawrence Berkeley Laboratory, Berkeley, USA 1989.

[13] A.M. Sessler, AIP Conf. Proc. No. 91, American Institute of Physics, New York, 1982, pp.154.

[14] H. Henke, 1988 Linear Accelerator Conf. Proc., Newport News, USA, 1988, pp. 481.

[15] V.E. Balakin, A.V. Novokhatsky and V.P. Smirnov, presented at 12th Internat. Conf. of High Energy Accelerators, Chicago, USA, 1983, pp. 119.

[16] W. Schnell, CLIC Note 34, CERN, Geneva, Switzerland, 1987.

[17] D.U.L. Yu and P.B. Wilson, SLAC-PUB-5062, Stanford, USA, 1989.