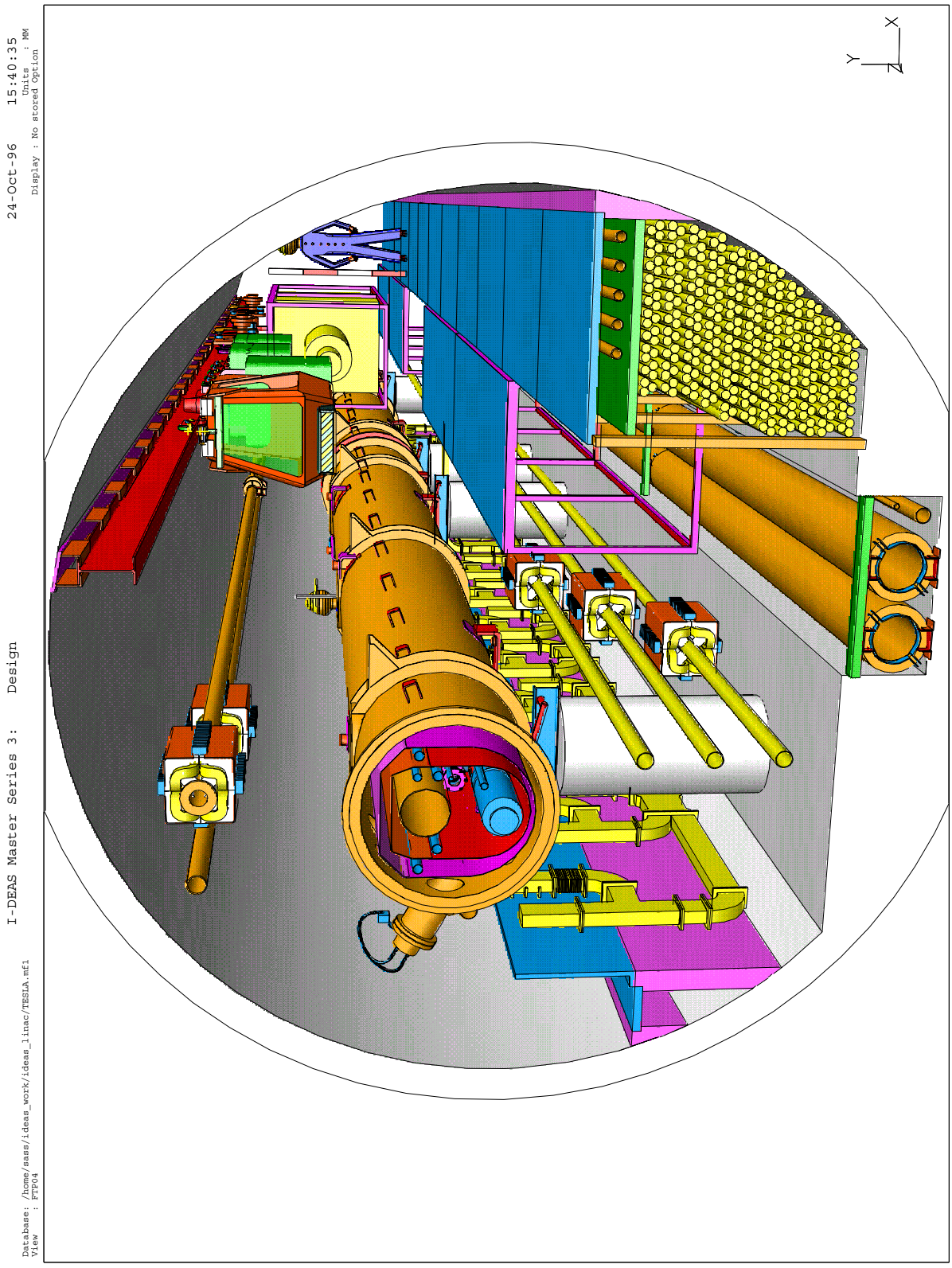
Chapter 3

TESLA Superconducting Linear Collider

TESLA Collaboration


24-Oct-96 15:40:35
Units : MM
Display : No stored Option
I-DEAS Master Series 3: Design
Database: /home/sass/ideas_work/ideas_linac/TESLA.mf1
View : FTP04
Y
X
Z

Author List

DESY (coordinating Institute), Notkestr. 85, D-22603 Hamburg: I. Altmann, R. Bacher, R. Bandelmann, W. Bialowons, W. Bothe, R. Brinkmann (editor), H.-D. Brueck, H. Burmeister, M. Clausen, G. Csuka, G. Deppe, H. Dinter, A. Drozhdin (now at FNAL), B. Dwersteg, J. Eckoldt, D. Edwards, U. Engelke, K. Escherich, W. Eschricht, K. Flöttmann, P.D. Gall, A. Gamp, R. Glantz, A. Goessel, G. Grygiel, K. Hanke, O. Hensler, R. Hensler, G. Hoffmann, N. Holtkamp, G. Horlitz†, D. Hubert, J.P. Jensen, H. Kaiser, R. Kaiser, U. Knopf, R.-D. Kohaupt, G. Kreps, R. Lange, M. Leenen, H. Lierl, T. Limberg, F. Löffler, A. Matheisen, G. Meyer, W.-D. Möller, M. Pekeler, O. Peters, B. Petersen, P. Pillat, J. Pflüger, D. Proch, K. Rehlich, D. Renken, D. Reschke, H.-O. Roggenbuck, J. Roßbach, J. Rümmler, M. Sachwitz, T. Schilcher, M. Schmitz, P. Schmüser, B. Schoeneburg, H. J. Schreiber, S. Schreiber, D. Schulte (now at CERN), W. Schwarz, M. Seidel (now at SLAC), J. Sekutowicz, D. Sellmann, S. Simrock, W. Singer, K. Sinram, B. Sparr, M. Stolper, T. Stoye, K. Tesch, D. Trines, F.R. Ullrich, N. Walker, R. Wanzenberg, H.P. Wedekind, G. Weichert, H. Weise, J. G. Weisend, B. H. Wiik, S. G. Wipf, K. Wittenburg, J. Wojtkiewicz, S. Wolff, K. Zapfe-Düren;

Academy of Mining and Metallurgy, Krakow: K. Pieczora, Z. Sanok;

Beijing Institute of Vacuum Electronics: Lu Fuhai;

Budker INP, 670090 Novosibirsk: B. Grishanov;

Budker INP, Protvino Branch, 142284 Protvino, Moscow Region: V. Balakin, A. Sery;

CE Saclay, F-91191 Gif-sur-Yvette Cedex: B. Aune, S. Chel, J. Gastebois, M. Juillard, A. Mosnier, O. Napoly, A. Novokhatsky (on sabbatical leave from INP Novosibirsk), J.M. Rifflet;

CERN: D. Bloess, G. Cavallari, R. Fortin, E. Haebel;

Cornell University, Ithaca, NY 14853: H. Padamsee, M. Tigner;

FNAL, P.O. Box 500, Batavia, IL 60510: M. Champion, E. Colby, H. Edwards, D. Finley, M. Kuchnir, T. Nicol, T. Peterson, V. Shiltsev;

FZ Karlsruhe, D-76021 Karlsruhe: K.P. Jüngst, H. Salbert;

Helsinki University of Technology, Dep. of Mathematics, Otakaari 1, SF-02150 Espoo: E. Somersalo;

IHEP, 142284 Protvino, Moscow Region: P. Chevtsov, V. Gubarev, S. Goloborodko, O. Kourneev, M. Maslov, V. Sytchev, Y. Tchernoousko;

IHEP Beijing, PR China: Kong Xiang-Cheng, Yi Ding;

INFN Frascati, Via E. Fermi 40, I-00044 Frascati: M. Castellano, M. Ferrario, M. Minestrini, P. Patteri, F. Tazzioli;

INFN Milano, L.A.S.A., Via Fratelli Cervi 201, I-20090 Segrate: D. Barni, A. Bosotti, D.

Giove, P. Michelato, C. Pagani, P. Pierini, L. Serafini, G. Varisco;

INFN Sezione Roma2, Via della Ricerca Scientifica 1, I-00133 Roma: L. Catani, S. Tazzari;

INP Krakow: J. Baran, A. Bienkowski, M. Talach;

I.P.N. Orsay, F-91406 Orsay Cedex: S. Bühler, T. Junquera;

Institute of Modern Physics Lanzhou, PR China: , Ang Zheng-Ting;

LAL, F-91405 Orsay Cedex: R. Chehab, J. Gao, T. Garvey;

Max Born Inst., Postfach 1107, D-12474 Berlin: W. Sandner, I. Will;

Polish Academy of Science, Inst. of Physics, Al. Lotnikow 32/46, Pl-02-668 Warsaw: J. Krzywinski;

Rolf Nevanlinna Inst., University of Helsinki, P.O Box 4, SF-00014 Helsinki: J. Sarvas, P. Ylä-Oijala;

SEFT, Siltavuorenpenger 20c, SF-00170 Helsinki: R. Orava;

SLAC, P.O. Box 4349, Stanford, CA 94309: P. Emma;

Soltan Institute, Krakow: T. Plawski;

Tampere University of Technology, P.O. Box 692, SF-33101: A. Koski;

TH Darmstadt, Inst. f. Hochfrequenztechnik, Schloßgartenstr. 8, D-64289 Darmstadt: U. Becker, P. Schütt, M. Timm, T. Weiland;

Thomas Jefferson National Accelerator Facility, 12000 Jefferson Av., Newport News, VA 23606: P. Kneisel, C. Rode;

TU Berlin, Einsteinufer 17, D-10587 Berlin: H. Henke, R. Lorenz (now at DESY), I. Reyzl;

TU Dresden, Inst. f. Kälte u. Kryotechnik, D-01062 Dresden: M. Kauschke, H. Quack;

Johann-Wolfgang-Goethe Universität, Inst. f. Angewandte Physik, Robert-Mayer-Str. 2-4, D-60325 Frankfurt: H.-W. Glock, P. Hülsmann, H. Klein, W.F.O. Müller, C. Peschke;

Universität Wuppertal, FB Physik, Gauss-Str. 20, D-42119 Wuppertal: G. Müller, H. Piel;

University of California, Adv. Accelerator Physics Department, Los Angeles, CA 90024-1547: J. Rosenzweig;

Contents

3.1 Overview

3.1.1 Introduction

As it has been discussed in chapter 1 of this report, studying electron-positron collisions at energies well beyond the reach of LEP ($E_{cm} \approx 200\,\mathrm{GeV}$) has a high potential for the future development of Particle Physics, in many respects complementary to the LHC. There is broad agreement in the High Energy Physics community that a linear e^+e^- collider with an initial center of mass energy of 350...500 GeV and a luminosity above $10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ should be built as the next accelerator facility.

The feasibility of a linear collider has been demonstrated by the successful operation of the SLC, the only existing facility of this kind. Nevertheless, meeting the requirements for a next generation linear collider is by no means an easy task. In particular, high beam powers and very small spot sizes at the collision point are needed in order to obtain a sufficiently high luminosity. Several groups worldwide are pursuing different linear collider design efforts [1]. The fundamental difference of the TESLA approach compared to other designs is the choice of superconducting accelerating structures. The challenge of pushing the superconducting linac technology to a high accelerating gradient and at the same time reducing the cost per unit length, both necessary in order to be competitive with conventional approaches, is considerable, but the advantages connected with this technology (as summarized below) are significant and we are convinced that the potential for the machine performance is unrivaled by other concepts.

TESLA uses 9-cell Niobium cavities (Fig. 3.1.1) cooled by superfluid Helium to T=2K and operating at L-band frequency (1.3 GHz). The design gradient for a 500 GeV collider is g=25 MV/m with an unloaded quality factor of $Q_0 = 5{\cdot}10^9$. The design of the accelerating structures is described in detail in section 3.2.1.1. This technology provides several important advantages for the design of a linear collider. The power dissipation

	S-band	X-band	Two-beam	TESLA
Accelerating gradient [MV/m]	17	29...55	78	25
RF-frequency [GHz]	3	11.4	30	1.3
RF-peak power [MW/m]	12	50...100	144	0.2
Av. beam power [MW]	7.3	4	4	8.2
Beam pulse length [μs]	2	0.1	0.006	800
Bunch spacing [ns]	6	1.4	0.67	708
Vertical spot size at IP [nm]	15	5	7	19
Luminosity [$10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$]	5	6	5	6

Table 3.1.1: *Parameters for different 500 GeV linear collider designs. A more detailed discussion of the TESLA parameters is given in section 3.1.3.*

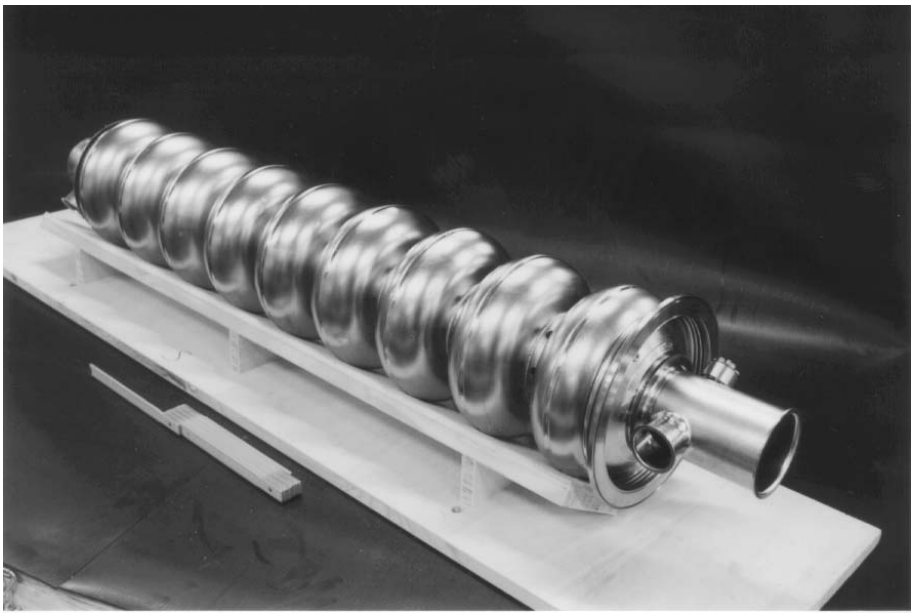


Figure 3.1.1: *The 9-cell Niobium cavity for TESLA.*

in the cavity walls is extremely small which allows to produce the accelerating field with long, low peak power RF-pulses and yields a high transfer efficiency of RF-power to the beam. With a high average beam power, the required luminosity can be achieved with a spot size at the interaction point (IP for short) only moderately (about a factor of 3.5) smaller than what has been achieved at the Final Focus Test Beam experiment performed at the SLAC linac [2]. At the same time the AC power consumption remains within acceptable limits (100 MW). The long RF-pulse allows for a large bunch spacing (see Table 3.1.1), making it easy for the experiment to resolve single bunch crossings. In addition, a fast bunch-to-bunch feedback can be used to stabilise the orbit within one beam pulse, which makes TESLA practically immune to mechanical vibrations which could otherwise lead to serious luminosity reduction via dilution of the spot size and separation of the beams at the IP. Further benefits of the long pulse are the possibility to use a head-on collision scheme with large-aperture superconducting quadrupoles in the interaction region and to employ a safety system which can "turn off" the beam within one pulse in case an emergency is indicated by enhanced loss rates.

The choice of a low drive frequency for TESLA results in very small transverse and longitudinal wakefields in the accelerating structures. This leads to relatively relaxed alignment tolerances for the linac components required for the transportation of a low emittance beam. Beam dynamics issues are discussed in detail in section 3.3, but it is instructive here to compare the different linear collider design concepts on a basis of simple scaling arguments [3]. One of the most essential contributions to emittance dilution results from short-range transverse wakefields due to random offsets of the accelerating structures with respect to the beam orbit. The emittance dilution from this effect can be written as

$$\frac{\Delta\epsilon}{\epsilon} \propto F \cdot \bar{\beta} \cdot \delta y_c^2 \tag{3.1.1}$$

where $\bar{\beta}$ denotes the average β-function in the linac (the stronger the focussing, the smaller $\bar{\beta}$), δy_c the rms-offset of the structures and F the dilution factor which depends on the beam parameters and very strongly on the linac frequency f_{RF}:

$$F \propto \frac{N_e^2 \sigma_z f_{RF}^6}{g^2 \epsilon_y} \tag{3.1.2}$$

The considerable variation of F for different linear collider designs is shown in Fig. 3.1.2. It becomes clear that TESLA can afford very much relaxed requirements for the alignment tolerances and for the beam optics. We consider it a crucial point to have a large safety margin concerning beam stability in order to be able to guarantee efficient and stable operation of such a future facility. In addition, part of the safety margin can be used to upgrade the machine performance beyond the "standard" requirements for the next generation linear collider. This point is discussed in more detail in sections 3.1.3 and 3.1.4.

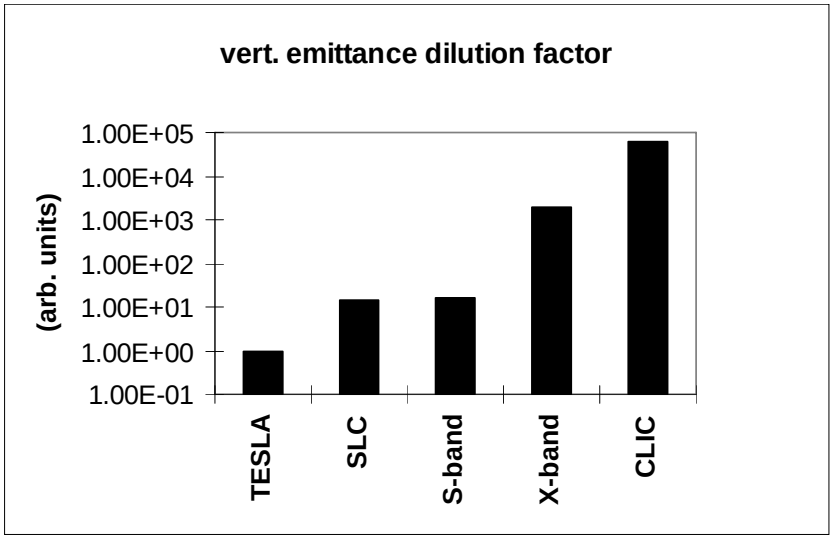


Figure 3.1.2: *Wakefield emittance dilution factor for different 500 GeV linear collider designs. The SLC is included for comparison.*

In order to demonstrate the feasibility of the s.c. cavity technology, the TESLA collaboration decided to start an R&D program and to build the TESLA Test Facility (TTF) several years ago [4]. In the formative stages of the TESLA collaboration, three 5-cell L-band cavities were built and tested to reach accelerating fields of 26-28 MV/m [5]. The TTF [6] includes the infrastructure for applying different processing techniques to the Niobium cavities obtained from industrial series production. At the time of writing this report, 12 cavities have been processed and tested at the TTF. Several of the cavities have reached or even surpassed the TESLA goal, see Fig. 3.1.3

for an example. A test linac is under construction incorporating 32 cavities in order to demonstrate acceleration of a beam up to 800 MeV. Operation with an RF- and beam-pulse structure as in the linear collider design is foreseen, so that a full integrated system test relevant for TESLA-500 is possible. While construction work on the TTF linac and commissioning of its first stage is in progress, several important results have already been obtained (see section 3.2.1.1 for more details on the status of cavity research):

- Several of the cavities have reached or even surpassed the TESLA goal, see Fig. 3.1.3 for an example.
- Field emission has been overcome by handling the cavities in a dust-free environment and by proper treatment.
- Gradients and quality factors achieved in horizontal tests were similar to those achieved in vertical tests.

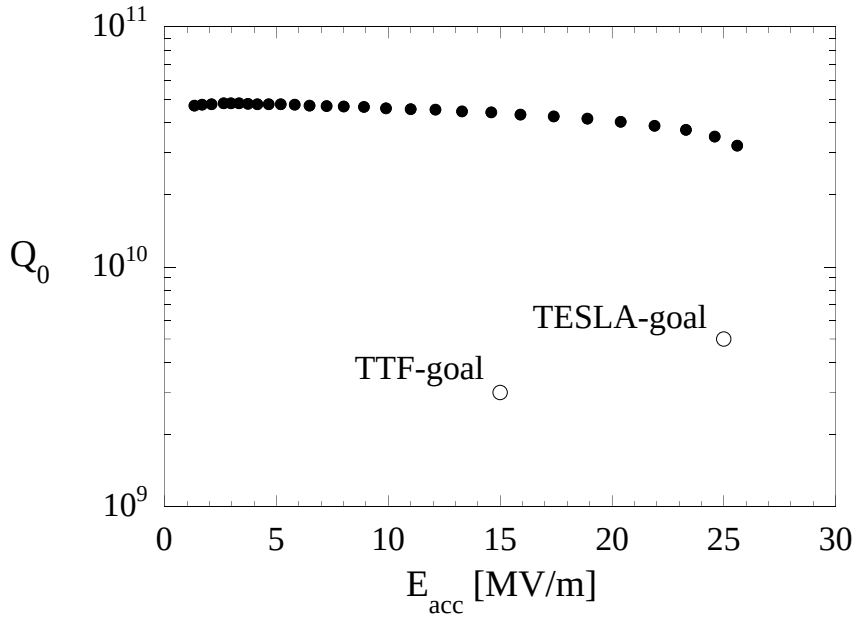


Figure 3.1.3: *Quality factor vs. accelerating gradient for one of the 9-cell TESLA cavities tested with CW-RF in a vertical cryostat. The initial goal for the TTF and the design goal for TESLA are indicated.*

While demonstration of a successfully working superconducting linac system is the primary goal at the TTF, its construction will also provide a sound basis for a cost analysis. Studies of cost effective component design are under way and the preparation of a detailed cost model for all major components and sub-systems of TESLA will be

the next step of our design work. There are, however, already important ingredients for a substantial cost reduction in the design of the systems.

One measure is to build cavities with a larger number of cells than done before (9 cells instead of typically 4-5 cells). Longer structures with higher number of cells lead to a reduced number of input couplers, HOM-couplers, tuning devices, RF-waveguides, etc. Building long cryostats with many cavities (there are 8 cavities per cryostat in the TTF design) also contributes to a reduction of the system costs. A large number of cryostats is grouped together in a 2.5 km long cryogenic unit for the TESLA linac. The helium distribution system is incorporated in the cryostats with only one feedbox every 2.5 km. The number of cold-to-warm transitions, which are costly because of the penetration of the heat shields and the vacuum vessel and which also contribute to the heat load, is reduced to a minimum.

X-Ray FEL and Further Options

Due to its ability to sustain highest beam quality during acceleration, the TESLA linac is the ideal driver for an X-ray Free Electron Laser (FEL). Such a device is now considered by the major part of the synchrotron radiation community as a truely new generation of X-ray sources. It is explained in chapter 5 of this report how an X-ray FEL can be integrated into TESLA, allowing to operate this facility in parallel with the linear collider.

The X-ray FEL concept represents a considerable extrapolation of present day FEL technology. It has therefore been decided to perform a proof-of-principle experiment at the TTF. After a successful test at about 50 nm wavelength, an energy upgrade of the TTF will open up the possibility to build a soft X-ray FEL user facility. Thus operational and scientific experience can be gained which seems indispensable for the construction of a large-scale X-ray FEL laboratory. Furthermore, a long-term adequate usage of the TTF will be provided.

If the linear collider is constructed close to an existing large proton storage ring (as the HERA-p ring or the TEVATRON), a linac-ring electron-proton collider with a center-of-mass energy of about 1 TeV is a possible option. There is also the possibility to convert the electron beam into a γ-beam by Compton scattering, thus obtaining γ-p collisions. The most serious issue here concerns the achievable luminosity (see e.g. refs. [8, 9, 10]). More detailed studies are required before a conclusion concerning the feasibility of a linac-ring e-p collider with reasonably high luminosity can be drawn.

Recently, the concept of $\mu^+\mu^-$ colliders has attracted much attention. While many design aspects of such a facility are still under study [11], it is already clear that a superconducting linac, operated in a multi-pass recycling mode, would be required. One may thus consider TESLA as the only linear collider which can be extended into a $\mu^+\mu^-$ collider in the long-term future, provided this approach towards multi-TeV lepton colliders is found to be feasible and superior to e^+e^- machines in this energy range.

In case TESLA is constructed at the DESY site, an additional option becomes attractive. The HERA electron ring can be used as a pulse stretcher to deliver a con-

tinuous electron beam of 15...25 GeV for Nuclear Physics experiments (see appendix B). The ring is filled by additional beam pulses generated in the lower energy part of one of the TESLA linacs. The additional investment required for this option would be minimal, still the quality of the beam delivered from the stretcher ring being comparable to the original ELFE proposal [7] for a continuous beam facility.

3.1.2 General Layout

A sketch of the overall layout of the TESLA linear collider is given in Fig. 3.1.4. We start the description of the various subsystems of the machine at the beam sources.

With the relatively large design beam emittance, it may be possible to obtain the required electron beam quality directly from a laser-driven RF gun. This design aspect is not fundamental but considered as a possible low-cost version of the electron injection system for an initial stage of operation. An additional damping ring, similar to the one on the positron side (see below), is likely to be required in order to exploit the full potential of the collider with smaller beam emittance. A polarized electron source with a larger emittance is used in that case (section 3.4.3). The requirements for longitudinal bunch compression are rather moderate and can be met with a single stage compressor, as described in section 3.6. After the bunch compressor, injection into the main linear accelerator takes place at a beam energy of 3.2 GeV.

The positron injection system has to provide a total charge of about $4 \cdot 10^{13} e^+$ per beam pulse, which does not seem feasible with a conventional source. The method considered here is to produce the positrons from γ-conversion in a thin target (section 3.4.5). The photons are produced by passing the spent high-energy electron beam after the IP through a wiggler. This requires a special beam-optical system and various collimators in order to deliver an electron beam of sufficiently small size to the wiggler. After the wiggler the electron beam is deflected, bypasses the photon target and is sent on a dump. The positrons are preaccelerated in a conventional 200 MeV L-band linac, followed by a 3 GeV superconducting accelerator. The method proposed here has a lower limit for the electron beam energy required to generate photons of sufficiently high energy. This limit is at about 150 GeV, which puts a lower boundary on the center-of-mass energy accessible with TESLA at maximum luminosity. In case lower energy running at high luminosity becomes important, solutions with additional electron beam pulses and bypass-beamlines are conceivable.

A low-intensity auxiliary e^+ source will be needed for commissioning and machine study purposes. We assume that the auxiliary source should be capable of generating single bunches at full design intensity and bunch trains at a few per cent of the design intensity.

Changes of the beam parameters at the IP can introduce fluctuations of the positron production efficiency, which in turn feed back to the spent beam properties via the collision at the next pulse. It has been checked that the coupled dynamical system is damped and stable and that the expected fluctuations are acceptable.

Besides providing a sufficiently high positron beam intensity, this concept offers additional advantages. With a thin target, the positron beam tends to have a smaller

transverse emittance than from a conventional source. The considerable investment and operating costs for a high-power electron linac needed in a conventional scheme are avoided. Furthermore, it is conceivable to obtain a polarized positron beam by using a helical undulator instead of a wiggler. This option is more ambitious concerning the required quality of the spent beam passing through the undulator.

The positron beam is injected into the damping ring at an energy of 3.2 GeV. The bunch train is stored in the ring in a compressed mode, with the bunch spacing reduced by about a factor of 14. Still a large ring circumference of about 17 km is required. We choose a damping ring design (section 3.5) here which avoids having to build an additional ring tunnel of this size. The layout has two 8 km straight sections, entirely placed in the main linac tunnel. Additional tunnels are only required for the small loops at the ends. Despite its unconventional shape (it has been named the "dogbone ring"), no particular difficulties concerning single particle and collective beam dynamics were found. A 400 m long wiggler section is foreseen to achieve sufficient damping. Special kicker devices are required for compression and decompression of the bunch train at injection and extraction, respectively. The two main linear accelerators comprise about ten thousand one meter long superconducting cavities each. Groups of 8 cavities are installed in a common cryostat, a cost-efficient modular solution which is already applied at the TTF linac (see section 3.2.2). Superconducting magnets for beam focussing and steering, beam position monitors and absorbers for higher order modes induced by the beam are integrated in the modules.

The RF-power is generated by some 300 klystrons per linac, each feeding 32 9-cell cavities. The required peak power per klystron is 8 MW, which includes a margin for correcting phase errors which occur during the beam pulse due to mechanical deformations of the cavities (Lorentz force detuning, microphonics). The high-voltage pulses for the klystrons are provided by conventional type modulators, or, as an alternative option presently under study, in "SMES" devices where the pulse energy is stored as magnetic field energy in superconducting solenoids. The TESLA RF-system is described in detail in section 3.2.4.

The cryogenic system of the TESLA linac is divided into 2.5 km long subsections, each supplied by a cryogenic plant, see section 3.2.5 for details.

The beam transport between the linac and the IP (so-called beam delivery system, described in section 3.7) consists of collimation, beam diagnostics and correction, and final focus sections. In the interaction region, no crossing angle is required, because the beams can be separated by an electrostatic deflector well before the first parasitic interaction with the subsequent bunch would occur (at 150 m from the IP). Large aperture superconducting quadrupoles can be used in the IR, with one benefit being that collimation requirements upstream for protection of the experiment from background are rather relaxed. With gas scattering practically absent in the TESLA linac and wakefields being very small, the expected amount of beam halo which must be collimated is small so that background from muons originating at the collimators is unlikely to be a problem. In case the loss rates at the collimators exceed an acceptable limit (e.g. due to mis-steering or possible failures upstream), the large bunch spacing allows to trigger a safety system which sends the remaining bunch train to a beamdump.

The concept of the final focus sytem is essentially the same as for the successfully tested FFTB system at SLAC. Beam size demagnification and chromatic corrections for the TESLA design parameters are even somewhat less ambitious than at the FFTB. The beams can be kept in collision at the IP with high accuracy by using a fast bunch-to-bunch feedback which measures the beam-beam deflection and steers the beams back into head-on collision within a time small compared to the beam pulse length. A similar system is foreseen at the entrance to the main linac and at its end, removing possible pulse-to-pulse orbit jitter generated in the injection system or in the linac.

The layout of the beam delivery system chosen here is optimized for a single interaction point. It can be entirely accomodated in a straight tunnel and is kept as simple as possible, with benefits for the required tunnel length (about 2.3 km for the total system between the two linacs) and for facilitating commissioning and operation of the machine. There is no fundamental problem arranging a 2nd IP by adding another beam line as sketched in Fig. 3.1.4. The detailed geometry and beam optics will depend on whether the 2nd experiment is to study e^+e^- collisions as well, or is aiming at the investigation of $\gamma\gamma$ and $e\gamma$ collisions. A discussion of the latter option for TESLA is given in appendix A. The two linear accelerators as well as the beam delivery system will be installed in an underground tunnel of 5 m diameter, see Fig. 3.1.5 and section 3.10. A $30 \times 80m^2$ experimental hall is foreseen to accomodate the detector, with an option to be extended in case a 2nd separate experiment is to be installed. Six additional surface halls are required at a distance of about 5 km along the linacs, each housing two cryogenic plants. These halls are connected to the underground tunnel by access shafts. They will also contain the modulators which generate the HV pulses for the klystrons. The pulse transformers are placed in the tunnel close to the klystrons and are connected to the modulators by cables. The long cables contribute to a few % of power losses, but it is advantageous that maintenance of the modulators will be possible while the machine is running without the need of access to the tunnel. Exchange of klystrons requires to interrupt operation of the machine. With an energy surplus of 2% foreseen in the design, such maintenance breaks of one day are necessary only every few weeks, assuming an average klystron life time of 40,000 h.

The layout of all subsystems of TESLA takes safety margins into account, as will become more clear by the detailed descriptions given in the respective sections of this report. Examples for safety margins are a 50% excess of the cooling capacity of the cryogenic plants and a 25% higher klystron peak power than what is needed according to the design parameters. In addition, margins concerning the beam parameters as emittances and bunch intensity are foreseen to allow for dilutions and losses at different stages from the source to the IP.

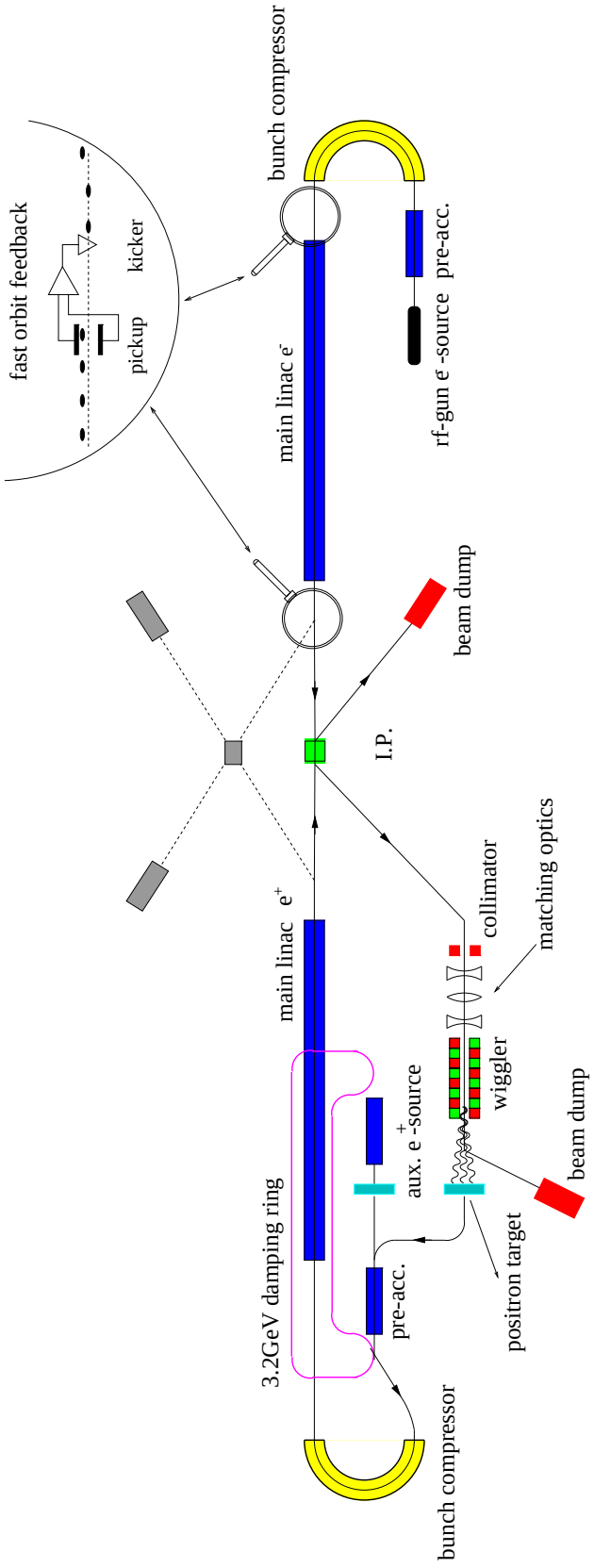


Figure 3.1.4: *Sketch of the overall layout of TESLA.*

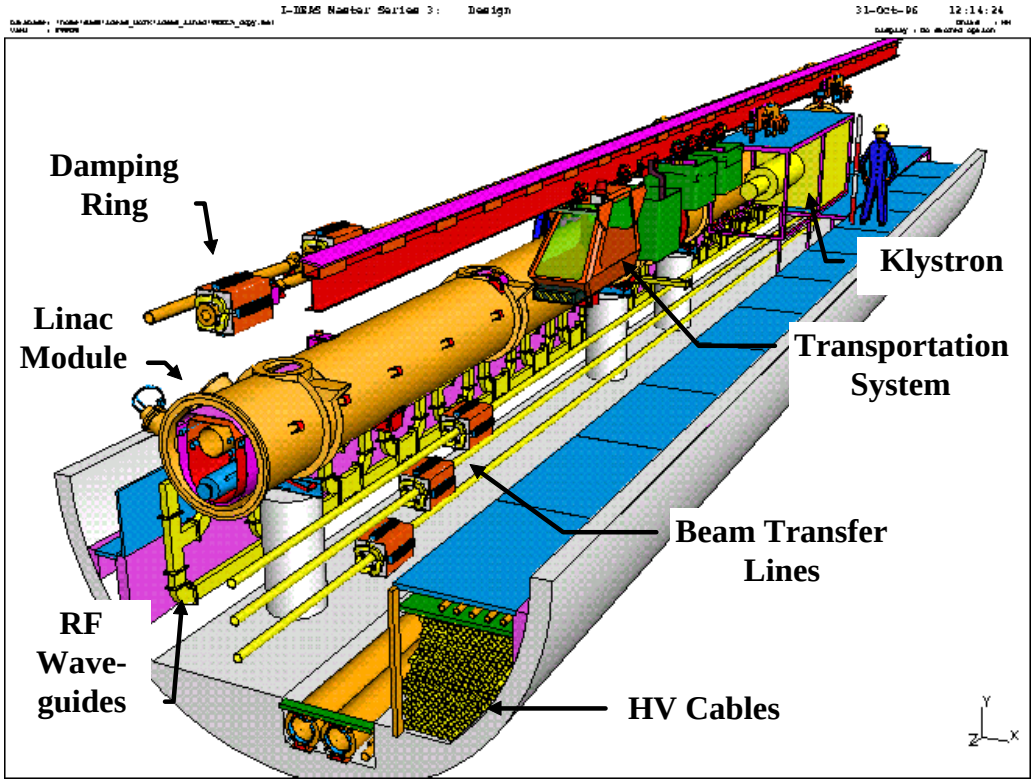


Figure 3.1.5: *Sketch of the 5 m diameter TESLA linac tunnel.*

3.1.3 Parameters at 500 GeV

The second key parameter for a linear collider, besides the center-of-mass energy of the colliding beams, is the luminosity L. It is given by

$$L = \frac{n_b N_e^2 f_{rep}}{4\pi\sigma_x^*\sigma_y^*} \times H_D \tag{3.1.3}$$

n_b	number of bunches per pulse
N_e	number of electrons (positrons) per bunch
f_{rep}	pulse repetition frequency
$\sigma^*_{x,y}$	horizontal (vertical) beam size at interaction point
H_D	disruption enhancement factor (typically $H_D \approx 1.5$)

Introducing the average beam power $P_b = E_{cm} n_b N_e f_{rep}$, the luminosity can be

written as

$$L = \frac{P_b}{E_{cm}} \times \frac{N_e}{4\pi\sigma_x^*\sigma_y^*} \times H_D \tag{3.1.4}$$

An important constraint on the choice of interaction parameters is due to the effect of beamstrahlung: the particles emit hard synchrotron radiation in the strong space-charge field of the opposing bunch. The average fractional beam energy loss from beamstrahlung is approximately given by [12]:

$$\delta_E \approx 0.86 \frac{r_e^3 N_e^2 \gamma}{\sigma_z(\sigma_x^* + \sigma_y^*)^2} \tag{3.1.5}$$

r_e classical electron radius
γ relativistic factor E_{beam}/m_0c^2

Beamstrahlung causes a reduction and a spread of the collision energy and can lead to undesirable experimental background. The energy loss δ_E therefore has to be limited to typically a few percent. By choosing a large aspect ratio $R = \sigma_x^*/\sigma_y^* >> 1$, δ_E becomes independent of the vertical beam size and the luminosity can be increased by making σ_y^* as small as possible. Since $\sigma_y^* = (\epsilon_{y,N}\beta_y^*/\gamma)^{1/2}$, this is achieved by both a small vertical beta function at the IP and a small normalized vertical emittance. The lower limit on β_y^* is given by the bunch length ("hourglass effect"). Setting $\beta_y^* = \sigma_z$, the luminosity can be expressed as:

$$L \approx 5.74 \cdot 10^{20}\mathrm{m}^{-3/2} \times \frac{P_b}{E_{cm}} \times \left(\frac{\delta_E}{\epsilon_{y,N}}\right)^{1/2} \times H_D \tag{3.1.6}$$

Due to the small wakefields, the superconducting linac is ideal for transporting a beam with an extremely small emittance $\epsilon_{y,N}$. In order to define a standard parameter set for the 500 GeV collider, we do not have to make full use of this potential, though. The high AC-to-beam power conversion efficiency of TESLA of about 16 % allows to produce a high beam power and to obtain the required luminosity with a total AC-power consumption of 99 MW and a normalized emittance achievable with tolerances (between 0.1 and 0.5 mm) in the linac and in the damping rings which can realistically be obtained at the time of machine installation. The additional use of beam-based correction and alignment methods will then allow to push $\epsilon_{y,N}$ to much smaller values, opening up options for upgrading the machine as schematically shown in Fig. 3.1.6. The other basic parameters, not directly entering into eq. 3.1.6, have been chosen such that freedom in parameter space is maintained, not being at a technical limit in any of the collider subsystems. For instance, the large bunch spacing (708 ns) leaves a considerable safety margin for technical components like the damping ring injection and extraction devices, bunch-to-bunch orbit measurement and correction, and for resolving single bunch crossings in the experiment. As another example, the relatively large horizontal emittance ($\epsilon_{x,N} = 1.4 \cdot 10^{-5}$m) relaxes the beam optics requirements in the damping rings. An overview of the TESLA standard parameters is given in Table 3.1.2.

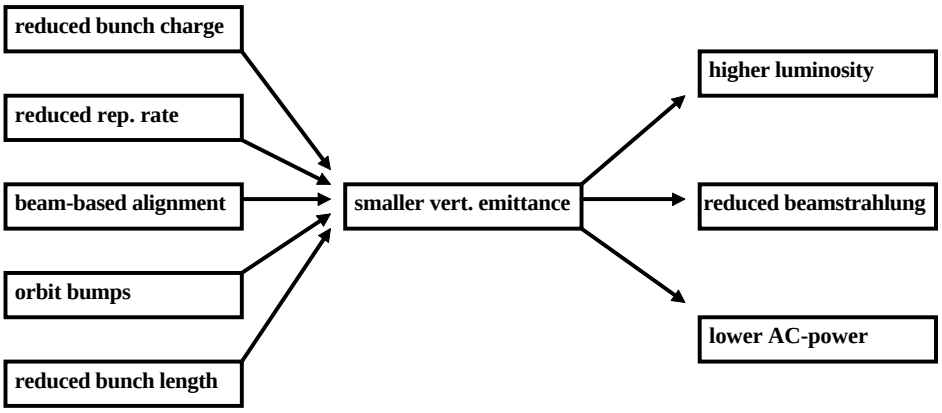


Figure 3.1.6: *Methods to reduce the vertical emittance in TESLA, and potential benefits.*

In Table 3.1.2 we also show a possible low-ϵ_y parameter set, in order to give an example for the large flexibility in parameter space available for TESLA. The version shown here combines a very low level of beamstrahlung with a somewhat higher luminosity and reduced AC-power. Such a parameter set with low δ_E would e.g. be desirable for operation around the top quark threshold (at $E_{cm} \approx 350\,\mathrm{GeV}$) or for a polarized positron source. Obtaining the small emittance will require to apply beam-based techniques, especially to improve the accuracy of the beam position monitor alignment w.r.t. the focussing magnets. The necessary accuracy (several 10's of μm, both in the linac and the damping rings) is easy to achieve and, once established, expected to be stable over long periods of time.

3.1.4 Upgrade to Higher Energy

The center-of-mass energy reach of TESLA is clearly limited by the maximum gradient achievable with the superconducting cavities and cannot, unlike designs using conventional accelerating structures, be increased by upgrading the RF-pulse power. In order to go to energies beyond $E_{cm} \approx 1\,\mathrm{TeV}$, the length of the machine would therefore have to be increased. We do believe, though, that a significant energy upgrade of TESLA will be possible within the site length for the 500 GeV design by operating the cavities at a gradient above 25 MV/m. The main reasons which justify this optimism are:

- The fundamental limit for the gradient in a Nb-structure at 2 K is well above 50 MV/m.
- Cavity tests at the TTF have shown that field emission can be very effectively suppressed and a quality factor far in excess of the design value ($Q_0 = 5 \cdot 10^9$) at 25 MV/m is feasible.
- Accelerating gradients of 30...40 MV/m [13, 14, 15] have been achieved with single-cell L-band cavities.

	Standard	Reduced ϵ_y
Accelerating gradient g [MV/m]	25	25
RF-frequency f_{RF} [GHz]	1.3	1.3
Fill factor	0.67	0.67
Total site length L_{tot} [km]	33	33
Active length [km]	20.4	20.4
# of acc. structures	19712	19712
# of klystrons	616	616
Klystron peak power [MW]	8	8
Repetition rate f_{rep} [Hz]	5	4
Beam pulse length T_P [μs]	800	800
RF-pulse length T_{RF} [μs]	1330	1330
# of bunches p. pulse n_b	1130	2260
Bunch spacing Δt_b [ns]	708	354
Charge p. bunch N_e [10^{10}]	3.63	1.82
Emittance at IP $\gamma\epsilon_{x,y}$ [10^{-6}m]	14, 0.25	12, 0.03
Beta at IP $\beta^*_{x,y}$ [mm]	25, 0.7	25, 0.5
Beam size at IP $\sigma^*_{x,y}$ [nm]	845, 19	783, 5.5
Bunch length at IP σ_z [mm]	0.7	0.5
Beamstrahlung δ_E [%]	2.5	1.0
Luminosity L [10^{33}cm^{-2}s^{-1}]	6	8.4
Beam power P_b [MW]	16.3	13
AC power P_{AC} [MW]	99	83

Table 3.1.2: *TESLA parameters for E_{cm}=500 GeV. The machine length includes a 2% overhead for energy management. The klystron power and AC power quoted include a 10% regulation reserve. The standard parameter set is the basis for the design study documented in this report. The additional parameter set (right-hand column) is given as an example for the potential of TESLA operating with a lower beam emittance.*

We are aware that a number of questions have to be adressed before a definite conclusion can be drawn (e.g. stronger Lorentz-force detuning, potentially more severe dark current problems), but with further R&D it seems conceivable that a maximum gradient of about 40 MV/m at $Q_0 = 5 \cdot 10^9$ can be reached with the 9-cell cavities. The average gradient for the entire linac may be lower initially, so that a possible scenario for an energy upgrade in steps could be to exchange groups of modules containing the "weakest" (i.e. lowest gradient) cavities with new, high-performance modules. This upgrade path leads to a maximum energy of E_{cm}=800 GeV. The subsystems for TESLA (in particular the beam delivery system) are designed such that the energy upgrade can be accomodated without further hardware modifications. In order to obtain optimum luminosity at higher energy, an upgrade of the RF-system is also necessary, though. At higher gradient, either the RF-pulse power or pulse length must be increased to maintain a reasonable ratio of beam-on time to RF-on time. We follow the first ap-

	800 GeV	800 GeV low ϵ_y	1.6 TeV
Accelerating gradient g [MV/m]	40	40	40
Total site length L_{tot} [km]	33	33	≈ 62
Active length [km]	20.4	20.4	40.8
# of acc. structures	19712	19712	39424
# of klystrons	1232	1232	2464
Klystron peak power [MW]	8	8	8
Repetition rate f_{rep} [Hz]	3	3	3
Beam pulse length T_P [μs]	640	640	640
RF-pulse length T_P [μs]	1330	1330	1330
# of bunches p. pulse n_b	1130	2260	2260
Bunch spacing Δt_b [ns]	566	283	283
Charge p. bunch N_e [10^{10}]	3.63	1.82	1.82
Emittance at IP $\gamma\epsilon_{x,y}$ [10^{-6}m]	14, 0.25	12, 0.025	12, 0.025
Beta at IP $\beta^*_{x,y}$ [mm]	25, 0.7	25, 0.5	25, 0.5
Beam size at IP $\sigma^*_{x,y}$ [nm]	669, 15	618, 4.0	436, 2.8
Bunch length at IP σ_z [mm]	0.7	0.5	0.5
Beamstrahlung δ_E [%]	5.2	2.2	6.7
Luminosity L [10^{33}cm^{-2}s^{-1}]	5.7	11	20
Beam power P_b [MW]	15.6	15.6	31.2
AC power P_{AC} [MW]	115	115	230

Table 3.1.3: *TESLA parameters for an upgrade to 800 GeV and for a 2nd stage at 1.6 TeV with doubled linac length (see text).*

proach, which has the advantage that the modulators providing the pulse power for the klystrons do not have to be modified. The 800 GeV version of TESLA would then have twice as many klystrons and modulators as the initial stage at 500 GeV. The pulse repetition rate is reduced so that the dynamic load for the cryogenic system remains almost constant and the AC-power consumption is only slightly increased. As in the previous section, we show two sets of beam parameters in Table 3.1.3. The alignment tolerances in the damping rings and in the main linac are for both versions similar as for the respective parameter sets of TESLA-500.

The potential for operating with a very small vertical beam emittance becomes particulary important when we consider a linear collider at energies beyond 1 TeV. According to eq. 3.1.6, scaling the luminosity as $L \propto E^2_{cm}$, as desirable, becomes difficult unless a high level of beamstrahlung is accepted. Assuming that the same vertical emittance as for the 800 GeV low-ϵ_y version can be maintained in a machine doubled in length, we arrive at the parameter set for a 1.6 TeV collider shown in the last column of Table 3.1.3. It demonstrates that the TESLA approach has the capability of maintaining reasonable interaction parameters at an energy scale well above 1 TeV. We have not yet investigated this 2nd stage upgrade in detail, but several aspects are already included in the initial stage design. The tunnel is kept straight in the beam

delivery section, so that it can later be used for the installation of linac structures. The two linacs of the 1st stage would be used as one linac of the 2nd stage (note that the superconducting cavities allow to accelerate a beam in both directions) and the machine length be extended to one side. Thus the part of the linac being used for the FEL facility and potentially other options would remain unchanged. A new beam delivery system and interaction region would have to be constructed. The entire injection system can be kept basically unchanged, except for modifications concerning the spent beam capture system and an additional transfer line from one of the damping rings to the beginning of the new 800 GeV linac.

Bibliography

[1] G. A. Loew (ed.), *International Linear Collider Technical Review Committee Report*, SLAC-R-95-471, 1995.

[2] D. Burke for the FFTB Collaboration, *Results from the Final Focus Test Beam*, Proc. of the IVth European Particle Accelerator Conf., London 1994, Vol. I, p. 23.

[3] R. Brinkmann, *Beam Dynamics in Linear Colliders- What Are the Choices ?*, DESY-M-95-10, 1995.

[4] *Proposal for a TESLA Test Facility*, DESY-TESLA-93-01, 1992.

[5] C. Crawford et al., Particle Accelerators 49, p. 1, 1995.

[6] D. A. Edwards (ed.), *TESLA Test Facility Linac - Design Report*, DESY-TESLA-95-01, 1995.

[7] J.M. De Conto (ed.), *Electron Laboratory for Europe - Accelerator Technical Proposal*, 1993.

[8] M. Tigner, B. H. Wiik and F. Willeke, Proc. Particle Accelerator Conf., San Francisco 1991, Vol. 5, p. 2910.

[9] Z. Z. Aydin et al., *HERA+LC Based γp Collider: Luminosity and Physics*, DESY-95-014, 1995.

[10] R. Brinkmann and M. Dohlus, *A Method to Overcome the Bunch Length Limitation on β_p^* for Electron-Proton Colliders*, DESY-M-95-11, 1995.

[11] *$\mu^+\mu^-$ Collider – A Feasibility Study*, BNL-52503, FNAL-Conf.-96/092 and LBNL-38946, July 1996.

[12] P. Chen and K. Yokoya, *Beam-Beam Phenomena in Linear Colliders*, KEK-report 91-2, 1991.

[13] T. Higuchi et al, *Investigation of Barrel Polishing for Superconducting Niobium Cavities*, KEK-report 95-220, 1996.

[14] P. Kneisel, R. W. Röth and H.-G. Kürschner, *Results from a Nearly "Defect-Free" Niobium Cavity*, Proc. 7th Workshop on RF-Superconductivity, p. 449,Saclay 1995.

[15] M. Ono, S. Noguchi, K. Saito, E. Kako, T. Shishido, H. Inoue, Y. Funahashi, T. Fujino, S. Koizumi, T. Higuchi, T. Suzuki, H. Umezawa and K. Takeuchi, *Achievement of High Acceleration Field (40 MV/m) in L-Band Superconducting Cavity at KEK*, Proc. 21st Linear Accelerator Meeting, Tokyo 1996, NUP-A-96-10, p. 38.

3.2 Linac Technology

3.2.1 Cavity Design and Characteristics

3.2.1.1 Introduction

The first multicell niobium cavities were fabricated at Stanford in the early 70's and used in an electron linac. The accelerating gradient was limited to about 2 MV/m due to electron multipacting [1]. A significant step towards higher gradients was the finding that multipacting could be virtually eliminated by changing the cavity geometry from a cylindrical to a spherical or elliptical cross section. The rounded shape permitted a simplified cavity production based on deep-drawing or spinning techniques instead of machining of bulk niobium. Starting in the early 1980's first beam tests with niobium cavities were carried out in the electron-positron storage rings CESR (Cornell), PETRA (DESY-Karlsruhe and DESY-CERN collaboration) and TRISTAN (KEK). In the second half of the 1980's 32 cavities were installed in TRISTAN and successfully used in seven years of high luminosity operation with an operating gradient of 3.5 – 4.5 MV/m. In the HERA electron ring 16 cavities are routinely in operation. Two superconducting linacs with niobium cavities have been built: the S-Dalinac at Darmstadt using 20-cell 3 GHz structures developed at Wuppertal and the CEBAF linac at Newport News using 5-cell 1.5 GHz structures developed at Cornell. The cavities for these machines were produced by industry.

An alternative route has been chosen at CERN for the 352 MHz cavities used in LEP. The resonators are deep drawn from copper and then sputtered with niobium. The major advantage is a considerable savings in the expensive niobium material and an improved heat conductivity in the cavity wall. A total of 272 four-cell cavities will be installed in LEP until 1998.

The largest SRF system presently in operation is the CEBAF linac with 338 cavities. All cavities have been subjected to tests in a vertical cryostat prior to installation in the linac [1]. The typical limitation was either a thermal breakdown or a severe degradation in quality factor due to field emission of electrons. The average accelerating gradient just below quench was about 13 MV/m but with a wide spread from 7 to 20 MV/m. The *usable gradient* in the vertical test, defined by a field emission loading of less than 1 Watt and an X ray level of less than 1 rad/h outside the cryostat, was about 3 MV/m lower. Another loss in gradient has been experienced after installation in the horizontal cryostats and mounting of the couplers: the usable gradient during commissioning ranged from 5 to 14 MV/m with an average of ≈ 8 MV/m.

The CEBAF experience is of great relevance for TESLA since the cavities are based on the same design principles and fabrication techniques. Obviously a significant progress is needed to achieve the TESLA design goal of 25 MV/m usable gradient or even the more moderate 15 MV/m in the TESLA Test Facility TTF. Numerous tests with single- and multicell cavities have shown that this is indeed possible. The recently

[1] "Multipacting" is a commonly used abbreviation of "multiple impacting" and disgnates the resonant multiplication of free electrons which gain energy in the RF electromagnetic field and impact on the resonator surface where they induce secondary electron emission.

introduced technique of cleaning the surface by high-pressure water rinsing [2, 3] has led to a considerable supression of field emission loading. Moreover, extreme care in cavity handling is essential to avoid dust particles that might act as field emitters and to preserve a low surface resistance. An elaborate infrastructure has been set up at DESY for cavity treatment and handling: industrial clean rooms according to semiconductor standards, an automated chemistry system, a high-pressure water rinsing system and an ultrahigh-vacuum furnace for heat treatment. These facilities are in use for the preparation of the TTF cavities.

The cavity test area comprises a bath cryostat for vertical tests at 1.5 to 4.2 K and a horizontal cryostat for the test of cavities that are welded into their liquid helium container and equipped with the input and higher-order-mode couplers used in the linac. A special feature is the possibility to apply short RF pulses with an instantaneous power of up to 1 MW to destroy field emitters in a cavity (the so-called high power processing technique HPP). Two temperature mapping systems to find "hot spots" are also available.

At present it appears that only cavities made from solid pure niobium have the potential of achieving accelerating fields of 25 MV/m or more while the gradients in niobium-sputtered copper cavities are typically in the 10 MV/m range. For this reason the decision was taken to consider only pure niobium as cavity material.

3.2.1.2 Basic principles of RF superconductivity

Surface resistance

In contrast to the d.c. case superconductors are not free of energy dissipation in microwave applications. The reason is that the RF magnetic field penetrates a thin surface layer and induces oscillations of the electrons which are not bound in Cooper pairs. The number of these "free electrons" drops exponentially with temperature. According to the Bardeen-Cooper-Schrieffer (BCS) theory of superconductivity the surface resistance is given by the expression

$$R_{BCS} \propto \frac{\omega^2}{T} \exp(-1.76\, T_c/T) \,. \tag{3.2.1}$$

where $f = \omega/2\pi$ is the frequency of the RF. In the two-fluid model of superconductors one can derive a refined expression for the surface resistance ([4, 5])

$$R_{BCS} = \frac{C}{T}\, \omega^2\, \sigma_n\, \Lambda^3 \exp(-1.76\, T_c/T) \,. \tag{3.2.2}$$

Here C is a constant, σ_n the normal-state conductivity of the material and Λ an effective penetration depth, given by

$$\Lambda = \lambda_L \sqrt{1 + \xi_0/l} \,.$$

λ_L is the London penetration depth, ξ_0 the coherence length and l the mean free path of the unpaired electrons. The fact that σ_n is proportional to the mean free path l leads

to the surprising conclusion that the surface resistance does not assume its minimum value when the superconductor is as pure as possible ($l \ll \xi_0$) but rather in the range $l \approx \xi_0$. For niobium the BCS surface resistance at 1.3 GHz amounts to about 800 nΩ at 4.2 K and drops to 1 nΩ at 1.5 K, see Fig. 3.2.1. This exponential dependence is the reason why cooling with superfluid helium is essential for achieving high accelerating gradients and at the same time very high quality factors. In addition to the BCS term there is a residual resistance R_{res} caused by impurities or lattice distortions. This term is temperature independent and amounts to a few nΩ for very pure niobium but may readily increase by large factors if the surface is contaminated.

Heat conduction in niobium

The heat produced at the inner cavity surface has to be guided through the cavity wall to the superfluid helium bath. Two quantities characterize the heat flow: the heat conductivity λ of the bulk niobium and the temperature drop at the niobium-helium interface caused by the Kapitza resistance. The heat conductivity of niobium at cryogenic temperatures scales approximately with the residual resistivity ratio[2] RRR, a rule of thumb being

$$\lambda(4.2\mathrm{K}) \approx 0.25 \cdot RRR \quad [\mathrm{W}/(\mathrm{m} \cdot \mathrm{K})]$$

but one has to keep in mind that λ is strongly temperature dependent and drops by about an order of magnitude when lowering the temperature to 2 K (Fig. 3.2.2). A good heat conductivity is the main motivation for using high purity niobium with $RRR \approx 300$ as a starting material. The RRR may be further improved by heat treatment of the finished cavity (see below).

Critical magnetic field

Superconductivity breaks down when the RF magnetic field exceeds the critical field of the superconductor. In the radio frequency case the so-called "superheating field" is relevant which for niobium is about 20% higher than the thermodynamical critical field of 200 mT. Niobium is in principle a soft type II superconductor without flux pinning. Hence weak magnetic fields like the Earth's field should be expelled upon cooldown but in practice a severe degradation in performance is observed in cavities without magnetic shielding.

3.2.1.3 Layout of the TESLA cavities

Choice of frequency

The losses in a radio-frequency cavity are proportional to the product of conductor area and surface resistance. For a given length of a multicell resonator, the area scales with $1/f$ while the surface resistance of a superconducting cavity goes like f^2 for

[2] RRR is defined as the ratio of the resistivities at room temperature and at 10 K (just above T_c).

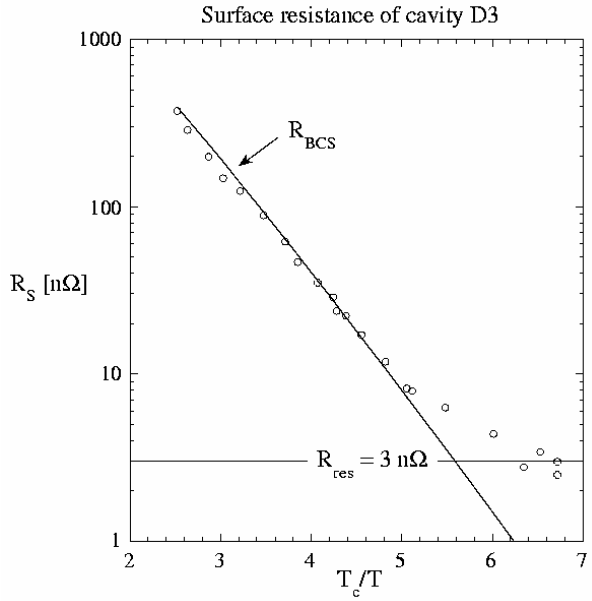


Figure 3.2.1: *The surface resistance of a 9-cell TESLA cavity plotted as a function of T_c/T. The residual resistance of 3 nΩ corresponds to an "unloaded" quality factor $Q_0 = 3 \cdot 10^{10}$.*

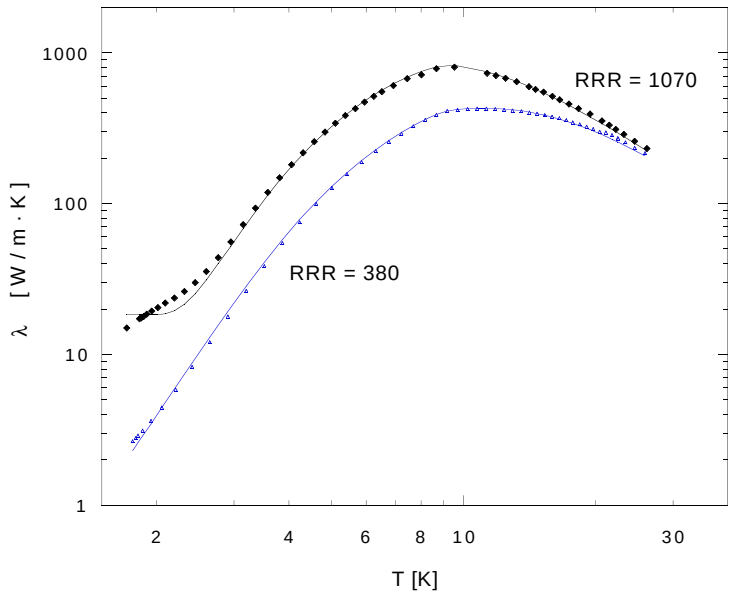


Figure 3.2.2: *Measured heat conductivity in niobium as a function of temperature [6]. Continuous curves: parametrisation by B. Bonin, using only RRR and the average grain size as input parameters [7].*

$R_{BCS} \gg R_{res}$ and is independent of f for $R_{BCS} \ll R_{res}$. Above 3 GHz the BCS term dominates and hence the losses grow linearly with frequency whereas below 300 MHz one obtains a growth with $1/f$. To minimize the Ohmic losses in the cavity wall one should therefore select f in the range 300 MHz to 3 GHz.

Cavities in the 350 to 500 MHz regime are in use in electron-positron storage rings. Their large size is advantageous to suppress wake field effects and higher order mode (HOM) losses. However, for a linac of 10 km length the niobium and cryostat costs for these bulky cavities are prohibitive, hence a higher frequency has to be chosen. Considering material costs $f = 3$ GHz might appear the optimum but there are compelling arguments for choosing about half this frequency.

a) At 1.5 GHz the iris diameter is a factor of two larger than at 3 GHz. The wake fields scale inversely with the second to third power of this diameter and are hence much weaker at the lower frequency. Beam emittance growth and cryogenic losses are correspondingly lower.

b) For very pure niobium the surface resistance grows with f^2, so a lower frequency is favourable to prevent thermal breakdown. In fact, a quality factor above 10^{10} can only be reached at 3 GHz if the helium temperature is lowered to 1.5 K which implies a significant increase in cooling costs in comparison with the 2 K cooling scheme of 1.3 GHz cavities. Moreover, the BCS resistance sets an absolute limit of about 30 MV/m at 3 GHz, hence choosing this frequency would definitely preclude a possible upgrade of TESLA to 35 – 40 MV/m [8].

c) The maximum number of cells per cavity is limited by the requirements that higher order modes must be extractable and no trapped modes be present. With this boundary condition the number of input and HOM couplers increases with frequency.

In summary, the arguments for a frequency around 1.5 GHz are far more compelling than for 3 GHz. The value of 1.3 GHz was chosen since high power klystrons are commercially available for this frequency.

Choice of geometry

A multicell structure is advantageous for maximizing the beam energy in a linac of a given length. With increasing number of cells per cavity, however, difficulties arise from trapped modes, uneven field distribution in the cells and too high power requirements on the input coupler. Based on experience with 4- and 5-cell cavities a choice of 9 cells was made for the TESLA cavities. A side view of the TTF cavity with the input and HOM coupler flanges is given in Fig. 3.2.3.

The shape of a cell is governed by the following considerations:

- the contour should strongly suppress multipacting,
- the electric and magnetic fields at the cavity wall should be minimized to reduce field emission and Ohmic heating,
- a sufficiently large iris diameter is desirable to reduce wake field effects.

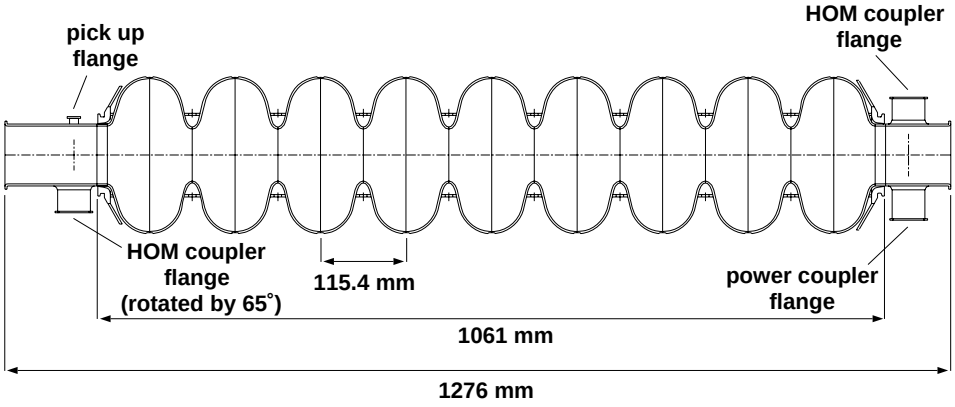


Figure 3.2.3: *Side view of the 9-cell TTF cavity with power coupler and HOM coupler ports.*

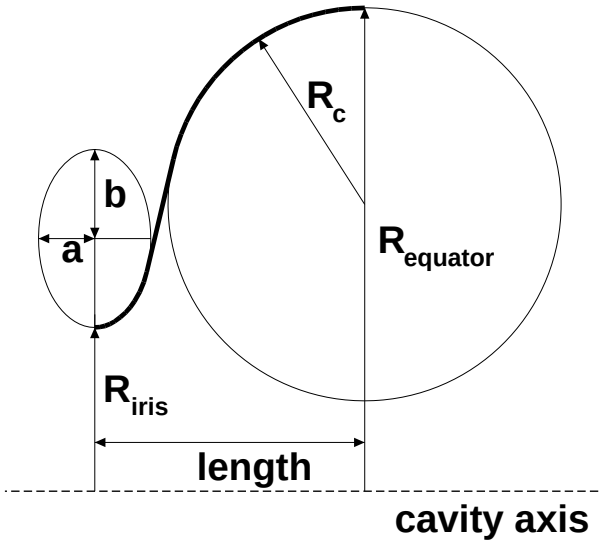


Figure 3.2.4: *Contour of a half cell.*

The shape of the TESLA cell was optimized using URMEL. The resonator is operated in the π mode (180° phase difference between adjacent cells). The longitudinal dimensions are determined by the condition that the electric field has to be inverted in the time a relativistic particle needs to travel from one cell to the next; the separation between two irises is therefore $c/(2f)$. The iris diameter influences the cell-to-cell coupling κ, the excitation of higher order modes and other important cavity parameters, such as the ratio of the peak electric (magnetic) field at the cavity wall to the accelerating field and the ratio (R/Q) of shunt impedance to quality factor. For the TTF cavities an iris diameter of 70 mm was chosen, leading to $\kappa = 1.8\,\%$ and $E_{peak}/E_{acc} = 2$. The most important parameters are listed in Table 3.2.1.

The contour of the half-cells, called "cups", is shown in Fig. 3.2.4. It is composed

Table 3.2.1: *TESLA cavity design parameters*

type of accelerating structure:	standing wave
accelerating mode	TM0 π mode
fundamental frequency	1300 MHz
design gradient E_{acc}	25 MV/m
quality factor Q_0	$> 5 \cdot 10^9$
active length L	1.038 m
number of cells	9
cell-to-cell coupling	1.87 %
iris diameter	70 mm
geometry factor	270 Ω
R/Q	1036 Ω
E_{peak}/E_{acc}	2.0
B_{peak}/E_{acc}	4.26 mT/(MV/m)
tuning range	$\pm$ 300 kHz
$\Delta f/\Delta L$	315 kHz/mm
Lorentz force detuning at 25 MV/m	$\approx$ 600 Hz
Q_{ext} of input coupler	$3 \cdot 10^6$
cavity bandwidth at $Q_{ext} = 3 \cdot 10^6$	430 Hz
RF pulse duration	1330 μs
repetition rate	5 Hz
fill time	530 μs
beam acceleration time	800 μs
RF power peak/average	208 kW/1.4 kW
number of HOM couplers	2
cavity longitudinal loss factor $k_{\parallel}$ for $\sigma_z = 0.7$ mm	10.2 V/pC
cavity transversal loss factor $k_{\perp}$ for $\sigma_z = 0.7$ mm	15.1 V/pC/m
parasitic modes with the highest impedance : type	TM_{011}
$\pi/9$ $(R/Q)/$ frequency	160 Ω/2454 MHz
$2\pi/9$ $(R/Q)/$ frequency	135 Ω/2443 MHz
bellows longitudinal loss factor $k_{\parallel}$ for $\sigma_z = 0.7$ mm	1.54 V/pC
bellows transversal loss factor $k_{\perp}$ for $\sigma_z = 0.7$ mm	1.97 V/pC/m

of a circular arc around the equator region and an elliptical section near the iris. The dimensions are listed in Table 3.2.2. The half cells at the end of the 9-cell resonator need a slightly different shape to ensure equal field amplitudes in all 9 cells.

3.2.1.4 Cavity treatment and test results

Summary of cavity production

The cavities for TTF have been or are being produced by four European companies. The present design is based on welding the 9-cell resonator from half cells that are

Table 3.2.2: *Half-cell shape parameters (all dimensions in mm).*

cavity shape parameter	midcup	endcup 1	endcup 2
equator radius $R_{equ.}$	103.3	103.3	103.3
iris radius R_{iris}	35	39	39
radius R_c of circular arc	42.0	40.3	42
horizontal half axis a	12	10	9
vertical half axis b	19	13.5	12.8
length l	57.7	56.0	57.0

deep-drawn from niobium sheet metal. The starting material is a niobium ingot which is highly purified by five- to six-fold electron beam melting by which the O_2, N_2 and C contamination is reduced to a few ppm. After rolling, the Nb sheets are degreased, a surface layer is removed by etching and then the Nb sheets are heat-treated at 800 °C in a vacuum furnace to achieve full recrystallization and a uniform grain size. Half cells are produced by deep-drawing, are then machined at the iris and equator weld zones, and cleaned by a suitable combination of degreasing, etching and clean water rinsing. They are welded together in a multistep process by electron-beam welding in a vacuum of $< 5 \cdot 10^{-5}$ mbar. A demanding requirement is to produce a smooth weld seam at the inner surface of the cavity which is particularly difficult for the equator welds because they are usually made from the outside. Also the requirements on mechanical precision are very demanding for a structure with about 20 weld connections.

Cavity treatment

A thin surface layer is damaged by the rolling of the Nb sheets and contaminated with oxygen by the subsequent 800 °C heat treatment. The damage layer, which may be 100-200 μm thick, has to be removed to obtain a good superconducting surface. Different approaches exist for this procedure, here we describe the method presently used at DESY. After degreasing the cavity receives chemical etching (also called Buffered Chemical Polishing BCP) using HF (48 %), HNO_3 (65 %) and H_3PO_4 (85 %) in the ratio 1:1:2. A layer of 80 μm is removed from the inner surface and about 30 μm from the outer surface. After careful rinsing and drying the cavity is tuned to obtain the correct fundamental frequency and equal field amplitudes in all cells. Since the tuning procedure may have introduced some dirt into the cavity, it is followed by a chemical etching of about 20 μm, a clean water rinsing and three steps of high pressure (100 bar) rinsing. The resistivity of the water leaving the cavity must be more than 18 MΩ cm to ensure that no drying stains are left on the surface. Several cavities have been tested after this step. The best result is shown in Fig. 3.2.5 where the quality factor of a nine-cell cavity is plotted as a function of accelerating field (this type of plot will be called "excitation curve" in the following).

While the cavity shown in Fig. 3.2.5 exceeds the 15 MV/m requirement for TTF

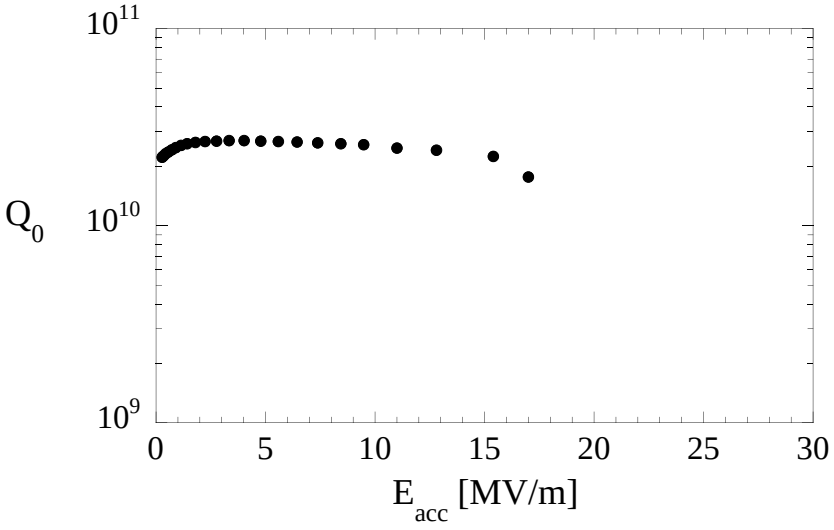


Figure 3.2.5: *Quality factor as a function of accelerating field for a cavity without heat treatment.*

already without heat treatment, this is not the case for the majority of cavities. Empirically it has been found [9] that a several-hour heat treatment at 1350-1400°C in an ultra high vacuum furnace leads to a significant improvement of the high-field capabilities of Nb cavities. Two effects play a role: the material is homogenized, and dissolved gases like hydrogen, oxygen and nitrogen diffuse out of the niobium thereby increasing the heat conductivity by about a factor of two. This high-temperature heat treatment, though quite successful in many cases, has a number of disadvantages:

The residual gas pressure in the ultra high vacuum furnace is too high to achieve the desired reduction of oxygen in the bulk niobium unless a getter material like titanium is applied. A titanium layer is deposited on the cavity surface which must be removed later by chemical etching. Titanium may diffuse deep into the niobium at grain boundaries. Moreover, the niobium material becomes extremely soft, the yield strength is reduced by almost an order of magnitude. Heat treated cavities must be handled with extreme care to avoid plastic deformation and detuning. Finally, the procedure is very time consuming and costly.

For these reasons alternative schemes are being explored like reducing the oven temperature to 800°C which is sufficient to drive hydrogen out of the niobium but not oxygen. This way the *RRR* will not be improved above the starting value of 300 but tests with single-cell cavities at KEK [10] have demonstrated that gradients of more than 30 MV/m can be realized with this moderate heat treatment.

In the following we present several "typical" test results to demonstrate the effect of the various cavity treatment steps. Fig. 3.2.6a shows the excitation curve of a cavity with heavy field emission. Starting from an excellent value at low field, the quality factor exhibits a steep decrease for fields above 8 MV/m. The reason is probably a single strong field emitter. In this and similar cases the HPP (high power processing)

technique [11] has been very successful: the high instantaneous power destroys the emitter (which can be seen on temperature maps) and afterwards the cavity can be excited to much larger fields. A drawback of this technique is also visible from the figure: the quality factor has dropped by a factor of 2-3 probably because of surface contamination from the evaporating emitter. Some of the loss can be recovered by warming up the cavity and cooling down again.

Fortunately, the long-standing problem of field emission can be largely eliminated by rinsing the cavity with ultraclean water at a pressure of about 100 bar. In several rinsing steps most of the particles sticking to the surface are removed. If the subsequent handling is done with sufficient care in a class-100 clean room, then field emission is no longer the limiting factor. An impressive example of the benefits of high pressure rinsing is shown in Fig. 3.2.6b. In spite of the progress achieved by this cleaning method, the possibility of processing with high RF power should still be foreseen for the cavities mounted in the linac to remove field emitters that might occur after some vacuum accident. This is important to provide a high operating gradient since demounting of the cavities from the cryomodule, high pressure water rinsing and reinstallation would certainly imply too much effort.

Fig. 3.2.6c demonstrates that a substantial gain in accelerating field can be achieved by heat treatment at 1400 °C. However, care must be taken that the titanium getter layer is completely removed. This is evident from the two curves in Fig. 3.2.6d. After the furnace treatment at 1400° a total of 80 μm was removed but the $Q(E)$ curve exhibited a steep drop at 10 MV/m. Removing an additional 75 μm led to the best test result obtained so far.

Also in non-heat-treated cavities an additional etching may lead to an enhanced performance. In a sample of 15 CEBAF 5–cell cavities the average gradient was 8 MV/m after the standard removal of 60 –80 μm, well below the average of 13 MV/m in the majority of the CEBAF cavities. Applying an additional chemical polishing of 70 μm raised the average gradient of these "substandard" cavities to the much improved value of 16 MV/m. Four of the 15 cavities received a 4-hour heat treatment at 1400°C with titanium getter, followed by 120 μm chemical etching to remove the Ti layer. All of these exceeded 20 MV/m.

Summary of TTF cavity tests

The design goal for the cavities of the TESLA Test Facility is a gradient of 15 MV/m at $Q = 3 \cdot 10^9$. The majority of cavities built and tested up to now have reached or exceeded this specification when operated in pulsed mode, as required in the linac (1.3 ms long pulses at a 10 Hz repetition rate).

Seven cavities from 2 manufacturers exceed 20 MV/m even in cw (continuous wave) mode and are actually close to the TESLA goal of 25 MV/m at $Q = 5 \cdot 10^9$. Their excitation curves are shown in Fig. 3.2.7. R & D work for TESLA has also been carried out at Cornell. One 4-cell and three 5-cell cavities reached 25 MV/m after a 1400 °C heat treatment.

A number of cavities reveal limitations and problems which, however, are now well

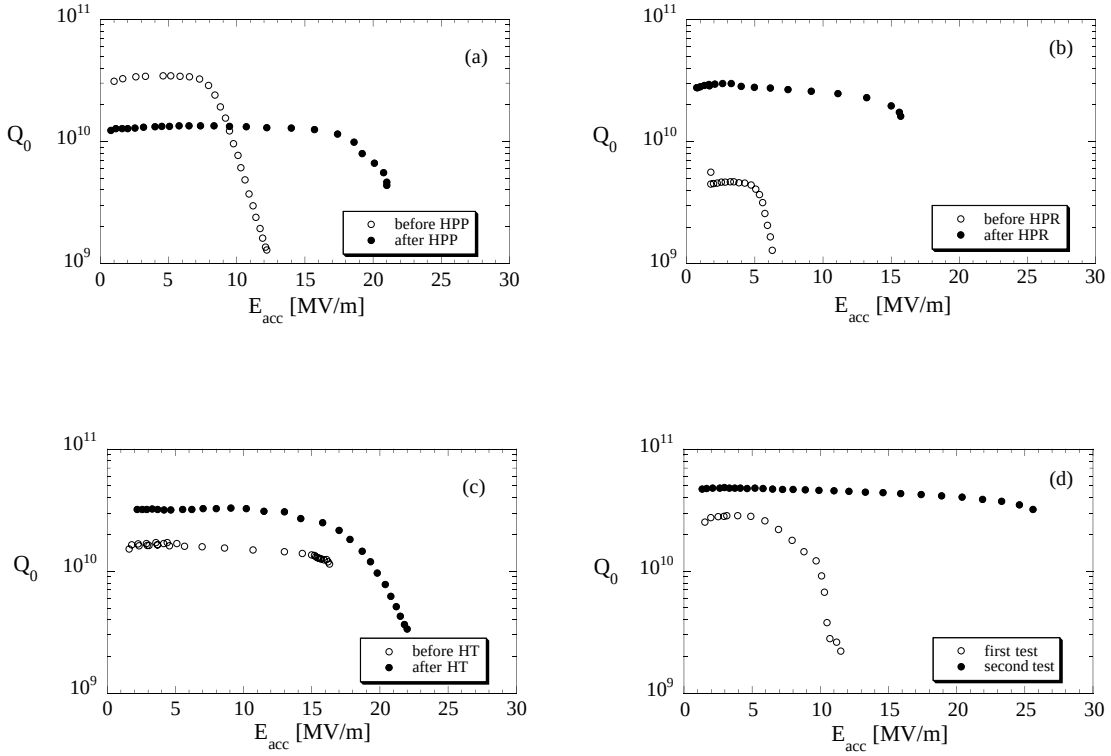


Figure 3.2.6: *Improvement in cavity performance due to various treatments: (a) High power processing, (b) high pressure water rinsing, (c) heat treatment for 4 hours at 1400 °C, (d) removal of surface defects or titanium in grain boundaries. In all four cases shown the cavities were limited by thermal breakdown.*

understood. A material defect has been found in the center cell of a resonator by means of the temperature mapping system. It has been identified as a ≈ 0.5 mm thick tantalum grain in the bulk niobium whose rim is touching the RF surface. As a conseqence the cavity was limited to 13 MV/m in the center cell. This cell has been cut out and the resonator will be repaired by welding in a new cell.

Measurements of all nine coupled modes of the 9-cell resonator are useful to determine the high-field performance of individual cells[3]. The single-cell statistics is shown in Fig. 3.2.8. It is seen that the majority of cells is well above 20 MV/m while only four cells are below the TTF specification of 15 MV/m. The observation that only a few percent of the niobium sheets are impaired by material defects is very important for future cavity production. With the eddy-current scanning method described below it will be possible to eliminate the defective sheets and thereby improve the overall performance.

The cavities from a third manufacturer exhibit a different type of limitation: the

[3]There is an ambiguity owing to the symmetry of the modes with respect to the center cell: it is not possible to distinguish between cell 1 and 9, cell 2 and 8 etc.

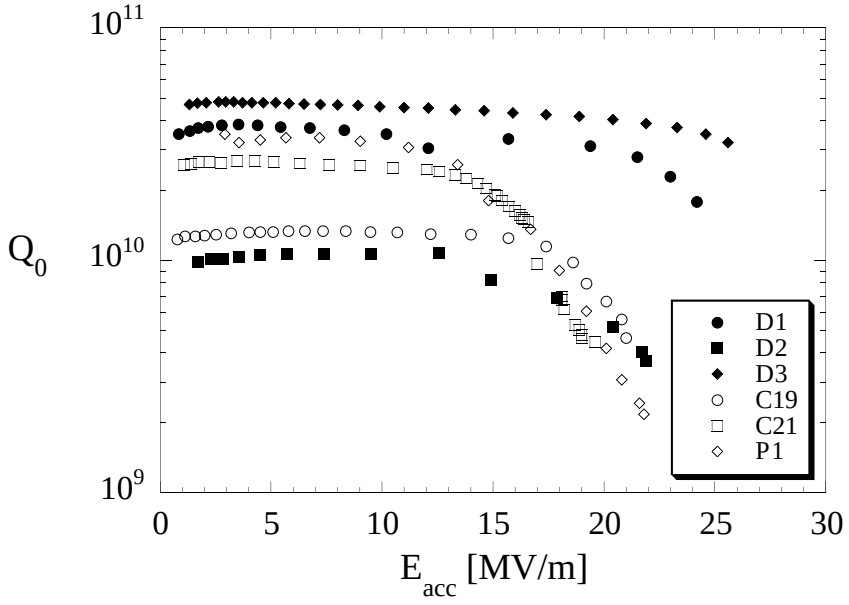


Figure 3.2.7: *The excitation curves $Q(E_{acc})$ of good TTF cavities*

quality factor starts at a high value but drops by a factor of 2 – 3 with increasing field until a quench occurs at about 13 – 14 MV/m in cw operation. Such "Q slopes" are routinely observed in niobium-sputtered copper cavities and are attributed to grain boundary effects. In pulsed mode most of these cavities can be excited beyond 15 MV/m and are hence useable for TTF, but fail the TESLA specification.

Using temperature mapping the origin of the reduced performance has been clearly identified. Spots with unusually strong heating were detected in various equator welds, leading to premature thermal breakdown and to a reduction of quality factor with increasing gradient. With an improved preparation of the weld regions and an optimized electron-beam welding procedure this type of degradation can be avoided in future cavities.

Nine cavities, fully equipped with main power and higher-order-mode couplers, have been tested in a horizontal cryostat prior to their installation in the capture cavity tank and the first cryomodule of the TTF linac, respectively. The average gradient in pulsed mode was 17 MV/m at $Q_0 = 3 \cdot 10^9$, almost as high as the average gradient of 17.5 MV/m achieved in the vertical test in continuous-wave mode.

Material research

Various diagnostic tools are in use to search for inclusions and to study surface defects. DESY collaborates with a number of universities and research centers which have expertise in materials science.

Scanning electron microscopes with energy-dispersive X ray analysis (EDX) have been and are being used to identify foreign elements on the surface. Only a depth of about 1 μm can be penetrated, so one has to remove layer by layer to determine the

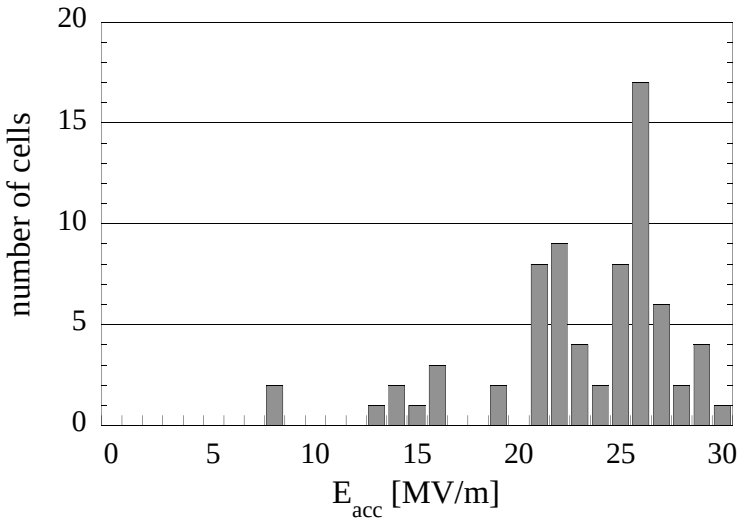


Figure 3.2.8: *Distribution of accelerating field achieved in the individual cells of eight 9–cell TTF resonators.*

diffusion depth of titanium or other elements. Alternatively one can cut the material and scan the cut region. The titanium layer applied in the high temperature treatment extends to a depth of about 10 μm in the bulk niobium. The sensitivity of the EDX method is rather limited, a Ti fraction below 0.5 % is undetectable. Auger electron spectroscopy offers higher sensitivity and using this method titanium migration at grain boundaries has been found to a depth of 60 μm. The microscopic methods are restricted to small samples and cannot be used to study entire cavities.

The narrow-band X ray beams at Hasylab permit element identification via fluorescence analysis. The cut-out cell was investigated this way and tantalum was identified at the spot detected by the temperature mapping system. In principle the apparatus allows the scanning of a whole niobium sheet such as used for producing a half cell, however this procedure is far too time-consuming.

A high-resolution eddy current system has been developed at the Bundesanstalt für Materialforschung (BAM) in Berlin. Using this system the tantalum inclusion mentioned above was easily detectable. For the future production it is foreseen to scan the niobium sheets before they go into cavity production.

Superconductor magnetization measurements offer the possibility to search for magnetic flux pinning. Pinning should be absent in pure niobium but will occur in a NbTi surface layer after the heat treatment with titanium getter. A working apparatus exists and will be routinely used to examine Nb samples which accompany all steps of cavity treatment. A new addition to this apparatus is a device to measure the upper critical field B_{c2} of the sample. This technique was developed at Wuppertal [12]. For pure niobium B_{c2} amounts to 0.27 T while with increasing titanium admixture B_{c2} will rapidly rise to values of many Tesla. The magnetic measurements are foreseen to provide on-line information on the required removal rate in the etching process.

Research and development for cavity improvement

The TTF cavity program and recent results from CEBAF and Cornell demonstrate that accelerating fields of 25 MV/m in multicell cavities are within reach. Two new limiting effects have been found that did not impair previous cavities with design gradients of 5 – 10 MV/m: tantalum or other foreign material inclusions in the niobium sheets and small irregularities in the electron-beam welds. Diagnostic instruments are at hand to avoid these defects. With some additional R&D effort the TESLA design goal of 25 MV/m can be safely achieved. Two routes are followed in parallel.

The first route is based on the present welded design. Here the emphasis is on improved cleanliness and quality control at all stages: niobium ingot production and sheet rolling in a clean environment; eddy current scan of Nb sheets to detect inclusions or defects at an early stage; tight quality control at the companies which are carrying out the deep-drawing and welding of the resonators; optimization of cavity treatment. An important improvement would be to raise the RRR of the niobium ingot material to values of more than 500. In that case a 1400 °C heat treatment of the cavity could probably be avoided.

In the second route a new method of cavity fabrication is investigated, namely hydroforming of the entire multicell resonator from a seamless niobium tube. First attempts with copper have been very promising. If this method should prove practical it would offer considerable advantages, concerning the simplicity of production, the quality of the inner surface and the costs. Increasing the iris radius from 35 to about 50 mm would greatly ease the hydroforming process. Cavities with this larger aperture feature smaller HOM losses and a lower sensitity to detuning [13]. On the other hand the peak electric and magnetic fields would be about 15 % higher than in the TTF cavities so a careful balance of the relative benefits and drawbacks must be found.

3.2.1.5 Cavity Tuning

Lorentz-force detuning and cavity stiffening

The electromagnetic field exerts a pressure on the cavity wall

$$P = \frac{1}{4}(\mu_0 H^2 - \epsilon_0 E^2) \tag{3.2.3}$$

which results in a deformation of the cells and a change ΔV of their volume. The consequence is a frequency shift

$$\frac{\Delta f}{f_0} = \frac{1}{4W}\int_{\Delta V}(\epsilon_0 E^2 - \mu_0 H^2)dV \ . \tag{3.2.4}$$

Here

$$W = \frac{1}{4}\int_V (\epsilon_0 E^2 + \mu_0 H^2)dV \tag{3.2.5}$$

is the stored energy and f_0 the resonant frequency of the unperturbed cavity. The computed frequency shift at 25 MV/m amounts to 900 Hz for an unstiffened cavity

of 2.5 mm wall thickness. The bandwidth of the cavity equipped with input coupler ($Q_{ext} = 3 \cdot 10^6$) is about 430 Hz, hence a reinforcement of the cavity is needed. Niobium stiffening rings are welded in between adjacent cells as shown in Fig 3.2.9. They reduce the frequency shift to about 400 Hz for a 1.3 ms long RF pulse[4]. The deformation of the stiffened cell is negligible near the iris where the electric field is large, but remains nearly the same as in the unstiffened cell near the equator where the magnetic field dominates. No simple scheme exists to reduce the deformation in this region except by increasing the wall thickness.

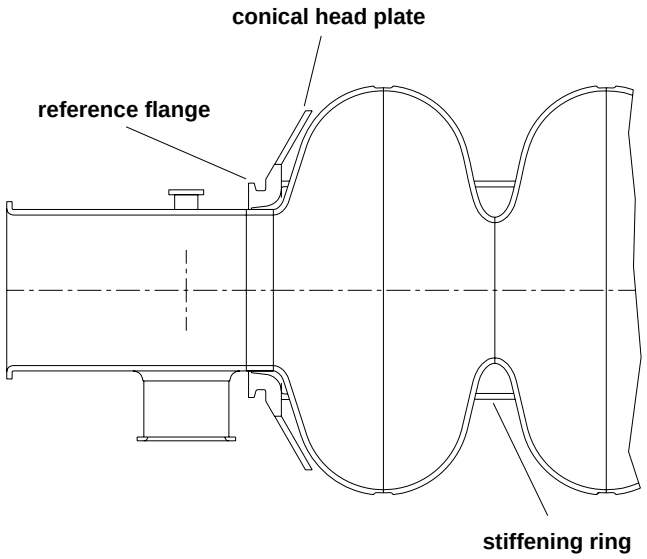


Figure 3.2.9: *End section of a cavity with stiffening ring, conical head plate for welding into the helium tank and reference flange for alignment.*

Tuning system

The goal of the tuning system is to adjust the cavity frequency to the desired value within a few % of the resonance width. Tuning is achieved via variation of the overall length of the cavity.

The tuning system consists of a stepping motor with a gear box and a double lever arm. The thrust is transmitted to the conical head plate at the diameter of the stiffening rings. The moving parts operate at 2 K in vacuum. The tuning range is about ± 1 mm, corresponding to a frequency range of ± 300 kHz. The resolution is 1 Hz. The

[4]Half of this shift is due to cell deformation, the other half to a shortening of the liquid-helium tank surrounding the cavity.

tuning system is adjusted in such a way that after cooldown the cavity is always under compressive force to avoid a backlash if the motion changes from pushing to pulling.

Microphonics

The cavity welded into its helium tank can be considered as a mechanical oscillator with many eigenmodes. The computed mechanical resonances are all above 10 Hz. Mechanical vibrations can be excited by external sources (ground motion, 50 Hz noise, vacuum pumps) but also by the pulsed electromagnetic field. Most of these vibrations will influence the resonance frequency and hence periodically detune the cavity. Details are presented in section 3.2.4.8.

3.2.1.6 Main Power Coupler

Design requirements

A critical component of a superconducting cavity is the power input coupler. For TTF two versions have been developed. The couplers consist of a "cold part" which is mounted on the cavity in the clean room and closed by a ceramic window, and a "warm part" which is mounted after installation of the cavity in the cryomodule. The warm section contains the transition from waveguide to coaxial line. This part is evacuated and sealed against the air-filled wave guide by a second ceramic window. The elaborate two-window solution was chosen to protect the cavity as well as possible against contamination during mounting in the cryomodule and against window failures during linac operation. An important mechanical requirement is that the couplers must allow for up to 15 mm of longitudinal motion inside the 10 m long cryomodule when the cavities are cooled down from room temperature to 2 K. For this reason bellows on the inner and outer conductors of the coaxial line are needed. Since the coupler connects the room-temperature waveguide with the 2 K cavity, a compromise must be found between a low thermal conductivity and a high electrical conductivity. This is achieved by several thermal intercepts and by using stainless steel pipes with a thin copper plating at the RF surface. The design heat loads are 6 W at 70 K, 0.5 W at 4 K and 0.06 W at 2 K. To reduce the danger of multipacting, future couplers will be equipped with a capacitive transition in the inner conductor to permit the application of a dc bias voltage.

Electrical properties

An instantaneous power of 210 kW has to be transmitted to provide a gradient of 25 MV/m for an 800 μs long beam pulse of 8.25 mA. At a repetition rate of 5 Hz this corresponds to an average power of 1.4 kW. The filling time of the cavity amounts to 530 μs and the decay time, after the beam pulse is over, to an additional 500 μs. At the beginning of the filling, most of the RF power is reflected. In the resulting standing-wave pattern the voltage is doubled at the points of constructive interference between incident and reflected wave. To verify that the coupler can safely stand these voltage

enhancements the tests in travelling wave mode have to be performed at four times the nominal power, namely 840 kW.

An adjustment of the external Q of the coupler is achieved by moving the inner conductor of the coaxial coupler section[5]. The range $Q_{ext} = 1 \cdot 10^6 - 9 \cdot 10^6$ is needed to compensate for variations in cavity geometry, couplers and beam current. The input coupler should be impedance-matched (power reflection less than 1%). It must also be capable of transmitting up to 1 MW at a reduced pulse length for *in situ* high power processing of the cavity if the performance should have degraded due to a vacuum failure.

Input coupler A

The coupler version A is shown in Fig. 3.2.10. It has a conical ceramic window at 70 K and a commercial planar waveguide window at room temperature. Variable coupling is achieved by the use of bellows on the inner and outer conductors of the coaxial line. The adjustment mechanism is located outside the cryomodule and permits 25 mm range of motion, which is more than adequate for the required range of Q_{ext}.

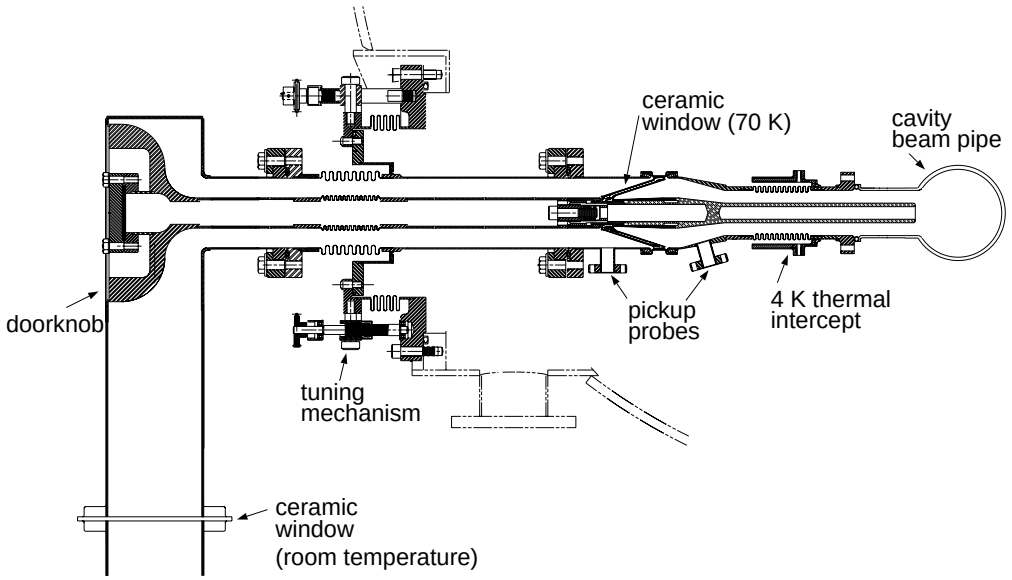


Figure 3.2.10: *Simplified view of the power input coupler version A.*

Flexibility for axial motion during cooldown is achieved via the inner and outer conductor bellows. The pivot point is outside the vacuum vessel. When the coupler

[5]Power coupler plus waveguide can be considered as an external load which extracts energy from the cavity once the pulsed RF has been switched off. The damping provided by the coupler is much bigger than that caused by the surface resistance, hence $Q_{ext} \ll Q_0$, where Q_0 is the quality factor of the "unloaded" cavity.

is mounted at room temperature, the center conductor will be off-center with respect to the input port of the cavity but center itself upon cooldown. The "cold" part of the coupler, together with the conical window, is permanently mounted on the cavity after it has been chemically processed, rinsed, and assembled in the clean room. The "warm" part of the coupler extends beyond the vacuum vessel of the cryomodule and can therefore be installed only after the cavity has been inserted into the vessel. The two coupler parts are joined by a metal-sealed flange at the outer tube and a screw at the inner conductor.

Conical ceramic window A conical shape was chosen for the cold ceramic window to obtain broad-band impedance matching. The Hewlett-Packard High Frequency Structure Simulator program HFSS was used to model the window and to optimize the shape of the tapered inner conductor. The reflected power is below 1 %. The ceramic penetrates into the inner and outer conductor to reduce the field strengths at the braze joints to approximately 50 % of their values at the surfaces of both conductors.

The ceramic window is made from 99.5% Al_2O_3. OFHC copper rings are brazed to the ceramic using Au/Cu (35%/65%) braze alloy. The inner conductors on each side of the ceramic are electron-beam welded, the outer conductors are TIG welded. The ceramic is coated on both sides with a 10 nm thick titanium nitride layer to reduce multipacting.

Doorknob transition and warm window The waveguide-to-coaxial transition is realized using a cylindrical knob as the impedance-transforming device. A planar waveguide window, such as used in commercial klystrons, is mounted upstream. A pair of matching posts is required on the air side of the window for impedance matching at 1.3 GHz. The shape of the doorknob was designed using HFSS. It is impedance-matched over a fairly broad frequency band.

Input coupler B

Coupler version B features two cylindrical ceramic windows, the first at the 70 K level near the cavity, the second at 300 K inside the transition from waveguide to coaxial line. The 70 K window has no direct view of the beam and is insensitive to electron multipacting.

A wave guide of reduced height and two stubs provide broad-band matching between the WR650 waveguide and the 50 Ω coaxial line. The warm ceramic window serves as a vacuum barrier towards the waveguide. Flexibility during cooldown is achieved by two bellow sections in the inner and the outer conductor. The tip of the inner conductor is movable to adjust the external Q. Low cryogenic losses and low RF losses are achieved by using stainless tubes of 0.5 mm wall thickness with a 10-20 μm plating with high-conductivity copper. The bellows of the inner conductor are made of hydroformed thin-walled stainless steel and have sufficient thermal resistance to permit the rest of the inner conductor to be made from copper.

Coupler tests and further development

Prototype couplers have been equipped with ports for diagnostics: a window to observe light effects with a photomultiplier and pick–up probes to measure charged particles. The presence of electrons is an indication of multipacting or plasma discharge and must be avoided since the resulting pressure rise may lead to material evaporation and permanent damage. Five resp. four couplers of each version have been successfully operated and reached a power level of 250 kW after some conditioning time. A simplified and improved version has been designed and is presently under construction. It features two ceramic windows, a 60 mm diameter coaxial line with reduced sensitivity to multipacting and the possibility to apply a dc potential to the center conductor of the coaxial line. In case of the LEP couplers a dc bias has proved very beneficial in suppressing multipacting.

3.2.1.7 Higher Order Modes

The intense electron bunches passing the cavity excite eigenmodes of higher frequency in the resonator which must be damped to avoid multibunch instabilities and beam breakup. This is accomplished by extracting the stored energy via HOM couplers and dissipate it in a load resistor outside the liquid helium container. The HOM couplers are mounted on the beam pipe sections of the nine-cell resonator. Therefore only those higher modes can be extracted which have a sizeable amplitude in the end cells. A problem is related to the so-called "trapped modes" which are concentrated in the center cells and have a low field amplitude in the end cells.

"Trapped" modes

The TE_{121} mode is an example of a trapped mode with a small amplitude in the end cells. In a multicell resonator the end cells are distorted by the presence of the beam pipe. Their shape is adjusted to ensure equal field amplitudes in all nine cells for the fundamental TM_{010} mode (the accelerating mode). This correction does not automatically improve the end-cell field of trapped higher modes. There is, however, an interesting way out. By an asymmetric tuning of the end cells one can enhance the field amplitude of the TE_{121} mode in one end cell while preserving the "field flatness" of the fundamental mode and also the good coupling to the "traveling" modes TE_{111}, TM_{110} and TM_{011}. For this purpose the curvature from equator to iris, the iris ellipse and the iris distance are modified. The effects of asymmetric end cell tuning are summarized in Fig. 3.2.11.

Polarization of dipole modes

Dipole modes can be excited in two orthogonal field patterns. Therefore a single-cell cavity should be equipped with a HOM coupler at each side with a relative angle of 90° between the couplers. In a multicell cavity the situation is more complicated, in principle two couplers on each side are needed to provide damping of trapped modes.

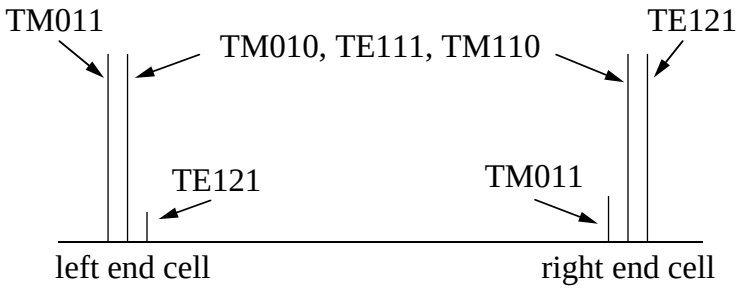


Figure 3.2.11: *Effect of asymmetric end cell tuning on various modes. The accelerating mode TM_{010} and the modes TE_{111} and TM_{110} are not affected while TM_{011} is enhanced in the left end cell, TE_{121} in the right end cell.*

This complexity can be avoided by breaking the azimuthal symmetry of the cavity, e.g. by giving it an elliptical cross section. Thereby the planes of polarization are fixed. For such a "polarized" multicell resonator, only one HOM coupler at each end is needed whose azimuthal orientation with respect to the main axes of the ellipse is 45° resp. 135°. Producing cavities with an elliptical instead of a circular cross section is technically feasible but increases the production costs. Fortunately, in a string of cavities dipole modes of arbitrary polarization can be damped even if the cavities possess cylindrical symmetry. To illustrate the principle we consider again the dipole mode TE_{121}. By asymmetric tuning it acquires a sizeable amplitude in the right end cell, for example. Sufficient damping for both planes of polarization is then provided by the right HOM coupler of the cavity in question and the left HOM coupler of the next cavity in the string, which has the orthogonal orientation. The viability of this idea was verified in measurements.

HOM couplers

The HOM couplers are mounted at both ends of the cavity with a nearly perpendicular orientation[6] to ensure damping of dipole modes of either polarization. A notch filter is incorporated to prevent that energy is extracted from the accelerating fundamental mode. Two types of HOM couplers have been developed and tested, one is mounted on a flange, the other one welded to the cavity.

The demountable HOM coupler is shown in Fig. 3.2.12a. Coupling to the RF field is provided by an open loop whose plane is orthogonal to the beam axis. The loop couples mainly to the magnetic field for TE modes and to the electric field for TM modes. It is capacitively coupled to an external load at 70 K. The rejection filter for the fundamental mode is formed by the inductance of the loop and the capacity at the 1.9 mm wide gap between the lower part of the loop and the wall of the coupling port. Tuning of the filter can be performed without opening the cavity vacuum with help of a niobium bellow. The loop is thermally connected to the 2 K helium bath.

[6] The angle between the two HOM couplers is not 90° but 115° to provide also damping of quadrupole modes.

The coupler has been tested in cw (continuous wave) operation up to an accelerating field of 21 MV/m (this was the field limit of the cavity). The temperature measured at the end of the coupling loop reached a maximum of 4 K which is totally uncritical. In the linac the coupler will be operated in pulsed mode with a duty cycle of 0.6 % hence heating should present no problem.

The welded version of the HOM coupler is shown in Fig. 3.2.12b. It resembles the coupler used in the 500 MHz HERA cavities which have been operating for several years without quenches. The good cooling of the inner conductor by two stubs makes the design insensitive to γ radiation and electron bombardment.

Many copper models of both versions have been built to optimize the coupler dimensions, the position of the ports, and the orientation and penetration of the coupling elements. Pairs of demountable and welded couplers have been tested on the asymmetrically tuned cavity models. The measured values of Q_{ext} are well below the tolerable limits. Both couplers provide sufficient damping for those higher order modes that might lead to emittance growth or beam breakup.

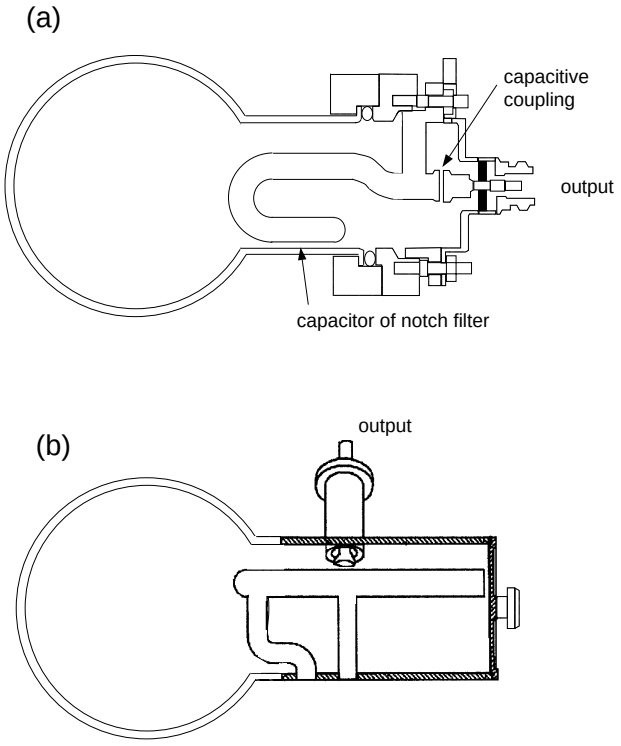


Figure 3.2.12: *The higher-order-mode couplers: (a) demountable HOM coupler, (b) welded HOM coupler.*

Fundamental mode rejection filter Both coupler versions permit tuning of the fundamental mode rejection filter when the HOM coupler is mounted on the cavity. It is possible to achieve a Q_{ext} of more than $1 \cdot 10^{11}$, limiting the extracted power to less than 50 mW at 25 MV/m. During the TESLA operation the rejection filter will remain stable since the couplers are placed in the insulating vacuum. In the welded HOM coupler the rejection filter is quite insensitive to detuning: a detuning by as much as $\pm$20 MHz raises the average extracted fundamental–mode power to just 1 W at $E_{acc} = 25$ MV/m.

3.2.1.8 Helium vessel

The helium tank contains the superfluid helium needed for cooling, supports the cavity and is part of the tuning mechanism. The tank is made from titanium since the thermal contraction relative to niobium is 20 times smaller than for stainless steel. Cooldown produces therefore negligible stress at the joints between cavity and tank and a stress of only $3\,\mathrm{N/mm^2}$ in a cavity that was stress-free at room temperature. With a stainless steel tank this number would rise to an untolerable $60\,\mathrm{N/mm^2}$ unless the tuning mechanism would be operated during cool-down. Titanium has the additional advantage that it can be directly electron-beam welded to niobium while stainless steel-niobium joints would require an intermediate metal.

The assembly of cavity and helium tank proceeds in the following sequence: the bellow unit is electron-beam (EB) welded to the conical head plate at one side of the cavity, a titanium ring is EB welded to the head plate at other side. The cavity is then introduced into the tank and the bellow as well as the titanium ring are TIG welded to the vacuum vessel. The two-phase He tube is rigidly connected to one end of the He tank and equipped with a bellow at the other end to permit longitudinal expansion relative to the He tank.

3.2.1.9 Magnetic Shielding

In principle, a perfect type II superconductor should expel the magnetic flux from its volume when cooled down in a static magnetic field smaller than B_{C1}. In niobium cavities, however, the magnetic flux is found to be trapped in the walls if the cavity is cooled down in a small external field of a few Gauss. The pinning of the flux lines is attributed to lattice imperfections in the niobium or to surface contamination. Trapped magnetic flux gives rise to an RF dissipation, which can be represented by a local surface resistance:

$$R_{surf} \simeq R_n \frac{B_\perp}{B_{c2}}$$

where R_n is the normal-state surface resistance of niobium and $B_\perp$ is the component of the ambient magnetic field that is perpendicular to the cavity surface. For a TESLA cavity this corresponds to an average surface resistance of 0.35 nΩ/mGauss, in good agreement with experimental results. In order to achieve $Q_0 = 3 \cdot 10^{10}$, the residual

surface resistance due to trapped flux must be smaller than 3 nΩ, corresponding to a remanent field of < 10 mGauss.

At the latitude of Hamburg, the earth magnetic field has a magnitude of 400 mGauss with roughly equal horizontal and vertical components. The necessary magnetic field attenuation is achieved by combining passive shielding and active compensation with Helmholtz coils. A two-step passive shielding is provided by the vacuum vessel of the cryomodule and by high-permeability cylinders around the individual cavities. The vacuum vessel is made from conventional steel with a remanent field in the order of several Gauss. To remove the remanence a "premagnetization" technique [14] has been applied. For this purpose a 40-turn toroidal coil is mounted on the cylindrical vessel and subjected to a current cycling between $\pm I_0$ at a rate of 6 A/s where I_0 starts at 400 A and is gradually reduced to zero. The resulting attenuation of the ambient field is found to be better than expected from a cylinder without any remanence. The explanation is that the procedure does not demagnetize the steel but rather *premagnetizes* it in such a way that the dangerous axial component of the ambient field is counteracted. This interpretation becomes obvious if the cylinder is turned by 180°: in that case the axial field measured inside the steel cylinder is almost twice as large as the ambient longitudinal field component. In Fig. 3.2.13a we have plotted the ambient field in the experimental hall and the field along the axis of the steel vessel after premagnetization. The striking observation is that the steel vacuum vessel by itself is almost capable of achieving the desired reduction to 10 mGauss, except near the ends. However, it is important to apply the premagnetization procedure at the exact location and orientation of the cryomodule.

The shielding cylinders of the TTF cavities are made from Cryoperm which retains a high permeability of more than 10000 when cooled to liquid helium temperature. Figure 3.2.13b shows the measured horizontal, vertical and axial components inside a cryoperm shield at room temperature, which was exposed to the Earth field of a few hundred mGauss. The combined action of premagnetized vacuum vessel and cryoperm shield is obviously more than adequate to reduce the ambient field to the level of a few mGauss. An exception are the end cells of the first and last cavity near the end of the cryomodule where the vessel is not effective in attenuating longitudinal fields. Here an active compensation is possible. With an arrangement of three Helmholtz coils, mounted on the end section of the steel vessel and the interconnection to the next cryomodule, the fringe field at the last cavity can be reduced to a harmless level.

3.2.2 The Module

The module described here is essentially identical to the one presently built for the TESLA Test Facility [15]. Based on the experiences that will be gained at the TTF a redesign will be made in due time.

The present TTF layout of the cryostat was largely driven by the goal to reduce the cost compared to existing superconducting cavity systems.

The concept is to build large modules incorporating many 9–cell cavities (eight at present), quadrupoles, steering coils and a beam position monitor and to combine

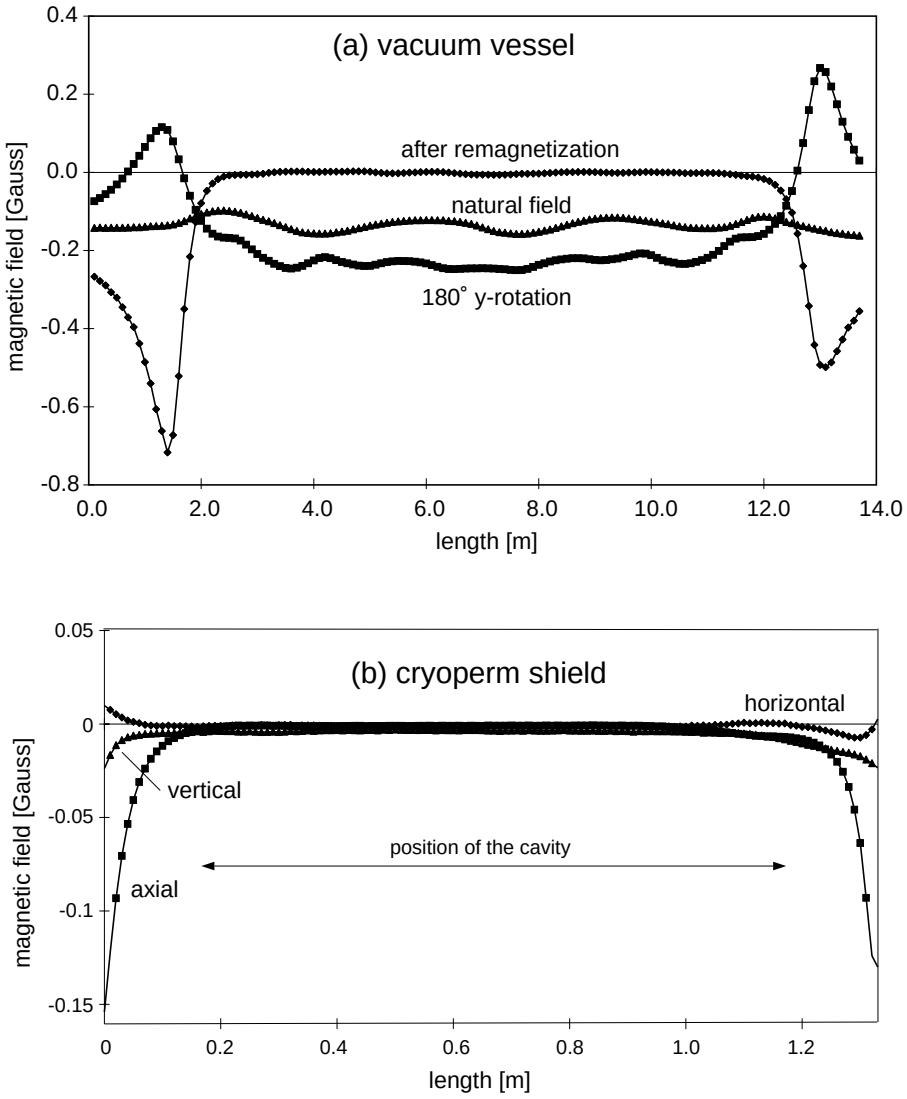


Figure 3.2.13: *Shielding of Earth magnetic field. (a) Shielding of axial component by the steel vacuum vessel of the cryomodule. (b) Shielding of axial, horizontal and vertical components by the cryoperm shield surrounding the cavity (without vacuum vessel).*

these modules to a continuous cryogenic string of about 2.5 km length. The helium distribution system needed to operate the superconducting cavities at 2 K and the magnets at 4.5 K at low thermal losses is also integrated into the cryostats. There will be only one cryogenic supply box at the end of the string. This basic concept avoids the large number of warm to cold transitions and the separate cryogenic supply boxes which constitute a major fraction of the cost of present superconducting cavity systems.

3.2.2.1 The Cryostat

Figure 3.2.14 shows a longitudinal view of the cryostat. The 300 mm diameter helium two-phase pipe is the main support beam for the accelerator module. The two-phase pipe is supported from above by three support posts that provide the necessary thermal insulation from room temperature. These posts are fastened to large flanges on the upper part of the vacuum vessel by adjustable suspension brackets. The purpose of the adjusting mechanism is to allow for the accelerator module axis to be brought into its proper position, independent of the absolute position of the vacuum vessel flanges on which it rests.

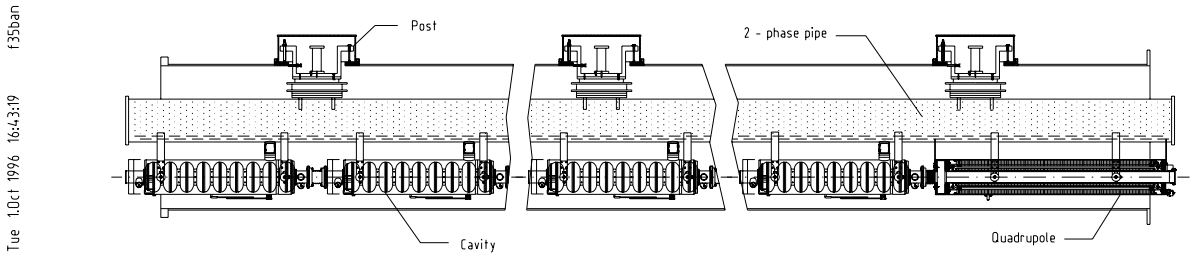


Figure 3.2.14: *Longitudinal view of the cryostat*

In the longitudinal direction the center post is fixed to the vacuum vessel while the two end brackets are allowed to move in the axial (z) direction to accommodate differential shrinkage during temperature cycling. The posts consist of a fiberglass pipe terminated by two shrink - fit stainless steel flanges. At appropriate intermediate positions, optimized to minimize the heat leak, two additional shrink - fit flanges are provided to allow intermediate heat sink connections to 4.5 K and 70 K. The post diameter has been chosen to push up the assembly main mechanical resonance frequency, corresponding to a pendulum-like motion in the x - y plane to 19 Hz and in the x - z plane to 20 Hz. These resonances are significantly above the dangerous 5 Hz (rep. rate) frequency. A further shift of the latter resonance away from the integer multiple of the rep. frequency seems necessary, though.

The 8 cavities (and the quadrupole package in the case of some of the modules) are attached to the two-phase pipe by means of stainless steel brackets equipped with adjusting screws for alignment.

The support system is designed to allow the accelerator module to move with respect to the vacuum vessel during thermal cycling. As the center point of the assembly is fixed to the vacuum vessel, the outer ends will be displaced longitudinally by up to 15 mm. The cavity main couplers must have sufficient flexibility to accommodate a displacement of the same magnitude. For a redesign of the cryostat rigid input couplers will also be investigated.

The cryostat includes two aluminum radiation shields at nominal temperatures of 4.5 K and 70 K. Each shield is divided in half longitudinally and each half consists of one upper section and six lower sections. The upper sections are supported by the post intermediate flanges. They are fixed to the center post but can slide on the end posts in the axial direction to avoid buildup of stress during thermal cycling. Experience with the earlier TTF cryostat design has shown that to reduce costs it is best to use cooling pipes that are directly welded to the aluminum shields and to weld the shields together where required. This design is similar to that successfully used on the SSC, LHC and HERA magnet cryostats.

Blankets of Multilayer Insulation (MLI) are placed on the outside of the 4 K and 70 K shields. The 4 K shield blanket is made of 10 layers while the 70 K blanket contains 30 layers. In addition, the cavity helium vessels, gas return pipe and 4 K pipes are wrapped with 5 layers of Multilayer Insulation to reduce heat transfer in the event of a vacuum failure.

The cryostat outer vacuum vessel is constructed of carbon steel and is evacuated to a pressure of less than 10^{-6} mbar during operation. Its design critical internal over pressure is 1.25 bar. Adjacent vacuum vessels are connected to each other by means of a cylindrical sliding sleeve. The vacuum vessel of the cryostat for the TTF are connected by sliding sleeves with O-Ring seals similar to the connection of the HERA s.c. magnets. For the linear collider a cheaper welding solution will be used (see also section 3.2.3). Radiation shield bridges are also provided. A relief valve on the sleeve and appropriate venting holes on the shields prevent excessive pressure build-up in the vacuum vessel in the event of accidental spills of liquid He from the cavity helium vessels. Wires and cables of each module are brought out at the module itself using flanges equipped with vacuum tight connectors. Figure 3.2.15 shows a cross section of the cryostat.

The following piping necessary for the cryogenic operation of the linac is incorporated in the cryostat:

1. The 2 K forward line: This line transfers single phase helium through the cryostat to the end of the string of modules furthest away from the refrigerator.

2. 2 K two-phase line: This line supplies the cavities with liquid helium and returns the cold gas pumped off the saturated He II baths to the refrigeration plant. This line is connected to the cavity helium vessels via a tube containing bellows for adjustment purposes. This line is also a key structural component of the cryostat (see the alignment and assembly section).

3. The 4.5 K forward and return lines: The 4.5 K forward line cools the quadrupole/-correction magnet package. The 4.5 K return line cools the 4.5 K radiation shield and the 4.5 K intercept on the main coupler.

4. The 40 K forward and 70 K return lines: The 40 K forward line cools the High Temperature Superconducting (HTS) current leads of the quadrupole and correction magnets. The 70 K return line cools the 70 K radiation shield, the 70 K

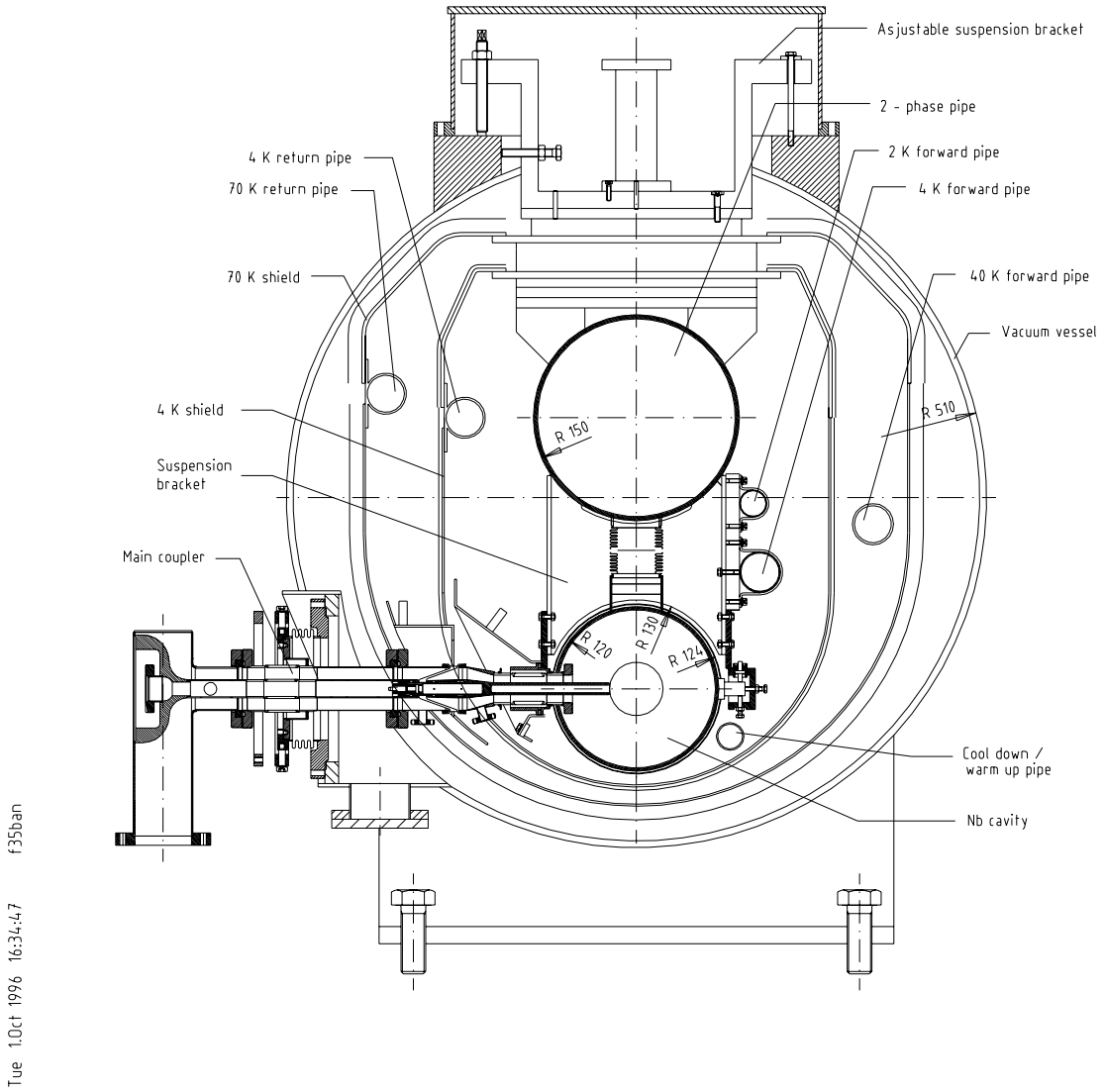


Figure 3.2.15: *Cross section of cryostat*

intercept on the main coupler and the HOM absorber in the cryomodule interconnection region.

5. Warm up/cool down line: This line parallels the two-phase supply line and connects to the bottom of each of the cavity helium vessels. It is used during cool down and warm up of the cryostat.

All cryogenic lines with the exception of those connected to the cavity He vessel must be able to withstand a test pressure of 20 bar absolute. The lines connected to the cavity He vessel must be able to withstand 4 bar absolute.

The helium lines of adjacent modules will be connected by welding similar to the

connection of HERA s.c. magnets making use of welding machines. The beampipes of adjacent mudules will be joined by flange connection (see section 3.2.3).

The cryostat must maintain the cavities and quadrupoles at their operating temperatures. For the cavities this temperature is 2 K and for the quadrupoles the operating temperature is 4.5 K. The static heat load is determined by the cryostat design while the total heat load is dominated by the RF losses and thus principally determined by the cavity performance. Table 3.2.3 shows the predicted heat load for a cryomodule.

The alignment requirements are driven by the beam dynamics. The result is that the axes of the 8 cavities must be aligned to the ideal beam axis to within 0.5 mm and those of the quadrupoles to within ± 0.1 mm within one betatron wavelength (the latter tolerance is less critical if beam-based methods to align the quadrupoles are applied). Additionally, the vertical mid plane of the quadrupole package must be aligned to the vertical direction to ≤ 0.1 mrad.

The alignment of the components is done while the cryostat is warm. In the TTF cryostats it will be investigated how well these alignments are maintained during installation, cool down and operation.

In order for the cavity performance not to be degraded by trapped magnetic fields, the cryostat must reduce the ambient magnetic field at the outer wall of the cavities to $\leq 0.5\,\mu$T.

In order to meet the $\leq 0.5\,\mu$T residual field requirement, a combination of active and passive magnetic shields is used. The active shield consists of a series of Helmholtz coils wrapped around the outside of the vessel. When powered, these coils will do a good job of canceling out the average longitudinal field. Passive shields made by cylinders of Cryoperm placed around the cavity helium vessels serve to reduce the remaining magnetic field. The vacuum vessel is made of soft steel and is demagnetised before assembly of the cryomodule.

The TESLA beam optics design requires different quadrupole strengths along the length of the linac. As a result, there are 7 different types of cryostats. The standard one without a quadrupole and 6 others with quadrupoles of different lengths (see Table 3.2.4)

The standard cryostat is 11.37 m long and has an outer diameter of roughly 1.2 m. Each cryostat contains 8 superconducting RF cavities (each enclosed within its own helium vessel).

3.2.2.2 Quadrupole and Correction Magnets

The quadrupole package is shown in Fig. 3.2.16. It consists of

- a superferric quadrupole enclosed in a stainless steel vessel cooled by supercritical helium at 4.5 K (at least less than 5.2 K),
- a pair of dipole correction coils wound on the beam pipe inside the quadrupole to provide horizontal and vertical correction dipole field,
- an RF beam position monitor (BPM) consisting of a pill box RF cavity, rigidly connected to the quadrupole,

	static (W)	dynamic (5 Hz)+static (W)
2 K		
RF load		7.32
radiation		
supports	0.60	0.60
input coupler	0.24	0.36
HOM load		1.54
HOM coupler	0.10	0.32
HOM coupler cable		0.35
beam tube bellows (9)		0.20
cables	0.12	0.12
Total	**1.06**	**10.81**
4 K		
RF load		
radiation	3.98	3.98
supports	2.40	2.40
input coupler	2.52	3.24
HOM load		
HOM coupler	1.00	1.69
HOM coupler cable		1.13
cables	0.37	0.37
Total	**10.27**	**12.81**
70 K		
RF load		
radiation	22.74	22.74
supports	6.00	6.00
input coupler	18.00	30.60
HOM absorber		16.11
cables	3.72	3.72
Total	50.46	79.17

Table 3.2.3: *Heat load for module of 11.37m length (without quadrupole)*

- three pairs of conduction cooled current leads where the cold part is made from High Temperature Superconductor (HTS). The current leads run from the liquid He vessel to a flange on the cryomodule vacuum vessel.

Quadrupoles Each linac contains quadrupoles for beam focusing, incorporated at the beam exit end of each second cryomodule up to an energy of 50 GeV, at each third cryomodule between 50 GeV and 150 GeV and at each fourth cryomodule between 150 GeV and 250 GeV. The required integral field gradient generally varies linearly

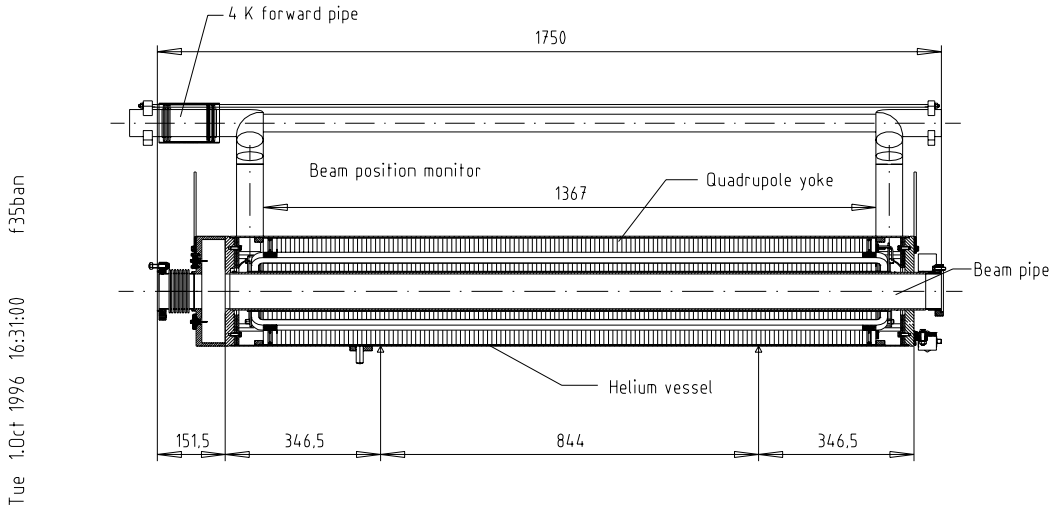


Figure 3.2.16: *Quadrupole package (longitudinal cut)*

with beam energy reaching a maximum of 31.03 T. This value already includes an upgrade of the accelerating voltage gradient in the cavities from the nominal 25 MV/m to 40 MV/m which would result in a 400 GeV beam energy for each linac. Because of the change in quadrupole per number of modules there is a step in the required integral field gradient at 50 (80) GeV and at 150 (240) GeV.

Due to easy manufacturing, easy alignment and low amount of superconductor required the quadrupoles are chosen of the superferric type with a maximum gradient of 22.7 T/m. As the integral field gradient scales with the beam energy and in order to provide maximum flexibility for beam-based alignment and beam optics, all quadrupoles need individual currents. Therefore, in order to limit the required cooling power, the maximum current is limited to less than 100 A.

With the above mentioned maximum field gradient of 31.03 T the maximum length of the quadrupole field is 1.367 m. The total length of the quadrupole package including coil ends, mirror plates, helium vessel end plates, beam position monitor, bellow and connection flanges amounts to 1.750 m. This length generally is needed only in the cryogenic unit of the high energy end. At lower energies the length of the quadrupole package is reduced in steps generally from unit to unit. This leads to 6 different lengths of quadrupole packages as shown in Table 3.2.4.

A later doubling of the linear collider energy could be achieved by doubling the whole length of the collider and shifting the collision point from between the first two linacs to the end of the second linac (the end of the whole collider in the first version). In this case, the electrons are accelerated in the second linac from 250 (400) to 500 (800) GeV running through it in a direction opposite to the previous one. For beam focusing in this extended version, quadrupoles of full strength (31.03 T) are needed in every fourth cryomodule up to 560 GeV, and in every sixth cryomodule up to the end (800 GeV). The same structure is foreseen for the first linac which would also permit

type	int. grad. (T)	max. grad. (T/m)	field length (m)	add. length (m)	quad. pack. length (m)
B	6.14	22.7	0.270	0.383	0.653
C	10.94	22.7	0.482	0.383	0.865
D	14.06	22.7	0.619	0.383	1.002
E	20.84	22.7	0.918	0.383	1.301
F	25.92	22.7	1.142	0.383	1.525
G	31.03	22.7	1.367	0.383	1.750

Table 3.2.4: *List of different quadrupole packages*

quadrupole type	number (standard version)	number (extended version)
B	108	72
C	96	64
D	146	78
E	136	68
F	224	40
G	102	490
Total	812	812

Table 3.2.5: *Required numbers of quadrupoles of different types (total collider)*

the beam to run in the opposite direction for collisions with protons from HERA. The required numbers of quadrupoles for the standard and extended versions of TESLA are listed in Table 3.2.5.

The superconducting coils are of the racetrack type, wound from a wire of rectangular cross section (about 200 turns per pole) and impregnated with epoxy. The yoke is made from 5 mm thick punched laminations assembled on a tool and locked through keys which give the required positional accuracy (0.02 mm). The keys are connected through pins and bolts to the outer stainless steel helium vessel tube surrounding the quadrupole and to the stainless steel end plates of the vessel.

Both end plates are covered with soft steel magnetic mirror plates for reducing the fringe field at the position of the superconducting cavities. The helium vessel endplate on the beam entrance side of the quadrupole is also a part of the beam position monitor (pill box type).

The main parameters of the quadrupoles are listed in Table 3.2.6.

Correction Dipoles The purpose of the correction dipoles is beam steering. They are layed out for a maximum orbit steering range of ± 3 mm per correction coil. The maximum integral field required for this purpose is 31 mT·m at the high energy end. However the integral field scales with energy in the same manner as the quadrupoles leading to 12.4 mT·m at 50 (80) GeV (with a dipole in every second cryomodule) and

quadrupole type	superferric
beam tube inner diameter	78 mm
pole tip (yoke inner) radius	45 mm
yoke outer radius	116 mm
yoke length	
B	0.270 m
C	0.482 m
D	0.619 m
E	0.918 m
F	1.142 m
G	1.367 m
coil cross section	16 mm × 13 mm
max. current	100 A
max. field gradient	22.7 T/m
max. field at conductor	2.1 T
b6 at r = 10 mm	$3.5 \cdot 10^{-4}$
b10 at r = 10 mm	$8.2 \cdot 10^{-7}$
conductor	rectangular cross section

Table 3.2.6: *Main parameters of quadrupoles*

24.7 mT·m at 150 (240) GeV (with a dipole in every third cryomodule). There are two dipole coils in each quadrupole, one for horizontal and one for vertical correction. The coils are one layer windings of 15 turns of round multifilamentary superconducting wire on the beam pipe which closes the helium vessel from the inside. They have the same length as the quadrupole. The maximum current is ~ 33 A. The main parameters of the correction dipoles are listed in Table 3.2.7.

Beam Position Monitor The beam position monitor (BPM) is of the pill box type attached to the beam entrance end plate of the helium vessel. It is equipped with two antennas for each direction (x and y). The beam position measurement resolution is aimed at 10 μm and the installation tolerance w.r.t. the quadrupole axis at 0.1 mm. The beam pipe part between the beam position monitor and the cavity contains a stainless steel bellow to compensate for misalignments and cooldown motions. With the exception of the beam position monitor cavity, the whole beam pipe of the quadrupole package as well as the bellows are copper coated inside.

Heat Loads The heat loads calculated for the quadrupole package of maximum length are listed in Table 3.2.8. As can be seen, due to the use of HTS current leads the heat load on the 4 K level is rather low. The load on the 2 K level takes the connection to the cavities into account. Dynamic losses occur at the beam tube bellows and at the beam position monitor.

dipole type	one layer cosΘ in superferric quadrupole
beam tube inner diameter	78 mm
pole tip (yoke inner) radius	45 mm
yoke outer radius	116 mm
yoke length	
B	0.270
C	0.482
D	0.619
E	0.918
F	1.142
G	1.367
average coil radius	42.5 mm
max. current	33.2 A
max. Ampere turns	1494 A
max. field	44 mT
number of turns	45
max. field at conductor	0.98 T
b3 at r = 10 mm	$7.8 \cdot 10^{-3}$
b10 at r = 10 mm	$9.5 \cdot 10^{-4}$
superconducting wire	round, 0.98 mm Ø

Table 3.2.7: *Main parameters of correction dipoles*

Quadrupole Supports and Alignment The quadrupole package is supported at the 300 mm diameter two-phase pipe in the same manner as the cavities. The system consists of two brackets (one at each end of the helium vessel). Bolts in the brackets allow the adjustment of the quadrupole unit with respect to the ideal beam axis inside the cryomodule with the aid of reference targets on temporary extension arms attached to each helium vessel end.

Alignment and survey play an important role because of the given tolerances (Table 3.2.9, rms values).

The alignment of the quadrupole is the most critical one. In order to achieve the required tolerances we have chosen:

- a superferric quadrupole where the accuracy of the field is mainly given by the accuracy of the yoke;
- a laminated yoke where the contours are most accurate by punching and due to punching keys locating the yoke laminations precisely;
- a helium vessel with pins to locate the keys and yoke inside the vessel;
- grooves in the endplates accurately machined to take over the position of the keys;

	static (W)	dynamic (5Hz) + static (W)
2 K		
bellows		0.02
BPM		0.07
radiation		
conduction	0.10	0.10
cable	0.01	0.01
Total	**0.11**	**0.20**
4 K		
BPM		0.10
radiation	0.61	0.61
current leads	0.02	0.03
cable	0.03	0.03
Total	**0.66**	**0.77**
70 K		
radiation	3.50	3.50
current leads	14.03	32.04
cable	0.26	0.26
Total	**17.79**	**35.80**

Table 3.2.8: *Loads of a full length (1.75 m long) quadrupole package*

component	tolerance (rms)
cavity position	0.50 mm
quadrupole position	0.10 mm
quadrupole field angle	0.10 mrad
quadrupole jitter	1 μm
BPM position	0.1 mm

Table 3.2.9: *Tolerances of alignment*

- reference targets at the helium vessel end plates machined in one set-up with the endplates;
- mechanical integration of the beam position monitor with the helium vessel end plate.

At the TTF accelerometers will be attached to the cold mass of the quadrupoles in order to measure vibrations in x and y direction. The measuring sensitivity at 2 Hz is about 20 nm and is mainly determined by the noise of the amplifier. The actual vibration tolerance in the TESLA linac is as large as 1 μm, because orbit jitter is

removed within one beam pulse by means of the fast orbit feedback.

3.2.2.3 Assembly and Alignment

The assembly and alignment of the cryostat are closely linked together.

First, the two-phase pipe (TPP), which is the basis for the support system, is straightened to better than ± 3 mm and heat treated to relieve any residual stress that might cause deformation later in the assembly process. Then this pipe is transferred to a work bench on which an ideal beam axis is defined. The pipe is placed on three supports that reproduce the effect of the three posts in the final cryostat. While on these supports the pipe is adjusted so that the centers of its end sections define the pipe reference axis to within ± 2 mm. The brackets and the support post flanges are then welded on to within ± 1 mm of their ideal position. The adjustable suspension brackets (ASBs) and the cryostat fiberglass support posts are also attached with this precision.

Next, the upper parts of the thermal radiation shields, cryogenic pipes and cabling are attached. This portion of the cryostat i.e. the structure excluding the cavity string and vacuum vessel is referred to as the cold support system. Three optical targets (Taylor - Hobsen spheres) are attached to the top of the three cryostat support posts. Using the ASBs these targets, and thus the posts and the TPP, are aligned within 0.1 mm to the previously defined ideal beam axis. They are also aligned to a horizontal reference plane to within 0.2 mm/m. The cold support system is flexible to a degree and may go out of this alignment during transport and further assembly. However, as long as the optical targets can be realigned using the ASBs, the alignment of the cold support system relative to the ideal beam axis can be recovered.

In the meantime, an ideal axis is defined for each cavity before it is welded into its helium vessel. An ideal axis is also defined for the quadrupole package which is also equipped with reference arms.

The objective of the final assembly of the cryostat is to attach the cavities and magnets to the cold support system and align the ideal axes of the cavities and magnets with the ideal beam axis defined on the cold mass. Of course, this also has the effect of aligning the cavity and magnet ideal axes with each other.

This work starts in the clean room where the cavities and quadrupole package are connected to each other by flanges. The cavities are aligned in the proper axial position (based on the position of the main coupler) and aligned so that the main coupler port axis is parallel to the floor. This cavity string is now moved out of the clean room to the first assembly station. At this station, the cold support system is lowered onto the cavity string attached to it via the support brackets and the connections between the helium vessels and the TPP are welded. The Cryoperm magnetic shielding is then installed. Next the cavities and quadrupole package are aligned to the ideal beam axis of the cold mass using adjustment screws at the support brackets. Bellows between the cavities on the beam tube and at the connections between the helium vessels and the TTP allow adjustment of the cavities relative to each other and to the quadrupole package. Once the alignment is finished the adjustment screws are locked in place via

set screws. The system is designed to keep the components in their aligned position during thermal cycling. The assembly is moved to the second station where the lower part of the thermal radiation screens, the MLI blankets, and remaining cabling are added.

This completed assembly is then moved to the third station where it is suspended between two cantilevers. The completed assembly is then inserted into the vacuum vessel. During this process, the ASBs are removed. Appropriate pins ensure that the ASBs can be replaced to $< \pm 0.1$ mm of their original position.

Once the assembly is installed in the vacuum vessel, the Taylor - Hobsen spheres on top of the ASBs are brought back into alignment thus returning all components back to their appropriate position relative to the ideal beam axis of the cold mass.

As a final step, the axis defined by the Taylor - Hobsen spheres is transferred to another set of reference marks on the side of the vacuum vessel. It is these reference marks that are used to align the vacuum vessel (and thus the cold mass) to the beam axis of the linac.

3.2.3 Vacuum System

There are three vacuum systems on the collider linacs with different requirements: The beam vacuum system, the insulating vacuum system and the vacuum system for the room temperature part of the RF input couplers. The layout concepts of the systems are based on the experiences gained at HERA and in the TESLA Test Facility. With more experience from the TTF linac the concept may have to be modified.

3.2.3.1 The beam vacuum system

One of the major objectives for the beam vacuum system is to preserve the cleanliness of the superconducting cavities' surfaces and thus the operation at high gradients and high Q. So contamination by any sort of dust or condensation of gases during assembly and operation has to be absolutely avoided.

Therefore the string of 8 cavities for a module is at present assembled and the connections between adjacent parts are done inside the clean room. All parts which are connected to the cavity, input couplers for example, have to fulfill the same requirements concerning cleanliness and have to undergo similar treatments before the assembly. The string of 8 cavities and the beam pipe inside the quadrupole is evacuated and closed off by manual all metal valves in the clean room after assembly. There is some concern on the production of metallic dust by these valves. However, as they are only operated a few times this seems not very probable. This will be investigated in the Test Facility.

The present solution to connect adjacent cavities by flanges and a rather long compensator, which is necessary for independent tuning and mechanical decoupling during thermal cycles, fulfills the requirements on cleanliness. The compensator is a copper coated stainless steel bellow. For the linear collider a high filling factor saves on tunnel length. It will be designed for to keep the cavities as close together as possible

and to keep the bellows between cavities as short as possible. RF-shielding of the connecting bellows has been avoided in the present test setup as to prevent rubbed off dust particles which eventually could enter the cavity and cause field emission. The wake fields generated by the bellows are tolerable as they constitute only 12 % of the unavoidable longitudinal wake fields from the cavity itself. However, shortening of the bellows for the linear collider will be beneficial.

Investigations are planned to eliminate the flanges at the interconnection between cavities and to the input coupler by an adequate welding technique.

The way the beam vacuum of adjacent modules is connected in the linac tunnel is again dictated by the necessity to prevent dust from entering the cavity vacuum. The beam pipe section between the modules is installed in a local clean room, similar to the one which is used in the TTF–linac to assemble the vacuum system. The connecting beam pipe contains stainless steel bellows to allow for the thermal shrinkage of the cavity string during cool down. It is planned to also incorporate a HOM absorber into this section, which is at present in the beampipe inside the quadrupole. Its purpose is to absorb power from travelling HOM fields at a temperature level of 70 K, thus reducing the thermal load of the 2 K level.

The connecting beam pipe is pumped via a pumping port and a manual all-metal valve by an oil free pump station to avoid any pollution of the cavity system by hydrocarbons, which increase the probability of field emission. The gate valves to the neighboring cavity sections will only be opened when the connecting beam pipe is clean and evacuated.

The present layout of the beam vacuum does not foresee any sector valve within the cryogenic section of about 2.5 km length. Only at the sector ends short room temperature sections between adjacent cryogenic sections are foreseen which contain sector gate valves.

Ten permanent pumping ports are foreseen for a sector. The connection to the pumping port of a connecting beam pipe between modules is done via a double knee shaped pipe which penetrates the heat shields and the vacuum vessel. These ports will be used to pump the vacuum sections when the cavities are at room temperature and to make sure that there are no helium leaks in the system. The pumps will not be connected to the beam vacuum when the cavities are cold as the base pressure of the pump station is higher than the pressure in the beam vacuum.

In case a module would have to be removed from the tunnel, the manual gate valves to the neighboring modules are closed. If venting of the module is necessary it will be done with filtered, dust free and pure argon at laminar flow to avoid any contamination or transport of dust by turbulences.

In a vacuum system which is in great part surrounded by superfluid helium, one has to worry about leaks. In the design of the cavities the input coupler and the HOM coupler are located outside the helium vessel, thus avoiding the situation of flange connections or feed-throughs being exposed to liquid helium. Thus only through leaks in the wall of the cavity itself can helium get into the beam vacuum. The cavities which have been tested in the test facility have welds made by electron beam welding on the equator and the iris of each cell. Up to now no leak has been detected at a sensitivity

of 10^{-14} mbar/sec. The risk of leaks will certainly get even less when hydro-formed cavities become feasible.

3.2.3.2 Insulating vacuum system

The layout of the insulating vacuum system is strongly influenced by the experiences gained with the insulating vacuum system of the superconducting magnets at HERA.

As the O-Ring seals in the sliding sleeves which connect adjacent vacuum vessels caused a lot of problems at HERA due to leaks and large permeation it is foreseen to connect adjacent vacuum vessels by welding. The connections will be done with welding lips which allow for several openings by cutting and consequent welding by special tools. The penetrations of the vacuum vessel for the input couplers are sealed by O-Rings in the present cryostat. For the final layout of the cryostat a welding connection is also aimed for at these penetrations. The vacuum vessel is made from cheap carbon steel protected with an appropriate paint.

The various helium pipes will be connected between modules by welding in a similar way as it was done at HERA. There the experience with welding of the helium pipes between neighboring superconducting magnets has been very positive. Only a fraction of less than 0.1 % had a leaky weld. Very similar procedures to the ones established at HERA will be used to assure leak tightness or to localise leaks.

At HERA there is a vacuum barrier in each quadrupole cryostat that means about every 25 m. This barrier essentially is for free as it was used as mechanical support of the cold mass. These barriers were very helpful during installation and commissioning of the vacuum system and speeded up the leak check of the outer vacuum vessel and localisation of helium leaks. During operation, however, these barriers are permanently bridged by bypass tubes. Due to the larger number of helium tubes and the large diameter of the 2 K gas return tube a vacuum barrier is complicated in the case of the linear collider modules. At present it is assumed that there will be no vacuum barrier at all within a 2.5 km long sector. However, this topic needs further discussions and more detailed layout of the whole cryogenic system.

Again from the experience gained at HERA permanent pump stations will be necessary only every 250 m during normal operation due to the large conduction of the cryostat. However, there will be additional pumping ports at about 50 m distance which allow to connect movable pump stations for initial pump down and in case of a larger helium leak.

During pump down and leak test, about 50 pump stations will be concentrated in one sector. Once the sector is commissioned they will be moved to the next sector with only 10 pump stations remaining per sector. There will be no gauges on the system except at the pump stations.

At present it is assumed that the probability of helium leaks to the insulating vacuum is not much different from HERA, where there are presently three larger leaks which were there right from the beginning. It was only needed to place two additional pump stations nearby. The size of the leaks stayed stable and no further leak developed. The layout of the helium vessel has avoided the potential of leaks at the penetrations

of the feed throughs and contains only circular welds.

3.2.3.3 Input coupler vacuum system

The present layout of the input coupler has two ceramic windows, one which is at a temperature of 70 K during operation and one at room temperature. There are several reasons for having a second ceramic window in the input coupler: It enables one to close off the cavity completely by mounting the coupler up to this window in the clean room, thus preventing any contamination during the further assembly of the module. If a crack of the ceramic window occurs during operation there will be no danger of gas and dust entering the cavities as there is vacuum on both sides of the window. In case the outer window develops a crack the inner window will protect the cavities.

The disadvantage of this concept is that one needs an additional vacuum system for the room temperature side of the input coupler. The tempting possibility of connecting the outer part of the coupler to the insulating vacuum is not possible as the pressure may not be low enough due to leaks in the insulating system. Also conditioning of the couplers would become necessary after venting of the insulating vacuum system.

In the present concept pumping of the outer coupler part is achieved through slots or holes in the wave guide directly connected to the input coupler. All eight couplers are connected to a common vacuum header which is pumped by one ion getter pump.

The vacuum system of the input coupler needs further study to simplify it as much as possible. Special care will be taken to select proper materials and treatments to achieve a low thermal out-gassing rate as well as low gas desorption during RF-operation. It is planned to bake and precondition the input coupler on the modules prior to the installation in the tunnel.

3.2.4 RF Generation and Control

3.2.4.1 Introduction

In each TESLA cryomodule there are eight cavities. The peak RF power needed for one nine cell cavity at full gradient and maximum beam current, i.e. 25 MV/m and 8.22 mA during the pulse, is 213.5 kW.

Based on the experience gained from the TTF we plan to supply a set of 4 cryomodules with a 10 MW RF source.The nominal peak power needed for 4 cryomodules is only 6.83 MW, but as we will show in section 3.2.4.8, a power margin of 10-30 % is needed for phase and amplitude control. The maximum available pulse power of 10 MW provides, therefore, an additional power reserve of more than one MW.

The RF source foreseen for TESLA is a 10 MW Multibeam-klystron (MBK) which has 2 output windows, each delivering 5 MW.

Since, especially at the beginning of the RF pulse, there is almost fully reflected power, circulators will be used to limit the reflected power seen by the klystrons to 1.2:1 VSWR or better.

The beam pulse consists of 1130 micropulses with a spacing of 0.708 μs, and 530 μs are needed to fill the cavity with RF. Hence the RF pulse length is 1.33 ms. The

repetition rate is 5 Hz for the major part of the linac. At the low energy end of the linac there are some sections running at the repetition rates of 10 and 20 Hz.

In the following we will describe the Multibeam-klystron which is specially being developed for TESLA. Subsequently we describe the modulator which is presently in use in the TTF and which would also be appropriate for TESLA. Then, an alternative modulator concept, the SMES, is presented, followed by a possible scenario for the primary power handling and the distribution of pulsed electrical power in the tunnel. The last sections are devoted to the wave guide RF distribution system and to the low level RF system for phase and amplitude control.

3.2.4.2 The 10 MW Multibeam-klystron

In the spectrum of the commercially available L-band klystrons there are several tubes, which can deliver about 5 MW peak power for pulse lengths up to 2 ms. For 10 MW peak power typical pulse lengths lie rather in the 100 μs range. For the TTF we had decided to power the four cryomodules by two 5 MW klystrons TH 2104 C from Thomson. These are conventionally pulsed cathode klystrons without modulating anode.

Our motivation to look for another RF source for TESLA was threefold:

In view of the total AC power consumption for the RF system of TESLA, which is of the order of 50 MW, we felt that any possibility to increase the 45 % efficiency of the 5 MW tubes should be exploited to reduce operation cost.

Furthermore, doubling the RF power to 10 MW will reduce the total number of klystrons to 604. From this a reduction of capital investment cost is expected, since the experience of existing klystrons shows that the solution of many smaller tubes is more expensive than fewer high power tubes.

Finally, we are heading for an increase in tube lifetime from today's 15-20000 hrs to at least twice as much. It is believed that this goal can be achieved by the use of specially coated cathodes with reduced operation temperature.

The electronic efficiency of a klystron is highest for minimum beam current I and maximum beam voltage U because the space charge forces of the beam counteract the beam bunching. A characteristic quantity, which allows to estimate the maximum achieveable efficiency of a klystron, is the perveance defined as $P = I/U^{\frac{3}{2}}$. For $P = 2 \cdot 10^{-6}$ an efficiency larger than 45 % is possible whereas at $P = 0.5 \cdot 10^{-6}$ one may hope for more than 70 % efficiency. For a single beam 10 MW klystron this implies a gun voltage of 240 kV. The design of a gun at this voltage level and pulse length would be very critical. The main limitation for high peak power and pulse duration in klystrons is electrical breakdown which can occur along the insulators or between the electrodes of the gun and in the RF circuits of the tube. For very high peak power arcing can occur across the gap of the output cavity since the gap voltage has to be comparable to the gun voltage for high efficiency. From existing tube performances the empirical relation $E \cdot V = 100/\tau^{0.34}$ between achieved pulse length τ[s], max. field strength E [kV/mm] in the gun region and gun voltage V [kV] has been established [16].

Therefore the favoured solution is a Multibeam–klystron with 7 beams, each having the microperveance of 0.5, proposed by Thomson. The gun voltage will be only 110 kV

and the total beam current 130 A. There will be two RF output windows. From the experience with the existing 5 MW klystrons we know that this window is uncritical.The solenoid focusing power will be only 4 kW. The layout of this klystron is sketched in Fig. 3.2.17.

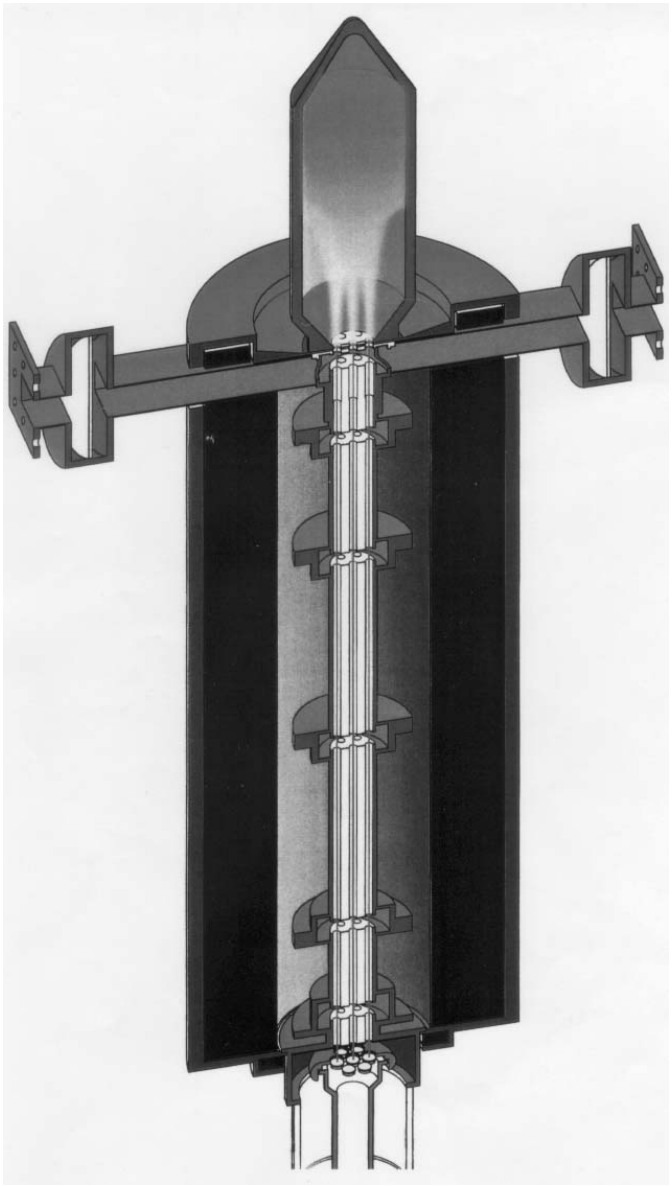


Figure 3.2.17: *Schematic layout of the 7 - beam klystron.*

A contract has been signed for the development of a 1.3 GHz 10 MW Multibeam-klystron with 1.7 ms pulse length and 10 Hz repetition rate. The specified efficiency is between 70 % and 75 %. Commissioning is forseen for fall 1997.

A prototype of a 64 kW 4-beam klystron operating at 425 MHz has demonstrated

an efficiency of 44 %. The cathode voltage was only 18 kV [17]. A proof of existence and feasibility of 5 MW L-band MBKs operating at ms pulse lengths comes from the Moscow Meson Factory, where 33 MBKs are running. These tubes have 6 beams and the operating voltage is only 75 kV. The microperveance of 1.4 per beam should allow for a higher efficiency than the 40 % of the actual tubes.

A fact which is often ignored is that reducing the RF output power by decreasing the drive power alone will also decrease the efficiency since reducing or switching off the RF drive power has no influence on the average dc electron beam current, and in the latter case the full energy of the beam is wasted in the collector. The efficiency of the device becomes zero.

Large power variations are performed more efficiently either by variation of the operating voltage or by means of a modulating anode or a gridded gun. For high power tubes which operate at cathode voltages well above 100 kV there are practically no tubes with a modulating anode because of isolation problems, and the technology of gridded guns for high power tubes is not yet very mature. Therefore our Multibeam-klystron will be a diode without modulating anode.

The design allows to decrease the saturation power limit by some 26 % by reduction of the gun voltage without loosing more than 5 % in efficiency. This is very important for the case where the maximum power needed, i.e. nominal power + regulation reserve, is below the maximum possible output power of the klystron.

The klystron 1 dB bandwidth is 3 MHz. This is more than adequate for fast response of phase and amplitude regulation loops.

Frequency:	1300 MHz	
RF Pulse length:	1500 μs	
Repetition rate:	10 Hz	
Beam Voltage:	110 kV	
Beam Current (total):	130 A	
Perveance per beam:	$0.5 \cdot 10^{-6}$	
Drive Power:	200 W max.	
Output Power:	10 MW peak	150 kW average
Body Dissipation:	15 kW max.	
Efficiency:	> 70 %	
Gain:	48 dB	
Bandwidth (1 dB):	3 MHz min.	
No. of beams:	7	

Table 3.2.10: *Design parameters of the Thomson multibeam-klystron*

The gain of 48 dB means that the drive power is below 160 W. This power can be generated at low cost with solid state amplifiers.

3.2.4.3 Other Possible RF Sources for TESLA

An interesting possible alternative to the MBK could be the Magnicon, a source developed at Budker Institute (Novosibirsk). It consists of a continuous electron beam which is deflected by the RF magnetic field of a circular deflection cavity [18].

In the subsequent drift space electrons deviate from the device axis and get into a stationary magnetic field of a solenoid. While entering the magnetic field the longitudinal velocity of the electrons is transformed into a rotational transverse one.

Further on, traveling along a helical trajectory in the output cavity, the electrons excite a TM_{110} oscillation mode. If the cyclotron frequency is equal to the one of operation, then the interaction can remain effective during many periods of RF oscillation resulting in a very high efficiency. Feasibility of the magnicon has been demonstrated by two prototypes. One has delivered 2.6 MW RF pulses of 30 μs length at 915 MHz. Here the electronic design efficiency was 85 %. The efficiency actually reached was 73 % due to the quality of the aluminum alloy used for cavity fabrication. The other magnicon was a frequency doubler which has demonstrated 25 MW output power at 7 GHz for 2 μs pulses. In this case the design efficiency of 60 % was reached.

It has been claimed that a 10 MW Magnicon with an efficiency close to 78 % could be built [19]. The disadvantage of the Magnicon is its high operating voltage. For the TESLA Magnicon it would be 250 kV.

There are also attractive features for TESLA from Klystrodes which are also called **I**nductive **O**utput **T**ubes (IOT). Here the beam is modulated with the RF frequency by a thin grid close to the cathode such that no drift space and bunching cavities as in a klystron are necessary. Without RF drive power the beam current is zero. The extra amount of RF power needed for phase and amplitude control in the superconducting cavities does not reduce the effective efficiency as in the case of a normal klystron. A disadvantage of this tube is its low gain which is well below 30 dB. The feasibility studies of such a tube for TESLA are, however, only in a very early stage.

3.2.4.4 Klystron Modulator Schemes

Already for the TTF we were confronted with the task of designing a cathode modulator for the 5 MW diode klystron which produces 130 kV pulses at a current of almost 100 A.

The solution of a switch tube in series with the klystron is not attractive. To our knowledge, at most one tube exists, which could switch currents of about 100 A for intervals like 2 ms. We were concerned whether this solution, if feasible at all, would be very expensive and unreliable.

The more exotic solution of a **S**uperconducting **M**agnetic **E**nergy **S**torage (SMES) seemed promising, but was too far from maturity to be used for the TTF in the first place. It will be described in the next section.

A proven solution consists of a **P**ulse **F**orming **N**etwork (PFN) in conjunction with a pulse transformer. This technique has been successfully used at SLAC, FNAL and at other places, but only with pulse lengths up to some hundred μs.

Evaluation of the costs of a 1.3 ms PFN stimulated search for other solutions. A very elegant solution, which has finally been adopted was proposed by Quentin Kerns

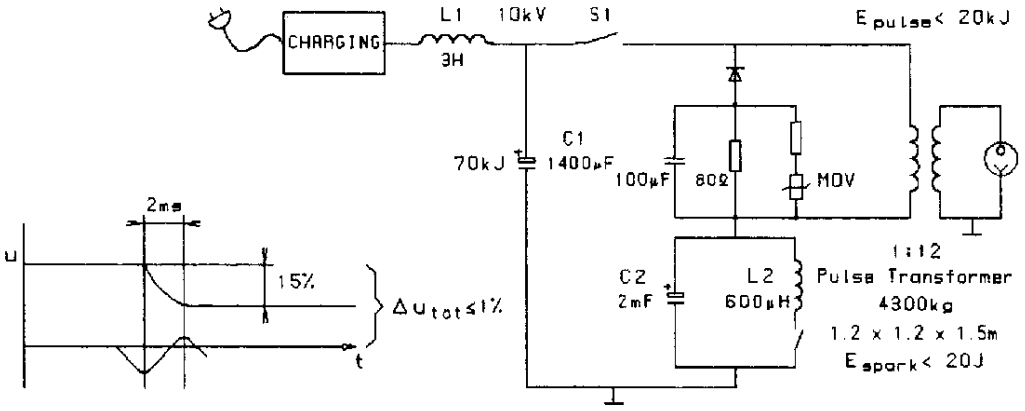


Figure 3.2.18: *Diagram of the TTF modulator with bouncer circuit. The bouncer circuit reduces the fluctuations of the pulse flat top to the* $\pm 1\,\%$ *level.*

from FNAL and is sketched in Fig. 3.2.18.

In operation, the DC power supply keeps capacitor C1 charged to the 9 kV level. The output pulse is started by turning on the **I**solated **G**ate **B**ipolar **T**ransistor (IGBT) switch S1, which connects C1 to the pulse transformer primary. The pulse is terminated after 1.5 ms (1.3 ms flat top + 0.2 ms rise time) by turning the IGBT off. The nominal current switched by the IGBT's is 1.56 kA. The primary pulse is stepped up to the klystron operating level by the 12 : 1 pulse transformer.

During the pulse, capacitor C1 discharges by 13 % of its initial voltage, putting an intolerable slope on the output pulse. To decrease the slope to the 1 % level without resorting to an 18 mF capacitor in the C1 location, the slope is corrected with a bouncer circuit. This is a resonant LC circuit which creates a single sine wave with a period of 7 ms. The bouncer is triggered slightly before the main pulse so that the linear, bipolar portion of the cycle is playing during the main pulse. The bouncer wave form cancels out the 13 % slope from C1, and reduces it to less than $\pm$ 1 %.

The need for a 1.3 ms pulse transformer, which is a fairly unique device, might be considered a disadvantage of this solution. For the TTF three such modulators were, however, built by FNAL and several suppliers for appropriate pulse transformers have been found. Here the pulse length was even 2.2 ms. The modulators have been running very reliably for several years for the TTF operation. A very detailed description of this modulator is given in [20] and [26].

The pulse rise time of these modulators is of the order of 150 μs, and the voltage droop during the pulse is $\pm$ 1 %. The measured overall wall plug power to useful HV conversion efficiency was 86 % for a pulse length of 2.2 ms. "Useful" here means that only the flat top of the HV pulse, which is good for RF generation, has been taken into

Electronic klystron efficiency	70 %
Electronic modulator efficiency	92 %
Rise time efficiency $\frac{T_p}{T_p+T_R}$	86.6 %
HV cable losses	2 %
Wave guide losses	4 %
Overall efficiency	52.5 %
Duty cycle	0.65 %
Total peak RF power	$4.125 \cdot 10^9$ W
Average RF power	26.8 MW
Wall plug power	51 MW
10 % power reserve for regulation	5.1 MW
Solenoid power	2.4 MW
Total wall plug power	58.5 MW
Nominal RF power to total wall plug power	
Conversion efficiency	45.8 %

Table 3.2.11: *Efficiencies*

account.

The average rise time of the HV pulse will, however, be closer to 200 μs because of the long cables between the pulse forming units in the service buildings and the pulse transformer-klystron units in the tunnel (see below). With the 1.3 ms length of the RF pulse an overall modulator efficiency close to 78 % is expected. The specified electronic efficiency of the klystron is above 70 %.

Assuming 10 % power reserve for amplitude and phase regulation and 4 % power loss in the circulators and wave guides, and taking the klystron focusing solenoid, which needs a continuous power of 4 kW into account, the wall-plug power to RF power conversion efficiency of the TESLA Linac will be about 46 % for a duty cycle of 0.65 %. All this is summarized in Table 3.2.11.

One may hope to reduce the modulator costs by replacing the energy storage capacitor by a SMES. In order to demonstrate the feasibility of such a system and to get a reliable cost estimation we have decided to build a prototype SMES modulator for the TTF which will be briefly described in the following.

3.2.4.5 Design of a 25 MW SMES Modulator for the Generation of Long Power Pulses

A SMES based power modulator for providing 25 MW pulsed power at TESLA conditions is described in this chapter. The maximum electrical output pulse power is 25 MW. It can be used to supply either two 5 MW klystrons or one 10 MW klystron. The modulator is thought to be used for the free electron laser at TTF and serves as a demonstrator of a larger modulator which should then supply about 10 klystrons in the collider. The requirements for the 25 MW modulator are listed in Table 3.2.12.

Output Voltage	130 kV
Output Current	190 A
Pulse width	1.7 ms
Flat Top	± 1 %
Repetition Rate	10 Hz

Table 3.2.12: *Output characteristic of a 25 MW SMES modulator for TESLA*

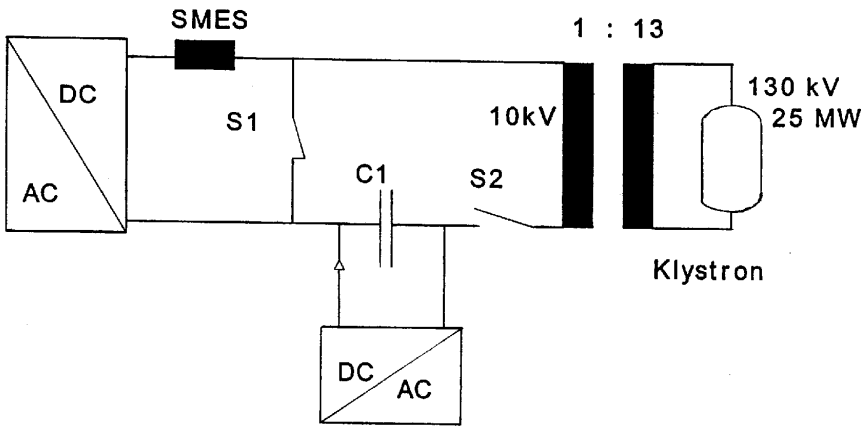


Figure 3.2.19: *Basic SMES modulator circuit.*

Principles of Operation

The basic modulator is shown in Figure 3.2.19. First it was described in [21]. The operation of this system can be devided into a charging mode and the pulse generation mode.

During the conduction time of the switch S1, which can be realised as a thyristor-switch proposed in [22] and the opening time of S2, which is made up of IGBT's, the SMES and the capacitor bank C1 are being charged. The SMES current and the capacitor voltage increase to their nominal values of 2500 A and 12 kV, respectively. For generation of an output pulse the switch S2 is closed. The thyristor S1 will be opened by the current of C1. In this pulse generation mode the SMES, the capacitor bank, and the klystron impedance (transformed to the primary side of the step up transformer) are forming a serial ringing circuit with the initial conditions I_{SMES} and $U_{capacitor}$.

The current in the circuit and in the klystron during the pulse generation time, referred to the primary winding of the step up transformer are described by the following equations, if we neglect the power supplies:

$$i_{klystron}(t) = I_{SMES}(0)(e^{-at}cos(bt) - \frac{a}{b}e^{-at}sin(bt)) - \frac{U_{Capacitor}(0)}{L_{SMES}}\frac{1}{b}e^{-at}sin(bt)$$

where

$$a = \frac{R_{klystron}}{2L_{SMES}}$$

$$b^2 = \frac{1}{CL_{SMES}} - a^2$$

After the 1.7 ms pulse generation time the thyristor will be closed and S2 will be opened so the charging mode begins again. During the generation time the SMES is first charged by the capacitor bank because the initial voltage of the capacitor is higher than the voltage across the primary side of the pulse transformer. Due to the TESLA requirement on the pulse shape and the series connection between the klystron (transformed to the primary side of the transformer) and the SMES, the current variation in the SMES is lower than 2 %. The versatile design of this modulator allows for a variable use such as the generation of pulses at a lower power (for supplying just one klystron, e.g.) and a varying repetition rate as well.

Simulation and Calculation of the Circuit

The calculated parameters of a 25 MW SMES modulator for TESLA conditions are listed in Table 3.2.13. The corresponding electrical circuit diagram is shown in Fig. 3.2.19.

SMES	L_{SMES}	= 70 mH	$I_{SMES}(0)$	= 2500 A	E_{SMES}	= 220 kJ
Capacitor	C	= .6 mF	$U_C(0)$	= 12 kV	E_C	= 43 kJ

Table 3.2.13: *Calculated modulator components*

The circuit has been investigated by simulations. In Figure 3.2.20 the output voltage across the klystron is shown. The amplitude of a single pulse varies within a tolerance band from +0.2 % to -1.4 % during a 1.7 ms pulse.

The change of 1.6 % of the magnitude of the pulse voltage also applies to the current in the SMES. The SMES current increases during the charging mode of the circuit from 2470 A to the initial value of 2500 A.

The maximum change of the SMES current during the pulse is 1.6 % corresponding to a decrease of the SMES energy of about 10 kJ. This energy drop of 10 kJ has to be reloaded during the 100 ms of the charging mode. Therefore a SMES power supply with an average power of 100 kW is necessary. The SMES power supply used in the simulations was a 30 V/2500 A voltage source.

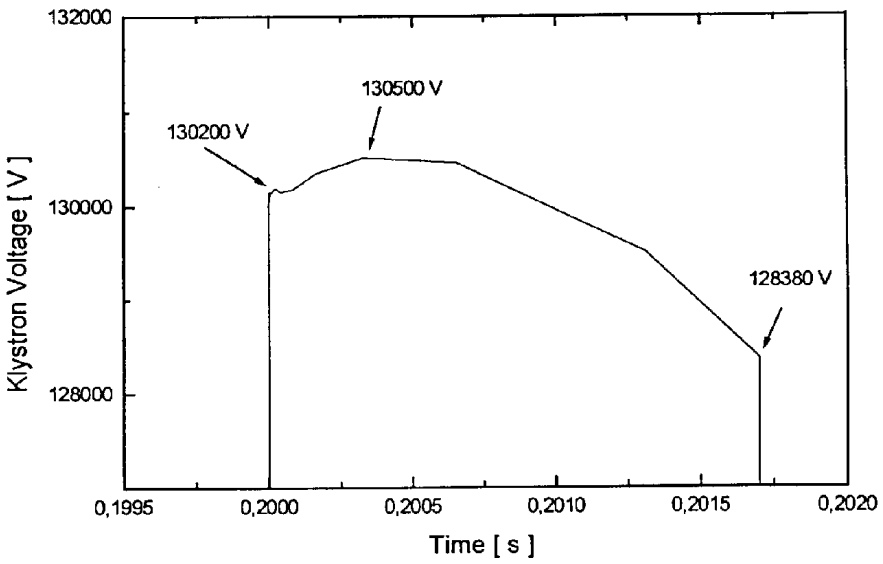


Figure 3.2.20: *SMES modulator output voltage. Single pulse across the klystron.*

There is a huge variation of the capacitor voltage during a cycle. The initial capacitor voltage of 12 kV drops during a pulse to about 5 kV. The voltage blocking capability of switch S2 has to be at least the maximum voltage across the capacitor bank. The capacitor energy decreases to 25 % of its value at the beginning of the pulse generation. The discharge is about 40 kJ. It is four times higher than the discharge of the SMES. This energy has to be recharged during the 100 ms cycle. Therefore the supply for the capacitor bank requires an average power of 400 kW. The average current of this supply was calculated to be 44 A. The properties of the required power supplies for the modulator are listed in Table 3.2.14.

SMES Supply	I_{av} = 2500 A	V_{load} = 30 V	P = 75 kW
Capacitor Supply	I_{av} = 44 A	V_{max} = 12 kV	P = 530 kW

Table 3.2.14: *Power supplies*

Summary and Conclusions

The calculations and simulations show that a SMES-based 25 MW modulator can be built with components that have been made available recently. A fast 250 kJ SMES was built in [23]. A bigger capacitor bank than required here was realised and described in [20].

There are several enterprises that can build the required power supplies. Dominant properties of the semiconductor switches needed for this modulator were proven in [23] and [20]. Problems concerning a reliable klystron protection have been solved in [20]

and will be adapted to this system.

This modulator concept may be applied also for higher output power and simultaneous supply of a larger number of klystrons.

3.2.4.6 Modulator Primary Power Handling and Pulsed Power Distribution

It is impressive to note that the peak RF power needed during each pulse is 4.125 GW for the whole linac. For a duty cycle of 0.65 % and with the efficiencies mentioned in 3.2.4.4, including 10 % regulation reserve, this results in an average wall plug power close to 59 MW.

Clearly the primary power handling system must be such that the mains see only the average power consumption. A possible scenario is shown in Fig. 3.2.21.

Each klystron together with its own pulse transformer forms a unit which is installed in the tunnel and which is exchanged as a whole in case of failure. The pulse forming circuits, i.e. capacitor banks with a bouncer or SMES units and their charging power supplies are located in the cryo-buildings. The distance between the cryo-buildings is 5 km, hence modulators for 2×50 klystrons requiring 2×12 MW total average mains power (at 10 Hz rep. rate) must be installed in each building.

Resonant charging of each of the 50 capacitor banks and their bouncer circuits will be accomplished by multiplexing a common power supply. This can be done by firing thyristors at the anticipated begin of the (re)charging process. The zero crossing of the charging current blocks the thyristors and terminates the charging process.

Each pulsing unit has an additional fast power supply with some 3 % charging capacity for individual matching. There is also a simple supply for initial charging of units which had tripped.

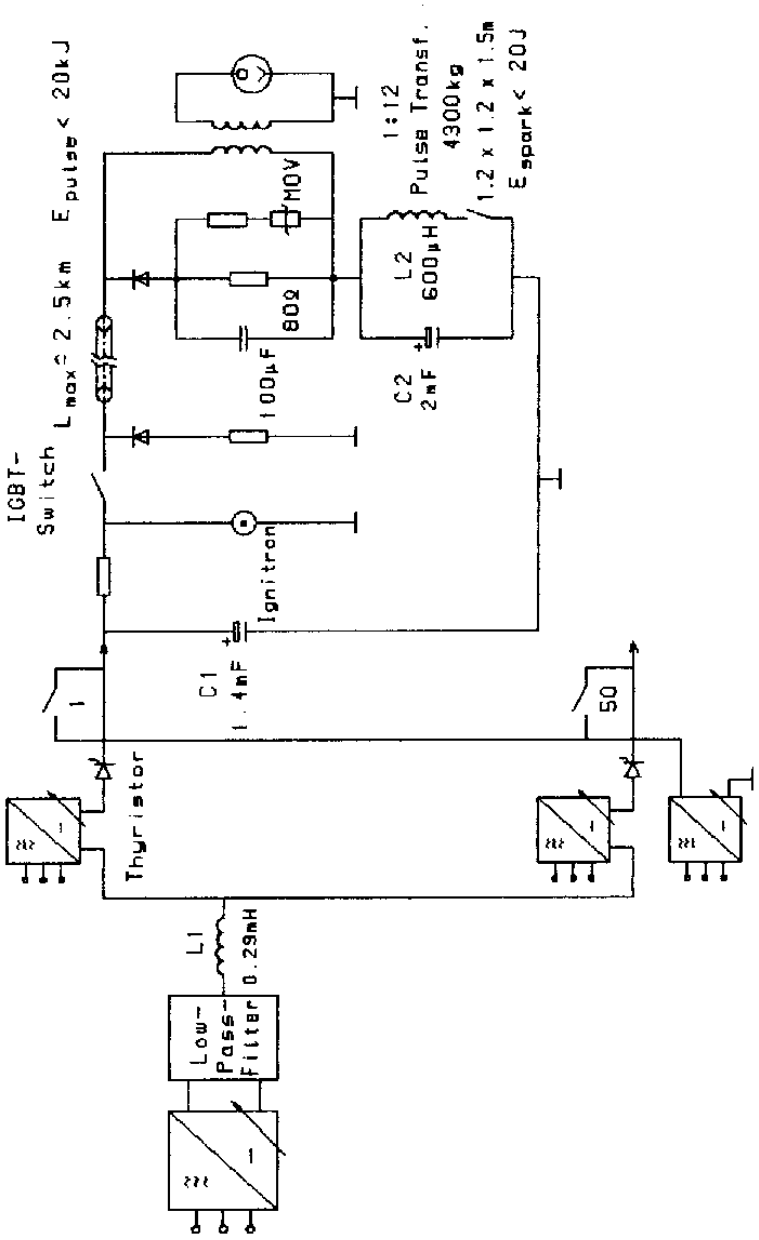


Figure 3.2.21: *Possible scenario for resonant sequential charging of 50 pulser units and power distribution over the average length of 1.25 km into the tunnel. In the tunnel only the klystrons and pulse transformers are located.*

For the stations running at 10 Hz repetition rate 50 current half waves are needed within 100 ms to subsequently charge each of the the 50 units in the time interval between two pulses. Therefore the charging frequency 250 Hz has been chosen. It is realized by the common 0.29 mH choke in conjunction with the 1.4 mF capacitor of each unit.

For the majority of the stations running at 5 Hz repetition rate the charging frequency can be halved by using a 1.16 mH choke.

The alternative solution of charging all the 50 units simultaneously by dimensioning the choke such that the variation of the charging current, which is shaped like the top of a cosine curve, does not exceed ± 1 % seems less attractive. In this case a 0.89 H choke is required at 10 Hz operation and 3.56 H at 5 Hz repetition rate. There would be no multiplexing, but thyristors would still be needed to disconnect units which have failed. These thyristors, however, never see a zero crossing of the charging current in this scenario, and would, therefore, require additional networks for controlled interruption. In addition, the cables connecting the pulser units with the pulse transformers in the tunnel can be disconnected rather than having extra manual switches.

The fluctuations seen by the mains are reduced to the 2 % level by a low pass filter. The capacitance of this filter is 0.2 mF and its inductance is 0.1 H. It also limits the current rise time in the case of a short circuit in one of the 1.4 mF capacitor banks to 100 A/ms. This gives sufficient safety margin for the primary thyristors to open and disconnect the mains.

The thyristor of the damaged capacitor bank will stay open, and a reserve RF system must be started. Then operation can restart immediately. If the short circuit occurs behind the IGBT switches, for example in a klystron gun, it can be handled by opening the corresponding IGBT switch within a few μs.

The power transmission lines between the pulser units and the pulse transformers deserve special attention. Their average length in the tunnel is 1.25 km and the maximum length is 2.5 km. Due to the inductance of these cables, which is of the order of 300 μH/km and hence comparable to the leakage inductance of the pulse transformer, the rise time of the pulse is significantly increased. Simulations [24] have shown that the optimum rise time and pulse shape is obtained for the cable impedance close to 10 Ω. This is close to the transformed klystron impedance of 8.3 Ω. Then the pulse rise time is increased from 150 μs for the case of a pulse transformer directly connected to the bouncer unit to 250 μs when a cable of 2.5 km length is present. These results refer to the case of a bouncer pulse forming unit. In the SMES case the optimum cable impedance is closer to 40 Ω [25].

The cable cross section must be at least 400 mm^2 to keep the average power loss on the cables below 2 %. Frequency components up to 6 kHz should be transmitted without significant loss, otherwise there will be furher increase of the pulse rise time. Since the penetration depth at 6 kHz is only 0.82 mm it seems appropriate to use 3 coaxial cables of 30 Ω impedance and about 55 mm diameter in parallel. The inner conductor must be stranded wire.

Klystron Interlocks

In the event of a klystron gun spark, which is detected in three different ways for redundancy, the energy deposited in the spark must be kept below 20 J to avoid damage of the gun.

The response to a spark will be an immediate opening of the concerned IGBT switch (see Fig. 3.2.21) to disconnect the capacitor bank from the sparking klystron. The energy stored in the transformer leakage inductance and in the power transmission cable is dissipated in two networks, one at the cable end near the IGBT consisting essentially of a reverse diode and a resistor. The second one is made up by an 80 Ω resistor across the transformer primary and by a 100 μF capacitor which limits the peak inverse voltage at the primary to 800 V when the IGBT is opened.

Should it fail to open the ignitron crowbar is fired and, in addition, a second independent switch installed in series with the IGBT is opened. Important interlocks are control of cooling water flow and temperature, the focusing solenoid current, and a vacuum interlock. Other interlock conditions result from sparks in the RF distribution system, reflected power, RF leaks, power couplers and from cryogenics.

3.2.4.7 The RF Distribution System

The 10 MW Multibeam-klystron has two RF output windows. Therefore the RF distribution system is based on two cryounits per 5 MW output window. Each unit contains eight 9-cell cavities. In principle, a tree-like RF distribution can be used, but a linear system branching off identical amounts of power for each cavity from a single line by means of directional couplers matches the linear tunnel geometry better. Such a system is already in use for the HERA superconducting RF system and has also been successfully tested in the TTF. All required components are readily available. This geometry permits to install additional power splitters to reduce the RF power sent into an individual arm if this particular cavity or its coupler can not handle the full nominal power.

One might object that unlike the case of the tree-like system, where all paths have the same electrical length, thermal expansion could be a problem for the linear distribution system. Fortunately the effect is only 0.24° of RF phase per centigrade and per cryomodule and is therefore uncritical.

The maximum power of 5 MW is symmetrically devided by a three dB power splitter and feeds the two cryomodules, see Fig. 3.2.22 and 3.2.23. The individual three stub wave guide tuners for impedance- and phase matching of each cavity provide the possibility of ±30° phase adjustment and to correct for VSWR up to 2:1.

Circulators are indispensable. They have to protect the klystrons against reflected power and, in conjunction with load resistors and the power input coupler, they define the loaded cavity impedance as seen by the beam. One possibility consists of placing the circulator between the klystron and the 3 dB splitter. Since, due to cavity filling, there is almost total reflection of the RF power at the beginning of the pulse, the standing wave RF voltage which occurs in the circulator and in some parts of the wave guide system may double. In other words, with respect to sparking protection, the equivalent

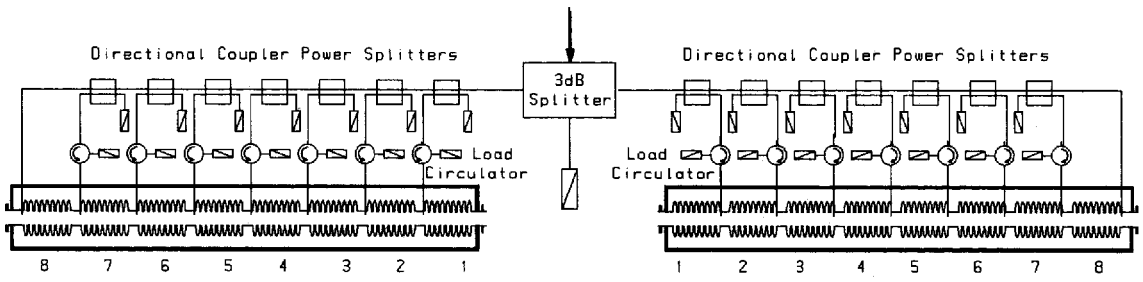


Figure 3.2.22: *Layout of the RF distribution scheme. One 5 MW arm from the 10 MW Multibeam-klystron feeds two cryounits. There are eight 9-cell cavities in each unit. The coupler hybrids branch off equal amounts of power into the individual cavities. There is an individual circulator for each cavity.*

power to be considered is 4 times the real power entering into the circulator, i.e. almost 20 MW. Assuming a breakdown field of 30 kV/cm, the recommended power rating for WR 650 wave guide is 14.8 MW for straight wave guide under ideal conditions, which means dry and clean.

Of course, the wave guide lengths can be arranged so that the reflected powers from each cryomodule will add up constructively in the load associated with the 3 dB power splitter rather than propagating to the circulator. This can work perfectly only if both power signals arrive with equal amplitude and proper phase at the 3 dB splitter. Clearly it is unrealistic to hope that this will always be the case.

Therefore the wave guide section between the klystron and the 3 dB splitter must be filled with a protective gas. In the TTF SF_6 of normal pressure, which can increase the power handling capability by up to 8 times, was used. The disadvantage of SF_6 is that, if sparking occurs, aggressive radicals are formed, and one has to change the gas in order to avoid wave guide corrosion. For TESLA SF_6 seems impractical because small amounts of radiation may also result in decomposition of the gas. Other gases would require an overpressure of several bars to get the same relative power capability as SF_6 at normal pressure.

Other problems stem from the fact that the development of a low insertion loss (0.1 dB) circulator which is capable of handling the 5 MW peak power under total reflection conditions is problematic. So far no circulator which works reliably at 2.5 MW at total reflection has been realised. In addition, a wave guide RF window is needed to separate the SF_6 from the normal air in the rest of the system. Here we have succeeded in building three pill-box windows capable of handling total reflection at any phase angle of RF pulses of some 3 MW peak power and 2 ms length. It does not seem possible to go up to 5 MW with this design. Other designs might become much more complex.

Therefore the alternative solution of using individual circulators for each cavity,

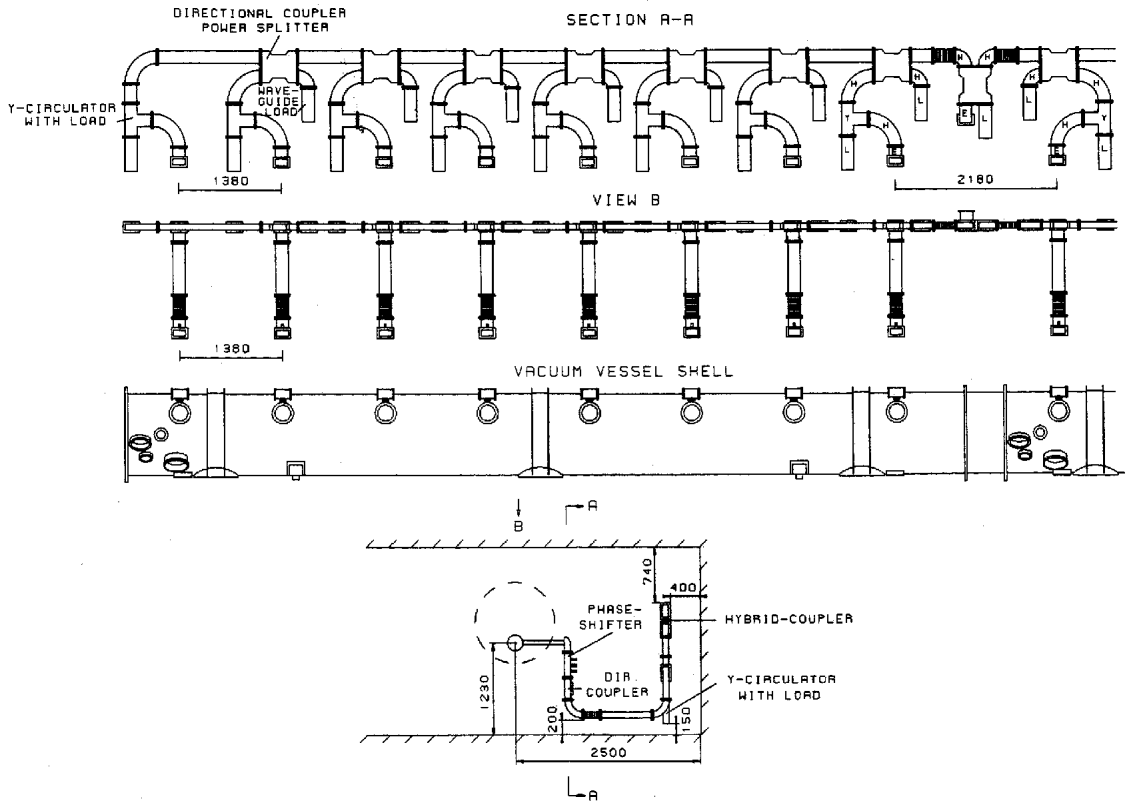


Figure 3.2.23: *Detailed layout of RF distribution scheme.*

which has been successfully tested in the TTF seems increasingly attractive. All the above mentioned technical problems disappear in this case. No protective gas and hence no window would be necessary in any part of the wave guide. These small circulators exist and have excellent characteristics. The insertion loss is 0.1 dB, the VSWR at total reflection better than 1.2 depending on phase angle. It has been measured that the total VSWR seen by the klystron windows can be kept below 1.2 with the RF distribution scheme shown in Fig. 3.2.22. With the help of an additional three post wave guide impedance matcher identical to the ones close to each cavity power input coupler this VSWR can still be minimised.

A further improvement due to the individual circulator solution is the increased decoupling of the cavities. Since there are RF pick-up antennas in the forward- and reflected power ports of the circulators, no extra directional couplers are needed in the wave guide arms which results in a cost reduction. Another cost reduction comes from the fact that the circulator loads need to be dimensioned only up to the average power per cavity and the loads connected to the directional coupler hybrids see only less than 100 W of average power and can be very simple. If there were no individual circulators some of these loads could see almost 4 times the average power sent to an individual cavity.

Our experience from the TTF with the investment cost for the small circulators

suggests that the overall cost of this solution will not be more expensive but much more reliable than the one of high peak power circulators.

Frequency:	1300 MHz $\pm$ 10 MHz
Power:	5 MW peak during 2 ms, 20 kW average
Isolation:	30 dB typ., 26 dB min.
VSWR:	1.1:1 max. for all ports
Flange:	CPR 650 F
Waveguide:	WR 650, normal wall thickness

All aluminum components shall be fabricated using 6061-T6 alloy.
Finish MIL C5541 class 1A.
We need seven different types of power splitters with coupling ratio:

1	.	9.03 dB $\pm$ .1 dB
2	.	8.45 dB $\pm$ .1 dB
3	.	7.78 dB $\pm$ .1 dB
4	.	6.98 dB $\pm$ .1 dB
5	.	6.02 dB $\pm$ .1 dB
6	.	4.77 dB $\pm$ .1 dB
7	.	3.00 dB $\pm$ .1 dB

Table 3.2.15: *Specification of WR 650 directional coupler power splitters*

Power:	.5 MW peak during 2 ms, 2 kW average.
VSWR:	1.05 : 1 by full power
RF leakage:	-95 dB
Physical length:	$<$ 500 mm
Flange:	CPR 650 F

Table 3.2.16: *Specification of WR 650 wave guide load for circulators.*

Frequency:	1300 MHz $\pm$ 5 MHz
Peak power:	0.5 MW. Due to total reflection of the RF signal during an important fraction of the pulse duration the maximum standing wave voltage will correspond to 2 MW
Average power:	2 kW
Insertion loss:	.1 dB
VSWR:	better than 1.2:1 for total reflection.
Isolation:	30 dB, min. 26 dB
RF pulse length:	2 ms
Repetition rate:	10 Hz

Table 3.2.17: *Specification of circulator*

3.2.4.8 Low Level RF Control

The cavities in the TESLA linacs are operated in pulsed mode at an average gradient of 25 MV/m with each klystron driving 32 cavities. The pulse repetition rate is 5 Hz and the RF pulse length is 1300 μs from which 500 μs are required to fill the cavity and 800 μs are needed for the acceleration of the beam. During the beam pulse the fluctuations of the accelerating field - defined as the vector-sum of the fields in 32 cavities - must be kept small to reduce the uncorrelated contributions to energy spread to less than $5 \cdot 10^{-3}$ for an ensemble of 32 cavities.

The major sources of field perturbations which have to be controlled by the low level RF system are caused by fluctuations of the resonance frequency of the cavity and by fluctuations of the beam current. Fluctuations of the resonance frequency are a result of deformations of the cavity walls induced by mechanical vibrations (microphonics) or the gradient dependent Lorentz force (also referred to as radiation pressure or ponderomotive force).

Mechanical vibrations caused by roughing pumps or the compressors of the cryogenic system are always present in an accelerator environment. These vibrations induce a modulation of the resonance frequency. The spectrum of the resulting phase and amplitude errors is a convolution of the spectrum of excitation and the mechanical resonances of the cavity. The microphonic noise is not predictable from pulse to pulse and is expected to be uncorrelated over the length of the linac.

The cavity detuning due to the Lorentz force acting on the walls of the cavity is proportional to the square of the cavity field. The TESLA cavities are stiffened to reduce the Lorentz force detuning constant to -1 Hz/(MV/m)^2. The error due to Lorentz force is predictable due to the repetitive excitation of the cavity and considered as correlated noise.

The bunched structure of the beam current induces a sawtooth like profile on the accelerating gradient. At the gradient of 25 MV/m a single bunch with a charge of $3.6 \cdot 10^{10}$ e$^-$ will induce a transient with a magnitude of 10^{-3}. Bunch-to-bunch charge fluctuations will modulate the amplitude of the transients thereby inducing bunch-to-bunch energy fluctuations.

A digital feedback system will provide control of the accelerating field. The accelerating field of each of the 32 cavities is measured, the vector-sum is calculated, and the amplified error signal drives the vector-modulator of the incident wave which is common to all cavities since they are driven by the same klystron. The digital approach allows for maximum flexibility in the choice of the control algorithms, automated calibration of the vector-sum, and extensive diagnostics and exception handling capabilities. The main features of the digital feedback are a sampling rate of 1 MHz for the individual cavity probe signals, digital in-phase and quadrature detection of cavity field and amplitude of incident wave and reflected wave, calculation of the vector-sum which includes gradient calibration and the correction of phase offsets for each cavity, and the feedback algorithm. The feedback algorithm includes time-optimal control and a Kalman filter to correct for loop delays on the order of several μs and provides an optimal state estimate in presence of detector noise. Adaptive feed forward is added

to minimize the control effort by compensation of repetitive errors.

RF Control Requirements Usually the requirements for an RF control system are stated in terms of achieved amplitude and phase stability. In addition, the extra power needed for control as a result of the Lorentz force and microphonics must be minimized. It is, however, also important to address issues such as reliability, operability, robustness and maintainability in the design. The low level controls must support procedures for the turn on of the RF system (commissioning and routine operation), calibration of gradient and phase for the vector-sum, and control of the frequency tuner.

Amplitude and Phase Stability The requirements for amplitude and phase stability of the vector-sum of 32 cavities are dictated by the maximum tolerable energy spread at the interaction point. The goal is an rms energy spread of $\sigma_E/E = 2 \cdot 10^{-3}$. Gradient and phase fluctuations will result in bunch-to-bunch energy spread while finite bunch length combined with the sinusoidal time dependency of the RF wave and induced bunch wake potential will result in intra-bunch energy spread. The intra-bunch energy spread can be minimized by proper phasing of the cavities i.e. the beam is accelerated a few degrees off-crest[27]. It is desirable to keep the bunch-to-bunch energy spread below the intra-bunch energy spread of $5 \cdot 10^{-4}$ in order to assure that the bunch-to-bunch chromatic effects will not be a dominating factor. Considering that correlated field fluctuations can be suppressed to better than $3 \cdot 10^{-4}$ the requirements for uncorrelated gradient and phase have been set to $\sqrt{N} \cdot 3 \cdot 10^{-4} = 5 \cdot 10^{-3}$ and 0.5^o respectively. This is due to the fact that the effect of uncorrelated errors on the final energy spread is reduced by $\sqrt{N}$, where $N = 302$ is the number of klystrons in each linac, and that a phase control to 0.5^o results in an energy spread contribution of the order of $3 \cdot 10^{-4}$ when operating 4^o off crest.

Operational Requirements The turn on procedure for the RF system is fairly complex due to the large number of cavities driven by the same klystron. It is therefore important to have fully or semi-automate procedures for the calibration of the vector-sum, phasing of cavities, and operation of the frequency tuner. Also gradient and phase changes should be controlled by a single knob as well as the RF system be turned on and off with a single button. In case of operational problems the control system must provide the operator guidance for trouble shooting and allow RF experts for completely manual operation.

RF Control Issues The amplitude and phase errors to be controlled are of the magnitude of 5 % and 20 degrees respectively as a result of Lorentz force detuning and microphonics. These errors must be suppressed by a factor of at least 10 which implies that the loop gain must be adequate to meet this goal. Fortunately, the dominant source of errors is repetitive (Lorentz force) and can be reduced significantly by use of feed forward.

Lorentz Force The Lorentz force yields a steady state change of the resonance frequency of $-625\,\mathrm{Hz}$ when increasing the gradient from zero to $25\,\mathrm{MV/m}$. Due to the dynamics of the cavity which has a mechanical time constant of about 1 ms which is comparable to the RF pulse length, the actual frequency change during a 1.3 ms RF pulse is only 400 Hz. The dynamics of the mechanical system are described by:

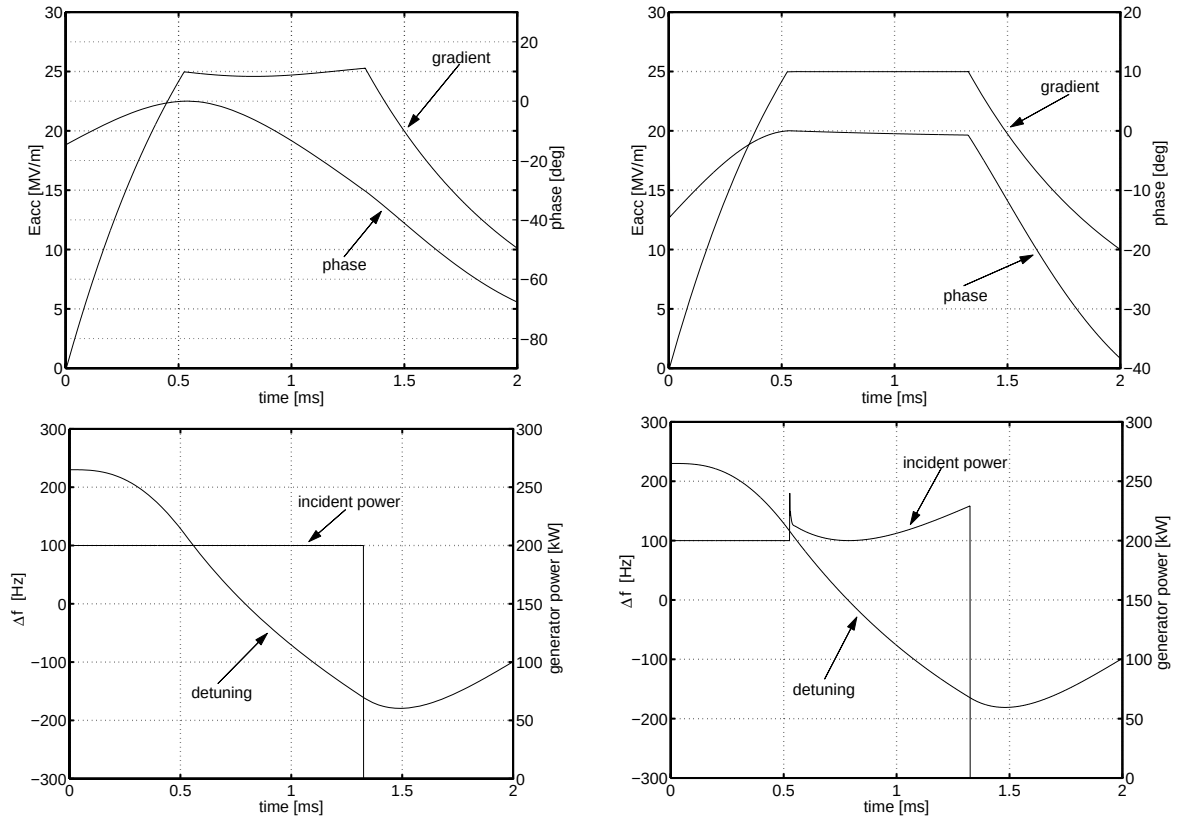
$$\dot{\Delta\omega} = -\frac{1}{\tau_m} \cdot \Delta\omega + \frac{2\pi}{\tau_m} \cdot K \cdot E_{acc}^2$$

with the Lorentz force detuning constant $K = -1\,(\mathrm{Hz}/(\mathrm{MV/m})^2)$. If the tuning at the beginning of the RF pulse is choosen such that resonance frequency and operating frequency agree in the middle of the beam pulse ($\Delta\omega = 0$), the peak detuning will not exceed the bandwidth of the cavity of 215 Hz (HWHM). This choice of optimum pre-detuning of approximately $-230\,\mathrm{Hz}$ also minimizes the RF peak power needed for control. The time dependency of the resonance frequency is shown in Fig. 3.2.24.

Microphonics The resulting phase errors can be as large as several ten degrees due the to the high loaded Q of the cavities of $Q_L = 3 \cdot 10^6$ corresponding to a bandwidth (HWHM) of 215 Hz. The sensitivity for the TESLA cavities is about 0.5 Hz per nanometer change in cavity length. Measurements in the TTF have shown typical microphonic noise levels of $\pm 30\,\mathrm{Hz}$ with a dominant modulation frequency of 27 Hz. The microphonics modulate the pre-detuning of the cavity and therefore the RF peak power requirement during the RF/beam pulse. A study of RF peak power requirements versus the amplitude of microphonics has shown that an amplitude of 100 Hz is acceptable.

Beam Loading The beam induced cavity voltage is $25\,\mathrm{MV/m}$ for the nominal average beam current of 8.3 mA while the generator induced voltage on resonance is $50\,\mathrm{MV/m}$ for an incident power of $213\,\mathrm{kW/m}$. These conditions result in the nominal steady state accelerating voltage of $25\,\mathrm{MV/m}$. Bunch-to-bunch charge fluctuations of 10 % rms will result in rms energy gain fluctuation of $1.4 \cdot 10^{-4}$ which can not be controlled by feedback due to gain bandwidth and power limitations. A gain bandwidth product of the feedback loop of 20 kHz allows, however, for a gain of greater than 10 at 1 kHz hence errors at this frequency can be suppressed by an order of magnitude.

Low Level RF Design The main components of the low level RF are the detectors for the cavity field, the controller for feedback, and actuators which control the incident wave to the cavity. For feedback one can choose the traditional amplitude and phase control system, one can apply feedback to in-phase (I) and quadrature (Q) component of the cavity field, or use direct RF feedback where the cavity probe signal is compared to an RF reference. The cavity can be operated in a self-excited loop configuration or it can be driven by a fixed frequency generator. All systems have in common an electrical signal which represents the cavity field that is compared to a reference signal. The resulting error signal is amplified and drives a modulator for the incident wave



a) open loop b) closed loop

Figure 3.2.24: *Simulation of cavity response in a) open and b) closed loop including beam. The cavity parameters are the Lorentz force constant* $K = -1\,Hz/(MV/m)^2$ *and the mechanical time constant* $\tau_m = 1\,ms$*. The simulated beam is injected after* $500\,\mu s$*. In open loop the best flat top response is achieved with a pre-detuning of* $230\,Hz$*. The sum of generator and beam current during beam acceleration will induce a steady state gradient of* $25\,MV/m$ *when the cavity is operated on resonance. The closed loop response shows the additional power needed to control the time varying detuning caused by the Lorentz force.*

to the cavity. These can be amplitude and phase controllers, I/Q controllers, or other means to control the vector of the incident wave.

For the TESLA RF system a digital RF control scheme has been adopted. The concept of digital control provides enormous flexibility in the control algorithms, diagnostics, and exception handling.

Digital RF Control Design For the TESLA RF system the following signals are measured to determine the state of each RF station consisting of 32 cavities driven by one klystron:

- field vector of each of the 32 cavities
- vector of the incident wave to each of the 32 cavities

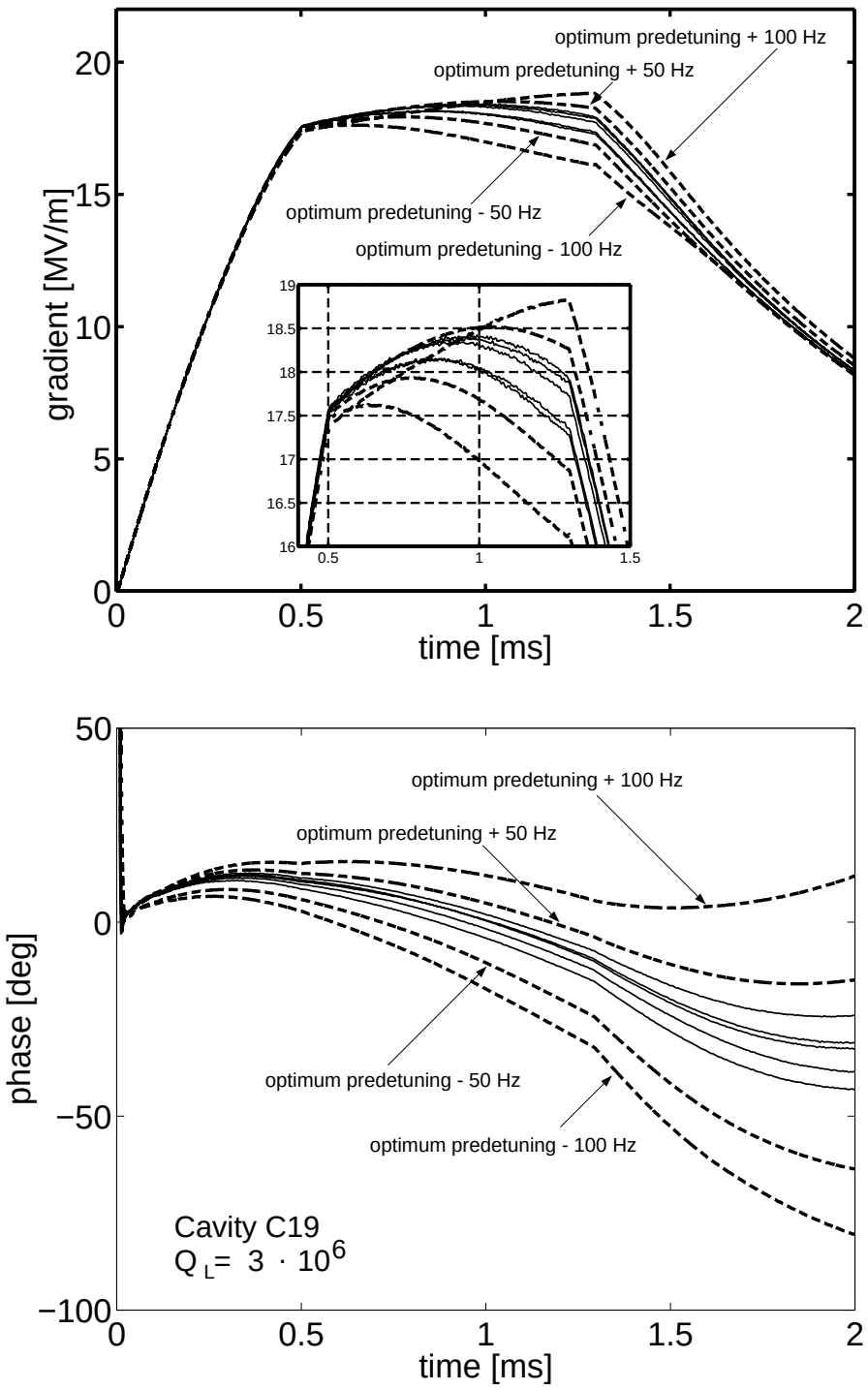


Figure 3.2.25: *Microphonics measured in the horizontal cryostat Chechia. The cavity is pulsed in open loop with simulated beam. The gradient and phase are modulated by the microphonics. The dashed lines show simulation results for 50 Hz and 100 Hz deviation from the optimum pre-detuning.*

- vector of the reflected wave from each of the 32 cavities
- vector of total incident and reflected wave measured at the klystron output

The actuators used for field control are:

- vector-modulator to control the incident wave common to 32 cavities
- cavity frequency control (motor activated tuner) for each of the 32 cavities
- manual adjustment for loaded Q and phase of incident wave for each of the 32 cavities using a three stub wave guide tuner and possibly a variable input coupler.

Fast control of the accelerating field can only be accomplished by modulation of the incident wave which is common to 32 cavities. Therefore fast control of an individual cavity field is not possible. The modulator for the incident wave is designed as an I/Q modulator to control the in-phase (I) and quadrature (Q) component of the cavity field instead of the traditional amplitude and phase modulators. The coupling between the loops is therefore minimized and control in all four quadrants is guaranteed.

The detectors for cavity field, and incident and reflected waves are implemented as digital I/Q detectors. The RF signals are converted to an IF frequency of 250 kHz and sampled at a rate of 1 MHz, i.e., two subsequent data points describe I and Q of the cavity field. The I and Q component which describe the cavity field vector are multiplied by 2x2 rotation matrices to correct the phase offsets and to calibrate the gradients of the individual cavity probe signals. The vector-sum is calculated and a Kalman filter is applied. The Kalman filter provides an optimal state (cavity field) estimate by correcting for delay in the feedback loop and by taking stochastic sensor and process noise into account. Finally the set point is subtracted and a time optimal gain matrix is applied to calculate the new actuator setting (I and Q control inputs to a vector modulator). Feed forward is added from a table in order to minimize the control effort. The feed forward tables are adaptively updated to reflect slowly changing parameters such as average de-tuning angle, microphonic noise level, and phase shift in the feed forward path.

Hardware The hardware of the digital feedback consists of 14-bit, 2 MHz ADCs (AD929 from Datel), the interface circuit to the input of the digital signal processors (DSP), the 6 DSPs (TMS320C40 from Texas Instruments), and two 16-bit, 10 MHz DACs which drive the vector modulator. Eight DSPs are used to acquire the probe signals of the 32 cavities and perform the matrix multiplication while one DSP is calculating the vector-sum and performing the feedback algorithm. The clock frequency of the DSPs is 50 MHz which allows for 25 instructions per microsecond. An additional DSP can be added to perform a more complex algorithm. The overall data processing time will be on the order of 2-3 μs. A cable delay of 600 ns, ADC conversion of 500 ns, and DAC conversion of 200 ns increase the total delay to 3-4 μs thereby adding less than 30 degrees phase shift at the unity gain crossing of 20 kHz. This leaves a phase margin of more than 60 degrees.

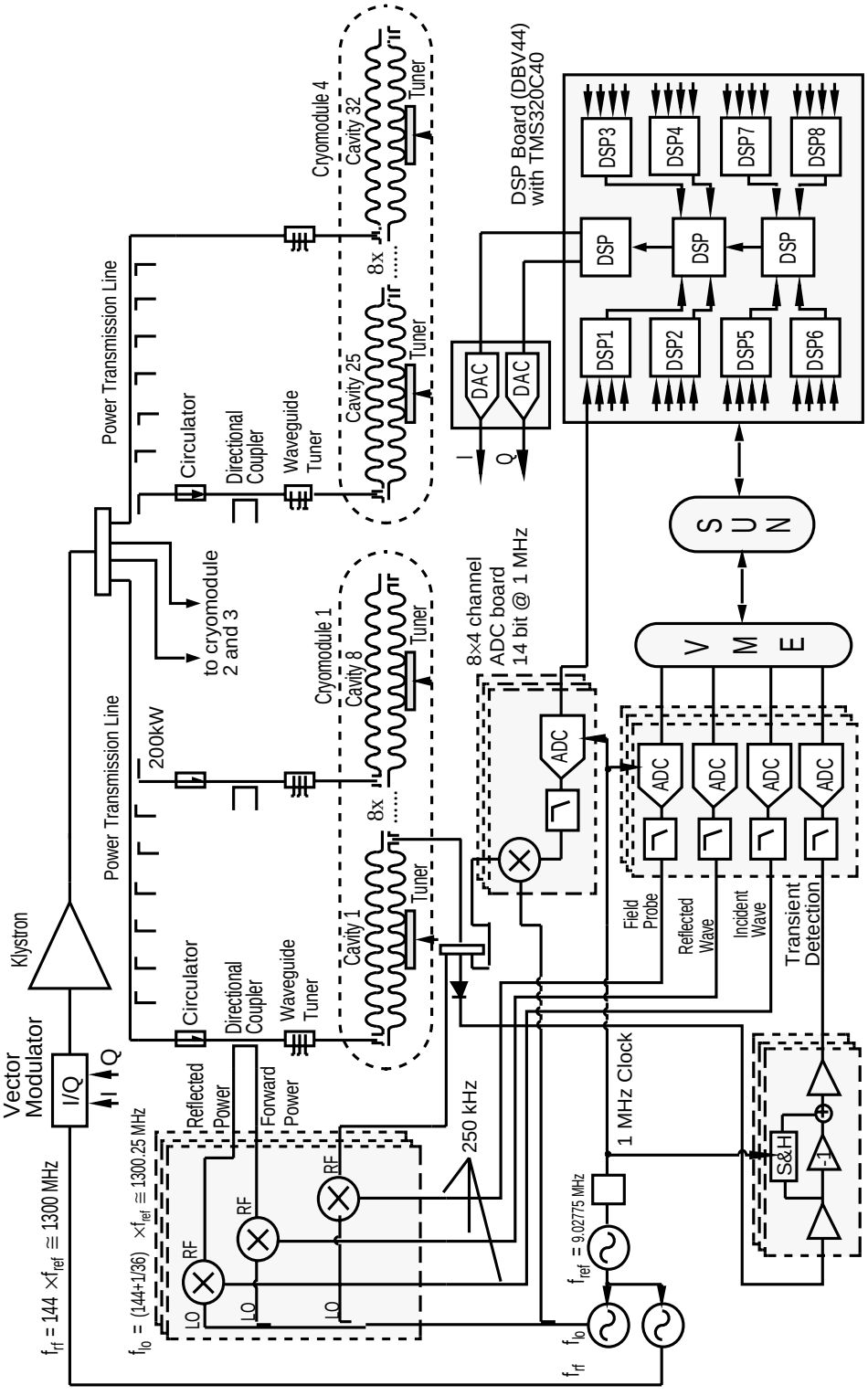


Figure 3.2.26: *Schematic of the Digital RF System for the TESLA Test Facility*

With the addition of the digital I/Q detectors for incident and reflected wave which are read by the control computers - these signals are not used in the feedback algorithm - the complete information about each cavity is available. Graphs of the individual signals are displayed for monitoring, optimization, and trouble shooting purposes. The tuner control is derived from the time dependence of the gradient signal. The calibration of the vector-sum and the phasing of the cavities is based on beam loading. The beam induced transients are observed at zero crossing and phase offsets are determined by nulling of the average transients. Dedicated hardware for transient detections allow for precise calibration of the vector-sum.

Feedback Algorithm The non-linear dynamics of a single cavity are described by a system of first order differential equations presented in state space form. The equations describe the envelope of the accelerating field as function of the driving terms which are composed of generator current and beam current. The last equation describes the coupling between accelerating gradient and the resonance frequency of the cavity.

$$\begin{pmatrix} \dot{\nu}_r \\ \dot{\nu}_i \\ \dot{\Delta\omega} \end{pmatrix} = \begin{pmatrix} -\omega_{1/2} & -(\Delta\omega+\Delta\omega_T) & 0 \\ \Delta\omega+\Delta\omega_T & -\omega_{1/2} & 0 \\ -\frac{2\pi\cdot K}{\tau_m}\nu_r & -\frac{2\pi\cdot K}{\tau_m}\nu_i & -\frac{1}{\tau_m} \end{pmatrix} \cdot \begin{pmatrix} \nu_r \\ \nu_i \\ \Delta\omega \end{pmatrix} + \frac{R\omega_{RF}}{2Q} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} I_r \\ I_i \end{pmatrix}$$

where ν_r, ν_i : real and imaginary component of the cavity voltage,
I_r, I_i : real and imaginary component of the total current (generator + beam)
$\Delta\omega + \Delta\omega_T = \omega_0 - \omega_{RF}$,
$\omega_0 = \omega_0(t)$: resonance frequency of the cavity,
ω_{RF} : operating frequency, $\omega_{1/2} = \frac{\omega_{RF}}{2Q}$: half width of resonance,
$\Delta\omega_T$: detuning by frequency tuner, τ_m : mechanical time constant,
K : Lorentz force detuning constant.

Linear and time invariant systems of the form $\dot{x}(t) = A \cdot x(t) + B \cdot u(t)$ are well understood in modern control theory and solutions for an optimum controller can be found in textbooks. For the above non-linear system which describes the cavity dynamics one can also find various solutions for a controller. It is, however, critical to keep the computational delay as small as possible to allow for sufficient gain bandwidth product and avoid instabilities due to delay. Since the data processing time (a few μs) will be short compared to the time constant of the cavity (700 μs), it is possible to correct for these delays and reach significantly higher feedback loop gains than in analog systems with the same delay.

The computational delay is minimized by treating the ensemble of 32 cavities like one cavity and applying the feedback algorithm only to the vector-sum. Simulations with realistic spreads in system parameters such as loaded Q_L, mechanical time constant τ_m, Lorentz force detuning constant K, power and phase of incident wave, calibration errors of the vector-sum, and initial detuning angle modulated by microphonics,

have demonstrated that the ensemble of 32 cavities is well described by the dynamics of a single cavity.

The present feedback algorithm applies a proportional gain to the error signal and uses a digital low-pass filter to reduce the sensor noise. More sophisticated algorithms such as time varying Kalman filter and time-optimal control are planned to be used in the near future.

The design of a feed forward algorithm is as important as the feedback algorithm. The Lorentz force induced gradient and phase fluctuations result in a repetitive error signal which can be compensated by feed forward which adds a fixed and negligible amount of CPU time. The feed forward tables are optimized by measurement of the average correction signal from the feedback loop and adjustment of the feed forward tables such that the average feedback signals are close to zero.

Measured Performance of the LLRF for the TTF

The digital feedback system has been extensively tested in the TTF with simulated beam to ensure that the open loop conditions are similar to the operating conditions under nominal beam loading. For simplicity the drive frequency has been varied to adjust the pre-detuning instead of the mechanical frequency tuner which must be used in routine linac operation.

Measured cavity parameters The cavity parameters for cavity D2 have been determined from open loop measurements. The cavity is pulsed and the transients for amplitude and phase are measured for different pre-detunings. The measured data is compared to simulation results in which the Lorentz force detuning constant and the mechanical time constant of the cavity are varied to achieve good agreement with the measurement.

Single cavity performance of the digital feedback system The performance of the digital I/Q controller for a single cavity operated at 20 MV/m is shown in Fig. 3.2.28. With feedback loop gains of 40 dB, the measured amplitude and phase fluctuations of the accelerating field are less than 0.5% rms and 0.5^o rms respectively thereby exceeding the uncorrelated noise requirements. The measured signal is dominated by 250 kHz noise as a result of systematic errors on the local oscillator signal caused by the gain imbalance and orthogonality errors in the vector modulator. This noise will be reduced significantly in the future with a new driver board which allows for individual adjustment of the four reference phases to guarantee precise 90^o steps.

The overshoot on the gradient signal at the time of (simulated) beam injection is due to inadequate adjustment of the feed forward tables at the time of the measurement. In the future, adaptive algorithms will be used to adjust the feed forward tables automatically to minimize the control effort of the feedback loops. The fast fluctuations on the gradient and phase signal on a time scale of a few ten microseconds are

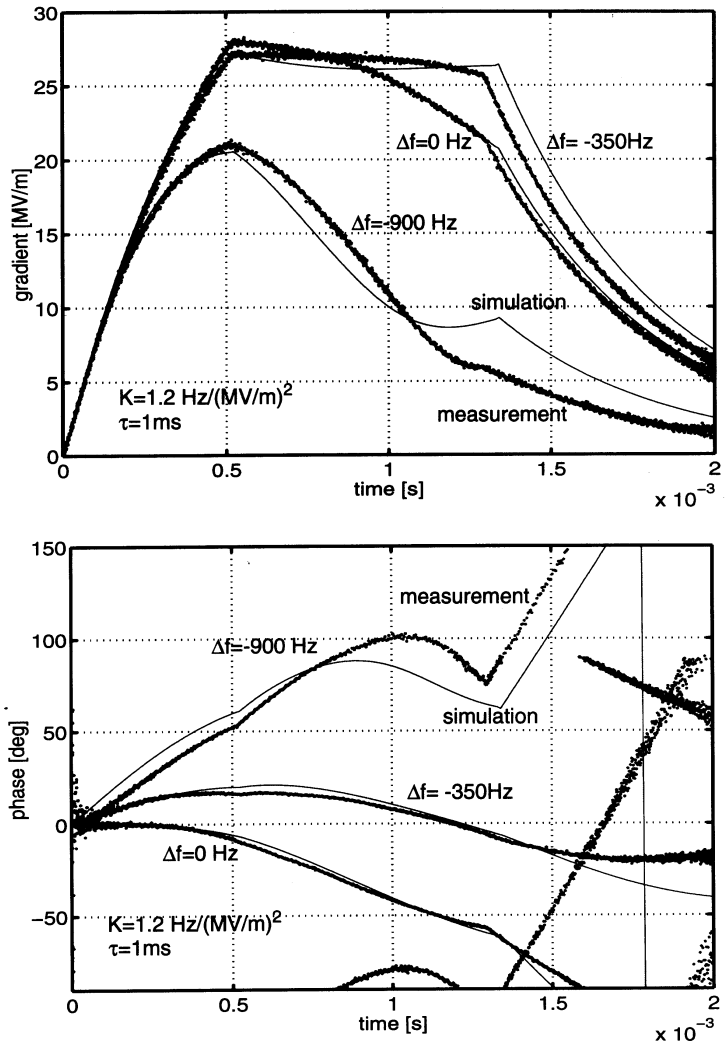


Figure 3.2.27: *Comparison of measurement and simulation results for cavity D2 in pulsed mode. The pre-detuning is varied. The Lorentz force detuning constant and the mechanical time constant of the cavity have been adjusted for good agreement.*

probably a result of phase noise on the reference line but further investigations are necessary to confirm and eliminate this noise.

It has been demonstrated that the digital RF system for the TTF provides the field regulation required for the TESLA linacs when controlling a single cavity. Simulations with measured cavity and other RF system parameters have shown that the regulation of the vector-sum of 32 cavities will not be a problem since the transfer function of an ensemble of many cavities is similar to that of a single cavity. It is however critical that the vector-sum is calibrated to better than 10 % in gradient and better than 1^o in phase, and that the klystron operated in its linear regime at all times.

Assuming realistic parameter spreads in loaded Q, the cavity parameters K and τ_m, errors in the calibration of the vector-sum, and microphonic noise levels of ± 50 Hz,

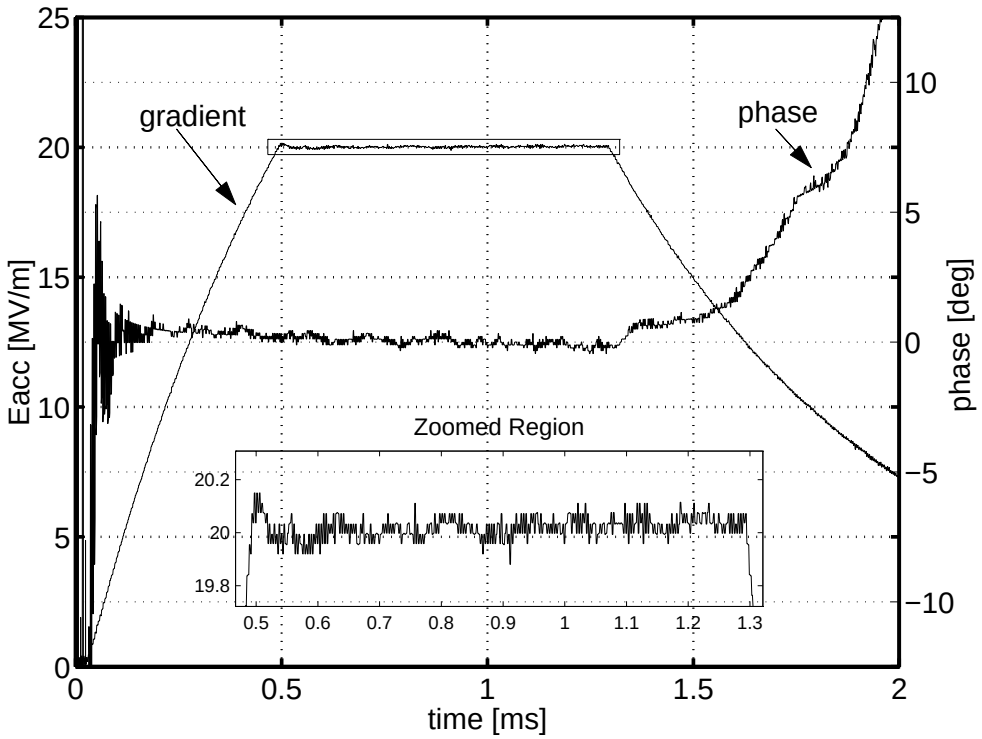


Figure 3.2.28: *Performance of the digital feedback system. Amplitude and phase stability exceed the requirements.*

the achievable average gradient will be approximately 2 MV/m lower than the average operable gradient. For an average energy gain of 25 MeV/m, the average operable gradient must be therefore of the order of 27 MV/m. The discrepancy is explained by the fact, that the combination of microphonics and parameter variations lead to modulated deviations from the ideal flat-top characteristics in the individual cavities thereby reducing the available energy gain.

3.2.5 Cryogenics

The present TESLA design requires the cooling of about 20,000 superconducting cavities down to $\leq 2\,\mathrm{K}$ by saturated superfluid helium in a bath. Grouped in numbers of 8 into 2448 cryomodules of about 12 m length the cavities are assembled along the 32 km length of the linear collider. Sections (units) of about 2.5 km length are supplied from a single refrigerator which has to provide also 4.5 K one-phase cooling for superconducting magnet packages (quadrupoles and correction dipoles) and for a radiation shield as well as 40 K to 80 K cooling for a second radiation shield and HOM absorbers. Generally two neighbouring refrigerators are housed in a common refrigerator hall above ground. Because of doubling the dynamic heat load due to the use for a Free Electron Laser (FEL), a special situation occurs in the first part of the linear accelerator near the DESY site.

Each refrigerator has a size roughly comparable to the ones used for HERA. Heat loads have been calculated for the static (without RF and current) and for the dynamic (full RF for 5 Hz beam bunch repetition rate respectively 10 Hz in the first two cryogenic units and 100 A current for magnet packages) cases. Dynamic nominal heat loads at 2 K dominate, reaching a nominal value of about 2 kW for units 3 to 12 and about 4 kW for units 1 and 2. A complete overview of the cryogenic loads at different temperature levels is given below.

Cooling schemes have been developed allowing stable steady state operation and cool down of smaller sections (half strings) partly in sequence. Refrigeration at below 2 K requires the use of cold compressors in the cold box. System and cold box designs have been studied by industry [28], [29] (2 firms), the results of which have been incorporated here as far as possible.

Within a year the cooling will have to run continuously over a period of about 10 months including cool down and warm up times (about 10 days each). Two months are foreseen for maintenance. An important aspect is the availability which is aimed at close to 100 % for the whole refrigeration system. Studies and proposals from industry have been used here too.

The whole cryogenic system will be generally operated from a central control room using workstations with X-Windows for visualisation which are connected to front end systems via high speed communication links for process control and data collection.

The total space required for the refrigeration system above ground in each refrigeration hall is about $2500\,\mathrm{m}^2$.

3.2.5.1 Layout

The two linacs consist of 19584 superconducting 9-cell cavities, each contained in a vessel made from titanium, in which it is cooled to a temperature of $\leq 2\,\mathrm{K}$. For ease of installation and cost minimisation 8 cavities are grouped and assembled into a common cryostat (module). A high order mode (HOM) absorber is housed in the beam tube at each module interconnection. A magnet package containing a superconducting quadrupole of varying length for beam focusing, a vertical and a horizontal correction dipole for beam steering and a "pill box" type beam position monitor is housed in every

second module from 3 GeV to 50 GeV, in every third module from 50 GeV to 150 GeV and in every fourth module above 150 GeV (see Chapter 3.2.2). Thus the length of a module varies from 11.37 to 13.12 m.

For cooling purposes 12 modules are grouped in a "string" of 141.69 to 142.552 m length (see Fig 3.2.30).

Finally, 17 strings with string interconnection boxes (SCBs) between neighboured ones and with feed and end boxes form a unit of about 2500 m length containing 204 modules cooled from a single refrigerator (see Fig. 3.2.29). The maximum number of magnet packages in a unit is 102.

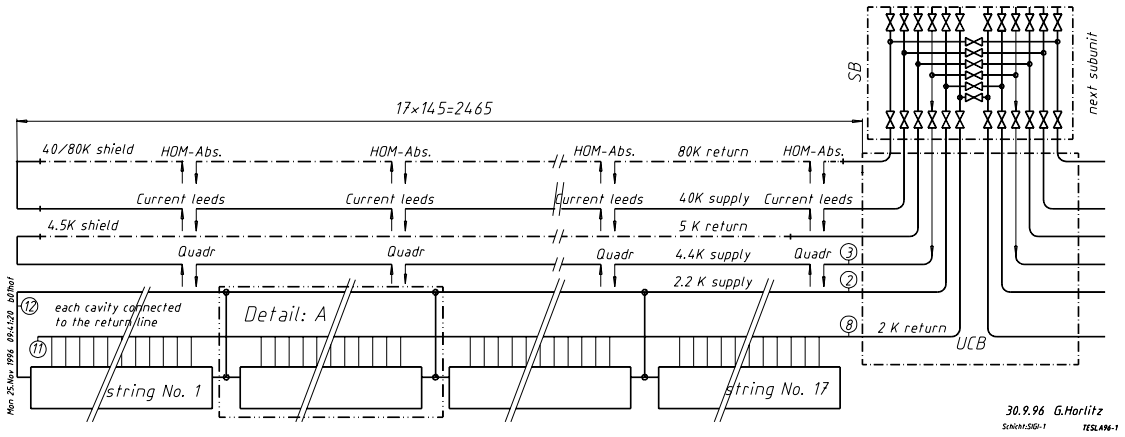


Figure 3.2.29: *Cryogenic supply of a unit*

The whole linear collider consists of 12 units placed in an underground tunnel. All modules are mounted horizontally (along an equipotential line with respect to gravity) with the exception of the modules of the first 2 units at the DESY site end which follow a line with a slope of –5 mrad (Fig. 3.2.31).

3.2.5.2 Overview of Cryogenic Load

Static and dynamic cryogenic heat loads of a module without magnet package and of the magnet package of maximum length have been estimated and are listed in Tables 3.2.3 and 3.2.8. Static load in this case means the load without any RF and beam in the cavities and without current in the superconducting magnets. Dynamic loads are shown for maximum accelerating fields (25 MV/m), quality factor $Q_0 = 5 \cdot 10^9$, 5 Hz beam repetition rate and 100 A current in the magnets. The dynamic heat load dominates (about a factor 10 higher than the static load) at 2 K, is relatively small at 4 K and is considerable at 70 K. While most of the accelerator will be operated at 5 Hz beam bunch repetition rate the cavities in the first two units will run at 10 Hz because this part will be used for coherent γ ray production in a Free Electron Laser (FEL) in addition.

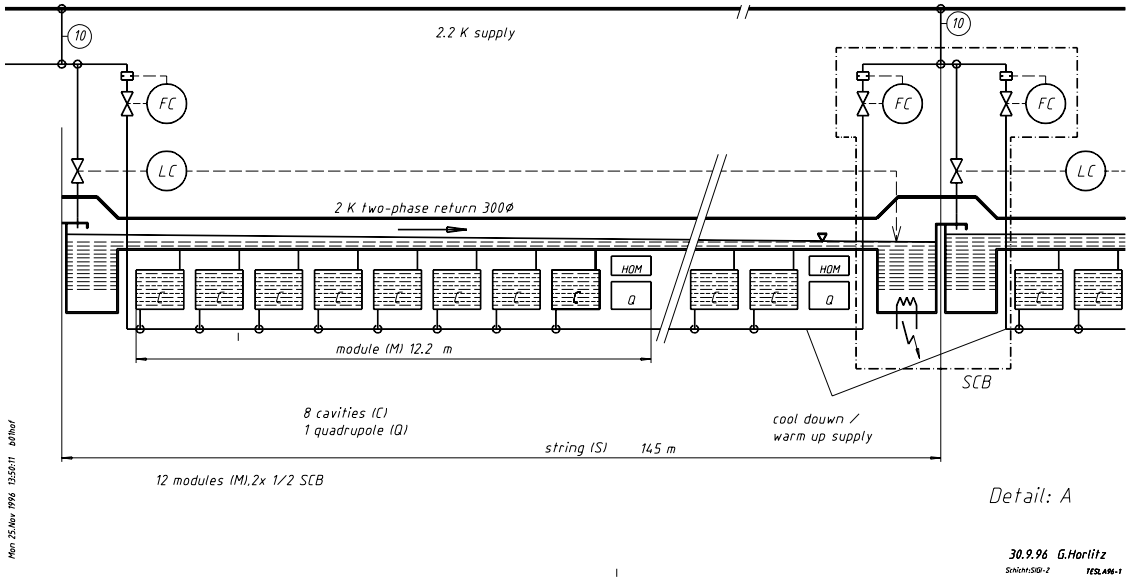


Figure 3.2.30: *Cryogenic supply of a string (detail A of Fig.3.2.29)*

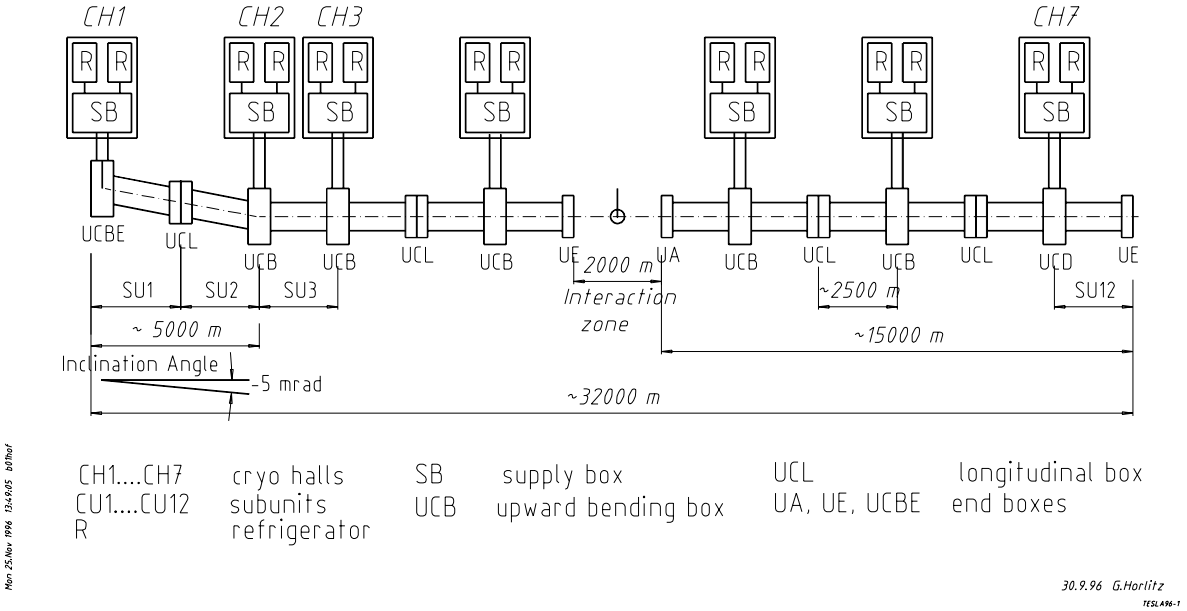


Figure 3.2.31: *Cryogenic supply of the collider*

Heat loads have been estimated also for the string interconnection, feed and end boxes as well as for the transfer lines up to the valve box at the refrigerator above ground (Table 3.2.18). Nominal heat loads for the standard units 3 to 12 and the special ones 1 and 2 are contained in Table 3.2.19. There are small differences (due to different lengths of magnet packages) among the standard units and between unit 1 and 2, but these have been neglected here. The total loads in the first two units

are higher by a factor of 1.9 at 2 K, a factor of 1.2 at 4 K and a factor of 1.4 at 70 K. Table 3.2.19 shows also the design heat loads for the layout of the cryogenic plants which contain a safety factor of 1.5.

From the heat loads the total electrical power needed has been calculated and listed in Table 3.2.20. The conversion factors used are 800 W_{el}/W at 2 K, 250 W_{el}/W at 4 K and 25 W_{el}/W at 70 K.

Additional heat loads, units 1-2, 10 Hz (W)						
	2 K	**2 K**	**4 K**	**4 K**	**70 K**	**70 K**
	static	**static + dynamic**	**static**	**static + dynamic**	**static**	**static + dynamic**
16 interconn. boxes	19	27	83	83	736	736
end box A	1	1	4	4	40	40
end box B	1	1	4	4	40	40
distribution box	16	16	67	67	844	844
transferline	13	13	10	10	105	105
vacuum barrier	5	5	20	20	200	200
sum	**55**	**63**	**188**	**188**	**1965**	**1965**
Additional heat loads, units 3-12, 5 Hz (W)						
	2 K	**2 K**	**4 K**	**4 K**	**70 K**	**70 K**
	static	**static + dynamic**	**static**	**static + dynamic**	**static**	**static + dynamic**
16 interconn. boxes	19	23	83	83	736	736
end box A	1	1	4	4	40	40
end box B	1	1	4	4	40	40
distribution box	16	16	67	67	844	844
transferline	13	13	10	10	105	105
vacuum barrier	5	5	20	20	200	200
sum	**55**	**59**	**188**	**188**	**1965**	**1965**

Table 3.2.18: *Additional nominal heat loads for cryogenic supply of a unit*

3.2.5.3 Cooling

The cavities have to be cooled to $\leq$ 2 K in order to assure the required acceleration voltage (25 MV/m) and quality factor $Q_0(> 5{\cdot}10^9)$. For this purpose the helium vessels of the cavities are filled with superfluid helium (Helium II) at a pressure of $<$ 31 mbar.

One-phase helium of 2.2 K and about 1.2 bar is flowing from the refrigerator through

Units 3 –12 (5 Hz) (W)						
	2 K	**2 K**	**4 K**	**4 K**	**70 K**	**70 K**
	static	**static + dynamic**	**static**	**static + dynamic**	**static**	**static + dynamic**
204 cryomod.	216	2205	2095	2613	10294	16151
quads	11	20	44	55	1679	3516
add. heat	55	59	188	188	1965	1965
sum (nomimal)	**282**	**2284**	**2327**	**2856**	**13938**	**21632**
plant design	**423**	**3426**	**3491**	**4284**	**20907**	**32448**
Units 1 – 2 (10 Hz) (W)						
	2 K	**2 K**	**4 K**	**4 K**	**70 K**	**70 K**
	static	**static + dynamic**	**static**	**static + dynamic**	**static**	**static + dynamic**
204 cryomod.	216	4194	2095	3131	10294	22008
quads	11	29	44	66	1679	5353
add. heat	55	63	188	188	1965	1965
sum (nominal)	**282**	**4286**	**2327**	**3385**	**13938**	**29326**
plant design	**423**	**6429**	**3491**	**5078**	**20907**	**43989**

Table 3.2.19: *Nominal heat loads of the units (incl. additional heat loads of Table 3.2.18) and values assumed for the cryo plant design (incl. a 50 % safety margin).*

AC-power (MW)	**static**	**stat. & dynam.**
units 3-12		
nominal	1.156	3.083
plant design	1.734	4.624
units 1-2		
nominal	1.156	5.009
plant design	1.734	7.514
total (units 1-12)		
nominal	**13.874**	**40.844**
plant design	**20.811**	**61.266**

Table 3.2.20: *Nominal operating AC-power requirement for the cryogenic system and requirement for full capacity (incl. 50 % safety margin).*

all cryomodules of a unit (see Fig.3.2.29). At the end of each cryogenic string a fraction of this helium is expanded in a JT valve into a liquid/gas separator inside the string connection box or inside the end box of the unit (see Fig. 3.2.30). Liquid as well as gas is flowing back to the refrigerator through the 300 mm (350 mm in case of units 1 and 2) diameter 2 K helium two-phase pipe. At the position of each cavity liquid helium is flowing into the helium vessel. Heat generated at the cavity is transported through the liquid to the surface where liquid is evaporated. Excess liquid – if any – at the end of each string is collected in a helium bath and evaporated there with the aid of an electric heater while the helium gas is passing forward through all strings back to the refrigerator. Proper control of the JT valve will assure that there is nearly no excess liquid helium entering the bath.

As in the first unit supplied from a refrigerator on the DESY site the beam has a slope of –5 mrad, the JT valves for expansion of the one-phase 2.2 K helium sit in the string interconnection boxes at the string entrances (seen from the refrigerator). The liquid Helium II thus flows off from the refrigerator while the gas returns in opposite direction. It is believed that this counter flow of gas and liquid in the 350 mm diameter helium two-phase pipe is possible in this flow regime (306 g/s at maximum) without creating slugs and mist in the liquid. However, detailed measurements in a test set-up need to be performed.

Pressures, temperatures and liquid levels in the 300 mm (350 mm) diameter 2 K two-phase helium tube are governed by the pressure drop of the flowing gas and liquid, by the evaporation of the liquid and by the hydrostatic pressures for the different cases: horizontal and inclined beam tube. Extensive calculations have been done by G. Horlitz [30]. The pressure (temperature) in the tube decreases from 31 mbar (2.0 K) at the far end of the unit down to 28 mbar (1.97 K) respectively 21 mbar (1.88 K) at the near end (seen from the refrigerator) of the unit for the standard respectively the first two units. The liquid level in the 300 mm (350 mm) diameter tube has its maximum (about 2 cm high) at the liquid entrance just behind the phase separator and decreases down to zero towards the end of a string.

One-phase helium of 4.5 K and 5 bar is flowing from the refrigerator through all quadrupoles of the unit in series, warming up to about 4.7 K towards the far end of the unit. On its way back it is cooling the 4.5 K radiation shields, returning to the cold box at a pressure of $\leq$ 2.3 bar.

Helium gas of 40 K and 19 bar is flowing from the refrigerator through heat exchangers at the HTS current leads of the superconducting magnet packages of the unit and returns to the refrigerator cooling the 70 K radiation shields and the HOM absorbers at the module interconnections. As the cavities used for TESLA have been heat treated and thus freed from dissolved hydrogen, fast cool down will not be necessary which eases the operation of the cryo-plant. However, in case fast cool down becomes necessary, it can be provided.

3.2.5.4 Cryogenic Plant

There will be seven refrigeration halls each containing two refrigerators (see Fig. 3.2.31). Generally each refrigerator is cooling one unit. Only the first two units because of their enhanced heat load at 10 Hz operation will be cooled from two refrigerators each. This concept allows to standardise plant design (for instance by using identical heat exchangers, cold boxes and piping) and thus standardise operation and maintenance routines as well as spare part inventory. Differences between the refrigerators for the first two and the remaining units are only in the turbines, the cold compressor sizes [28] and the numbers of warm compressors [29].

Each refrigerator has its own compressor system consisting of seven compressor skids (2 low pressure, 3 medium pressure, 2 high pressure ones) and three oil removal skids [28] (see Fig. 3.2.32). There is a complete redundant compressor system in each refrigeration hall. The low pressure (LP) skid operates with a suction pressure of 1.05 bar and compresses gas returning from the 4.4 K bath, from the 5 K shield and from the exhaust of the cold compressor line. The medium pressure (MP) skid operates at variable suction pressure and compresses gas returning from the turbine exhaust keeping constant the pressure ratio across the turbines. The high pressure (HP) skid compresses gas from the LP and MP skids to a pressure of 19 bar. In this scheme the efficiencies of both compressors and turbines are very high.

During steady state operation after entering the cold box the gas is running through a simple guard adsorber which removes small impurities (see Fig. 3.2.33 for cold box layout). It can be bypassed for regeneration [28]. Further down in the refrigerator the gas passes through several heat exchangers and is expanded in turbines. There are four expansion stages [28]. The first consists of 2 turbines in series and cools and expands the gas from 80 K, 17 bar of the shield cooling return, to about 40 K and medium pressure. If the expansion ratio of the turbine is high enough only one turbine is sufficient here [29]. The second stage also consisting of 2 turbines in series expands from high pressure at 21 K to medium pressure at 11 K. The third stage expands from high pressure at 10 K to medium pressure at about 5.5 K. The fourth stage expands from high pressure at about 5.9 K to about 5 bar at 5.4 K. This gas flow is then used for cooling the 102 (at maximum) quadrupoles and the 4.5 K shield and will later on flow through the cavities [28]. Before entering into the quadrupoles this gas is subcooled in a coil type heat exchanger inside a helium bath. The bath is filled by liquid obtained after JT expansion of the gas returning from the 4.5 K shield. The liquid level is controlled by an electric heater. At the far end (seen from the refrigerator) of a unit the temperature of the quadrupole cooling gas will be ≤ 4.8 K. Returning to the cold box this gas cools the 4.5 K shield.

For cavity cooling 4.5 K liquid at 1.2 bar is taken out of this bath, subcooled to 2.2 K in a heat exchanger by the returning gas stream and expanded at each cavity string to a pressure of about 26 mbar [28]. The low pressure is obtained by running the returning gas through 5 cold compressors. Between the first two and the last three cold compressors the gas flow is recooled to 5.6 K for ease of cold compressor control. The fifth cold compressor is needed only in the RF-off mode in order to achieve the required

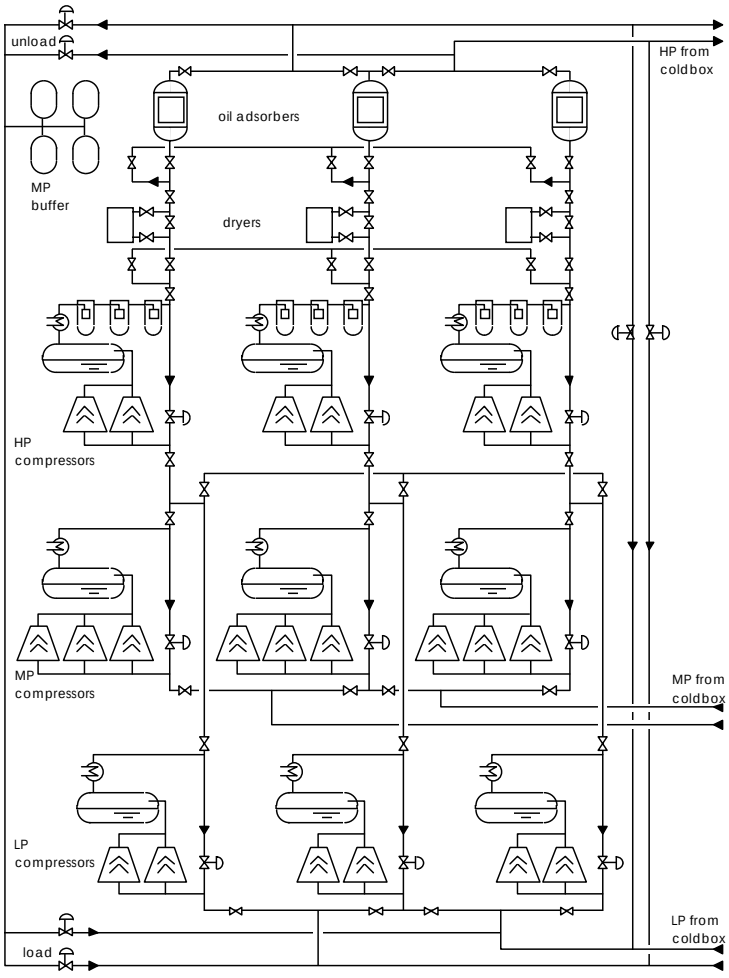


Figure 3.2.32: *Warm compressor system of a refrigeration hall [28].*

outlet pressure of 1.2 bar. When the RF is in full operation this cold compressor is bypassed [28].

During cool down liquid nitrogen is used for precooling. The purified helium gas can be fed into the refrigerator cold box either at room temperature (during the first cool down phase) or at 80 K [28]. For purification during the cool down phase there is a full flow switchable adsorber unit outside the refrigeration cold box.

3.2.5.5 Components and Subsystems

Warm compressors are oil lubricated screws with casing and rotors made of cast iron utilising radial sleeve bearings. To prevent leakage of process gas double mechanical shaft seals are provided. Each compressor has a hydraulically operated slide valve for stepless capacity control of the compressor flow between 10 % and 100 % [28].

The dehydration skid consists of switchable dual bed adsorbers filled with a molecular sieve [28]. Regeneration is accomplished by closed loop gas circulation at elevated

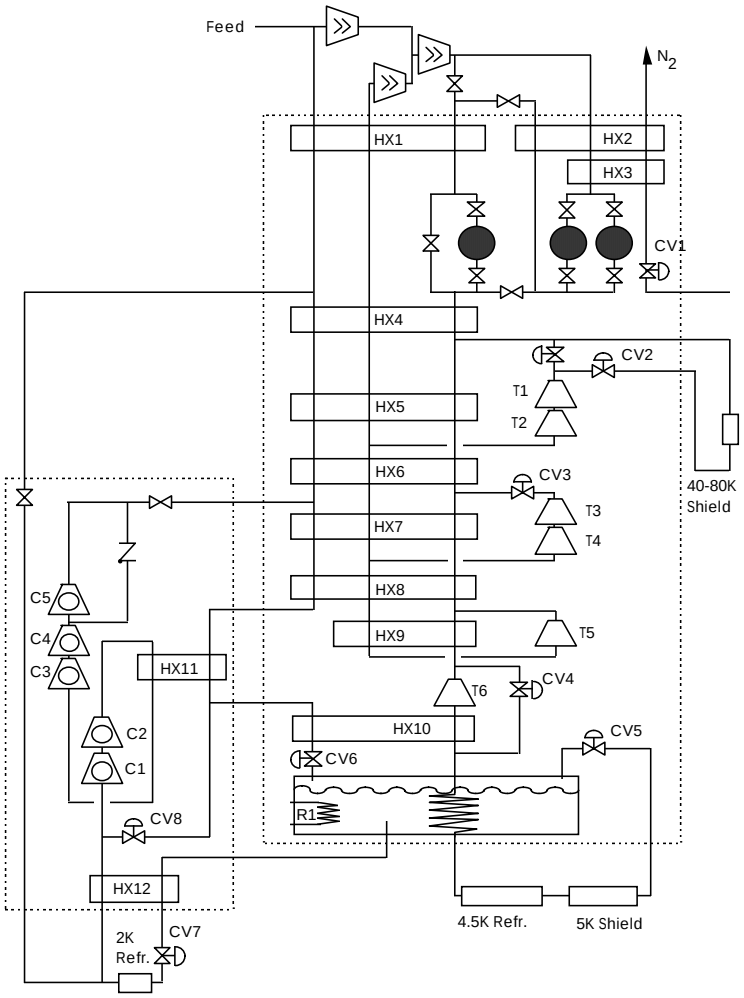


Figure 3.2.33: *Cold box layout [28].*

temperature, subsequent evacuation and cool down.

The oil removal system is skid mounted and consists of a three stage coalescing filter system, a charcoal adsorber and a dust filter [28]. The adsorber will remove both, vapour phase and liquid impurities to less than 10 ppb per weight. At present horizontal installed cylindrical tanks with a total volume of 1000 m^3 for every two cold boxes are foreseen for storing warm medium pressure (20 bar) helium gas. They may be installed outdoors.

Helium is stored as a liquid in a dewar of 120 m^3 volume serving two cold boxes each. During shut down periods the evaporating gas is compressed into the warm buffer by one compressor unit. The total storage capacity has a margin of 26 %.

The outer shells of cold boxes, cold compressor boxes and valve boxes are designed as vertical vacuum tanks made from carbon steel [28]. For easy access the vacuum jacket can be disconnected and lowered from the top.

Heat exchangers are very compact aluminum brazed counter flow plate fin types of

up to five different gas streams in the same core [28].

Different types of turbines may be used. The Linde turboexpander is a small, single stage centripetal turbine which is braked with a directly coupled, single stage centrifugal brake compressor [28]. It is equipped with dynamic gasbearings which function at ambient temperature. For start up auxiliary magnetic bearings are used. The speeds range from 75000 rpm to 282000 rpm. The gas cooler at the warm compressor side is cooled with water. The turbine housing is fastened to the top plate of the cold box and projects vertically into its vacuum space. The bearing cartridge forms an integral part of the turbine and can easily be removed from outside without breaking the vacuum of the cold box. The main features of the L'Air Liquide turbines are static gas bearings operating at room temperature, relatively large clearances not very sensitive to impurities, high expansion ratios and easy way of speed control [29]. Several turbine sizes exist with speed between 50000 rpm and 600000 rpm.

There are also different types of cold compressors under consideration. The Linde cold compressor is a new design which has been tested at CERN successfully. It is a single stage radial flow turbo-compressor [28]. This concept offers high reliability, high efficiency and a compact design. The standardised cartridge design allows easy access to the machine without interfering with the cold box vacuum. Lubricated ceramic bearings are used to carry the vertical shaft. The electric motor is surrounded by process atmosphere. The maximum operating speed would be about 1000 rpm. The L'Air Liquide cold compressor successfully used for Tore Supra and CEBAF has active magnetic axial and radial bearings and roll bearings for emergency [29]. A new development for CERN is using fully pneumatic bearings [29].

There are several filters within the cold box to protect the turboexpanders from being damaged by mechanical impurities. Cryogenic adsorbers are installed inside the cold box at 80 K and 20 K level [28]. They protect the colder equipment from impurities. They consist of an activated charcoal bed fixed between rigid upper and lower filter packages. Replacement of the bed is possible. Regeneration is performed by the circulation of warm gas.

The string interconnection boxes (SCBs) contain a JT valve expanding the 2.2 K forward flow into a phase separator, a barrier for flow of liquid between neighboured strings, a liquid collecting box with electric heater, valves and instrumentation [28].

Feed and end boxes of each unit contain in addition cold/warm transitions and pneumatic valves for the beam tube.

Transferlines will contain sleeves for thermal and vacuum separation.

The vacuum of each cold box will be pumped continuously by a system consisting of a two stage primary pump, an oil diffusion pump with water cooled baffle and all necessary valves and instrumentation.

3.2.5.6 Control System

The whole system consisting of 7 refrigerator halls (14 refrigerators), and valves and instrumentation in the collider tunnel generally will be operated and controlled from one central control room. Workstations with X-Window are used for visualisation.

These workstations are connected to the front-end systems through a high speed communication link. The front-end systems collect all I/O data and include all the process control functionality based on standardized program blocks. Every major subsystem like cold box, or compressor unit will have its own I/O rack in order to collect all data from the plant and the string of cryomodules. Intelligent sensors are connected directly to field buses. All temperatures have to be monitored precisely. Temperatures of the cavities are measured accurately using pressure gauges for the 2 K two phase flow.

Local control terminals (operator interfaces) will allow access to the plant control, especially during commissioning, for service and maintenance at every single unit.

3.2.5.7 Availability of Cryogenics

Availability of the cryogenic system for high energy experiments is a very important issue. It is anticipated that the availability of the whole cryogenic system of TESLA will be close to 100 %. This is achieved by several methods. Full redundancy will be installed for all process steps except for the refrigeration cold box. In each refrigerator hall containing two refrigerators there will be a redundant third warm compressor line as well as a redundant third cold compressor line. Utilities like instrument air and cooling water supply will be installed as parallel independent systems with redundancy. Electrical power will be buffered with uninterruptable power supply units (UPS) wherever possible.

Cold boxes will have individual availability of $> 99\,\%$. Important components like turbines, cold compressors, adsorbers and filters will have their own vacuum boxes and will be accessible from outside the cold box without breaking the cold box insulation vacuum [28].

In case that a refrigerator fails completely the concept of having two refrigerators in each refrigeration hall allows to cool two units from one refrigerator. Both units can be kept cold without running the RF until the defective plant has been repaired. If the heat loads stay within the calculated limits one could even run the collider at reduced bunch repetition rate (for instance 3 Hz) if the remaining plant in that refrigeration hall is running at full capacity, which can be further increased by using liquid nitrogen and liquid helium from the storage dewar for pre-cooling.

3.2.5.8 Utilities

The expected power consumption per cryogenic hall (two refrigerators) is listed in Table 3.2.21 [28].

3.2.5.9 Space Requirements

The space required at each refrigeration hall for the compressor system including oil and water cooler is about 40 m $\times$ 45 m. For cold boxes, cold compressor boxes, oil adsorbers and a liquid helium dewar the required space is about 25 m $\times$ 25 m.

Electrical power	3300 V	10750 kW
	380 V	87 kW
	220 V	3 kW
	24 V	3 kW
Cooling water		
cold box (for turbine and baffle cooling):		
	max. temp.	20 °C
	Δ T	10 °C
	consumption	15 m^3/h
compressors:		
	max. Temp.	30 °C
	Δ T	10 °C
	consumption	1050 m^3/h
Air for instrumentation		
	minimum pressure	7 bar
	consumption	150 Nm^3/h

Table 3.2.21: *Utility requirements per cryogenic hall.*

3.2.6 Failure Handling

As has been mentioned in Section 3.1, the long macropulses and also long bunch to bunch distance in the TESLA Superconducting Linear Collider have extremely positive consequences with respect to failure handling: a safety system can be employed which can "turn off" the beam even within one macropulse in case an emergency is indicated by enhanced loss rates. And even more, superconducting accelerating structures as well as the used superconducting quadrupole magnets have typical quenching times much longer than the single macropulse. Therefore the operation of the TESLA Linear Collider would be extremely safe.

In this section main emphasis is put on the discussion of possible beam transport failures during the passage through the main linac with its 15 km length. A beam hitting the beam tube wall, i.e. in most cases the iris of a superconducting accelerating cavity, can easily destroy the cavity if it is well focused. Similar to a beam dump window problem the beam macro pulse power will not only heat the hit spot. It will cause an "explosion" or cracking of the material. The system would be "vented" with more or less clean Helium from the surrounding He vessel. A shutdown including a warm up and cool down cycle would be necessary. This has to be avoided.

Since the bunch train is 240 km long, only 6.2 % of the bunch train is in the 15 km long linac during the macropulse. Producing a beam inhibit signal at any of the subcomponents should not take longer than approximately 10 μs. The maximum signal travelling time is determined by the linac length. At a typical speed of 5 ns/m (standard coaxial cable) the length of 14.57 km corresponds to roughly 75 μs, thus the triggering of a beam injection inhibit can start after 85 μs.

In a superconducting linac changes of the magnetic fields or of the accelerating fields happen on a time scale long compared to this delay time until injection inhibit. Quenching of superconducting magnets might start within a macropulse but the only consequence could be a slight drop in the field strength. The magnets need typically much more than 800 μs for a total quench. Nevertheless, the consequence of a switched off quadrupole is shown later (see below). Similar to this a quench of a group of accelerating cavities (e.g. 32 cavities) needs time. Here it is more important to study the consequences of a jump in the reference phase or the consequence of a klystron failure. This has been done, and the results are described in section 3.2.6.2.

3.2.6.1 Philosophy for beam injection

The general philosophy for beam injection should be as follows. An exception handler will check the "o.k."-state of all subcomponents along the linac. For the magnets it would be a check of the stationary operation, i.e. basically a check if a magnet has quenched. The case of having a short circuit in one of the coils is not as dangerous as for a warm magnet. The current would still take the superconducting path (i.e. the winding), and if not (e.g. during ramping) the magnet would start quenching because of heat production in the short circuit. This would be detected.

For the superconducting accelerating cavities the exception handler can do much more. The filling time before beam injection is about 500 μs, thus there is plenty of

time to check the actual amplitude and phase after e.g. 300 to 400 μs.

During the macropulse the beam transmission can be measured as well as the beam position along the linac. The results can be evaluated between macropulses (200 ms $\hat{=}$ 5 Hz repetition rate are enough for a very complete analysis) and can be used to detect slow changes.

As a summary it can be stated that a system check can be performed about 100 to 150 μs before beam injection. And the achieved "o.k." is valid until the end of the following macropulse. Nothing severe can happen as long as the beam injection parameters are fixed by a collimator, and as long as the linac operator does not overwrite safe limits for the settings of magnets.

3.2.6.2 Magnet and Cavity Failure's Consequence on Beam Optics

The consequences of switched off qaudrupoles and switched off groups of cavities have been studied. For the beam optics in the linac a scheme essentially identical to the one described in section 3.3 has been used.

Figure 3.2.34 shows the envelopes along the linac. The matching between the different FODO sections is possible by using five quadrupoles each, three of them at the end of the preceding section, and two of them at the beginning of the following section.

Figure 3.2.35 shows the consequence of switched off quadrupoles. The maximum beam size increases by a factor of three to five, depending on the position along the linac. The picture shows only two selected quadrupoles but switching off any of the other quadrupoles gives similar results although the combination of several switched off quadrupoles can clearly yield bigger beam sizes. This is shown in Fig. 3.2.36. Thus it will be necessary to monitor all quadrupoles before beam injection, e.g. together with the accelerating cavity's status during cavity filling. This is possible. The only consequence is that a magnet failure has to be repaired or compensated by using preceding quadrupoles before injection of any further macro pulses.

A klystron failure has only a minor influence on the beam optics. Figure 3.2.37 reflects the insensitivity to a variation of the injection energy. Similar to this, a switched off klystron (32 cavities) can be hardly seen in the enveloppes. Figure 3.2.38 gives the comparison between normal operation (solid lines) and irregular conditions (dashed lines) simulated by switching off the klystron output 50 μs after the end of cavity filling, i.e. at t = (0.5 ms + 50 μs) instead of t = (0.5 ms + 800 μs). After switching off the klystron the cavity field starts decaying but finally increases because of beam induced voltage. Taking the amplitude $E_{acc}(t)$ and the phase into account, one can easily calculate the energy gain per cavity over the beam pulse. The subsequent bunches get a time dependent energy gain. If one would not interrupt beam injection, the difference between the first and last bunch would be about 32 MeV per cavity or about 1 GeV per 32 cavities, i.e. per klystron. The consequent use of the above mentioned exception handler reduces this "energy spread" to about 0.15 GeV which is only 6×10^{-4} of the final beam energy.

A jump in the linac reference phase during the beam pulse can be treated in the

same way. The klystron power would reach its maximum and the exception handler would switch off the klystron after about 50 μs. The energy spread over the bunchtrain in this case reaches a peak value of 19 %, still small enough for safely transporting the beam to the end of the linac. The beam then hits the energy collimators and the enhanced loss rates trigger the safety dump system.

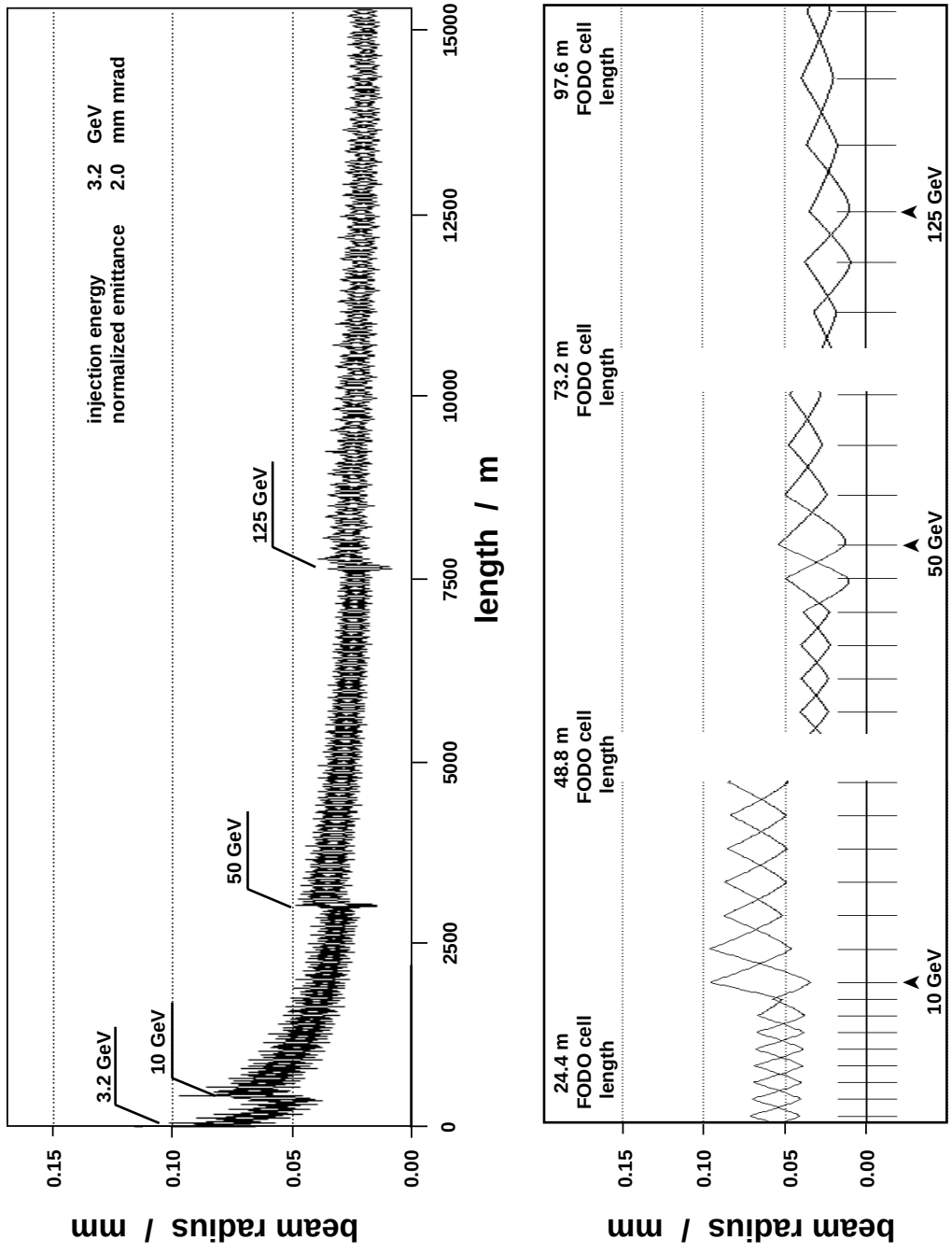


Figure 3.2.34: *Beam envelopes along the linac. The matching between the FODO sections is done by five quadrupoles each.*

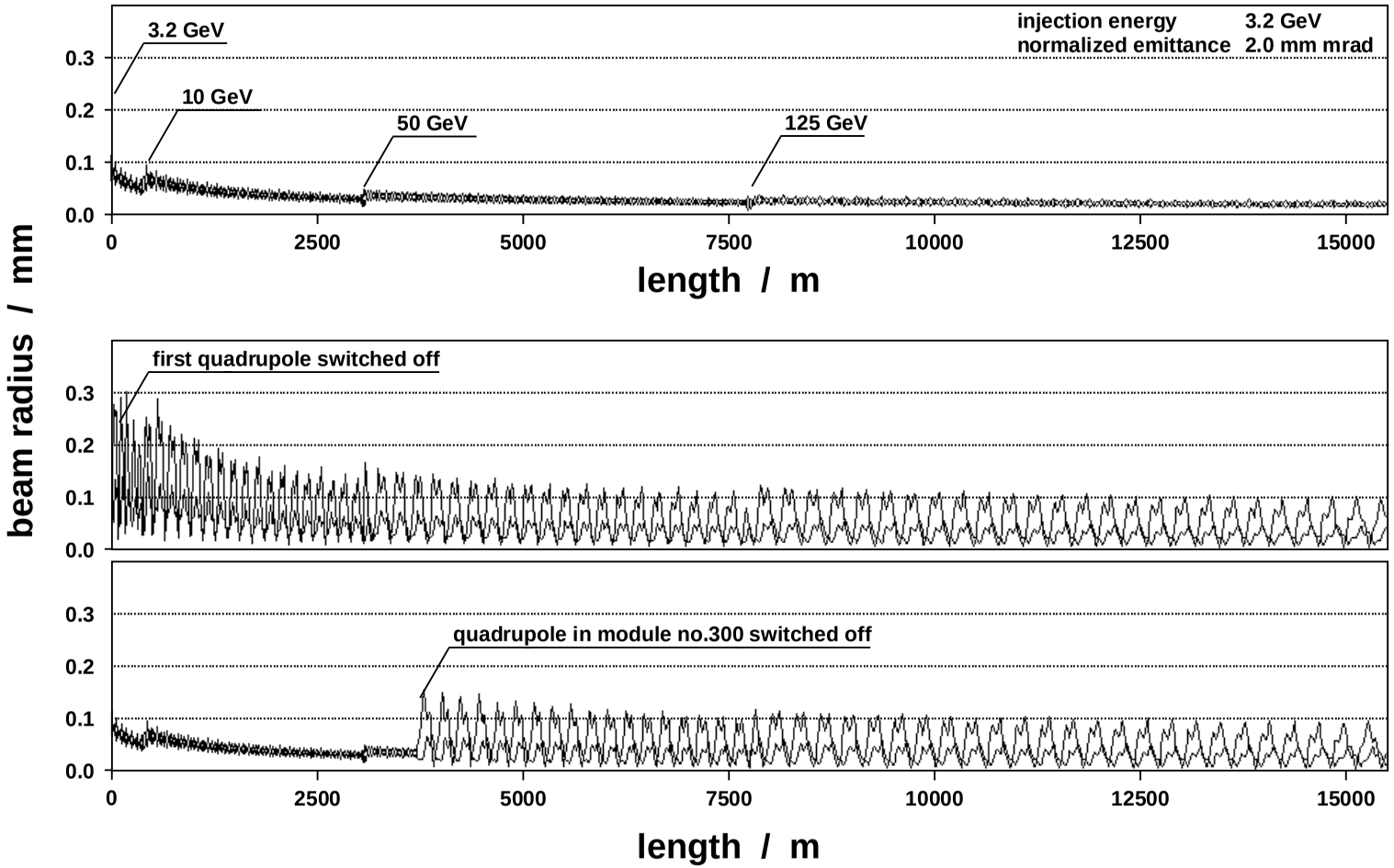


Figure 3.2.35: *The consequence of switched off quadrupoles. The maximum beam size increases by a factor of three to five, depending on the position of the quadrupoles.*

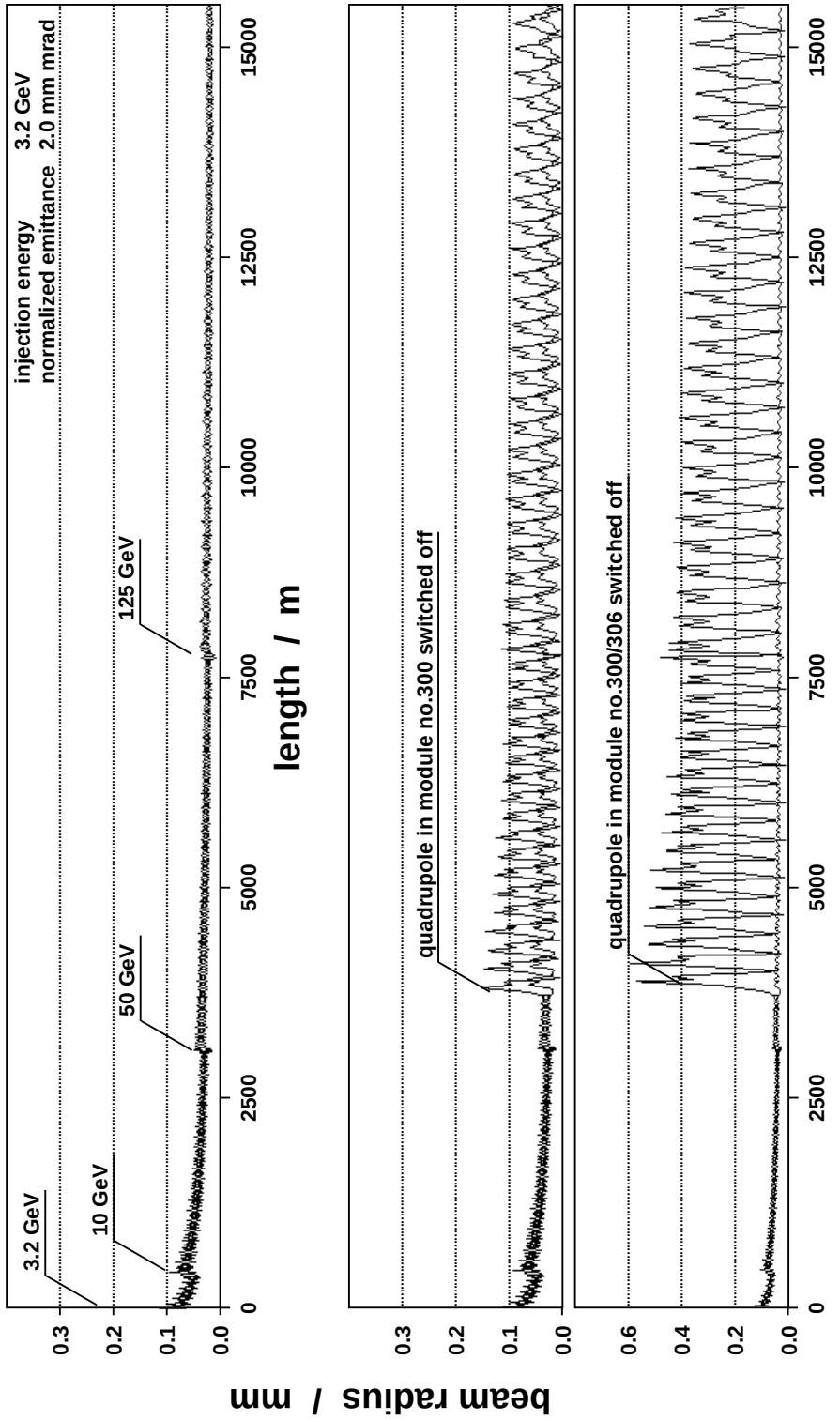


Figure 3.2.36: *An unfortunate combination of switched off quadrupoles yields larger beam sizes. The available aperture still corresponds to more than 50 standard deviations.*

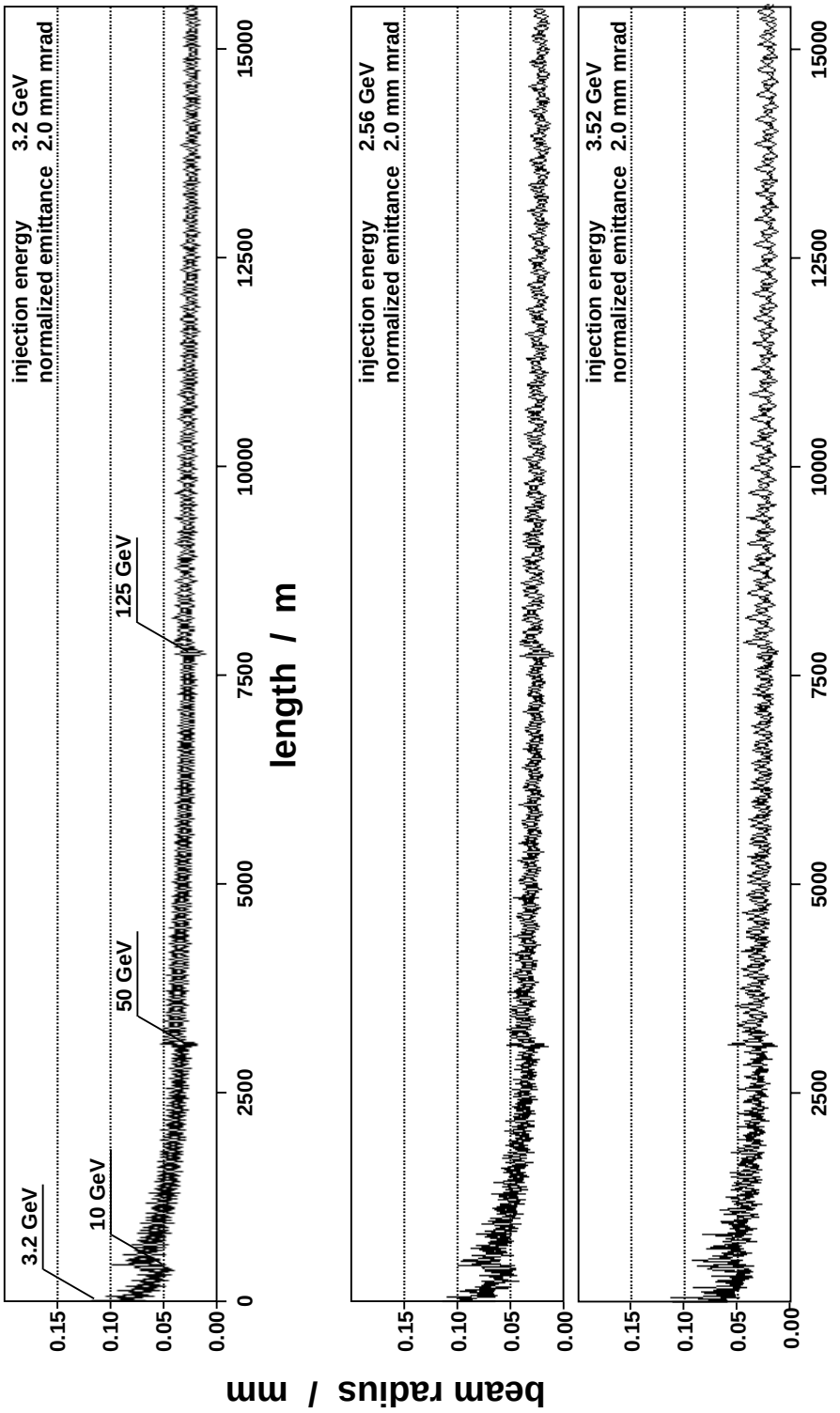


Figure 3.2.37: *The impact of different injection energies.*

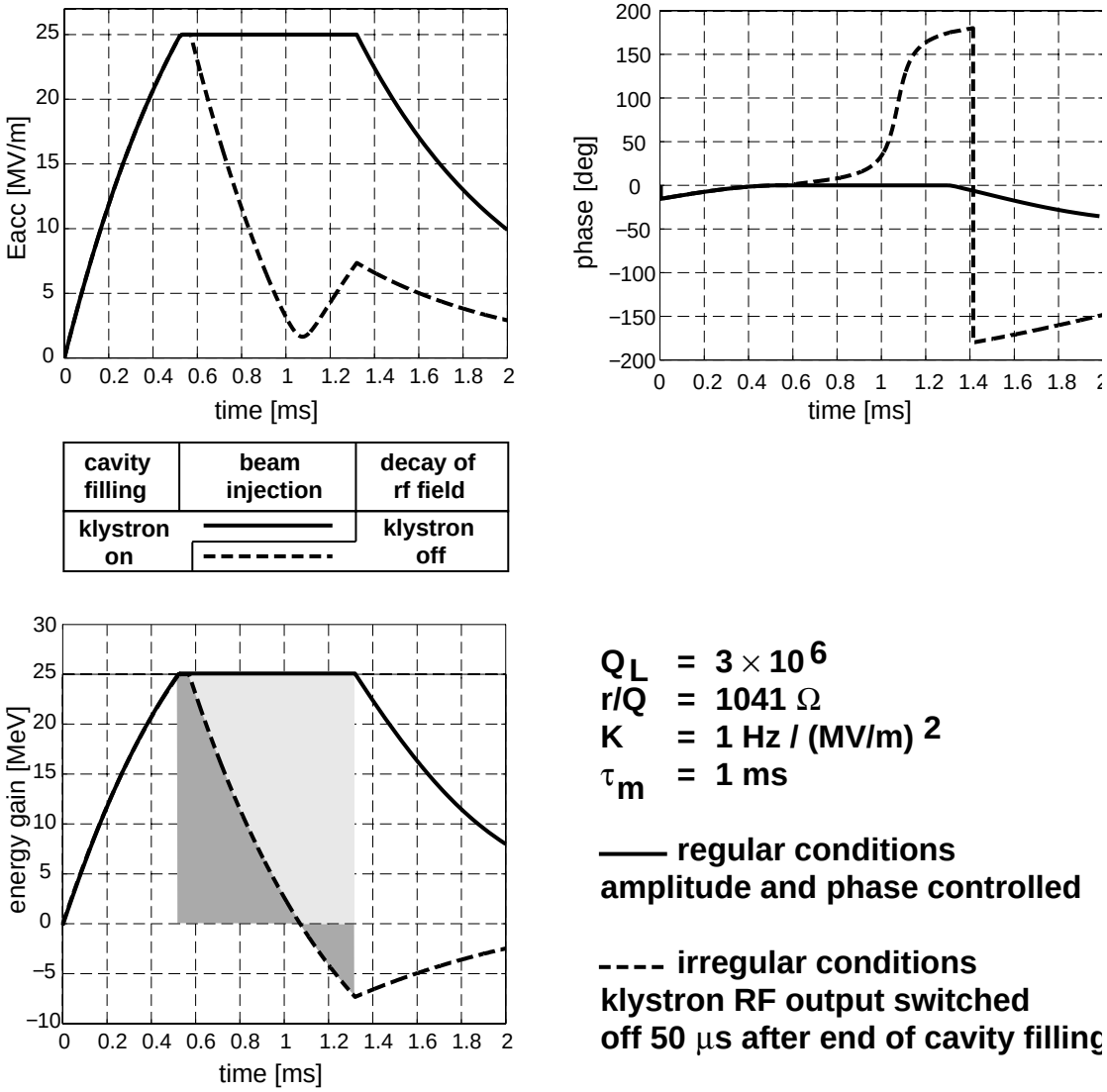


Figure 3.2.38: *Comparison between normal operation (solid lines) of a single accelerating cavity and irregular conditions (dashed lines) in which the klystron output is switched off 50 μs after the end of cavity filling. After switching off the klystron the cavity field starts decaying but finally increases because of beam induced voltage. The subsequent bunches get a time dependent energy gain.*

Bibliography

[1] C. Reece et al., Part. Acc. Conf. Dallas (1995)

[2] P. Kneisel, B. Lewis, Part. Acc., 53 (1996), 97,

[3] Ph. Bernard et al., Proc. Eur. Part. Acc. Conf. Berlin 1992

[4] B. Bonin, DESY-CERN School "Superconductivity in Particle Accelerators", Hamburg 1995

[5] W. Weingarten, DESY-CERN School "Superconductivity in Particle Accelerators", Hamburg 1995

[6] Th. Schilcher, TESLA-Report 95-12 (1995)

[7] B. Bonin, Supercond. Sci. Technology 9 (1996), pp 453 - 460

[8] H. Padamsee, Proc. 5th Workshop on RF Superconductivity, Hamburg 1991, p. 904

[9] H. Padamsee et al., Proc. 4th Workshop on RF Superconductivity 1990, KEK, Ed. Y. Kojima

[10] E. Kako et al., Proc. 7th Workshop on RF Superconductivity, 1995, Saclay, Ed. B. Bonin

[11] J. Graber et al., Nucl. Instr. Meth.A350 **(1994)** 572

[12] G. Müller, Proc. 3rd Workshop on RF Superconductivity, 1988, Argonne, Ed. K. Shepard

[13] J. Sekutowicz, TESLA-Report 97-02 (1997)

[14] W. Albach, G.A. Voss, Zeitschr. Angew. Phys. **9** (1957) 111

[15] F. Alessandria et al., *Design, Manufacture and Test of the TTF Cavity Cryostat*, Adv. Cryogenic Engineering 41A, p. 855, 1996.

[16] *High Voltage Breakdown in the Electron Gun of Linear Microwave tubes*, A.M. Schroff, and A.J. Durand in "High Voltage Vacuum Insulation" by R.V. Latham (Ed), Academic Presss, March 1995, ISBN 0124371752

[17] J.C. Terrien, G. Faillon, and P. Guidee, *RF Sources for Recent Linear Accelerator Projects*, 16th International Linac Conference, August 1992, Ottawa, Ontario

[18] Oleg A. Nezhevenko, *The Magnicon: A New RF Power Source for Accelerators*, IEEE Particle Accelerator Conference Vol. 5, May 6-9, 1991, San Francisco, CA, USA.

[19] Oleg A. Nezhevenko, private communication

[20] H. Pfeffer, C. Jensen, S. Hays, L.Bartelson, *The TESLA Modulator*, TESLA Report 93-30

[21] K.P. Juengst and H.Salbert, *Fast SMES for Generation of High Power Pulses*, 14th International Conference on Magnet Technology, 1995 Tampere Finland.

[22] W. Bothe, *Pulse Generation for TESLA, Considerations on SMES-Variants*, TESLA-Collaboration Report, April 1996.

[23] H. Salbert, *Theoretisches und experimentelles Modell einer gepulsten Hochleistungspulsversorgung mit supraleitendem magnetischem Energiespeicher*, Dissertation, Universitt Karlsruhe, 1996.

[24] W. Bothe, *Pulse Generation for TESLA, a Comparison of Various Methods*, TESLA Report 94-21

[25] W. Bothe, private communication.

[26] *The TESLA TEST FACILITY LINAC-Design Report*, Ed. D.A. Edwards, Tesla Report 95-01

[27] A. Mosnier and O. Napoly. *Energy Spread Induced in the TESLA Linac,* TESLA Report 93-07.

[28] Linde Kryotechnik AG, *Study for the Cryogenic System TESLA*, October 1996.

[29] L'Air Liquide, private communication, October 1996.

[30] G. Horlitz, *A Study of Pressures, Temperatures and Liquid Levels in the 2 K Kelvin Refrigerator Circuit of a 1830 m TESLA Unit under Different Conditions and System Configurations*, TESLA 94-17, December 1995 and G. Horlitz, *The Cryogenic System for the Superconducting e^+e^- Linear Collider TESLA*, Proceedings of the NIFS-Symposium on Cryogenic Systems for Large Scale Superconducting Applications, National Institute for Fusion Science, Toki, Nagoya, Japan, May 1996, to be published.

3.3 Beam Dynamics

3.3.1 Introduction

To achieve the desired high luminosities, long trains of bunches with very small transverse sizes are required at the interaction point. In this chapter, the issues, pertinent to emittance preservation during the acceleration in the main linac, are discussed. The primary sources of transverse emittance dilution in a high energy linear accelerator are the transverse wakefields excited in the accelerating sections and the dispersive errors caused by the focusing magnets. However, thanks to the low RF frequency and the large iris holes of the TESLA accelerating structure, wakefield effects are much smaller than in higher frequency designs working at room-temperature. Furthermore, the dispersive effects can be easily controlled because the beam energy spread is kept at a small value along most of the accelerator. In turn, that means that the alignment tolerances on the different elements - focusing magnets, beam position monitors, accelerating structures - which are achievable by nowadays survey techniques, are sufficient to ensure beam quality preservation without recourse to other very elaborated and complex correction techniques.

3.3.2 Beam Optics

The focusing of the beam is ensured by means of FODO lattices, which are characterized by the phase advance and the beta function. A constant phase advance of $\mu = 90°$ and an energy scaling of the beta function, $\beta \propto E^{\alpha}$ with $\alpha = 1/2$, are usually chosen. This beta scaling is optimal in the sense that the BNS damping technique, used to suppress the wakefield effect induced by a coherent betatron oscillation, is efficient all along the linac for a bunch of constant correlated energy spread. We will see thereafter that for the TESLA linac, BNS damping is not required because of low wakefields. A larger beta function and a weaker beta scaling with energy, $\alpha = 0$ to 0.2, is more appropriate. The wakefields effects are dominant for large α values (less quadrupoles) whereas the chromatic effects are dominant for small α values (more quadrupoles). The TESLA lattice was then optimized in such a way that a balance of both effects is achieved, leading to the smallest transverse emittance growth. Two beta steps in total, at 50 and 150 GeV, giving minimal emittance growth, are enough. Consequently, the number of matching sections, needed at each beta step, is reduced. On the other hand, at the low energy part of the linac, the focusing lattice must make the best of both TESLA and FEL beams, which have different energies. The difference in energy between the two beams is accomplished by decreasing the accelerating gradient from 25 MV/m to 15 MV/m. A lower phase advance of $\mu = 60°$ for the high energy beam is chosen in order to make the low energy beam stable. The same phase advance is kept all along the linac, because a higher phase advance after the extraction point of the FEL beam does not improve the results. The adopted beta function along the linac of TESLA 500 is shown on Fig. 3.3.1. The linac is divided into 3 sections, with quadrupole spacings of 2, 3 and 4 cryo-modules, respectively. The last quadrupoles

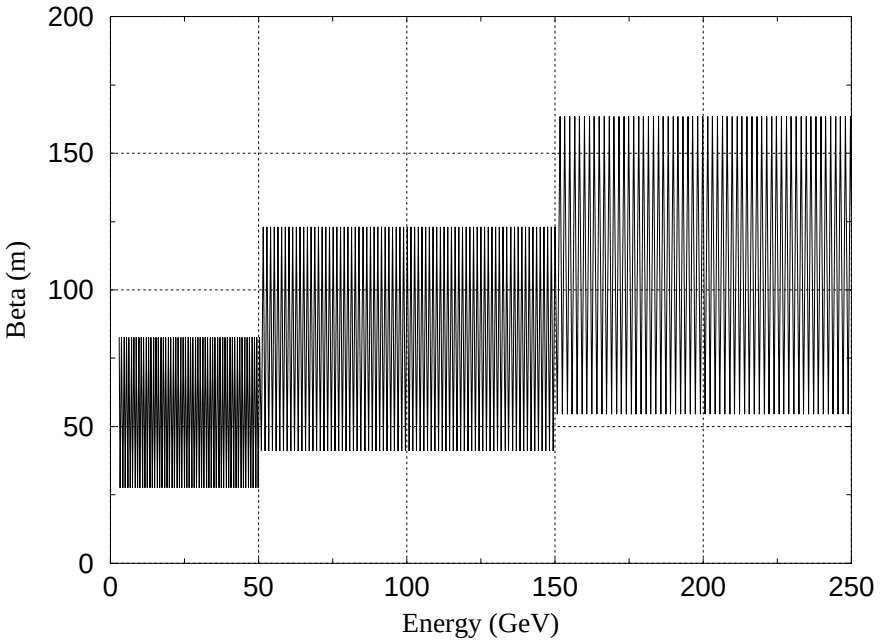


Figure 3.3.1: *Beta function along the TESLA 500 linac.*

located at the high energy part have to provide an integrated field gradient of about 17 T. For superconducting quadrupoles, allowing 1 T fields at the pole, the length of the magnet is still modest. With regard to the beam optics, being faced with the energy upgrade of the linac, the strategy is the following : the beta steps are kept to the same location and the same quadrupole arrangement is used. For TESLA 800, the linac is unchanged till the first beta step, coinciding with the extraction point of the FEL beam. From this 50 GeV point, the accelerating field rises to 40 MV/m and the second beta step occurs now at 210 GeV. For TESLA 1600, the beam travels down successively the two linacs and thus, the low energy part becomes the high energy part in one of the linacs. The first linac is unchanged, but in the second one, the quadrupole spacing of 4 modules is conserved, while the quadrupole spacing of 3 and 2 modules is enlarged to 6 modules. A schematic drawing of the beam optics for the three energies is represented on Fig.3.3.2. Table 3.3.1 sums up the maximal integrated field gradients, assuming a constant phase advance of 60 degrees.

	TESLA 500	TESLA 800	TESLA 1600
Linac energy (GeV)	250	400	800
Quad spacing (modules)	2-4	2-4	2-6
Max int. gradient (Tesla)	17	27	36

Table 3.3.1: *Maximum integrated gradient for the main linac upgrades.*

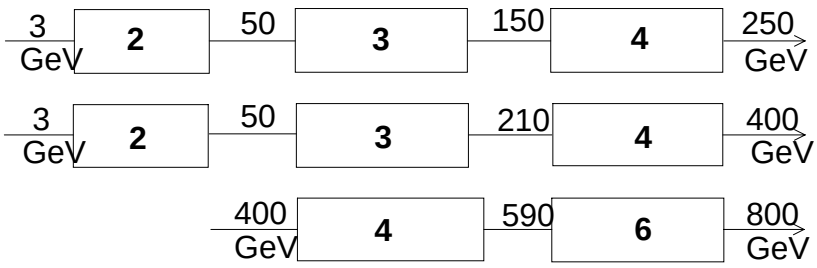


Figure 3.3.2: *Quadrupole spacing (in unit of cryo-modules) for the TESLA linac upgrades.*

3.3.3 Emittance Preservation

To achieve very flat and tiny beams at the collision point, beams of very small emittances, especially in the vertical plane, must be first generated by the injector and second accelerated by the main linac, while preserving the high quality beam to the final focus system. Since the beam pulse is composed of many bunches, single bunch and multibunch effects must be studied. Once the tolerable emittance growth is specified, all the tolerances on the various errors, like injection jitter and misalignment of the linac components, i.e. accelerating structures, focusing magnets and beam position monitors, can be in principle determined. Simple analytical expressions can help to estimate the individual tolerances. However, the net effect of simultaneous errors in a real machine is not the simple summation of the individual effects of the errors, therefore numerical simulations must be performed. The emittance growth was studied with the tracking code DILEM, which simultaneously combines single and multibunch effects of multiple errors and includes numerous correction techniques, like simple and elaborated orbit steering or wakefields compensation methods.

3.3.3.1 Single Bunch Effects

We consider first the short-range wakefield effects: As a bunch travels down accelerating structures, wakefields are excited and will act back on the bunch itself. The longitudinal wakefields affect the energy of the particles along the bunch, whereas for off-axis beam trajectories, the transverse wakefieds tend to deflect the beam further away from the axis. The deflection is stronger for the trailing particles, giving rise to a "banana shape" of the bunch and an increase of the transverse emittance.

Longitudinal single bunch dynamics

Though not easy for short bunches - bunch length smaller than 1 mm - passing through TESLA 9-cell cavities, the longitudinal wakes were carefully computed. Furthermore,

since the beam tubes, larger than the iris openings, which connect two adjacent cavities, form themselves also cavities, a chain of at least 5 cavities is necessary to get the convergence of wake computation for a 0.7 mm long bunch, as we can see in Fig. 3.3.3. In addition to the final longitudinal bunch potential, computed for a chain of 8 cavities and normalized for 1 cavity, the individual wakes, induced in each of subsequent cavities in the module, are plotted as a function of the axial position in the bunch (head of the bunch is on the right side). A bunch, travelling through the accelerator, will gain energy from the externally driven accelerating mode and will lose energy due to the longitudinal wakefields. The net energy spread can be reduced by running the bunch off the crest of the accelerating wave, such that the RF wave slope tends to cancel the average slope of the beam induced potential. As we will see thereafter, the BNS damping technique, which uses additional correlated energy spread to reduce the transverse wakefield effects on a bunch performing a betatron oscillation, is not needed in the TESLA linac. Minimal energy spread can then be achieved all along the linac by running the bunch at the optimal phase. For the parameters of TESLA 500, the optimal phase is 3.4 degrees, leading to a final correlated energy spread lower than 5.10^{-4}. Fig. 3.3.4 shows the energy variations due to the longitudinal wake, the RF wave and the sum of both contributions within the bunch. Bunch charge fluctuations are not critical: for a 10 % charge variation, the bunch energy spread still remains well below 10^{-3}. In addition to this correlated energy spread induced by the structures of the main linac, the energy spread of the injected beam, generated by the injector, and decreasing adiabatically with the energy, is taken into account in all the simulations, with an initial value of 1%.

Transverse single bunch dynamics

For beam dynamics studies in the transverse plane, we need the transverse delta wake excited by a point-like bunch, while cavity codes give only the transverse bunch wake of a displaced bunch. Fig. 3.3.5 shows the transverse impulse factor $k_\perp$, quantity proportional to the total transverse momentum given to the bunch by its own wake, computed for different bunchlengths. A fit of $k_\perp$ suggests a $\sigma^{1/2}$ asymptotic dependance. The corresponding delta wake per cavity, and used for single bunch instability calculations, is directly deduced from the fit of the transverse impulse factor,

$$W_\perp(s) = 1290\sqrt{s} - 2600\,s \qquad (\mathrm{V/pC/m})$$

where s is the position with respect to the exciting point-like charge. When a beam is injected off-axis, performing a betatron oscillation about the central orbit, the tail particles are driven on resonance by the wakefield excited by the head particles and the amplitude of the oscillation grows dramatically as the beam proceeds down the linac. This resonant effect leads to a tight tolerance on the amplitude of the betatron oscillation, unless BNS damping is used to cure this resonant effect. By making the head more energetic than the tail of the bunch, the transverse wake effect is then partially compensated by the chromatic phase advance of the trailing particles. However, the bunch energy spread is then larger, giving rise to a larger dispersive dilution, due to

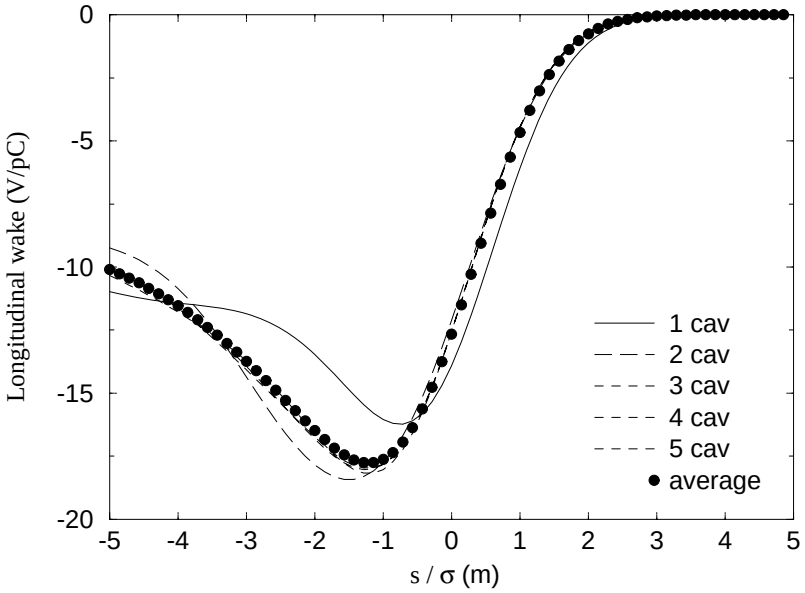


Figure 3.3.3: *Longitudinal bunch wake potentials induced in each of the subsequent TESLA cavities in a module (bunchlength σ = 0.7 mm)*

quadrupole misaligments. Thanks to the low wakefields of the TESLA cavities, BNS damping is not necessary and the energy spread can then be kept to its minimal value all along the linac. Fig. 3.3.6 shows the development of the emittance growth along the linac for the TESLA 500 lattice, assuming an initial beam offset equal to a full rms beam size and no BNS damping. We note that even for this large beam jitter, the final emittance growth is lower than 6%. The transverse wakefields driven by random misalignments of the cavities will deflect the beam from side to side around the central axis. The two particle model, describing the bunch by two macro particles, two bunchlengths spaced, gives an analytical expression of the final emittance growth at the linac exit

$$\frac{\Delta\varepsilon}{\varepsilon} \propto \frac{N^2 W_\perp^2}{\varepsilon}\,\beta_o \frac{1}{\alpha}\left[\left(\frac{\gamma_f}{\gamma_i}\right)^{\alpha} - 1\right] y_c^2$$

where y_c denotes the rms offset of the cavities, N the bunch population, $W_\perp$ is the transverse wake evaluated at 2 σ_z, β_o the beta function at the linac entrance and α the exponent of the beta scaling with energy, normally set between 0 and 1/2. We note that the more the beam is focussed (β_o and α small), the less the beam is diluted by the wakefields.

In order to avoid a continually growing trajectory in presence of quadrupole position errors, the beam must be steered all along the linac. The simplest trajectory correction is the 'one-to-one' correction, in which the dipole correctors steer the beam towards the quadrupole centers after reading of the BPM measurements. Since the actual orbit will then follow the BPM misalignments, significant dispersive errors will arise. Once again, a two particle model can be used to find an analytical estimation of the

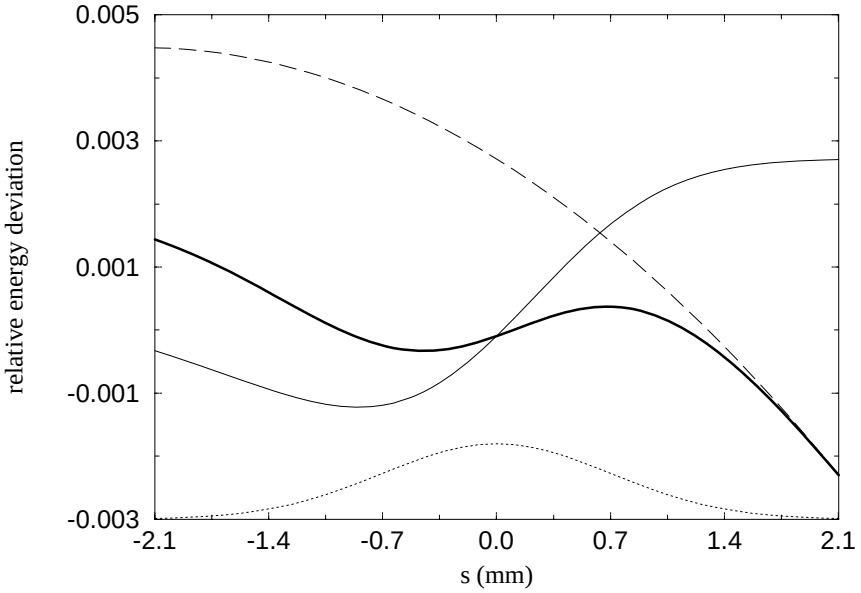


Figure 3.3.4: *Energy profiles within the bunch due to longitudinal wake (thin line), RF wave (dashed line) and both effects (thick line). The dotted line denotes the Gaussian bunch profile.*

emittance growth. For example, for a correlated energy spread, the emittance growth can be written as

$$\frac{\Delta\varepsilon}{\varepsilon} \propto \frac{\delta_c^2}{\varepsilon}\,\frac{1}{\beta_o^2}\,\frac{1}{1-\alpha}\left[\left(\frac{\gamma_f}{\gamma_i}\right)^{2-2\alpha} - 1\right] y^2$$

where δ_c is the relative rms correlated energy spread, y is the rms offset of the BPMs (including quadrupole offset and BPM offset relative to the quadrupole center). Contrary to the previous dilution driven by the cavity misalignments, the dispersive emittance dilution is lower for a less focussed beam (β_o and α large). An optimal focusing lattice, in which the dispersive effects and the wakefield effects are balanced, can then be found. Fig. 3.3.7 shows for example the vertical emittance growth, due to simultaneous chromatic and wakefields effects, driven by random quadrupole and cavity alignment errors, with different scalings of the beta function with energy. The rms quadrupole, BPM and cavity offsets are $\sigma_q = 100\ \mu m$, $\sigma_{bpm} = 100\ \mu m$ and $\sigma_c = 500\ \mu m$, respectively. We conclude that a small beta scaling, $\alpha = 0$ to 0.2, provides the smallest emittance dilutions. The mean emittance growth, computed with the chosen TESLA 500 lattice, the same misalignments and the simple "one-to-one" correction, is 12%. To alleviate the dispersive dilution, more elaborated correction algorithms have been proposed. All these methods refer to beam based alignment (correction) techniques, because the magnet misalignments (corrector settings) are deduced from BPM measurements with different beam trajectories. For example, the "shunt" technique consists in moving each quadrupole until no beam shift is observed at the next BPM or

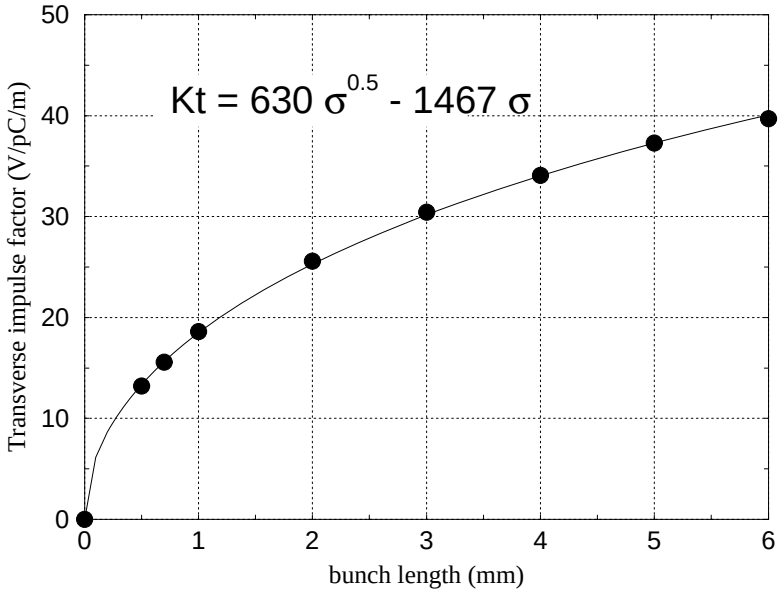


Figure 3.3.5: *Transverse impulse factor $k_\perp$ per cavity for different bunchlengths.*

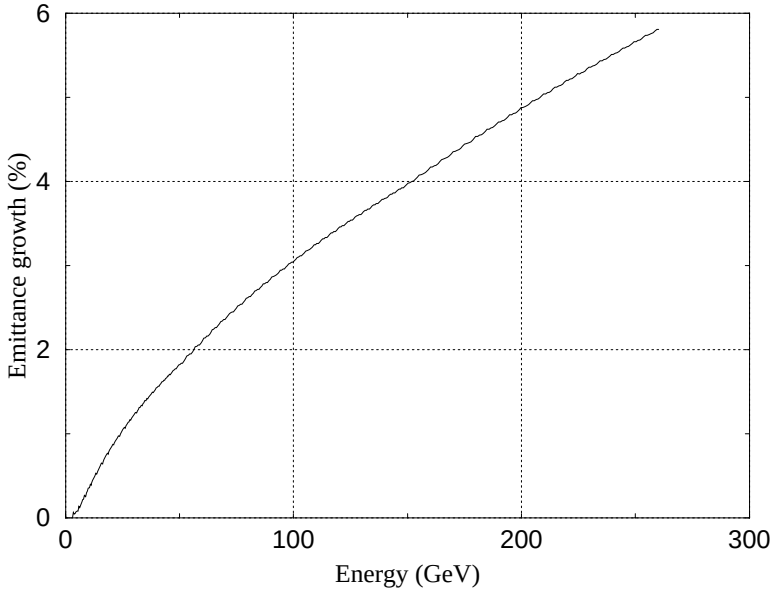


Figure 3.3.6: *Vertical emittance growth with a beam offset equal to the rms beam size.*

the 'Dispersion Free' correction, where at least two trajectories with effective different energies are used to cancel the dispersive errors. In this case, the final dispersive error is much less dependent of the magnitude of the misalignments, but is still limited by the BPM resolution, the reading-to-reading jitter of the BPM measurement. With the help of these beam based corrections, which could be applied after some machine

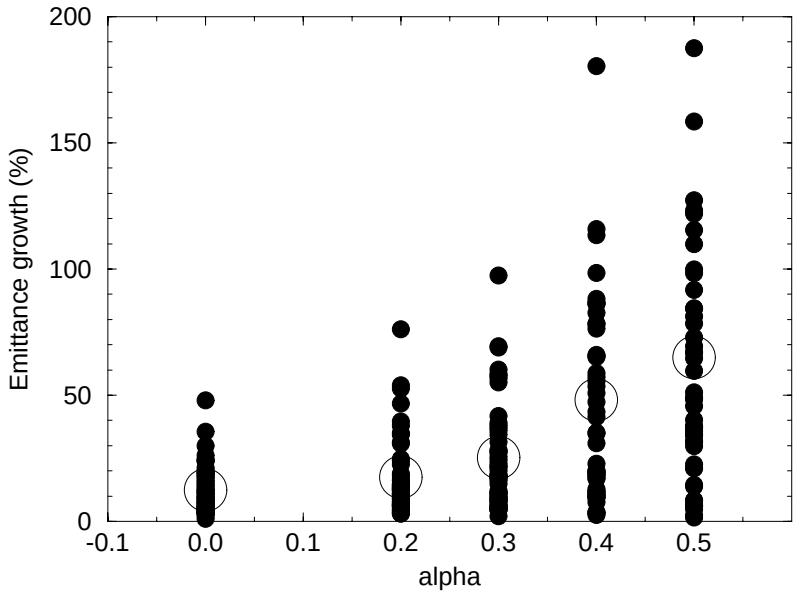


Figure 3.3.7: *Emittance growth vs. beta scaling parameter a with energy.*

experience, the tolerance on quadrupole misalignment can, of course, significantly be relaxed. For example, let now quadrupole alignment errors be equal to the cavity ones ($\sigma_q = \sigma_c = 500\ \mu m$), the tolerance on the BPM resolution is $\sigma_{res} = 15\ \mu m$ by using the "shunt" correction technique.

For correlated errors, the dispersive dilution depends on the correlation length of alignment errors. For example, Fig. 3.3.8 gives the quadrupole tolerance as a function of the wavelength, without wakefield effects and assuming an energy spread of 5.10^{-4}. We note that for small wavelengths, the tolerance of 100 μm, corresponding to random errors, is recovered whereas for the long wavelength components the tolerance is much looser. In addition to a position error, the cavity can be tilted with respect to the horizontal beam line and the accelerating field gives an additional transverse kick to the beam, with a strength which depends on the longitudinal position along the bunch. The tolerance on this cavity angle error has been estimated at 1 mrad from numerical simulations.

3.3.3.2 Multi Bunch Effects

The TESLA beam pulse is composed of many bunches, which will feel the long-range wakefields of all the preceding bunches and will be more and more deflected as they proceed down the linac. This phenomenon, called multibunch beam breakup, would require to place the cavities very precisely with respect to the beam axis. In addition, similar to the single bunch dispersive dilution, the bunch-to-bunch energy deviation will introduce a multibunch beam filamentation.

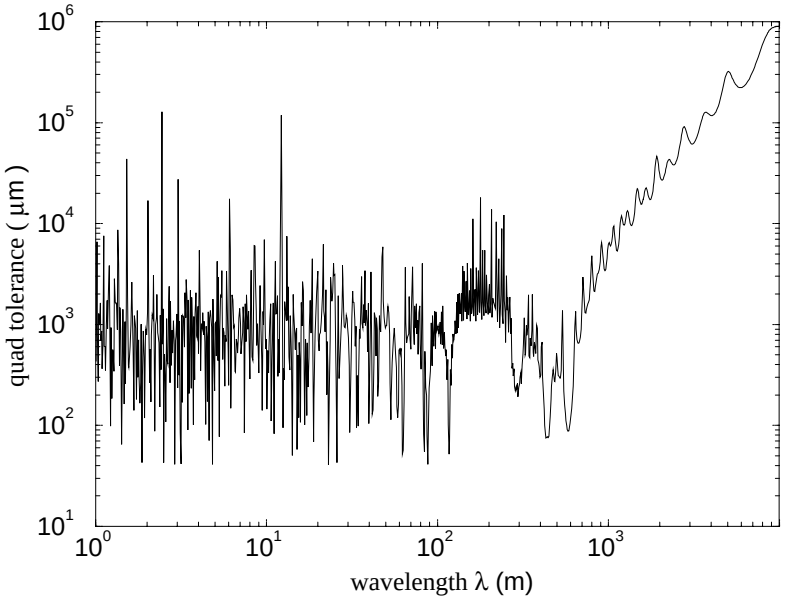


Figure 3.3.8: *Quadrupole alignment tolerance vs. wavelength for a 3% dispersive emittance growth with an energy spread of* 5.10^{-4}.

Longitudinal multi-bunch dynamics

The inter-bunch energy spread has to be smaller than the intra-bunch energy spread in order to preserve the transverse beam quality from chromatic effects. In particular, an 'optimal' trajectory (also called a 'gold' trajectory) for the first bunch of the train, giving the smallest single bunch emittance dilution, as discused above, will be lost for the following bunches if the bunch to bunch energy deviation is too large. This multibunch energy spread is mainly produced by the transient beam loading in the accelerating structures. Fortunately, the high beam loading will be compensated in the standing-wave TESLA cavities by matching the power extracted by the beam to the RF power supplied by the klystron. The residual energy gain variations along the beam pulse, caused mainly by Lorentz forces or microphonics detuning and RF mismatches, will be suppressed by the feedback system down to a few 10^{-4}. The extra power required for the field stabilization, of the order of 10%, could be strongly decreased with the help of harmonic cavities installed at a few locations along the linac. The systematic energy fluctuations would then be locally cancelled, in such a way that the beam filamentation has no time for growing.

On the other hand, the different energy kicks, imparted to the bunches of the train by the longitudinal higher order modes, cancel out because of the natural frequency spread of the modes, and are therefore harmless.

Transverse multi-bunch dynamics

With a train of bunches, the long-range wakefields excited by each bunch will act on the subsequent bunches. The major effect comes from the multibunch beam break-up (or cumulative BBU) : a beam passing off-axis through an accelerating structure excites transverse dipole modes, causing transverse deflections of the following bunches. This displaced beam will excite similar modes in the downstream structures, which will, in turn, deflect even more the following bunches of the train, and so on, resulting in an increase of the beam emittance. To suppress the cumulative BBU, the conjunction of two cures are used, namely the detuning and the damping techniques. Unlike long normal conducting sections with many cells, where the detuning usually takes place inside each structure, the deflecting modes frequencies change from cavity to cavity over the whole linac. The Gaussian detuning is naturally provided by the fabrication errors. The effective wake is then the average of the individual wakes over many structures, as long as the bunch displacement is small. For a Gaussian distribution and an infinite number of cavities, the wake would be reduced by the factor $\exp\left[-1/2(\sigma_\omega \tau)^2\right]$, where σ_ω and τ are the rms angular frequency spread and the time slipped away. In reality, the averaging is performed over a finite number of sections and the wake envelope does not vanishes, due to the recoherence effect. In order to reduce the effective wake for the rest of the long TESLA bunch train, a slight damping is necessary. The lower the quality factor Q, the less the excited field will persist over the bunchtrain. The quality factors of the modes are reduced by means of higher order modes dampers, which are mounted at both ends of the TESLA cavity. The parameters used for the multibunch calculations - frequency, geometric impedance and Q - of the ten most dangerous dipole modes are shown in Table 3.3.2. The assumed dampings, Q values of around or below 10^5 for the highest R/Q modes are in agreement with measurement on copper models. The effective dipole wake, resulting from the detuning and damping effects, is plotted

Frequency (GHz)	R/Q Ω/m^2	Q_{loaded}
1.874	$8.61\ 10^4$	$1.20\ 10^5$
1.865	$6.38\ 10^4$	$7.56\ 10^4$
1.880	$1.92\ 10^4$	$1.94\ 10^3$
1.736	$1.48\ 10^5$	$7.20\ 10^3$
1.887	$1.60\ 10^3$	$6.73\ 10^5$
1.709	$1.14\ 10^5$	$8.40\ 10^3$
1.837	$5.00\ 10^3$	$4.68\ 10^4$
1.884	$5.00\ 10^2$	$3.31\ 10^5$
1.853	$3.00\ 10^3$	$5.16\ 10^4$
1.763	$1.70\ 10^4$	$7.20\ 10^3$

Table 3.3.2: *Parameters of dipole cavity modes.*

on Fig. 3.3.9. The theoretical wake and the more realistic wake, when averaged on the finite number of cavities over a few betatron wavelengths, and assuming no damping at all, are also shown for comparison. Thanks to the modes detuning, the dipole wake decreases by a factor of 12 with respect to the peak value when the second bunch comes (0.7 ms later) and by a factor of about 90 for the other bunches. Though the mode damping does not provide a significant additional reduction factor on the second bunch, it is, on the other hand, very efficient for the end of the bunch train. In other words, thanks to the large bunch spacing and the natural mode detuning, a weak damping is sufficient for preserving the beam quality of the whole bunch train. The multi-bunch emittance growth for the whole train can be computed for different

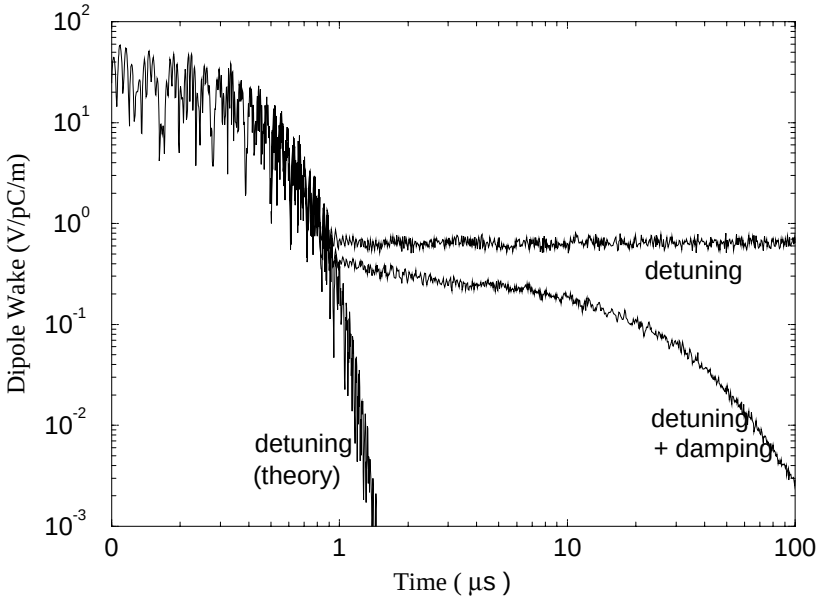


Figure 3.3.9: *Theoretical and real effective dipole wake, without and with damping.*

damping levels. It is shown on Fig. 3.3.10, as a function of the scaling factor by which the real Q of all the dipole modes of the TESLA cavity have been multiplied. For these calculations with the lattice of TESLA 500, an inter-bunch energy spread of $3.5 \cdot 10^{-4}$ is taken into account, the simple 'one-to-one' trajectory correction is applied and the quadrupole and cavity alignment tolerances, specified by single bunch dynamics (σ_q = 100 μm and σ_c = 500 μm) are chosen. The mean value of the emittance growth is only few percents for the real Qs (scaling factor = 1) of the cavity modes. This low multi-bunch dilution prevents any further dilution of the single bunch emittance. Calculations, including simultaneous single and multi-bunch effects, confirmed that the overall emittance dilution for the whole bunch train is very close to the single bunch emittance dilution. Once the multi-bunch beam break up has been well controlled, the beam energy has to remain as stable as possible to prevent any further filamentation.

In case of klystron trips, especially in the low energy part of the linac, the relative energy variation can be large (32 cavities are fed by one klystron) leading to a large beam orbit variation. However, thanks to the orbit feedback, the initial orbit can be rapidly recovered after a few beam pulses.

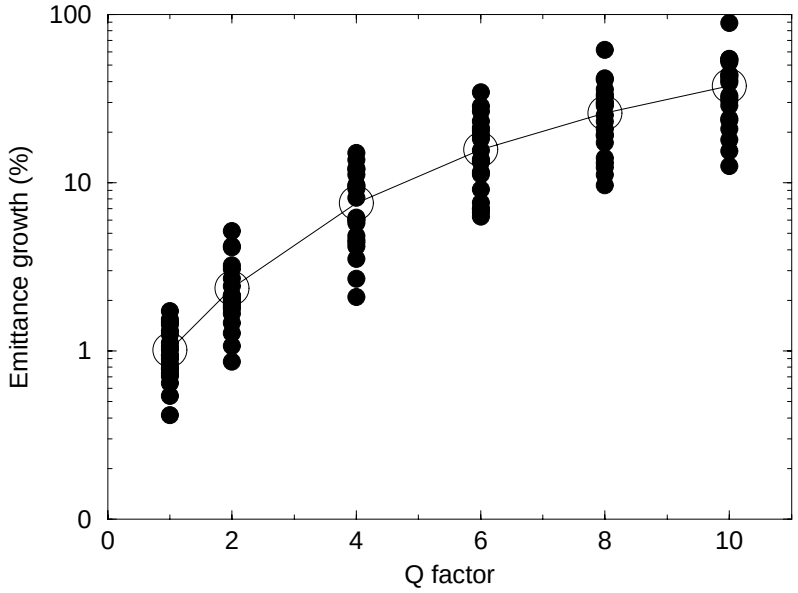


Figure 3.3.10: *Multi-bunch emittance growth as a function of the damping level (TESLA 500 lattice, σ_q = 100 μm and σ_c = 500 μm).*

3.3.4 Orbit Stability

Ground motion is of major concern because it will displace focusing magnets, which, in turn, will dilute the beam emittance in the linac through dispersive effects. Numerous power spectra of ground motion have been measured at different places. At low frequency, the motion can be quite large and is usually described by the 'ATL' law. Beam orbit feedback, involving dipole correctors and BPMs, can be used, at least up to a frequency of about $f_{rep}/20$, corresponding to 0.25 Hz for TESLA. At high frequency, the motion falls off rapidly, but is too fast to use beam orbit feedback with simple steering magnets.

The diffusive ground motion, leading to large displacements after long time intervals, can be described by the 'ATL' law, which states that the relative rms displacement Δy after a time T of two points separated by a distance L is given by

$$\Delta y^2_{rms} = A \cdot T \cdot L$$

The parameter A ranges from 10^{-7} to 10^{-5} μm^2/s/m, depending on the site. Beam-based alignment techniques will recover either the proper alignment of the elements or the 'gold' trajectory, this one which minimizes the dispersion. Nevertheless, these dispersion free corrections require measurements of the beam orbit with different quadrupole settings can only be used periodically, with some rather long time intervals. In between, correction feedback loops have then to steer the beam continuously to the original gold orbit such as it was previously measured by the BPMs. Fig. 3.3.11 shows the resulting emittance growth as a function of the product $A \cdot T$. We note that the diffusive ground motion produces a linear growth of the dispersive dilution with time. Limiting a maximal emittance growth to 15% and assuming a conservative value of the parameter $A = 10^{-5}$, the beam orbit must be re-steered about once per hour, a very comfortable value for the cut-off frequency of 0.25 Hz of the beam orbit feedback. With correctors steering periodically the beam to the original orbit, the quadrupole displacements are so large after a long time that a new gold trajectory must be found again by means of a beam base correction technique. Assuming again a tolerable emittance growth of 15%, the time interval between two such procedures, which necessary stops the physics experiments for a few hours, is several years. Fast vibrations, or jitter in quadrupole

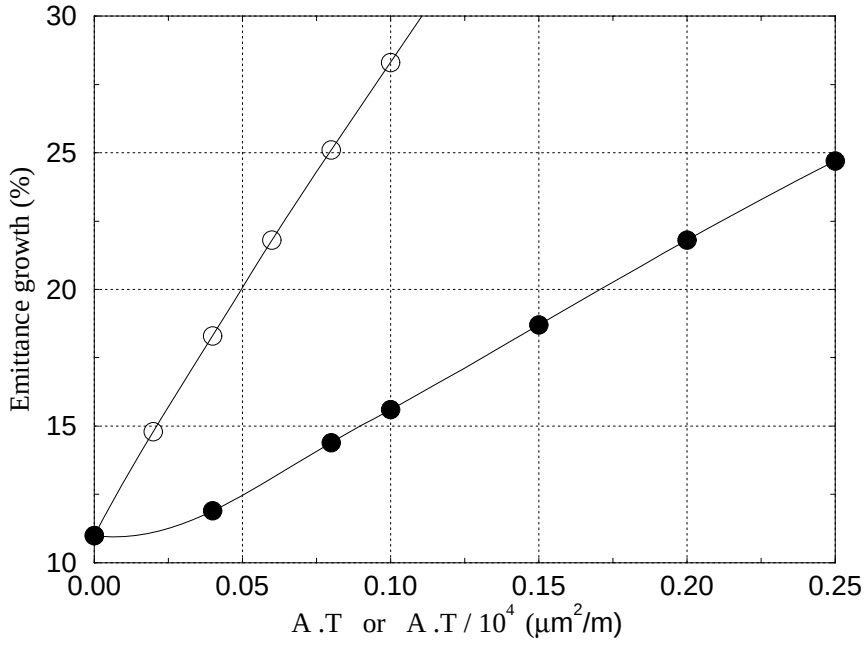


Figure 3.3.11: *Averaged emittance growth vs. time for diffusive ground motion without (white circles) and with orbit feedback (black circles).*

position, cannot be compensated by conventional dipole correctors and will scatter the bunch centroids at the end of the linac, thus preventing the bunches from colliding

at the interaction point. Allowing a maximum luminosity reduction of 3%, the rms amplitude of the fast quadrupoles motions we could tolerate can be written as

$$y_{qrms}^2 = \frac{\varepsilon_y}{16} \frac{\overline{\beta}_f^2 \, sin\mu \cos^2\mu/2}{4L_{\rm tot}} \, 2\,(1-\alpha)$$

where α is the exponent of the beta scaling with energy. With the lattice of TESLA 500, the jitter tolerance would be about 80 nm. Fortunately, since the bunches within the train have very large spacing, feedback can be applied on the bunch centroids within the train. The fast kickers at the end of the linac will re-align the bunches, after reading the displacement of the first several bunches. In addition to this bunch offset problem at the exit of the linac, the deviation of the bunch trajectories from the normal corrected orbit will induce beam filamentation, through dispersive effects. In this case, the tolerance on quadrupole jitter is about 1 μm, larger than the expected displacements at high frequency.

3.3.5 Summary Tolerances and Emittance Dilution

Beam dynamics studies allowed to find an optimized focusing lattice, for which the wakefield and dispersive effects are balanced, and to estimate the final emittance dilution for given tolerances. The initial alignment of the elements, with a tolerance on both cavities and quadrupoles of the order of 500 μm, can be accomplished by a conventional survey. In case of the quadrupoles and the BPM's w.r.t. the quadrupoles, the expected accuracy is actually smaller, about 0.1...0.2 mm. Later on, there is no need to move any quadrupole or cavity for re-alignment. Estimations of beam dilution with the simple 'one-to-one' orbit steering point out that the quadrupole and BPM offsets must then be located within a 100 μm range. Even if slightly exceeded at installation time, this tolerance can be easily achieved with the help of beam based correction procedures and was already met on the SLC at Stanford with BPM resolutions of about 15 μm. For example, numerical simulations, showed that the 'shunt' technique, which locally cancels the dispersion by varying the quadrupoles one-by-one, without requiring neither linac modelling nor matrix computations, restrict the bunch emittance growth to 10% with the same BPM resolution of 15 μm. The expected BPM-resolution is better than that and more advanced beam-based methods can be applied, so that it will be possible to achieve a smaller emittance dilution than required for the TESLA design parameters.

With regard to the energy upgrades of the main linac to 400 and 800 GeV (TESLA 800 and TESLA 1600), the emittance growths amount to about 20% for the same cavity and quadrupole tolerances.

Thanks to large bunch spacing and to strong reduction of the transverse wake, due to complementary effects of natural cavity-to-cavity detuning and of slight mode damping, the multibunch beam breakup is well controlled. Assuming an inter-bunch energy spread lower than $5 \cdot 10^{-4}$, the multibunch emittance growth is much smaller than the single bunch one. For that reason, tracking calculations involving single and multibunch effects give very similar results to those solely involving single bunch effects.

For slow ground motion, feedback loops adjust the beam orbit to the 'optimal' trajectory, previously found by beam based correction. According to the phenomenological 'ATL' law, predicting the relative displacement with time and position, the steering feedback has only to be faster than one hour. When the elements displacement becomes too large, beam based correction can be used again, every few years. For fast motion, fast deflecting correctors, located at the linac end, will be used to re-align the bunch centroids. The tolerance on quadrupole jitter is then simply determined by the dispersive errors arising through the linac and is of the order of one micron.

Bibliography

[1] V. Balakin, A. Novokhatski, V. Smirnov, *VLEPP: Transverse Beam Dynamics*, Proc. 12^{th} Int. Conf. on High Energy Accelerators, Fermilab, 119 (1983).

[2] T. Raubenheimer, *The Generation and Acceleration of Low Emittance Flat Beams for Future Linear Colliders*, SLAC-387 (1991).

[3] B. Baklakov, P. Lebedev, V. Parkhomchuk, A. Sery, A. Sleptsov, V. Shiltsev, *Investigation of Seismic Vibrations and Relative Displacement of Linear Collider VLEPP Elements*, Proc. 1991 Part. Accel. Conf., San Francisco, 3273 (1993).

[4] A. Mosnier, *Instabilities in Linacs*, CERN Accel. School, Rhodes, CERN 95-06, 459 (1993).

[5] A. Novokhatski, A. Mosnier, *Short Bunch Wake Potentials for a Chain of TESLA cavities*, Int. Report, CEA Saclay, DAPNIA/SEA-9608 (1996).

[6] A. Mosnier, *Beam Instabilities Related to Different Focusing Schemes in TESLA*, Proc. 1993 Part. Accel. Conf., Washington, 629 (1993).

[7] A. Mosnier, *Single and Multi-bunch Wakefields Effects in the TESLA Linac*, Proc. 4^{th} European Part. Accel. Conf., London, 1111 (1994).

3.4 Injection System

3.4.1 Introduction

The injection system must provide the sources of electrons and positrons for the two linac beams. The low energy beams must be accelerated to 3.2 GeV where, if their emittance is not sufficiently small, they must be sent to damping rings for further emittance reduction in the x- and y-planes by synchrotron radiation damping. After the damping rings the beams proceed through bunch compressors and are injected into the main linacs.

For the positron source it is planned to use the "spent" high energy electron beam after it passes through the collision point in the interaction region. This beam is passed through a wiggler and the photons produced impinge on a rotating disc target where the resulting electromagnetic shower yields the required positrons. The positrons are focused in a solenoid and pre-accelerated to about 200 MeV in a normal conducting RF-structure. After that the beam is injected into a superconducting 3 GeV linac similar to the one used at the TTF. Since the damping ring is layed out for a normalized transverse acceptance of 0.05 m and and energy acceptance of $\pm 0.5\,\%$, the pre-linac must provide transport and acceleration of beams of at least this admittance.

The wiggler photon production scheme potentially can be modified to provide polarized positrons by using a helical undulator. This scheme is considered as an upgrade option.

The electron source is much more straightforward than what is needed for positrons. Even so, at this juncture there are several options and technical challenges to address. These hinge on the ability of an RF-gun to produce (a) sufficiently small and asymmetric ($\epsilon_y \ll \epsilon_x$) emittances, (b) polarized beams. The first issue can be resolved by having a damping ring for the electrons as well as the positrons. Concerning the second issue, if polarized beams are not possible with an RF-gun, then DC-guns are a reasonable backup. The following discussions will amplify these issues.

Once the electron beam has been produced and accelerated to a few MeV, it can be further accelerated in a 3.2 GeV injector linac similar to the TTF superconducting linac and either sent to the damping ring or passed directly through the bunch compressor to the main e^- linac.

3.4.2 Unpolarized Electron Source

The recent developments and efforts on laser driven RF guns make them attractive to consider for the electron source. The TTF Injector II can be used as a model for this type of injector. The injector II layout consists of the gun system followed by a single Tesla cavity and a magnetic bunch compressor (chicane). Various diagnostics and optical elements for beam analysis and matching into the TTF linac follow the chicane. The gun system consists of a laser, a normal conducting 1 1/2 cell L band gun cavity, and a focusing solenoid system. The present injector being developed for TTF uses a symmetric gun and should provide normalized emittances of $20 - 30 \cdot 10^{-6}$ m

and bunch lengths of about 1 mm for bunch charges of 8 nC ($5 \cdot 10^{10}$ particles). The momentum spread needed for the bunch compression is about 250 keV. The bunch train consists of 800 bunches spaced at 1μs intervals for a macropulse length of 0.8 ms. The design has also been looked at for 1 nC and is expected to provide $< 5 \cdot 10^{-6}$ m emittance.

Laser systems being developed for the injector must provide stable intensity and phase light pulse trains as this determines the injection phase and charge variations of the bunches in the linac. The desired light pulse is flat temporally with 10-30 psec full width. Spatially, for the symmetric gun, a flat disk of 1.5-3 mm radius is required.

Two different laser systems under development use Nd:YLF flash lamp pumped lasers at 1054 nm with fourth harmonic generation to 263 nm for the photocathode illumination. The energy delivered to the photocathode per bunch is 5-10 μJ. The bunch to bunch energy stability is expected to be better than 10%.

The gun uses a photocathode of Cs_2Te. This cathode has a relatively long operating lifetime (months) at high quantum efficiency. Efficiencies as high as 3-5% are typical. Therefore for 10 nC out of the gun, at 1% efficiency and 4.7 eV photon energy the light pulse energy required per bunch is 5 μJ.

The gun photocathode is prepared (Cs_2Te coated on a molybdenum cathode plug) in a separate chamber and a number of cathodes can be interchanged without breaking vacuum of the gun system. An operating vacuum of 10^{-10} mbar is required.

The symmetric gun under development for TTF is a copper one and one half cell disk structure with a waveguide input on the full cell. Space charge emittance compensation is implemented with focusing solenoid lenses surrounding the gun. A field strength of 1.2 T is used. The peak RF field gradient is 35 MV/m for design operation. This is approximately the field on the cathode as well. The cavity iris radius is 20 mm, the shunt impedance Z = 28 MΩ/m, and RF peak power 2.2 MW. The beam energy out of the gun is about 4 MeV. It is hoped to be able to operate the gun at gradients as high as 50 MV/m which would require 4.5 MW RF peak power and provide better emittance quality.

The gun is followed by a cryostat with one 9-cell cavity to increase the energy to 15-20 MeV. Downstream of this cavity a magnetic chicane bunch compressor decreases the bunch length by about a factor of two to σ_z=1 mm. This injector system is designed to allow measurements of HOM losses and wakefields in the TTF cavities.

The superconducting pre-linac would look basically like the TTF linac. For an energy of 3.2 GeV, 120 9-cell cavities (15 modules) are required and the total length of the pre-linac is about 180 m.

3.4.3 Polarized Electron Source

At present, all operating polarized electron sources are using DC-guns generating long pulses in the ns range, usually followed by a bunching system (see e.g. [1]). A GaAs cathode is illuminated by suitable laser pulses to extract pulses of polarized electrons.

RF gun An attractive source of electrons here is the use of an RF gun rather then a dc gun. Rf guns have produced short bunches with high charge similar to the bunches required for TESLA. At the TTF [2] an injector for unpolarized electrons is under construction, which uses an L-band RF gun. So far, GaAs was not used in RF guns. One reason was the high sensitivity of the activated GaAs to contaminations [3]. An ultra-high vacuum of better than $1{\cdot}10^{-11}$ mbar must be maintained, which is not the case for existing RF guns. Suggestions have been made to overcome this problem leading to a first proposal using RF guns as a source for polarized electrons [4]. A successful attempt has recently been made to test a GaAs crystal in an RF gun [5]. Other potential problems when using GaAs as a cathode material have to be considered as well (see also [5]): the space charge and current density limit, multi bunch effects, the response time, the quantum efficiency and the laser system.

Space charge limit The space charge limit for short bunches can be estimated by Gauss' law for infinite separated electrodes: $\sigma = \epsilon_0 E$ with σ being the charge density on the cathode. For a reasonable extraction field E of 40 MV/m the extractable charge density is limited to 35 nC/cm^2, which is well above the requirements for Tesla.

Current density limit A more severe limitation specific to GaAs-type cathodes is the so called "cathode charge limit" [6], which is in the case of short excitation pulses practically a current density limit [5]. This limitation is typical for semiconductor cathodes with a negative electron affinity (NEA) surface, where a high electron extraction probability is achieved by the application of alkalis and oxides to the surface of the GaAs crystal. The NEA surface (together with the p-dopant, which bends the energy bands downwards at the surface) lowers the vacuum level below the conduction band minimum in the bulk, resulting in high extraction rates. The limit can be understood as a competition between the rate of near-thermal electrons arriving from the conduction band at the surface and the discharge rate of electrons from the surface [7].

Since the current density limit would increase with the discharge rate from the surface, it is preferable to use the highest possible extraction field. Also a larger cathode surface and longer pulses will lower the current density.

Because the maximum acceleration field in an RF gun varies linearly with the RF frequency, an S-band gun would be preferred. But on the other hand, the dimensions of an RF gun scale inversely with the RF frequency. Therefore, the maximum usable cathode area varies quadratically with the RF wavelength favoring L-band guns.

The current density limit for the Tesla case can be estimated using SLC data [8]: the SLC gun produces about 10 nC in a 2 ns long pulse from a cathode of an area of 1.5 cm^2 close to the cathode charge limit. A 100 nm medium-doped ($5{\cdot}10^{18}$/cm^3) strained-lattice GaAs cathode layer is used. The acceleration voltage is 1.8 MV/m. Thus, the current density is limited to 3.3 A/cm^2. Since the current density scales linearly with the acceleration voltage E [9], it is convenient to express the current density limit for the Tesla case (with bunch charge of 6 nC) by the following expression

$$E \ A \ \sigma_t \geq 3.3\mathrm{A/cm}^2 \tag{3.4.1}$$

with E in MV/m, the cathode surface A in cm^2 and the FWHM pulse length σ_t in ns. Table 3.4.1 shows examples of combinations of E, A and σ_t for which the current density is at the limit of 3.3 A/cm^2, as measured at the SLC gun. For an L-band gun, several parameter sets fulfill the condition 3.4.1. A reasonable choice is an acceleration field of 40 MV/m and a pulse length of 26 ps with a cathode radius of 1 cm, which is not too large. The radius of the iris of an L-band gun is around 2 cm. It is not completely impossible to find a parameter set for an S-band gun, but one has to choose all the parameters at their limits. An L-band gun is clearly the prefered choice.

E [MV/m]	A [cm^2]	σ_t [ps]	remarks
40	3.1	26	reasonable L-band choice
40	13	6	largest cathode L-band
40	2.1	40	longest pulse length L-band
50	3.1	21	highest field L-band
70	2.5	20	S-band: all parameters at limit

Table 3.4.1: *Shown are examples of parameter sets, for which the current density is at the limit of 3.3 A/cm^2 as measured at the SLC gun. E is the accelerating voltage, A the cathode surface, and σ_t the pulse length (FWHM).*

Multibunch operation For multibunch operation, there is an additional effect of the cathode charge limit: the temporary flattening of the bands in the band bending region due to trapped electrons in the surface states, preventing further electrons from being extracted. Experiments at the SLC have shown, that the recovery time of the cathode is in the order of 10 to 100 ns depending on the cathode thickness [1]. This is much less than the Tesla bunch spacing of 708 ns.

Response time The response time of NEA cathodes is usually large. The theoretical model described in [7], which is consistent with SLC data, gives an estimation of 20 ps to 7 ns. Experiments indicate response times as low as 10 ps for cathodes with a thin layer of 150 nm [10].

Polarization, quantum efficiency and lifetime A polarized electron beam is produced by applying circular polarized laser light on a strained lattice GaAs cathode. SLC data show a polarization maximum of more than 80 % at a laser wavelength of 845±10 nm [11]. Unfortunately, the quantum efficiency drops sharply for highest polarization to 0.1 ... 0.3 %. In addition, the quantum efficiency depends strongly on the operating conditions of the gun. High dark current of more than 50 nA [1] and poor vacuum conditions reduces the quantum efficiency significantly. In addition, activation of the cathode by cesiation is regularly required. To obtain a reasonable up time of

the source, a loading system which enables loading and removal of the cathode without breaking the gun vacuum is essential. Conditioning of the gun, regular cesiation and other cathode manipulations will be done with the cathode removed. A similar system will be tested for the TTF RF gun.

The SLC gun is operated with the same cathode for more than a year with a quantum efficiency between 0.1 and 0.3 % at highest polarization. The regular cesiation required to maintain this efficiency is done in situ without breaking the vacuum [11].

Laser Most polarized electron sources use lasers based on Ti:Sapphire, which are commercially available, especially for cw and pulsed applications with low repetition rates of a few tens of Hz. These lasers are tunable in a wide range around 800 nm and exceed easily the power or pulse energy requirements for most sources.

In our case, a long train of 0.8 ms length of 1130 bunches has to be generated with a repetition frequency of 5 Hz. This is very unusual for common lasers and requires a special design, which is not yet commercially available. For the TTF RF gun, such a laser system is under construction [2]. For the unpolarized source at TTF the requirements on the laser wavelength are different. Therefore, this laser will provide a test of the generation of the bunch train, but can not directly be used for a polarized source [12].

For the generation of long pulse trains, a laser material with good thermal properties like low thermal lensing is preferred. The strong thermal lensing of Ti:Sapphire would lead to a variation of the focusing of the laser beam during the pulse train resulting in an undesired non-uniformity. Other laser materials, which are tunable in the required wavelength range are Cr:LiCAF and Cr:LiSAF. Their advantage is the low thermal lensing and the possibility to be pumped by flash lamps and laser diodes [13]. Mode locked Cr:LiSAF lasers and regenerative amplifiers have been successfully built [14]. In the proposal for the TTF laser system, a laser based on Cr:LiSAF/Cr:LiCAF, which would fulfill the energy and bunch structure requirements, has already been discussed [15].

Emittance The emittance can be estimated using the formulae given in [16]. The thermal emittance scales linearly with the cathode radius. For a cathode radius of 1 cm one gets $\epsilon_{\mathrm{th}} = 5\,\pi$mm mrad. The RF contribution to the emittance growth is larger, it scales quadratically with the cathode radius: $\epsilon_{\mathrm{rf}} = 10\,\pi$mm mrad. Assuming that emittance dilution due to space charge can be largely compensated, the total emittance is expected to be below $20\,\pi$mm mrad.

DC gun Studies have been made for a design of a high charge dc gun which meets the Tesla requirements in terms of bunch charge and train structure [17]. The proposed injector consists of a 500 kV electron gun followed by a bunching system. The emittance achieved was in the order of 200 to $300\,\pi$mm mrad, which is well in the acceptance of the damping ring.

The previous discussion concerning the use of a GaAs cathode in such a gun, namely

the charge limit, multi bunch operation, polarization, quantum efficiency, and laser is similar for the RF and dc gun option.

The main difference to an RF gun is that the acceleration field is much lower in a dc or pulsed dc gun. This limits the extractable beam current such that for the required charge of 6 nC the bunch length has to be in the ns scale. In addition, due to high space charge forces, the emittance is very large.

Let us consider the proposed 500 kV gun and a gun similar to the inverted-geometry gun proposed for SLC [18] which operates at 200 kV.

The current limit of a dc gun due to space charge can be estimated using Child-Langmuir's law. Taking a typical perveance of $0.3\,\mu\mathrm{A/V^{3/2}}$, the current is limited to 100 A for a 500 kV gun and 27 A for a 200 kV gun.

Scaling the cathode charge limit seen by the SLC gun to a field at the cathode of 10 MV/m for the 500 kV gun and 3 MV/m for the 200 kV option, the extractable current is then limited to 18 A/cm^2 and 5.5 A/cm^2, respectively. Taking a large cathode area of 3 cm^2, the current is limited to 54 A and 16 A resp, which is lower than the limit due to space charge.

The required bunch charge for Tesla is 6 nC. Allowing a charge overhead of a factor of 2, the bunches must be longer than 250 ps and 750 ps, respectively.

A shorter bunch length is preferred, since it would simplify the bunching system. On the other hand, to avoid damaging of the cathode by ion bombardment, the dark current should be kept as low as possible. Besides a careful processing of the gun this requires low fields at the electrodes. The SLC gun operates at 120 kV with a dark current of less than 50 nA. The highest field on the electrodes is 7 MV/m [1]. The 500 kV gun designed for the TJNAF-FEL uses a GaAs cathode and will operate with fields of about 12 MV/m at the electrode surfaces [19].

As mentioned above, the gun vacuum has to be better than 10^{-11} mbar to allow a long cathode lifetime and thus a reliable operation of the gun. Processing and conditioning of the gun must be done without the cathode in place. A loading system will be attached to the gun, which allows loading and removal of the cathode without breaking the vacuum. This is also necessary for regular cesiations which will be performed at the removed position. A transport chamber will allow to insert a cathode prepared in a separate preparation chamber.

3.4.4 Flat Beam RF Photocathode Electron Source

The expected performance of cylindrically symmetric RF photoinjector systems is well understood from theoretical analysis and simulations which have been benchmarked by experiment. An extension to a fully three-dimensional approach has recently been undertaken by investigators at UCLA and Fermilab. This effort has been motivated by the demands of the TESLA linear collider design, in which the normalized emittances are $\epsilon_{x,y} = 14,\ 0.25\,\mathrm{mm\,mrad}$ at Q = 5.8 nC. In this machine it may be possible to have asymmetric emittances at the injection point of an electron linac which meet the constraints set by the interaction point. In this way, the electron damping ring may be unnecessary, and in the case of $e-\gamma$ or $\gamma-\gamma$ collider options, no damping rings at

all are needed. While the emittances may not be low enough to eliminate the need for an electron damping ring in the TESLA design under consideration, benefits in design and operation of the damping ring may be derived from injection of lower, asymmetric emittance electron beams.

To simply state the approach to design of an asymmetric RF photoinjector source [20], one uses the fact that the temperature of the emitted beam particles from a photocathode is only a function of cathode material and laser photon energy. This allows creation of asymmetric emittances $\epsilon_y \ll \epsilon_x$ by illuminating the photocathode with a ribbon laser pulse ($\sigma_y \ll \sigma_x$). Typically, the dominant emittance growth mechanism at relevant current densities is space-charge. This is conveniently mitigated by the ribbon geometry of the electron beam pulse. One must also pay attention to the potential for RF induced emittance growth, especially in the horizontal dimension where the beam is large. In the spirit of the problem for cylindrically symmetric systems performed by K.J. Kim [21], we have examined the scaling of emittances with respect to beam dimensions, accelerating gradient and phase, and beam charge. The "thermal" contribution to the emittance given by the finite spread in transverse momenta of photoemitted electrons is, assuming 0.1 eV emission temperature (typical of semiconductors)

$$\epsilon_{x,y}^{th}[mm \cdot mrad] \approx \sqrt{\frac{kT_\perp}{m_e c^2}} \cdot 0.25\sigma_{x,y}[mm] \tag{3.4.2}$$

For TESLA this implies beam sizes of $\sigma_x < 5\,\mathrm{cm}$, and $\sigma_y < 1\,\mathrm{mm}$.

Under the assumptions of a cylindrically symmetric RF cavity structure, it is easy to derive the minimum RF contribution to the rms emittances which arise from the time dependent transverse focusing forces in the RF structure; they can be simply written as

$$\epsilon_{x,y}^{rf} \approx \frac{eE_{rf}}{\sqrt{8}m_e c^2} \cdot \sigma_{x,y}^2 (\Delta\Phi)^2, \tag{3.4.3}$$

where $\Delta\Phi = (\omega_{rf}/c)\sigma_z$ is the phase extent of the beam. Minimizing these quantities implies operation at low frequency, low accelerating gradient, and with small bunch sizes in all dimensions. For our case, we are considering a fairly large horizontal beam size as well as high gradient, and so this is an issue for the horizontal emittance.

There are a number of proposed ways to remove or minimize the emittance diluting phase dependent defocusing due to the transverse RF forces in the gun. Because of the use of emittance compensation to deal with the space-charge contribution to the emittance, it is perhaps wisest to use a scheme which mitigates the transverse kick in the x-direction. This can be accomplished by removing as best as possible the dependence of the RF fields on x. It should be noted in this regard that the defocusing kick induced at the end of a cylindrically symmetric cavity has approximately equal electric and magnetic parts. The electric component of the kick is mainly due to the fringing fields near the iris at the exit of the gun. The fringe field can be made mainly vertical simply by making the iris opening a slit, long in the horizontal dimension. The magnetic component, however, cannot be so easily diminished, as it is not dependent on the iris, but the accelerating field in the interior of the cavity. The only way to diminish the magnetic force in the horizontal direction is to break the symmetry of

the cavity outer wall. The simplest asymmetric structure is a rectangular box cavity, but this only diminishes (does not eliminate) the horizontal RF forces, and produces a sinusoidal dependence of the accelerating field on the x-dimension. A better choice is to use an "H-shaped" structure, which has been investigated in the context of sheet-beam klystron development at SLAC. In this structure, the accelerating field in the "bar" region of the H-cavity has no dependence on x and therefore has acceleration independent of transverse position, as in a symmetric structure. It should be noted that the vertical forces (which are due mainly to the backwards speed-of-light space harmonic in this p-mode structure) in this case in the bar region are now entirely in the vertical dimension, and are twice as large as the equivalent symmetric cavity. The function of the side regions is to allow the longitudinal field to go zero sinusoidally, to satisfy the boundary condition, and choose (along with the vertical dimension) the cavity resonant frequency. An example of this structure, along with the fields that it supports are shown in Fig. 3.4.1, which displays the results of a 3-D electromagnetic simulation for a 2856 MHz cavity under development at UCLA, using the code GDFIDL [22].

An analysis of the space-charge contribution to the emittance gives the following scaling, derived from both approximate analysis and PARMELA simulation [20]:

$$\epsilon_{x,y}^{sc} \approx \frac{2N_b r_e}{7\sigma_{x,y}W} \exp(-3\sqrt{W\sigma_y})\sqrt{\sigma_y/\sigma_z}, \qquad W = \frac{eE_{rf}\sin(\Phi_0)}{2m_e c^2}. \tag{3.4.4}$$

Here, the scaling $\epsilon_x^{sc} = (\sigma_x/\sigma_y)\epsilon_y^{sc}$ is approximately obeyed; this is a consequence of the fact that for an ellipsoidal charge distribution, the normal electric field at any bunch boundary is approximately the same. It should be noted that minimizing the space charge contributions to the emittances implies operation at high accelerating gradient. It should also be emphasized at this point that these emittances, induced by the space charge at low energy, are mainly due to the different orientations of each "z-slice" of the beam. A scheme is discussed below, emittance compensation, which can remove these correlations, effectively lowering the emittances of the bunch. Before taking up this discussion, it should also be pointed out that one cannot arbitrarily reduce σ_y, because of the limit on surface charge density emitted from the cathode before the retarding space-charge cuts off emission ($\sigma_x\sigma_y > 2Q/E_z$). This effect has been observed at UCLA.

It is instructive to look at the product of the emittances, to see if there is some chance of achieving the desired emittances without need for emittance compensation,

$$\epsilon_x^{sc}\epsilon_y^{sc} \approx \left[\frac{2N_b r_e}{7W}\right]^2 \cdot \frac{\exp(-6\sqrt{W\sigma_y})}{\sigma_x\sigma_z}. \tag{3.4.5}$$

The argument in the exponential in this expression is generally smaller than one for reasonable accelerating gradients, and so one can see that in general a beam large in the longitudinal and horizontal dimensions is needed to minimize the product of the space charge emittances. Unfortunately, these are precisely the dimensions that the RF contributions are sensitive to, and thus one cannot consider making them arbitrarily

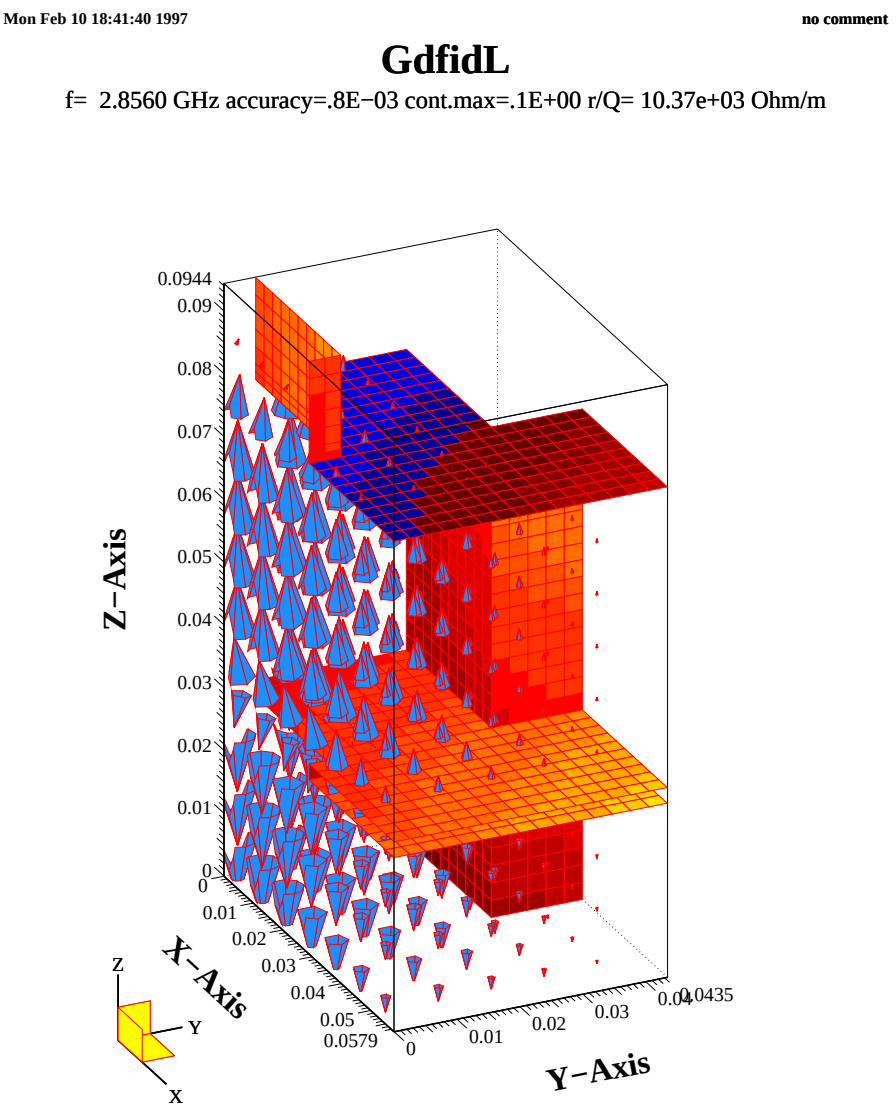


Figure 3.4.1: *Three-dimensional electromagnetic simulation of S-band, H-shaped asymmetric RF gun structure, from program GDFIDL.*

large. As an example of a design where the RF contribution is not yet significant, we take $N_b = 3.6 \cdot 10^{10}, W = 80m^{-1}, \sigma_z = 2mm, \sigma_x = 2.5cm, \sigma_y = 1mm$. In this case, we have $\epsilon_x \epsilon_y \approx 600(mm \cdot mrad)^2$ as opposed to about $3(mm \cdot mrad)^2$ as needed by TESLA, and we miss the design goal by an order of magnitude in both dimensions.

It is apparent that further efforts must be made in order to achieve the emittances

needed for linear collider designs, that is, one must employ an emittance compensation scheme for sheet beams accelerated in asymmetric RF structures. Emittance compensation is essentially a process by which the correlations (mainly as a function of the longitudinal position in the bunch) in the beam transverse phase space that are induced by space charge forces in the photocathode gun, are removed by focusing the beam, and allowing the correlations to be removed after one surface-plasma oscillation of the space-charge dominated beam. When the orientation of the transverse phase space ellipse of each "z-slice" has the same angle, the linear component of the emittance due to this effect is removed. Since the space-charge forces must still be large enough after the focusing lens to allow for compensation, this scheme tends to work better at lower accelerating gradients for convenient focusing geometries. This fortunate direction in the design of the system allows us to mitigate concerns about the RF contribution to the emittance, and about peak and average RF power (source and heat dissipation problems), in both the symmetric and asymmetric cases. In the asymmetric case, we must have the RF contribution to the transverse phase space trajectories be small, otherwise the space charge compensation will suffer from interference. This condition is achieved by the choice of large aspect ratio (sheet) beam profile, and the asymmetric structure described above.

While the serious design calculations must be performed with simulation programs which include as many experimental effects as possible, these are very time consuming. One needs to have a model which predicts the approximate behavior of the system before proceeding to simulation. This has been developed recently as an extension of the analytical understanding of the emittance compensation process [23]. This process proceeds by first focusing and then accelerating, in order not to "freeze out" emittance oscillations, as previously conjectured, but to drive down the emittance in a secularly diminishing fashion. This behavior is obtained when one operates on the invariant envelopes, which are analogous to the Brillouin flow obtained in a non-accelerating system. In the case of asymmetric beams, one must break the symmetry of the external focus in order to suppress a quadrupole oscillation which tends to immediately symmetrize the x- and y-beam envelopes. This can be done by using a vertically focusing quadrupole field super-imposed on the final acceleration section (solenoidal focusing is not allowed in this device, because of a coupling of the x and y phase planes due to an $\vec{E} \times \vec{B}$ rotation which is dependent on longitudinal position in the beam), such that the quadrupole strength is chosen as

$$K_x = -\frac{1}{4}\left(\frac{\gamma'}{\gamma}\right), \tag{3.4.6}$$

where γ' is the average normalized accelerating gradient in the linac.

In order to illustrate this process, we show the results of a calcuation which integrates the envelope equations for a number of different z-slices, at positions z_i. The transverse focusing forces arise from external quadrupole focusing, and from the first order transient kicks experienced by electrons at the exit and the entrance of the RF cavities, and the second order (alternating gradient focusing) in the interior of the RF cavities. Emittance terms included in the envelope calculations are only the thermal

components.

The results of a typical calculation are shown in Fig. 3.4.2. In this case, the vertical emittance is well compensated, while there is essentially no improvement in the horizontal emittance. This is because the effective defocusing strength associated with the space charge is different in the two dimensions, as can be seen from the envelope equation,

$$K_{x,y}^{sc} = \frac{4I(z_i)}{I_0\gamma^3(\sigma_x + \sigma_y)\sigma_{x,y}}. \tag{3.4.7}$$

Because of this, the compensation process generally proceeds much faster in the small dimension, and thus it is difficult to design a system which simultaneously compensates in the vertical and horizontal dimensions.

The solution to this problem is straightforward: since the beam near the cathode is much larger in x than in y or z, one can effectively remove the dynamics in the x direction (and the dependence on x of the y and z components of the space charge forces), as it is already done for the RF contributions, from consideration by making the beam distribution uniform in x. The effect of this is illuminated by Fig. 3.4.3, which displays the x component of the force along the x- axis of the bunch, and the y component of the force at one σ_y, in the $y-z$ plane, for two cases: (a) where the beam distribution is gaussian in all three dimensions, and (b) where it is gaussian in y and z , but uniform up to a hard boundary in x . It can be seen that for the trigaussian beam the horizontal forces rise approximately linearly to a large value at about one σ_x. For the uniform distribution in x, however, there is very little horizontal field over about 85-90% of the beam, and the field rises steeply and nonlinearly near the beam boundary. As it was noted before, the maximum field is nearly the same, but the region of the beam which is affected is much smaller. Note that the vertical field is also nearly uniform over almost all of the beam, again degrading near the beam edge. In practice, one uses this final 10-15% of the beam as "guard charge" during the compensation to remove the horizontal field and homogenize the vertical field, then removes it by collimation after the linac section. In this way, one effectively removes the x-dimension from the compensation problem, reducing the problem to two dimensions, as in the (by now well understood) cylindrically symmetric case. It should be noted in this regard that the multiple envelope model for emittance compensation dynamics has been compared extensively to simulation as well as to the analytical model of Serafini and Rosenzweig [23], and has been essentially validated for cylindrically symmetric beams. Benchmarking of the models in the asymmetric case awaits further refinements of three-dimensional computational tools. Point-by-point simulations with PARMELA have displayed emittance compensation, but proven too noisy and time-consuming to explore the physics in detail. A three-dimensional particle-in-cell code for modelling this type of photoinjector is presently under development at UCLA, which should be better able to predict the performance limitations of this device quantitatively.

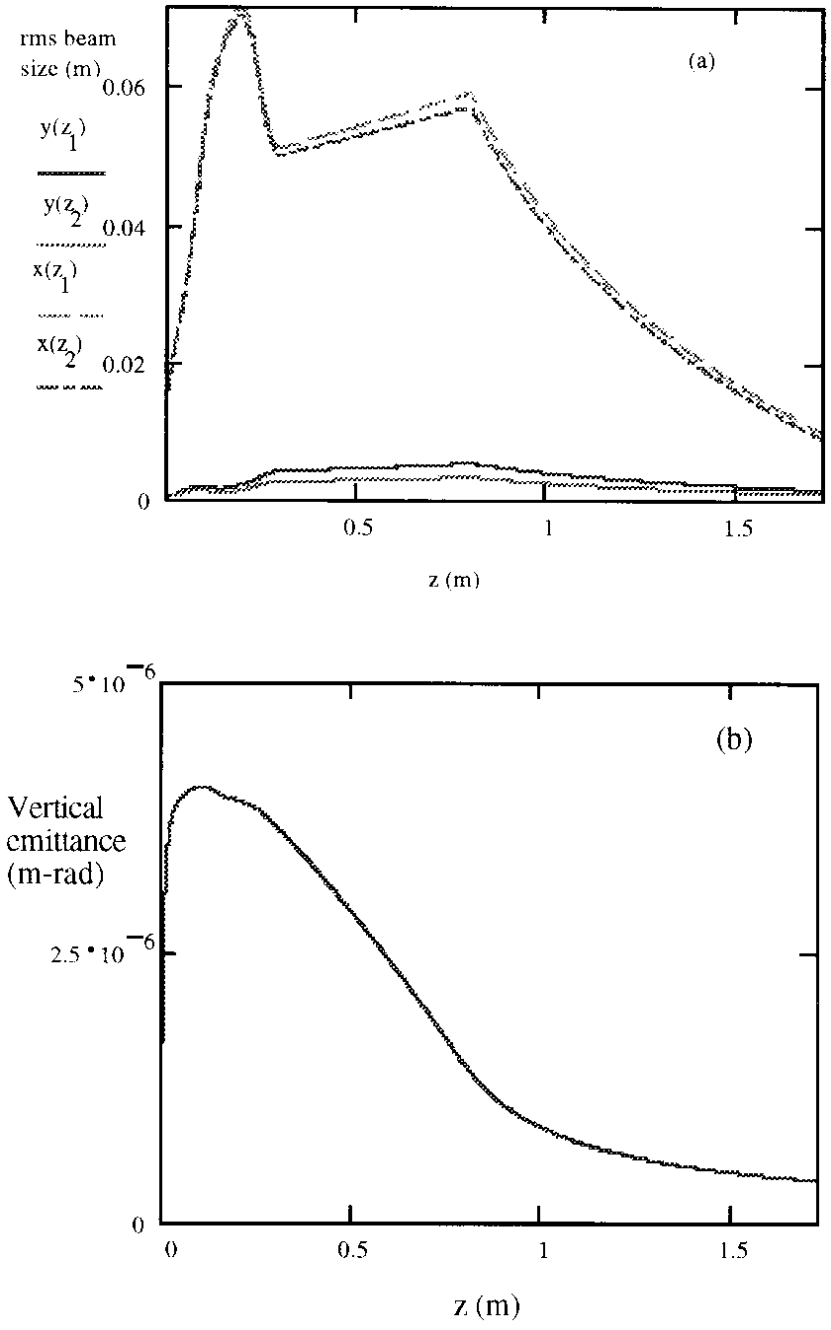
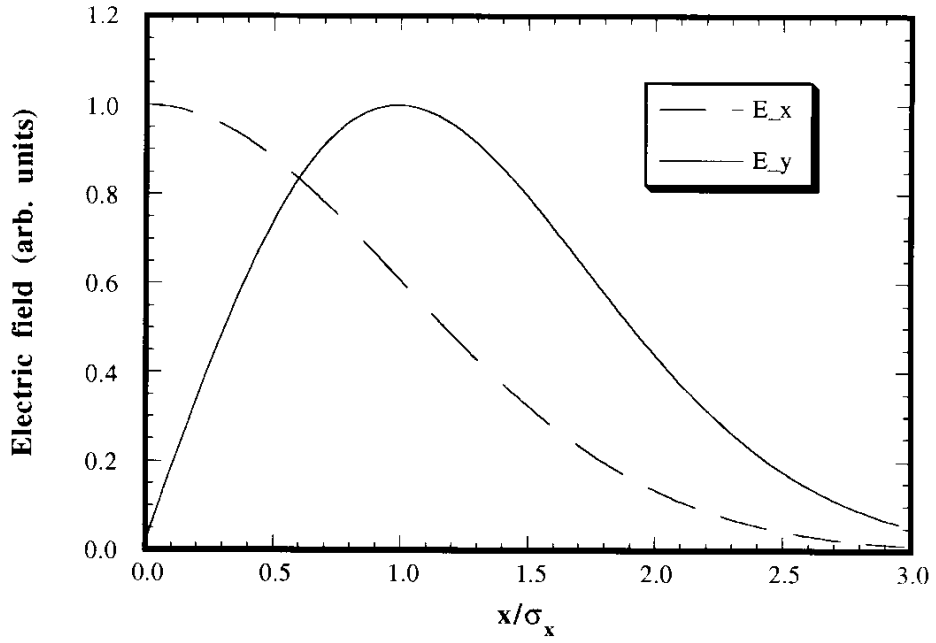
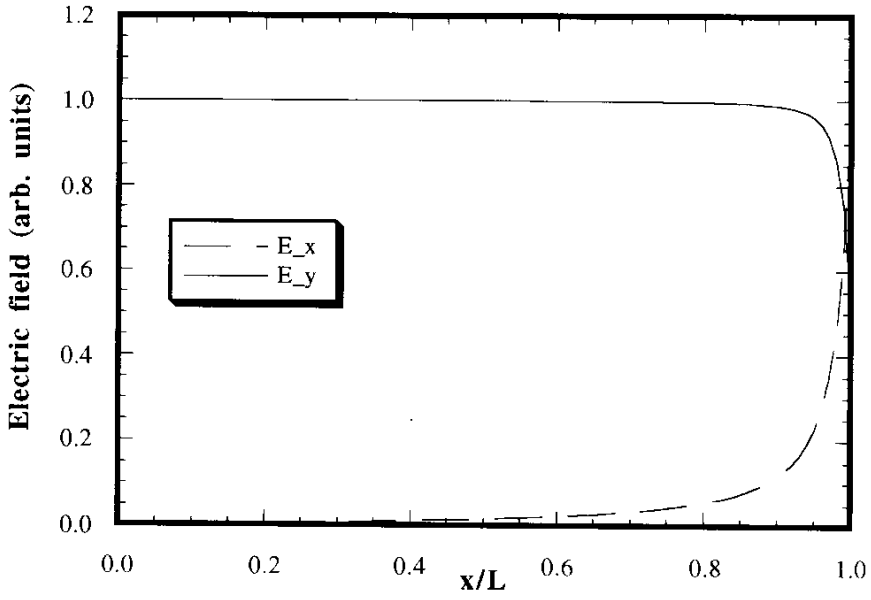


Figure 3.4.2: *(a) Rms beam sizes for two beam slices in multiple-slice model calculation of emittance compensation for 10 nC beam. A vertically focusing quad is placed symmetrically over the cathode plane, followed by a horizontally focusing quad, another doublet, a drift, and a TESLA cavity linac. (b) Vertical emittance evolution in this case.*



(a) Electric field components for gaussian in x, y and z.



(b) Electric field components for beam uniform in x, gaussian in y and z.

Figure 3.4.3: *Electric field components for (a) uniform, as opposed to (b) horizontally gaussian density profile.*

3.4.5 Positron Source

3.4.5.1 Introduction

A fundamental intensity limit for positron sources is given by the thermal stress which is built up in the conversion target due to the energy deposition of the primary electron beam. In case of TESLA the bunch train is so long that the target can be rotated even within the bunch train. With a target velocity of 50 m/s the heat load of about 50 bunches contributes to the thermal stress at a given position of the target. Table 3.4.2 compares the design parameters of TESLA with parameters reached at the SLC positron source which is the positron source with the highest intensity operating so far.

The SLC source is operating close to the stress limit of the target. An extension of this kind of source by two orders of magnitude in intensity seems to be impossible with a reasonable effort of technological development. Therefore a new concept based on the conversion of high energy wiggler radiation in a thin target has been developed [24, 25]. The scheme allows also to produce polarized positrons but the technological demands for a polarized source are higher than for an unpolarized source. The present design is for an unpolarized source but allows for a later upgrade to produce polarized positrons.

parameter	SLC	TESLA
# of positrons p. pulse	$3...5 \cdot 10^{10}$	$4.1 \cdot 10^{13}$
# of bunches p. pulse	1	1130
pulse duration	3 ps	0.8 ms
bunch spacing	8.3 ms	708 ns
repetition frequency	120 Hz	5 Hz

Table 3.4.2: *Comparison between TESLA and SLC parameters.*

3.4.5.2 General Layout

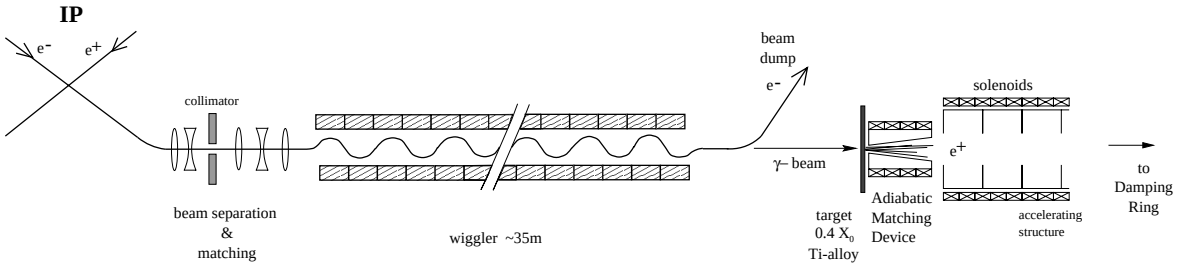


Figure 3.4.4: *Sketch of the positron source layout.*

A schematic layout of the source is shown in Fig. 3.4.4. The source is designed to produce twice as many positrons as required. The 250 GeV electron beam is used after collision as primary beam for the positron production. Due to the strong beam-beam forces during the interaction process the emittance of the beam is increased and the energy distribution has developed a long tail of low energy particles. The outgoing electron beam is separated from the incoming positron beam and then collimated in order to remove the low-energy tail. Still more than 80% of the particles pass through the optics (see next section). The final beam emittance is $\epsilon_x < 10^{-8}$m and $\epsilon_y < 10^{-9}$m, sufficiently small to match the beam to the successive wiggler section. The wiggler is a conventional design of 35 m length. Here photons are generated with a broad energy distribution extending up to above 70 MeV. A dipole section of ≈30 m length separates the electron beam from the photon beam and guides it to the beam dump. The photons are used to produce electron-positron pairs in the thin conversion target. A minimum spot size of the photon beam on the target of σ_{min}=0.5 mm can be realized. The heating of the target is dominated by the ionization losses of electrons and positrons given by $E_{dep} \approx 2\mathrm{MeVcm}^2/\mathrm{g}$ per charged particle. Thus the temperature rise of the target ΔT can be estimated as:

$$\Delta T[K] = \frac{2eN \cdot 2 \cdot 10^6}{c \cdot A \cdot \eta} \tag{3.4.8}$$

where c denotes the heat capacity in $Jg^{-1}K^{-1}$, A the source area in cm^2 and η the efficiency of positron capture after the target. In order to get N positrons, $2N/\eta$ particles (electrons and positrons) have to emerge from the target. The heat load of the target is determined by three more or less free parameters:
- the heat capacity of the target material c,
- the efficiency of the capture optics η,
- the source area A.
An increased source area counteracts a high capture efficiency, because the phase-space density of the emerging positrons is reduced, thus a small source area is favorable. Since in the present design the photons are generated in a wiggler, rather than by bremsstrahlung in the target, a high positron yield is reached with a thin target of only 0.4 radiation length (X_0) thickness. Compared to a conventional source the thermal stress of the target can be significantly reduced by two effects:
- one is able to use a material with low nuclear charge Z which in general have a high heat capacity.
- the divergence of the positron beam is small due to the reduced scattering in the target and hence the capture efficiency is increased.
A conventional target requires many radiation lengths for the full development of the electromagnetic cascade. Positrons which are produced in the first steps of the cascade will not emerge from the target due to the ionization losses inside the material. The ionization loss per radiation length depends on the material and is lower for high Z materials. Hence a high-Z material has to be used in a conventional source in order to reach a high positron yield. In thin targets, however, the conversion efficiency is in first order independent of the material. Therefore it is possible to use a low-Z material

which has a higher heat capacity (Dulong-Petit-rule). In the present design a Titanium alloy is foreseen as conversion target which allows to increase the particle density in the target by about an order of magnitude as compared to a high-Z Tungsten-Rhenium alloy.

The second advantage of a thin target is the reduced multiple scattering which determines the transverse beam emittance of the positrons. The rms scattering angle scales with the square root of the path length of the particle in units of X_0. The rms scattering angle scales non-linearly with the target thickness since in a conventional source most of the positrons are produced close to the target exit. Nevertheless, the transverse momenta of the positrons produced in a thin target are considerably smaller than the transverse momenta of positrons from a conventional source. Therefore the capture efficiency is increased by a factor 4-5 in the present design as compared to a conventional source.

The capture optics behind the target is of a conventional design. Since the positrons have a broad distribution of transverse and longitudinal momenta they have to be accelerated in an acceleration section embedded in a solenoid field. The acceptance of a solenoid channel is characterized by a large spot size and small angles while the positrons emerging from the target have a small spot size and large angles. To match the positrons to the acceptance of the solenoid, an Adiabatic Matching Device (AMD) is used. It consists of a tapered solenoid field which starts with a high initial field and tapers down adiabatically to the constant end field.

After acceleration to an energy of 150-200 MeV the positrons are separated from the electrons and the photons in a magnetic chicane. The electron and the photon beams are dumped and the positron beam is accelerated to its final energy and transferred to the damping ring.

The safety margin of a factor of two in the positron production rate is valid for 250 GeV electron beam energy. For operation at lower energy this margin is reduced. Without modifications, the source can be used down to about 160 GeV. Operation at lower energy with full design positron beam intensity requires to increase the length of the wiggler.

In the following the components of the positron source will be discussed in more detail.

3.4.5.3 Spent Beam Capture System

Due to the very strong beam-beam force the particles which do not pass through the center of the opposing bunch are strongly deflected. This disruption effect causes a broadening of the angular distribution after the interaction. The beam size remains nearly constant or even decreases due to the focusing pinch effect.

For all investigations a data set of roughly 12000 particles was created with a beam-beam simulation code using the parameters at the interaction point. Fig. 3.4.5 shows the phase space distributions both in the horizontal and vertical plane. The differences in the distribution are caused by the flatness of the beam.

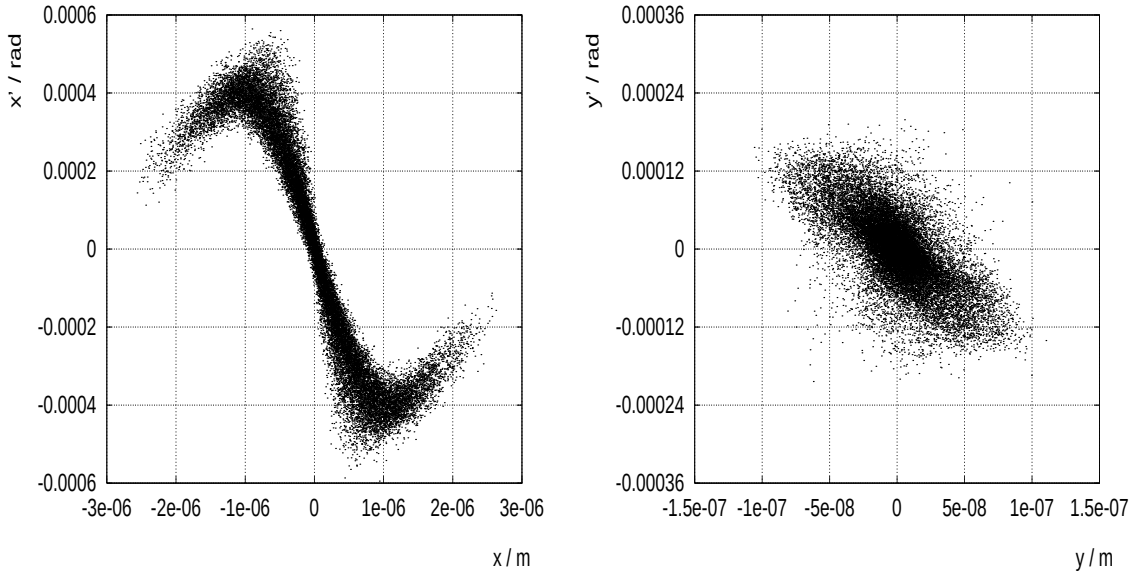


Figure 3.4.5: *Horizontal and vertical phase space distribution after the interaction.*

The deflected particles emit high-energy synchrotron radiation (beamstrahlung). This statistical process causes a broad energy spread of the disrupted beam and a mean energy loss of 3% of the initial energy. A non negligible part of the beam has lost more than 10% energy. Fig. 3.4.6 shows the energy distribution after the interaction.

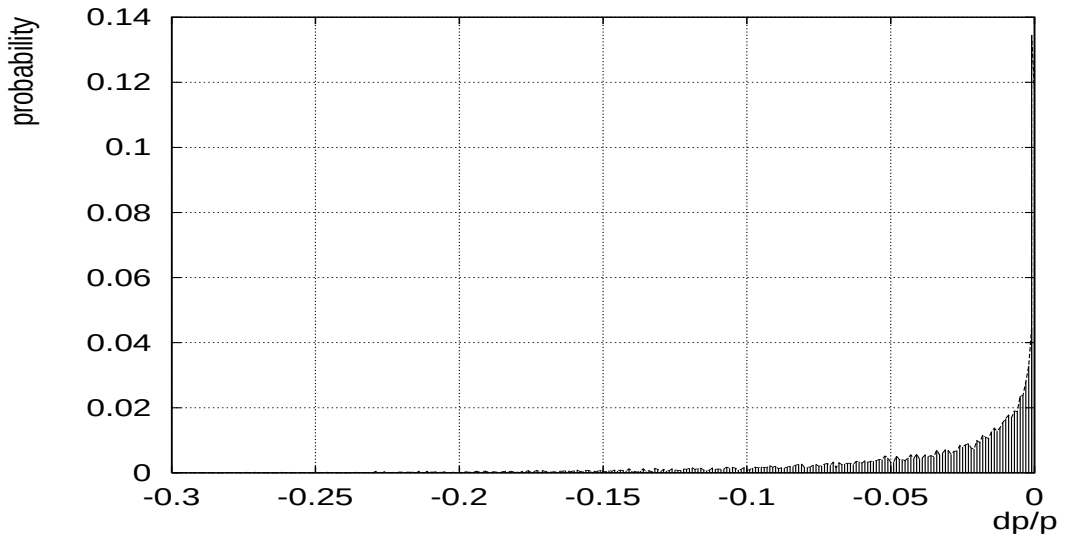


Figure 3.4.6: *Energy spectrum of the disrupted beam*

The spent beam is focused by the final doublet (FD) of the FFS for the incoming beam. As mentioned above the disrupted beam has a very small diameter and a

broad angular spread. Thus the matched β-functions after the IP are small and are rising to very high beta values inside the FD. This yields strong chromatic effects which cause an emittance growth by three orders of magnitude horizontally and two orders of magnitude vertically if uncompensated. Since the vertical emittance after the interaction is about two orders of magnitude smaller than required, it is sufficient to compensate only the horizontal chromaticity in a chromatic correction system (CCS).

Chromatic Correction System

After initial separation of the spent beam as described in section 3.7 a magnet lattice basically very similar to the horizontal CCS for the incoming beam follows. The boundary conditions for the design of this beam-optical system are:

- The geometry of the beamline should be such that it fits together with the FFS into the same tunnel.
- The bending magnets have to be sufficiently weak to avoid strong emittance growth from synchrotron radiation.
- The momentum bandwidth has to be large to accomodate a large fraction of the broad momentum distribution of the spent beam.

Starting from a basic layout which fulfills the first two conditions an optimization of the bandwidth is performed by adding sextupoles at strategic positions in the lattice. An ensemble of representative trajectories is tracked through the system. Optimum sextupole strengths which lead to a minimum beam emittance at the end of the lattice are determined. The procedure is similar to the one applied in ref. [26]. The amount of spent beam accepted by the system is directly related to the maximum tolerable emittance in the wiggler. As shown in Fig. 3.4.8, the horizontal emittance starts growing rapidly if more than 85...90% of the beam is passed through the beamline, indicating that the low-energy tail falls outside the bandwidth of the optics. So even with an optimized achromatic system it is not possible to transport the whole beam from the IR on through the capture optics and afterwards through the wiggler or undulator. Thus the low energy tail which still contains a considerable fraction of the beam power has to be collimated.

After the end of the CCS the optics is continued with a drift space of 20m and two quadrupole doublets with a distance of 18.7 m. The beam is focused such that the radiation emitted in a 35 m long wiggler hits the target in a focal point 30 m behind the wiggler i.e. 70 m behind the last quadrupole doublet.

Fig.3.4.9 shows the geometry of the whole capture system with the wiggler and the drift space to the target with regard to the beam delivery system of the oncoming positron beam. The separation between the two beamlines is sufficiently small to fit into a tunnel of 5 m diameter.

The optimization of the optics and the need to collimate the major part of the low energy particles mainly in regions without optical elements (see tab.3.4.3) results in a capture efficiency of 84.3% of the disrupted electron beam.

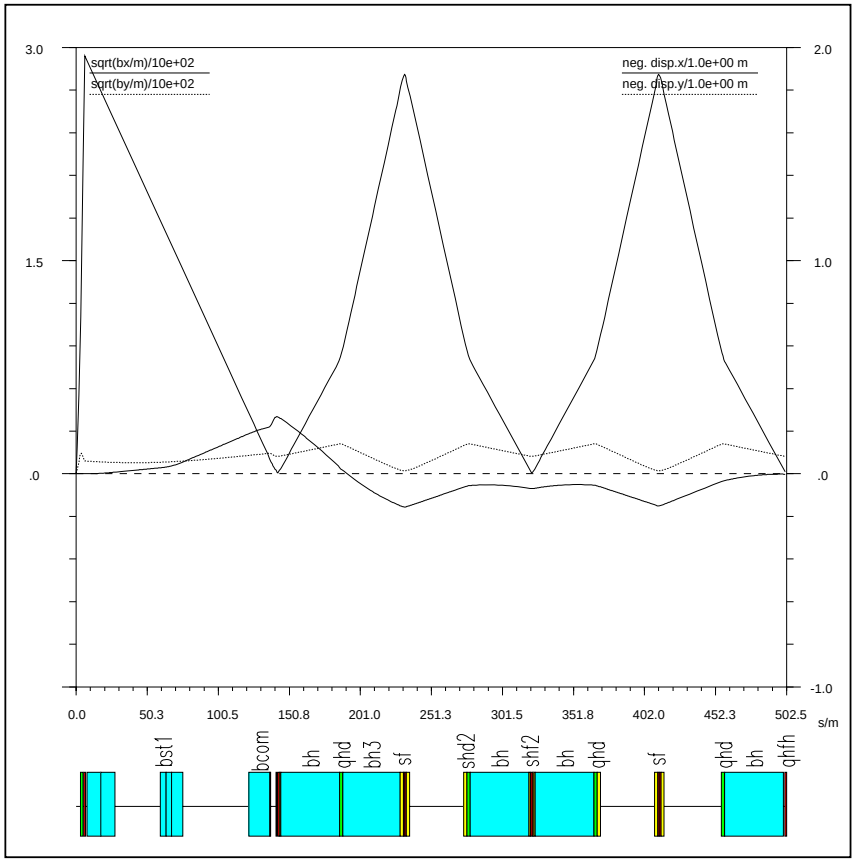


Figure 3.4.7: *Optics of the spent beam capture system.*

Although the rms-beamsize along the optics does not exceed values of 7 mm horizontally and 1mm vertically the maximum horizontal amplitudes can grow above 6 cm. To reduce the maximum horizontal beamsize a small fraction of the beam had to be collimated inside the second bending magnet in the first FODO cell. Fig. 3.4.10 and Fig. 3.4.11 shows the maximum beam envelopes of the disrupted beam along the capture optics with the position of the collimators. The 35 m long wiggler is positioned between 553 m and 588 m behind the IP.

Stability of the positron source

The efficiency of the capture optics for the disrupted electron beam depends on the parameters of the interacting bunches. The optics and the collimation system shown in Fig. 3.4.7 and Fig. 3.4.9 is developed for interacting bunches with design parameters. In reality, the capture efficiency for the disrupted beam may not be constant but depends on the energy spread due to beamstrahlung. As explained above, the beamstrahlung is a function of the bunch charge and other interaction parameters.

If we consider a perturbation of the interaction, for example a displacement of the orbit, the spent electron beam would not be disrupted like in a head-on interaction. Hence the number of captured electrons passing the collimation system can be higher.

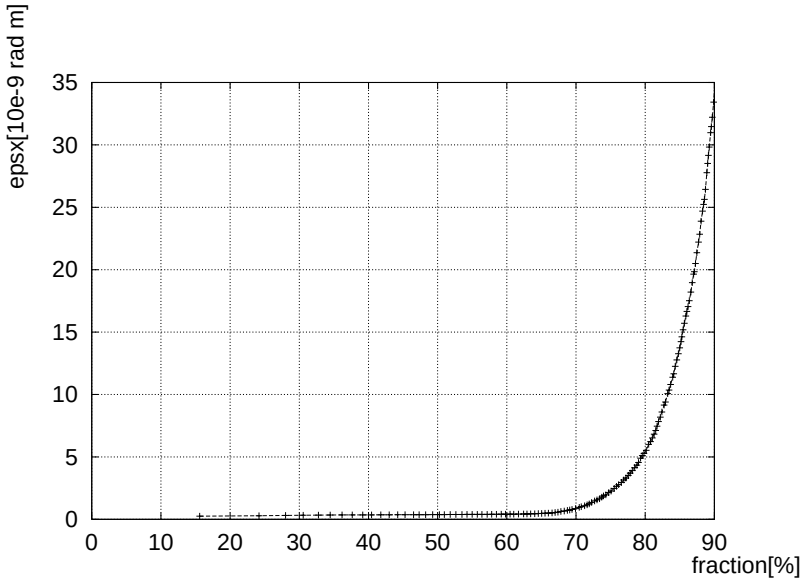


Figure 3.4.8: *Horizontal emittance at the end of the system vs. percentage of the captured spent beam.*

Collimator	Position behind IP [m]	Collimated beam power [KW]	Remarks
COL1	27.51-59.51	356.19	between the separator and the septa
COL2	75.51-122.12	372.52	between the septa and the first bend
COL3	188.64-229.14	74.6	inside the second bend in the 1st FODO cell
COL4	235.89-274.14	144.48	inside the CCS
COL5	502.52-522.52	212.33	after the CCS
COL6	547.74-552.74	1.88	5m before the wiggler to avoid vertical amplitudes $> 2\,\mathrm{mm}$

Table 3.4.3: *Collimator positions and collimated beam power.*

The positron bunches in the next pulse have a higher charge which in turn causes a higher electron beam energy spread after the IP. The fixed energy acceptance of the capture optics then leads to a lower intensity electron beam for the photon production. In the second pulse after the perturbation the positron bunch has a lower charge and so the system can oscillate. For a stable operation of the linac, this oscillation must be damped.

The investigation of stability is started with the maximum perturbation, i.e. the case that the bunches do not interact at all. The electrons pass the collimation system completely and the positron bunches for the next pulse have a 19% higher charge than

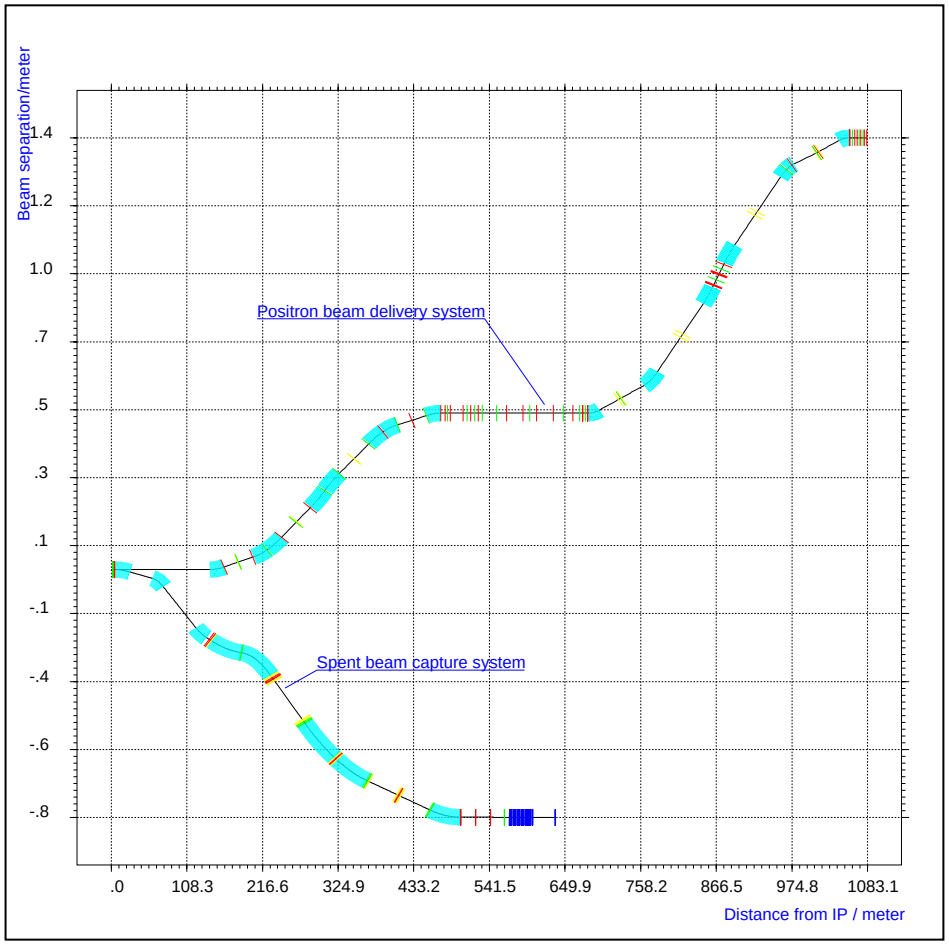


Figure 3.4.9: *Beam line for the disrupted beam with the 35 m long wiggler (lower branch) and the incoming positron beam (upper branch). At the end of the capture optics the separation is 1.3 m.*

designed. The next interaction is simulated with this asymmetric bunch charges. The outgoing disrupted electron beam was tracked through the capture optics to obtain the efficiency for the following pulse. This yields a new (reduced) positron intensity for the next interaction, etc. Fig. 3.4.12 shows the resulting behavior of the bunch charge as a function of interactions after the perturbation. The oscillation decreases and so the positron source can be operated stably from this point of view.

Efficiency Gain due to Parameter Change at the IP

The efficiency of the capture system is limited by higher order geometric and chromo-geometric aberrations. With a relatively moderate adjustment of the interaction parameters, both the horizontal beam emittance and the energy spread can be reduced, thus decreasing the population of the tails in the spent beam which have to be colli-

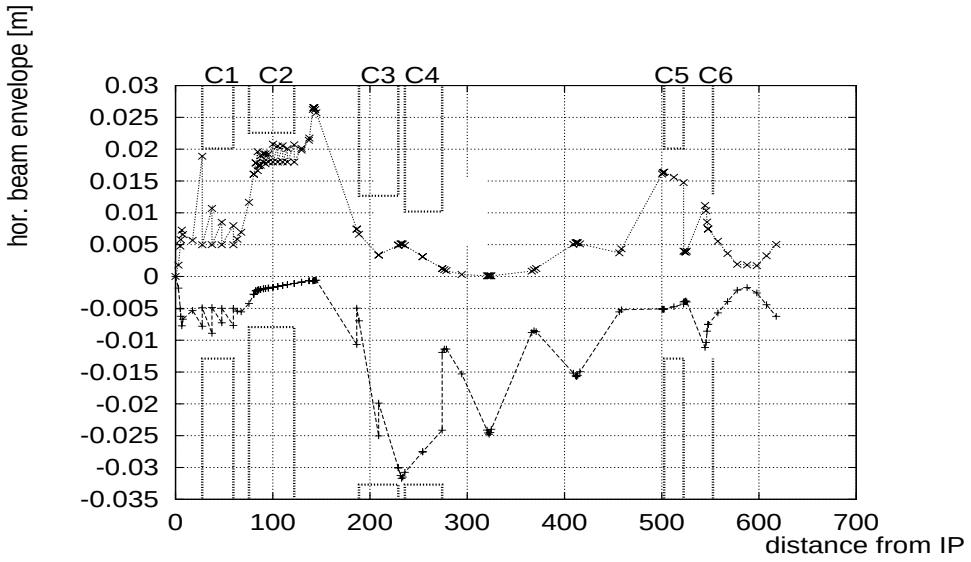


Figure 3.4.10: *Horizontal envelope for an ensemble containing 84% of the spent beam.*

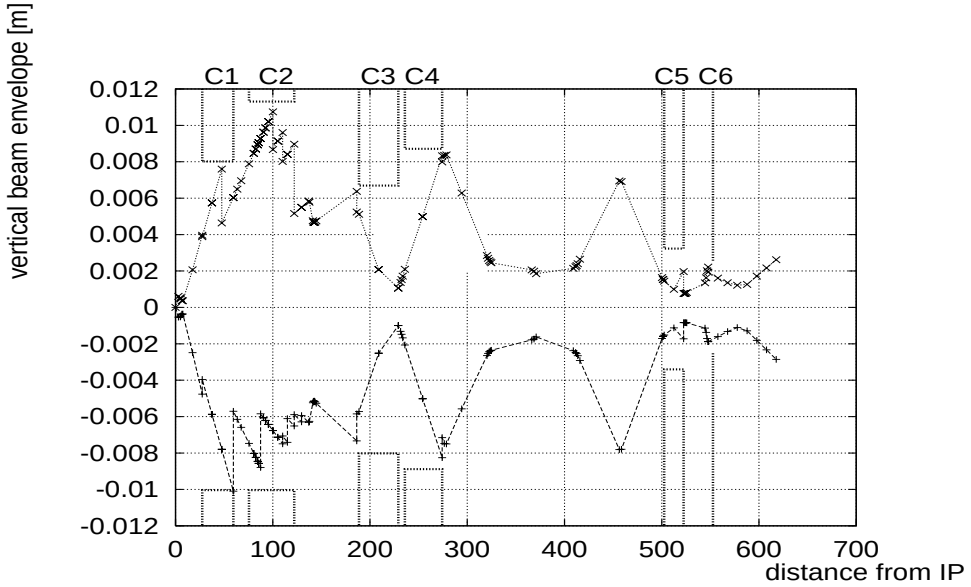


Figure 3.4.11: *Vertical envelope for an ensemble containing 84% of the spent beam.*

mated. This is demonstrated in Fig. 3.4.13 where the spent beam capture efficiency is shown as a function of the horizontal beta of the positron beam at the IP. An increase of $\beta_x^*(e^+)$ by a factor of two already reduces the amount of collimated spent beam from 16% to 6%. At the same time, the luminosity is reduced by only about 20%, which can be compensated by a slight reduction of the vertical beam emittance. With such a modification, the operation of the source becomes very insensitive with respect to fluctuations of the interaction parameters. Furthermore, an efficient operation of a polarized positron source would be greatly facilitated.

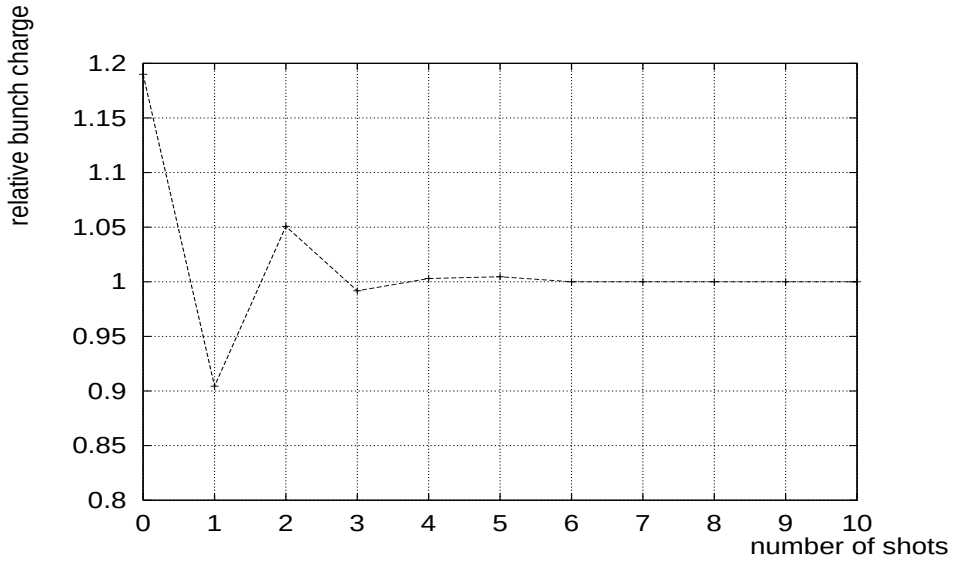


Figure 3.4.12: *Oscillations of the positron bunch intensity after one missing interaction. The bunch charge reaches the design value after 4 pulses.*

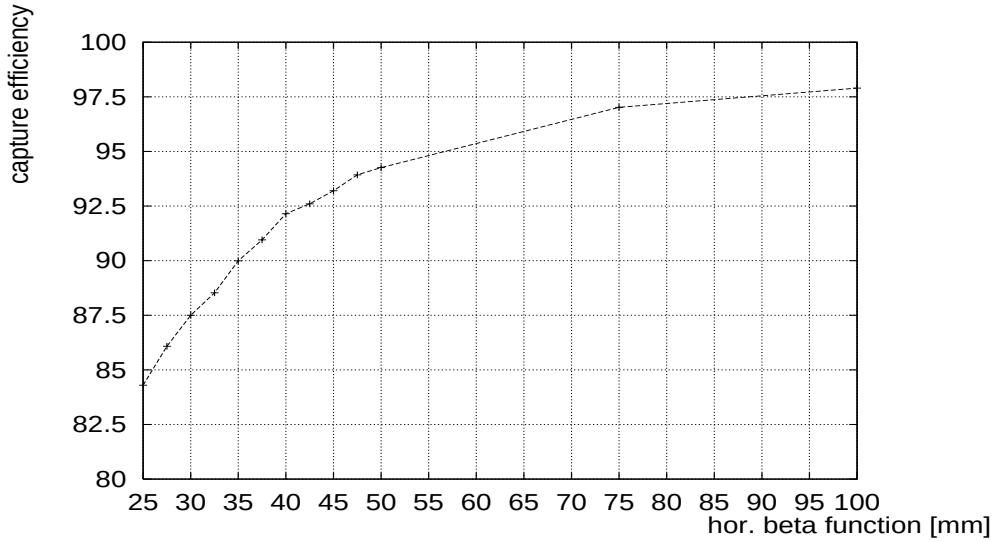


Figure 3.4.13: *Capture efficiency as a function of the horizontal β-function of the positron beam at the IP. Starting from the design value of 25 mm the capture efficiency reaches 97.9% of the disrupted beam for $\beta_x^*(e^+) = 100$ mm.*

3.4.5.4 Wiggler

A planar wiggler consists of a row of short dipole magnets with constant field strength but alternating field direction. For a large period length λ_w the photon number spectrum can be approximated by the well known photon number spectrum of a dipole magnet. Fig. 3.4.14 shows the photon number spectrum of a 1m long dipole with a magnetic field of B=1.7 T.

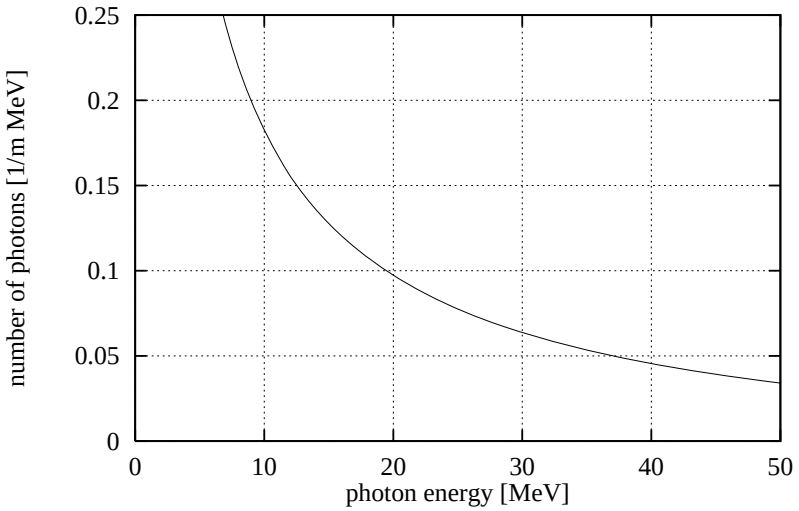


Figure 3.4.14: *Photon number spectrum of a planar wiggler with B=1.7 T at a beam energy of 250 GeV.*

In order to reach a small spot size of the radiation on the target the electron beam is focused through the wiggler onto the target. The contribution of the natural opening of the wiggler radiation to the spot size can be approximated by:

$$\sigma_{SR} \approx \frac{eB\lambda_w}{2\pi p_e} \cdot (L/2 + D) \tag{3.4.9}$$

Here p_e denotes the electron momentum, L the wiggler length and D the driftspace between the wiggler and the target.

For $L \approx$35 m and $D \approx$30 m the period length λ_w should not exceed 3.2 cm in order to reach a spot size contribution of 0.5 mm. Since the electron beam emittance is quite small a gap height of only 5 mm is sufficient to transport the beam through the wiggler. Today wigglers and undulators based on permanent magnet technology are widely used in both synchrotron radiation sources and free electron lasers. Period length and magnetic fields similar to the values discussed above are not unusual for these devices. In contrast to the sinusoidal field distribution of a synchrotron radiation device a more rectangular distribution would be favorable for the positron source in order to keep

the length of the wiggler shorter. Various inexpensive proposals of wigglers with a rectangular field distribution and similar period length are discussed in ref. [27].

3.4.5.5 Target

A high energy photon passing through a material can create an electron positron pair in the field of a nucleus (pair production). These charged particles again lose energy on their way through the material via collisions with electrons (ionization) and radiation processes in the field of a nucleus (bremsstrahlung). The ionization losses account for the majority of heat deposition in the material, while photons produced via bremsstrahlung may once more produce electron positron pairs. This sequence continues, i.e. the number of electrons, positrons and photons increases exponentially, while the mean energy of the particles decreases. The development of the shower decays when the particle energy drops below ≈10 MeV because the energy loss due to ionization exceeds the production of photons at low energies. Bremsstrahlung and pair production are essentially inverse processes, hence they can be characterized by a common parameter called the radiation length X_0. The radiation length of a material is approximated by:

$$X_0^{-1} = \frac{4\alpha r_e^2 N_A \rho}{A} \cdot Z(Z+1) ln(183 Z^{1/3}) \qquad (3.4.10)$$

where A is the atomic weight, Z the nuclear charge, ρ the density of the material and N_A Avogadro's number.

Since for important electromagnetic processes (bremsstrahlung, pair production, multiple scattering) some or all of the dependence upon the medium is contained in the radiation length, it is convenient to measure the target thickness in terms of the radiation length. Besides bremsstrahlung and pair production the shower development is influenced by many other processes like Compton scattering, Møller scattering, Bhabha scattering, photo effect etc. A complete treatment of the development of an electromagnetic cascade is possible only with a numerical simulation code. The following calculations have been performed with the code EGS4 [28]. The EGS code is a general purpose package for the Monte-Carlo simulation of electromagnetic showers. Since its formal introduction in 1978 it has been extended and improved and a lot of comparisons with experiments have verified the performance.

Fig. 3.4.15 shows the positron yield per electron as function of the target thickness for different materials. The incoming photon beam has been generated in a 1m long wiggler of 1.7 T magnetic field. For a target thickness above $0.5X_0$ the yield increases slowly before it starts to decrease again. Since the multiple scattering increases the beam emittance with increasing target thickness, the optimum overall efficiency, i.e. including the capture efficiency, is found at about $0.4X_0$. At this thickness the yield for a low Z material like Titanium is only about 16% lower than for Tungsten.

In Fig. 3.4.16 the temperature distribution in a Titanium target induced by a single train of wiggler photons from a 35 m long wiggler of 1.7 T is shown. As a safety margin the number of positrons produced in this simulation exceeds the nominal value behind the capture optics by a factor of two.

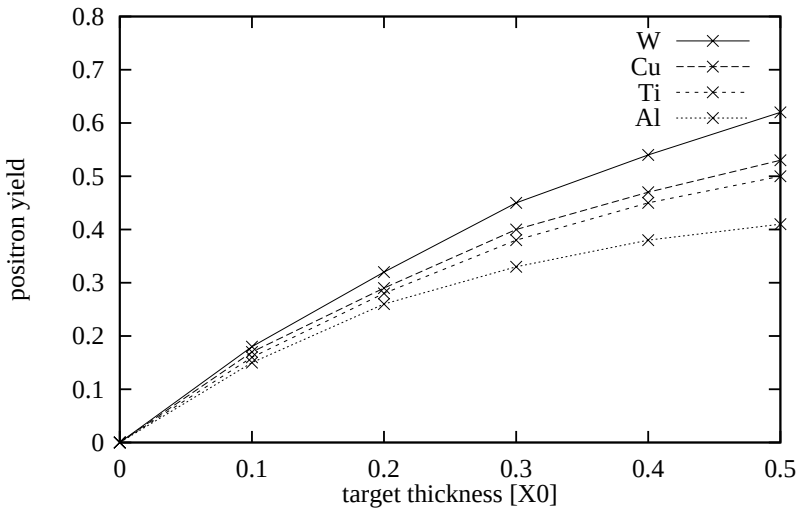


Figure 3.4.15: *Positron yield vs. target thickness for different materials. The photons are generated by a 250 GeV electron beam in a 1 m long wiggler of 1.7 T magnetic field.*

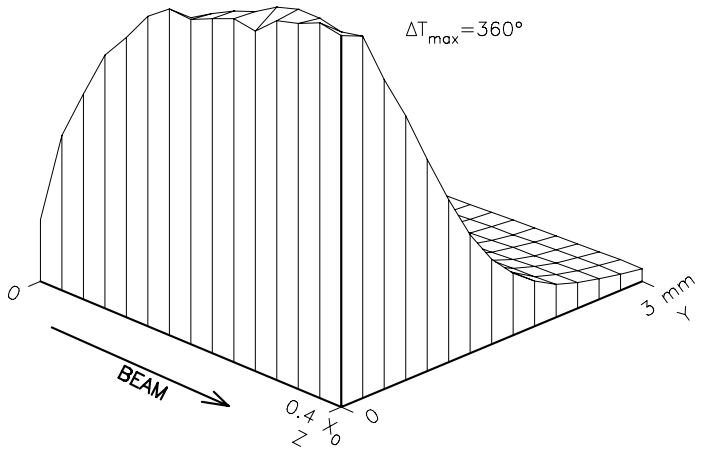


Figure 3.4.16: *Radial and longitudinal temperature distribution in the Titanium target. The spotsize of the incoming beam is $\sigma_{x,y}$=0.7 mm.*

The maximum temperature rise is only 360 °C, lower than the maximum temperature reached in a Tungsten-Rhenium target close to the stress limit given in ref. [29]. The density of charged particles is higher in the Titanium target but due to the 5 times higher heat capacity the temperature does not increase.

Besides the temperature rise the mechanical constants of the target material have to be taken into account. Experimental data for the maximum heat load of a low

Z positron conversion target are not available even though the material is used in high power collimators. In order to scale the results obtained for high Z targets to the case of low Z targets the material constants can be combined with respect to their proportionality to the mechanical stress $P_t \approx \Delta T \alpha E$ induced by heating (α: expansion coefficient, E: elastic modulus).

The induced stress has to remain below the 0.2% yield strength $P_{0.2}$ which is the stress where the deformation of a material ceases to be elastic, $P_t < P_{0.2}$.

Table 3.4.4 lists material constants for the Tungsten-Rhenium alloy used in the SLC positron source and various Titanium alloys.

The geometrical dimensions of the SLC target (thickness: 2.1 cm) and a Titanium target (thickness: 1.4 cm) are comparable while the temperature distribution in direction of the rotation of the target is quite different. The ratio $P_{0.2}/\alpha E = \Delta T_{crit}$ can be used to compare the maximum temperature rise that the target can withstand. It is seen from Table 3.4.4 that the maximum temperature in a Titanium target can be up to twice as large as in a Tungsten target.

material	c [J/kg]	α $10^{-6}K^{-1}$	E $10^{11}N/m^2$	$P_{0.2}$ $10^{7}N/m^2$	ΔT_{crit} K^{-1}	x_0 cm	ρ g cm^{-3}
W-26Re	0.13	5.0	4.00	91.0	455	0.35	19.3
Ti	0.52	8.4	1.16	55.0	564	3.56	4.54
Ti-5Al-2.5Sn	0.52	9.3	1.17	86.3	793	3.5	4.46
Ti-13V-11Cr-3Al	0.62	8.8	0.98	96.1	1114	3.5	4.82

Table 3.4.4: *Material constant of Tungsten-Rhenium and various Titanium alloys.*

The target has to rotate with a high velocity in order to avoid an overlapping of all bunches within one RF pulse at the same target spot. A velocity of about 50 m/s at the circumference of the target is necessary to spread out the beam spot over a distance of 4 cm. This can be achieved with a target of 80 cm diameter rotating at 1210 revolutions per minute. The maximum heat load reached in this case corresponds to 50 bunches overlapping at the same location or a temperature rise of about 380 °C. The diameter of the target ring is chosen in a way that subsequent bunch trains are placed next to each other on the target so that all parts of the target ring are loaded equally. The average heat load amounts to 7 kW. Cooling by radiation might be sufficient in case of the large target wheel but needs further careful investigation. Cooling with water is possible with a vacuum feed-through based on differential pumping as it has been designed at CERN [31].

3.4.5.6 The Capture Optics

The particles which emerge from the target have to be accelerated in a cavity embedded in a solenoid field for focusing. Here the final emittance and the efficiency of the positron source are defined. Since, compared to the thick target of the conventional

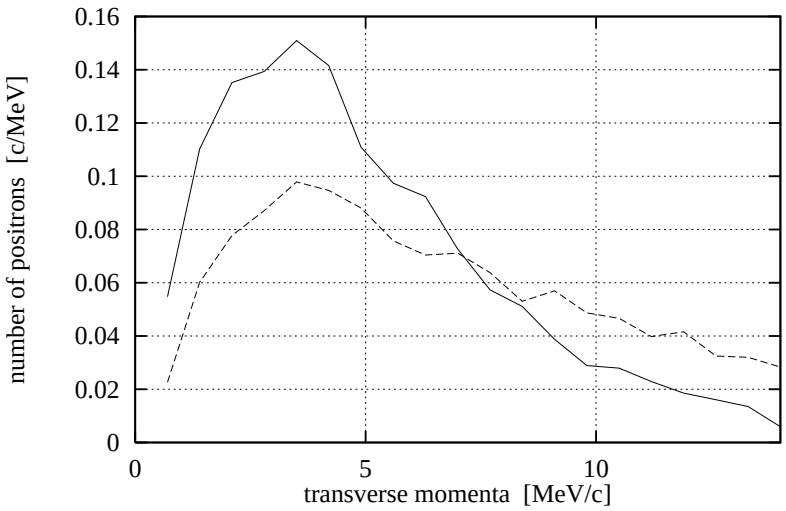


Figure 3.4.17: *Comparison of transverse momentum distribution for a SLC-like source ($6X_0$, W, dotted line) and a thin target as proposed here ($0.4X_0$, Ti, solid line).*

source the multiple-scattering is reduced in a thin target, the transverse momenta of the positrons emerging from the target are smaller. Fig. 3.4.17 compares transverse momenta of a SLC-like source with those of a thin target source driven by photons from a wiggler.

The smaller transverse momenta lead to a higher capture efficiency of the positrons behind the target. In order to match the emittance of the positron beam, characterized by a small spot size and a large divergence, to the acceptance of the solenoid which is determined by a large spot size and a small divergence a matching section is introduced between the converter target and the first accelerating cavity. The matching section consists of a so-called Adiabatic Matching Device (AMD), a tapered solenoid starting with a high initial field and tapered adiabatically down to the constant end field. The acceptance of the system is matched to the acceptance of the damping ring so that no particle losses occur later in the damping ring. Besides the acceptance of the damping ring that limits the useful acceptance of the matching device two mechanisms lead to an emittance growth of the positron beam in the matching device and hence to additional particle losses:
- emittance growth due to non adiabatic fields.
- bunch lengthening due to path length and velocity differences.
R. Helm has found a solution for the particle motion in an adiabatically varying solenoid field [32] from which he derived the optimum on-axis field distribution for the matching device along the longitudinal coordinate z as:

$$B(z) = \frac{B_i}{1 + g \cdot z} \tag{3.4.11}$$

where B_i is the initial solenoid field and g the taper parameter.

The condition for an adiabatic field variation is then given as $(gp_e)/(eB_i) \ll 1$. In order to fulfill this condition for particles with higher energy the taper parameter g has to be small. However, this means that the matching section becomes long and that the bunch lengthening becomes stronger. An optimum is reached with g=30 m^{-1} in the present design. Figure 3.4.18 shows the energy distribution of the positrons as they emerge from the target and the fraction of captured particles for the optimized optics. Figure 3.4.19 shows the longitudinal distribution of the positrons behind the AMD. While the core of the distribution represents the incoming electron beam (σ_z=0.7 mm or 1° of RF-phase), a long tail has developed due to bunch lengthening effects. The positrons in the tail will be lost in the damping ring since the curvature of the RF leads to a large energy deviation of these particles. In the simulation an energy acceptance of ±1 % corresponding to an RF phase of ±7.5° is assumed.

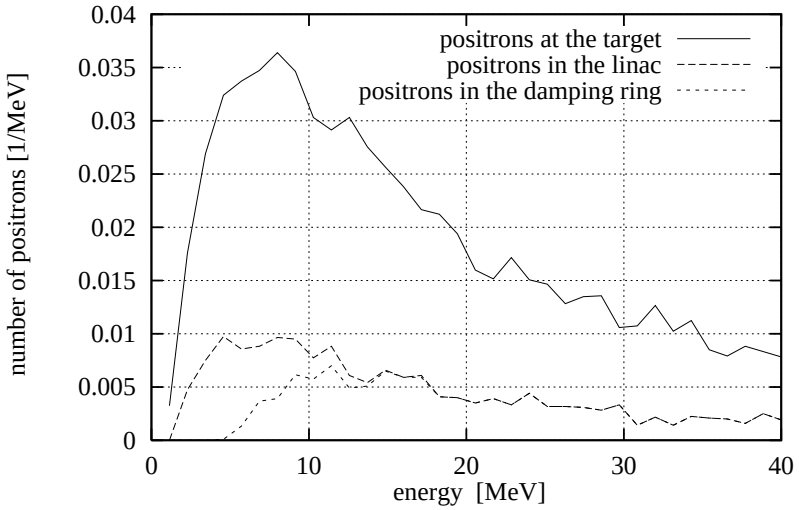


Figure 3.4.18: *Energy distribution of the positrons after the target (solid line) and the fraction of captured positrons (dashed lines).*

A capture efficiency of 17% can be reached with an initial field of 6 T. Fields of up to 8 T have already be realized [33] and 10 T seem to be feasible. However, since the bunch train is long the realization of a higher field is more difficult with a pulsed device and a lower field seems to be more reasonable. Table 3.4.5 collects important parameters of the positron source.

3.4.5.7 Positron PreAccelerator

The accelerating structure downstream of the positron source accelerates both positrons and electrons emerging from the titanium target. The number of electrons is 1.7 times that of the positrons. Both particles are accepted in the Adiabatic Matching Device

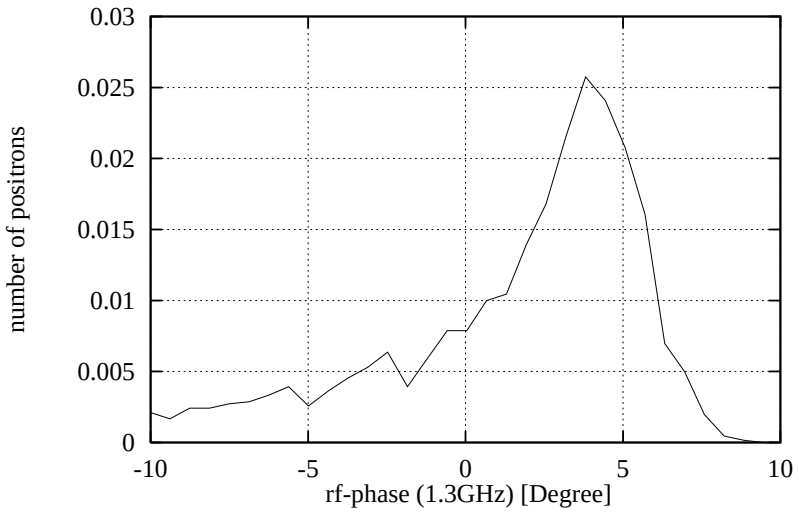


Figure 3.4.19: *Longitudinal beam profile of the positrons after the matching device.*

(AMD) where they undergo helical trajectories with identical characteristics, provided their energy and transverse momenta are the same. The only difference concerns the sense of rotation of the helical trajectories. The maximum transverse momentum, for particles moving close to the axis ($x_0 = y_0 \approx 0$), is about 3 MeV/c at the target and 0.5 MeV/c at the exit of the AMD. The corresponding maximum radius of the helix in a constant field solenoid surrounding the accelerator section is then approximately 10 mm. Hence positrons and electrons with the same transverse momentum are moving on helices with the same radius but with different path shifts, depending on their energy.

Energy differences between positrons and electrons

In an accelerating section optimised for positron acceleration, the electrons are first decelerated and then re-accelerated after having lost almost all of their kinetic energy. The energy difference between positrons and electrons, together in the first RF bucket at the entrance to the accelerating section reaches, therefore, about twice the average value of the energy of the electrons. This could be verified easily if the central energy of the accepted positrons in the Matcing Device is well defined. Experimental evidence of such behaviour has been observed at LAL-Orsay and CERN with a Quarter Wave Transformer as a Matching Device. Such a system, exhibiting a narrow energy acceptance makes the verification easier. In this case, at LAL as well as at LIL (CERN), the energy difference corresponded to twice the central energy associated with the matching system. We might then expect a similar behaviour for the electrons and positrons with an energy difference of about 15-20 MeV.

Choice of the pre-accelerator structure

The transverse and longitudinal acceptances for the positron source depend on the pre-accelerator parameters. Choosing a large iris aperture helps in improving the transverse

Wiggler	
peak field	1.7 T
period length	31 mm
gap height	5 mm
γ-spot size on target	0.7 mm
# of photons p. electron	368
mean photon energy	22 MeV
photon beam power	250 kW
Target	
material	Ti-alloy
thickness	1.42 cm ($0.4X_0$)
pulse temperature rise	360 K
av. power deposition	7 kW
Adiabatic Matching Device	
initial field	6 T
taper parameter	30 m^{-1}
end field	0.16 T
capture cavity iris radius	23 mm
General	
capture efficiency	17%
# of positrons p. electron	2
norm. e^+-beam emittance	0.01 m
total energy width	$\pm$30 MeV
required D.R. acceptance	0.048 m

Table 3.4.5: *Overview of the positron source main parameters.*

acceptance; however, bunch lengthening due to particle spiralling in the longitudinal magnetic field is made easier in large apertures. The limitation in bunch lengthening is imposed by the limitations on energy dispersion at the entrance of the Damping Ring (DR). Figure 3.4.20 shows the acceptance region in the (pL, pT) phase plane. The borders are defined as follows; - Energy: $(p_L)_{min}, (p_L)_{max}$
- Transverse momentum: $(p_T)_{max}$
- Bunch lengthening: Angular limits are represented for the case of S-band and L-band structures and for a maximum RF phase spread of 15 degrees.

Here the energy domain is given between 5 MeV and 25 MeV, the maximum transverse momentum at the target is about 6 MeV/c (3 MeV/c for near axis particles). It can be seen that in order to maximise the overall acceptance for the positron source an L-band structure with a rather large iris is preferable. This frequency choice (1.3 GHz) also matches with the main linac frequency.

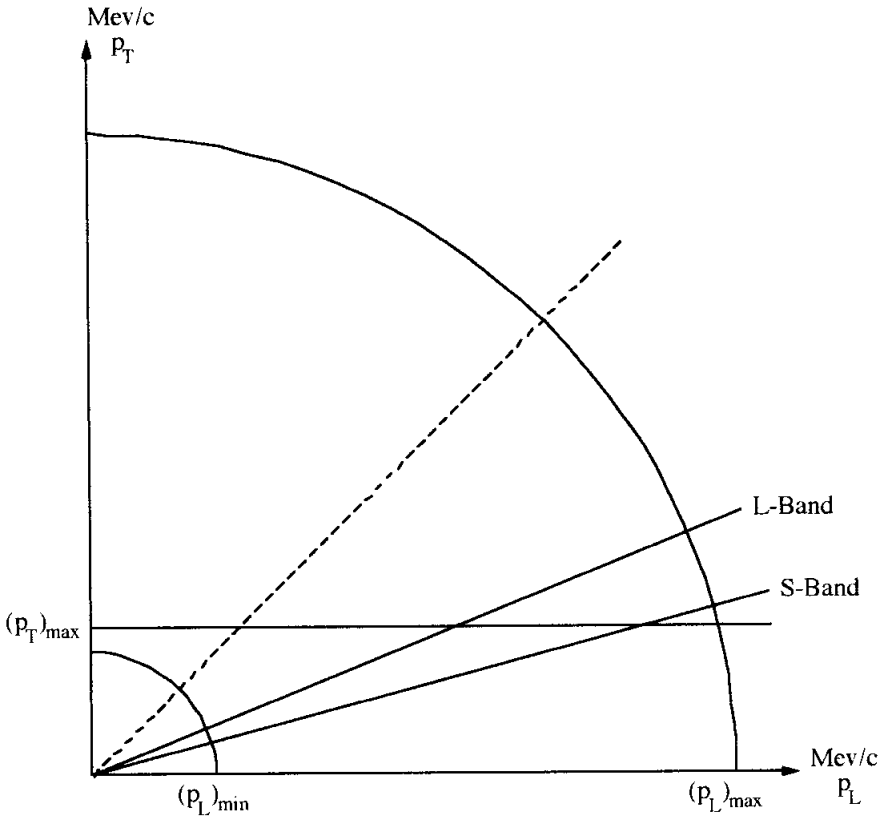


Figure 3.4.20: *Acceptance limitations for longitudinal and transverse momenta for two different choices of the injector frequency.*

Beam Loading

Exact calculations of the beam loading in the pre-accelerator with both electrons and positrons with different energies but with approximately the same transverse momenta is rather complicated. Simulations, using the RF characteristics of the chosen structure might shed some light on this important topic. Some studies have been started and results should appear in the near future. In the meantime we might expect some type of beam charge compensation to occur during the macropulse.

The Pre-Accelerating Structure

The positron pre-accelerator linac (3.2 GeV) would simply use superconducting L-band cavities such as those used for the main accelerator. However the optical characteristics of the positron beam emerging from the convertor preclude the use of a superconducting structure before the beam attains an energy of approximately 200 MeV. Therefore a copper structure is needed immediately behind the conversion target. As noted above, an L-band structure is preferred over a higher harmonic frequency from the point of view of positron acceptance and despite the loss of shunt impedance. The use of a nor-

mal conducting structure immediately raises the question of the maximum permissible accelerating gradient for 0.5 % duty cycles and the associated thermal dissipation problems. It is clear that the largest shunt impedance possible is desired in order to reduce cavity losses and to minimise the number of klystron- modulator assemblies required. A survey of copper accelerating structures indicates that a frequency scaled version of the MAMI injector linac structure [34] would be a good candidate for the positron pre-accelerator. This structure has a standing wave shunt impedance of 77 MΩ/m at 2445 MHz which indicates that we might achieve 56 MΩ/m at 1.3 GHz. Operation of the linac at a field of, say, 15 MV/m would then imply an average dissipation of 20 kW/m. Although this dissipation is slightly higher than that of the existing MAMI structure (15 kW/m continuous) it is no larger than the typical average power dissipation in single cell cw cavities used for electron storage rings (albeit at lower frequency). In addition, the smaller frequency of TESLA in comparison to MAMI and the resulting larger cavity dimensions, should easily allow increased cooling capabilities for an L-band version of the MAMI geometry. It should be noted that the problem of high power, long pulse operation of copper cavites might benefit from the existing studies being made for the TESLA RF gun which has to support a much higher field for the same pulse width and frequency. Operation at 15 MV/m implies a peak power of 4 MW/m. Consequently, the 10 MW klystrons currently under development for TESLA would comfortably feed two 1 m sections with 20% power in reserve. Six such klystrons would provide a positron beam of 180 MeV. One disadvantage of the MAMI structure is that its iris aperture, when scaled to 1.3 GHz is only 23 mm in diameter. This could seriously diminish the acceptance of the linac behind the matching device. Opening the iris would, in turn, reduce the shunt impedance, so requiring increased peak power for the same gradient. One could consider using increased aperture cells at the beginning of the pre- accelerator, followed by 23 mm apertures at an appropriate energy. Further cavity design work and simulation studies are needed here.

The 3 GeV positron pre-linac using standard TESLA cavities could start with doublet focussing between each cavity, yielding a beta function of about 1.5 m. The beam size at the entry of this linac would be $\sigma_{x,y}$=6 mm so that the iris aperture accomodates more than 5 standard deviations.

3.4.5.8 Low Intensity Auxiliary Source

For the commissioning of the positron damping ring, the main positron linac, etc., a source that works independent of the main electron linac would be desirable. Most of this work can be done or even has to be done at low intensities. Therefore a conventional low intensity source will be integrated into the high intensity source. An electron beam of only 500 MeV is sufficient to produce a few % of the design positron current using the thin Titanium target and the capture optics of the high intensity source. In this case the heat load of the target is high but does not reach the limits that have been previously discussed. The electron source for the low intensity positron source can also be used to provide electrons for the main linac for electron-electron and gamma-gamma collisions, respectively.

3.4.5.9 Potential Upgrade to a Polarized Positron Source

The proposed positron source allows also to produce polarized positrons as it was originally proposed by V. E. Balakin and A. A. Mickhailichenko [25, 24]. The technological demands are, however, much stronger in this case. Therefore the polarized source is regarded as a potential upgrade option.

In order to produce circularly polarized photons a short period helical undulator of about 100 m length has to be used instead of the planar wiggler. Since only the on-axis photons are completely circularly polarized off-axis photons have to be scraped off. This requires that the spot size of the photon beam on the target is dominated by the natural opening of the radiation rather than by the emittance contribution of the electron beam. Therefore the distance between the undulator and the target has to be increased to 150 m and the electron beam emittance has to be decreased to $\epsilon_x = 5 \cdot 10^{-10}$ m and $\epsilon_y = 10^{-10}$ m. The emittance requirements can easily be fulfilled for the vertical emittance, but can be achieved in the horizontal plane only with a strong collimation in the beam separation optics. Therefore only 60% of the electron beam can be matched to the undulator. Better results can be achieved if the horizontal β-function of the beams at the IP is increased with the disadvantage of some luminosity loss. Therefore improvements of the beam separation optics would be desirable. The parameters of the helical undulator are demanding and have not been reached so far. With a period length of 1 cm an on-axis field of 1.3 T has to be reached. These parameters can be realized only with superconducting technology at a gap radius of only about 2 mm. A detailed technological design of an undulator with the required parameters has not been worked out yet.

The polarization of the circularly polarized photons is transferred to the electron-positron pairs during pair production. For the calculation of the processes in the target the EGS4 code has been extended. The code includes polarization effects for pair production, bremsstrahlung and Compton scattering and yield a maximum longitudinal polarization of a polarized positron source of 50-60%.

Bibliography

[1] R. Alley et al., *The Stanford Linear Accelerator Polarized Electron Source*, Nucl. Instrum. And Meth. A365 (1995) 1.

[2] *Tesla Test Facility Linac – Design Report*, Ed.: D.A. Edwards, DESY Print, March 1995, TESLA 95-01.

[3] J.S. Fraser et al., Proc. of 1987 IEEE Particle Accelerator Conference, Washington DC, p. 1705.

[4] J. Clendenin et al., *Prospects for generating polarized electron beams for a linear collider using an RF gun*, Nucl. Instrum. And Meth. A 340 (1994) 133.

[5] K. Aulenbacher et al., *RF Guns and the Production of Polarized Electrons*, CLIC Note 303 and NLC-Note 20, May 1996.

[6] M. Woods et al., J. Appl. Phys. 73 (1993) 8531.

[7] A. Herrera-Gómez and W.E. Spicer, Proc. SPIE 2022 (1993) 51.

[8] D. Schultz et al., *The Polarized Electron Source of the Stanford Linear Accelerator Center*, SLAC-PUB-6606, August 1994; presented at 17th International Linear Accelerator Conference (LINAC 94), Tsukuba, Japan, 21-26 Aug 1994.

[9] H. Tang et al., *Experimental Studies of the Charge Limit Phenomenon in NEA GaAs Photocathodes*, SLAC-PUB-6515, June 1994; presented at 4th European Particle Accelerator Conference (EPAC 94), London, England, 27 Jun - 1 Jul 1994.

[10] P. Hartmann et al., *Picosecond Polarized Electron Bunches from a Strained Layer GaAsP Photocathode*, submitted to Nucl. Instrum. And Meth.

[11] J.E. Clendenin, *Polarized Electron Sources*, SLAC-PUB-95-6842, May 1995; Talk given at 16th IEEE Particle Accelerator Conference (PAC 95) and International Conference on High Energy Accelerators, Dallas, Texas, 1-5 May 1995.

[12] Although a conversion of the second or the third harmonic of the TTF laser system to 840 nm by optical parametric generation is in principle possible, it will be very difficult to achieve a reasonable efficiency to obtain the required laser energy per pulse.

[13] See e.g. W. Koechner, *Solid-State Laser Engineering*, Berlin, ^{4}1996, p. 400 ff.

[14] White et al., Opt. Lett. 17 (1992) 1067 and Ditmire, Perry, Opt. Lett. 18 (1993) 426.

[15] I. Will, P. Nickles, W. Sandner, *A Laser System for the TESLA Photo-Injector*, Internal Design Study, Max-Born-Institut, Berlin, October 1994.

[16] C. Travier, *On the Possibility of a Normal Conducting Photo-Injector for Tesla*, LAL/SERA/92-432, November 1992.

[17] B. Dunham, M. Jablonka, *Modeling of a High Charge Injector for the TESLA Test Facility*, DESY Print April 1994, TESLA 94-11.

[18] M. Breidenbach et al., *An inverted-geometry, high voltage polarized electron gun with UHV load lock*, Nucl. Instrum. And Meth. A 340 (1994) 1.

[19] C.K. Sinclair, *A 500 kV photoemission electron gun for the CEBAF FEL*, Nucl. Instrum. And Meth. A 318 (1992) 410.

[20] J.B. Rosenzweig et al., *Design of a High Duty Cycle, Asymmetric Emittance RF Photoinjector for Linear Collider Applications*, Proc. 1993 IEEE Particle Accel. Conf. 3021 (IEEE, 1994).

[21] K.J. Kim, Nucl. Instr. Methods A 275, 201 (1988).

[22] Warner Bruns, private communication.

[23] L. Serafini and J.B. Rosenzweig, *Envelope Analysis of Intense Relativistic Quasi-laminar Beams in RF Photoinjectors: A Theory of Emittance Compensation*, submitted to Phys. Rev. E.

[24] V. E. Balakin and A. A. Mikhailichenko, *The Conversion System for Obtaining High Polarized Electrons and Positrons*, Preprint INP 79-85, 1979.

[25] K. Flöttmann, *Investigations Toward the Development of Polarized and Unpolarized High Intensity Positron Sources for Linear Colliders*, DESY-93-161, 1993.

[26] R. Brinkmann, *Optimization of a Final Focus System for Large Momentum Bandwidth*, DESY-M-90-14, 1990.

[27] R. Brinkmann, J. Pflüger, V. Shiltsev, N. Vinokuro, P. Vobly, *Wiggler Options for TESLA Damping Ring*, DESY-TESLA-95-24, 1995.

[28] W. Nelson, H. Hirayama and D. Rogers, *The EGS4 Code System* SLAC-265, 1985.

[29] S. Ecklund, *Positron Target Materials Tests*, SLAC Collider Note-128, 1981.

[30] E. M. Reuter and J. A. Hodgson, *3D Nummerical Thermal Stress Analysis of the High Power Target for the SLC Positron Source*, SLAC-PUB-5370, 1991.

[31] P. Sievers and M. Höfert, *A Megawatt Electron Positron Conversion Target - A Conceptual Design*, Proc. European Accelerator Conference, Rom 1988.

[32] R. H. Helm, *Adiabatic Approximation for Dynamics of a Particle in the Field of a Tapered Solenoid*, SLAC-4, 1962.

[33] H. Ida, *Present R&D Status of Positron Sources for JLC-1 and ATF*, presented at Sixth International Workshop on Linear Colliders, Tsukuba 1995.

[34] Compendium of Scientific Linacs, CERN/PS 96-32, 1996.

3.5 Damping Ring

3.5.1 Introduction

The normalized transverse emittance of the beam produced at the positron source is several orders of magnitude larger than the design emittance for collider operation. The only conceivable way to achieve the required emittance is to store the beam temporarily (between linac pulses) in a damping ring. In case of the electron beam, using a laser rf-gun may be a way to produce a beam with sufficiently small emittances. This could provide a cost saving first stage of TESLA with only one damping ring. However, for optimum performance of the collider with a vertical emittance as small as possible, an electron damping ring will likely be required as well. Since there are only minor differences between the electron and the positron version of the damping ring, with somewhat relaxed requirements for the e- ring, we focus in this section on the positron ring design.

Assuming a normalised emittance of the positron beam injected into the ring of $\gamma\epsilon_i = 0.01\,\mathrm{m}$ (see section 3.4.5), the required emittance reduction factors are $\epsilon_{f;x}/\epsilon_i = 1.2\times10^{-3}$ and $\epsilon_{f;y}/\epsilon_i = 2\times10^{-5}$ for the horizontal and vertical plane, respectively. Here, a 25 % emittance dilution budget between extraction from the ring and the interaction point is included. The final emittance in the ring after a storage time equal to the linac cycle time $T_c = 200\,\mathrm{ms}$ is given by

$$\epsilon_f = \epsilon_{eq} + (\epsilon_i - \epsilon_{eq})exp(-2T_c/\tau_D) \tag{3.5.1}$$

where ϵ_{eq} and τ_D denote the equilibrium emittance and transverse damping time, respectively. The resulting requirements for these basic parameters are summarized in table 3.5.1.

In order to keep the demands for the bunch compressor moderate, the damping ring should also fulfill limitations concerning the longitudinal phase space. Furthermore, the ring must safely accomodate the injected positron beam, which sets lower limits on the transverse and longitudinal acceptance. These additional parameters are included in table 3.5.1.

The main complication which arises for the TESLA damping ring design is due to the pulse structure of the linac. The train of 1130 bunches per pulse has a length of 0.8 ms, or 240 km. Since a damping ring of that size would be unreasonable, the bunchtrain must be stored in the ring in a compressed mode, with a bunch spacing much smaller than in the linac. The bandwidth of the systems required for injection into and extraction from the ring are inversely proportional to the bunchspacing in the ring. Furthermore, the demands for the fast multibunch feedback system grow with decreasing ring length. We choose here a ring design with a circumference of C=17 km, leading to rather conservative bandwidth requirements with a bunch spacing of 50 ns.

In the next subsection the general layout, the magnet lattice and the beam optics of the positron damping ring are discussed. It is followed by descriptions of the injection/extraction devices, the vacuum system and the rf-system. Furthermore an analysis of collective beam dynamics and of requirements for feedback systems is given.

e^+ emittance at injection $\gamma\epsilon_{i;x,y}$	$0.01m$
hor. equilibrium emittance $\gamma\epsilon_{eq;x}$	$1.2 \times 10^{-5}m$
hor. emittance at extraction $\gamma\epsilon_{f;x}$	$1.2 \times 10^{-5}m$
vert. equilibrium emittance $\gamma\epsilon_{eq;y}$	$1.0 \times 10^{-7}m$
vert. emittance at extraction$\gamma\epsilon_{f;y}$	$2.0 \times 10^{-7}m$
cycle time T_c	$200ms$
transverse damping time $\tau_{D;x,y}$	$35ms$
equilibrium bunch length σ_z	$< 1cm$
equilibrium momentum spread σ_p/p	0.1%
transverse acceptance $\gamma A_{x,y}$	$> 0.05m$
momentum acceptance A_p	$> 0.5\%$

Table 3.5.1: *Goal parameters for the positron damping ring*

3.5.2 General Layout

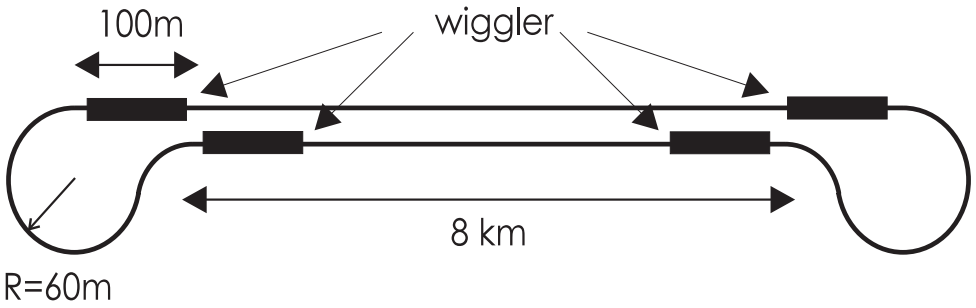


Figure 3.5.1: *General layout of the "dogbone"-shaped damping ring*

In order to avoid the need for a costly additional 17 km ring tunnel, the TESLA damping ring is designed such that about 95 % of the lattice can be installed in the linac tunnel. Additional tunnels are only required for the two "loops" at the ends of the 8 km long straight sections (see fig. 3.5.1). The geometry of these arcs is shown in more detail in fig. 3.5.2. They consist of a negative bending section with a bend angle of $\Phi_- = -11/48 \times 2\pi$, followed by a short straight section and a positive bending section with $\Phi_+ = 35/48 \times 2\pi$. The length of the straight section is adjusted such that a horizontal spacing of the long straight sections in the linac tunnel of 0.8 m results. The basic unit of the arcs is a FODO cell with a bend angle of $\pi/24$ and an average bending radius of 60m (the negative and positive bending sections have 11 and 35 FODO cells, respectively). In order to achieve a sufficiently small damping time, wiggler sections are foreseen at the end of the long straights (fig. 3.5.1). The transverse damping time can be written as (partition numbers $J_{x,y} = 1$)

$$\tau_{x,y} = \frac{2E_bC}{U_0c} \tag{3.5.2}$$

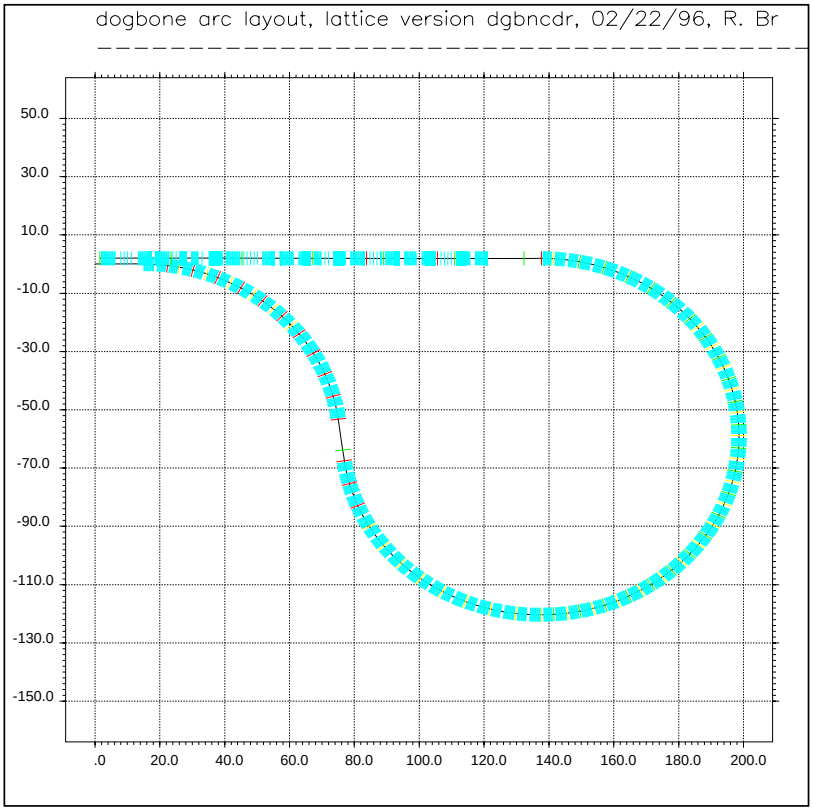


Figure 3.5.2: *Geometry of the "dogbone" damping ring arcs*

where E_b is the beam energy and U_0 the energy loss per revolution. The latter is given by

$$U_0[MeV] = 1.26 \cdot 10^{-3} \times E_b^2 \oint B^2 ds \tag{3.5.3}$$

with E_b in GeV and the magnetic field B in T. For the beam energy we choose $E_b = 3.2\,\mathrm{GeV}$ as the result of a parameter optimization, taking into account considerations such as:

- The RF-system. The required circumferential voltage U_{RF} grows with increasing E_b.
- The emittance contribution from the arcs. For a given beam optics, the normalized horizontal emittance scales approximately as $\gamma\epsilon_{arc} \propto E_b^6$, when the dominant contribution to the damping comes from the wigglers.
- The bunch compressor, which favors a low beam energy.
- The total length L_w of the wigglers. For given field strength B, L_w scales like E_b^{-1}, thus favoring a high beam energy.

Using eqs. 3.5.2 and 3.5.3 yields for the required damping integral:

$$\oint B^2 ds \approx \int_{\text{wiggler}} B^2 ds = 805\ T^2 m \tag{3.5.4}$$

and for the energy loss per revolution

$$U_0 = 10.4\ MeV \tag{3.5.5}$$

The wiggler design described below has $B_w = 1.5\,\text{T}$, so that the length of each of the four wiggler sections is about 100 m. The relative beam energy spread σ_E is determined from E_b and B_w as:

$$\sigma_E \approx 0.468 \cdot 10^{-3} \times \sqrt{E_b[\text{GeV}] B_w[\text{T}]} = 1.02 \cdot 10^{-3} \tag{3.5.6}$$

The roughly 8 km long straight sections are layed out as simple periodic FODO channels. Injection and extraction takes place at opposite ends of the straight section, so that the damping ring itself acts as part of the beamline required to transfer the positrons to the entrance of the main linac.

3.5.3 Beam Optics

Despite the somewhat unconventional layout of the "dogbone"-shaped damping ring, the beam-optical design is rather straight-forward and simple. In the following, the beam optics for the different parts of the ring are discussed. A compilation of the main parameters of the damping ring is given at the end of this section (table 3.5.6).

Straight Sections

The basic unit of the long straight sections is a 60° phase advance FODO cell of 100 m length. This choice of focussing strength is a compromise between keeping the β-functions within limits and avoiding an excessive contribution to the chromaticity, which can only be compensated in the relatively short arcs. For the quadrupoles, an aperture radius of 50 mm is foreseen, which provides a linear acceptance well in excess of the value required for the injected beam. The required gradient is small, so that a simple, low-cost design can be chosen (small amount of iron, no water cooling required). A summary of the beam optics and magnet parameters is given in table 3.5.2.

Arcs

The FODO cell in the arcs has a length of 8.5 m. Each cell comprises two quadrupoles, two dipoles, two sextupoles and two correction magnets. The design value of the phase advance is set to 90° per cell. The corresponding magnet strengths and optics parameters are listed in table 3.5.3. The quadrupoles and sextupoles are designed to allow for an increase in focussing strength, which provides sufficient flexibility to provide a smaller horizontal emittance and momentum compaction factor when desirable. In

FODO cell length	100 m
cell phase advance	60°
total # of cells	152
chromaticity contribution $\xi_{x,y}$	27.9
maximum $\beta_{x,y}$	175 m
average $\beta_{x,y}$	110 m
aperture radius	50 mm
norm. linear acceptance	0.089 m
quadrupole length	0.2 m
quadrupole gradient	1.05 Tm^{-1}
Ampère windings	1000
power per quad	70 W

Table 3.5.2: *Optics and magnet parameters for the straight sections*

FODO cell length	8.5 m
cell phase advance	90°
total # of cells	92
chromaticity contribution $\xi_{x,y}$	29.3
maximum $\beta_{x,y}$	13.5 m
average $\beta_{x,y}$	8.3 m
half aperture ($hor. \times vert.$)	25×15 mm
norm. linear acceptance	0.10 m
dipole length	3 m
dipole field	0.205 T
Ampère windings	5720
quadrupole length	0.2 m
quadrupole gradient	19.2 Tm^{-1}
Ampère windings	4300
power per quad	1 kW
sextupole length	0.2 m
sextupole gradient (max.)	113 Tm^{-2}

Table 3.5.3: *Optics and magnet parameters for the arc sections*

the design optics, the arc contribution to the normalized emittance amounts to $\gamma\Delta\epsilon_x = 0.23 \times 10^{-5}$ m (the main contribution comes from the wigglers, see below). The total natural chromaticity of the arcs is comparable to the one from the straights, but it has to be taken into account that the entire ring chromaticity must be compensated in the arcs in order to define the required sextupole strengths. All magnets have conservative strengths requirements and can be built as standard storage ring components (e.g. of the type used in the HERA electron ring, scaled down in size).

At the end of the arcs, a dispersion suppressor and matching section is foreseen adjacent to the wiggler sections. The matching between the positive and negative

bending sections in the arcs is provided by a "$-I$ transformer", see fig 3.5.3.

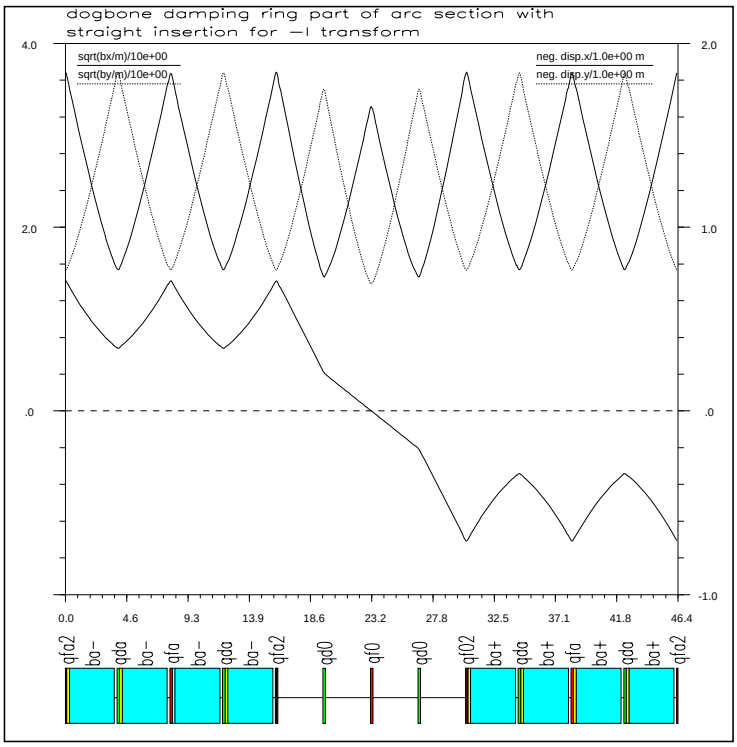


Figure 3.5.3: *Beam optics in the arc section showing the sign reversal of the dispersion function by inserting two "empty" FODO cells.*

Wiggler Sections

Besides fulfilling the requirement for the damping integral, as described above, the wiggler section must be layed out such that the required horizontal emittance is safely achieved. For a wiggler with approximately rectangular field shape (period length λ_w large compared to gap height g_w), the emittance is given by:

$$\gamma\epsilon_x = 2.22 \cdot 10^{-6} m \times B_w^3 \lambda_w^2 \beta_{x,w} \tag{3.5.7}$$

where $\beta_{x,w}$ denotes the horizontal beta function averaged over the wiggler length. With our choice $B_w = 1.5\,\mathrm{T}$ and $\lambda_w = 0.4\,\mathrm{m}$, we need $\beta_{x,w} \approx 10\,\mathrm{m}$. This requirement is easily met by arranging the lattice as a FODO structure with a cell length of about 10 m and a phase advance of 60°. The same quadrupoles as in the arcs can be used here. Sufficient reserve for the focussing strength is provided, thus allowing to reduce $\beta_{x,w}$ by about a factor of two if a smaller horizontal emittance is desirable. In total, 48 wiggler cells are installed, placed between the arcs and the long straight sections.

The contribution to the ring chromaticity (see table 3.5.4) is much smaller than from the long straight sections so that a tighter focussing in the wiggler cells will not have a large influence on the required sextupole strength in the arcs. Today wigglers and

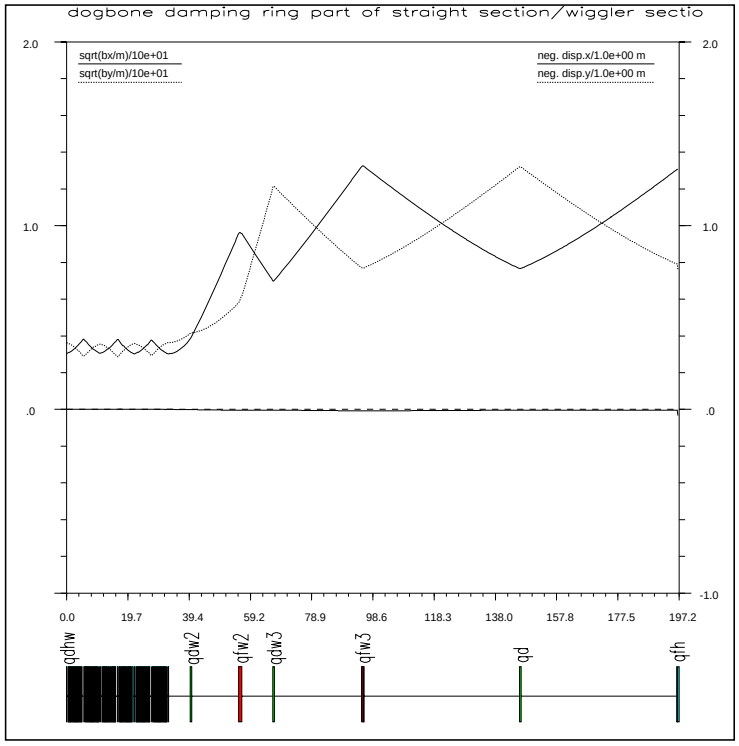


Figure 3.5.4: *Beam optics in the wiggler-to-straight matching section.*

undulators are widely used in both synchrotron radiation sources and Free Electron Lasers. During the last 15 years permanent magnet (PM) technology based initially on SmCo and later on NdFeB material has been developed to build these devices. At intermediate field levels up to about 2 T this technology is now almost exclusively used. The advantages of PM technology are evident: the devices are passive, which means they do not need any power supply, do not cause operational costs and require almost no maintenance. PM material is characterized by its $B-H$ demagnetization curve. The optimum operational point is chosen when the energy product $B \cdot H$ is maximum. In this case the volume of magnet material is minimized and given by :

$$V_m \simeq \frac{gh}{\mu_0 (BH)_{max}} \int B^2 dL, \tag{3.5.8}$$

The demagnetization curve of NdFeB is essentially a straight line throughout the second quadrant with a slope very close to one. In this case

$$(B \cdot H)_{max} \simeq \frac{1}{2} B_r \frac{1}{2} H_c \approx \frac{B_r^2}{4\mu_0}, \tag{3.5.9}$$

where B_r is the remanent field, and H_c is the coercive force of the material. Depending on material quality the maximum energy product for NdFeB is in the range of typically 240 – 340 kJ/m^3. Assuming 300 kJ/m^3 the absolute lower limit of material for the TESLA DR is about 2.8 m^3. A cross section of the wiggler aperture of $g \times w = 25 \times 50\,\mathrm{mm}^2$ is assumed here. Leakage flux and stray fields generally lead to larger volumes. We are using a wiggler design which has been optimized to have the point of operation close to the $(B \cdot H)_{max}$ value. The schematic layout of 1.5 T wiggler is shown in fig. 3.5.5. Conically shaped pole pieces are chosen which minimize leakage flux. A field distribution along the beam axis close to rectangular is produced. The effective fill factor is approximately given by $\eta_w = 1 - g_w/\lambda_w$ in order to account for the deviation from a rectangular field distribution in the transition between positive and negative field sections. This correction of about 7 % is included in the total wiggler length quoted in table 3.5.4.

Damping integral I_D	815 T^{2}m
Wiggler length	388 m
Fill factor	93 %
# of wiggler units	96
Gap height g_w	25 mm
Period length λ_w	0.4 m
Peak field B_w	1.5 T
Volume of material	4.8 m^3
Required NeFeB	36,000 kg
Chromaticity contribution $\xi_{x,y}$	-4.8, -9
norm. linear acceptance	0.06 m

Table 3.5.4: *Parameters of the wiggler section.*

Tolerances and Correction Procedures

The vertical equilibrium emittance in the damping ring is essentially determined by the quadrupole, sextupole and BPM alignment tolerances and by the orbit correction procedures applied. We assume here tolerances as listed in table 3.5.5, which are close to state-of-the-art alignment accuracies achievable at installation time. The ring with errors has been simulated using the PETROS code. Standard orbit correction with the MICADO algorithm [1] is applied in several iterations, until the rms-orbit error does not improve significantly any more. This procedure is repeated for 25 different random seeds of errors. The typical residual rms-orbit deviation after correction is 0.30 mm. The resulting vertical beam emittance at extraction time is shown in fig.

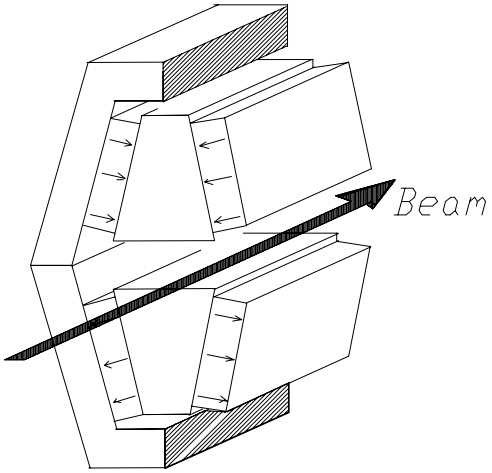


Figure 3.5.5: *Sketch of the 1.5T wiggler design.*

Element	Error type	Tolerance (rms)
Quadrupoles (arcs)	Transverse position	0.1 mm
(straight sections)	Transverse position	0.2 mm
	Roll angle	0.2 mrad
Sextupoles	Transverse position	0.1 mm
	Roll angle	0.5 mrad
Dipoles	Roll angle	0.2 mrad
BPM's vs. quads	Transverse position	0.1 mm
BPM resolution	Transverse position	0.01 mm

Table 3.5.5: *Alignment tolerances for the damping ring components.*

3.5.6. From the scattering of the simulation data, one can conclude that the design value of the emittance can be realistically achieved with just this standard method of orbit correction. By applying additional beam-based methods, it is possible to further reduce the vertical emittance. By improving the alignment of BPM's w.r.t. the focussing elements, a significant reduction of the spurious vertical dispersion can be achieved. This method has been applied at the HERA electron ring, in that case in context with spin polarization optimization [2]. Preliminary studies based on a BPM-alignment accuracy of 0.03 mm (rms) and the same orbit-kick minimization procedure used at HERA show that an equilibrium vertical emittance of $\gamma\epsilon_y \approx 2 \times 10^{-8}$ m is feasible. As an additional method, empirical emittance optimization using dispersive orbit bumps can be applied.

An important question concerns the long-term orbit stability and the typical time between necessary orbit corrections. From the experience at HERA [3], a diffusion-like orbit drift presumably caused by uncorrelated slow ground motion must be taken into account. From the HERA data we assume a ground motion according to the ATL-rule with $A \approx 10^{-5}\mu\mathrm{m}^2\mathrm{s}^{-1}$. This will cause a vertical emittance increase at a rate of $\gamma\Delta\epsilon_y/\Delta t \approx 2.5 \cdot 10^{-11}\mathrm{ms}^{-1}$. Steering the orbit back to the "golden" orbit every

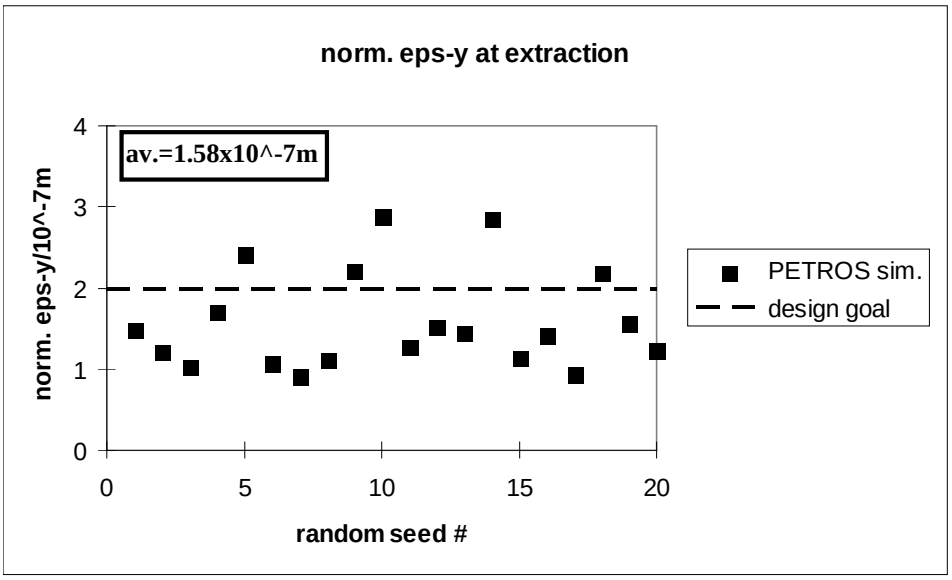


Figure 3.5.6: *Vertical emittance at extraction after applying orbit correction (see text). Results of PETROS simulations for 25 different random seeds of magnet position errors are shown. The dashed line denotes the design emittance according to table 3.5.1.*

20 min. will limit the relative emittance increase to about 4 %. This correction of an orbit change can be done at an accuracy limited only by the resolution of the BPM's, which is much better than their absolute alignment precision. In the HERA electron ring, the BPM resolution is about 20μm if the orbit is averaged over 64 revolutions. The goal of 10μm BPM resolution for the damping ring should be attainable without problems by longer averaging.

It can be concluded that from the point of view of beam optics and tolerances the design emittances in the damping ring can be safely achieved and that there is room for an improvement of the vertical emittance by about an order of magnitude.

Beam-Gas and Intra-Beam Scattering

Incoherent scattering processes set upper limits on the beam lifetime and lower limits on the beam emittances in the damping ring. The dominant contribution to vertical emittance growth is due to elastic Coulomb scattering with the gas nuclei. The rms emittance growth from this effect is given by [4]:

$$\gamma\epsilon_{y,BG} = \frac{\pi c r_e^2}{\tau_y^{-1}\gamma} < \beta_y N_{gas} > l_{sc} \qquad (3.5.10)$$

Here, N_{gas} is the number density of gas molecules and l_{sc} depends logarithmically on the ratio of maximum to minimum scattering angle ($l_{sc} \approx 5$). Due to the large beta-function, the vacuum pressure in the long straight sections is most critical. For $p_{gas} =$

10^{-9} mbar in the straight sections and $p_{gas} = 10^{-8}$ mbar (with Z=7.3, assuming CO-molecules) in the arcs and wiggler sections, we find an emittance increase of 16 % w.r.t. the design value. The rms emittance according to eq. (3.5.10) strongly overestimates the actual effect on the core of the beam, because the distribution function resulting from a small number of scattering events is very different from Gaussian [5]. Therefore, the resulting luminosity loss from beam-gas scattering in the damping ring is practically negligible.

Intrabeam scattering (IBS) has an effect on both the transverse and longitudinal emittances and can be viewed as imposing a lower limit on the 6-dimensional phase space density in the bunch. The IBS-growth rates are estimated in an approximation given in [6]. For the design emittance values from table 3.5.1 we find: $\tau^{-1}_{\epsilon,IBS} = 0.6\,s^{-1}, \tau^{-1}_{x,IBS} = 2.2\,s^{-1}, \tau^{-1}_{y,IBS} = 0.3\,s^{-1}$. This yields an increase of the horizontal emittance by 7 %, whereas the effect in the longitudinal and vertical plane can be neglected. It becomes clear that the vertical emittance can be further reduced by a large factor before the IBS-limitation sets in, especially if such a machine parameter upgrade is connected with a reduction of bunch charge, as described in section 3.1.

The most important limitation of the beam lifetime at high particle densities in the bunch is due to the Touschek effect (see e.g. [7]). This lifetime turns out to be several hours, so that the Touschek-effect is not even a limitation to operating the ring in a storage-mode, as may be desirable for commissioning and machine study purposes.

Dynamic Acceptance

The dynamic acceptance of the damping ring is investigated by particle tracking using the RACETRACK simulation code. The sextupole strengths in the arcs are adjusted for zero chromaticites. Random offsets of the sextupoles w.r.t. the beam orbit, the dominant source of distortions of the linear optics, are taken into account as well as synchrotron oscillations. The optimization of the working point is relatively uncritical, as long as the strongest resonances (up to 4th order) are avoided. An ensemle of 16 particles with different initial betatron phases is tracked over 1,000 revolutions (about 2 damping times). The resulting dynamic acceptance for different initial ratios ϵ_x/ϵ_y is shown in fig. 3.5.7. It is found to be larger than the linear acceptance, so that the non-linear effects of the sextupoles do not represent a limitation of the acceptance for the injected positron beam. Other nonlinearities can in principle also reduce the acceptance. The effect of the wigglers, with the rather pessimistic assumption of a sinusoidal field shape, is found to be insignificant. We also expect that the effects from multipole errors in the dipoles and quadrupoles can be kept small by proper design and standard manufacturing tolerances.

We conclude the section on the damping ring beam optics by summarizing the main parameters in table 3.5.6.

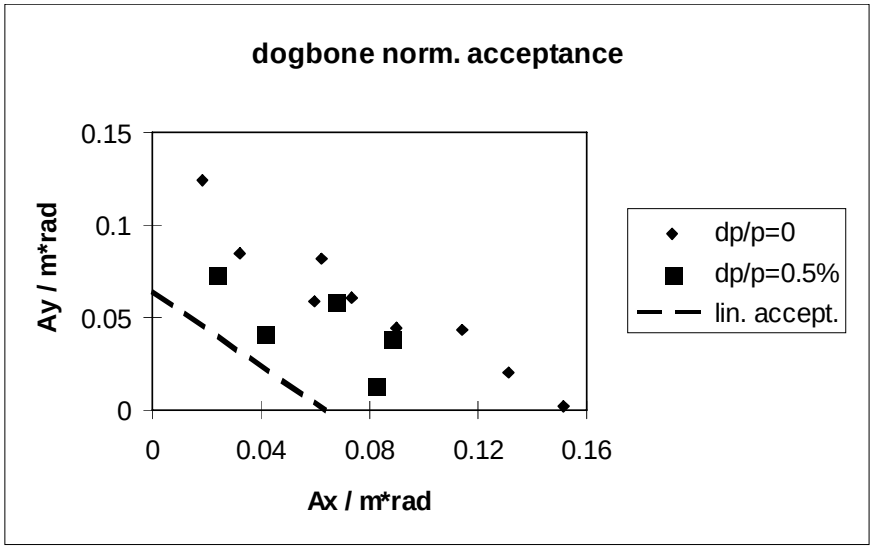


Figure 3.5.7: *Dynamic acceptance of the damping ring. The simulations include a 0.2 mm rms offset of the sextupole magnets.*

3.5.4 Rf-System and Collective Effects

According to the design parameters in table 3.5.6 the rf-system has to provide a circumferential voltage of 25 MV and an average power of 1.24 MW to the beam. The choice of frequency, f_{rf}=433 MHz, is motivated by having an integer fraction of the main linac frequency and by the availability of high power klystrons. The choice of accelerating cavities has to take into account an analysis of collective effects, in particular coupled bunch instabilities driven by higher order cavity modes (HOM).

Cavities and Coupled Bunch Instabilities

We focus here on a concept which uses 4-cell superconducting cavities very similar to the ones installed at HERA (but scaled in frequency from 500 MHz to 433 MHz). Other concepts are also conceivable and will be briefly addressed below.

We assume a power limit per cavity of about 150 kW so that in total 8 cavities are needed. The length of one 4-cell cavity is $4(\lambda_{rf}/2)$=1.39 m and the required accelerating gradient amounts to 2.25 MV/m. From the experience with the HERA-cavities including HOM-couplers, the strongest modes can be estimated to have the following impedances (per 4-cell cavity):

$$Z_{||}(f_{HOM} \approx 0.8\,\mathrm{GHz}) = 60\,rmk\Omega, Z_{\perp} = 13\,\mathrm{M\Omega/m} \tag{3.5.11}$$

The longitudinal instability growth rate due to higher order modes in the RF system can be calculated as [8]:

$$\frac{1}{\tau_{||}^{k}} = \frac{\alpha I_B f_0}{2(E/e)\nu_s} \sum_{p=-\infty}^{+\infty} (pN_b + k + \nu_s) \cdot Z_{||}(\omega_0(pN_b + k + \nu_s)) \tag{3.5.12}$$

Circumference C	17 km
Beam energy E_b	3.20 GeV
# of bunches n_b	1130
Bunch charge N_e	$3.63 \cdot 10^{10}$
Bunch spacing Δt_b	50 ns
Stored current I_b	116 mA
Norm. hor. emittance $\gamma\epsilon_x$	$1.10 \cdot 10^{-5}$ m
Transverse damping time $\tau_{x,y}$	33.8 ms
Hor./vert. tune Q_x/Q_y	59.85 / 57.71
Hor./vert. chromaticity ξ_x/ξ_y	-68.6 / -73.3
Norm. acceptance $A_{x,y}$	0.06 m
Energy loss p. turn U_0	10.7 MV
RF-frequency f_{RF}	433 MHz
RF-voltage U_{RF}	25 MV
Energy spread σ_E	0.10 %
Compaction factor α_c	$3.4 \cdot 10^{-4}$
Bunch length σ_z	9.5 mm
Synchrotron tune Q_s	0.095
Energy acceptance $\Delta E_{max}/E$	1.5 %
Incoh. Space Charge ΔQ_y	-0.17

Table 3.5.6: *Main design parameters of the positron damping ring.*

The upper limit of the growth rate is obtained for that multibunch mode with a frequency coinciding with the mode frequency of the strongest HOM. Using 3.5.11 and scaling the impedance linearly with the number of cavities, we find $\tau_{||}^{-1} = 272\,s^{-1}$. Since the damping rate from synchrotron radiation is only $60\,s^{-1}$, a feedback system is required. The systems installed in PETRA and HERA [9] work with a bunch spacing of 96 ns and provide a damping rate above 1 kHz. It is expected that a similar longitudinal feedback for the damping ring with twice the bandwidth but only about one third the damping rate of the PETRA/HERA systems can be built without problems.

For beam loading compensation, the cavity resonance frequency has to be detuned by several kHz. This can give rise to an instability of the multibunch mode nearest in frequency, with a rise time potentially smaller than the other HOM-driven modes. To avoid this, a simple feedback loop (like the "nearby-mode" feedback in HERA) is foreseen.

In the transverse plane, the instability growth rate is given by [8]:

$$\frac{1}{\tau_\perp^k} = \frac{I_B f_0}{(E/e)} \sum_{p=-\infty}^{+\infty} \beta \mathtt{Re} Z_\perp(\omega_0(pN_b + k + \nu_\perp)) \qquad (3.5.13)$$

With β=10 m at the cavities and an impedance of $8\times13\,\mathrm{M}\Omega$ for the strongest mode, the growth rate is $\tau_\perp = 648\,s^{-1}$, to be compared with a damping rate of $29\,s^{-1}$. The multibunch feedback system in PETRA and HERA is capable of providing a damping

rate of several kHz. With a similar system for the damping ring, suppression of HOM-driven transverse instabilities is guaranteed.

In the transverse plane, coupled bunch instabilities are also strongly driven by the resistive wall effect of the vacuum chamber. The transverse resistive wall impedance of the Al vacuum chamber of the ring yields $\beta \mathtt{Re} Z_\perp [M\Omega] \approx 1300/\sqrt{n}$ (dominated by the straight sections, with about 10% contribution from the wiggler and the arc sections).

The maximum growth rate occurs at the mode number which is equal to the integer part of the tune $k = [\nu_\perp]$, thus, to derive the impedance one should replace $\sqrt{n}$ by the square root of the fractional part of the transverse tune $\sqrt{\Delta\nu_\perp}$. Taking $\Delta\nu_\perp = 0.15$, we get $\tau_\perp^{-1} \approx 2000\ s^{-1}$ for the 17 km long "dog-bone" ring. This instability is much easier to damp than the one driven by the cavity-HOMs, because it requires only a rather low-bandwidth feedback. It is conceivable, though, that a broad-band feedback can be made strong enough to damp the instability driven by the resistive-wall effect as well so that an additional system may not be necessary.

With different concepts for the RF-cavities, a stronger suppression of HOMs is possible. Using about 40 single-cell Cu-resonators of the type foreseen for the PEP-II B-factory, the HOM impedance [10] could be reduced by roughly a factor of two in the longitudinal plane and a factor of five in the transverse plane. Feedback systems are still required, but with relaxed damping rates. As yet another alternative, strongly HOM-damped superconducting single-cell cavities similar to the one tested at CESR [11] could be employed. In that case, the Q-values of the HOMs may be as low as 100 and a broadband feedback could become obsolete, at least in the longitudinal plane.

Single Bunch Effects

In the longitudinal plane, the microwave instability sets a constraint on the maximum tolerable effective broadband impedance [8]:

$$\Big(\frac{Z_\parallel}{n}\Big)_{eff}^{thr} = \frac{(2\pi)^{3/2}\alpha_c (E/e)\sigma_E^2 \sigma_z}{eN_e c}, \tag{3.5.14}$$

For parameters of the dogbone damping ring we get

$$\Big(\frac{Z_\parallel}{n}\Big)_{eff}^{thr} = 86 m\Omega. \tag{3.5.15}$$

The purely inductive part of the impedance $Z_\parallel^{ind} = -i\omega L$ does not contribute to the microwave instability but still can give rise to bunch lengthening, which is approximately given by [10]:

$$\frac{\Delta\sigma_z}{\sigma_z} \approx \frac{L}{2\sqrt{2\pi}\sigma_z} \frac{N_e r_e}{C\alpha_c \gamma \sigma_E^2}. \tag{3.5.16}$$

An estimate of the different contributions to the broadband impedance is given in ref. [12]. We only summarize the results here, shown in table 3.5.7. The microwave instability is safely avoided (by an order of magnitude) and the inductive part of the impedance is expected to cause a bunch lengthening of about 10%.

Component	Z/n, mΩ	k_{loss}, V/pC
RF Cavities	2.3	22.0
Resistive wall	5.1	6.5
Total non-inductive	7.4	
Bellows	7.8	2.8
BPMs	4.3	0.8
Kickers	1.3	2.8
Other induct.	~ 4	$\sim$2
Total inductive	$\simeq 17$	
Loss factor		$\simeq$38
Loss power, kW		$\simeq$25

Table 3.5.7: *Contributions to the non-inductive and purely inductive part of the longitudinal broadband impedance according to the estimate given in ref. [12].*

The transverse broadband impedance has also been estimated in ref. [12]. It is found that the mode-coupling instability is not an issue of concern for the damping ring.

3.5.5 Injection and Extraction System

Injection into and extraction from the damping ring has to be done bunch-by-bunch in order to cope with the different bunch spacing in the linac and in the ring. This means that the full pulse length of the kicker device must be smaller than twice the bunch spacing in the ring (Δt_b(ring)=50 ns for the reference design and 25 ns for the upgrade options). Furthermore, a burst of n_b=1130 pulses spaced by Δt_b(linac)=708 ns has to be provided per injection (extraction) cycle. Assuming that the kickers and septa are installed in the straight section where the beta-function is large ($\approx$150 m), the required kick angle for beam injection with 3 standard deviations clearence at the septum bar is 0.6 mrad (the requirement for extraction is much relaxed due to the smaller beam emittance). This corresponds to a field integral of 0.006 Tm.

The technical solution for the kicker is based on a travelling wave stripline device (see ref. [13]). The effective field integral during the flat top of the pulse (see Fig. 3.5.8) is equal to $2e\hat{U}l/a$ where $\hat{U}$ is the applied voltage, l the kicker length and a the half-aperture. With $\hat{U}$=5 kV, a=50 mm and l=0.5 m, a single kicker provides a deflection strength of $3 \cdot 10^{-4}$ Tm, so that 20 of these devices are required for the extraction system. A prototype kicker has been built and tested using a high voltage burst pulser and it was shown that an effective pulse length as short as about 10 ns could be achieved [13]. With this approach, a sufficient safety margin allowing for further reduction of the bunch spacing is thus provided.

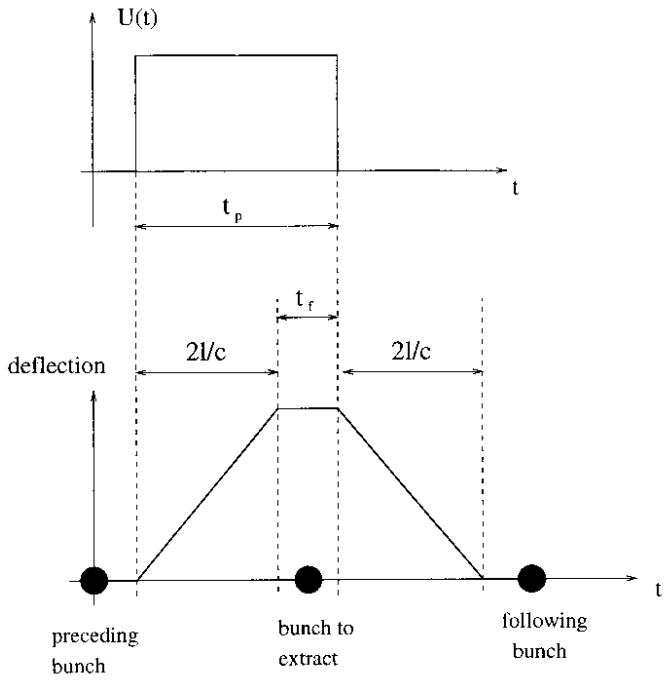


Figure 3.5.8: *Sketch of the travelling wave kicker time structure.*

3.5.6 Vacuum System

The vacuum system for the damping ring consists of three parts: two 8 km long straight sections, two arcs and four wiggler sections.

Straight Sections

For the straight sections 9.5 m long vacuum chambers will be made from 100 mm diameter aluminum tubes with aluminum CF-flanges welded to the ends. In order to reach the required average pressure of $\bar{p} = 10^{-9}$ mbar titanium sublimation pumps with 1000 l/s are installed every 10 m supplemented with small (60 l/s) ion getter pumps every 20 m. The getter pumps will be also used for pressure read out. The pumps are integrated into 0.3 m long stainless steel chambers with bellows, which are installed between the long aluminum chambers as indicated in fig. 3.5.9. The pump ports and the bellows are rf-shielded by CuBe spring fingers. The flange connection between aluminum-aluminum and aluminum-stainless steel flanges are done by CF-aluminum gaskets.

A third type of chamber is used for the quadrupoles located every 50 m. These 1 m long aluminum or copper coated stainless steel chambers are equipped with beam position monitors (BPM). After mounting them rigidly to the quadrupole, the BPM's will be calibrated with respect to the quadrupole axis. In total 1520 long aluminum chambers, 1520 pump units with bellows and 304 aluminum quadrupole chambers are needed for the straight sections.

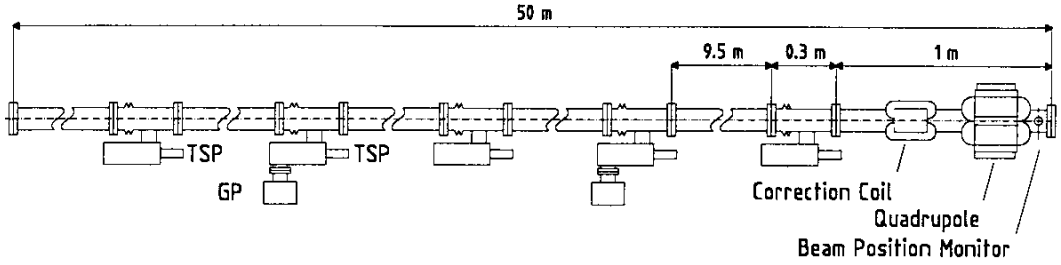


Figure 3.5.9: *Pattern for the vacuum chambers in the straight sections of the TESLA damping ring.*

Pneumatic valves with rf-shields segment the straight sections into several long vacuum sectors. To start pumping and for leak check manual valves are foreseen at some of the ion getter pumps to connect movable pump stations.

The vacuum chamber in each 8.5 m arc cell is split in the following way: two long chambers passing the dipole magnet and the correction coil and two short chambers through the quadrupole and sextupole magnets. All chambers will have the same elliptical aperture of $52 \times 28\,\mathrm{mm}^2$. The arrangement of the vacuum chambers is shown in fig 3.5.10.

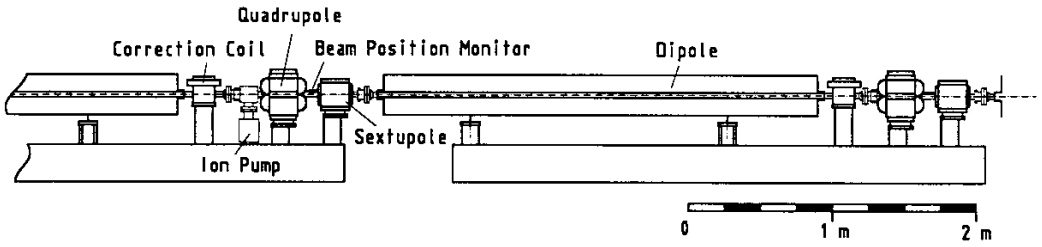


Figure 3.5.10: *Arrangement of the vacuum chambers in the arcs of the damping ring.*

The required vacuum of $p = 10^{-8}$ mbar for the arcs is somewhat less stringent than in the straight sections. However, due to the synchrotron radiation of 170 W/m the gas load from desorption of gas molecules from the chamber walls will be much higher. Therefore the aluminum vacuum chamber in the 3 m long dipole magnets has a separate channel for integrated pumps, e.g. NonEvaporable Getter (NEG) vacuum pumps. Fig. 3.5.11 shows a tentative layout of the profile of this chamber including a channel for water cooling. The chamber will be made by extrusion and bent according to the bending radius of the dipole. The integrated pumps are mounted on a support, which also serves as rf-shield of the pump channel as indicated in fig. 3.5.11. This structure will be slid into the aluminum chamber. The pump channel is terminated on both sides of the dipole magnet. Straight elliptical aluminum tubes with transition into

round aluminum CF-flanges are welded on both sides. The water channel is bridged between two dipole magnets.

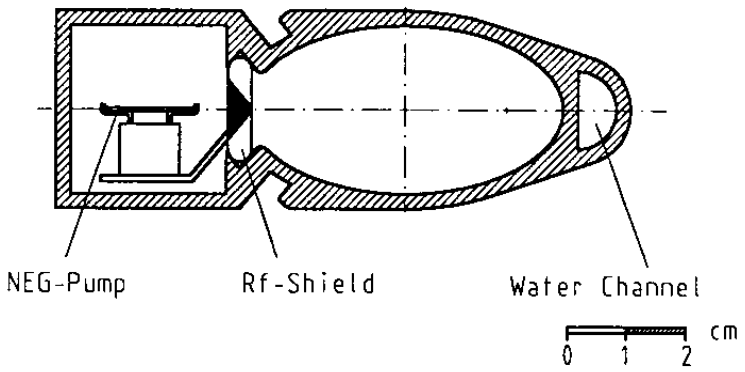


Figure 3.5.11: *Tentative layout of the vacuum chamber through the dipole magnets in the arcs of the damping ring. Integrated pumps are installed in a separate pump channel. The chamber is water cooled.*

Each quadrupole must be equipped with a BPM. Therefore short straight stainless steel or copper chambers of the same elliptical profile as the other chambers, formed by pressing of a tube, are rigidly mounted in the quadrupoles. The BPM's will be calibrated with respect to the quadrupole axis. Bellows are integrated into these chambers on both sides of the quadrupoles. Small ion getter pumps (60 l/s) attached to pump ports in every second unit supplement the integrated pumps. As in the straight sections rf-shielding of the bellows and pump ports is done by CuBe spring fingers. In case of a stainless steel chamber the remaining part of the pipe will be copper coated.

The magnets and vacuum chambers of each half cell are mounted on one girder, prealigned and installed as one unit in the tunnel for the arcs.

In each arc one pneumatic valves with rf-shield will split the vacuum system into two 200 m long sectors to facilitate activation and regeneration of the NEG pumps. Movable pump stations can be connected through manual valves attached to some ion getter pumps to start pumping.

Wiggler Sections

The four wiggler sections are each 120 m long. The vacuum chamber in each 10.5 m wiggler cell is split into 4 m long chambers through the wiggler magnets and 1 m long chambers through the quadrupole magnets and correction coils. About 15 kW of synchrotron radiation will be deposited in the vacuum chamber of each half cell.

The aluminum chamber through the wiggler magnets has a beam chamber with vertical aperture of about 20 mm and an antechamber on either side with a slot height of 25 mm as shown in fig. 3.5.12. Vacuum pumping will be performed by NEG pumps

located at the end of each slot. The antechamber slots are sloped to protect the pumps against direct synchrotron radiation. The chamber will be water cooled .

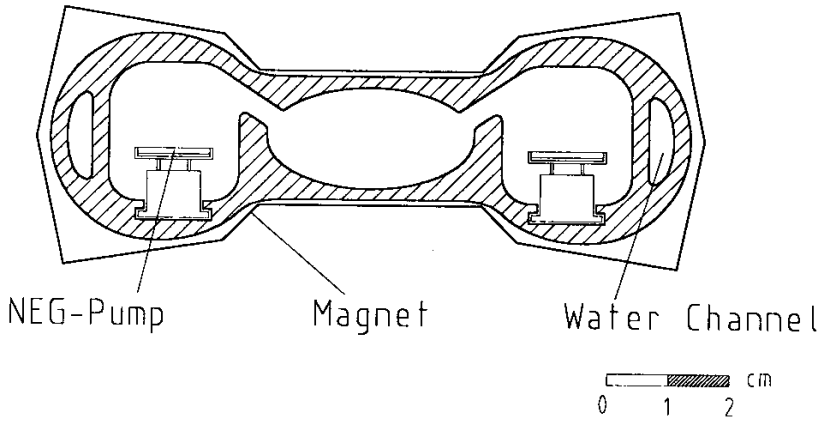


Figure 3.5.12: *Schematic of the cross section of the wiggler vacuum chamber for the damping ring.*

The vacuum chamber through the correction coil and the quadrupole magnet is equipped with BPM's, bellows and a small ion getter pumps (60 l/s) in every second unit to supplement the integrated pumps. Again, rf-shielding of the bellows and pump ports will be done by CuBe spring fingers. In front of the quadrupole magnets absorbers are needed to shield the synchrotron radiation. Pneumatic valves on both sides of each wiggler will separate the wiggler vacuum from the arcs and straight sections.

Bibliography

[1] B. Autin and Y. Marti, *Closed Orbit Correction in A.G. Machines Using a Small Number of Magnets*, CERN-ISR-MA-73-17, 1973.

[2] E. Gianfelice et al., *Application of a Beam-Based Alignment Technique for Optimizing the Electron Spin Polarization at HERA*, Proc. 5th Europ. Part. Acc. Conf., Barcelona 1996.

[3] R. Brinkmann and J. Roßbach, *Observation of Closed Orbit Drift at HERA Covering 8 Decades of Frequency*, Nucl. Instr. Meth. A350, p. 8, 1994.

[4] T. O. Raubenheimer, *The Generation and Acceleration of Low Emittance Flat Beams for Future Linear Colliders*, Thesis, SLAC-387, 1991.

[5] T. O. Raubenheimer, *Emittance Growth due to Beam-Gas Scattering*, KEK report 92-7, 1992.

[6] A. Piwinski, *Intrabeam Scattering with Vertical Dispersion*, Proc. SSC Workshop, Ann Arbor, MI, 1983.

[7] H. Wiedemann, *Low Emittance Storage Ring Design*, Proc. US/CERN Part. Acc. School, South Padre Island, TX, 1986.

[8] A. Chao, *Physics of Collective Beam Instabilities in High Energy Accelerators*, John Wiley & Sons, Inc., 1993.

[9] M. Ebert et al., *Transverse and Longitudinal Multi-Bunch Feedback Systems for PETRA*, DESY-91-036, 1991.

[10] S. Heifets et al., *Impedance Study for the PEP-II B-Factory*, SLAC/AP-99, 1995.

[11] S. Belomestnykh et al., *Wakefields and HOMs Studies of a Superconducting Cavity Module with the CESR Beam*, IEEE Proc. Part. Acc. Conf., Vol. 5, p. 3391, Dallas 1995.

[12] V. Shiltsev, *Beam Dynamics Issues in the TESLA Damping Ring*, DESY-TESLA-96-02, 1996.

[13] B.I. Grishanov, F.V. Podgorny, J. Rümmler and V.D. Shiltsev, *Very Fast Kicker for Accelerator Applications*, DESY-TESLA-96-11, 1996.

3.6 Bunch Compressor

3.6.1 Introduction

A detailed design for a single stage beam bunch length compressor for the TESLA Linear Collider is presented. Compression is achieved by introducing an energy-position correlation along the bunch with an RF section at zero-crossing phase followed by a short bending section with energy dependent path length (momentum compaction). The motivation for a wiggler design is presented and many of the critical single bunch tolerances are evaluated. A solenoid based spin rotator is included in the design and transverse emittance tuning elements, diagnostics and tuning methods are described. Bunch length limitations due to second order momentum compaction and sinusoidal RF shape are discussed with options for compensation.

3.6.2 Bunch Compressor Design Issues

The bunch compressor design is influenced by several criteria:

- The compressor must reduce the bunch length extracted fom the damping ring to the appropriate size in the linac.
- The system must perform a 90 degree longitudinal phase space rotation so that damping ring extracted phase errors do not translate into linac phase errors which can produce large final beam energy deviations.
- The system must not significantly dilute the transverse emittances and should include tuning elements for correction.
- With its low energy and initially small energy spread, the bunch compressor is a convenient place to include solenoids for full control of the spin orientation.
- The compressor should be short, simple and as error tolerant as possible.

3.6.3 Bunch Compressor Parameters

The rms bunch length at the entrance to the linac (for uncorrelated Damping Ring energy spread σ_{δ_0}) is

$$\sigma_{zl} = \sqrt{(1+\alpha k)^2\sigma_{z_0}^2 + \alpha^2\sigma_{\delta_0}^2} \tag{3.6.1}$$

with $\alpha \equiv \int \frac{\eta(s)}{\rho(s)} ds$ and $k = \frac{2\pi e V_0}{\lambda E_0}$, where η and ρ are dispersion and bending radius along the beam line while V_0 and λ are amplitude and wavelength of the RF system. E_0 is the beam energy.

The minimum bunch length is achieved when $\alpha = -1/k$ which also decouples the linac phase from the DR phase. The momentum compaction is chosen to achieve the desired linac bunch length and the rf voltage is then given:

$$\alpha = -\frac{1}{k} = \frac{\sigma_{z_l}}{\sigma_{\delta_0}} \qquad eV_0 = -\frac{\lambda_{rf} E_0}{2\pi\,\alpha}. \tag{3.6.2}$$

parameter	symbol	unit	
Energy	E_0	GeV	3.2
rms horizontal emittance	ϵ_x	mm-mrad	8.0
rms vertical emittance	ϵ_x	mm-mrad	0.13
rms DR bunch length	σ_{z_0}	mm	9
rms DR energy spread	σ_{δ_0}	%	0.1
rms linac bunch length	σ_{z_l}	mm	0.6
rms linac energy spread	σ_{δ_l}	%	1.4
rf wavelength	λ_{rf}	m	0.231
rf voltage	V_0	MV	190
rf gradient	G_{rf}	MV/m	25
length of rf section	L_{rf}	m	7.6
momentum compaction	α	m	0.64

Table 3.6.1: *Bunch compressor parameters. Emittance values represent those immediately after DR extraction and are smaller than the design values at the IP in order to have a safety margin.*

The resulting energy spread at the linac entrance is approximately amplified by the inverse of the bunch compression factor.

$$\sigma_{\delta_l} = \sigma_{\delta_0}\sqrt{\frac{\sigma_{z_0}}{\sigma_{z_l}} + 1} \approx \sigma_{\delta_0}\frac{\sigma_{z_0}}{\sigma_{z_l}} \quad , \quad (\sigma_{z_l} \ll \sigma_{z_0}) \tag{3.6.3}$$

The damping ring (DR) and initial linac parameters relevant for bunch compressor design are listed in Table 3.6.1.

3.6.4 Bunch Compressor Optics

A simple way to produce the necessary momentum compaction is with a chicane made of four horizontal bending magnets. A chicane of total length 24 m for the TESLA design is possible, however, the peak dispersion required is large at 1.8 m. With 1.5 % rms energy spread the horizontal beam size becomes very large (σ_x=27 mm) and dipole field quality tolerances become tight. A more error tolerant design is achievable with a wiggler section [1] made up of N_p periods each with two bending magnets and two quadrupoles-one at each zero crossing of the dispersion function. If the bends are rectangular and each of length L_B with angle θ_B and separated by ΔL, and the quadrupoles alternately reverse sign to produce a phase advance per cell of $\Psi_x(=\Psi_y)$, the momentum compaction, α_w, and the horizontal emittance growth due to synchrotron radiation (SR), $\Delta\gamma\epsilon_x$, are given by [1]:

$$\alpha_w = \frac{1}{4}\theta_B^2[N_p(2\Delta L + L_B) - \frac{1}{6}(2N_p + 1)L_B] \tag{3.6.4}$$

$$\Delta\gamma\epsilon_x \approx (1.25 \cdot 10^{-8} m^2 \cdot GeV^{-6}) \cdot E_0^6 N_p \frac{|\theta_B|^5(\Delta L + L_B)}{L_B^2 |sin\Psi_x|} \tag{3.6.5}$$

parameter	symbol	unit	
momentum compaction	α_w	meters	0.600
dipole magnet length	L_B	meters	2.0
bend-to-bend separation	D_L	meters	3.22
bend angle / dipole	ϕ_B	degrees	16.0
total length	L_{tot}	meters	47
maximum x-dispersion	η_{max}	mm	308
SR x-emittance dilution	$\Delta\epsilon_x/\epsilon_{x0}$	%	1.8
dipole magnetic field	B_0	kG	15.37
quad. pole tip fields - F (D)	B_Q	kG	2.0 (1.7)
number of periods	N_p	-	4
phase advance per period	q_x	degrees	90

Table 3.6.2: *Wiggler design parameters (NOTE: All non-correction quadrupoles in the entire beamline are of length 15 cm and pole-tip radius 1 cm).*

By setting the necessary momentum compaction [2] and choosing a horizontal SR emittance dilution of 1-2%, a dipole field strength of 15 kG, and roughly minimizing dipole magnet length and peak dispersion, the wiggler design parameters listed in Table 2 were developed. With $|\sin\Psi_x|^{-1}$ minimized at $\Psi_x = \pi/2$, the number of periods is chosen so that each dipole at peak dispersion has an opposing dipole separated by π in betatron phase advance: all orders of dispersion generated by equal multipole field components cancel. The peak horizontal beam size in the wiggler is 4.3 mm. The β and dispersion functions and a schematic beamline layout of the wiggler (including a 12 meter, 8 cavity TESLA superconducting rf module) are plotted in Fig. 3.6.2.

3.6.5 Peripheral Sections and the Full Beamline

In addition to the primary function of bunch compression, the beamline described here also includes sections for spin rotation, cross-plane coupling correction, and phase space diagnostics. This section describes these modules (for reference, the entire beamline is depicted in Fig. 3.6.2).

3.6.5.1 Spin Rotator

Using the TESLA damped vertical emittance, it can be shown [3] that for a chicane type "half serpent" $\pi/2$ spin rotator [4] constructed from horizontal and vertical bending magnets, and for $< 2\%$ SR vertical emittance dilution, the relationship between the length of the bending magnets and the beam energy is discouraging: $L[m] > 30 \cdot E[GeV]$. Even at 3.2 GeV the dipoles are about 100 m long. Therefore a spin rotator based on superconducting solenoids is proposed prior to bunch compression. Since the damped beam is flat ($\epsilon_y/\epsilon_x \ll 1$) the cross-plane coupling induced by the solenoids must be compensated. This is achieved by rotating the spin by $\pi/2$ around the longitudinal

parameter	symbol	unit	
momentum compaction	α_w	meters	0.600
effective solenoid length	L_{sol}	meters	2.275
maximum solenoid field	B_z	kG	38.0
total arc bend angle	ϕ	degrees	12.0
total length of rotator system	L_{tot}	meters	52
number of dipole magnets (12 degree arc)	N_{dipole}	-	4
dipole field	B_0	kG	6.1
SR x-emittance dilution	$\Delta\epsilon_x/\epsilon_{x0}$	%	$1 \cdot 10^{-3}$
momentum compaction of all arcs	α_a	m	0.02
chromatic ϵ_y dilution ($\sigma_{\delta 0}$) = 0.1%)	$\Delta\epsilon_y/\epsilon_{y0}$	%	0.7
chromatic ϵ_y dilution ($\sigma_{\delta 0}$) = 0.2%)	$\Delta\epsilon_y/\epsilon_{y0}$	%	4.6
chromatic ϵ_y dilution ($\sigma_{\delta 0}$) = 0.3%)	$\Delta\epsilon_y/\epsilon_{y0}$	%	10.9

Table 3.6.3: *Spin rotator parameters*

axis with a pair of $\pi/4$ solenoids separated by a 'reflector' beamline which causes the solenoid pair's coupling to cancel while their spin precession adds. In this way the solenoid coupling is always canceled regardless of the solenoid settings as long as the solenoid pairs have equal strength ($\pm 2\%$). By separating two such solenoid pair systems with a $\pi/2$ spin rotation around the vertical axis (i.e. an arc of total bend angle $\theta = \pi/2\alpha\gamma$) full arbitrary control of the spin orientation is possible. The focusing effect of the solenoids is corrected with four matching quadrupoles per paired solenoid section. The matching quadrupoles are positioned between the paired solenoid sections and the central arc (see Fig. 3.6.2). At $\sim 0.1\%$ rms DR energy spread the emittance dilution due to chromatic coupling of the four solenoids together at maximum field is negligible at 0.7%. The parameters of the spin rotator system are given in Table 3.6.3.

3.6.5.2 Coupling Correction Section

In order to empirically correct anomalous cross-plane coupling due to either damping ring extraction errors or spin rotator errors, a skew correction section (SCS) is included immediately following the final rotator solenoid. The system is constructed from four small skew quadrupoles with zero nominal strength. The first and second skew quadrupoles are separated by betatron phases of $\Delta\Psi_x = \Delta\Psi_y = \pi/2$. This is also true for the third and fourth skew quadrupoles. The second and third are separated by $\Delta\Psi_x = \pi, \Delta\Psi_y = \pi/2$ such that the four skew quadrupoles orthogonally control the four coupling coefficients in the beam: $\langle x'y\rangle$, $\langle xy'\rangle$, $\langle xy\rangle$ and $\langle x'y'\rangle$, respectively. The four skew corrections are actually orthonormal in that the emittance sensitivity to each skew quadrupole is designed to be a constant (i.e. $\beta_x\beta_y$ is equal at each skew quadrupole). The relative emittance increase for one thin skew quadrupole of focal

length f at beta functions of $\beta_x\beta_y$ is

$$\frac{\epsilon_y}{\epsilon_{y_0}} = \sqrt{1 + \frac{\epsilon_{x_0}}{\epsilon_{y_0}}\frac{\beta_x\beta_y}{f^2}}. \tag{3.6.6}$$

With skew quadrupoles of 10 cm length, 1 kG pole tip fields and 1 cm radii, a factor of about two in emittance increase can be corrected per skew quadrupole.

3.6.5.3 Diagnostic Section

It is important to provide measurement capability after the bunch compressor so that the beam emittance and matching can be monitored and corrected before entrance into the main linac. Following the wiggler section is a simple set of three FODO cells with phase advance per cell of $\Delta\Psi_x = \Delta\Psi_y = \pi/4$. By placing four wire scanners [5] near the vertical focusing quadrupoles the ideal four-wire scanner phase sampling is available to make high precision emittance and beta function measurements in each plane. The nominal rms beam size at the scanners is chosen as 15 μm vertically and 80 μm horizontally. A matched beam is easily identified since in this case the beam sizes per plane are identical at each of the four wire scanners. The cross-plane coupling in the beam is not fully measurable with this section. However, by simply minimizing the vertical emittance (flat beam) with the four SCS skew quadrupoles all coupling can be corrected (with some iteration necessary in extreme cases). A direct measurement capability is achievable, if desired, by using six wire scanners with phase advances similar to the SCS described above [6].

3.6.5.4 Compressor RF Section

The energy-position correlation is introduced with one 12 meter superconducting TESLA accelerating module with eight cavities each of 1.035 meter length with gradient of 25.0 MV/m in order to produce the necessary 190 MV for full bunch compression.

3.6.5.5 The Combined Beamline

The four sections described thus far (spin rotator, coupling correction, bunch compressor and diagnostics) are combined using various appropriate matching quadrupoles to preserve the periodic beta functions in the various FODO cells. Fig. 3.6.2 shows the beta and dispersion functions of the combined system. The combined system floorplan is laid out in Fig. 3.6.1. The system, including the 3.1 meter spin rotator bump, should fit inside a 5 meter diameter tunnel. The angle of the first arc may be reduced and the angle of the third arc equally increased if it is necessary to horizontally displace the outgoing beamline relative to the incoming one. It is only necessary that the 12 degree arc remains fixed (for 3.2 GeV) so that the electron or positron spin orientation is fully controllable over a sphere.

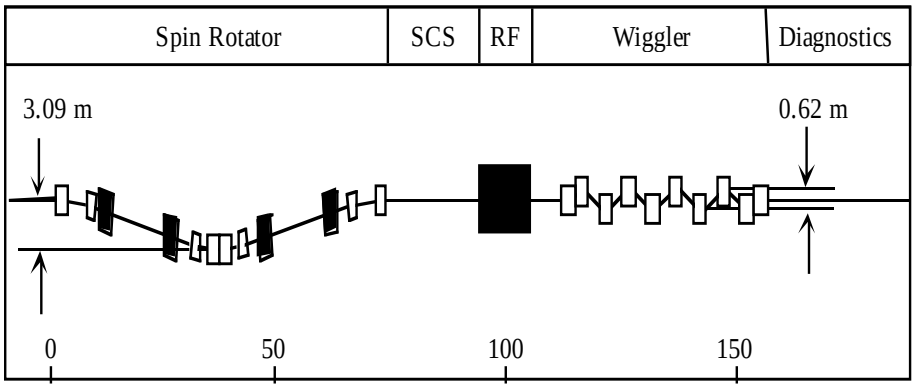


Figure 3.6.1: *Schematic layout of complete beamline floorplan. Solenoids and the RF-module are in black, dipole magnets are in white (quadrupole magnets not shown).*

3.6.6 Tolerances

3.6.6.1 Single Bunch Tolerances

The large energy spread and strong bending of the wiggler present some challenges for transverse emittance preservation, especially in the vertical plane. Several important tolerances are examined below. Table 4 evaluates these tolerances:

- A rolled dipole magnet introduces vertical dispersion which, if not corrected, will chromatically filament in the main linac and dilute the vertical emittance.

- A quadrupole (sextupole) field component in a dipole with horizontal dispersion will dilute the horizontal emittance through the generation of first (second) order horizontal dispersion. The symmetry of the wiggler optics introduces a strong cancellation of these two effects if the dipole field errors are similar for all magnets. This tolerance, when applied to the wiggler dipoles, can therefore be considered as a constraint on the field uniformity over the eight inner wiggler dipoles (the first and tenth have near zero dispersion).

- A misaligned quadrupole will generate vertical and/or horizontal dispersion.

- A rolled quadrupole in a non-dispersive section will couple the beams while a rolled quadrupole in a dominantly dispersive section will generate vertical dispersion. Table 4 lists a range of roll tolerances with the tightest tolerance on one of the matching quadrupoles just after the RF section and also on the quadrupole in the center of the 12 degree arc (see Fig. 3.6.2). The loosest tolerances are on the wiggler quadrupoles which are weak.

- There are also tolerances on the difference of the field strengths between the paired solenoids and also on the absolute gradient of the eight quadrupoles which make up the reflector beamline between the paired solenoids. These tolerances

tolerance	symbol	unit	
wiggler dipole magnet roll tolerance	$\|\phi\|$	μrad	80
quad. field in wiggler dipole (r0=20 mm)	$\|b1/b0\|$	-	$4.0 \cdot 10^{-5}$
sext. field in wiggler dipole (r0=20 mm)	$\|b2/b0\|$	-	$1.4 \cdot 10^{-4}$
F-quad. x-misalignment tolerance	$\|\Delta_x\|$	μm	460
D-quad. y-misalignment tolerance	$\|\Delta_y\|$	μm	70
Quadrupole roll tolerance range	$\|\phi\|$	mrad	3.5 - 14
Paired solenoid field difference tolerance	$\|\Delta b_z/B_{z-max}\|$	%	2
reflector quad. grad. tol. range (max. sol.)	$\|\Delta G/G_0\|$	%	0.4 - 1.7
compression RF phase stability tolerance	$\|\psi_{RF}\|$	degrees	0.33

Table 3.6.4: *Single element tolerances for TESLA wigglers and spin rotators for* $< 2\%$ *emittance dilution or* $|\langle z_l \rangle| < \sigma_{z_l}/3$ *longitudinal stability at the IP (no correction assumed).*

have been evaluated with tracking and are included in Table 3.6.4. These tolerances assume no correction. The reflector beamline quadrupole gradient tolerances are evaluated at the maximum solenoid field and vary over the eight quadrupoles. The tightest of the eight are for the two quadrupoles which are most centered between the paired solenoids.

- If the phase, ψ_{RF}, of the compression RF were to deviate from the zero crossing this energy offset will advance or delay the phase of the bunch as it enters the linac and upset the collision timing at the interaction point (IP).

The roll tolerances are quite tight and probably difficult to achieve. However, these tolerances can be relaxed by an order of magnitude by including small normal and skew quadrupoles of zero nominal field to be used for dispersion correction.

3.6.6.2 Damping Ring Extraction Phase Tolerances and Limits for Bunch Compression

The compression arguments presented thus far are limited to linear effects. In fact the bunch is not infinitely compressible and at some point second order momentum compaction limits the achievable bunch length. the second order momentum compaction, α_2, the linac bunch position of a particle has a quadratic dependence on its DR bunch position.

$$z_l = [\alpha + \alpha_2\delta_0]\delta_0 + (1 + k[\alpha + \alpha_2\delta_0])z_0 + \alpha_2 k^2 z_0^2 \approx \alpha\delta_0 + \alpha_2 k^2 z_0^2 \tag{3.6.7}$$

The approximation made in eq. 3.6.7 is possible since $\alpha_2\delta_0 \ll \alpha$ and $k = -1/\alpha$. For a wiggler, the approximate relationship between the second and first order momentum compaction is

$$\alpha_2 \approx -\frac{3}{2}\alpha. \tag{3.6.8}$$

For a DR bunch length distribution which is Gaussian (i.e. $\langle z_0^4 \rangle = 3\langle z_0^2 \rangle^2$) and 'full compression' ($k = -1/\alpha$), the linac rms bunch length is

$$\sigma_{z_l}^2 \approx \langle z_l^2 \rangle - \langle z_l \rangle^2 = \alpha^2 \sigma_{\delta_0}^2 + \frac{9}{2} \cdot \frac{\sigma_{z_0}^4}{\alpha^2} \tag{3.6.9}$$

If α is reduced in order to compress the bunch toward zero, the second term on the r.h.s. of eq. 3.6.9 begins to dominate the final bunch length. The minimum achievable bunch length is given by:

$$\sigma_{z_l - min} = \sqrt{3\sqrt{2}\sigma_{\delta_0}}\sigma_{z_0} \tag{3.6.10}$$

In addition to limiting the achievable bunch length, the second order momentum compaction also gives rise to a quadratic linac "phase" dependence on DR "phase". Thus, DR extraction "phase" errors, $\langle z_0 \rangle$, will transform into final beam energy deviations roughly depending on the mean accelerating RF phase in the main linac, $\langle \psi \rangle$ (=0 at RF crest). For small "phase" errors, $z_l \ll \lambda$, and including only the RF curvature in the main linac, final beam energy deviations, $\langle \delta_E \rangle$, can be approximated by

$$\langle \delta_E \rangle \approx -\frac{2\pi \tan\langle \psi \rangle}{\lambda} \cdot \langle z_l \rangle. \tag{3.6.11}$$

The DR extraction "phase" jitter tolerance is then:

$$|\langle z_0 \rangle| < \sqrt{\frac{|\langle \delta_E \rangle| \lambda \alpha}{3\pi |\tan\langle \psi \rangle|}}. \tag{3.6.12}$$

The necessary stability of the longitudinal beam position at the interaction point (IP) also sets a tolerance on DR "phase" jitter. The DR "phase" jitter tolerance for fraction bunch position IP stability of $|\langle z_l \rangle| < \sigma_{z_l}/n$ (e.g. n =3) is

$$|\langle z_0 \rangle|_{tol} < \sqrt{\frac{2\alpha\sigma_{z_l}}{3n}} \tag{3.6.13}$$

The bunch length is also limited by the sinusoidal character of the compressor RF. Including the 3rd order term of the $\sin(z_0)$ expansion and not including the second order momentum compaction, the bunch position of a particle in the linac, for full compression where $k = -1/\alpha$, is

$$z_l \approx \alpha\delta_0 + \frac{1}{6}\left(\frac{2\pi}{\lambda}\right)^2 \cdot z_0^3, \tag{3.6.14}$$

and for a DR bunch length distribution which is initially gaussian, the rms bunch length is

$$\sigma_{z_l}^2 \approx \alpha^2 \sigma_{\delta_0}^2 + \frac{5}{12} \cdot \left(\frac{2\pi}{\lambda}\right)^4 \cdot \sigma_{z_0}^6 \tag{3.6.15}$$

parameter	symbol	unit	
linac linear bunch length	σ_{zl}	mm	0.56
linac frequency	f_{RF}	GHz	1.3
linac mean acc. phase	$\langle\psi\rangle$	degrees	3.3
The following include the effect of second order momentum compaction only.			
linac rms 2nd order bunch length	σ_{zl}	mm	0.63
min. 2nd order linac bunch length	σ_{zl-min}	mm	0.56
DR "phase" tol. for $\langle\delta_E\rangle < 0.01\%$	$\lvert\langle z_0\rangle\rvert_{tol}$	mm (deg)	5 (8)
DR "phase" tol. for $\lvert z_l\rvert/\sigma_{z_l} < 1/3$	$\lvert\langle z_0\rangle\rvert_{tol}$	mm (deg)	9.2 (14)
The following include the effect of the sinusoidal RF only.			
linac rms sinusoidal bunch length	σ_{z_l}	mm	0.65
min. rms sinusoidal linac b.l.	σ_{z_l-min}	mm	0.34
DR "phase" tol. for $\langle\delta_E\rangle < 0.01\%$	$\lvert\langle z_0\rangle\rvert_{tol}$	mm (deg)	8 (12.5)
DR "phase" tol. for $\lvert z_l\rvert/\sigma_{z_l} < 1/3$	$\lvert\langle z_0\rangle\rvert_{tol}$	mm (deg)	12 (18)
The following includes both non-linear bunch length limitations.			
linac rms bunch length	σ_{z_l}	mm	0.70

Table 3.6.5: *Bunch length limits and DR phase jitter tolerances, due to second order momentum compaction and sinusoidal compressor RF each taken in isolation. The DR extracted bunch length distribution is taken as a Gaussian for these calculations. The phase tolerances specified in deg. are w.r.t. the compressor-RF. They are smaller by a factor of three for the damping ring RF-system.*

In this case the minimum achievable bunch length is in the direction $\alpha \to 0$. The DR extraction "phase" jitter tolerance due to the sinusoidal RF is given by

$$|\langle z_0\rangle|_{tol} < \frac{\lambda}{2\pi}\left(\frac{6|\delta_E|}{|\tan\langle\psi\rangle|}\right)^{\frac{1}{3}} \tag{3.6.16}$$

Required IP longitudinal stability sets a DR "phase" jitter tolerance of

$$|\langle z_0\rangle|_{tol} < \left(\frac{6\sigma_{z_l}}{n}[\frac{\lambda}{2\pi}]^2\right)^{\frac{1}{3}} \tag{3.6.17}$$

Bunch length limitations and DR phase jitter tolerances for both second order momentum compaction and sinusoidal RF effects are summarized below in Table V. It can be shown that the two non-linear bunch length limitations discussed above approximately add in quadrature. The net rms bunch length, including both effects, is listed at the end of the table. Note, with these non-linear effects the rms bunch length results are somewhat dependent on the initial DR extracted bunch distribution-taken as a Gaussian here.

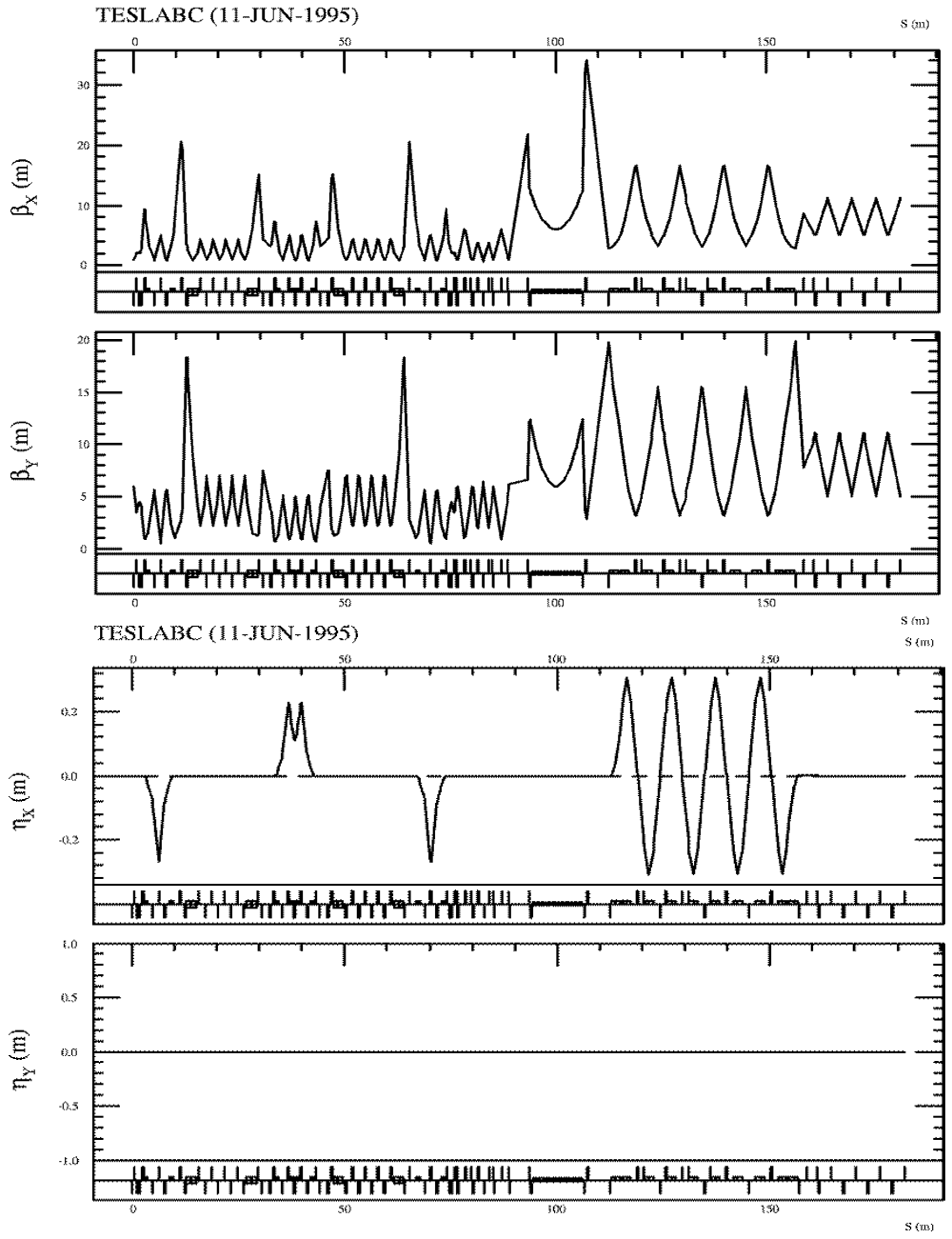


Figure 3.6.2: *β and dispersion functions for wiggler and RF section. Dipole magnets are shown as low rectangles biased above the beam line while the eight RF structures are centered. Horizontal focussing quadrupoles are biased above, vertical focusing quads below the beam line.*

Bibliography

[1] T. O. Raubenheimer, P. Emma, S. Kheifets, *Chicane and Wiggler Based Bunch Compressors for Future Linear Colliders*, Proc. of the 1993 Particle Accelerator Conference, Washington, D.C., 1993.

[2] The momentum compaction has been chosen slightly smaller than the linear design would require in order to somewhat compensate for RF non-linearities.

[3] P. Emma, *A Spin Rotator System for the NLC*, NLC-NOTE 7, December 1994.

[4] T. Fieguth, *Snakes, Serpents, Rotators and the Octahedral Group*, SLAC-PUB-4195, January, 1987.

[5] The term "wire scanner" may refer to a laser wire rather than a metallic filament, depending on survivability details. However, it will be prudent to include simple, well tested metallic filament type wire scanners for single bunch measurement to initially commission the beamline.

[6] P. Emma, *A Skew Correction and Diagnostic Section for Linear Colliders*, NLC-NOTE in preparation.

3.7 Beam Delivery System

3.7.1 Introduction

The Beam Delivery System (BDS) is the transfer line which transports the beam from the linac exit to the interaction point (IP) of the collider. Its main function is to demagnify the beams and to bring them to collisions with spot sizes down to 845 nm $\times$ 19 nm at the center of the detector. The overall 67×167 demagnification from linac to IP is mostly produced by a strongly focusing quadrupole doublet 3 m distant from the IP. Together with a weaker upstream doublet, it composes a final transformer (FT), or telescope, with 15×50 demagnification.

Even with the small 34 μrad $\times$ 27 μrad beam angular divergences at the IP corresponding to the design emittances listed in Table 3.1.2, the chromaticity of this last doublet, which is roughly given by the ratio $\hat{\sigma}_{\rm Q}/\sigma^*$ of the maximum beam size in the last doublet to the beam size at the IP, is too large in comparison with the expected 10^{-3} incoming beam relative energy spread. The FT is therefore preceded by a Chromatic Correction Section (CCS) where sextupoles, located in dispersive regions, create the necessary quadrupole correction for the off-energy particles. The energy bandwidth of this correction determines the energy acceptance of the whole system.

The optics of the Final Focus System (FFS), namely the combination of the Chromatic Correction Section and the Final Transformer, derives from the one of the SLAC Linear Collider FFS [1] with one essential difference consisting in separate, rather than interleaved, horizontal and vertical chromaticity correction sections. The non-interleaved CCS optics which requires more space but in return reduces considerably the amount of non-linear aberrations, has been adopted and operated successfully by the Final Focus Test Beam (FFTB) facility [2] at SLAC.

The transverse beam tails passing through the doublet quadrupoles produce synchrotron radiation which in turn can be a source of background in the detector. The photon energies and angles are maximum in the last doublet where masking is difficult. As seen in Fig.3.7.1, the collimation requirements for a complete clearance of the interaction region is about $14\,\sigma_x \times 46\,\sigma_y$. The beam halo, possibly generated in the linac, is therefore scraped off in the collimation section (CS) by an arrangement of spoilers and absorbers described in Section 3.7.8.

A short Linac Matching Section (LMS) matches the beam from the linac exit to the entrance of the collimation section.

The essential ability to measure the beam emittances after the linac and before final focus system and to correct the incoming transverse coupling is provided by the Tuning and Diagnostic Section (TDS), described in Section 3.7.5, following the Collimation Section.

Finally, changing the beam sizes and angular divergences at the IP in order to optimize the collision parameters without modifying the FT and CCS sensitive optics can be done in the Beta Matching Section (BMS), located between the TDS and the CCS.

After colliding head-on, the spent beams are deflected by an electrostatic separator

at the exit of the detector, and further downstream by septum magnets. In order to maintain a straight orbit for the incoming beam in the separator, a weak magnetic field is superimposed to exactly cancel the electrostatic force acting on the incoming beam while doubling its effect on the outgoing one. After extraction, the positron beam is guided to its dump, and the electron beam is used for positron production.

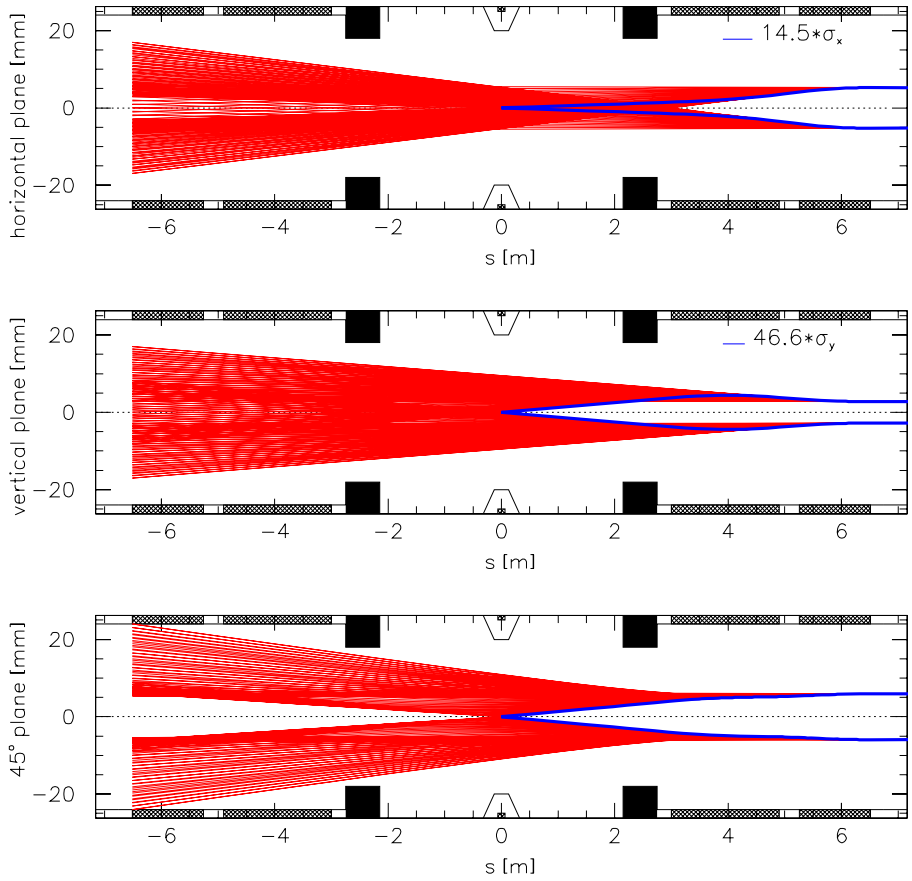


Figure 3.7.1: *Synchrotron Radiation emitted in the last doublet quadrupoles through the vertex detector and opposing doublet apertures.*

3.7.2 Magnet Lattice and Optics

The magnet layout and beam sizes along the beam delivery system are shown in Fig.3.7.2. The BDS is composed of the 6 sections whose function has been described above. These sections, listed in Table 3.7.1 with some of their characteristics, are described thereafter.

• The **Linac Matching Section** (LMS) is composed of six quadrupoles, following the last focusing linac quadrupole, which provide enough free parameters to match the

	L [m]	s [m]	β_x [m]	β_y [m]
Linac Matching Section (LMS)	17.3	0.0	113.445	19.489
Collimation Section (CS)	389.5	17.3	13.081	7.244
Tuning and Diagn. Section (TDS)	182.2	406.8	13.081	7.244
Beta Matching Section (BMS)	31.5	589.1	21.383	9.655
Chromatic Corr. Section (CCS)	331.4	620.6	5.614	1.767
Final Transformer (FT)	142.5	951.9	5.614	1.767
Interaction Point (IP)		1094.4	0.025	0.0007

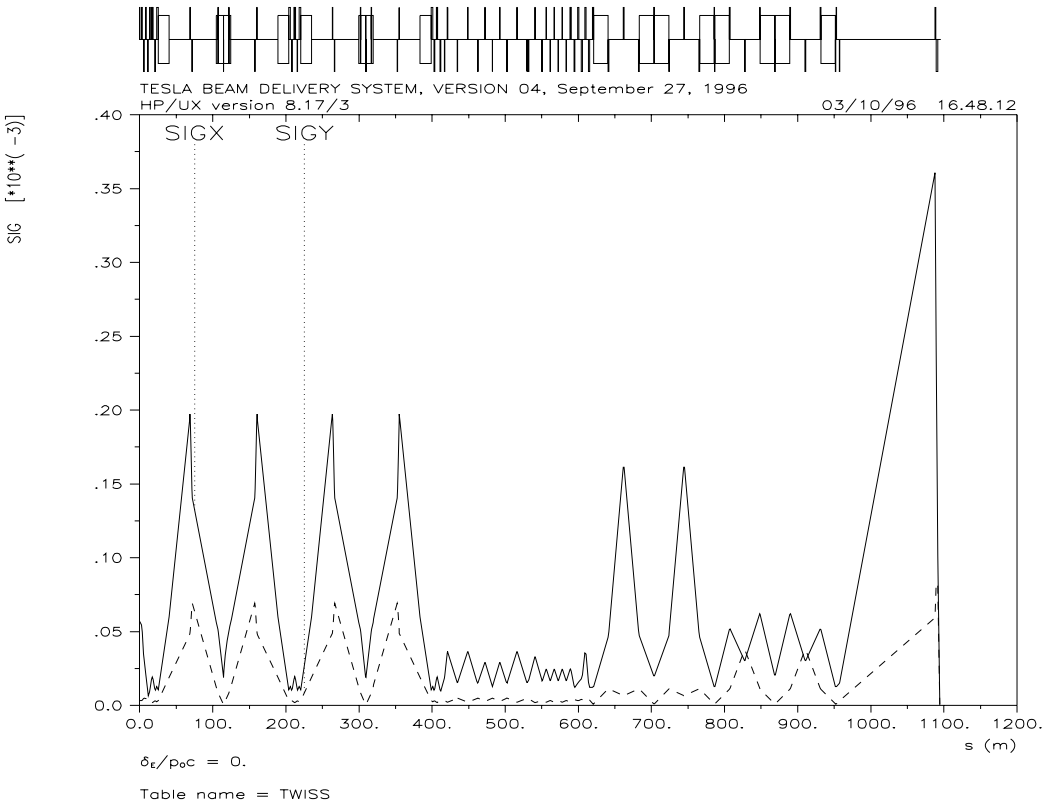
Table 3.7.1: *Orbit lengths and beta functions at the entrance of the BDS sub-sections. The overall x and y demagnification factors are* 67×167.

beam to the entrance of the collimation section and therefore to tune the beam sizes at the collimator locations.

• The **Collimation Section** (CS) provides 4 regions of combined high horizontal and vertical betas as well as large horizontal dispersion. Each region hosts a pair of Titanium collimators to intercept off-momentum and large (x and y) amplitude particles. In normal operation, their aperture is set to $8\,\sigma_x$ and $32\,\sigma_y$ and they collimate particles with more than 2% momentum offset. The last two collimator pairs are in phase with the last doublet: they intercept the sine-like trajectories (with respect to the IP) which have the largest amplitudes in this doublet and therefore produce the most harmful part of the synchrotron radiation in the last quadrupoles. The two upstream collimator pairs are in phase with the IP and thus intercept the cosine-like trajectories. This second phase collimation is necessary to protect the detector from phase mixing, of chromatic or non-linear origin, occurring between the collimation section and the last doublet. To minimize the chromatic phase mixing in the CS itself, the chromaticity created by the CS high-beta regions is corrected locally by embedded sextupoles pairs, in a way similar to the one described below for the CCS.

• The **Tuning and Diagnostic Section** (TDS) is a section devoted to measuring and correcting the sensitive transverse coupling, inherited from the Linac transport and/or created in the Collimation Section. The coupling correction is performed in the Skew Quadrupole Correction Section which hosts four skew quadrupoles located at the proper relative phase advances to cancel the four coupling coefficients of the transverse beam matrix. Tuning is done by minimizing the vertical emittance which is measured, together with the horizontal one, in the Emittance Measurement Section (EMS). The EMS is a simple FODO lattice with three and a half 45° cells along which four wire scanners are regularly disposed every 45° in phase space. Each wire scanner measures the horizontal and vertical spot sizes, expected to be about equal to 17 μm and 3.3 μm. This is therefore in the range of solid carbon wires as the ones currently and reliably operated in the SLC. The regularity of the beam phase advances and spot sizes at the wire locations optimizes the reconstruction of the transverse emittances from the spot size measurements.

• the **Beta Matching Section** (BMS) contains 7 quadrupoles which allow to



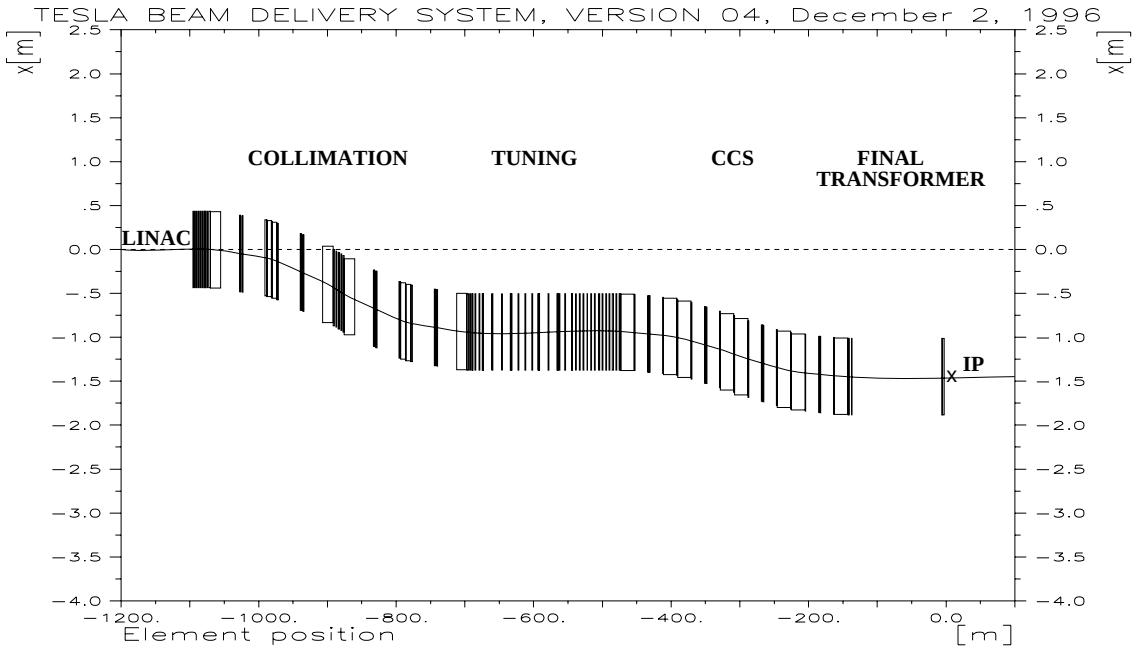


Figure 3.7.2: *Magnet layout and beam spot size (in meters) in the BDS (top), and horizontal survey of the BDS (bottom).*

match the beam matrix at the entrance of the final focus system over a wide range of parameters. Since the final focus system is a pure demagnifying telescopic system, the BMS can therefore be used to fine tune the beam at the IP as well as the phase advance from the collimators to the last doublet. It can also be used, for the start-up operation, to blow up the beam by as much as a factor 10 with a relative variation of the quadrupole gradients limited to about $\pm 20\%$. The BMS magnets need therefore to be tunable over this range.

• The **Chromatic Correction Section** (CCS) is a one-to-one transformer composed of two FODO sections with four 90° cells each. Both sections contain four bending magnets symmetrically disposed to generate a closed horizontal dispersion bump, and a pair of sextupoles to tune the chromaticity. Since the two sections have opposite quadrupole polarities, the first one corrects the horizontal chromaticity and the second one the vertical chromaticity. Each of the four sextupoles is located next to the central quadrupole where the phase advance is a multiple of $\pi/2$ and the relevant beta-function is maximum. In this way the sextupoles do not generate geometric or chromatic aberrations other than the ones needed to cancel the chromatic aberrations from the last doublet. The bandwidth of this system, shown in Fig.3.7.3 is more than sufficient to accept the foreseen 0.1 % bunch energy spread and bunch to bunch energy deviations. With opposite dipole polarities, the two sections create an S-shaped CCS orbit with two opposing deflection angles of about 3 mrad generated by four 19.7 m long dipoles with a field of 32 mT. The beam relative energy spread due to synchrotron radiation along these 8 dipoles and the relative energy loss between sextupoles of the same pair are both very small, lower then 2×10^{-5}.

• The **Final Transformer** (FT) is a telescopic system which achieves the final demagnification of the beam spot sizes by 15 horizontally and 50 vertically. It is composed of a weak and a strong doublet separated by 130 m. The strong final doublet, located inside of the detector, is composed two 2 K superconducting quadrupoles of 250 T/m gradient, 48 mm physical aperture diameter and 1.9 m and 1.25 m lengths, as shown in Fig.3.7.1. Their design is based on successfully tested prototypes for the LHC lattice quadrupoles [3]. Their cross- section is shown in Fig.3.7.4. The two opposing final doublets are separated by 6 m around the IP and their apertures provide enough clearance for the spent beam to allow collisions with zero crossing angle.

• **Spent Beam Extraction**

Because of the 700 ns (about 200 m) bunch separation, the beams can collide head-on and be separated outside of detector in the long drift space between the two doublets of the FT. This is done by a 20 m long electrostatic separator with a field of 4.8 MV/m starting about 10 m after the IP. To prevent synchrotron radiation from hitting the detector, a weak 16 mT magnetic field is superimposed so that the incoming trajectory is not bent. With an overall angular deflection of 0.77 mrad, the outgoing beams are offset by 3 cm from the main beam axis at 60 m from the IP. This separation is large enough for septum magnets to further extract the electron beam towards the positron production system and the positron beam towards its dump.

• **Tracking Results**

Results of tracking simulations through the entire BDS are shown in Fig.3.7.5. They

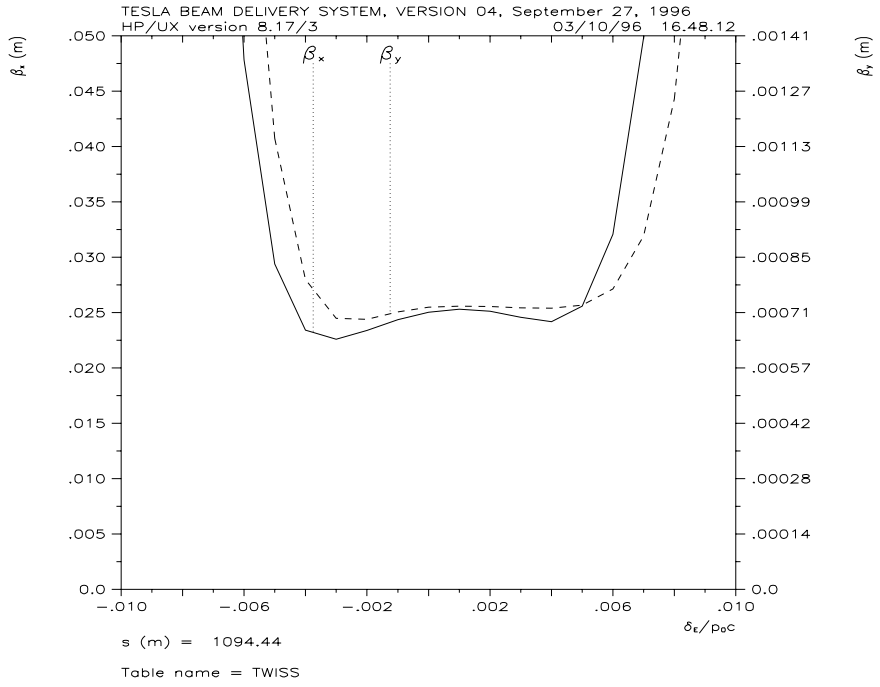


Figure 3.7.3: *Energy bandwidth of the 500 GeV beam delivery system calculated from the energy dependence of the beta-functions at the IP.*

have been performed with MAD [4] and they include the effect of synchrotron radiation in the magnets. Luminosities are calculated from the beam distributions without pinch effect. Up to 2.5×10^{-3} relative energy spread the effect of non-linear aberrations on the beam IP spot sizes is negligible and the expected luminosity is 97 % of the design one.

• **Magnets and Beam Pipe Apertures**

The required aperture radii of the BDS quadrupoles (final doublet excepted) and sextupoles are shown in Fig.3.7.6 for a 1 T pole-tip field which is an achievable field for room temperature magnets. With this assumption the tightest apertures are about 5 mm radius. A 1 cm diameter stainless steel beam pipe is acceptable from the impedance point of view since the relative energy spread and energy loss induced by the resistive wakefields along 1.1 km of beam pipe are about 3×10^{-5}. It also provides a sufficiently large linear acceptance of about 25 horizontal and 100 vertical standard deviations. In principle, a larger aperture is possible by increasing the length of some of the magnets.

• **The 800 GeV Upgrade**

The 500 GeV cm-energy Beam Delivery System optics described above can be

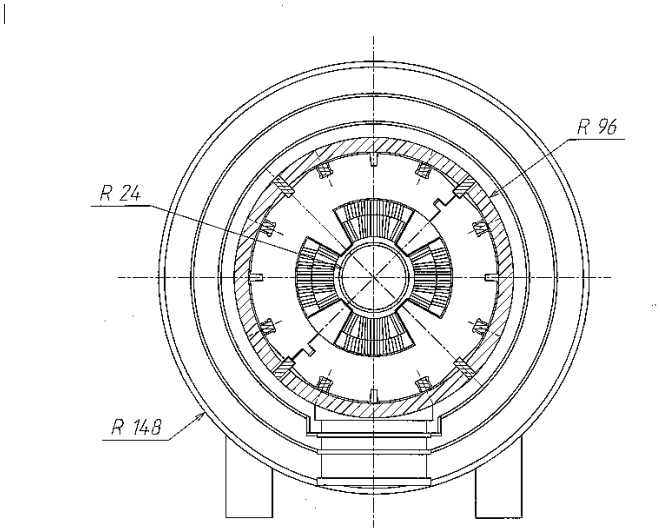


Figure 3.7.4: *Last doublet superconducting quadrupole cross-section.*

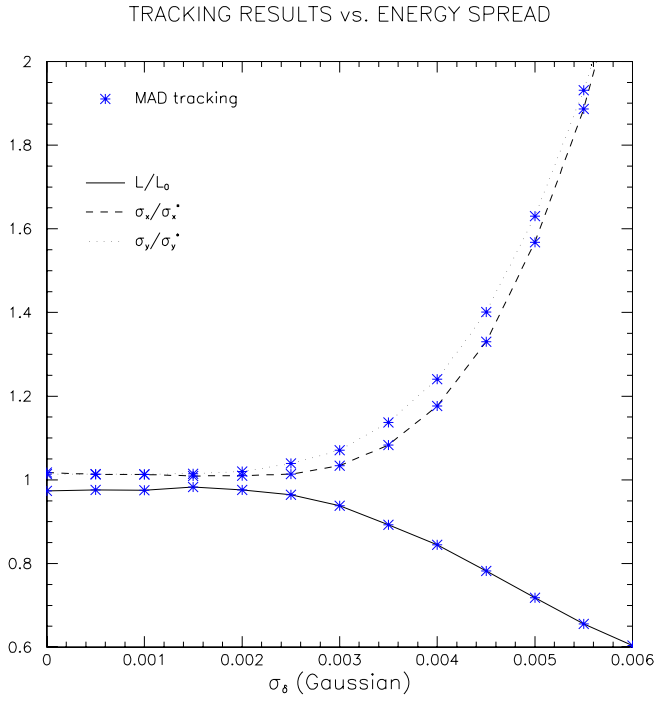


Figure 3.7.5: *Dependence of the spot sizes and luminosity, normalized to their design values, on the Gaussian RMS relative energy spread.*

adapted to 800 GeV cm-energy with the same magnets, except for the last doublet quadrupoles. Since the beta-functions at the IP are identical to the ones for 500 GeV cm-energy, the Twiss parameters over the BDS are kept essentially unchanged by in-

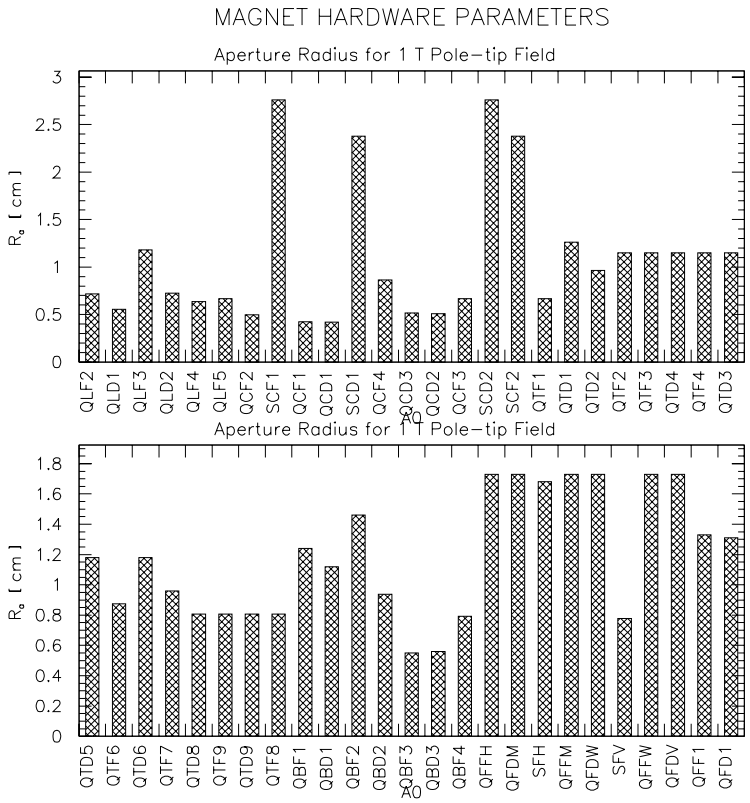


Figure 3.7.6: *Aperture radii of the BDS magnets (dipoles and the final SC quadrupoles excepted) assuming 1 T pole tip field.*

creasing the BDS quadrupole focusing strengths in proportion to the beam energy. However, in order not to increase the gradient of the last doublet superconducting quadrupoles, while keeping a 3 m drift space to the IP, the FT is re-designed with longer, namely 2.71 m and 1.67 m long, SC magnets. The FT demagnification is also reduced to 15×30 to keep the upstream doublet at the same location. This reduction is exactly compensated by increasing the demagnification of the Beta Matching System with some retuning of the BMS quadrupoles. The resulting energy bandwith of the system is shown in Fig.3.7.7. The strengths of the dipole magnets in the entire system are sufficiently low to keep emittance growth from synchrotron radiation at a level of a few percent.

Spent beam separation is achieved by lengthening the electrostatic separator to 32 m which gives 2.6 cm outgoing beam offset in the septum magnets.

3.7.3 Sensitivity to Errors

Displacement and magnet errors occuring upstream of the beta matching section (BMS) will be accurately detected in the diagnostic section and corrected. Downstream errors

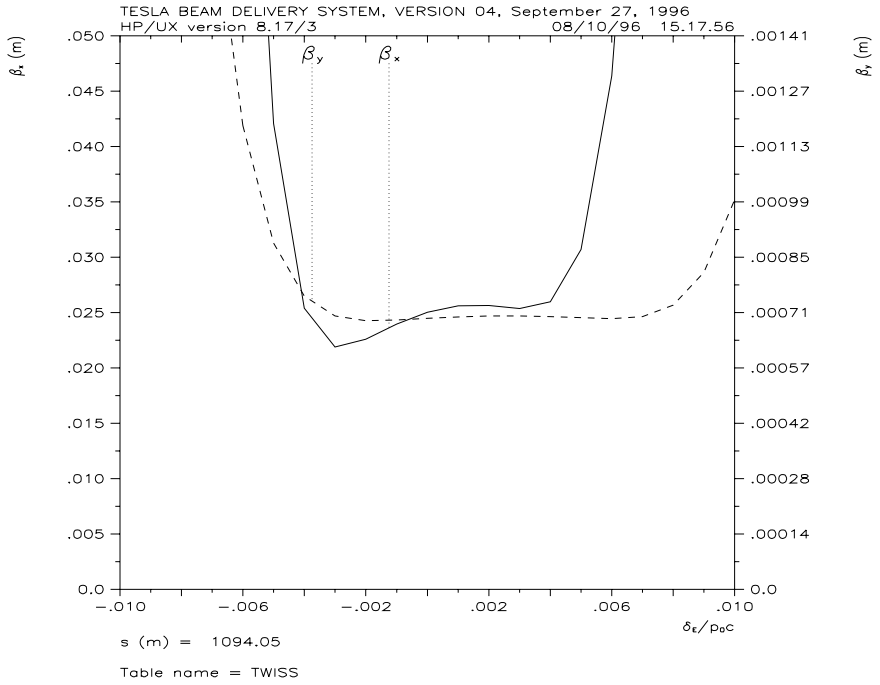


Figure 3.7.7: *Energy bandwidth of the 800 GeV beam delivery system calculated from the energy dependence of the beta-functions at the IP.*

will be more difficult to measure with the fewer and less precise beam size monitors equiping the IP and its image point at the exit of the CCS. This section is thus devoted to the study of the sensitivity to errors of the BMS and FFS final sections.

This error sensitivity is calculated from the first and second order derivatives of the IP orbit and transfer matrix with respect to magnet errors as well as from the first order derivatives of the T-matrix of the transfer map. This information, stored in large arrays, allows one to calculate the sensitivity of the optical system to any kind of misalignments, vibrations or magnet field instabilities.

Once translated in terms of luminosity loss and spot size growth, this information can also be used to calculate tolerances of the final focus system to misalignment or field errors. To do so, one has to specify a model that describes the source of these errors, their spectrum and correlations over time and distance. In this section, the following two simple situations are considered: first, fixed (or static) errors affecting only one magnet, assuming all other elements are perfectly aligned and tuned, and second, uncorrelated random transverse vibrations affecting all magnets of both final focus beam lines. More realistic models taking into account ground motion are studied in Sect.3.7.4.

• **Static Tolerances**

The static tolerances are shown in Figs. 3.7.8 and 3.7.9 assuming a beam energy spread of 0.1 %. They are derived from a 2% reduction of the luminosity calculated by assuming that one beam is perfectly matched while the other beam is affected by the error considered. In all pictures, the inverse tolerances are plotted in such a way that the highest bars correspond to the tightest tolerances. Black bars give the tolerances with no steering corrections, while the white bars give tolerances with IP offset corrections included. Besides the last doublet quadrupoles which, as expected, set the tightest tolerances, the CCS bending magnets show also a high sensitivity to skew rotations and field errors. Dispersion errors, especially vertical ones, are mainly due to the last CCS quadrupoles and the first FT doublet.

• **Uncorrelated Vibration Tolerances**

For the model of mechanical vibrations where magnet transverse motions are uncorrelated and have the same horizontal and vertical RMS, $\sigma_x^{(mag)}$ and $\sigma_y^{(mag)}$, the RMS of the resulting relative beam offset, crossing-angle and the single beam dispersions at the IP are then given by:

$$\begin{array}{rcl} \sigma_{\delta x^*} = 0.88 \times \sigma_x^{(mag)} & , & \sigma_{\delta y^*} = 0.24 \times \sigma_y^{(mag)} \\ \sigma_{\theta_x^*} = 8.95\,\mathrm{rd}/m \times \sigma_x^{(mag)} & , & \sigma_{\theta_y^*} = 44.7\,\mathrm{rd}/m \times \sigma_y^{(mag)} \\ \sigma_{D_x^*} = 98.8 \times \sigma_x^{(mag)} & , & \sigma_{D_y^*} = 19.2 \times \sigma_y^{(mag)} \end{array}$$

Actually the offset RMS $\sigma_{\delta x^*}$ and $\sigma_{\delta y^*}$ do not include the dominant contribution from both the opposing last doublets which can be directly read from Figs. 3.7.8 and 3.7.9.

In this model, small luminosity losses can be parametrized as

$$\delta\mathcal{L}/\mathcal{L}_0 = -2\% \times \{(\sigma_x^{(mag)}/318\,\mathrm{nm})^2 + (\sigma_y^{(mag)}/35.2\,\mathrm{nm})^2\} \quad \text{(no offset correction)}$$

for the fast jitter with no steering correction done (again excluding the contribution from the last doublets), and

$$\delta\mathcal{L}/\mathcal{L}_0 = -2\% \times \{(\sigma_x^{(mag)}/558\,\mathrm{nm})^2 + (\sigma_y^{(mag)}/204\,\mathrm{nm})^2\} \quad \text{(with offset correction)}$$

for the case where steering correction is done. In the first case, the luminosity loss is dominated by the beam relative offsets. In the second case, this contribution is removed and the luminosity loss is mainly coming from spot-size growth with a small contribution from the vertical crossing angle. Since a fast IP orbit steering feedback is foreseen, the latter case is the relevant one here. The 2 % luminosity loss tolerances can be read immediately from the above expressions. With the 0.1 % energy spread considered, the horizontal spot-size growth induced by horizontal misalignments comes 36% from waist-shift and 64% from horizontal dispersion while the vertical spot-size growth induced by vertical misalignments comes half from yx'-coupling and half from vertical dispersion. The vertical spot-size growth induced by horizontal misaligments comes totally from waist-shift, and the horizontal spot-size growth induced by vertical misalignments comes totally from xy'-coupling.

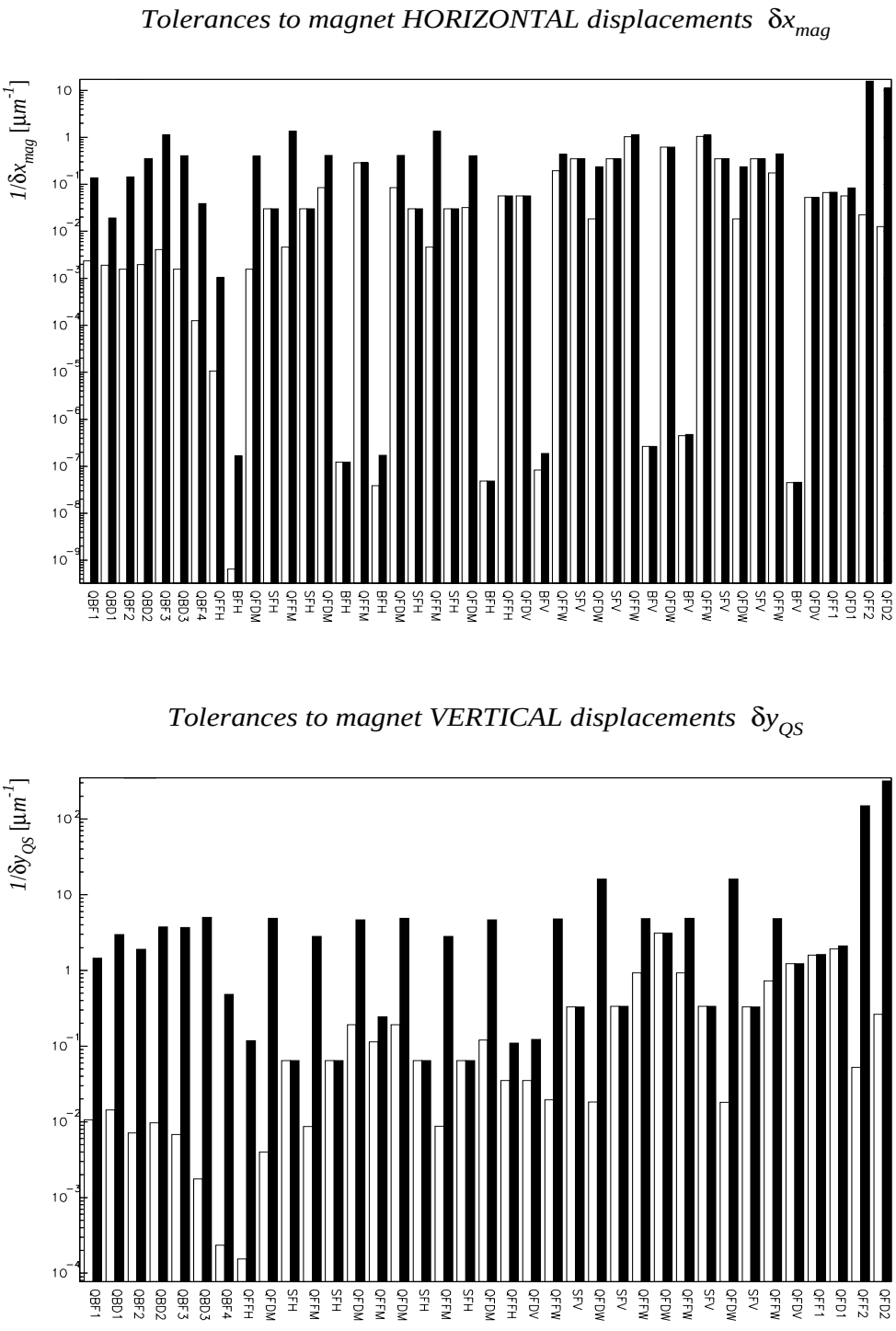


Figure 3.7.8: *FFS Sensitivity to transverse displacements (black bars: with IP offset ; white bars: IP offset corrected).*

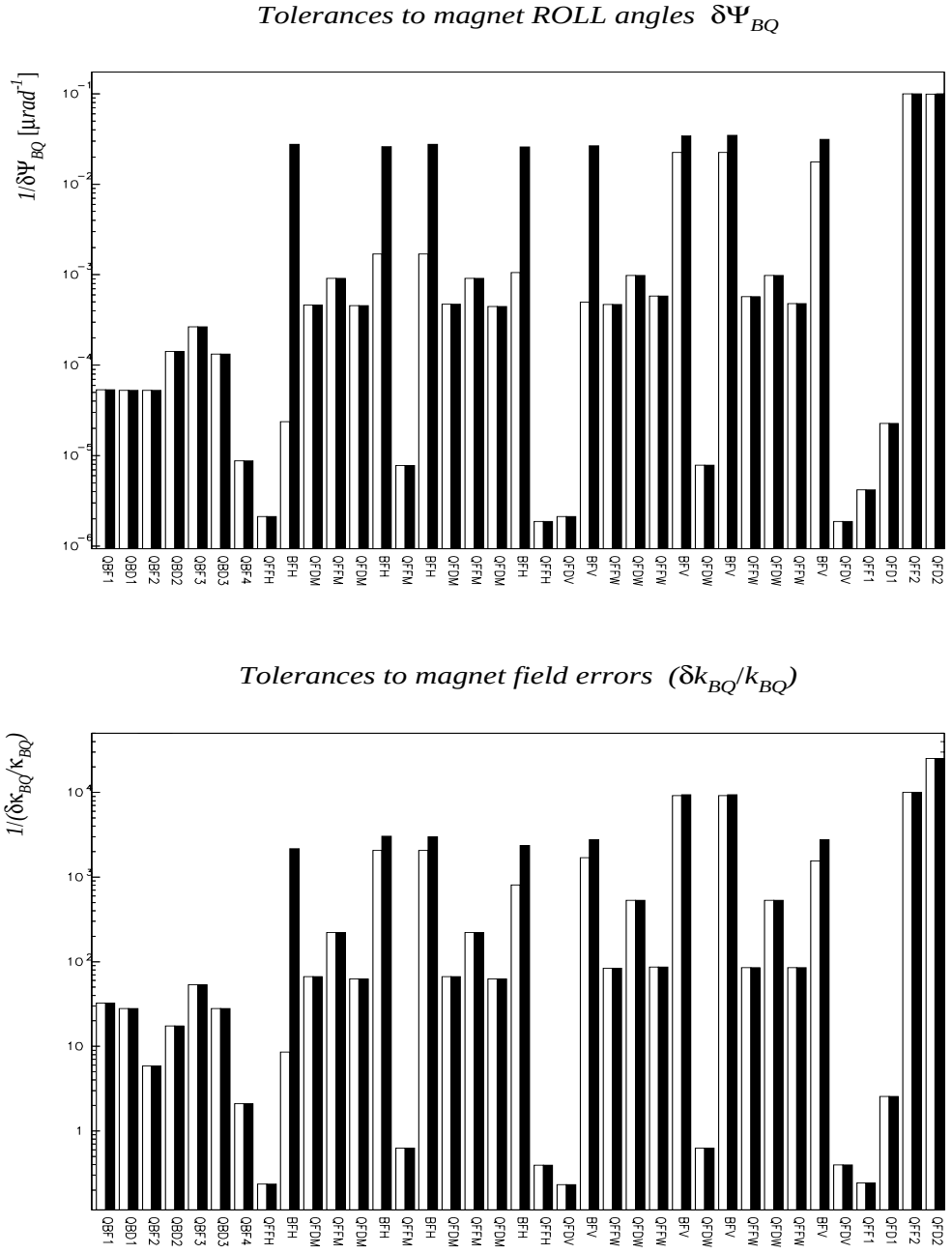


Figure 3.7.9: *FFS Sensitivity to skew rotations and field errors (black bars: with IP offset ; white bars: IP offset corrected).*

3.7.4 Ground Motion

Ground motion is an important issue for the linear collider, because it may result in misalignment of focusing and accelerating elements and thus in trajectory offset and emittance dilution. Ground motion influence is more severe in the beam delivery system, where tolerances are the tightest. It is therefore natural to discuss it in detail in this chapter. Some relevant considerations appear in the chapter describing beam dynamics in the main linac as well.

• **Mathematical description of ground motion** consists in spectral representation, because the relevant quantity $y(t,s)$, the vertical position of the ground surface, can be considered to move in a "random" fashion and should be characterized therefore by the corresponding power spectrum $P(\omega,k)$ [11]. Other characteristics (such as the measurable temporal power spectrum of absolute motion taken in a single point $p(\omega)$, the spectrum of relative motion of two separated points $\rho(\omega,L)$, the correlation $C(\omega,L)$ and the power spectrum $P(t,k)$ of the misalignment $y(t,s)-y(0,s)$ etc...) are all related to $P(\omega,k)$.

While spectral properties of the ground motion are described by the power spectrum, spatial spectral properties of the considered focusing structure of the linear collider can be described by a spectral response function [11]. The errors induced by misalignments, such as the rms beam offset or the rms beam dispersion at IP, can be determined by the integral of the corresponding spectral response function with the power spectrum of the misalignment. For example the rms beam dispersion at the IP is

$$\langle \eta^2(t) \rangle = \int_{-\infty}^{\infty} P(t,k)\, G_\eta(k)\, \frac{dk}{2\pi}$$

The spectral response functions $G_{\rm off}(k)$ or $G_\eta(k)$ show, in terms of rms relative offset or beam dispersion, the spectral response of the considered focusing section to the misalignment having spatial period $2\pi/k$ (Fig.3.7.10).

The spectrum of misalignments, if produced by ground motion, exhibits the following evolution

$$P(t,k) = \int_{-\infty}^{\infty} P(\omega,k)\ \ 2\,[1-\cos(\omega t)]\, \frac{d\omega}{2\pi}$$

In most cases it is essential to have the orbit stabilization feedback taken into account. The equilibrium beam characteristics, for example the rms equilibrium beam offset at the IP, can then be estimated using the function $F(\omega)$, which characterizes the spectral performance of the feedback, convoluted with the ground motion spectrum and the corresponding spectral response function.

However, orbit stabilization or tuning algorithms may be quite complex and their performance may depend on many factors such as BPM or mover resolution etc... In these cases analytical methods based on the spectral response function cannot be applied, thus direct simulations of ground motion [17] and particle tracking should be used to determine the beam degradation.

• **Measured information on ground motion** is available from studies performed in many laboratories all over the world [5]-[10].

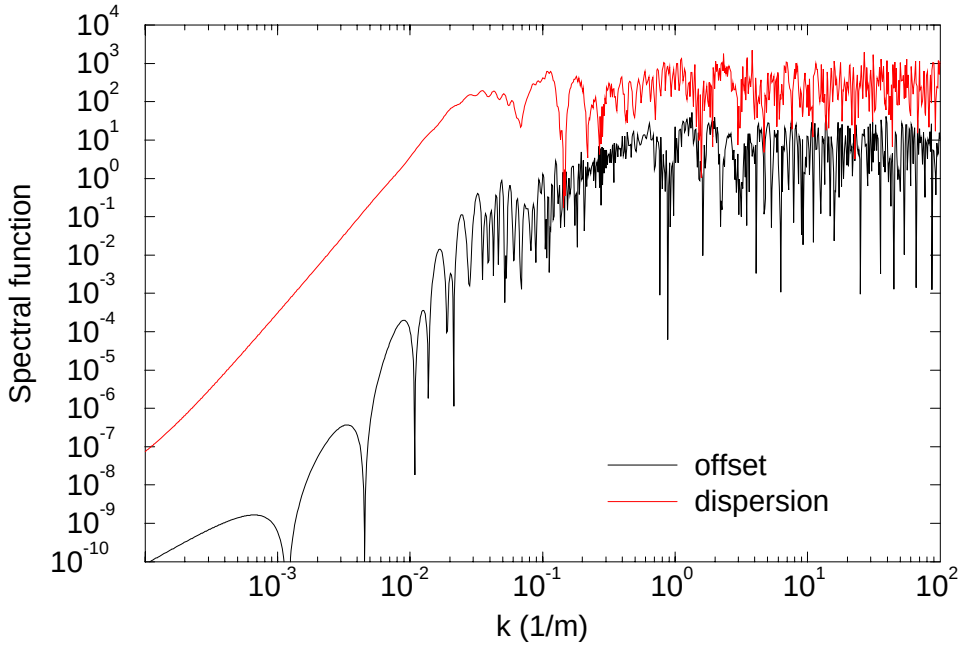


Figure 3.7.10: *Spectral response functions for the relative offset $G_{\rm off}(k)$ and for the beam dispersion $G_\eta(k)$ for the TESLA Beam Delivery system.*

It is known that the power spectrum $p(\omega)$ of absolute ground motion (which contains contribution of all k) grows very fast with decreasing frequency. In quiet conditions it behaves approximately as $p(\omega) \propto 1/\omega^4$ in a wide frequency band. The motion is unavoidable as it consists of seismic activity. At low frequency $f < 1$ Hz significant contributions to absolute ground motion come from remote sources like water motion in the oceans, atmospheric activity, temperature variations etc... A well-known example of the ocean influence is the peak in the band 0.1 – 0.2 Hz with a few micrometers amplitude (Fig.3.7.11).

On the other side, in the band $f > 1$ Hz the human produced noise is usually dominating over the natural noise and the power spectrum depends very much on the local conditions (location of sources, depth of tunnel etc.). Locally generated noise can be much bigger than remotely generated. For example the spectrum measured at the tunnel of an operating circular accelerator (DESY HERA collider) presents high amplitudes at $f > 1$ Hz due to noise generated by different technical devices (Fig.3.7.11). This noise may have a big amplitude and poor spatial correlation. Technical devices of the future linear collider therefore should be properly designed in order to pass as low vibration as possible to the tunnel floor.

It is known from correlation measurements [6, 10] that in quiet conditions the motion in the band $f > 0.1$ Hz can be considered as wave-like, i.e. the frequency ω and wave number k are connected via phase velocity v. At $f \approx 0.1$ Hz the value of v was found to be close to the velocity of sound in the surrounding media: about 3000 m/s at LEP and about 2000 m/s at the SLC tunnel. For cultural noise this value decreases rapidly: the value of v determined from the SLAC data behaves as

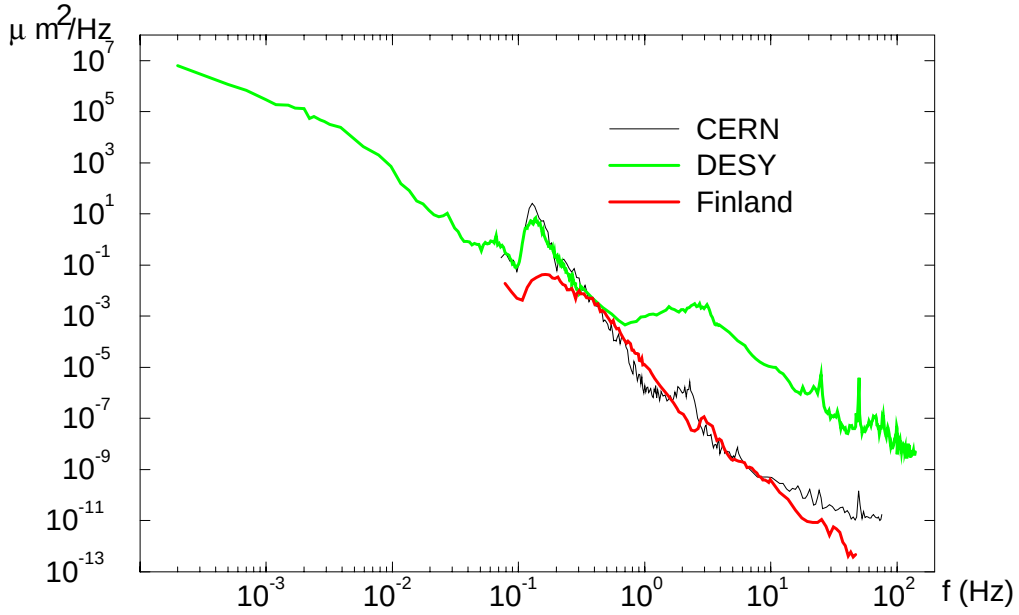


Figure 3.7.11: *Absolute power spectrum measured in quiet (CERN LEP tunnel, shutdown [6]; Finland, cave [7]) and noisy (DESY HERA tunnel, operational condition [8]) condition.*

$v \approx 450 + 1900\exp(-f/2\text{Hz})$ m/s (for $f > 0.1$ Hz) [10], HERA measurements gave $v \approx 400$ m/s [8]. The SLAC measurements have shown (at least at this place and these conditions) that the contribution of non-wave motion is negligible for $f > 0.1$ Hz.

The motion at $f < 0.1$ Hz is different. The elastic motion (produced by the moon, for example) is present here also, but of much bigger relevance is the inelastic diffusive motion, probably fed by the elastic motion and caused by its dissipation. The motion is believed to be described by the "ATL law" [5], which states that the relative rms displacement after a time T of the two points separated by a distance L is

$$\langle \Delta Y^2 \rangle = A\,T\,L$$

The parameter A was found to be $A \approx 10^{-6\pm1}$ $\mu\text{m}^2\text{s}^{-1}\text{m}^{-1}$ at different places. One can see that this displacement is proportional to the square root of the time and separation: this stresses the random, non wavelike, diffusive character of the slow relative motion.

The parameter A was observed to be smaller in tunnels built in solid rock. It also depends on the method of tunnel construction: in the tunnel bored in granite $A \approx 10^{-6}$ $\mu\text{m}^2\text{s}^{-1}\text{m}^{-1}$ was observed, while in a similar tunnel built by use of explosions, the parameter A is found to be 5 times larger probably because of the rock fragmentation, artificially increased during construction [9]. The high frequency correlation in the second case is also poor. The parameter A tends also to be smaller in deeper tunnels [13].

The ranges of T and L where the "ATL law" is valid are very wide. In [13] it was shown that "ATL law" is confirmed by measurements of ground motion in different

accelerator tunnels in the range from minutes to tens of years and from a few meters to tens of kilometers. The measured relative power spectra, presented in [9] and in [12], exhibit the "ATL" behavior for $f < 0.1$ Hz (for $L \approx 30$ m). These measurements indicate that the transition region from wave to diffusion motion is placed at rather short times (a few seconds). Measurements of the closed orbit motion in the HERA circular collider have shown that the power spectrum of this motion corresponds to the "ATL law" in a wide frequency range, from $f \approx 0.1$ Hz down to $f \approx 10^{-6}$ Hz and the parameter is $A \approx 10^{-5}$ $\mu\mathrm{m}^2\mathrm{s}^{-1}\mathrm{m}^{-1}$ [12].

The very slow motion can be systematic (not described by power spectrum) as well. Such motion has been observed at LEP Point 1 and PEP [14] where some quadrupoles move unidirectionally during several years with rate about 0.1 – 1 mm/year. Points a few tens of meters apart can move in opposite directions. The amplitude of the motion can be larger than the one of diffusive motion. The motion is probably due to geological peculiarities of the place or due to relaxation if the tunnel was bored in solid rock.

The elements of the linac will be placed not on the floor, but on some girder, which could amplify some frequencies due to its own resonances. It is not only this amplification that is dangerous, more important is that the different girders can amplify or change the floor motion in different ways, and hence spoil correlation of the floor motion. It is therefore preferable to push the girder resonances to high frequencies where the correlation is poor anyway and floor amplitudes are smaller. This requires firm connections of the girder to the floor. The active systems [16], which can help to isolate the quadrupoles from high frequency human produced floor motion, should be made insensitive to slow motion (below 0.1...1 Hz), otherwise the correlation of long wavelength motion may be destroyed by imperfections of the active feedback when operating at low frequencies.

• **A model of the ground motion spectrum** $P(\omega, k)$ can be built [11] with the assumption that the low frequency part of motion is described by the "ATL law" [5], and the high frequency part is produced mainly by waves. The waves are assumed to be elastic, transverse, propagating at the surface of the ground with uniform distribution over azimuthal angle. The "ATL law" is included to the model in a modified form, in order to prevent overestimation of fast motion, which is an intrinsic feature of the pure "ATL law". The spectrum, which include both the waves and the modified "ATL law", is the following

$$P(\omega,k) = \frac{A}{\omega^2 k^2}(1-\cos(kB/A/\omega^2)) + \sum_i \frac{a_i}{1+[d_i(\omega-\omega_i)/\omega_i]^4}\,\frac{2}{\sqrt{(\omega/v)^2-k^2}}$$

We consider two set of parameters to cover two extreme cases of seismic conditions. Parameters of the first model are the following: $A = 10^{-5}$ $\mu\mathrm{m}^2\mathrm{s}^{-1}\mathrm{m}^{-1}$, $B = 10^{-6}$ $\mu\mathrm{m}^2\mathrm{s}^{-3}$. The single peak described by $\omega_1 = 2\pi \cdot 0.14$ Hz for the frequency of the peak, $a_1 = 10\mu\mathrm{m}^2/\mathrm{Hz}$ for its amplitude, $d_1 = 5$ for its width, and $v_1 = 1000$ m/s for the velocity. This parameters represent quiet conditions.

The second model corresponds to seismic conditions with big contributions from cultural noises (as in the HERA tunnel in operating conditions). The parameters are

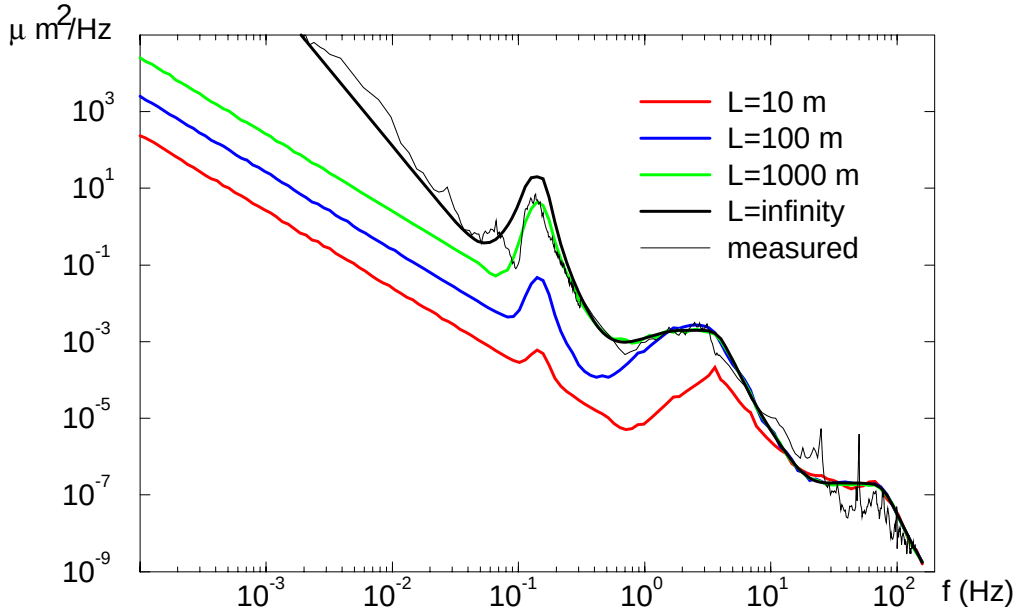


Figure 3.7.12: *Absolute power spectrum measured in a noisy conditions (HERA during operation [8]) compared to absolute and relative spectra for different spatial separation L calculated from the second model.*

the following: $A = 10^{-5}$ μm^{2}s^{-1}m^{-1}, $B = 10^{-3}$ μm^{2}s^{-3} and three peaks: $f_1 = 0.14$, $f_2 = 2.5$, $f_3 = 50$ Hz; $a_1 = 10$, $a_2 = 10^{-3}$, $a_3 = 10^{-7}\mu$m^2/Hz; $d_1 = 5$, $d_2 = 1.5$, $d_3 = 1.5$; $v_1 = 1000$, $v_2 = 400$, $v_3 = 400$ m/s. The corresponding spectrum of relative motion is shown on Fig.3.7.12.

• **Beam offset and spot size at the IP of TESLA** have been calculated analytically using the models of ground motion spectrum $P(\omega, k)$ and the corresponding spectral response functions, determined in linear approximation [11].

The most harmful effect of misalignments of focusing elements is the relative vertical offset of the opposite beams at the IP (Fig.3.7.13). The tolerance on the offset, corresponding to 2% luminosity loss, is around 5 nm for TESLA. It corresponds to critical times around 1 s for the quiet model and 1 ms for the (likely somewhat too pessimistic) noisy model. This means that even in the extremely noisy conditions the fast orbit feedback, described in section 3.8.3, will be able to keep the beams colliding.

Transverse displacements of focusing elements generate spot size growth at the IP induced by dispersion, longitudinal shift of the beam waists and xy-coupling generated by offsets of the beams in quadrupoles and sextupoles. Corresponding spectral response functions have been calculated and the beam size growth due to all these effects has been determined (Fig.3.7.14). The main contribution to the vertical beam size growth comes from yx' coupling and from vertical dispersion. One can see that the critical time scale for 2% luminosity loss is about 200 s that imposes quite relaxed requirements to the beam size stabilization feedback described in the next section.

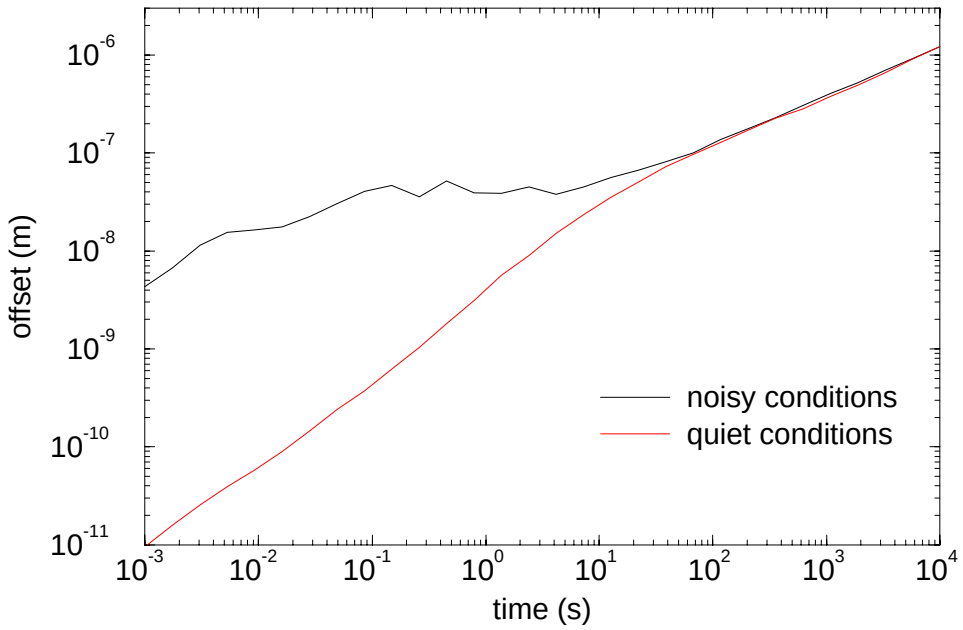


Figure 3.7.13: *Relative vertical rms offset of the beams at IP of the TESLA Beam Delivery system for the noisy and quiet models of the power spectrum of ground motion.*

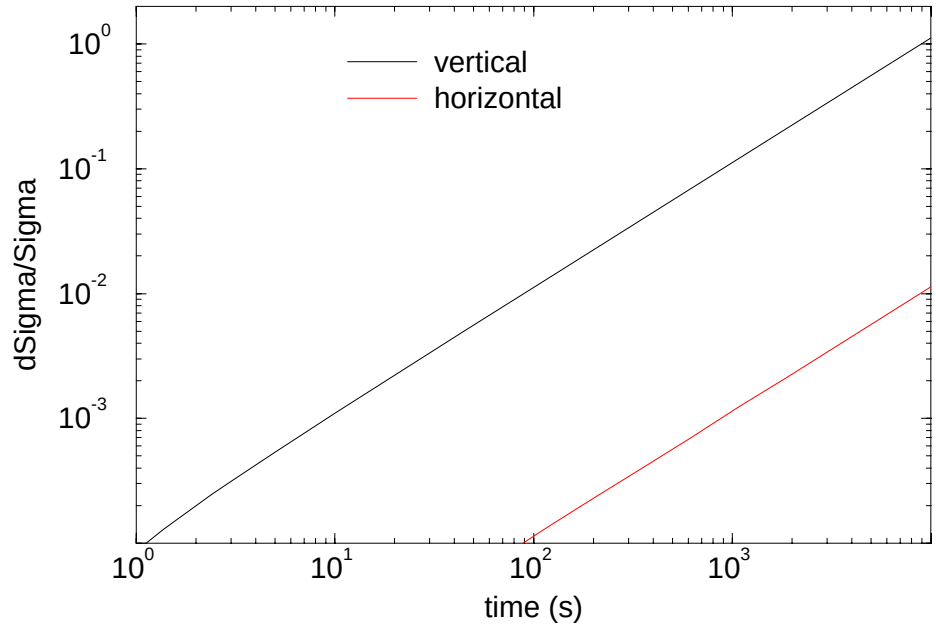


Figure 3.7.14: *Vertical and horizontal rms beam size growth* $\delta\sigma^*/\sigma^*$ *versus time for the "ATL" ground motion with* $A = 10^{-5}\ \mu m^2 s^{-1} m^{-1}$. *Rms energy spread of the beam is* $\delta p/p = 10^{-3}$.

3.7.5 Phase Space Measurement and Tuning

3.7.5.1 Introduction

With the experience gained from two existing final focus systems (SLC and the final focus test beam, FFTB), the various tuning techniques which are required to produce the small beam sizes at the IP are already mature and well tested. For the next generation of linear colliders, the tuning used at the FFTB [18] is particularly relevant. In the following sections, we will discuss the application of these and other tuning techniques to the TESLA beam delivery system.

3.7.5.2 Energy Measurement and Correction

The beam delivery system offers several near optimum positions for the precise measurement of the relative energy (momentum) error of the bunches. Both the collimation sections and the chromatic correction sections have symmetric high dispersion points positioned by a *-I* transformation apart in betatron phase space, allowing any betatron component of the horizontal beam offset at these points to be effectively canceled out by taking the sum of the two measurements. The resolution of such a system is given by (a) the value of the dispersion function (b) the resolution (noise) of the BPM and (c) the accuracy of the cancelation of the betatron component. Table 3.7.2 gives the relevant parameters for the four sets of symmetric dispersion points available in the TESLA beam delivery system. The resolution is calculated assuming the given BPM precision (5 μm) and a (conservative) random beam jitter of $\sigma/4$.

section	D_x (mm)	β_x (m)	σ_x (μm)	σ_{BPM} (μm)	resolution
collimation	47	1157	182	3	10^{-4}
CCS X	35	881	159	5	10^{-4}
CCS Y	35	26.5	28	5	10^{-4}

Table 3.7.2: *Key parameters for -I paired BPMs used as spectrometers in the beam delivery section. The collimation section refers to both the sine- and cosine-like sections which have identical parameters.*

Either of the two BPM pairs in the collimation sections can be used in a fast energy feedback system to correct the bunch to bunch energy error before entering the final focus section. Such an arrangement is shown schematically in figure 3.7.15. By placing the correction RF cavities downstream of the measurement BPMs, a signal delay time of the order of 0.1 μsec can be achieved (assuming a cable signal velocity of approximately $0.8c$). By performing all the necessary signal processing in hardware it should be possible to make an energy correction within the TESLA bunch spacing (0.7 μsec). The system can also be used as a low bandwidth feedback to compensate slow drifts. In this case the correction can be performed at the exit of the linac using

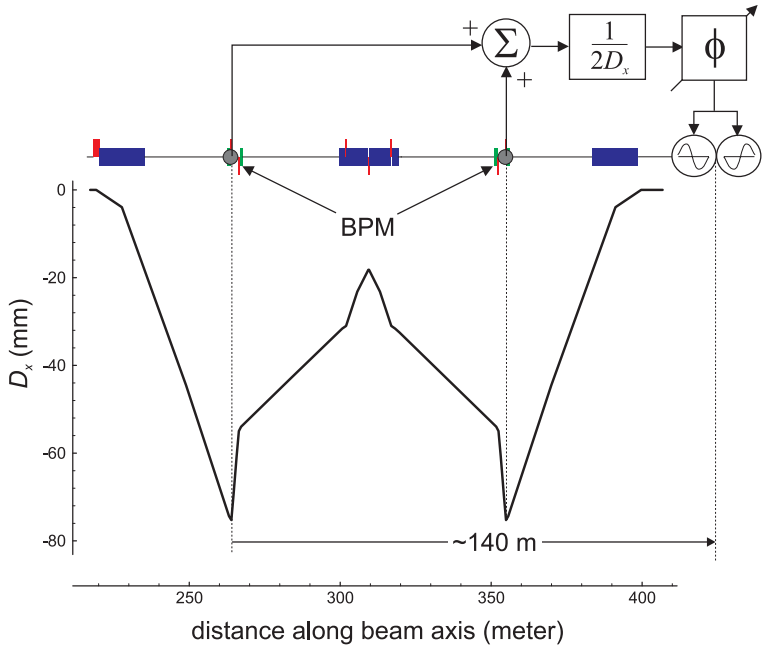


Figure 3.7.15: *Fast energy correction scheme (feedback). Two back-phased cavities are used to cancel the energy spread induced along the bunch length from the (nominally) zero phase crossing of the RF*

existing cavities, and no additional (fast) cavities are required in the beam delivery section.

3.7.5.3 Beam Based Alignment

The most critical aspect of the initial final focus tuning is the alignment of the magnets. The strong quadrupoles and sextupoles within the system have tight tolerances that must be met in order to reduce optical aberrations below the acceptable values. Since the alignment tolerances are tighter than can probably be achieved using conventional survey techniques, the position of the magnets must be corrected using the beam based alignment techniques already successfully demonstrated at the FFTB [19].

By using individual power supplies, the excitation of each magnet is changed and the resulting difference in downstream orbit is recorded, from which the relative offset of the magnet with respect to the beam can be estimated. Each magnet (excluding dipoles) is placed on a remotely translatable stage (magnet mover) which enables the precise (sub-micron) positioning of the magnet [20]. With high precision BPMs , alignment precisions on the order of 3 μm should be feasible. The strong Sextupoles can be aligned using the waist shift techniques first demonstrated successfully at the SLC [21].

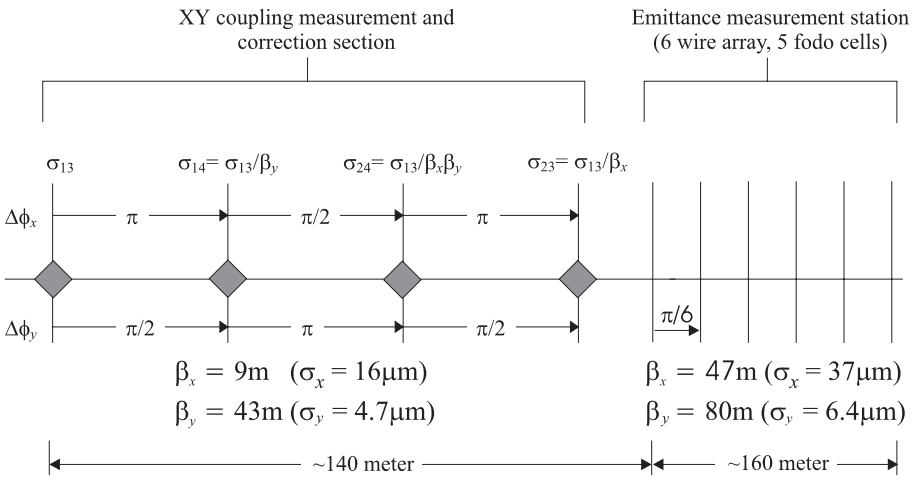


Figure 3.7.16: *Schematic of coupling measurement and correction section together with the downstream six wire array emittance measurement section. The fodo parameters for the emittance measurement station are chosen to give a 1% measurement resolution (assuming a 1 μm rms wire).*

3.7.5.4 Coupling Correction

With emittance ratios of 50:1 it is important to reduce any cross plane (X-Y) correlations in the 4-dimensional beam phase space. The linear correlations ($\langle xy \rangle$, $\langle xy' \rangle$, $\langle x'y \rangle$ and $\langle x'y' \rangle$ can be completely removed with four skew-quadrupoles placed at non-degenerate phases. There are two possible correction schemes that can be applied:

1. measure the correlations and adjust the four skew-quadrupoles accordingly.

2. generate a set of four orthogonal coupling knobs, and minimise the vertical emittance as a function of each in turn [23],

In TESLA, both possibilities are supported in a single tuning section. The lattice shown in figure 3.7.16 is chosen to make each single skew-quadrupole correct an independent coupling term (i.e an orthogonal knob) [23]. Immediately downstream of the skew-quadrupoles is a 6-wire emittance measurement section as described in [22], which is capable of a relative emittance measurement error of approximately 1%. The vertical emittance is minimised as a function of each skew-quadrupole strength in turn. It should be noted that the skew-quadrupoles are only orthogonal over a limited range around zero: should a large amount of correction be required, then the procedure must be iterated.

The orthogonal split-phase lattice utilised for the coupling correction also affords an elegant de-coupled measurement of the four individual X-Y correlations. Four wire scanners are placed at the same locations as the skew-quadrupole as discussed in [22]. Each wire scanner has an additional angle wire (the u-wire), which together with the horizontal and vertical wire measurements can be used to determine the $\langle xy \rangle$ correlation

of the beam at that wire scanner. At each wire scanner location, the measured $\langle xy \rangle$ correlation relates directly to an individual correlation term of the beam as measured at the first wire scanner (as shown in figure 3.7.16). Together with the downstream emittance measurement station, the coupling measurement station allows the complete determination of the 4-dimensional phase volume, the results of which can be used to determine the settings of the matching quadrupoles (skew and normal) with the aid of some suitable optics program, preferably on-line.

3.7.5.5 Beta Matching

By design, a final focus system for a linear collider contains large chromatic elements which are so positioned as to cancel the resulting aberrations. Given the delicate balance of this cancelation and the subsequent tolerances of the components, it is highly desirable to limit the amount of tuning which takes place within the final focus to a minimum: the IP tuning should be considered as a final fine tuning of the IP phase space (see section 3.7.5.6). To facilitate this, care must be taken to make sure that the incoming phase space is matched correctly to the final focus. The wire array system described in section 3.7.5.4 allows the determination of the full 4-dimensional phase volume, from which the setting of the a) skew-quadrupoles in the coupling correction section and b) the normal quadrupoles in the downstream β-matching section can be determined. As an additional cross check, a single wire scanner can be placed at the IP image point at the entrance of the horizontal chromatic correction section. If the final focus lattice is correctly adjusted, then a waist at this wire scanner corresponds directly to a waist at the IP. Waist scans can be made at the wire by driving six upstream quadrupoles in the correct (non-linear) fashion[26]. Assuming the coupling has already been corrected, the β-function at the image point can be obtained from the waist scan information, and a subsequent adjustment made to obtain the correct beam size. For the design vertical emittance the vertical beam size at the image point is 952 nm ($= 50 \times \sigma_y^*$), the measurement of which is well beyond the carbon fiber limit and will require laser technology or possibly a gradient undulator beam size monitor[24].

3.7.5.6 IP tuning

The final focus optics must first be checked and if necessary adjusted to given the correct box-structure lattice. Specifically the -I sections between sextupoles in the chromatic correction sections must be measured using the corrector-BPM null method outlined in [18]. Additional wire scanners at the intermediate IP image points can be also used to check the consistency of the optics (since they should all measure the same beam size). Unfortunately the beam sizes at these downstream image points have a large contribution from the chromaticity generated by the CCS sextupoles, and so it will be necessary to turn the sextupoles off during the measurement.

Final (linear) tuning of the IP phase space requires in principle six orthogonal knobs:

1. X and Y waist IP shift,

2. X and Y IP dispersion,

3. $\langle xy \rangle$ and $\langle x'y \rangle$ coupling.

For the last case (coupling), only the $\langle x'y \rangle$ is important, since there is no generating term for the $\langle xy \rangle$ correlation due to the single phase nature of the lattice (this assumes that the incoming coupling has been corrected using the techniques described in section 3.7.5.4). The remaining coupling term can be tuned out using a single skew-quadrupole placed close to the final doublet, or using the sextupole mover technique discussed below.

All five required tuning knobs can be easily generated using transverse displacement of the sextupole pairs in the CCS. This technique is particularly appealing since errors in alignment of the sextupoles are expected to be a significant source of the aberrations we wish to use them to correct. For a given sextupole pair, the transverse offset (horizontal and vertical) can be driven symmetrically or anti-symmetrically. The two sets of sextupoles movers can be ganged together to produce orthogonal X and Y knobs (for the waist and dispersion correction). Table 3.7.3 shows the various knob combinations while table 3.7.4 gives the numerical parameters. Exactly the same effect can be achieved by placing small normal and skew quadrupoles at the sextupoles and ganging the power supplies in pairs.

	symmetric	a-symmetric
horizontal	waist shifts	horizontal dispersion
vertical	$\langle x'y \rangle$ and $\langle xy' \rangle$	vertical dispersion

Table 3.7.3: *The four combinations of sextupole movers and their effect at the IP.*

		X	Y
X CCS	waist	830	2.4
	dispersion	6.2	0.33
	coupling	44.5	
Y CCS	waist	129	355
	dispersion	0.34	3.3
	coupling	215.	

Table 3.7.4: *Coefficients for the various mover combinations given in table 3.7.3. The coefficients are per unit sextupole offset.*

The typical amount by which the sextupoles must be displaced depends on the size of the scan range and subsequent correction necessary. As a good guideline, we assume

that a scan range is such that it increases the effective IP β-function by a factor of three (i.e. it increase the beam size by a factor of $\sqrt{3}$). Combining the above coefficients into orthogonal knobs, table 3.7.5 gives the amount of transverse motion of each sextupole to complete a 3β scan (for the coupling, only the vertical beam size is scanned using the vertical CCS sextupoles). Mechanical movers of a similar design as those developed for the FFTB [20] should be adequate given the requirements in table 3.7.5.

	CCS X (μm)	CCS Y (μm)	motion
x-waist	42.	-0.29	horizontal
Y-waist	-0.44	2.82	horizontal
x-dispersion	191	-19.1	horizontal
y-dispersion	-0.42	7.6	vertical
coupling		3.5	vertical

Table 3.7.5: *Expected motions of the sextupole pairs in the X and Y CCS to give orthogonal 3β scans at the IP.*

All the tuning described above requires some measurement of the beam size at the IP. While colliding beams are available, beam-beam scans can be used as a tuning signal, as can any fast luminosity signal. It is considered essential that a single beam diagnostic be available at or near the IP for initial tuning and fault diagnostics. A laser interferometer beam size monitor[25] will be placed some 80 cm along the beam axis from the actual IP. The focal point of the beam must be moved to the location of the the laser monitor requiring an adjustment of the β-match (a waist shift). Such a waist shift is easily produced by the sextupole pairs as previously mentioned (a symmetric horizontal displacement of 360 μm for X and 845 μm for Y).

3.7.6 Orbit and Spotsize Stabilisation

The alignment tolerances in the beam delivery system are some of the tightest in the machine. While it is expected that initial beam-based alignment can achieve these tolerances, the question still remains how stable the system is over time, given that ground motion will slowly displace all the components from their ideal positions. The effects of ground motion on the luminosity can be separated into three distinct time scales:

1. displacement of the beams relative to each other at the IP
2. dilution of the IP beam sizes by linear optics aberrations
3. dilution of the IP beam sizes by non-linear optics aberrations.

The relative displacement of the beam at the IP has the shortest time constant: if the effects of ground motion are left uncorrected, the luminosity drops rapidly as

the beams move out of collision. Assuming that this offset is corrected using the proposed fast feedback system (section 3.8.3), the next effect is that of linear optics aberrations increasing the beam size at the IP. As the ground motion displaces the magnets, the RMS of the beam trajectory through the system will increase rapidly, generating spurious dispersion, IP waist shifts and X-Y coupling (the latter two effects coming from the strong sextupoles in the chromatic correction sections). These effects can be dramatically reduced by routinely re-steering the beam delivery systems. In addition, the linear IP tuning knobs can be scanned during routine luminosity operation to remove any residual linear aberration not corrected by the steering (see section 3.7.5.6). Once (1) and (2) are corrected, the only remaining effects are non-linear aberrations (most notably non-linear dispersion). When these effects become large enough to reduce the luminosity below an acceptable limit, then the beam delivery system will require invasive tuning, and most probably a re-alignment of the magnets (using beam-based alignment).

To study the effects of ground motion and how effective the various tuning and orbit correction strategies are, a purpose built simulation tool known as MERLIN has been developed at DESY. The MERLIN model of the TESLA beam delivery system has the following key points:

1. both electron and positron systems are simulated, and the luminosity is calculated from the results of particle tracking through both (independent) models

2. ground motion (ATL) is applied to both electron and positron models in a consistent way to obtain the correct correlations between the two

3. components can be placed on girders (supports), and will then move together under the influence of ground motion

4. each magnet has an associated BPM, both of which are placed on magnet movers (which in turn are placed on a support).

Figure 3.7.17 show recent result of the MERLIN TESLA BDS model. The three graphs shown are the averaged results of running twenty random seeds for the ground motion, with $A = 10^{-5}\mu m^2/m/s$. The logarithmic time scale represents thirty days of running. The first graph (a) shows the effect of ground motion alone with no tuning, from which it is clear that the luminosity drops below an acceptable level within seconds. The second graph (b) shows how the effect of the IP steering feedback stabilises the luminosity by keeping the beams in collision. The critical time for luminosity loss is now of the order of one hour. The final graph (c) shows the effects of so-called one-to-one orbit correction: the beam is steered to zero the beam offset in all the BPMs (and consequently the centres of the their associated magnets). Particular care is taken to steer to the centre of all sextupole magnets and the final doublet (in the latter, a single BPM placed between the two doublet magnets is used). The result indicates that a 5% drop in the luminosity after after thirty to forty days can be expected (due to residual non-linear effects, most notably non-linear vertical dispersion). Result (b) suggests

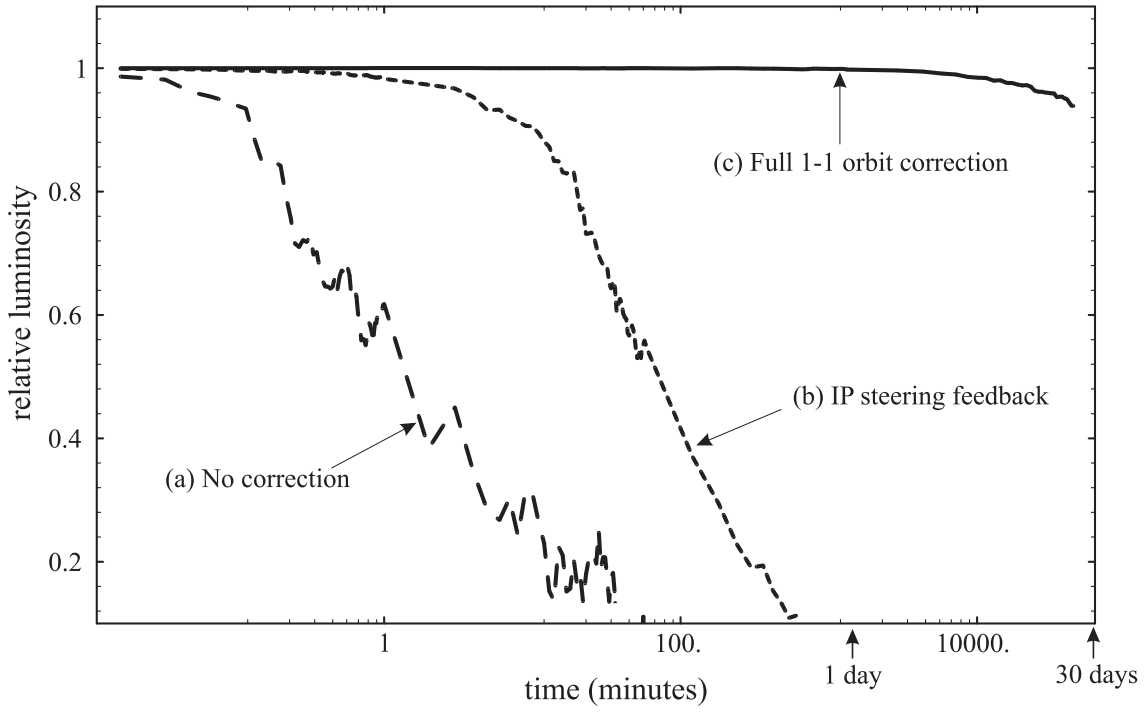


Figure 3.7.17: *Simulations of luminosity (relative to the design value) as a function of time. Three separate simulation runs are shown depicting (a) no correction, (b) IP beam offset correction (IP steering feedback) and finally (c) full 1-to-1 orbit correction of both electron and positron beams (see text for additional details).*

that such an orbit correction is only necessary approximately once every thirty minutes or so, giving the correction system plenty of time to average many pulses, eliminating the need for extremely high precision (i.e. nanometer-type) BPMs.

Given the current results, it is expected that the TESLA beam delivery system will require an invasive re-tuning (i.e. beam-based alignment) at intervals of approximately thirty to forty days. While this is acceptable, it is expected that use of better steering algorithms, tuning schemes and non-linear knobs can extend the time between re-alignment.

3.7.7 Beam-Beam Effects

The two colliding bunches focus each other at the interaction point, reducing the effective beam cross section compared to the nominal one. The actual luminosity is thus larger than the geometrical one by the pinch enhancement factor H_D. Due to the bending of their trajectories the particles emit a radiation called beamstrahlung. The resulting energy loss smears the luminosity spectrum and has thus to be limited, which in turn limits the achievable luminosity. The resulting spectrum is especially important for studies of the top quark at threshold and is discussed in the appropraite section in more detail.

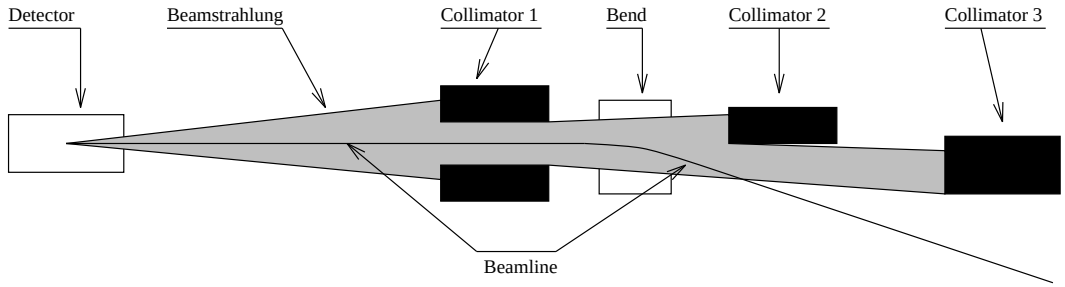


Figure 3.7.18: *Sketch of the beamstrahlung collimation*

The strong bending of the trajectories requires a numerical simulation to estimate the actual luminosity and the beamstrahlung. The numbers in the following are obtained using the computer code GUINEA-PIG [27]. For the standard beam parameters of TESLA the simulations yields a luminosity of $6 \cdot 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ ($H_D = 1.6$) and an average relative energy loss of the beam particles of $\delta = 2.5\%$. The total power of the beamstrahlung of one beam is about 200 kW.

The beamstrahlung is emitted into a small cone in the forward direction of the beams. Its angular spread is dominated by the angular spread of particles in the bunches during the interaction and the core is well below 500 μrad. The photons therefore pass the final quadrupoles safely. Since the beams collide head-on in TESLA, the magnets of the final focus system on the other side have to be protected from the beamstrahlung. Therefore three collimators are used, the first of which has to intercept most of the power—see Figure 3.7.18. The large distance between the final doublet and the closest magnets in the final focus system allows to collimate at a distance of 90 m from the interaction point. The collimator has a small inner radius of 4 mm to also collimate the synchrotron radiation emitted by the incoming beams, which form a significant background source in SLC.

The collimators consist of a titanium beam pipe with a thickness of 0.5 mm surrounded by water longitudinally flowing at a speed of 3 m/s. Simulations show that the stress induced when the photons heat up the titanium, is an order of magnitude below the endurance limit [27]. The direct temperature increase of the water is well below 10 K.

In the collimator neutrons are produced by the photons in the electro-magnetic showers. The flux at the position of the detector would be about $8 \cdot 10^{11}\,\mathrm{cm}^{-2}$ per year. The main detector be shielded with a concrete wall, 5.5 m reduce the flux to less than $4 \cdot 10^{6}\,\mathrm{cm}^{-2}\mathrm{y}^{-1}$ or 1 μSv/h. The vertex detector is less easy to shield since the wall must have a hole for the beams. With the mask inside the detector described in chapter 2 the flux of neutrons in the detector can be suppressed by more than three orders of magnitude to less than $10^{9}\,\mathrm{cm}^{-2}\mathrm{y}^{-1}$.

The beamstrahlung also increases the background produced in the interaction point. An important contribution to the background are the three processes creating e^+e^--pairs: $ee \to ee(e^+e^-)$, $\gamma e \to e(e^+e^-)$ and $\gamma\gamma \to e^+e^-$, where the real photons are due to beamstrahlung. The pair particles are produced mainly at small angles with respect to the beam axis but since their energy is small their trajectories are strongly bent by

E_{cm}	[GeV]	500	500 low ϵ_y	800	800 low ϵ_y	1600
L	$[10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}]$	6.0	8.4	5.7	11.1	22.4
δ	[%]	2.5	0.95	5.2	2.2	6.1
N_P	$[10^3]$	96	51	180	104	230
E_P	$[10^3\,\mathrm{GeV}]$	164	71	650	310	2200
$N_\perp$		31	15.8	46.5	28	50
$N'_\perp$		31	15	49	27	41
N_{Hadr}		0.13	0.06	0.46	0.19	0.87
N_{MJ}	$[10^{-2}]$	0.3	0.15	1.6	0.7	6.9

Table 3.7.6: *Luminosity L average energy loss δ and background for different TESLA parameter sets. N_P pair particles with a total energy of E_P are produced per bunch crossing, $N_\perp$ of which have transverse momentum in excess of* 20 MeV *and an angle with respect to the beam axis of more than* 150 mrad. *$N'_\perp$ is the number of particles per bunch crossing that will not enter the masks—without taking the deflection by the beams into account (which gives a value about* 20% *higher in the case of TESLA500). The number of hadronic events is N_{Hadr} including N_{MJ} minijet pairs with a transverse momentum of more than* 3.2 GeV.

the fields of the beams. If a particle moves in opposite direction of the bunch with the same sign of charge, it is deflected by the fields to a maximal angle depending on its energy.

Simulation with GUINEA-PIG predicts the production of about 10^5 particles with a total energy of $1.6 \cdot 10^5$ GeV per bunch crossing in TESLA. Most of these particles hit the final quadrupoles where they produce secondary photons and neutrons. The detector is shielded from the photons by a tungsten mask. In Table 3.7.6 the number of particles with a transverse momentum of more than 20 MeV and an angle of more than 150 mrad is given. This number is about 31 per bunch crossing. The numbers for other parameter sets can also be found in Table 3.7.6.

Single bremsstrahlung ee $\rightarrow$ eeγ is another process leading to low energy particles. It can be used to measure luminosity, but since the particles can loose a significant amount of energy they are strongly over-focussed by the oncoming beam and can therefore be lost in the final quadrupoles. The simulation yields a total energy deposition in the final quadrupoles by these particles of about $2 \cdot 10^5$ GeV per bunch crossing.

The rate of hadronic two photon events is 0.13 per bunch crossing with a photon-photon centre of mass energy in excess of 5 GeV. The visible energy of each of these events is about 10 GeV.

3.7.8 Beam Collimation

In order to avoid that the detector can be blinded by background produced from large amplitude halo particles in the interaction region, the trajectory amplitudes have to be restricted such that (a): the entire beam remains within the aperture of the final

quadrupoles in front of the IP and (b): the synchrotron light produced by particles with large offsets in the final doublet (FD for short) can pass freely through the aperture of the doublet downstream from the IP. The second condition is more stringent than the first and we lay out the collimation system according to this boundary condition. With its large aperture radius of 24 mm, the superconducting doublet can accomodate outgoing synchrotron light produced by particles with amplitudes as large as 12 horizontal and 48 vertical standard deviations for the standard parameters at 500 GeV. It is unlikely that in the TESLA linac tails reaching out to such large amplitudes can be populated. Scattering from gas molecules is almost not present in the superconducting structures and wakefield effects are small. Dark current is expected to be low with field emission being suppressed by careful processing of the cavities and furthermore field emitted particles which are accelerated are swept away by the next quadrupole(s) downstream. Beam halo produced in the damping ring can be scraped off before injection into the main linac. Therefore, including an efficient collimation system in the beam delivery design provides a large amount of safety to protect the interaction region from unwanted background.

3.7.8.1 Layout of the System

In principle, only large amplitude trajectories at the FD-phase are harmful, since the acceptance for IP-phase trajectories is orders of magnitude (in units of σ's) larger. FD-phase collimation can easily be provided at the high-β positions in the final focus system (FFS). However, being relatively close to the IP, muons originating from the FFS-collimators can reach the detector with high probability and the tolerable amount of beam halo scraped off in the FFS would be small. We therefore include a first collimation stage farther away from the IP, inserted between the end of the linac and the tuning and diagnostic section. Only a small fraction of halo particles escaping from the primary system will then hit the secondary collimators in the FFS.

In this scheme, one must take into account the possibility of trajectories changing amplitude and phase due to aberrations in the beam optical system between the 1st stage collimators and the FD. This leads to the requirement that collimation at both the FD and the IP phase is necessary and that the collimator aperture must be smaller by about a factor of 1.5 compared to the FD acceptance defined above. The resulting layout of the 1st stage system is sketched in fig. 3.7.19. Two pairs of spoilers, placed at high β and high dispersion points in the lattice, provide simultaneous collimation in x, y and $\Delta p/p$. The first pair collimates at the IP-phase, the second one at the FD-phase. We follow here a concept of mechanical collimation originally developed at SLAC [28]. The spoilers are thin (about 2 R.L. Titanium, see below) and act to increase the angular and momentum spread of the part of the beam outside the acceptance while 20 R.L. Cu-absorbers downstream catch the halo emerging from the spoilers. An overview of the spoiler and absorber parameters is given in table 3.7.7. The secondary collimators in the FFS are set to apertures corresponding to $12\sigma_x$ and $48\sigma_y$ and are safely in the shadow of the primary system.

The overall efficiency of the system is tested by particle tracking calculations using

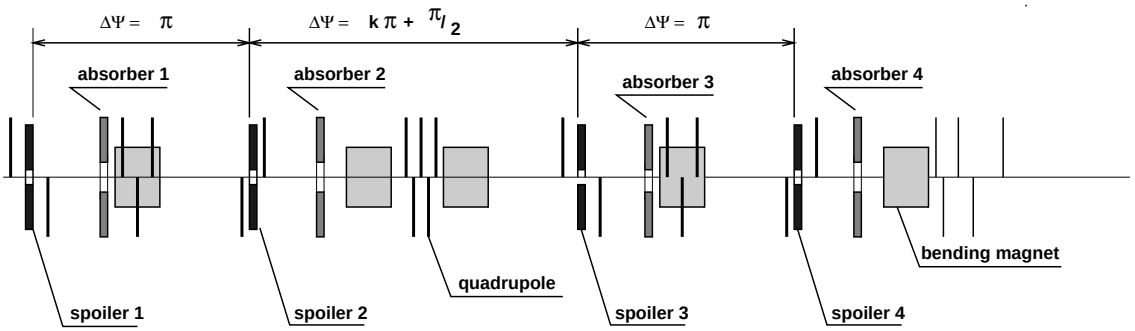


Figure 3.7.19: *Sketch of the 1st stage (main) collimation system.*

element	SPOILERS				ABSORBERS			
number	1	2	3	4	1	2	3	4
half gap g_x [mm]	1.3	1.3	1.3	1.3	2.30	2.73	2.30	2.73
half gap g_y [mm]	1.9	1.9	1.9	1.9	3.30	2.68	3.30	2.68
# of σ_x	8	8	8	8	30	30	30	30
# of σ_y	32	32	32	32	100	100	100	100
acceptance $\Delta p/p$ [%]	3.2	3.2	3.2	3.2				

Table 3.7.7: *Apertures of the main collimation system spoilers and absorbers.*

the STRUCT code [29]. The beam halo is simulated by a $|x|^{-1} \cdot |y|^{-1}$ particle distribution outside the collimator aperture. The efficiency of the primary stage is shown in fig. 3.7.20 as a function of spoiler length. For 2 R.L. Titanium, a fraction of 5×10^{-5} of the simulated beam halo escapes from the system. In the simulation all of these particles are stopped in the 2nd stage collimators. From the tracking results we estimate that less than 10^{-7} of beam halo particles have a chance to produce a hit in the interaction region quadrupoles with the conservative assumption that the efficiency of the 2nd stage collimators is 10^{-3}. For one hit per bunch crossing, a beam loss of up to about 0.03% in the collimators would be tolerable, far in excess of a conceivable halo population of the beam. The average synchrotron radiation power deposited in the interaction region for such an unrealistically high beam loss is estimated as 0.006 mW, roughly 6 orders of magnitude smaller than without any collimation.

3.7.8.2 Spoiler Protection

Though the beam is enlarged at the spoiler position by optical magnification it still exhibits an enormous power density of $P_{tot}/2\pi\sigma_x\sigma_y \approx 140$ MW/mm^2. Of course there is no material which could withstand such a power density for a longer time. On the other hand the spoilers are close to the beam and one has to take into account the possibility that a mis-steered beam hits the spoiler head on. In that case the enhanced loss rates trigger a kicker upstream which extracts the bunchtrain from the main beamline. The extracted beam is then guided to the main beam dump (see section 3.7.9). In order to allow a few μs reaction time for the safety system, the spoilers still have to accept the

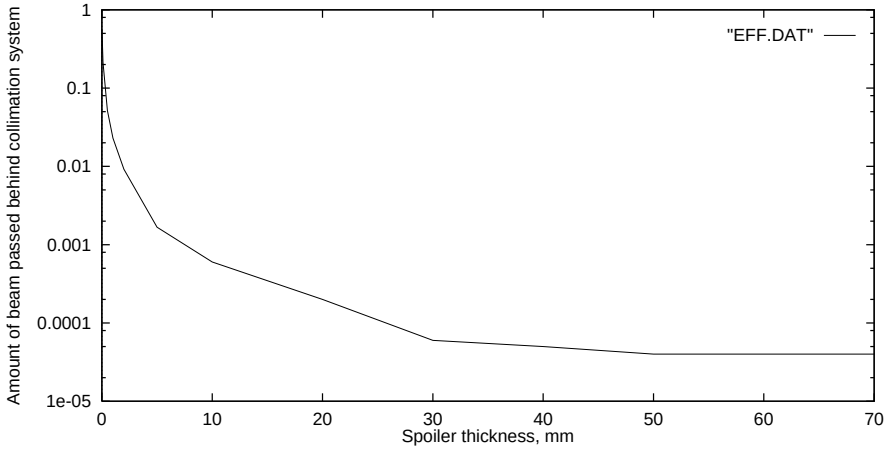


Figure 3.7.20: *Fraction of beam halo escaping from the primary collimation system as a function of spoiler thickness. For Titanium 2 R.L. correspond to 70 mm.*

	$(dE/dx)_{min}$ [MeV/cm]	ρ [g/cm^3]	$c(T = 298\text{K})$ [J/gK]	λ [W/cmK]	X_0 [cm]
Ti	6.85	4.54	2.0	0.15	3.56
C	4.0	2.20	0.71	2...3	19.3

Table 3.7.8: *Some properties of pyrolytic graphite and titanium.*

loss of several bunches.

We consider here two possibilities - a titanium spoiler and a graphite spoiler. Properties of both materials are given in table 3.7.8.

Due to the large radiation length the graphite spoiler is much longer compared to titanium. This leads to a larger transverse spread of the induced shower in graphite and results in a lower maximum energy deposition. In order to determine the instantaneous temperature jump due to the deposited energy one has to take into account the temperature dependence of the heat capacity:

$$\frac{dE}{dm} = \int_{T=T_0}^{T_0+\Delta T_{inst}} c(T)dT. \tag{3.7.1}$$

The instantaneous temperature jump ΔT_{inst} can be calculated by solving (3.7.1) numerically. Parameterizations for $c(T)$ can be found in [30], temperature curves for both materials are shown in fig. 3.7.21.

The heating induces thermal stresses in the material which, if too high, will lead to cracks and damage the spoiler. The temperature limit can be estimated from the induced stresses:

$$\sigma_{UTS} > \frac{1}{2}\alpha E \Delta T_{inst}. \tag{3.7.2}$$

Here one should take into account that the material dependent parameters σ_{UTS} ultimate tensile strength, α linear expansion coefficient and E elastic modulus are also

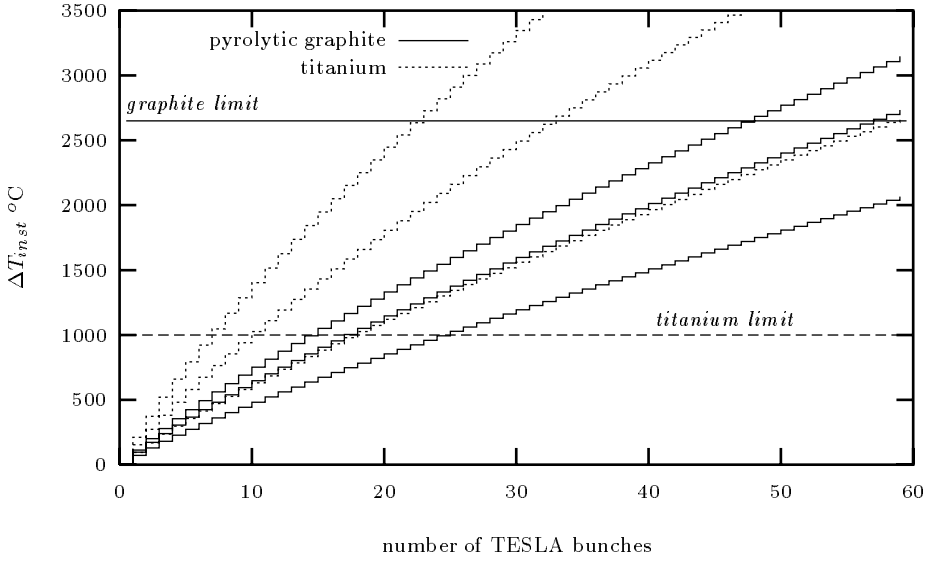


Figure 3.7.21: *Instantaneous temperature rise in graphite and titanium as a function of the number of TESLA bunches hitting the spoiler. The three curves for each material correspond to a thickness of 1.0, 1.5, 2.0 radiation lengths.*

functions of temperature. From (3.7.2) we estimate limits of $\approx$ 1000 °C for Ti and 2650 °C for pyrolytic graphite. From the temperature curves in Figure 3.7.21 it can be concluded that both materials are suited for a beam spoiler and can withstand a sufficient number of bunches. However, a two radiation length thick pyrolytic graphite spoiler can accept nearly 50 bunches on the same spot whereas the equivalent titanium spoiler stands only 6 bunches. The passage time of 6 bunches is 4.2 μs which is sufficient to trigger the safety system. Still a graphite spoiler would promise a higher inherent safety in case of accidental beam losses. Another advantageous property of graphite is its large thermal conductivity λ which becomes important if relatively large steady beam losses on the spoilers have to be handled. Such a situation is imaginable for instance during set-up and tuning of the machine.

3.7.8.3 Emittance Dilution by Wakefields

Ultrarelativistic particles moving in a perfectly conducting pipe experience no transverse kick since the forces caused by the electric and magnetic field cancel exactly. However, at discontinuities as the entrance or the exit of a beam spoiler this cancelation is distorted and particles in an off-axis bunch will experience an effective kick due to the induced fields. The so called geometric wake field kick varies along the longitudinal direction of the bunch and can be estimated for rectangular spoilers by [31]:

$$\Delta y'(z) = \theta_1 \exp\left(-\frac{z^2}{2\sigma_z^2}\right)^1; \qquad \theta_1 = \frac{\pi^2 r_e N_B}{2\sqrt{2\pi}\gamma\sigma_z}\frac{\langle\Delta y\rangle}{g}, \tag{3.7.3}$$

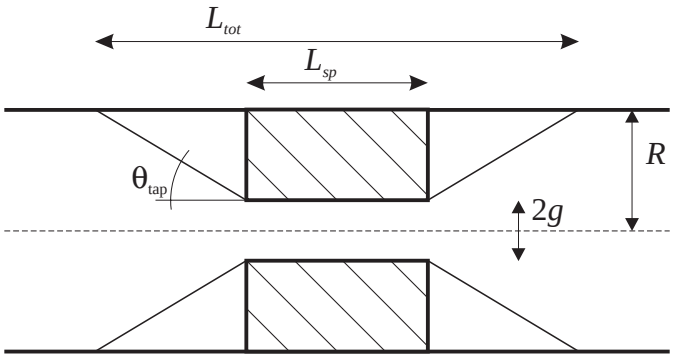


Figure 3.7.22: *Sketch of the tapered spoiler.*

where $r_e = 2.8 \cdot 10^{-15}$ m the classical electron radius, $\sigma_z = 0.7$ mm the rms bunch length, $N_B = 3.6 \cdot 10^{10}$ the bunch population, $\langle \Delta y \rangle$ the transverse bunch offset and g the half gap between the spoilers.

Since the strength of the kick varies along the bunch it dilutes the beam emittance. We assume that the mean kick can be corrected with steering elements and demand a luminosity reduction of less than 2 % for one σ_y offset at the spoiler:

$$\frac{\Delta L}{L} = \frac{1}{2}\frac{\Delta \varepsilon}{\varepsilon} = \frac{1}{2}\frac{\Delta y'^2_{rms}}{\sigma'^2} \leq 0.02, \text{ or } \quad \Delta y'_{rms} \leq \frac{1}{5}\sigma'.$$

The rms-wakefield kick $\Delta y'_{rms}$ is obtained by averaging of $\Delta y'$ and $\Delta y'^2$ over the longitudinal charge distribution (which is assumed here to be Gaussian):

$$\Delta y'_{rms} = \left(\langle \Delta y'^2 \rangle - \langle \Delta y' \rangle^2 \right)^{\frac{1}{2}} = \left(\frac{2-\sqrt{3}}{2\sqrt{3}} \right)^{\frac{1}{2}} \theta_1 = 0.278 \cdot \theta_1.$$

In order to reduce the geometric wakefield the spoiler can be tapered as shown in fig. 3.7.22. For taper angles $\theta_{tap} \ll 1$ the reduction factor is a linear function of θ_{tap} [32]:

$$k_1 = \frac{6\theta_{tap}}{\pi} \approx \frac{12(R-g)}{\pi(L_{tot}-L_{sp})}. \tag{3.7.4}$$

Of course the beam spoilers are not perfect conductors and for narrow gaps the resistive losses act back on the bunch. The so called resistive wall wakefield kick can be estimated by [33]:

$$\Delta y'(z) = \theta_2 \frac{1}{\sqrt{2\pi}} \int_{s=0}^{\infty} \frac{ds}{\sqrt{s}} \exp\left(-\frac{(z-s)^2}{2} \right); \qquad \theta_2 = \frac{\pi^2 r_e N_B L_{sp}}{2\gamma} \sqrt{\frac{\lambda_{sp}}{\pi \sigma_z}} \frac{\langle \Delta y \rangle}{g^3}, \tag{3.7.5}$$

where L_{sp} is the length of the spoiler and λ_{sp} is a characteristic length calculated from the conductivity σ_{sp} of the spoiler material:

$$\lambda = \frac{1}{c\mu_0\sigma_{sp}}.$$

	$\gamma\varepsilon$ [mm mrad]	β [m]	σ [μm]	σ' [nrad]	g [mm]	g/σ
X:	14	868	158	181	1.3	8
Y:	0.25	6797	59	8.7	1.9	32

Table 3.7.9: *Some beam related parameters at the spoiler locations.*

Here the rms-kick is given by [28]

$$\Delta y'_{rms} = \left(\frac{K(\sqrt{3}/2)}{2\sqrt{\pi}} - \frac{\Gamma^2(1/4)}{8\pi} \right)^{\frac{1}{2}} \cdot \theta_2 = 0.292 \cdot \theta_2 .$$

Tapering increases the length of the spoiler and therefore the resistive wall kick. The increase can be estimated by replacing the factor $\sqrt{\lambda_{sp}} L_{sp}/g^3$ in (3.7.5) by an integral over the total length:

$$\frac{\sqrt{\lambda_{sp}} L_{sp}}{g^3} \approx \int_{l=0}^{L_{tot}} \frac{\sqrt{\lambda(l)}}{g^3(l)}\, dl .$$

Thereby we obtain an enlargement factor for the rms-resistive wall wakefield kick of the tapered spoiler of (see geometry in fig. 3.7.22)

$$k_2 = 1 + \frac{1}{2} \sqrt{\frac{\lambda_{tap}}{\lambda_{sp}}} \left(\frac{L_{tot}}{L_{sp}} - 1 \right) \left(1 + \frac{g}{R} \right) \frac{g}{R} . \tag{3.7.6}$$

Here λ_{tap} is the characteristic length of the taper material which could be a thin copper foil for instance. Note that the heating of the taper material is not critical since the beam spot is enlarged by $1/\theta_{tap}$.

As long as the resistive wall wakefield kick is smaller than the geometric wakefield kick, tapering of the spoiler is helpful. The optimum total length of the tapered spoiler L_{tot} is achieved when both kicks have equal strength:

$$0.278 \cdot k_1(L_{tot}) \cdot \theta_1 = 0.292 \cdot k_2(L_{tot}) \cdot \theta_2 .$$

Now we can apply these formulas to the TESLA parameters. We consider three possible layouts for the spoiler - a spoiler made of pyrolytic graphite, the graphite spoiler with copper coating and a spoiler made of titanium. For all kinds of spoilers we assume a thickness of 2 radiation lengths and a radius of the beam-pipe of $R = 12.5$ mm. TESLA parameters relevant for the beam spoilers are given in table 3.7.9, material parameters in table 3.7.10 and results for the different layouts in table 3.7.11. It turns out that the graphite spoiler, even untapered, misses the goal of $\Delta y'_{rms} \leq \sigma'/5$ already due to the resistive wall effect (see table 3.7.11). So tapering will not improve the situation in that case. However, the Ti spoiler is uncritical if one applies some tapering. If coating of the inner surface of the graphite spoiler with a thin layer of copper or another material with high electrical conductivity is possible, also such a design could be used.

The conclusion from these considerations is that mechanical collimation of the TESLA beam at amplitudes of $8\sigma_x$ and $32\sigma_y$ can be realized with a tapered Ti spoiler.

	$C_\perp$	$C_\parallel$	Ti	Cu	Au
λ [m]	$2.65 \cdot 10^{-6}$	$1.06 \cdot 10^{-8}$	$1.47 \cdot 10^{-9}$	$4.5 \cdot 10^{-11}$	$6.1 \cdot 10^{-11}$

Table 3.7.10: *Characteristic length of several materials. Note the anisotropic properties of pyrolytic graphite. We assume that the material is aligned along the beam in the direction of the higher conductivity.*

	geometric		resistive wall		opt. tapered		opt. len. [m]	
	$\frac{\sigma'_x}{\Delta x'_{rms}}$	$\frac{\sigma'_y}{\Delta y'_{rms}}$	$\frac{\sigma'_x}{\Delta x'_{rms}}$	$\frac{\sigma'_y}{\Delta y'_{rms}}$	$\frac{\sigma'_x}{\Delta x'_{rms}}$	$\frac{\sigma'_y}{\Delta y'_{rms}}$	L^x_{tot}	L^y_{tot}
C	9.0	1.7	9.4	4.0	-	-	-	-
C,Cu	9.0	1.7	142.1	60.7	130.2	48.1	1.0	1.5
Ti	9.0	1.7	132.8	56.7	122.8	45.9	0.66	1.2

Table 3.7.11: *Wakefield kicks for different spoiler materials and taperings for a beam offset of 1 σ. The table contains the geometric and resistive wall kicks for the untapered spoiler and the geometric kicks for an optimal tapered spoiler. Given is always the value of $\sigma'/\Delta y'_{rms}$ which should be larger than 5 for a luminosity reduction of less than 2%. Furthermore the total length for an optimal tapered spoiler is given.*

Emittance growth from wakefields is practically negligible in that case. A graphite spoiler would be advantageous in view of its thermal properties, provided that the problem of coating with highly conductive material can be solved.

3.7.9 Beam Extraction and Dump

3.7.9.1 Extraction Beamline

After passing through the electrostatic/magnetic separator adjacent to the FD, the separation of the outgoing beam from the incoming beam is enhanced by septum magnets in order to stay clear from the next beamline elements of the FFS (see fig. 3.7.23).

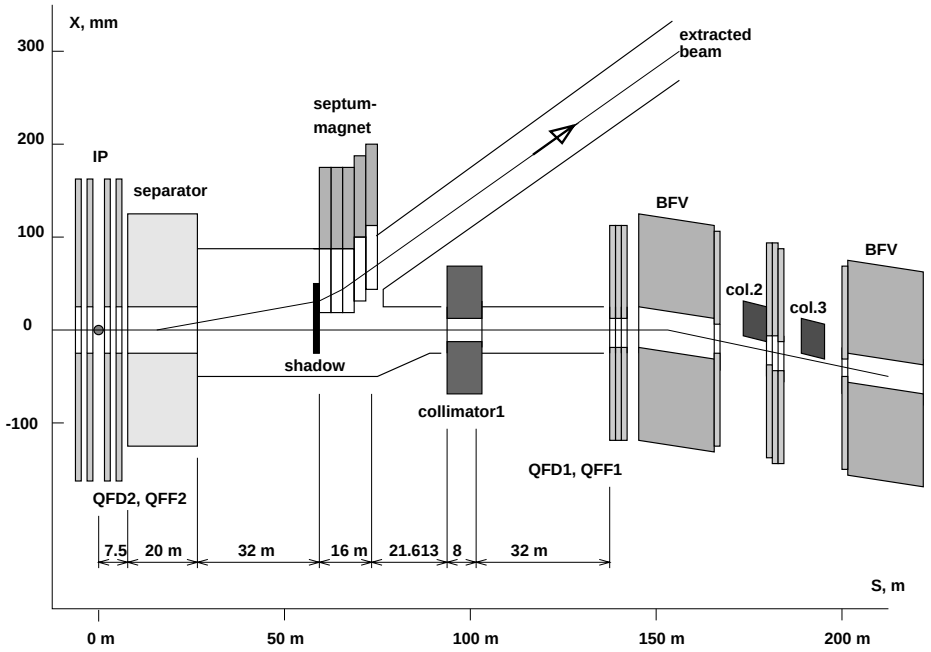


Figure 3.7.23: *Layout of the beam extraction system after the IP.*

A low-field design ($\approx$ 0.1 T) for the septum is chosen to allow for a thin septum bar. This helps to reduce the heat load from beamstrahlung and tails of the disrupted beam. The septum is shielded by a watercooled screen as sketched in fig. 3.7.24. The distribution of the heat load has been calculated using computer simulations of the beam-beam interaction and electromagnetic shower simulations in the screen, see fig. 3.7.25. After the septum, the positron beam is guided to the beam dump whereas the electron beam is captured in a special beam optical system, as described in section 3.4.5. The transfer line to the dump does not require any focussing elements, but a horizontal and vertical deflection of 10 mrad each are foreseen to increase the separation from the incoming beam and to have the muon beam emerging from the dump going downwards, away from the surface. The synchrotron radiation in these dipole magnets is sufficient to blow up the beamsize to about $2\,mm^2$ if the dump is placed 250 m away from the IP. Thus damage to the beamdump window and to the absorber block is excluded (see the following section), even for the case of an "unspoiled" beam which has not experienced the beam-beam interaction.

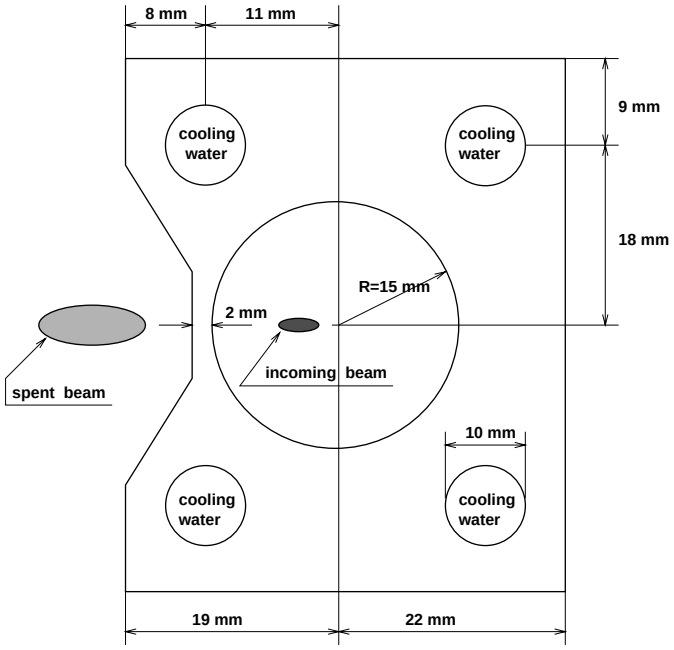


Figure 3.7.24: *Cross section of the septum magnet screen which absorbs losses from beamstrahlung and spent beam tails.*

3.7.9.2 Beam Dump

This section focusses on the question how to dispose of both the electron- as well as the positron beam after the collision. Based on shower, heat transfer and mechanical stress calculations a beam dump design capable of handling the $250\,GeV$ beams in TESLA is presented. After introducing the requirements on such a beam dump simple estimations on energy deposition and heat transfer answer the question of adequate absorber materials. Using shower simulation codes a beam absorber is calculated in more detail with respect to absorption efficiency and local energy densities the latter one giving a lower limit for the effective beam size when entering the absorber. The problem of induced radioactivity (production and handling of long lived isotopes) is mentioned very briefly, since the whole complex of radiation safety is discussed separately in section 3.11.

Requirements

The energy E_t carried in one bunchtrain, i.e. macropulse, is $1.6\,MJ$ for the TESLA scheme. This would suffice to melt $2.2\,kg$ of iron, $2.8\,kg$ of aluminium, $4.1\,kg$ of copper or $40\,kg$ of lead. Although being rather high, solutions exist to handle macropulse energies of that amount, e.g. at the proton machine of HERA, where the beam dump is capable of absorbing a complete fill, i.e. $820\,GeV$ protons with a total energy of $2.6\,MJ$ in a train of $20\,\mu s$, but this repeats typically less than one time in 8 hours only. Contrary to that the linear collider operates with high repetition rate, leading

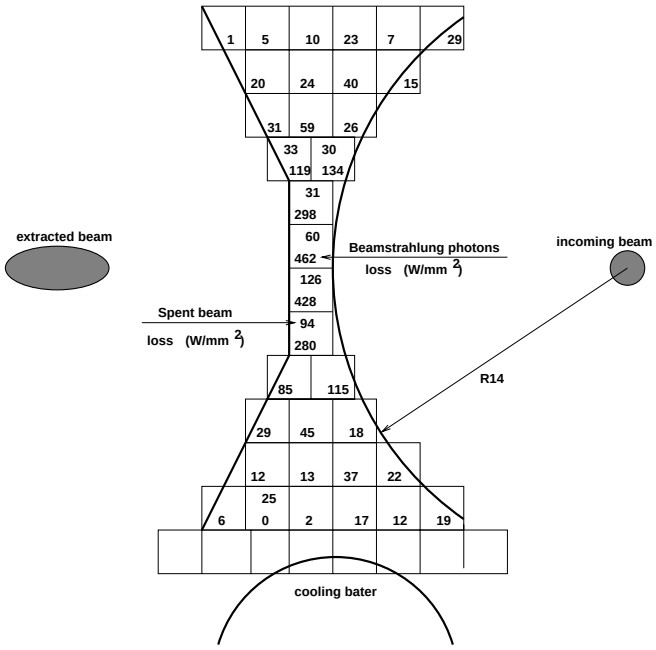


Figure 3.7.25: *Power distribution on the face of the septum screen from spent beam losses and beamstrahlung. In a 1 m long screen the temperature rise in steady state operation remains below 20 K (most of the produced shower particles are scattered out to the sides).*

to an enormous average beam power of about $8\,MW$. This represents not only a huge heat load, but is also a source of high radiation and isotope production. In addition to the thermal capabilities the integral radiation leakage should not exceed 1 % of the incident beam power, which is still $80\,kW$. Of course the shielding around the dump will take some fraction of it and the contribution from the leaking particles to the total activation is less than 1 % because of their low energy.

The decision on the location of the beam dump depends strongly on its size. From the point of view of civil engineering costs and simplicity a solution of putting the absorber inside the tunnel is preferred instead of building an extra hall for it. On the other hand, additional shielding around the tunnel would be required to avoid activation of ground water. The final layout for the dump region still needs to be worked out.

Assuming the absorber would fit into the tunnel figure3.7.26 illustrates a very schematic top view of the area between interaction point (IP) and beam dump. From the point of view of radiation and induced radioactivity there is no strong argument to put the beam dump outside if it would fit inside, since there will be already an area of higher dose rate level in the vicinity of the IP due to the presence of two collimation systems, that can not be avoided there. Getting different deflection in the separation region the low energy tail of the disrupted beam can not be transported through the beamline towards the dump. Distributed collimators have to scrape this fraction of the beam (see also section 3.4.5), which is between 10 % and 20 % of the total

intensity corresponding to a power in the order of $1\,MW$. In addition beamstrahlung in the order of $200\,kW$ coming from the IP must not hit components of the incoming beamline. Therefore it is equipped with beamstrahlung collimators.

Besides the absorber for the spent beam there need to be a possibility to dispose of the beam when not passing through the IP but being extracted after the main linac during its commissioning. If it is designed symmetrically the same absorber could be used for both purposes being hit either by the spent beam or the commissioning beam from opposite sides, but excluding simultaneous operation because of power reasons. Since this would not allow independent commissioning of one main linac while the other beam passes through the IP, two beam dumps with identical specification have to be put behind each other as indicated in figure 3.7.26 on either side of the IP. The commissioning dump will also be a part of an emergency dump system. Whenever the beam tends to get lost in the collimation system behind the main linac, which protects the experimental area, the particle source at the injector has to be switched off and the beam which is still in the main linac will be deflected onto this dump using fast extraction kickers at the end of the main linac.

Apart from these main dumps there have to be additional beam absorbers in the whole complex of the linear collider facility, e.g. at the FEL laboratory. Nevertheless the requirements on the main dump are the most severe ones. Having found a feasible concept for the main dump will automatically give a solution for the other dumps. From the previous discussion the requirements on such a main dump can be summarized as follows:

1. Withstand absorption of $1.6\,MJ$ energy pulses in combination with $8\,MW$ average power
2. Energy absorption efficiency $\geq 99\,\%$
3. Compact as possible to be put inside the tunnel
4. Production rate of dangerous and long lived isotopes as small as possible

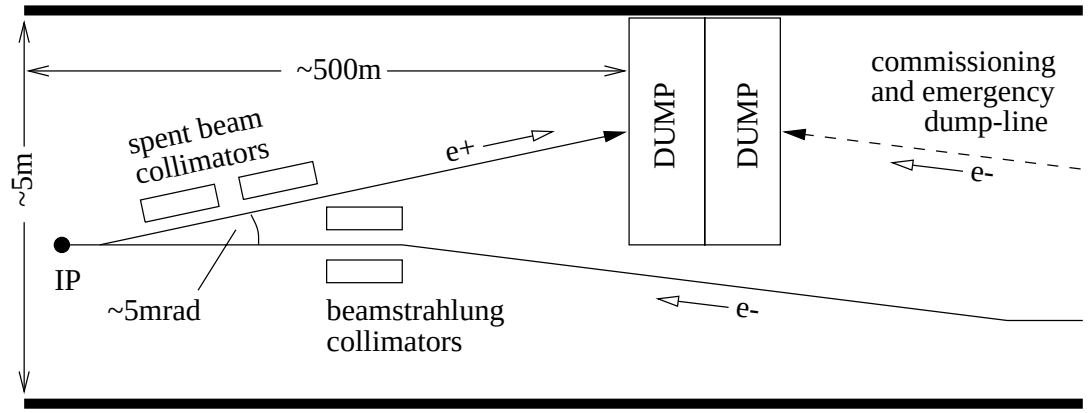


Figure 3.7.26: *Schematic topview of the area between IP and dump assuming the absorber fits inside the tunnel.*

5. Sufficient shielding to prevent groundwater and soil from being activated

Electromagnetic Showers

When hitting the absorber a high energy electron or positron beam initiates an electromagnetic shower (EMS), consisting of electrons, positrons and photons. In addition high energy photons can produce hadrons by means of photonuclear reactions. This is of major interest for the estimation of induced radioactivity, but less important for considerations on energy deposition since this is dominated by ionisation and exitation of atoms due to charged EMS components.

With respect to a cylindrical coordinate system an electron of energy E_0 enters a homogeneous material at $r = z = 0$ in positive z-direction. At first the combination of bremsstrahlung and pair production happening in average once in every radiation length for every electron, positron or photon leads to an exponential growth of the number of particles with z. Since the energy of the secondaries reduces similarly, this process stops when particles have reached the critical energy E_c, where ionisation reactions become dominant. The corresponding longitudinal position $z = t_{max}$ is called the shower maximum. That is where the longitudinal energy density defined as:

$$dE/dz = \int_{r=0}^{\infty} \int_{\phi=0}^{2\pi} dE/dV \; r \, d\phi \, dz$$

has its maximum value. Since this integration depends on the radial shower distribution it is obvious, that for a given material the maximum local energy density $(dE/dV)_{max} = \rho \cdot (dE/dm)_{max}$ appears at a position that depends on the incoming beamsize, while t_{max} is independent of it. With the radiation length X_0 and the critical energy E_c of a given material the development of the shower can be parametrized. Therefore the position of the shower maximum t_{max} and the particle multiplicity $M(t_{max})$ at that point can be expressed as [34]:

$$\begin{aligned} t_{max}(E_0) &= 1.01 \cdot \left(ln(\tfrac{E_0}{E_c}) - 1 \right) \cdot X_0 \\ M(t_{max}, E_0) &= 0.31 \cdot \tfrac{E_0}{E_c} \cdot \left(ln(\tfrac{E_0}{E_c}) - 0.37 \right)^{-\frac{1}{2}} \end{aligned} \tag{3.7.7}$$

where E_0 is the energy of the primary electron resp. positron. Assuming that at the shower maximum all the particles are still minimum ionising a rough estimate on the longitudinal energy density $(dE/dz)_{max}$ at the shower maximum created by one primary particle can be given as:

$$(dE/dz)_{max} \approx M(t_{max}, E_0) \cdot (dE/dz)_{min} \tag{3.7.8}$$

where $(dE/dz)_{min}$ is the energy deposited by one minimum ionising particle. The radial size of the shower is characterized by the Molière radius $R_M = 21.2MeV \cdot X_0/E_c$. In order to contain 98 % (1 % energy leakage radially and longitudinally) of the primary energy a cylindrical absorber has to have a radius $R_{99\%}$ and a length $L_{99\%}$ given by [35]:

$$L_{99\%} = \left[1.52 \cdot ln(\tfrac{E_0}{MeV}) - 4.1 \cdot ln(\tfrac{E_c}{MeV}) + 17.6 \right] \cdot X_0 \quad \text{and} \quad R_{99\%} \approx 5 \cdot R_M$$

		C	Al	Fe	Cu	Pb	H_2O
$X_0\ [cm]$		25.1	8.89	1.76	1.44	0.56	36.1
$R_M\ [cm]$		7.01	4.71	1.78	1.62	1.56	9.53
$E_c\ [MeV]$		75.9	40.0	21.0	18.8	7.6	80.3
$(dE/dz)_{min}\ [MeV/cm]$		3.03	4.37	11.6	12.9	11.7	2.03
$L_{99\%}\ [cm]$		470	190	42	35	16	670
$R_{99\%}\ [cm]$		35	24	8.9	8.1	7.8	48
$\rho\ [g/cm^3]$		1.7	2.7	7.87	8.96	11.35	1.0
$m\ [kg]$		3080	930	82	65	35	4850
$t_{max}\ [cm]$	from eq. 3.7.7	180	69	15	12	5.3	257
	MARS	200	80		14		260
$(\frac{dE}{dz})_{max}\ [\frac{GeV}{cm}\ per\ e]$	from eq. 3.7.7	1.11	2.93	14.3	17.6	37.7	0.707
	MARS	1.1	2.8		16.8		0.8

Table 3.7.12: *Basic properties of an electromagnetic shower initiated by a primary particle of energy $E_0 = 250\,GeV$ in different materials.*

From the given expressions most of the basic properties of the shower can be estimated. They are listed in table 3.7.12 for most likely absorber materials, where t_{max} and $(dE/dz)_{max}$ are compared with the result of the simulation code MARS, which has been developed over many years at IHEP (Protvino, Russia) and FNAL (Batavia, USA). As for calculating the maximum steady state temperature level, which is determined by $(dE/dz)_{max}$ a simulation code is not required since the calculated result agrees within 15 %. But when dealing with local energy densities dE/dm, designing a composite absorber or including hadronic effects the simulation can not be avoided. In addition to the shower size for 98 % energy containment the mass $m = \rho \cdot \pi \cdot R^2_{99\%} \cdot L_{99\%}$ of a corresponding absorber is given as well. Absorbers of that size would take 98 % of the shower but whether they withstand heating due to local energy densities and average power is subject to the following section 3.7.9.2. This is easy to judge in the case of the Pb absorber. With a mass of $35\,kg$ only it will be completely melted by the energy of one bunchtrain. It has to be mentioned, that the graphite parameters are valid for a material coming from a pyrolytic production process. Therefore the density is 75.9 % of the theoretical value only. As a consequence the radiation length is increased and the minimum ionisation loss is decreased by the same factor.

Basic Considerations on Heating and Choice of Material

The incident particle leads to a certain distribution of deposited energy with a maximum local energy density $(dE/dm)_{max}$ somewhere on the shower axis. Even during the long passage time $T_t = 800\,\mu s$ of a bunchtrain, heat conduction is negligible since the diffusion length $< d >$ given by:

$$< d >= \sqrt{\frac{\lambda \cdot T_t}{\rho \cdot c}} \quad \text{with} \quad \lambda := \text{heat conductivity} \quad c := \text{specific heat}$$

is only $0.3\,mm$ for copper or $0.26\,mm$ for graphite, which is small compared to the radial shower size respectively the required beamsize at the dump entrance as will be seen later. Therefore the instantaneous temperature rise ΔT_{inst} is directly proportional to the density of the deposited energy. Its maximum value $(\Delta T_{inst})_{max}$ is given by:

$$N_t \cdot (dE/dm)_{max} = \int_{T_0}^{T_0+(\Delta T_{inst})_{max}} c(T)\,dT \tag{3.7.9}$$

where T_0 is the temperature just before the passage of the bunchtrain. The integral form takes a temperature dependent specific heat $c(T)$ into account, which is quite important in the case of graphite, where it can be described using the following fit function:

$$c(T) = 1.48 \cdot \tfrac{J}{g \cdot K} \cdot (1.44 - e^{-\frac{T}{511^\circ C}}) \quad \text{with T in degree celcius}$$

After the passage of the bunchtrain this temperature on the shower axis decays until the next bunchtrain arrives. Therefore the temperature varies with time in a sawtooth-like manner. On the other hand for a continuous beam with the same average power as for the pulsed one an equilibrium temperature drop ΔT_{eq} is expected between the shower axis and the cooled edge.

Let us first deal with the equilibrium situation, i.e. how to get rid of the average beam power. For all solids besides their heat conductivity this problem is determined by the area and the thickness through which the heat has to flow. We assume that the beam is evenly distributed along a circle with radius r_1 on the face of a cylindrical dump. In order to contain the shower radially the outer radius r_2 has to be $r_2 = r_1 + R_{99\%}$. Assuming that all the deposited energy flows radially from r_1 to r_2 the temperature drop $(\Delta T_{eq})_{max}$ between both radii at the position of the shower maximum can be given as:

$$(\Delta T_{eq})_{max} = (\tfrac{dP}{dz})_{max} \cdot \tfrac{1}{2\pi\lambda} \cdot ln(r_2/r_1) \quad \text{with} \quad (\tfrac{dP}{dz})_{max} = \tfrac{I_{ave}}{e} \cdot (\tfrac{dE}{dz})_{max} \tag{3.7.10}$$

For each material of interest we set a maximum allowed temperature difference ΔT_{max}, either given by its plasticity limit $\sigma_{0.2}$ or by 20 % of its melting temperature T_{melt}, depending which of both represents the lower limit, i.e.:

$$\Delta T_{max} \approx \text{Min}\,[\sigma_{0.2}/(E \cdot \alpha) \text{ or } 0.2 \cdot (T_{melt} - 20^\circ C)]$$

where E is the elastic modulus and α the coefficient of linear thermal expansion.

For an average beamcurrent $I_{ave} = 32\,\mu A$, the power density at the shower maximum $(dP/dz)_{max}$ was derived from the calculated values of $(dE/dz)_{max}$ according to equation 3.7.8 as listed in table 3.7.12. They vary between $36\,kW/cm$ (C) and $1.2\,MW/cm$ (Pb). From these values the required radius $r_1 = r_2 - R_{99\%}$ and the resulting outer diameter of a dump $D = 2 \cdot (r_1 + R_{99\%})$ in order to keep $(\Delta T_{eq})_{max} \leq \Delta T_{max}$ can be calculated according to equation 3.7.10. The results are listed in table 3.7.13. Except for graphite the diameter of all other materials exceeding several $10\,m$ excludes their installation inside the tunnel and is even far off from a reasonable limit to be built

	C	Al	Fe	Cu	Pb
$(dP/dz)_{max}$ $[kW/cm]$	36	94	458	560	1200
λ $[W/(cm \cdot K)]$	1.6	2.0	0.8	3.8	0.35
ΔT_{max} $[K]$	800	60	240	150	30
r_1 $[m]$	1.4	30	34	13	1400
$D = 2 \cdot (r_1 + R_{99\%})$ $[m]$	3.5	60	68	26	2800

Table 3.7.13: *Minimum required diameter D of a dump in order to keep the maximum equilibrium temperature drop $(\Delta T_{eq})_{max} \leq \Delta T_{max}$ for $E_0 = 250\,GeV$ and $I_{ave} = 32\,\mu A$.*

anyhow. On the one hand graphite has a high radiation length leading to moderate power densities. On the other hand its excellent thermomechanical properties allow a high working temperature around $800°C$ not determined by the plasticity limit as it is for the other materials. In addition there is no safety problem in handling and machining graphite and its relatively low mass number limits the variety of produced isotopes. This all makes graphite the best candidate for a dump built from a solid material. When estimating the diameter of the solid cylindrical dump only cooling at the outer surface at $r_1 + R_{99\%}$ was assumed. Since r_1 is large enough additional cooling at an inner surface at $r_1 - R_{99\%}$ is possible. This will either reduce the equilibrium temperature or it allows to decrease the radial size. Nevertheless a graphite based absorber will have an outer diameter between $2\,m$ and $3\,m$ determined by heat conduction not shower containment and a length given by $L_{99\%}$ of about $5\,m$.

This is completely different for a liquid absorber such as water. Having even a little higher radiation length than graphite it is also a good candidate for a dump design. With $23\,kW/cm$ the maximum power density is not significantly less than in graphite, but instead of distributing the beam across the absorber in order to increase the area for heat conduction, one can introduce a sufficiently high water flow transversely to the shower axis, so that the water in the core of the shower is exchanged within the repetition time of bunchtrains. Under these circumstances the average power will be transported to an external heat exchanger and the size of the absorber is purely determined by the 99 % shower containment constraint. This will give a diameter of $2 \cdot R_{99\%} \approx 1\,m$ and a length of $L_{99\%} \approx 7\,m$ in the case of water. Being at least a factor of 2 more compact radially the length is only 40 % increased with respect to a graphite absorber. Although this is advantageous there are qualitative differences that will be discussed at the end of this section.

Now we have to consider the implications due to instantaneous heating during the passage of one bunchtrain. As described in equation 3.7.9 the maximum temperature rise $(\Delta T_{inst})_{max}$ depends on the number of particles N_t in a bunchtrain and the maximum local energy density in the shower $(dE/dm)_{max}$ deposited by one particle. For a given material only the size of the incident beam can be varied in order to reduce the local energy density and thus the instantaneous heating. Assuming a $250\,GeV$ gaussian beam with an aspect ratio of 1:1 the maximum local energy density $(dE/dm)_{max}$ and its longitudinal location on the shower axis was calculated for graphite and water

Graphite								
$\sigma_x = \sigma_y$ $[mm]$	0.1	0.5	1	2	3	5	7.5	10
$(dE/dm)_{max}$ $[GeV/g\ per\ e]$	13	2.8	1.1	0.45	0.27	0.13	0.07	0.05
z-position of $(dE/dm)_{max}$ $[cm]$	50	80	100	125	130	150	165	175

Water						
$\sigma_x = \sigma_y$ $[mm]$	1	3	5	10	30	50
$(dE/dm)_{max}$ $[GeV/g\ per\ e]$	1.05	0.25	0.12	0.05	0.009	0.004
z-position of $(dE/dm)_{max}$ $[cm]$	145	190	205	220	230	250

Table 3.7.14: *Maximum local energy density normalized to one primary particle in a graphite or water absorber deposited by a* $250\,GeV$ *beam as a function of its transverse beamsize.*

as a function of the transverse beamsize $\sigma_x = \sigma_y$ by means of the shower simulation code MARS. From these results as listed in table 3.7.14 a lower limit for the required beamsize at the dump entrance can be given. Due to steady state heating the temperature T_0 before the passage of a bunchtrain will be about $400°C$ in a graphite based dump (see next section). The remaining margin up to the working temperature limit of $800°C$ can be allowed for instantaneous heating. Taking into account the temperature dependent specific heat of graphite the corresponding difference in enthalpy per mass unit is $660\,J/g$ for $T_0 = 400°C$ and $(\Delta T_{inst})_{max} = 400\,K$. For a bunchtrain with $N_t = 4 \cdot 10^{13}$ this requires a $(dE/dm)_{max}$ of less than $0.1\,GeV/g\ per\ e$ corresponding to a beam size with $\sigma_x = \sigma_y$ not less than about $6\,mm$.

In the case of water $T_0 = 40°C$ and $(\Delta T_{inst})_{max} = 40\,K$ are reasonable assumptions in order to prevent it from boiling. Therefore about $170\,J/g$ can be used for instantaneous heating. From the same considerations as above the relevant limits are $0.027\,GeV/g\ per\ e$ and $\sigma_x = \sigma_y \approx 20\,mm$. The estimates on the required beamsize have to be regarded as lower limits for an effective spotsize over which the particles of one bunchtrain have to be distributed at the dump entrance. In most cases being rather large this spotsize is not automatically achieved by the betafunction and the emittance of the beam at the position of the dump. Additional measures such as a defocussing quadrupole, a thin foil for spoiling the emittance, a fast sweeping system which distributes the beam within the time T_t of one bunchtrain passage or any combination of it have to be taken. From this point of view a water absorber requires more effort.

As proven in this section even for the high average power of $8\,MW$ an edge cooled solid dump made from graphite is possible. Water can be used as an alternative as proposed for the NLC [36]. On the one hand a water absorber is more compact and due to the main contribution from short lived isotopes the dose rate in $1\,m$ distance after $1\,year$ of $8\,MW$ operation decays rather rapidly within a waiting time of $3\,hours$ down to a level of $20 - 30\,mSv/h$ ($1\,Sv = 100\,rem$), which can be reduced by a shielding factor of 10^2 with $5\,cm$ of lead [37]. The corresponding dose rate for a graphite absorber is around $70\,mSv/h$, i.e. only a factor of 2 higher. This would allow access

or at least walking by at the dump region. The remaining dose rate is determined by $^7Be\,(t_{1/2} = 54\,days)$ but can be removed in a resin filter. On the other hand there are severe safety problems correlated with a water absorber. Having an annual production rate of $11.4\,Ci/MW$ the activity of tritium ($t_{1/2} = 12\,years$) after 1 year of $8\,MW$ operation is $88\,Ci$ equivalent to $33\,cm^3$ of $^3H_2$ gas. Since distributed in the whole system of water cooling special care has to be taken to prevent leaking of it. There is no conclusive approach how to handle the tritium. Dilution of course is a bad idea. The presently allowed concentration of $800\,\mu Ci/m^3$ which can be put into waste water would require $3.5\,l/s$ continous exchange of the absorberwater with fresh water. Another safety problem arises from the radiolysis of water. Due to a production rate of $0.3\,l/s$ per MW deposited beam power [38], $2.4\,l/s$ hydrogen is produced at $8\,MW$ operation, which can be recombined to H_2O catalytically. Again any leakage of the system has to be avoided because a concentration of 4 % H_2 in air is explosive. Being proportional to the directly deposited beam power all the water related problems are less by a factor of 100 for the cooling circuit of an edge cooled graphite dump. Although the amount of ^{3}H production will not be significantly different in graphite, it will like all other produced isotopes be bounded and localized inside the solid and not spread around in a fluid with the danger of getting lost by leakage. For this qualitative difference the graphite based solution is preferred. An optimized design using shower simulations is presented in the following section. The solution using water can be regarded as a backup.

Absorber Design Based on a Graphite Core

From considerations in the previous section we know that the graphite absorber has to look like a hollow cylinder with a radial thickness of Δr between an inner resp. outer radius of $r_1 \pm \Delta r/2$, while the beam is distributed along a circle with radius r_1 at the face of this dump. Putting a material of higher density and better thermal conductivity in the region of the transverse shower tails, where it can survive, will reduce both Δr and r_1. Longitudinally the dump is not sectioned. Several runs of MARS simulations were made to optimize this layout. The result is shown in figure 3.7.27. A graphite tube with $r_1 = 1\,m$ and a radial thickness of $10\,cm$ is embedded between an inner and outer radial layer of aluminium $10\,cm$ thick each. For its smaller radiation length (see table 3.7.12) only $5\,cm$ of copper can be used alternatively. This compound has a total radial thickness of $\Delta r = 30\,cm$ with Al respectively $\Delta r = 20\,cm$ with Cu, an outer diameter D of $2.3\,m$ resp. $2.2\,m$, a total mass of $27\,tons$ resp. $40\,tons$ and will be cooled at the inner and outer surface at $r_1 \pm \Delta r/2$. A total length of $L = 6\,m$ guarantees 99 % absorption of the primary energy in both cases.

Distribution of the beam along a circle with radius $r_1 = 1\,m$ can be achieved by means of an orthogonal pair of dipoles harmonically excited at the same frequency but shifted in phase by 90°. If placed $100\,m$ upstream from the dump the maximum deflecting angle of each magnet has to be $10\,mrad$, which requires an integrated field-strength of about $9\,Tm$ for a $250\,GeV/c$ beam. To provide a homogeneous azimuthal heat source the oscillation period of this slow sweeping system has to be significantly

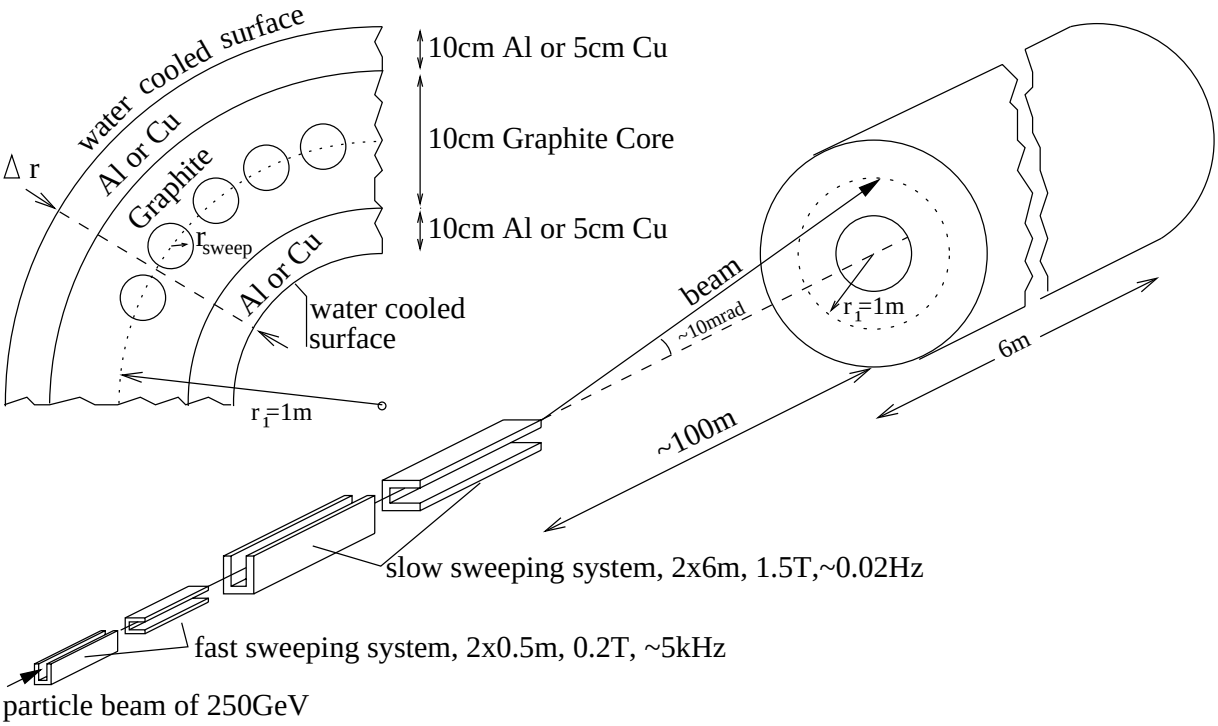


Figure 3.7.27: *Dump scheme based on a solid absorber with a graphite core.*

conduction in C between $100cm \leq r \leq 105cm$	transfer $C \rightarrow Al$ or Cu at $r = 105cm$	conduction in Al between $105cm \leq r \leq 115cm$	conduction in Cu between $105cm \leq r \leq 110cm$	transfer Cu or $Al \rightarrow H_2O$ at $r = 105$ or $110cm$
$190\,K$	$\approx 50\,K$	$135\,K$	$45\,K$	$\approx 50\,K$

Table 3.7.15: *Contributions to the equilibrium temperature rise due to heat conduction or heat transfer in the C-Al or C-Cu composite dump at the shower maximum for* $E_0 = 250\,GeV$ *and* $I_{ave} = 32\,\mu A$.

faster than thermal diffusion processes. Since the thermal time constant for this dump scheme is around $200\,s$ the frequency of the slow sweepers should be $1/50\,Hz$ or more.

The resulting equilibrium temperature rise at the shower maximum ($t_{max}(\mathrm{C}) \approx 2\,m$) for a $250\,GeV$ beam with $I_{ave} = 32\,\mu A$ is listed in table 3.7.15. The maximum heat flux at the water cooling boundary is roughly $27\,W/cm^2$. The temperature will rise by about $50\,K$ according to a typical heat transfer coefficient of $0.6\,\frac{W}{cm^2 \cdot K}$ for water at metallic walls. The graphite metal boundary is more crucial, because specific heat transfer is only about $0.2\,\frac{W}{cm^2 \cdot K}$ [39]. The effective surface has to be increased at least by a factor of 2, which was already taken into account when estimating a ΔT of $50\,K$ across this boundary. The technical design has to guarantee that the thermal contact of this boundary is not lost due to different expansion during warm up or cool down of the dump. Including the temperature drops from heat conduction and assuming

a temperature of $50°C$ for the cooling water, the graphite core at $r = 1\,m$ rises up to a temperature of $475°C$ when neighboured by aluminium. For its better thermal conductivity this temperature can be reduced to $385°C$ when using copper.

The following considerations on instantaneous heating are based on a beamsize with $\sigma_x = 3\,mm$ and $\sigma_y = 0.5\,mm$ at the dump entrance for a beam having not interacted (according to the preceding section the minimum beam cross section is $2\,mm^2$). From table 3.7.14 we know already that a minimum effective spotsize of $\sigma_x \cdot \sigma_y = 36\,mm^2$ is required in order to keep $(\Delta T_{inst})_{max} \leq 400\,K$. Therefore the above given beamsize is not large enough. A fast sweeping deflection system is proposed to increase the effective spotsize. Similar to the slow sweeping system it consists of a pair of orthogonal kicker magnets, again oscillating with a phase shift of 90°. It has to distribute the particles of one bunchtrain along a small circle with radius r_{sweep} within the bunchtrain passage time $T_t = 800\,\mu s$. Therefore the frequency of the fast sweeping system needs to be a few kHz. Table 3.7.16 gives the maximum local energy density $(dE/dm)_{max}$ normalized to one primary particle as a function of r_{sweep}. In order to

$r_{sweep}\ [mm]$	0	3	5	10	20
$(dE/dm)_{max}\ [GeV/cm\ per\ e]$	0.63	0.32	0.12	0.06	0.032

Table 3.7.16: *Maximum local energy density in graphite as a function of the sweep radius of the fast deflection system based on an incident beamsize with* $\sigma_x = 3\,mm$ *and* $\sigma_y = 0.5\,mm$.

keep $(\Delta T_{inst})_{max} \leq 400\,K$ less than $0.1\,GeV/cm\ per\ e$ and thus $r_{sweep} \approx 10\,mm$ is required. Located in the neighbourhood of the slow sweeping system a deflection angle of $0.1\,mrad$ is needed. The main parameters of the slow and fast deflection systems are summarized in table 3.7.17. The instantaneous temperature rise in the Al or Cu

	magnetic field in deflector	length of deflector	frequency $[Hz]$
slow deflection system	$1.5\,T$	$2 \times 6\,m$	$\geq 0.02\,Hz$
fast deflection system	$0.2\,T$	$2 \times 0.5\,m$	$\approx 5\,kHz$

Table 3.7.17: *Main parameter of both sweeping systems for a* $250\,GeV$ *beam.*

material is only about $20\,K$.

With this scheme the maximum temperature in graphite is kept below a maximum reasonable working temperature of $800°C$. As a consequence of the slow deflection system a large beampipe and if not directly flanged onto the dump a huge exit window is required. The angle of $10\,mrad$ needs to be taken into account by shaping the dump conically rather than cylindrically.

Bibliography

[1] Murray J.J., K.L. Brown and T. Fieguth, *The Completed Design of the SLC Final Focus System*, Proc.of Part. Acc. Conf., Washington, D.C., March 1987 and SLAC-PUB–4219 (Feb.1987).

[2] D. Burke for the FFTB Collaboration, *Results from the Final Focus Test Beam*, Proc. of the IVth European Particle Accelerator Conf., London 1994, Vol. I, p. 23.

[3] Rifflet J.M. et al., *Cryogenic and Mechanical Measurements of the first two LHC Lattice Quadrupoles Prototypes*, EPAC94 Conf., London, Great Britain, 1994.

[4] Grote H. and F.C. Iselin, *The MAD Program*, CERN/SL/90-13 (AP) Rev.3 (Jan.1993).

[5] B. Baklakov, P. Lebedev, V. Parkhomchuk, A. Sery, A. Sleptsov, V. Shiltsev, INP 91-15; Tech. Ph. **38**, 894, (1993).

[6] V. Juravlev, A. Sery, A. Sleptsov, W. Coosemans, G. Ramseier, I. Wilson, CERN SL/93-53, CLIC-Note 217, 1993.

[7] V. Juravlev, P. Lunev, A. Sery, A. Sleptsov, K. Honkavaara, R. Orava, E. Pietarinen, *Seismic Conditions in Finland and Stability Requirements for the Future Linear Collider*, HU-SEFT R 1995-01, Helsinki 1995.

[8] V. Shiltsev, B. Baklakov, P. Lebedev, J. Rossbach, C. Montag, *Measurement of Ground Vibrations and Orbit Motion at HERA*, DESY HERA 95-06, June 1995.

[9] Shigeru Takeda, K. Kudo, A. Akiyama, Y. Takeuchi, T. Katoh, Y. Kanazawa, S. Suzuki, *Slow Ground Motion and Large Future Accelerator*, Proc. EPAC 96, 1996.

[10] C. Adolphsen, G. Mazaheri, T. Slaton, in: Proc. of Workshop on Physics and Experiments with Linear Colliders, Tsukuba, 1995.

[11] A. Sery, O. Napoly, Phys. Rev. E. **53**, 5323, (1996).

[12] R. Brinkmann, J. Rossbach, *Observation of Orbit Motion in HERA Covering Eight Decades in Frequency*, Nucl.Instr.Meth.A **350**, 8, (1994).

[13] V. Shiltsev, *Space-Time Ground Diffusion: the ATL Law for Accelerators*, Proc. of Workshop on Accelerator Alignment, Tsukuba, 1995.

[14] R. Pitthan, *Re-Alignment: It is the Tunnel Floor which Moves, Isn't it?*, SLAC-PUB 95-7043, 1995.

[15] A. Sery, A. Mosnier, DAPNIA/SEA/96-06, CEA Saclay, 1996.

[16] C. Montag,*Active Stabilisation of Quadrupole Vibrations for Linear Colliders*, DESY 96-053, 1996.

[17] A. Sery,*Ground Motion Studies with Respect to Linac Performance*, Proc. of the Linac 96 Conference, Geneve 1996.

[18] P. Tenenbaum, D. Burke, S. Hartman, R. Helm, J. Irwin, R. Iverson, P. Raimondi, W. Spence, V. Bharadwaj, M. Halling, J. Holt, J. Buon, J. Jeanjean, F. Le Diberder, V. Lepeltier, P. Puzo, K. Oide, T. Shintake, and N. Yamamoto, *Beam Based Optical Tuning of the Final Focus Test Beam*, SLAC-PUB-95-6770, May 1995. 3pp. Presented at 16th IEEE Particle Accelerator Conference (PAC 95) and International Conference on High Energy Accelerators, Dallas, Texas, 1-5 May 1995.

[19] P. Tenenbaum, D. Burke, R. Helm, J. Irwin, P. Raimondi, K. Oide, and K. Flöttmann, *Beam Based Magnetic Alignment of the Final Focus Test Beam*, SLAC-PUB-95-6769, May 1995. 3pp. Presented at 16th IEEE Particle Accelerator Conference (PAC 95) and International Conference on High Energy Accelerators, Dallas, Texas, 1-5 May 1995.)

[20] G. Bowden, P. Holik, S.R. Wagner, G. Heimlinger, and R. Settles, *Precision Magnet Movers for the Final Focus Test Beam*, Nucl. Instrum. Meth.A368:579-592,1996.

[21] P. Emma, J. Irwin, N. Phinney, P. Raimondi, N. Toge, N.J. Walker and V. Ziemann, *Beam Based Alignment of the SLC Final Focus Sextupoles*, SLAC-PUB-6209, May 1993. 3pp. Presented at 1993 Particle Accelerator Conference (PAC 93), Washington, DC, 17-20 May 1993.

[22] N. J. Walker, *Beam Phase Space and Energy Measurements in TESLA and SBLC*, to be published as a TESLA report.

[23] *Zeroth-Order Design Report for the Next Linear Collider*, SLAC Report 474, May 1996.

[24] E. G. Bessonov, J. Pflüger, G.-A. Voss and N. J. Walker, *Beam Size Measurements in a Linear Collider using an X-ray Gradient Undulator*, DESY internal report DESY M 96-18, September 1996.

[25] T. Shintake, K. Oide, N. Yamamoto, A. Hayakawa, Y. Ozaki, D. Burke, R.C. Field, S. Hartman, R. Iverson, P. Tenenbaum, and D. Walz, *Experiments of Nanometer Spot Size Monitor at FFTB using Laser Interferometry*, KEK-PREPRINT-95-46, May 1995. 3pp. Contributed paper at 1995 Particle Accelerator Conference and International Conference on High-Energy Accelerators, May 1 - 5,1995, Dallas, Texas, U.S.A. "Zeroth-Order Design Report for the Next Linear Collider", SLAC Report 474, May 1996.

[26] N.J. Walker, J. Irwin, and M. Woodley, *Global Tuning Knobs for the SLC Final Focus*, SLAC-PUB-6207, Apr 1993. 3pp. Presented at 1993 Particle Accelerator Conference (PAC 93), Washington, DC, 17-20 May 1993.

[27] D. Schulte, *Study of Electromagnetic and Hadronic Background in the Interaction Region of TESLA.*, Ph.D. thesis, University of Hamburg 1996.

[28] N.Merminga, J.Irwin, R.Helm, R.D.Ruth, *Collimation Systems for a TEV Linear Collider*, SLAC Pub. 5165 Rev. (1994)

[29] I.Baishev, A.Drozhdin, N.Mokhov, *STRUCT Program User's Reference Manuel*, SSC-MAN-0034, February 1994.

[30] Handbook of Chemistry and Physics, Chemical Rubber Publishing Company (1976)

[31] K.L.F. Bane, P.L. Morton, *Deflection by the Image Current and Charges of a Beam Scraper*, SLAC Pub. 3983 (1986)

[32] Kaoru Yokoya, *Impedance of Slowly Tapered Structures*, CERN SL/90-88 (1988)

[33] A.Chao, *Physics of Collective Beam Instabilities in High Energy Accelerators*, J.Wiley & Sons, New York (1993)

[34] B. Rossi, *High-Energy Particles*, Prentice-Hall, New York, 1952.

[35] I.S. Baishev, M.A. Maslov, M. Seidel, *Design Study of a Beam Dump for the TESLA and S-Band Test Facilities at DESY*, DESY-TESLA-95-10, 1995.

[36] NLC-ZDR Design Group, *Zeroth Order Design Report for the Next Linear Collider*, SLAC, February 1996.

[37] K. Tesch, *A Note on Residual Radioactivity of Beam Absorbers for a Linear Collider*, DESY, March 1996.

[38] D.R. Walz, A. McFarlane, E. Lewandowski, *Beam Dumps, Stoppers and Faraday Cups at SLC*, SLAC-Pub. 4967, 1989.

[39] J. Kidd et al.,PAC IX, Washington, 1981.

3.8 Instrumentation and Controls

With the passage of time, the instrumentation and control systems have come to play a role of steadily increasing importance for the achievement of accelerator performance goals. They are no longer just the means for intermittent operator intervention; rather, they provide the complex of feedback loops necessary for performance. In this respect, the linear collider resembles the "fly-by-wire" characteristics of modern high performance aircraft. The low frequency TESLA approach ameliorates but does not avoid this situation.

The discussion to follow is comprised of three topics. The next subsection treats beam-related instrumentation: the sources of beam information and input to the feedback loops. Next the major orbit feedback schemes for TESLA are outlined. Finally, the approach to the control system is described. Despite the volatility of computer hardware and software, there are aspects that are likely to endure and these receive emphasis.

3.8.1 Beam Diagnostics

We begin with a summary of some of the basic ideas and numbers which lead to instrumentation needs, requirements, and device descriptions. Sections elsewhere in this report where detailed information can be found are referenced. This introduction concludes with a tabulation of issues and approaches.

The Linac

For the linac two main requirements have been addressed. They are the preservation or control of the transverse and longitudinal emittances.

Transverse emittance (Section 3.3) is preserved by control of alignment of the quadrupoles, BPM's and cavities. In addition energy variation and energy spread must be controlled (i.e. longitudinal emittance).

RMS alignment values of 100, 100, 500 μm for quad, BPM, and cavity have been chosen. These values are relative to the ideal beam line and are assumed to be valid over distances of up to 600 m. Beyond this distance, typical of a betatron wave length, sensitivities decrease rapidly.

Using the above alignment errors, single bunch vertical emittance growths of $\approx$ 12% are found on the average using the "one to one" beam steering technique. (i.e., steering from the center of one detector to another). In the more time consuming beam-based alignment method, the center of the beam is found relative to the center of the quad and to the BPM by noting beam position change downstream and its null point as the quad strength is varied. Then the 12% emittance growth can be achieved if the BPM resolution is 15 μm even when the quad position rms is as large as 500 μm (Section 3.3). With better quad alignment, as will likely be the case (Section 3.9), the beam-based method allows to further decrease emittance dilution which is desirable for upgrade of the luminosity. Therefore a BPM resolution and stability of 15 μm for the linac has

been chosen. In principle this accuracy has to be achieved only by averaging of entire bunch trains, but there seems no reason why individual bunch measurements can not be made and changes in the bunch train positions observed.

The effect of injection errors does not seem critical as a one sigma beam offset at the start of the linac makes only 6% vertical emittance dilution.

The above emittance growths are based on a bunch energy spread of 5×10^{-4} that results from the adjusting the bunch time to run at 3.4° before RF crest and makes for best cancellation of the logitudinal wake field. Multibunch emittance growths with inter-bunch energy spreads of 3.5×10^{-4} show very little difference from the single bunch results (Section 3.3).

The effect of ground motion on orbit stability has been estimated. Beam resteering (1 to 1) should be performed every hour in order to limit vertical emittance dilution to less than 20% for ground motion frequencies < 0.25Hz (i.e., 5Hz/20 - the high cutoff frequency of the conventional beam steering feedback system). Ground vibration can prevent the beams from colliding at the IP. As the jitter tolerance on the linac components would need to be < 80 nm for a 3% luminosity reduction, it is planned that a fast kicker will resteer the bunches during each pulse. The vibration tolerance on quadrupoles which would cause dispersive spot size growth and result in 3% luminosity reduction is $1\,\mu$m. This is an order of magnitude larger than the expected vibration levels.

The RF

The bunch length is controlled by the damping ring and compressor system. The energy and energy spread within the bunch and throughout the bunch train is controlled by the RF cavities phase and amplitude, their systematic errors and variations. The goals of the RF system regulation loops (Section 3.2.4.8) are to regulate the amplitude of each 32 cavity vector sum system to an rms amplitude of $5 \cdot 10^{-3}$ and an rms phase of 0.5° for uncorrelated field fluctuations. This should provide an intra-bunch rms energy variation term of $\approx 3 \cdot 10^{-4}$, smaller than the intra-bunch spread. Fluctuations in the phase will produce imperfect wake compensation leading to greater intra-bunch spread.

Coherent field fluctuations occuring through the system can arrise from bunch charge fluctuations. A 10% bunch charge flucation will product a $1.4 \cdot 10^{-4}$ gradient fluctuation for the next bunch. In general 10% charge fluctuations will produce energy fluctuations at the 10^{-3} level.

Systematic absolute gradient amplitude and phase measurement pecission, calibration and control against drift may be difficult at the 10^{-4} beam energy level. Ways to monitor the beam energy, energy spread, bunch length and bunch charge must be provided.

The beam energy can be measured to a resolution of 10^{-4} and corrected either by a fast RF system in the BDS Collimation Section or with slower linac cavity compensation. BPM resolution of 3-5 μm is required (Section 3.7.5.2). The beam optics in the BDS are also well suited to determine the intra-bunch energy spread. By measuring the beam width at a position of large dispersion and relatively small β-function,

the achievable accuracy for determining the relative energy spread is of the order of 10^{-4} (10-20% of the design bunch energy spread) with a 5 μm beam size measurement resolution.

The Beam Delivery System

The beam delivery system (BDS) reaches from the end of the linac to the IP and is 1100 m long. The about 50 quads, sextuploes, and 16 bends per beam will each be provided with BPM's attached to the magnets and the assemblies placed on magnet movers for beam based alignment and tuning. The beam pipe diameter is 10 mm. BPM precision and resolution has been specified at 3 μm throughout.

To make the following discussion easier to follow, functional sections are in order (see Section 3.7):

- (LMS) linac matching section,
- (CS) collimation section,
- (TDS) tuning and diagnostic section which is composed of the (SQSC) skew quad correction section and the (EMS) emittance measurement section,
- (BMS) beta matching section,
- (CCS) chromaticity correction section,
- (FT) final transformer.

The CS collimation section contains high beta and high dispersion regions with collimator pairs, energy measurement and correction, energy spread measurement, beam loss measurement for protection of the collimators.

The TDS tuning and diagnostic section provides for measurement and correction of transverse coupling in the SQSC and for measurement and minimization of emittance in the EMS. This is the major measurement and tuning area which sets, measures and adjusts the beam properties prior to the final transport to the IP. Six wire scanners are used for emittance measurement; four additional scanners are used at the skew quads are implemented with "u" angle wires as well as x, y wires. The emittance measurement station is capable of a few per cent measurement error. The carbon wires are 4 μm diameter.

The BMS provides for fine tune at the IP and adjustment of phase from the collimators to the final focus. Measurements made in the TDS give information for adjustment of these quads. Beam can be checked at the IP image point that occures (a) at the begining of the CCS and (b) within the CCS at the beginning of the FT. The vertical image of the 19 nm spot has a magnification of 50$\times$ to give a spot of 950 nm. This will require laser wire or gradient undulator (see below) instruments. The sextupoles in the CCS have to be turned off for this measurement at image point (b) in order not to dilute the image point spot size by large chromatic aberrations. The CCS " -I" optics sections must be checked by a BPM-corrector null method to assure the proper optics.

Final tuning of the IP spot size requires movement of the sextupole pairs in a variety of combinations. The crucial vertical spot size measurement must be carried out with a

laser interferometer (Shintake monitor [1]). The required resolution is about a factor of three better than what has already been demonstarted at the FFTB [2]. The horizontal spot is better measured with a laser wire [3]. These monitor laser beams pass through the electron beam 30-80 cm from the IP point and the low-β^* point must be shifted in z to this location. This is accomplished by the sextupole adjusters in the CCS.

The Shintake monitor has limited dynamic range of spot size that it can accurately measure for any specific laser wavelength. At a laser wave length of 262 nm spot sizes from 10 to 45 nm can be measured with 10% accuracy. On the other hand the laser wire spot can not be made smaller than ≈ 560 nm diameter using the same laser frequency. This is probably fine for the 845 nm sigma horizontal beam spot size.

The profile scans using laser beams must detect the scattered Comptons which come out straight ahead with opening angle $1/\gamma$. For systems within the IP detector, the Comptons must be detected dowmstream of the IP after bends in the CCS section.

The Shintake monitor has the potential for bunch length measurement as well. But as both it and the horizontal laser wire are deeply embedded in the detector, other less invasive instruments or instruments at other locations may be more practical for measurements that do not have to be done at the IP.

Other monitors are planned for the IP region (see also Chapter 2):

- Beam-Beam deflection monitor
- Fast lumonosity monitor
- Precision energy measurement
- Polarization measurement

These monitors are noninvasive and can be used during standard operation.

The steering from the beam-beam force is used to monitor the offset of the two colliding beams. A kick angle of 30 to 60 microradians per sigma y offset is expected. A BPM 2 to 3 m downstream of the IP would see typically 100 μm offset per beam sigma separation. Resolution of 10 μm is planned. Individual bunch measurements are possible. The signal can be used for beam steering feedback (see Section 3.8.3).

A fast luminosity monitor can be used to detect beam detuning at the IP to do non-invasive fine adjustments and horizontal beam scans. Radiative Bhabha pairs detected about 9 m from the collision point could provide the appropriate monitor information

Precision absolute energy measurements will be made of the beams in the extraction lines after the IP. The measurement used displacement synchrotron light fans from bend magnets to measure the bend angle of the beam in a spectrometer magnet. resolution of 10^{-4} is expected. However beams must not be colliding if precise before collision beam energy is to be measured.

The polarimeter uses laser beam scattering and detects Compton scattering asymetry. It is likely to be located before or after the IP.

Beam Steering – Initial Setup and Drift

Static tolarances of BDS magnet alignment indicate that typically y alignment tolarances of 0.5 to 10 μm are needed for many of the magnets if less than 2% luminosity

decrease (with collision beam steering) is to result from an individual magnet. A model for uncorrelated rms motions of the magnets in x and y predicts that rms motions of 204 nm in y and 558 nm in x will yield 2% luminosity reduction (again with collision beam steering) (Section 3.7.3).

As the alignment tolerances are tighter than is likely to be achieved with conventional survey, beam-based alignment using individual magnet power supplies will be utilized during initial beam set up. Magnet positions are adjusted with submicron positioners using BPM's with 3 μm precision.

A ground motion model based on a noisy environment predicts that 2% luminosity reduction from beam centroid motion will occure in 1 ms. Fast orbit feed back based on measurements of the first few bunches in a train will be necessary to keep the beams in collision. The beam centroid steering feedback is described in Section 3.8.3.

The same model predicts 2% reduction from vertical beam size growth in 200 sec. This then sets a time scale for the beam size stabilization feedback. The simulation program Merlin has been used to check possible noninvasive tuning proceedures to correct the deterioration of the beam size by linear optics aberrations (Section 3.7.6). It was found that simple careful one-to-one steering of the BDS elements using the BPM's and magnet movers every half hour or so was sufficient to restore the luminosity to almost its full value. A drop in peak (after retuning) lumimosity of 5% occured after 30-40 days indicating that an invasive retune would be necessary on this time scale.

	where	typical beam size [μm]	resolution required [μm]	freq. of adjustment	bunch structure	invasive y,n
POSITION x, y						
bpm	linac	300-60 x, 40-8 y	15 y	min.-1hr	pulse avg.	n
	BDS	360-13 x, 75-0.9 y	3 y	min.	pulse avg.	n
(disp. tuning)				30 days	pulse avg.	y
slow feedback	IP	0.845 x, 0.019 y	10 y	4 sec.	pulse avg.	n
fast feedback	IP		10 y	μs	bunch-by-b.	n
BEAM SIZE						
carbon wire	TDS	37 x, 6 y	1 x, y, u	30 days	one bunch	y
laser wire	IP image	13 x, 0.9 y	0.6 y	demand	bunch by b.	y
	IP$\pm$0.8m	0.845 x	0.6 x	30 days	scan over	y
Shintake mon.	IP$\pm$0.8m	0.019 y	0.01	30 days	one pulse	y
BEAM ENERGY						
bpm	CS or CCS	180 x	3 x	μs	bunch-by-b.	n
synch. light	extr. line			exp. demand	bunch-by-b.	n
ENERGY SPREAD						
wire	CS or CCS	13 x	5 x		one bunch	y
BUNCH LENGTH						
streak camera,	before	700 z	100 z	demand	one bunch	n
interferometer	linac			demand	or avg.	n
BUNCH CHARGE						
toroid/wall current	linac, BDS		1%	observe	bunch-by-b.	n

Table 3.8.1: *Overview of beam diagnostic requirements and tools. Additional instrumentation for luminosity and beam polarization measurement (not shown in the table) is foreseen in the IR.*

3.8.1.1 Beam Position Monitors

Overview

The beam diagnostics in *TESLA* will mainly be used for beam position monitoring. In this section we discuss the beam position monitor (*BPM*) system in the main linacs, which has to measure the average position of several bunches in a single pulse. In addition, a small number of high-resolution BPM's will be installed at selected points, e.g. in the Beam Delivery Section (*BDS*). Some of these special monitors must provide bunch-by-bunch measurements for the fast feedback system. The required *resolution*

machine	location	# of monitors	resolution
TESLA	linac	about 400	15 μm
	Beam Delivery Section	about 100	$1-3$ μm
	Interaction Point	1	≤ 10 μm

Table 3.8.2: *Required parameters for beam position monitors*

must be reached for a single bunch passage by a certain number of monitors. For all monitors the *dynamic range* must be one order of magnitude and the *sensitivity* has to be of the same order even for reduced bunch charges. It is a special problem for the electronics that the *long-time drift* of all systems has to be smaller than the resolution to avoid frequent beam based alignment.

The required signal-to-noise ratio of button or stripline monitors is roughly proportional to the ratio $R_o/3.4 \cdot \delta_{x,y}$ where R_o is the beam pipe radius and $\delta_{x,y}$ the resolution in x and y, respectively [8]. Resolutions of about 1 μm (rms) have been achieved using both buttons and striplines. In the former case this was at storage rings and light sources, where the bandwidth of the electronics can be reduced to some kHz. At the Final Focus Test System at SLAC, where striplines are used to detect beam offsets of less than 1 μm (rms), the beam pipe diameter is smaller by a factor of 2 (BDS) to 20 (main linacs) compared to *TESLA*.

Another important fact is that the accuracy of these monitors strongly depends on the precision of the installation of single electrodes. This is an essential problem for *TESLA*, where all monitors have to be attached to the quadrupoles inside a cryostat. Therefore the conclusion is that electrodes are not the best choice. In the following we will concentrate mainly on the discussion of resonant rf-monitors.

Circular Cavities

The simplest microwave BPM-structure is a circular cavity excited in the TM_{110}-mode by an off-axis beam. The amplitude of the TM_{110}-mode cavity yields a signal proportional to the beam displacement and the bunch charge, its phase relative to an external reference gives the sign of the displacement. In principle, both TM_{110}-polarizations have to be measured to obtain the displacements in x and y, respectively.

Cylindrical cavities are foreseen for the correction of all quadrupoles (inside the

cryostats) in the main linacs and in the BDS. The same type of monitor will be used to provide informations to the fast feedback systems discussed in section 3.8.2.

Signals The resolution near the electrical center of the cavity is limited by the thermal noise of the electronics and the excitation of common modes. Since the field maximum of these common modes is on the cavity axis, they will be excited much more strongly than the TM_{110} by a beam being only slightly off-axis. This gives the required frequency sensitive TM_{010}-rejection, S_{010}, to detect an offset δ_x. But the minimum detectable TM_{110}-signal is still limited by residual signals at the dipole mode frequency ω_{110}. The resulting resolution near the electrical center is roughly given by the ratio of the spectral densities at ω_{110}. Both limitations where estimated in [6]:

$$S_{010} = \frac{V_{110}(\omega_{110})}{V_{010}(\omega_{010})} \approx \frac{5.4}{\lambda_{110}} \cdot \frac{k_{110}}{k_{010}} \cdot \delta_x = K \cdot \delta_x \tag{3.8.1}$$

$$\delta_x{}^{\min} \approx \frac{v_{110}(\omega_{110})}{v_{010}(\omega_{110})} = \frac{1}{K} \cdot \frac{1}{Q_L} \cdot \frac{\omega_{110}^2}{\omega_{110}^2 - \omega_{010}^2} \tag{3.8.2}$$

This minimum resolution $\delta_x{}^{\min}$ can be enhanced by taking the difference of two opposite cavity outputs. The rejection of common field components is mainly limited by the finite isolation between the Σ- and the Δ-port of a hybrid or a Magic-T.

Main Linac Monitors On account of the limited longitudinal space, for the correction of all quadrupoles inside the cryostats, single circular cavities were designed. To avoid interference from the accelerating structures, the cavity was designed for a resonant frequency of $f_{110} = 1.517$ GHz (see also Fig. 3.8.1). A loaded Q of about 900 is required to measure the position of individual bunches spaced by 708 ns. CrNi was chosen as the cavity material to realize this, even at 4 K and with a coupling factor of $\beta \approx 1$ (no distortion of the electrical field). Detailed design data are given in [5].

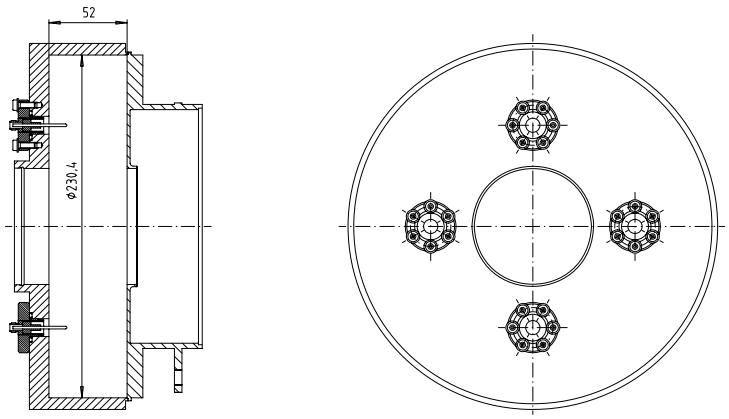


Figure 3.8.1: *Cylindrical cavity for quadrupole correction*

Assuming a cavity without beam pipes and a bunch charge of 37 pC, a TM_{110}-

amplitude of 2.34 mV (in the cavity) can be expected for an offset of $\delta_x = 10$ μm. The frequency sensitive TM_{010}-rejection, required to detect this displacement, is about 69 dB [7]. Because of the limited space, a standard stripline hybrid is used outside the cryostat to obtain a resolution of less than 6 μm (hybrid isolation of 20 dB).

Monitors in the BDS The same type of monitors will be used for measuring the position in the *BDS* and at the *IP*. Some parameters for both monitors are summarized in Table 3.8.3.

parameter	**BDS-monitor**	**IP-monitor**
type of monitor	cylindrical cavity	cylindrical cavity
beam pipe radius	5 mm	24 mm
bunch separation	708 ns	708 ns
required resolution	1 μm	10 μm
required bandwidth	5 MHz	5 MHz
position stability per day	1 μm	1 μm
cavity radius	$R_{res} = 90$ mm	$R_{res} = 15$ mm
cavity length	$l = 50$ mm	$l = 9$ mm

Table 3.8.3: *Requirements and preliminary design parameters for cavity-type monitors*

At a specific location the beam offset only needs to be measured in one polarization direction (x or y). Thus the monitors can be split physically into a x- and a y-monitor. Therefore both TM_{110}-polarizations can be individually tuned, which also increases the isolation between them and allows the measurement of very weak signals. The phase reference will be measured in an additional TM_{010}-cavity, operating close to the TM_{110}-mode (BPM-cavity) frequency.

Because of the required resolution of 1 μm in the *BDS* (single bunches) and a beam pipe diameter of 10 mm, a resonant frequency of 12 GHz was chosen. This monitor is under development for the TTFL-FEL to measure the position of the electron beam between two undulator modules. For the monitors at the *IP* a resonant frequency of 1.96 GHz was chosen. The reason for this particular frequency is the beam pipe diameter and to use standard electronic components (cellular telephone systems).

An important question for further design steps is how to realize the required field sensitive common mode rejection. Therefore different hybrid structures are under investigation. To get the desired loaded Q-values without any field distortions the monitors will be made of alumina (*BDS*) and CrNi (*IP*), respectively. The temperature of each individual cavity has to be stabilized.

The required frequency sensitive TM_{010}-rejection to detect a displacement of 1 μm is about 89 dB. This will be realized by different coupling factors for both modes and by bandpass filters at the input of the electronics. By using a field selective filter, e.g. by taking the difference of two opposite outputs, the theoretical resolution for this design will be better than 0.5 μm. After 20 m of coaxial cable, the signal at the analog

input of the electronics will be some mV per μm offset for single bunches with a charge of 1 nC.

Signal Processing The TM_{110}-amplitude is detected in a homodyne receiver by mixing the cavity output and a reference signal down to DC. When the beam is on the right, the system can be set up to give positive video polarity. The signal changes the phase by 180^o when the beam moves to the left, and for a centered beam it becomes zero. The signal processing electronics, as shown in Fig.3.8.2, are divided into two parts:

- analog electronics to filter, to amplify and to mix the signal down to DC; the reference signal (LO) will be provided by a TM_{010}-cavity
- Analog-to-Digital Conversion (ADC)

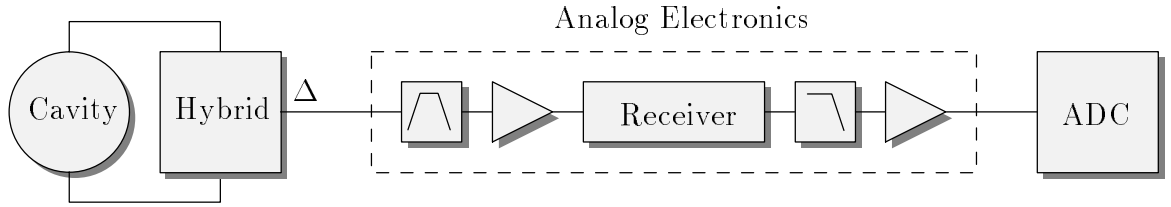


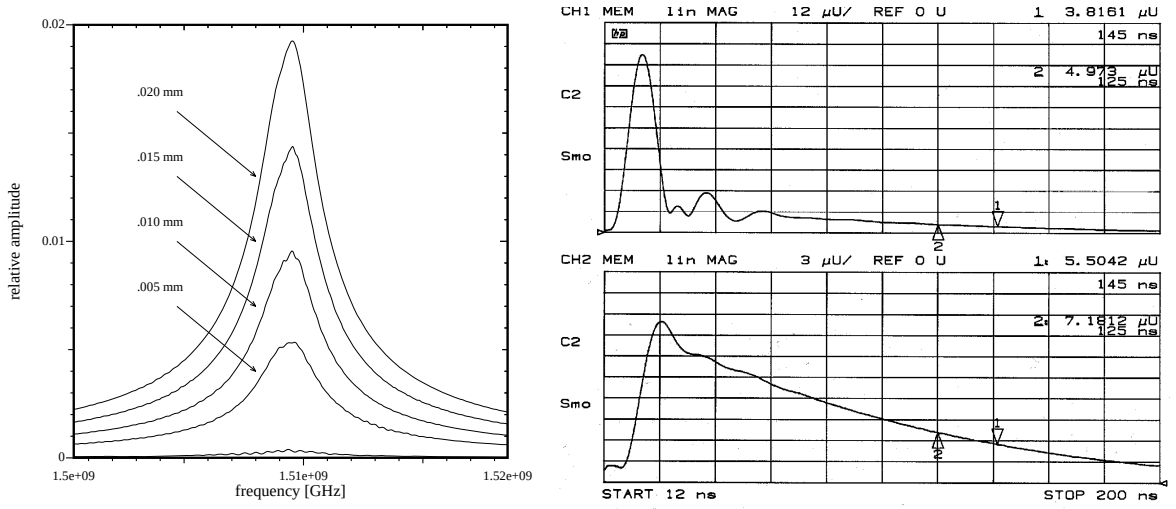
Figure 3.8.2: *Block diagram for the signal processing electronics*

The stop-band attenuation of the bandpass filter, together with the hybrid and different coupling factors for the TM_{010}-mode and the TM_{110}-mode, gives a frequency sensitive common mode rejection of about 100 dB. To reduce errors caused by temperature related phase shifts (long transmission cables), the mixing process will be performed in an I/Q-mixer.

After passing a low-pass filter and a bipolar video amplifier, the signal may be either viewed directly on an oscilloscope for adjustment, or digitized by a 12-bit ADC-board for quadrupole correction. The normalization will be performed by a computer.

Other Issues Bench tests were carried out on a stainless steel prototype to determine the resolution near the center and to test the electronics. Therefore the cavity was excited by an antenna fed by a network analyzer. A resolution of about 5 μm was measured in the frequency domain and in the time domain [7]. All measurements were performed at room temperature.

In addition, a prototype was tested at the CLIC Test Facility at CERN to demonstrate the principle single bunch response of the monitor and to measure the amplitude of the TM_{110}-mode as a function of the relative beam displacement [7].



a) b)

Figure 3.8.3: *a) Measured resolution in the frequency domain, b) Impulse response with (lower trace) and without a hybrid (upper trace)*

3.8.1.2 Bunch Length Measurement

The rms bunch length for the TESLA design is 700 μm (2.3 ps). For the FEL operation, the rms bunch length is 0.025 mm (80 fs) (see Chapter 5).

For bunch length measurements in this regime, different kinds of radiation emitted by the particle beam can be used. This can be transition radiation, synchrotron radiation or Čerenkov radiation. Transition radiation has the advantage that it can easily be produced by moving a thin metal target into the beam. However, at high current the target can be destroyed. This can be avoided either by doing measurements at low current only or by using diffraction radiation [10]. Alternatively, synchrotron radiation produced non-destructively by magnetic elements can be used [11].

Short radiation pulses can be measured in the time domain using streak cameras. A streak camera consists of a photocathode which converts light pulses into electron pulses. These are then deflected and displayed longitudinally on a phosphor screen (Fig. 3.8.4). These systems offer the unique possibility of single bunch observation in real time. The time resolution, widely assumed to be limited to 1 ps and longer, has been improved significantly by the manufacturers. Fast streak cameras are now commercially available with a resolution of down to 180 fs. This might be even improved in the future.

Alternatively, measurements in the frequency domain can yield information about the bunch length. This method works better the shorter the bunch length is. At wavelengths of the order of the bunch length and longer, coherent transition or synchrotron radiation is produced as the bunch acts as one single particle. The coherent part of the spectrum is much more intense than the incoherent part. It can be shown that the coherent radiation spectrum is just the Fourier transform of the longitudinal charge distribution. Hence, measuring this part of the spectrum means measuring the bunch

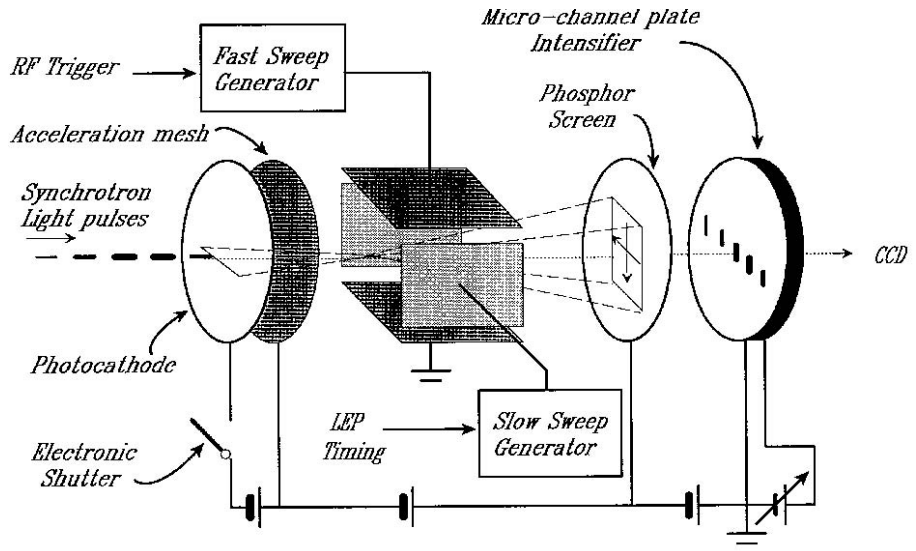


Figure 3.8.4: *Streak camera for measurements in the time domain as used at LEP.*

length.

From picosecond bunches, coherent radiation is produced in the millimeter wavelength range. For the observation of mm wavelength radiation, a photo-acoustic detector can be used. It works at room temperature and has a flat response from 0.1 mm wavelength up to 1 cm. For shorter bunch lengths, the onset of the coherence effect shifts to the far infrared where pyroelectric and bolometric detectors are available at low cost.

Two different types of spectrometers have been developed for use at the TTFL and as prototypes for the TESLA 500 collider. A series of high pass filters for mm waves serves, together with the broadband detector, as a simple spectrometer. This spectrometer has been successfully used for bunch length measurements at the CLIC test facility [12].

A polarizing Michelson interferometer (Martin-Puplett interferometer) can measure the autocorrelation function of the radiation pulses. The power spectrum can then be obtained as the Fourier transform of the autocorrelation function. The interferometer uses wire grids as beam splitter and roof mirror reflectors (Fig. 3.8.5). It has an intrinsic flat characteristic over the whole wavelength range of interest.

As a third option, the use of a new spectrometer based on superconductivity is being examined. A setup developed at the Forschungszentrum Jülich is being evaluated with respect to its possible application for bunch length measurements [13].

3.8.2 Beam Size Monitoring

An important class of diagnostics are the transverse beam size monitors which are required primarilly for the measurement of the beam emittance. Two modes of operation are expected:

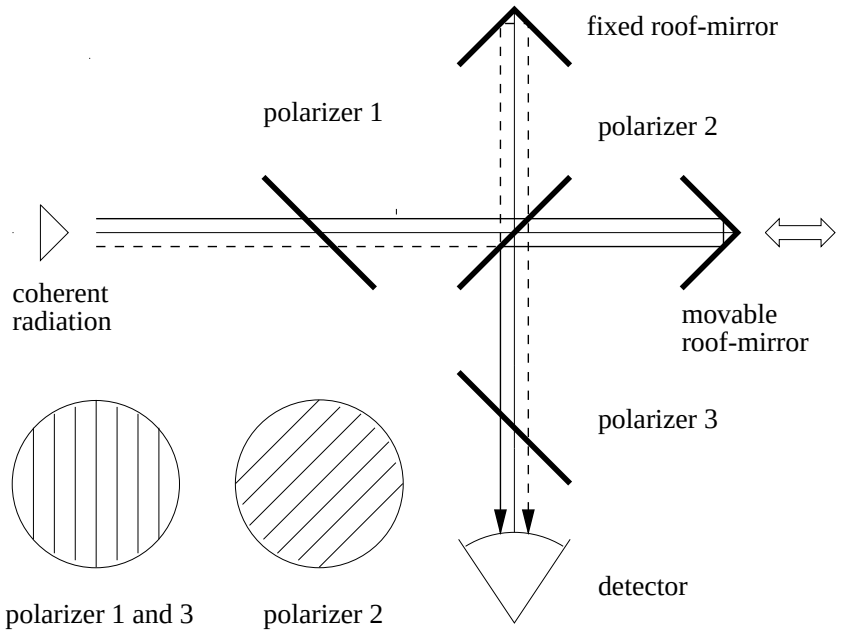


Figure 3.8.5: *Martin-Puplett interferometer for measurements in the frequency domain as to be used at TTFL.*

1. problem diagnosis, commissioning and boot-strapping, i.e. non-luminosity operation possibly with single bunch per pulse, and

2. continuous or semi-continuous monitoring of the beam during nominal luminosity operation.

For the former case, either Optical Transition Radiation (OTR) screens or SLC-type wire scanners[4] using 4 μm or 7 μm carbon filaments are expected to be the standard diagnostic.

Carbon filament wire scanners are used to sample the beam at different positions within the beam profile, and the beam width is then determined by fitting a Gaussian distribution to the data. The technique requires several bunches to perform a single measurement, so the result is an effective beam size of all bunches measured including any position jitter between them. However, the jitter can be measured by high-precision ($\approx 1\mu$m) BPM's so that this contribution to the effective beam size can be unfolded. Two types of signal can be used: i) bremsstrahlung from the wire can be detected some distance directly downstream, or ii) secondarys which are eventually scattered out of the beam pipe can be detected. The former requires a downstream dipole in order to position the bremstrahlung deletector on the downstream (photon) axis: such an approach is most applicable in the beam delivery system where bends occur naturally. In the linac and other straight sections, one must rely on the secondaries scattered out of the beampipe for detection. The wire scanners themselves are small compact devices

which can easily be incorporated into the machine with the mimimum of impact to the lattice.

At the full TESLA bunch population of 3.6×10^{10}, the charge density of a single bunch should be sufficiently low that a 4 μm Carbon wire should survive single bunch operation at the nominal repetition rate of 5Hz. The temperature rise in the wire caused by a single bunch (with charge N_e) passage can be estimated from the energy loss $\mathrm{dE/dx} = 4\,\mathrm{MeV/cm}$, the heat capacity $c_p = 1.18\,\mathrm{J/(g\,K)}$ and the specific density $\rho = 2\mathrm{g/cm}^3$ of Carbon:

$$\Delta T \approx 0.44 \cdot 10^{-3}\mathrm{K} \times \frac{N_e[10^{10}]}{\sigma_x \sigma_y [\mathrm{cm}^2]} \tag{3.8.3}$$

In the TDS we have $\sigma_x\sigma_y \approx 200\,\mu\mathrm{m}^2$ and for N_e=3.6·10^{10} we get ΔT=790 K, which seems tolerable. However, the wire will only survive a few bunches of the nominal bunch train in multi-bunch mode, and so Carbon wire devices cannot be used for monitoring the beam size during luminosity production. Even in single-bunch operation, it is still possible that the wires will fail, and so it is prudent to have each scanner equipped with 5 to 10 separate wires.

The resolution of a beam size measurement with a wire is expected to be better than the wire diameter. The sytematic error for a 4μm wire and 6μm beam size is about 5%.

The OTR device In the measurement of transverse beam charge distribution, OTR present a number of advantages: it is fully linear with the beam charge, without any saturation effect, it is emitted by a surface crossing and thus does not show any depth effect. Its temporal emission in the subpicosecond range allows the selection of a single bunch. The main practical limits to OTR use at high energies, when the low emittance allows small beam dimensions, are the diffraction limit that, due to the angular distribution and the polarization properties, is larger than for standard point optical sources and for 500 nm wavelength is roughly 10 μm, and the fact that it is derived by an intercepting device, so that the screen can be destroied by the high energy density lost in the material by well collimated beams. For an aluminum foil, which is the preferred material for its high and uniform reflectivity over the whole optical spectrum, the temperature rise caused by a microbunch of 3.6·10^{10} electrons is .23 °C/R^2 in which R is the radius (in mm) of a flat circular beam distribution, and assuming a rather conservative value of $2\,\mathrm{MeV/cm}^2$ as the single particle energy deposition (energy independent). Limiting the temperature rise to 300 °C, we obtain a minimum beam radius of 20 μm (that corresponds to an rms value of about 10 μm). It will be important to study different materials to find the best compromise between emission efficency and thermal and mechanical properties.

For multi-bunch operation, there are three possibilities:

1. laser wire scanner
2. pulse snatching scheme (off-axis monitoring).
3. gradient undulator

In principle a laser wire acts in the same manner as a Carbon-wire scanner. A laser wire can be focused to a sub-micron spot, which is then scanned through the beam. The situation is also similar to the Carbon-wire approach with respect to signal detection: either the Compton scattered photons are detected directly downstream (again requiring a downstream dipole), or the back-scattered electrons (which are eventually sweeped out of the machine) are used. For TESLA, it is feasible that a long pulse-train laser can be used to sweep through the entire bunch train in a single shot, thus allowing a fast effective beam size measurement of a single bunch train. Such an approach is necessary for the laser monitors close to the interaction point, where it is neccessary to mitigate the effects of the final doublet vibratration on the (particle) beam. In a short-pulse mode, the laser wire can be used to sample single bunches in the bunch train, sampling on successive trains the same bunch location. This mode of operation is directly analagous to the single-pulse Carbon-wire operation. As of writing, laser-wire technology is still in its infancy, although such a device now exists and operates at the SLC [3]. However, much work still remains to be done (especially for the back-scattered electron mode). It is advantageous (and prudent) to always place a carbon-wire device and a laser-wire device together, so that one can be used to calibrate and troubleshoot the other.

A more conservative approach (but slightly more invasive) is to use fast kickers to "steel" a single bunch from a bunch train and send it to an off-axis scanner. Here again Carbon-wire devices can be used (as well as any other fluorescent or OTR screen technique). The approach probably requires more tunnel space than the laser-wire option.

Of the three methods proposed, only the X-ray gradient undulator[9] offers a single-shot single bunch measurement of the beamsize. A gradient undulator is an undulator whose pole-pieces are displaced one-half period as shown in figure 3.8.6. The undulator field has now a linear gradient in the vertical direction. As a result, vertically off-axis particles radiate, while particles on-axis do not. The total power output from the undulator is directly proportional to the square of the beam height. Theoretical studies have shown that at the exit of the linac (250 GeV) in excess of 10^5 60 MeV photons can be expected for a 1 m device with a 1 cm undulator period, and a field gradient of 4.8 T/cm. The main disadvantage of the device is that it requires a dipole downstream in order to place the X-ray detector on the axis of the undulator (again making the device more suitable for the beam delivery section). If the device is to be used in the linac, then small chicanes will be required, and a length penalty will be incurred. It should be noted, however, that the gradient undulator is a conceptual device, and work is currently in progress to evaluate its performance and limitations.

3.8.3 Orbit Feedback

Overview

In this section we discuss the orbit feedback schemes for *TESLA*. In order to reduce the influence of vibrations which can cause luminosity reduction, several feedback loops

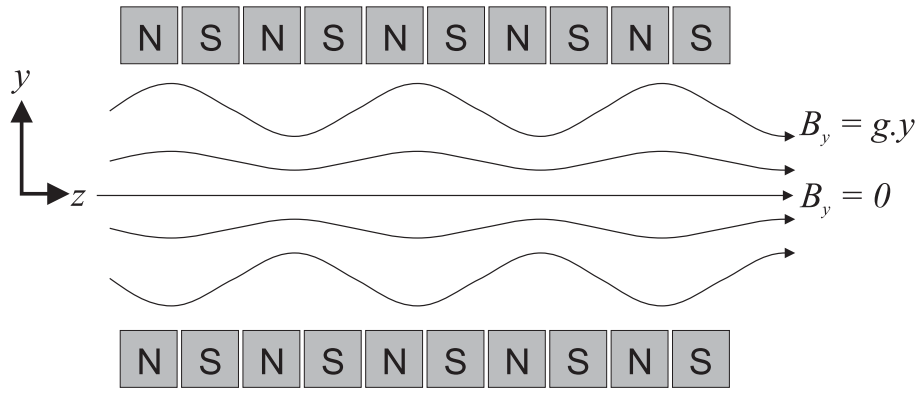


Figure 3.8.6: *Principle of the X-ray Gradient Undulator. The vertical undulator field is proportional to the vertical offset (y), being zero on the nominal axis (y = 0). Thus particles on-axis do no radiated, while particles off-axis do. The total power radiated by a centered beam is therefore proportional to* σ_y^2 *[9].*

are foreseen to regulate the beam orbit and to keep the beams colliding.

The main sources for the orbit movement can be divided into two types: perturbations due to ground motion in the lower frequency range and disturbances in the kHz range caused by the superconducting cavities, mainly microphonics and Lorentz-force detuning. In principal the control system for orbit correction can be divided into three different feedback systems:

- A slow feedback system
  in the main linac which is working at the repetition frequency of 5 Hz. It is foreseen to cancel perturbations in the low frequency range which are mainly due to ground motion. The beam will be steered by the use of corrector coils. The slow feedback is limited by the sampling frequency and the time neccessary for the corrector coils to change to the requested settings. Therefore this feedback is able to detect and to react on disturbances only up to .5 Hz.

- A fast feedback system
  has to be installed to compensate disturbances containing higher frequency components. It will correct the beam orbit from bunch to bunch which is possible because of the bunch spacing of 708 ns. The fast feedback system will be located in the tuning and diagnostic section of the the beam delivery system (*BDS*). Fast kickers are used to correct the beam offset.

- A fast IP-feedback
  which will keep the two beams in collision using the beam-beam-deflection.

The bunch spacing of 708 ns is one of the main advantages of *TESLA*: it is possible to use the pilot bunch to correct all following bunches of the same pulse.

All feedback systems can be digital, which allows the realization of a data base driven feedback. The data received are also recorded and can be used for further

statistics. The exact number of feedback loops neccessary to control the beam orbit in a satisfactory manner is not yet known; therefore one should always expect to add additional feedback loops.

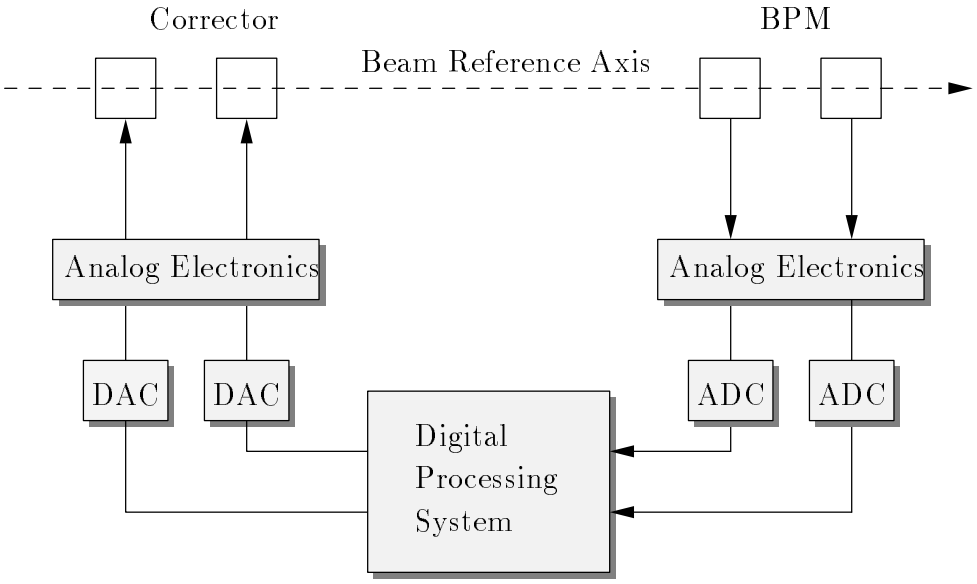


Figure 3.8.7: *Schematic diagram of the slow and the fast digital feedback system*

All digital feedback systems will consist of the same basic elements as shown for the slow and the fast system in Fig. 3.8.7:

- **beam position monitors** (BPM) to obtain a signal proportional to the beam position
- analog **signal processing electronics** and **analog-to-digital conversion** (ADC)
- a **digital processing system** to determine the correction signals
- **digital-to-analog conversion** (DAC) and **signal processing electronics** to applify the correction signals
- and **correctors** to apply the requested kicks.

Beam Position Measurement

The first device in the feedback loop is an electromagnetic pick-up to obtain a signal proportional to the beam position. Cylindrical cavities are foreseen for the detection of these signals, and the details are described in section 3.8.1.1. The principal scheme for the signal processing will be the same for both types of feedback systems (see also Fig.3.8.2). But for the fast feedback system the analog electronics and the analog-to-digital conversion have to be optimized with respect to the processing time.

Slow Feedback System

Disturbances in the low frequency range are mainly due to ground motion. They will cause quadrupole displacement which will lead to beam emittance growth. In order to eliminate the effects of displaced elements a slow feedback system working on the repetition frequency of 5 Hz has been developed.
For the orbit correction four states, the position and the angle of the beam trajectory in both transverse plans, will be controlled. For each transverse plane a feedback loop needs at least two beam position monitors which measure the actual offset of the beam and two simple steering magnets are used to regulate the beam at a certain position.

The conceptual design of this pulse-to-pulse orbit control is based on tools of Optimal Control Theory [16]. It takes the stochastical nature of the sensor noise and of the beam movement into account. The feedback loop can be divided into four main parts: measurement taken by BPM's – estimation of the beam's position and angle – calculation of control input – correction by corrector coils. It contains an optimal filter (Kalman Filter) to estimate the beam's position and angle and an optimal controller to calculate the corrector settings [17, 18].

At the maximum operating frequency of about .4 Hz, the assumed r.m.s. quadrupole motion from pulse to pulse is in the order of 200 nm. At the end of the linac e. g. a kick of 0.36 μrad is required; for a 250 GeV beam the correctors must provide a magnetic field of 3 Γm. This can be done by the standard correction coils installed in the linac.

Results of simulation runs show that the feedback algorithm provides a good orbit control in the requested range. The feedback loop damps frequency disturbances up to 0.4 Hz quite well, where a DC bias rejection of -50.77 dB was achieved. The r. m. s. resolution of the beam position monitor was assumed to be 1μm.

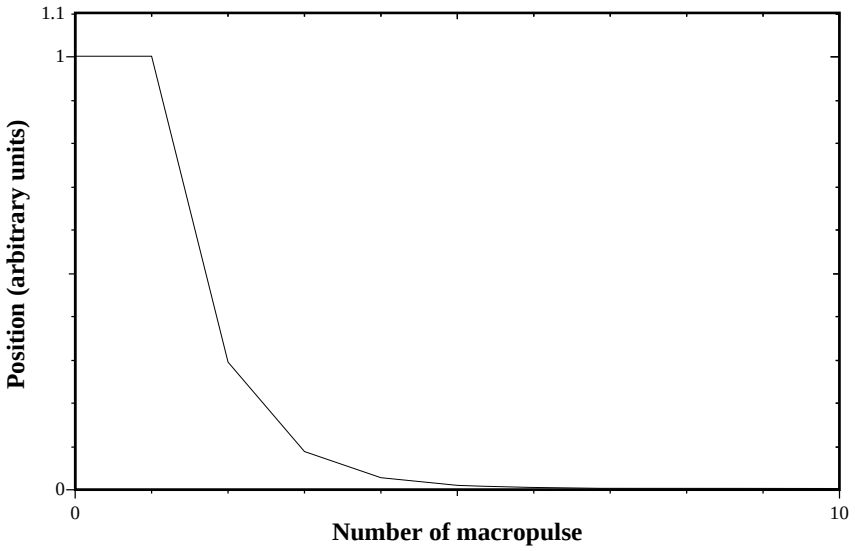


Figure 3.8.8: *Simulation plot of the slow feedback for TESLA: Feedback loop response on a step disturbance on the incoming beam*

A better DC bias rejection is achieved by including an integral part into the feedback loop. The disadvantage is a higher amplification in a range between 0.5 Hz and 1.9 Hz. A comparison of the control loop design with and without integral part in order to get the best feedback design will need work in future.

Fast Feedback System

Disturbances like fast vibrations of quadrupoles, microphonics and Lorentz-force detuning caused by the superconducting cavities cannot be detected and controlled by the slow feedback system. In order to correct the position and the angle of every bunch in a train, first investigations of a digital bunch-by-bunch feedback have been carried out. This fast feedback system is foreseen to be located in the tuning and diagnostic section of the *BDS*. It will consist of two BPM's and two fast kickers to steer the beam.

Taking into account the separation of optimum locations for pick-ups and kickers, long transmission line lengths and ADC/DAC-conversion times, the remaining time-interval of 200 ns should be sufficient for 10 instruction cycles (floating point operations) in a standard microcontroller. The rapidly increasing development of high speed analog and digital signal processing components leads to the assumption that this time interval for the digital processing will become much longer.

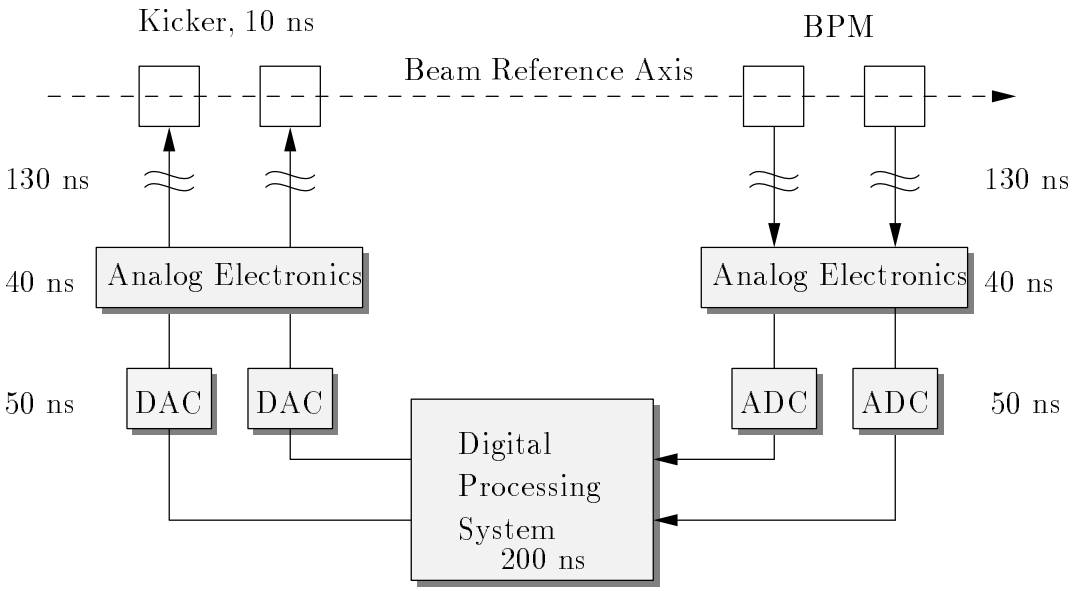


Figure 3.8.9: *Maximum time dedicated to each device for the bunch-to-bunch feedback*

IP Feedback

At the *IP*, the beam-beam scattering will be used to keep the beams in collision. This method has already been used in the SLC for single bunches to measure the offset of the two colliding beams from pulse to pulse precisely [15]. For *TESLA* the kick angle α per $\sigma_y^\star$ estimated from beam-beam simulations is approximately given by

$$\alpha \quad [\text{mrad}]: \quad \frac{\Delta y_{IP}}{\sigma_y^\star} \times 0.047 \text{ mrad} \tag{3.8.4}$$

with $\sigma_y^\star = 19$ nm beam spot size and Δy_{IP} beam-beam separation at the *IP*. The deviations from the linear relation 3.8.4 are not taken into account here, but they are not expected to have an essential influence on the feedback performance.

In order to reduce the processing time of the feedback loop, the locations of the BPM's and the kickers have to be as close as possible to the *IP*. If the BPM next to the *IP* is located at the position of the final doublet separated by 3 m from the *IP*, a $\sigma_y^\star$ separation will result in an offset of about 140 μm. Therefore the required resolution of the BPM can be relaxed and will be about 10 μm.

Within one pulse a bunch to bunch measurement of both outgoing beams can be used to get a precise measurement of the offset of the two colliding beams. The difference of the pick-ups taken before and after the IP of one bunch gives the sinusiodal part of the trajectory whereas the sum determines the requested kick to steer the beam.

Correctors

In the digital system, standard Digital-to-Analog Converters will be used to convert the signals back into analog values for the correctors. These signals have to be amplified in a fast power amplifier.

In the slow feedback system corrector coils are used to steer the beam from pulse to pulse. For the fast feedback the correctors will be magnets with a rise time of about 10 ns [20]. The magnetic field of 0.5 m long structure is about 2.7 Γm.

The orbit feedback at the *IP* shall be able to remove a separation up to $10\,\sigma_y^\star$. Assuming a distance of 4 m between kicker and *IP* this will result in a requested kick of 0.023 μrad for each kicker. Therefore a magnetic field of $2 \cdot 10^{-5}$ T is required for an energy of 250 GeV.

3.8.4 Control System

The computer hardware and software has been changing rapidly over the last several years and will continue to change during the course of contructing the linac. On the other hand, a software development for such a big machine has to start early to be ready in time. Most of the investment goes into computer software and into front-end electronics. For the front-end components, the investment in industrial standards is quite safe. Systems such as VME have a long lifetime, for example. Software developments should be based on new industrial methods. Object oriented tools have

recently been gaining a growing market share and have proved to be very powerful methods for control system design in the case of TTF. In addition to apply object orientation, an implementation on more than one computer platform leads to a better design and helps migrating to new operating systems. Special consideration of the huge amount of equipment, (due to the length of the machines) has to be taken. The software and the hardware layout of the control system must allow for this. On the software and communication side, an efficient access to device data is necessary. A lot of similar devices must be readable with little overhead and short response times. Higher levels in the control system should provide summary information on device groups. Good and automatic monitoring of all devices is essential. The hardware layout should be fault tolerant. This can be achieved by a modular design. The modularity of the beam based hardware for instance could follow the structure of the klystrons, and allows one klystron to be switched off in case of a failure. Operation of the linac is still possible if a few klystrons fail. Failing hardware and software components should not in general stop larger segments. Therefore, the control system has to follow the hardware modularity also.

3.8.4.1 Architecture

For most areas in the linac we expect low enough radiation level to be able to bring the front-end electronics, including the device servers, into to tunnel. In a few places some extra shielding against synchrotron radiation will have to be inserted. Low radiation also allows us to use modern industrial, fieldbus based sensor/actor electronics. The overall architecture consists of three layers of computers. A top level with display or client programs, a front-end layer with device servers and I/O and a middle layer with powerful group servers. The upper layer is the interface to the operators. These upper services should be available to the consoles in the control room, the experts working in the tunnel and the specialists in their offices. All of them should be able to use the same programs. Only the level of details presented should be different, but nonetheless available to all levels of users. Modifications of device data must be protected with access rights. Since a lot of users and client programs access a lot of data from the front-end and the middle layer, a fast network is necessary to decouple the display stations. A net switch that routes packets from the clients to their servers provide the important bandwidth. Client programs often need a go/no-go information from a group of devices only. In such a big system it is necessary to ask a large number of front-ends to get this information. A middle layer server should provide such collective reports. Likewise, the middle layer should supply frequently requested data of all its front-end computers in a single block transfer. Proxy services may also be added since the middle layer servers are faster than the front-end computers. But, a direct access to front-ends (and not just for debugging purposes) is still necessary. In general, the middle layer implements higher services and improves the performance of the system. Data bases, file servers, simulation servers and other kind of servers without direct device connections should be placed in the middle layer also.

Subsystems of the machine use their own subnet with front-end computers. The

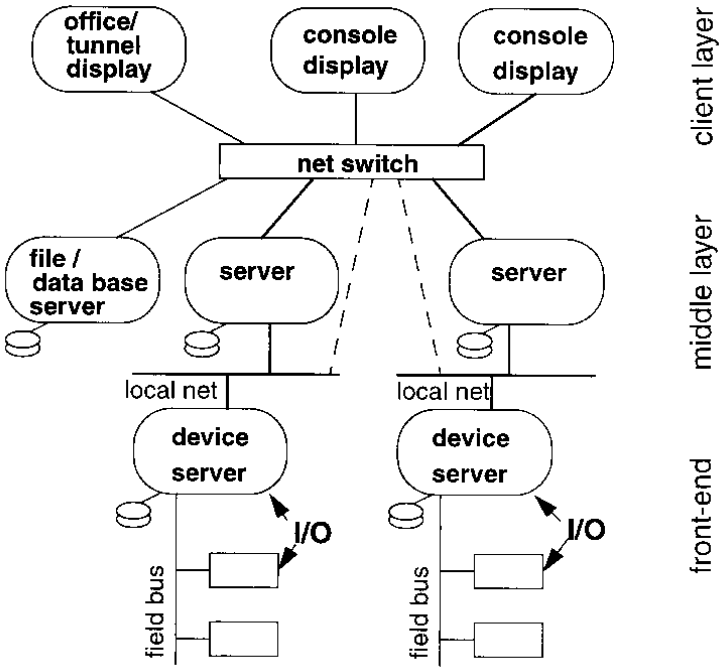


Figure 3.8.10: *Different layers of the machine control system.*

front-ends are distributed and close to the devices of the linac. Some hardware is directly connected to the device servers, other equipment use field busses to connect the front-end electronics to the device servers. Fieldbus electronics introduce a further level of computers in the system. Programmable Logic Controller (PLC) and 'intelligent' devices are examples of such stand-alone processors. All group servers and most front-end servers are equipped with disk drives. The other device servers mount their file systems from the corresponding group server. If a front-end has to run stand-alone in case of a network problem, it should have a direct connected disk or must have the program and data files in memory. Archiving of all data from the equipment can consumes a high portion of the network bandwidth if the storage is not in a local disk.

Front-end Computer

Most of the device input and output channels connect via VME modules or fieldbus electronics to the device servers. The VME system is designed as a reliable industrial electronic environment. Almost any computer, analog or digital input/output module and fieldbus interface is available in VME. The front-end servers should be grouped to separate different subsystems. This is a requirement for the installation, debugging and maintenance. A huge number of devices can reduce the overall operation time of the machine. Malfunctions of a single device should not stop machine operation. Grouping of devices or redundant layout must be used. The organisation depends very much on the individual device type. Low level RF systems for the klystrons can be grouped to serve one klystron only. A single failure would switch off the cavities of one

klystron only and the machine could continue to operate.

Server

A linac control system consists of several different kinds of servers. Servers is this sense are computers that provide a service to other processes. Besides the servers of the front-ends, servers are placed in the middle layer. They help client programs to get informations on the huge number of data acquisition equipment. Middle layer servers collect the front-end data and add proxy services. Since all device information of a subsystem is available in the middle layer, higher order functions can be implemented. Such device group servers build a new collective device that represents the group to client programs. Further services provided by the middle layer servers:

- User and program file systems
- Data bases for all machine parameters
- On-line machine optics calculations and simulations
- Sequencer to automate the linac operations
- Alarm and error logging

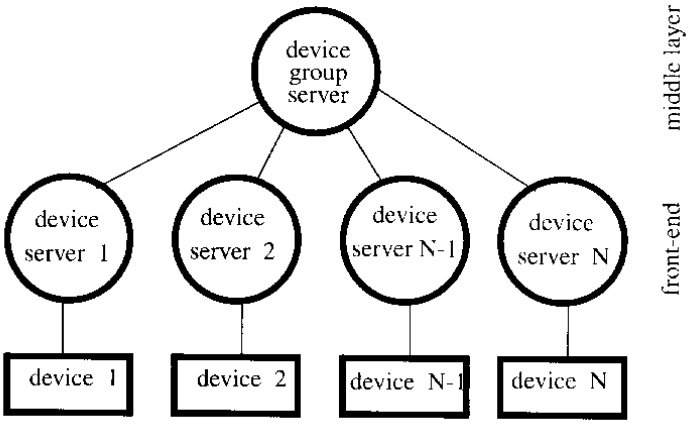


Figure 3.8.11: *Connection of front-end servers to group servers.*

Clients

A client is a program that uses services from the servers in the control system. An application program or a display process are typical examples of clients. These programs run on the operator consoles, in the tunnel during installation and maintenance and in the offices of the machine groups. Other applications run in the hardware and

software development labs. All these programs should use the same communication protocol and the same library to talk to the servers. This is important since it should be avoided to rewrite the same software for different environments. The protocol is described in the 'Communication Interface' chapter. Client software is a wide field and consists of some generic programs:

- Tools to display alarms and other device errors
- Tools to access all data in the system (some times called parameter page)
- Tools to display historical trends and other plots

Generic programs don't have to be modified in a changing environment. They allow access to all devices by standard methods. Other tools are standard programs that allow to create dedicated applications to:

- generate and display synoptic representations of devices
- store and recall machine parameters
- integrate device data into office products like text editors and spread sheets
- access device data for analytical programs like MATLAB
- interface to machine simulation programs
- provide control system I/O for industrial tools like LabVIEW
- define automated sequencing of machine operations

Some of the above applications just use the application program interface (API) of the control system with a few lines of interface code. Commercial tools will play an increasing role in the client programs.

3.8.4.2 Network

This section discusses the various networks in the control system. At the front-ends this are fieldbusses to connect the local input and output of the devices to the corresponding servers. Devices servers connect to the group servers via standard local area networks (LAN). The devices get their timing and synchronization signal on a dedicated network and the interlock system is connected to a further separate line. Standard LAN technology is also used in the links between the servers and all client processors. Clients are distributed in the control room, the tunnel and the offices. And the whole network is connected to the DESY and outside nets.

Fieldbus

A fieldbus is a network that connects up to 32 sensors, actors or complex front-end input/output devices to the local device servers. Because fieldbusses are standard industrial systems with a lot of suppliers they should be preferred compared to proprietary busses like SEDAC (SErial Data Acquisition and Control). With the new bus systems ProfiBus and CAN, we already have some experience in TTF and HERA. They work reliably and modules from a lot of manufactures are available already. There is no single common standard fieldbus on the horizon. It is still a fast growing market and in future further busses like LON might be added as standard parts. Today the market is immense with a lot of different systems. A control system on the other hand has to focus on very few standards to reduce the maintenance costs since modern fieldbusses are quite complex compared to SEDAC. Some instruments are equipped with the older standard interfaces GPIB and RS232. RF instruments usually come with the GPIB interface. This bus system is limited to about 12 meters only and restricted to measurement equipment and their peripherals. RS232 is a point to point link. Usually the software effort to communicate to the RS232 devices is underestimated. These two bus systems should be avoided in the front-end I/O.

Local Area Network

Networks between the consoles, group servers and front-end servers will be based on standard LAN technology. This is today the 10 Mbit Ethernet. Cheap interfaces for all computer types are available. Ethernet is certainly too slow for the envisaged size of the machine. But, in the near future this standard will be replaced by a factor of ten faster network (100 Mbit/sec.). FastEthernet, FDDI, ATM, Fibre Channel are such technologies and are already available. The servers in the middle layer should be connected by an other factor of ten faster network i.e. about 1 Gbit/sec. All these network components are or will be standard items from many suppliers and can easily be integrated in the control system since they are based on the TCP/IP communication. For the long distances fibre optical links will be used to prevent ground loops and to isolate the subsystems. Local connections use less expensive twisted pair cabling to the active hubs that link to the fibber optics. In order to keep the network load low and predictable, the network is segmented into several parallel lines to the front-ends. Maintenance and managing considerations also lead to a good segmentation of subsystems.

Communication

During the installation and maintenance of all components of the machines, mobile terminals will be used. These terminals connect to the standard network and allow access to all devices in the system. A wireless network should be used in order to be more flexible and to save installation costs compared to a wall plug system. Such a wireless system will be used for audio communications also. It can be based on the GSM standard which is supplied by a lot of manufactures and cheap units. A

few private receiver/transmitter inside the tunnels will connect the terminals to the controls network and DESY telephone system.

Interlock

The machine will be equipped with an interlock system at several points:

- differential current measurement of the beam
- beam loss monitors
- RF input coupler spark and field emission protection
- Klystron okay
- low level RF operational
- cryogenic okay
- vacuum okay and beam valves open
- power supplies okay
- experiments ready for beam

A detected failure in any part of the distributed system will switch off the guns and injectors. All sources connect to a common fibre optical line. The system can be considered as an independent hardware link from the interlock sources to the beam sources. Although it works without the control system, the interlock has to be readable from the control system to insure that the cause of an event can be detected. And because of the several thousand interlock points of origin, individual sources must be maskable i.e. the control system must be able to switch them off. A good bookkeeping is essential here. A further requirement is the ability to record the history of an inhibit event. Therefore the inhibit signal is feed back as an event in the timing system. This event is detected by all front-ends as a 'post-mortem' trigger to save their data before and after the interrupt time.

3.8.4.3 Software

Throughout the development time of the software for the machines the computer industry will introduce a lot of new products. A main challenge for the design will be to create the software modular with good interfaces to allow easy migrations to new environments. Modular means that some submodules must be changeable without interference with other part. It also means to design reusable objects. The 'heart' of the control system is the communication interface. All client and server programs communicate by it. When ever possible standard industrial components should be used for front-end hardware up to the display programs. And the software should be designed with the common tools and programming languages. Our experience in the

TTF control system showed that object oriented languages (like C++) are adequate design environments to create reusable modules. Object orientation leads to a clean frameworks and follows the industrial paradigm.

Operating Systems

From today's perspective the choice of operating systems suitable for a control system are UNIX and Windows NT and some real time systems for front-ends. In the future the operating system will probably develop to more object oriented systems. Because of maintenance reasons the number of different operating systems should be kept small. On the other hand a design which is based on more than one system leads to a better portability in the future. Therefore, the TESLA control system should be based on UNIX, Windows NT and VxWorks. The latter is a real-time system for front-ends only. The other two operating systems have 'soft' real-time capabilities and provide a multi-threaded, multi-tasking environment.

Communication Interface

A basic element in a control system is the communication interface. It defines the application program interface (API) for the data exchange between clients and servers. Because of it's important role in a control system, the API must be based on standard protocols. For the network layer this is the TCP/IP Internet protocol. Higher level protocols are transported over the basic TCP/IP services. Requirements for the top level are:

- has to transport a well defined set of data objects
- needs to convert the data between different computer architectures
- provides common methods e.g. translations of the data structures
- interface library for all CPU types in the control system
- access to all data (including archived data) of devices in a consistent way
- interface to standard desktop applications like spread sheets and to special controls applications
- has to provide an on-line names service to query existing devices and their properties
- defines a unique way for names of device addresses and their parameters
- allows access to single and multiple devices in one call
- sends parameters as data to the device in a read call to allow atomic requests
- client programs must get device address resolutions from a name server and not by recompilation

Database

Data bases are used to store the parameters of the machines, the configurations of the servers and the archived on-line measurements of devices. In addition, all data from the production process of the cavities need to be stored in a data base. For TTF this data is stored in ORACLE and it is planed to continue this work for the TESLA collider. ORACLE is a good choice since it is the standard data base for DESY, it has interfaces to several computer types and has a solid market share to trust in the future support. For on-line archiving and configuration data of front-end servers a big data base is too slow and voluminous to fit into a small front-end. A distributed data base access would not help because a server should be independent from network connections if possible. Beside this, the estimated network load could be very high for several thousand devices with several properties to archive at a 10 Hz update rate. Object oriented data bases or fast file based archives are better suited for this purpose. They may be located in the front-end and group servers.

Bibliography

[1] T. Shintake, *Proposal of a Nanometer Beam Size Monitor for e+e- Colliders*, NIM A311, p. 453, 1992.

[2] T. Shintake et al., *First Beam Test of Nanometer Spot Size Monitor Using Laser Interferometry*, KEK-Preprint 94-129, Oct. 1994.

[3] M. Ross, *High Performance Spot Size Monitors*, LINAC96 Conf., Geneva Aug. 1996.

[4] M.C. Ross, E. Bong, L. Hendrickson, D. McCormick, M. Zolotorev, *Experience with Wire Scanners at SLC*, SLAC-PUB-6014, December 1992. Invited talk given at 1992 Accelerator Instrumentation Workshop, Berkeley, CA, Oct 27-30, 1992. Published in Berkeley Accel.Wkshp.1992:264-270 (QCD183:A22:1992)

[5] *TESLA TEST FACILITY LINAC - Design Report*, edited by D.A. Edwards, DESY Hamburg, TESLA-Note 95-01, 1995.

[6] W. Schnell, *Common-mode rejection in resonant microwave position monitors for linear colliders*, CLIC note 70, CERN-LEP-RF/88-41, 1988.

[7] R. Lorenz, *Beam Position Monitors in the TESLA Test Facility Linac*, IEEE Conference Proceedings of the PAC 95, Dallas, May 1995.

[8] R. Shafer, *Beam Position Monitoring*, AIP Conference Proceedings 212, 1989, pp. 26-58.

[9] E. G. Bessonov, J. Pflüger, G.-A. Voss and N. J. Walker, *Beam Size Measurements in a Linear Collider using an X-ray Gradient Undulator*, DESY internal report DESY M 96-18, September 1996.

[10] M. Castellano, private communication.

[11] C. Bovet, M. Placidi, *A dedicated synchrotron radiation source for LEP beam diagnostics*, CERN LEP note 532, 1985.

[12] K. Hanke, V. Schlott, K. Aulenbacher, H.H. Braun, F. Chautard, *Beam diagnostics using coherent transition radiation at the CLIC test facility*, CERN CLIC note 298, 1996.

[13] Y.Y. Divin, H. Schulz, U. Poppe, N. Klein, K. Urban, V.V. Pawlowskii, *Millimeter-wave Hilbert-transform spectroscopy with high-T_c Josephson junctions*, Appl.Phys.Lett. 68, p.11, 1996.

[14] G.A. Voss, R. Brinkmann, N. Holtkamp, *A Beam Based Interaction Region Feedback for an S-Band Linear Collider*, presented at the LINAC96 Conference, Geneva, August 1996.

[15] P. Bambade, R. Erickson, *Beam-Beam Deflections as an Interaction Point Diagnostics for the SLC*, SLAC-PUB-3979, May 1986.

[16] I. Reyzl, *Simulation of Feedback for Orbit Correction*, to be published in IEEE Conference Proceedings of the EPAC 96, Sitges, June 1996.

[17] Gene F. Franklin, J. David Powell, Michael L. Workman, *Digital Control of Dynamic Systems*, 2nd ed., Addison-Wesley, 1990.

[18] Brian D. O. Anderson, John B. Moore, *Optimal Filtering*, Prentice-Hall, 1979.

[19] Alan V. Oppenheim, Ronald W. Schafer, *Discrete-Time Signal Processing*, Prentice-Hall, Inc., 1989.

[20] B. I. Grishanov, F. V. Podgorny, J. Ruemmler, V. D. Shiltsev, *Very Fast Kicker for Accelerator Applications*, TESLA-Report 96-11, November 1996.

3.9 Survey and Alignment

3.9.1 Network on the Surface of the Earth

First, the coordinates of reference points along the linear collider have to be determined with respect to the existing coordinate system at the respective site (in case of DESY, this would be the HERA coordinate system). The reference points are the base to mark out the planned halls. The demanded global accuracy (standard deviation) of the reference points should be better than 5 mm over the whole area of the linear collider of about 30 km. Today the coordinates will be measured by the satellite system GPS (GPS = Global Position System). With points on the top of the halls the coordinates have to be transferred into the tunnel to control the tunnel boring machine and later on to mark out the supports of the components of the beam transport system. With a precision levelling a vertical network also has to be established.

3.9.2 Requirements for the Aligment of the Components

In the groundplan the tunnel axis is a straight line approximately. In the vertical plane the axis of the tunnel follows the earth curvature except for the first 6.5 km. Then the height of the tunnel in this area is constant everywhere. The heights are referenced to the geoid. Therefore it is possible that the earth radius for the tunnel axis is not constant. Depending on the gravity of earth small changes in the radius are possible.

In the tunnel there are 5 separate beam lines to be aligned. The components of each beam line have to be aligned with a high accuracy. The standard deviation of any point over a range of 576 m (this is the maximum betatron wave length) in the transverse direction should be better than 0.5 mm horizontally and 0.1...0.2 mm vertically.

Therefore it is too inefficient to align each beam line separately. It will be more efficient to have only one alignment (basic alignment). Then the components of each beam line have to be connected to the basic alignment.

3.9.3 Basic Alignment

The reference points for the basic alignment are fixed on the tunnel wall. They are only target points, not suitable to put geodetic instruments on. The distances between the points are between 25 m to 50 m. The geodetic instruments will be set on a moveable carriage (Fig. 3.9.1). The carriage rolls on beams mounted on the tunnel wall above the causeway. The carriage can be fixed to the tunnel wall at any position with clamps.

The basic alignment can be carried out with Precision Total Stations (Tachymeters). A Tachymeter measures the horizontal angle, the vertical angle and the distance to the target point simultaneously. These measurements can be carried out either manually or automatically. The targets are Taylor–Hobson–Spheres for the manual and prisms in Taylor–Hobson–Spheres for the automatic methods. "Automatic" means that a computer controls the instrument, and the target recognition will be done automatically. The accuracy is for both instruments

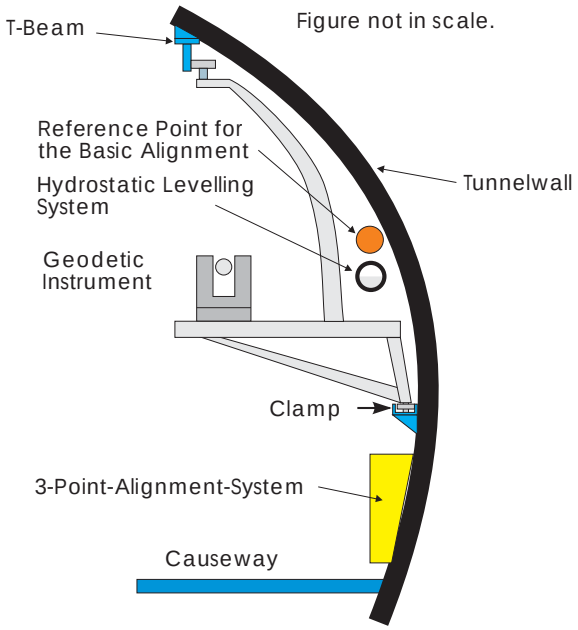


Figure 3.9.1: *Movable Carriage for the Geodetic Instruments and Locations of different Measuring Devices*

- horizontal direction 0.2 mgon (3 μrad),
- vertical angle 0.2 mgon (3 μrad) and
- distance 0.1 mm.

To coordinate the reference points on the tunnel wall the movable carriage with the Tachymeter can be set in front of each reference point. Now one can measure to several reference points, for example to two or three points back– and forward. How many points are to be measured is a question of the demanded accuracy. Fig. 3.9.2 shows the expected accuracy of a point in the middle of a 600 m long area in dependence on the number of measured target points and of the distances between the target points. The required accuracy can be achieved by using at least 3 reference points.

In the calculation only random errors are taken into account. What are the effects of systematic errors (for instance the refraction of air)?

3.9.4 Systematic Effects: Refraction of Air

Refraction of air means that if the density of the air is not constant then the line of sight to the target point is not a straight line exactly. The density depends mainly on the temperature of air. If there is a constant gradient in temperature for example of 0.1 degrees/m, a line of sight with a length of 600 m will follow a circular curve with an offset of 4.5 mm. These effects exceed the demanded accuracy by far and it is very difficult to determine their actual amounts. But there are some techniques available to reduce the effects of refraction of air:

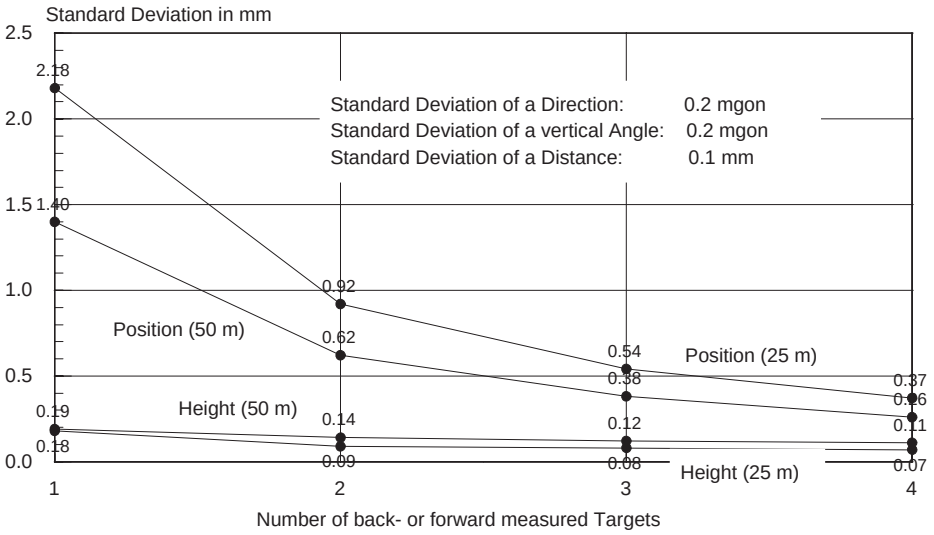


Figure 3.9.2: *Expected Accuracy in the Middle of a 600 m long section*

- hydrostatic levelling system for the vertical position,
- angle measurements without the effects of refraction (under development),
- Heelsche Alignment System (It does not work here, because the collider is not a straight line. In the vertical it follows the earth curvature.),
- stretched wires in an overlapping manner for the horizontal position,
- 3–point–alignment–technique for the horizontal position.

In the following the hydrostatic levelling system and the 3–point–alignment–technique will be described in detail.

3.9.5 Hydrostatic Levelling System

In the part of the tunnel which is horizontal, it is possible to install a pipe filled half with water. The pipe is mounted along the tunnel wall, so that the reference points are directly above the pipe (Fig. 3.9.3). A movable device with a capacitive distance sensor can be set in the reference point to measure the height of the reference point in respect to the water level. The water level is horizontal but it is not a curve with a constant radius. The accuracy of the sensor is better than 5 μm und therefore it seems possible to measure the heights of the reference points over a length of about 600 m better than 0.1 mm. If the water level makes vibrations one can reduce this effekt by breakwaters which are mounted in the pipe.

3.9.6 3–Point–Alignment–System

A technique to reduce the effects of refraction of air will be given by the "3–point–technique". For example there are mounted reference plates on the tunnel wall in

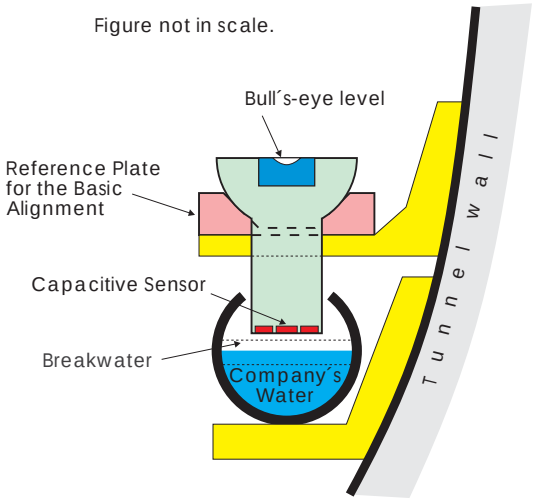


Figure 3.9.3: *Hydrostatic Levelling System*

regular distances (1 m oder 2 m). From a movable bar in each position the distances will be measured to the 3 opposite reference points (Fig. 3.9.4) to better than 1 μm by capacitive sensors. The distances are very small, about 0.5 mm. The distances define the horizontal angle of the reference point in the middle to the both other neighbouring points. Therefore a traverse is measured and the displacements of each reference point in respect to a reference line can be calculated. The accuracy depends mainly on the stiffness of the bar. If the accuracy of the distance sensor is 0.1 μm, over an area

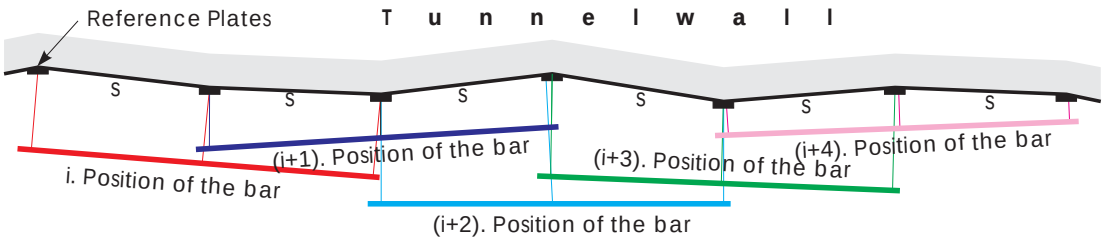


Figure 3.9.4: *Straightness measurements with the "3–Point–Technique"*

of 600 m the displacements of the straightness can be measured better than 0.5 mm (Fig. 3.9.5). In this calculation the distances between the reference plates are 1 m and 2 m. Now it is possible to have bars with lenghts up to 10 m or 20 m. But in this case the bars are not stable enough by themselves. With a stretched wire and a differential diode mounted in the bar the stiffness of the bar can be controlled. A 26 m long test assembly with reference plates in 1 m distance has shown that the calculated accuracy will be achieved.

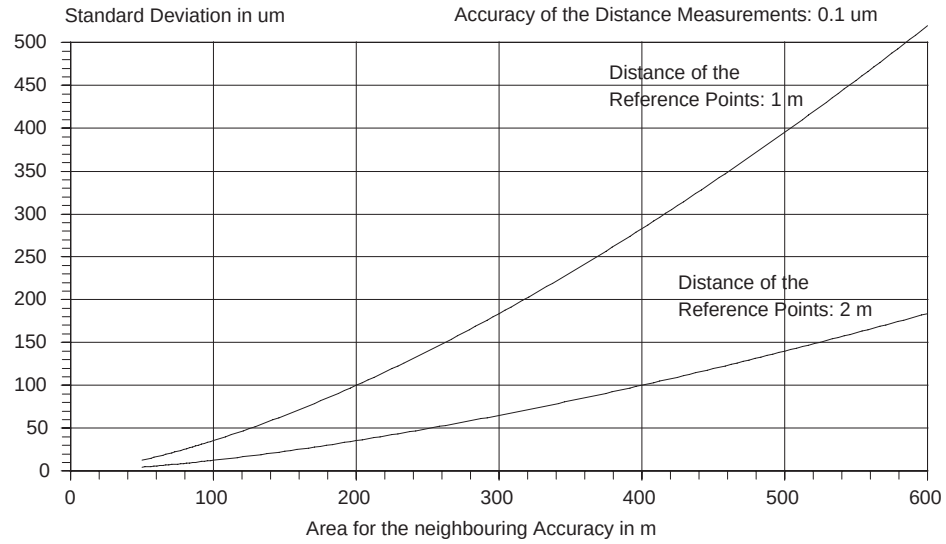


Figure 3.9.5: *Expected neighbouring accuracy for straightness measurements with the "3–Point–Technique"*

3.9.7 Transfering the Coordinates

After performing the measurements as discussed above, the coordinates of the reference points on the tunnel wall are available with the demanded accuracy. In the next step the coordinates of the reference points have to be transferred to the components of each beam line. For this each magnet should have two reference plates for targets and a support to measure the roll. The movable carriage is positioned at any place so that the neighbouring reference points and the points of the component which has to be aligned can be pointed out. The measurements to the reference points define the coordinates of the point on the movable carriage. Then the alignment of the component can start. For this one can use the following geodetic instruments

- Precision Total Stations (Tachymeters),
- Lasertracker and
- new instruments (under development).

The alignment will be carried out for each component individually. To control the adjustment it is advisable to make a control measurement in which the reference points and the magnet points are included simultaneously. To have a good redundancy each magnet point is measured twice from the neighbouring instrument station of the geodetic instrument.

3.10 Conventional Facilities and Site Considerations

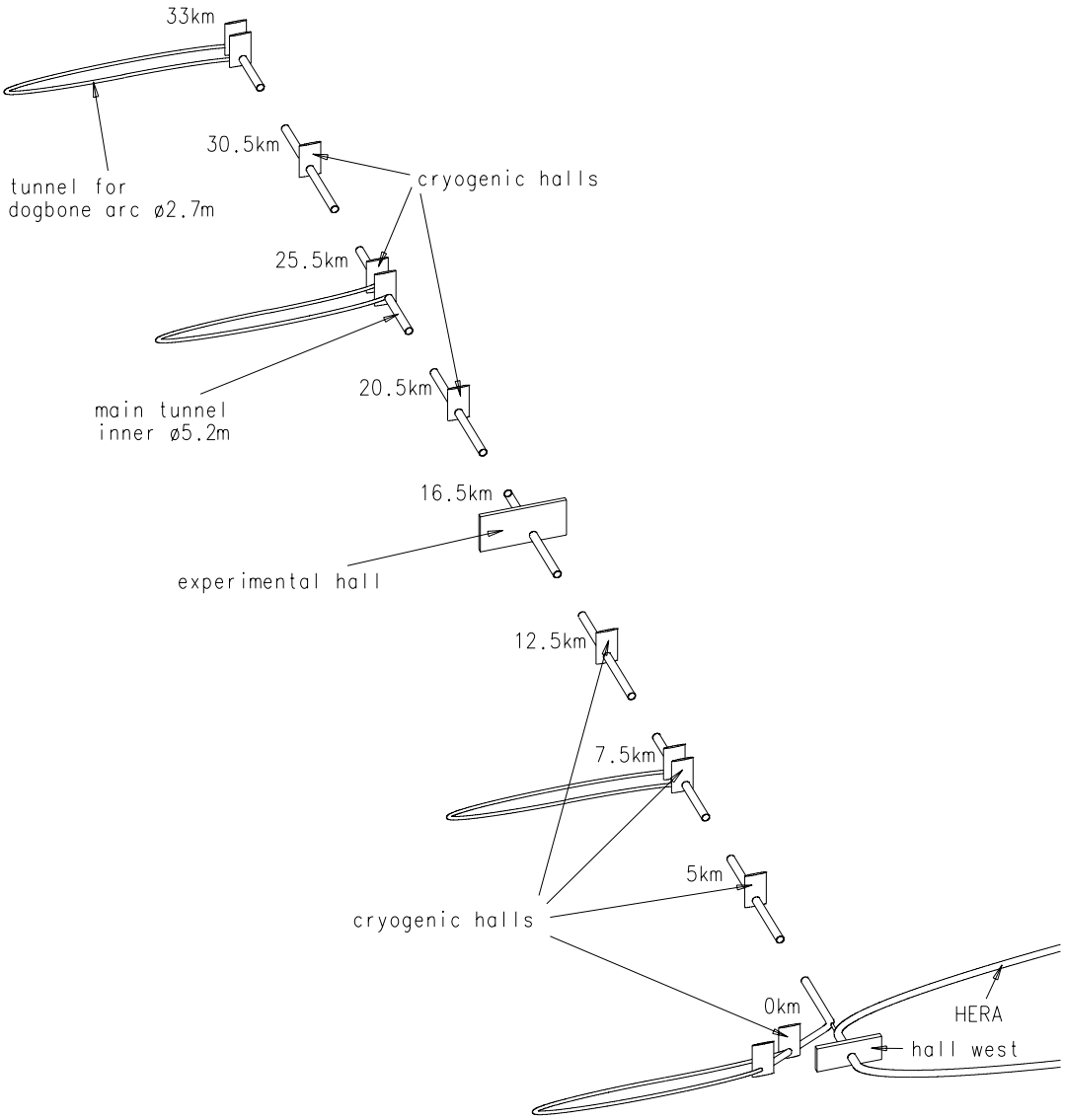


Figure 3.10.1: *Schematic view of the TESLA Linear Collider at DESY.*

3.10.1 Introduction

Different conventional facilities must be set up for the future Linear Collider. The detailed considerations in the following assume that the Linear Collider facility is built at DESY in Hamburg. Not all of them are valid at other locations without modifications. A discussion of Fermilab as a possible site is given in section 3.10.11. In any case coupling a new large accelerator complex to an existing laboratory site is advantageous from the point of view of investment cost, construction time and available know-how in accelerator technology, particle physics and synchrotron radiation research. Further boundary conditions we take into account here are the integration of the Free Electron Laser facility and the possibility to make use of existing accelerators and infrastructure at the DESY site.

These considerations predefine that the tunnel starts on the DESY site and runs in direction North-Northwest, parallel to the straight section West of HERA. By far the largest single construction is that of the tunnel for the accelerators, the beam lines and for the various auxiliary components. This tunnel has a total length of about 33 km and an inner diameter of 5 m. An underground experimental hall for the Particle Physics detector is located in the central area which is layed out to also accomodate the X-ray Free Electron Laser laboratory. Administrative, social and technical buildings are placed on this common site, too. Additional underground halls are located at the two ends of the linac tunnel. Seven cryogenic plants are distributed along the tunnel, with one plant at the beginning of the electron linac. The other six plants are connected to the tunnel by shafts. The four arcs for the two damping rings are located next to the main tunnel.

3.10.2 Overall Site Layout

The Linear Collider tunnel will be built underground at a depth of about 15 m. The main part of the tunnel is in sand and in ground water. With well-proven tunnel boring technology, buildings at the surface above the tunnel remain completely unaffected. As shown in the top view (see Fig. 3.10.2) the complete tunnel is straight (except for a short initial section, see below) and in a profile view (see Fig. 3.10.3) it is mainly horizontal (i.e. following the earth's curvature). HERA is rotated counterclockwise by an angle of 22.7^o out of the northerly orientation (approximately North-Northwest orientation). This orientation was chosen as the main direction of the Linear Collider to enable the option for producing collisions of protons stored in HERA with electrons from the main linac (note that the beam direction in the HERA-p ring has to be reversed for this option). The beginning of the tunnel starts laterally shifted with respect to the HERA hall West and after 250 m bends with a radius of 500 m into the main direction. The main electron linac starts at the beginning of the straight tunnel. The initial section of the tunnel houses the prelinac together with the beam preparation for the FEL operation.

The interaction point HERA West is about 20 m above sea level. The axis of the main part of the linac tunnel lies about 10 m below sea level. The straight section

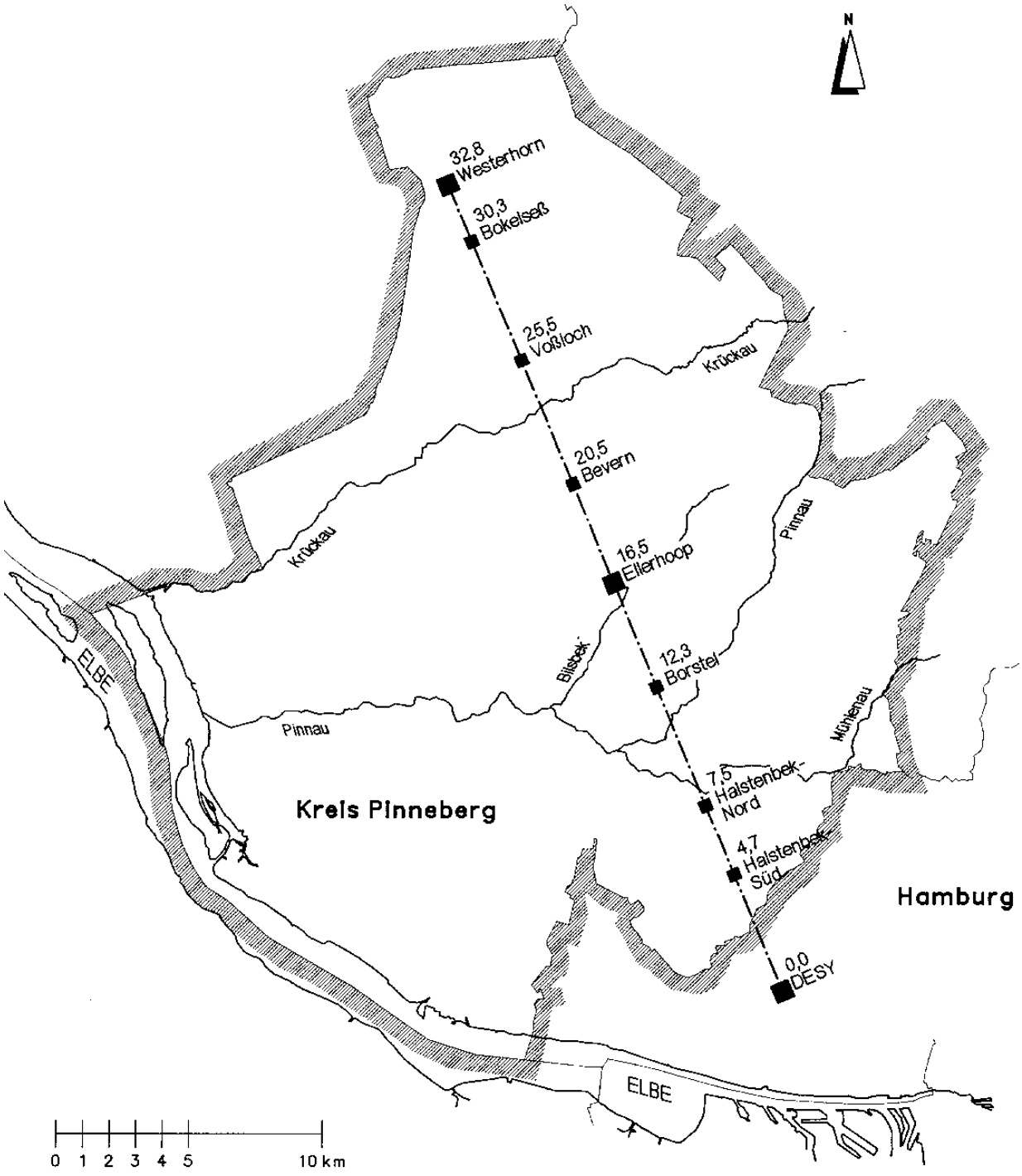


Figure 3.10.2: *Overall view of the Linear Collider site at DESY.*

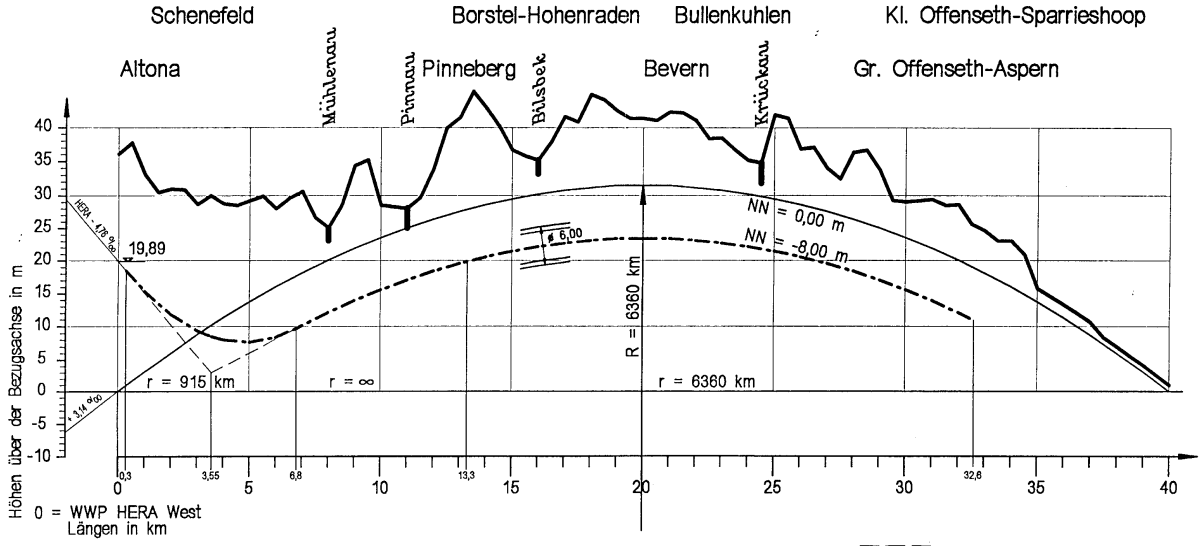


Figure 3.10.3: *Expanded profile of the Linear Collider area North-Northwest of DESY.*

HERA West has a slope of about 8 mrad out of the horizontal. Therefore there is a smooth transition from the initial slope into the horizontal direction with a bending radius of about 1000 km. Figure 3.10.3 shows an expanded profile of the linear collider area North-Northwest of DESY together with the collider tunnel. The depth of the tunnel below ground level of about 15 m is more than sufficient to guarantee shielding against radiation. The tunnel is below the ground water level over nearly the entire length.

The tunnel will be built by a tunnel boring machine similar to that used to construct the HERA tunnel. Identical concrete segments have been chosen for the lining of the main tunnel. The minimum inner diameter is 5 m, with the actual diameter increased to 5.20 m due to tolerances in the tunnel boring. An underground experimental hall is in the center of the Linear Collider complex, with additional underground halls at the two opposite ends. These three halls are built in open-cut excavations. The experimental hall and the end station are located on new areas. The start station together with one cryogenic plant are located on the existing DESY site. Additional areas are necessary for the six remaining cryogenic plants. The plants are connected with the collider tunnel through shafts. Two dog-bone damping rings are foreseen for the damping of the electron and positron beams before the main acceleration. The straight sections for the rings are located in the main tunnel, but additional tunnels must be bored for the four arcs. They are at the ends of the tunnel and at a distance of 7.5 km from these points at a cryogenic plant. These tunnels have an inner diameter of 2.5 m and the arcs have a radius of 80 m.

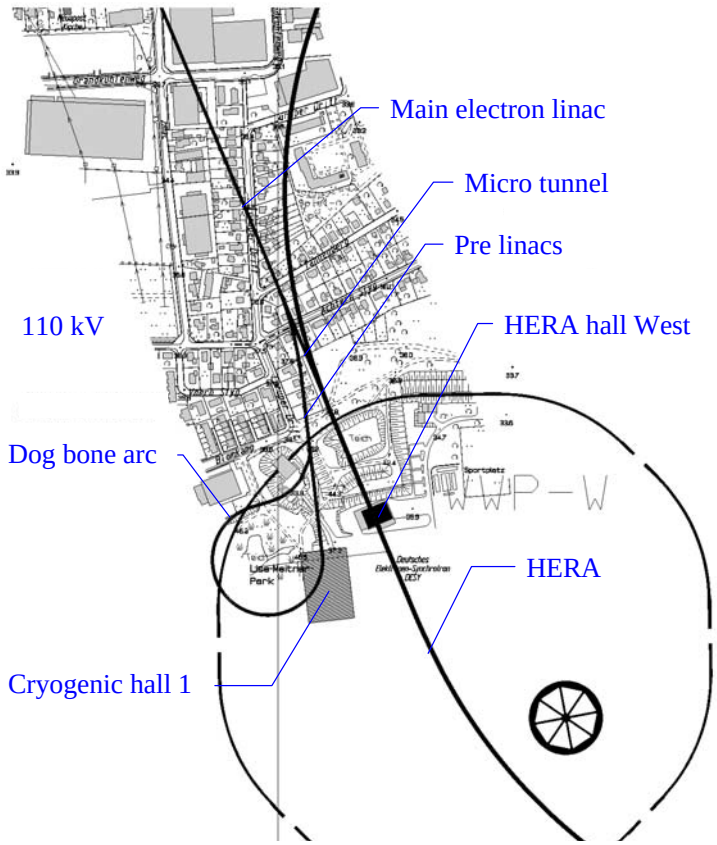


Figure 3.10.4: *DESY site of the TESLA Linear Collider.*

3.10.3 DESY Site

The Linear Collider tunnel starts on the existing DESY area next to the HERA hall West. At the beginning there is an underground hall. The tunnel boring machines for the main tunnel and for the dog-bone arc can start in this hall. An additional shaft is necessary for the other connection of the arc to the main tunnel. The starting point for the main tunnel boring machine is laterally shifted from the HERA hall West. After a distance of about 250 m the tunnel bends with a radius of 500 m into the main direction. From that point on the tunnel is straight in a top view up to the end station. The same point is the beginning of the main linac. The initial section houses the beam preparation (injector, bunch compressors) for the FEL and the collider operation. On ground level above the underground hall is a cryogenic plant for the first 2.5 km of the superconducting electron linac. The connection between the collider and the HERA tunnel can be made by a small tunnel. In Figure 3.10.4 an overview of this complex is shown.

3.10.4 Experimental Area

The experimental hall with the interaction point is located at the center of the Linear Collider, at a distance of 16.5 km from the hall HERA West. The underground hall has

an area of 35 m × 80 m and a height of 20 m below the portal crane in present planning. Space for a second hall with the same dimensions but laterally shifted is available in case a second detector would be installed. The X-ray Free Electron Laser laboratory is also located on this common area together with the administrative, social and technical buildings. The electron beam dump is located at a distance of about 600 m from the interaction point; the positron beam dump at a distance of about 300 m. Both dumped beams are orientated downwards with an angle of 20 mrad. Small underground halls with access from the ground level are foreseen for the dumps. The total length available for the beam delivery system is 3 km, the length of the central experimental site about 800 m. The positron dump is located outside the main experimental area below an additional external area. Sufficient concrete shielding around the beam dump is foreseen to avoid any activation of the surrounding soil and ground water.

Site	Position L/km		Power S/MVA
DESY	0	Cryogenic plant 1	12
		Modulators	17
Cryogenic hall 2	5	Cryogenic plant 2	12
		Modulators	17
Cryogenic hall 3	7.5	Cryogenic plant 3	12
		Modulators	11
Cryogenic hall 4	12.5	Cryogenic plant 4	12
		Modulators	11
Experimental area	16.5	Experiment	3
		FEL laboratory	3
Cryogenic hall 5	20.5	Cryogenic plant 5	12
		Modulators	11
Cryogenic hall 6	25.5	Cryogenic plant 6	12
		Modulators	11
Cryogenic hall 7	30.5	Cryogenic plant 7	12
		Modulators	11
End station	33		3
		Sum	182

Table 3.10.1: *Installed electric power for the TESLA Linear Collider. These values include redundancy and the reactive currents, and allow for the energy upgrade with a nominal power consumption of* $P = 130\text{MW}$.

3.10.5 End Station

The end station is a small underground hall. This hall is necessary as a starting point for the tunnel boring, for access to the tunnel during installation and operation and for one connection of the dog bone arc. An additional shaft is necessary for the other connection of the dog-bone arc.

3.10.6 Cryogenic Halls

In addition to a new cryogenic plant located on the existing DESY site six additional plants are distributed along the linac. At each of these points a hall with an area of 35m $\times$ 100m is planned. A hall houses a cryogenic plant which supplies 5 km of the superconducting linac. The first 5 km are designed for a higher linac repetition rate (including FEL and Nuclear Physics operation). Here the cooling power of one plant supplies 2.5 km of the linac. Each external hall is connected to the main tunnel through a shaft of about 12 m inner diameter. The modulators for the klystrons are on a second level in these cryogenic halls. These areas are also needed for the power distribution and water cooling plants.

3.10.7 Tunnel Layout

The main tunnel of the TESLA Linear Collider has a minimum inner diameter of 5 m (5.20 m minus construction tolerances). The main components installed in the tunnel are the cryomodules for the superconducting linac and the beam lines for the dog-bone damping rings, for the FEL laboratory and for the ELFE@DESY (Nuclear Physics) option. In addition there are numerous auxiliary components. In Figure 3.10.5 a tunnel cross section is shown, chosen at a position of the most complex situation. The cross section is divided vertically into three parts. One third is for sensitive accelerator and beam line components and for the rf distribution. A walk path is located on the outer side for installation and maintenance. The place for the transport system is in the center above the place for components which need easy access. On the bottom the hot and cold water pipes are installed in a small channel; this channel serves for the collection and detection of leakage water. The klystrons, electronics and vacuum pumps are mounted in special containers, the center of mass of which are just below the hook of the transport system. This allows for easy installation and exchange of components with a relative short lifetime. The transport system is a so-called monorail, utilizing an I-beam fixed at the top of the tunnel. The vehicle is used to transport equipment and persons.

The last part is the main walk path. The space above the grids is free at all times and the space below is reserved mainly for the cable connection between modulators and klystrons. The free space is for an escape route and for lines-of-sight for optical survey.

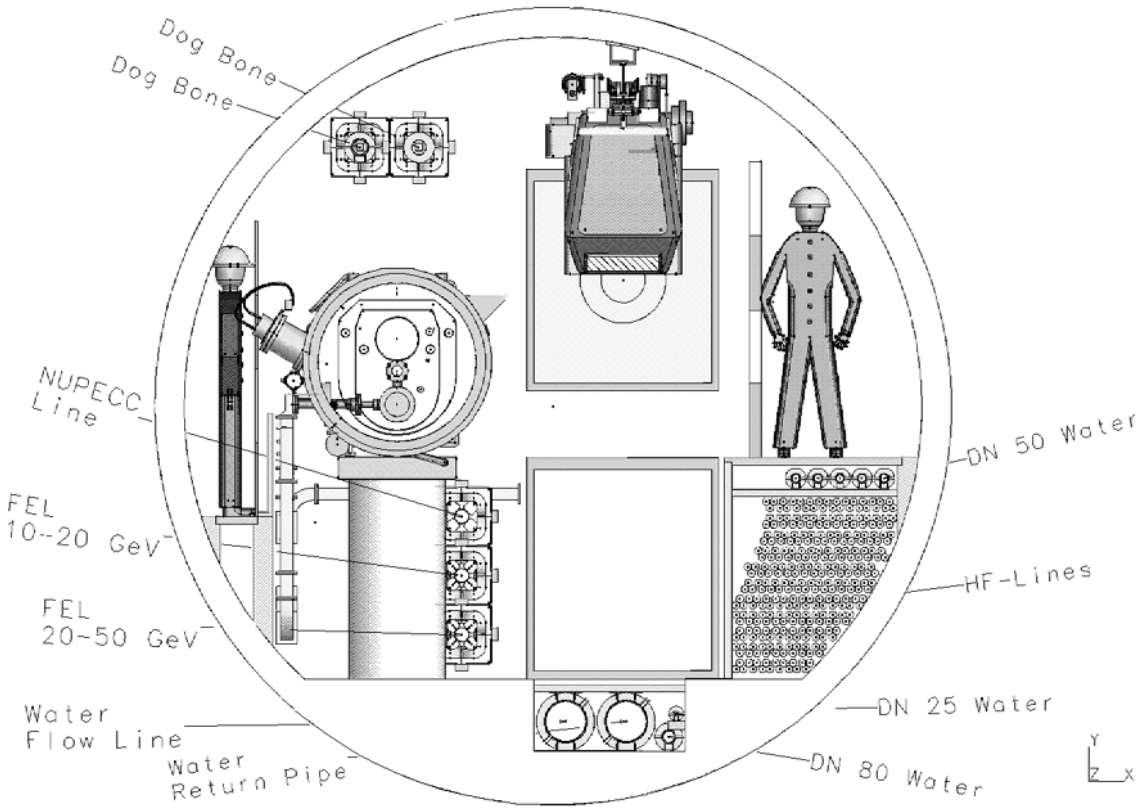


Figure 3.10.5: *Cross section of the TESLA Linear Collider tunnel.*

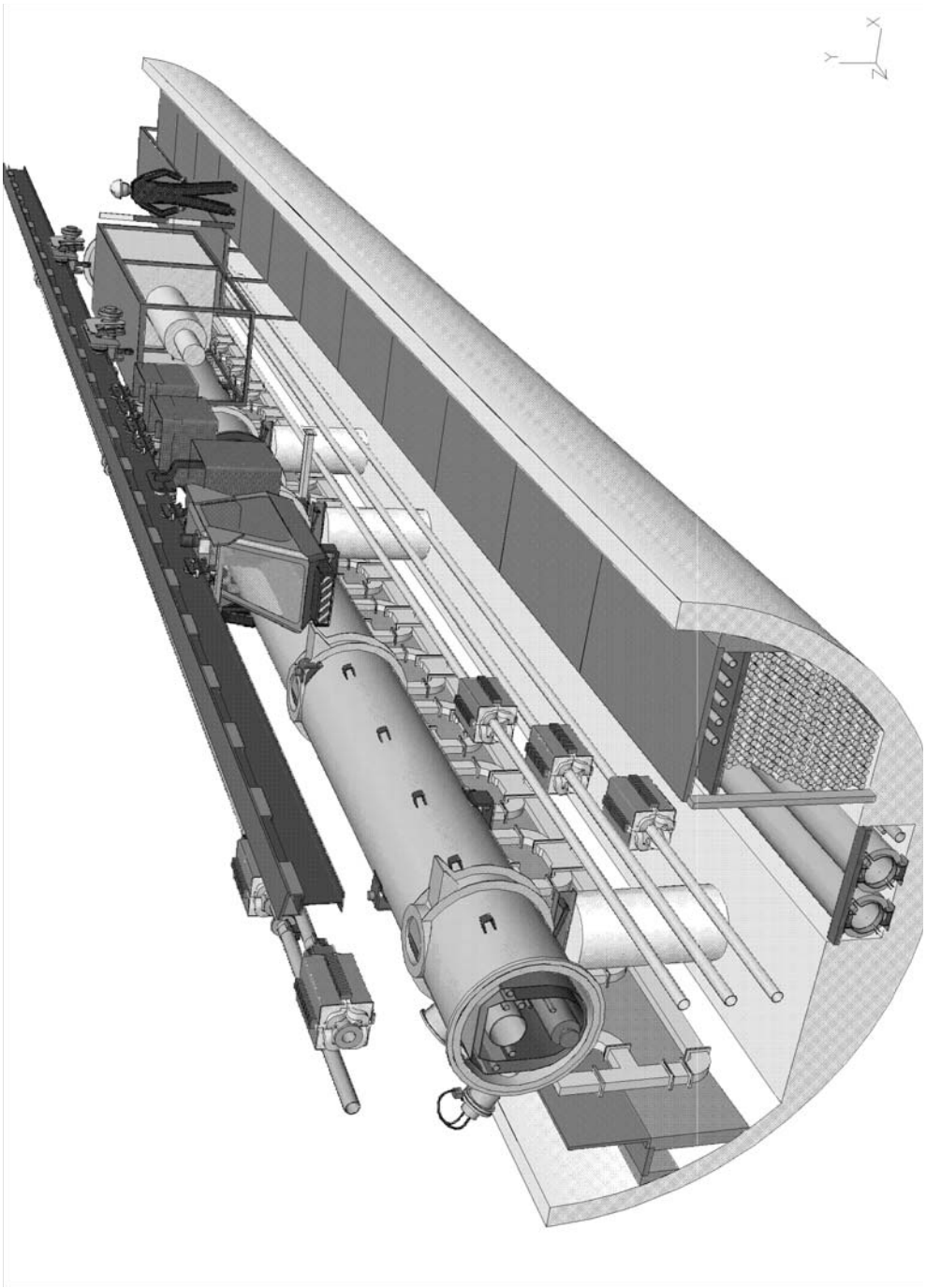


Figure 3.10.6: *View into the TESLA Linear Collider tunnel.*

3.10.8 Power Distribution

In order to provide sufficient redundancy and a margin for reactive power, the power distribution system is layed out for a total AC-power of 182 MVA. There are essentially two options to distribute the power to the cryogenic/modulator halls. First, the local medium voltage grid can be used to supply each hall individually. Second, the power can be obtained from 110 kV high voltage lines and distributed to the individual halls at medium voltage level (36 kV) through the accelerator tunnel. The latter option is sketched in Fig. 3.10.7. Guided by the availability of 110 kV lines in this part of the country, 4 high-to-medium voltage transformer stations are foreseen: one on the DESY site and three others at external halls along the linac. For the distribution to the other halls medium-high voltage cables ($3\times$ 240 mm^2 Al) are installed in the tunnel. The losses on these cables amount to about 50 W/m.

The decision which of the two options for the power distribution is preferable can only be taken once the availability of power lines and cost aspects have been analysed in more detail.

The amount of AC-power required inside the linac tunnel is relatively small. Power is needed for utilities such as klystron focussing magnets, water distribution pumps and tunnel light. In total, about 10 MW have to be distributed in the linac tunnel over the full length of 30 km. The supply voltage in this cases is chosen at 690 V. It is also foreseen to have the magnet power supplies installed in the tunnel. There are about 400 quadrupoles and 600 steering coils in each of the two linacs which have to be powered individually. The estimated total AC-power for the magnet supplies amounts to 5 MW and the heat load to the tunnel air from this source is about 20 W/m. The magnet power supplies are also connected to the 690 V system.

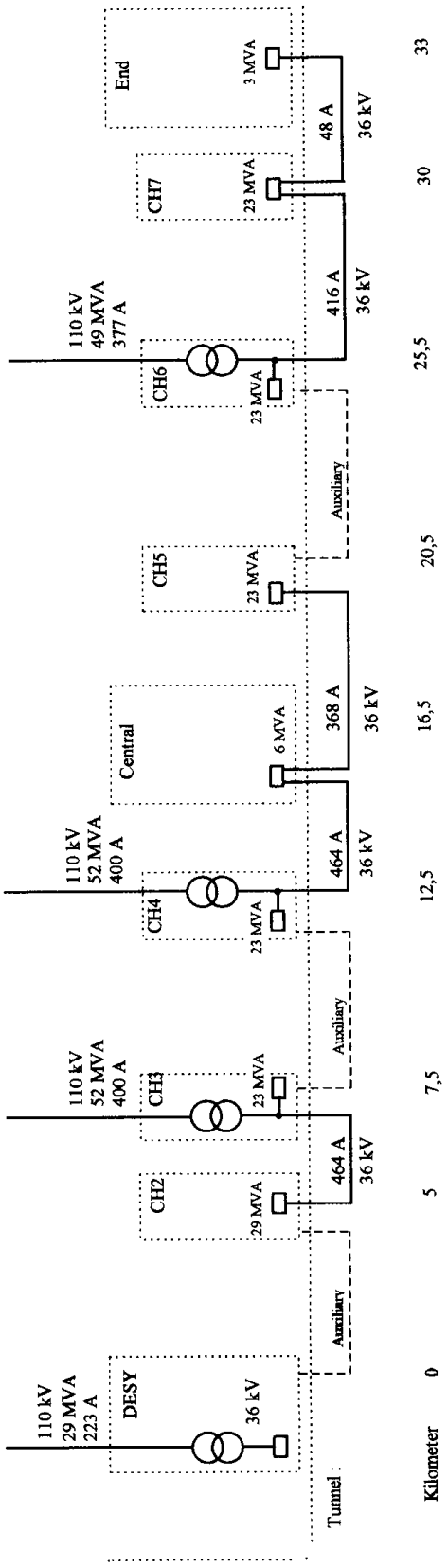


Figure 3.10.7: *Sketch of the power distribution system.*

3.10.9 Water Cooling System

The layout for the cooling circuit of the TESLA linear accelerator is divided up into separate sections in the same way as the cryogenics system. Every cooling station has cooling towers, heat exchangers, pumps and control systems. The layout of the cooling plant is depicted in Fig. 3.10.8.

The cooling tower is a so called wet cooling tower which works by evaporation of part of the falling water. Therefore an important number for such a cooling tower is the maximum value of the humidity in the air. A derived number of this is the wet bulb temperature. For this linac a value of 19.5 °C to 20.0 °C is preferred. There will be 15 towers for every station, placed beside the cryogenic halls. The fan velocity of the cooling towers will be controlled to maintain a constant temperature at the heat exchangers. The heat exchangers themselves will be controlled by a bypass system to provide a nearly constant temperature into the tunnel. The main tubes in the tunnel are distribution tubes. Because all elements are connected in parallel, there will be a subdivision system to avoid too many connections to the main tube. The material of the main tubes will be stainless steel and copper for the subdivision tubes. The length of the subdivision system will be 144 metres, this is the distance between six klystrons. The RF-power is either in the collector or in the loads, and therefore these elements can be connected in series. To make up for the pressure losses in a klystron/collector, a booster pump will be necessary.

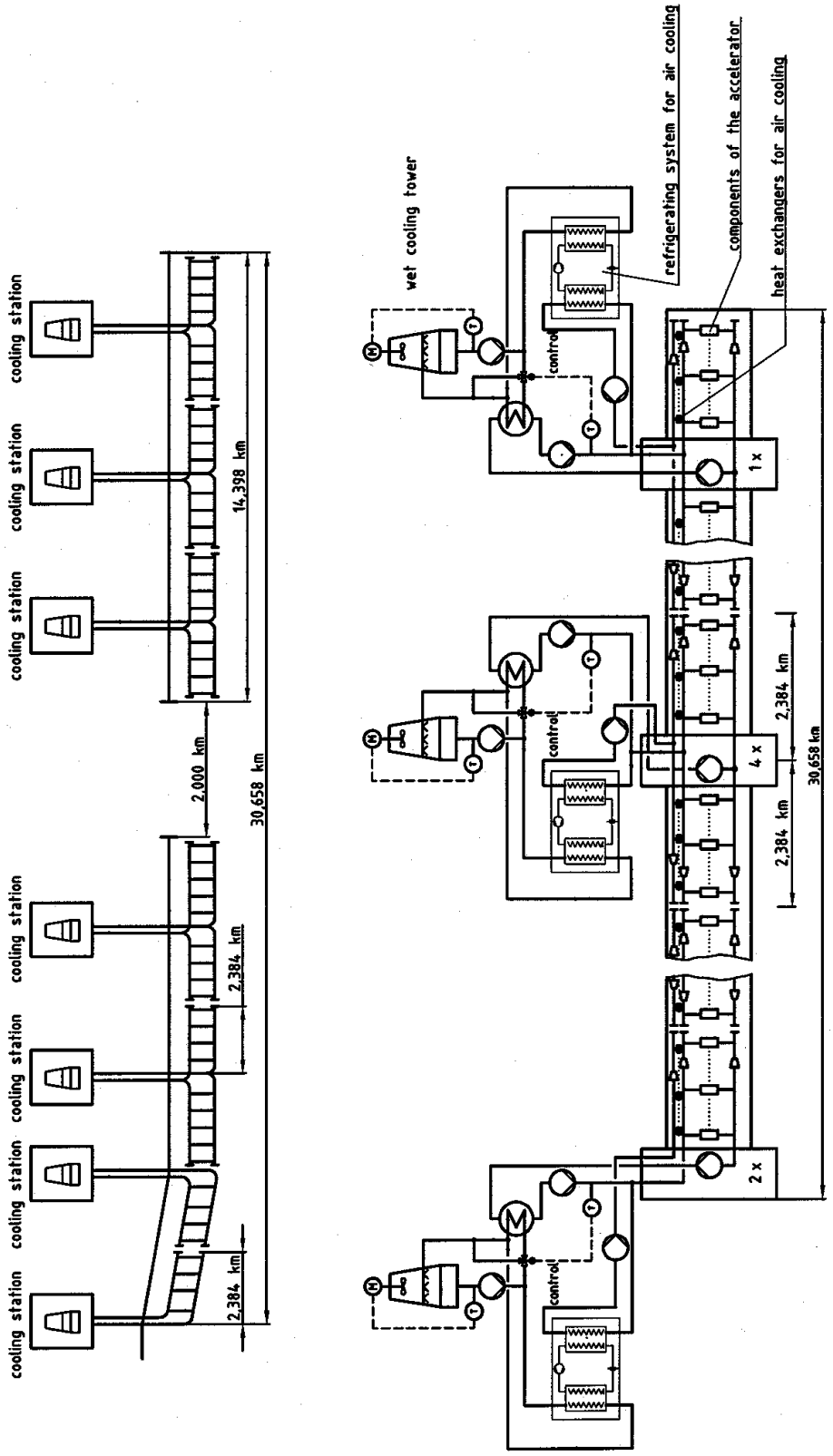


Figure 3.10.8: *Sketch of the cooling tower plant.*

The water will be de-ionised water with a conductivity of 1 mS/cm. Therefore the only allowed materials are stainless steel, copper and bronze (gunmetal) . The maximum pressure in some RF-elements in the present design is low. Therefore boosterpumps in the return water line will be necessary to keep the pressure in the return line as low as possible. The diameter of the tubes varies over the length, so that the pressure gradient has a nearly linear characteristic. The first diameter is 250 mm. The pressure loss over the length is 2.5 bar. The temperature of the return water is 60 °C - 70 °C . Therefore the return water tube will be insulated with mineral wool of 100 mm thickness. Then the heatloss from the return water into the tunnel will be 50 W/m. To minimise the mechanical stress in the tubes by thermal expansion there are compensators in the tubes. Every six metres will be a slide support point for the tube and between two compensators a fixed support point. For the slide elements PP-sheets are preferred. These elements are already used as slide sheets in the other DESY accelerators. The last or direct connections to all elements (e.g. magnets, klystrons etc.) will be done by EPDM-rubber tubes. The advantages are: resistance to radiation, high flexibility for the connections, good and quick installation, direct connection between elements at a high electrical potential and earthed copper tubes, etc.

3.10.10 Air Conditioning and Ventilation

The tunnel sections between cryo halls have a length of 5 km with no intermediate access. The tunnel is an underground building and therefore forced ventilation is required. Note that the ventilation does not contribute significantly to cooling, which is almost completely dominated by the tunnel walls. The heat emission to the tunnel air must be minimized in order to avoid excessive temperature rise over long periods of operation. Removal of the heat from the main sources by the water cooling system and shielding of hot surfaces will ensure this. In addition, a cold water cooling pipe can be installed.

During access persons have to work in the tunnel and in case of emergency a smoke-free escape must be provided. In case of a fire the ventilation removes the smoke in one direction, and fire fighting would begin from the opposite direction.

The regular velocity of the tunnel air should be 0.6 m/s (corresponding to a volume flow of 30,000 m^3/h) and for smoke removal >0.8 m/s. The pressure drop over 5 km is 500 Pa, which creates a force of 1,000 N on a 2 m^2 door. This must be taken into account in the design of the escape exit. The speed at which the smoke moves is sufficiently slow to be able to leave the tunnel safely in the direction of air flow.

The fire risk, determined by the product of fire load times fire probability, should be minimized. The fire load is given by the amount of inflammable materials. These are in particular the oil in the pulse transformers and the cables. The fire probability is related to heat sources and high-voltage devices (klystrons, power supplies, etc.). Since the first 15 min. after ignition of a fire are very important as fire fighting is most effective, a smoke detection system will be installed. After detection of smoke, klystrons and power supplies can be automatically switched off.

During access to the tunnel the fire probability can be drastically reduced by shut-

ting down the main ignition sources. There will be no access during machine operation.

A recirculation of the tunnel air is not possible. At the beginning of the linac tunnel fresh air is blown into the tunnel. The cryo halls are bypassed and only a small fraction of fresh air is added. At the endstation hall the air from the tunnel is extracted. In case of a fire, fresh air is blown directly into the respective section of the tunnel and the smoke is extracted at the next hall. The ventilators are switched to high speed in that case.

3.10.11 Fermilab as a Potential Site for the Linear Collider

3.10.11.1 Introduction

Fermilab like DESY is a potential location for a future Linear Collider. The existence of a functioning High Energy Physics Laboratory, which can provide the organizational structure, experienced staff, physicist and engineering teams and know-how, the basic infrastructure and plant, and the user support, is key to carrying out a project as large and technically challenging as TESLA in a cost effective and timely way. Even if major civil and technical components of the existing facility are not used, the stable organization and general infrastructure are of great value in proceeding effectively with project design and construction.

The Fermilab site was originally chosen in the 60's because it and the surrounding area fulfilled the major requirements of rational site selection: easy access to a major city and airport, good highway and rail transportation, work force availability, good and stable geology, good power and water availability, reasonably flat and open land (to west now), good community resources and infrastructure, weather, etc. Though the local area has been one of the fastest growing in the country, it still is reasonable to consider off site location of a large accelerator complex, as most construction will be underground. This is especially true if one looks to a site in the open farm lands that still start just a few miles to the west of Fermilab and the Fox River.

Two significant features of the area surrounding Fermilab are the geology and the availability of electric power. The dolomite bedrock of the area is ideal for underground tunnel construction and extensive experience exists from the Chicago water project TARP. This project has constructed almost 100 miles of tunnel throughout metropolitan Chicago.

Near to Fermilab is a large power distribution net of Commonwealth Edison with a hub centered at Electric Junction just 2 miles south of the laboratory. Commonwealth Edison is one of the largest generating systems in the country with over 20,000 MW of owned capacity. The transmission network provides two 345 kV power corridors of particular interest. One runs north south on the east side of the lab and the other runs from just south west of the lab, across the Fox river straight westward for almost 40 miles. It is advantageous to consider linear collider site locations in conjunction with these power lines, and to adopt the natural north, south, east, west, road grid of the region. Easements for the collider will undoubtedly be easier to obtain if associated with utility or road right of ways. This strategy has been used by TARP which has

extensively sited its tunnels under roads, canals, and rivers. In this sense linear colliders maybe much easier to site than circular accelerators.

In the discussion which follows substantial use has been made of work carried out, reports written, and information provided by the State of Illinois and its Illinois Department of Energy and Natural Resources (ENR). Extensive work was carried out by the State in its preparation of the Illinois SSC site proposal. This information remains invaluable for any future project proposal. Figures depicting the geology of the area have been supplied by Illinois State Geology Survey, Illinois Dept. of Natural Resources.

3.10.11.2 Overview of Possible Site Layouts

Fig. 3.10.9 shows two possible site areas which might be considered. a) the West Line site could be connected to Fermilab or not. The Fermi Main Ring tunnel could be used for the damping ring, for instance. The collision area would lay to the west about 10-15 miles near the same interstate highway that serves Fermilab. b) the North South site could have the interaction area on site and possible use the MR tunnel for the damping ring(s) as well.

West line proposed site This proposed line runs due west 28 miles (45 km) from Kane- Dupage County Line boundary (which runs north south through the west side of the Fermilab site). This site line is located about 1/2 to 1 mile south of Fermilab south boundary . It is chosen to align with a 345 kV power line which starts about 2 miles to the west of County Line and east of the Fox River. The 30 mile extent of the possible site is approximately defined by the Troy bedrock valley just west of Dekalb. This valley is cut deep into the bedrock and is filled with glacial sediments of clay, sand and gravel. (The 30 mile extent is not a firm limit but the top of the bedrock drops below an elevation of 450 ft in this region and ground elevation has risen to 900 ft .)

Fig. 3.10.10 depicts the geological cross section. The tunnel and experimental hall would be located completely in the Galena - Platteville dolomite and would be about 350 ft below the ground surface. The tunnel could as well be situated on a slope from west to east following the contour of the bedrock (150 ft/10 miles).

The 345 kV power line has a dogleg horizontal offset of 0.3 miles at one location alone its length. The bends of this dogleg are about 12 degrees. It would be interesting to see just what impact incorporating this offset into the collider design might have.

The center of the layout is near the RT. 47 exit on Interstate 88. The Interstate would pass over the tunnel near here (about 2 miles). The Interstate exit is about 10 miles from the one that serves Fermilab site.

The far end of the line is within easy access of US 30 which runs parallel about 2 miles to the south. A railroad track runs adjacent to US 30 throughout the area. This rail could be an asset for construction but must be evaluated as a potential ground vibration source.

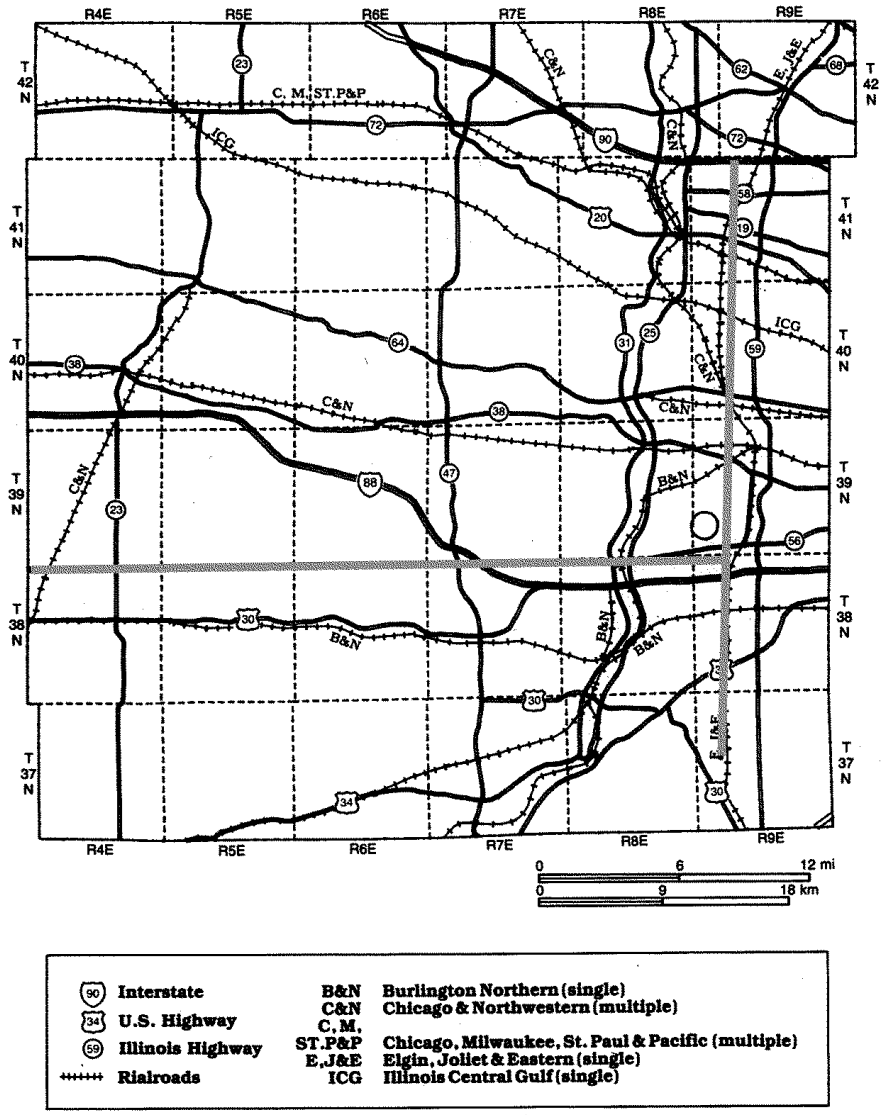


Figure 3.10.9: *Map of the possible site areas.*

North South line proposed site The line for this proposal would have the interaction region located near the eastern side of the Fermilab site. The tunnel line would follow or be closely associated with a combination of road, rail, and power corridors.

Fig. 3.10.11 shows the geological cross section. Two tunnel elevations are possible. One option has the tunnel and the hall completely within the Galena. The exact elevation would be chosen so that the roof of the hall would be about 25 to 30 ft below the top of the Galena. The tunnel depth would be about 370 ft below the surface at Fermilab.

The other option would put the roof of the experimental hall about 100 ft below the top of the Silurian dolomite. This keeps the hall out of the upper 75-100 ft of the

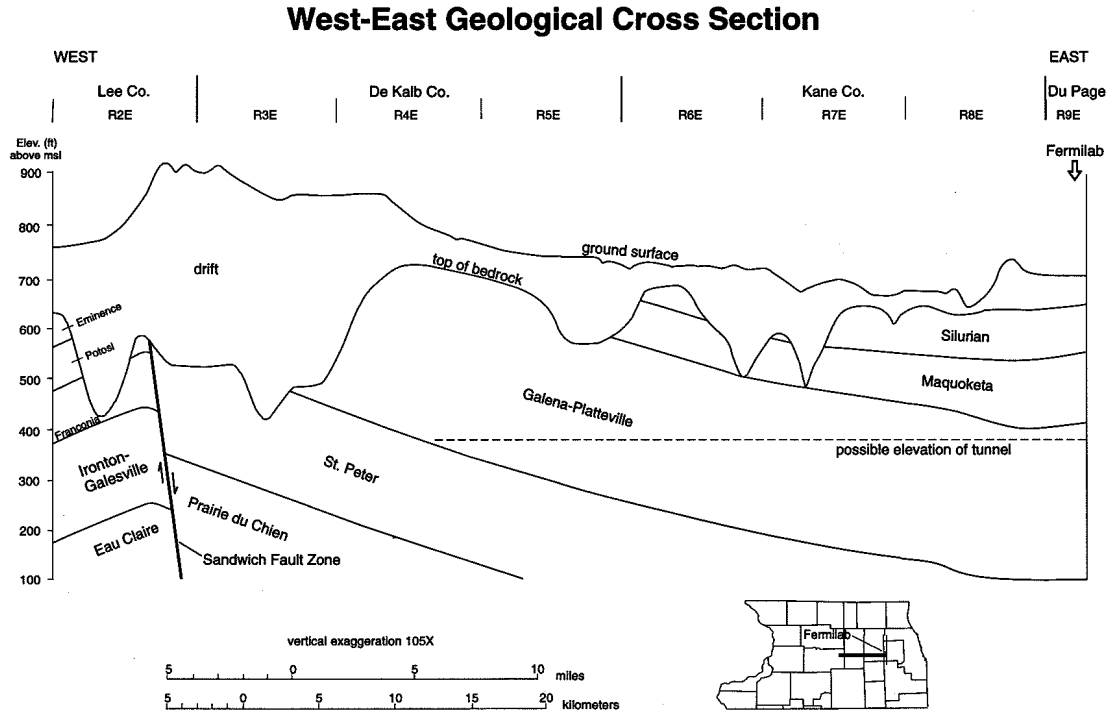


Figure 3.10.10: *West-East Geological Cross Section*

bedrock where more water can be present. The experimental hall could be entirely within the Silurian or the lower part of the hall could be in the dolomite portion of the Maquoketa. The tunnel would be about 175 ft below ground and entirely in the Maquoketa unit.

Cost estimates would need to be carried out to determine just what would be the best site choice.

3.10.11.3 Geology

The Illinois region consists of sedimentary bedrock units overlain with glacial deposits. The bedrock units consist of a number of carbonate dolomites (or limestones) and shales of considerable thickness overlying sandstone (Ancell Group). The bedrocks have not been significantly affected by folding or faulting. There is a general slope of the units of about 10 to 15 ft per mile to the southeast. There is negligible seismic risk and no potential for subsidence. There is an inactive fault system to the south west of the proposed siting areas. This fault system, the Sandwich Fault presents no major construction obstacle were it to fall in the site area. Sediments deposited over the fault indicate no movement has occured for 200,000 to 1 million years.

The dolomite bedrock is a relatively uniform and predictable medium. It provides an excellent and well understood tunneling environment. The dolomite itself is an aquatard. Where not cracked or fractured (e.g. surface interface) it is impervious to

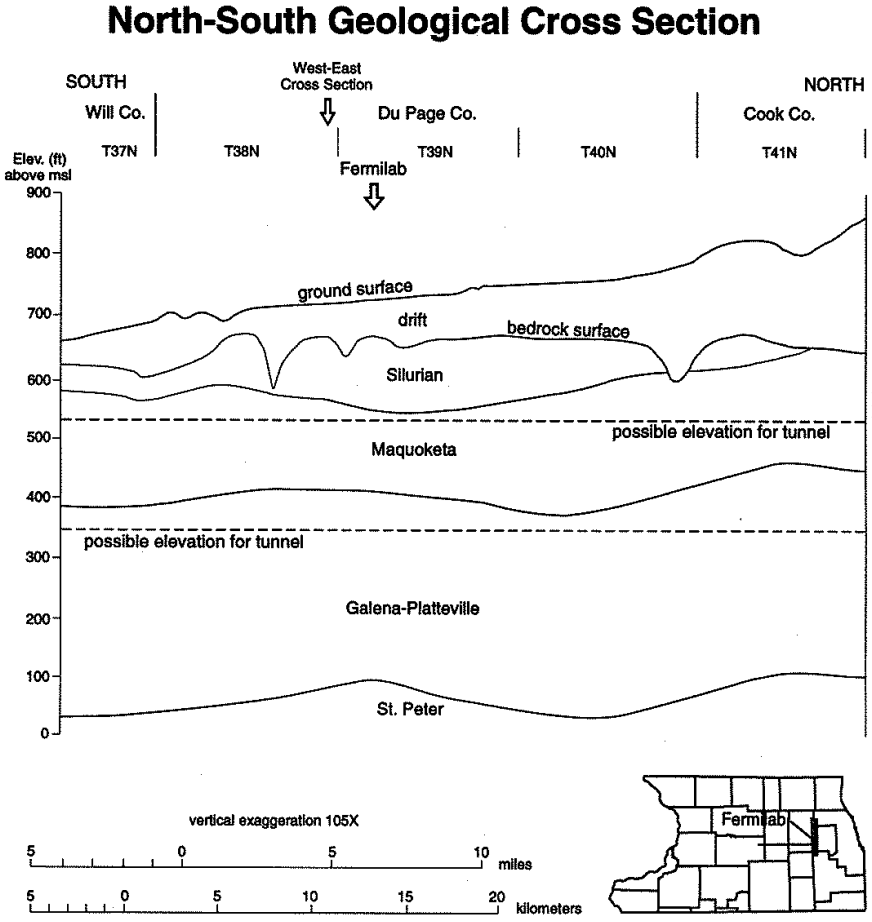


Figure 3.10.11: *North-South Geological Cross Section*

water. The glacial drift on the other hand is less predictable and not as desirable for tunneling. Water is found in the drift above the bedrock, at the very top of the bedrock at drift bedrock interface, and also in the deeper layers of sandstone below the dolomite. For the collider tunnel to stay located in uniform unfractured dolomite it may need to be located at typically 350 ft below the surface. (Considerably deeper than the about 15 m for the DESY site, but consistent with the LEP tunnel.)

The dolomite units consist of Silurian dolomite, Maquoketa- a shale with interbeded dolomite, and the Galena - Platteville dolomites. This latter group is of 300 to 350 ft thickness throughout the area. It is the preferred unit for tunnel construction. However both the Silurian and the Maquoketa can be viable alternatives. The Silurian does not cover the entire region but has been worn away to the west. For a North South collider alignment the Silurian may be an alternative where it exits as a thick enough layer above the Maquoketa.

The Maquoketa shales have medium to medium high slake durability and do not swell. It should be stressed that the Maquoketa though a shale has reasonable materials-

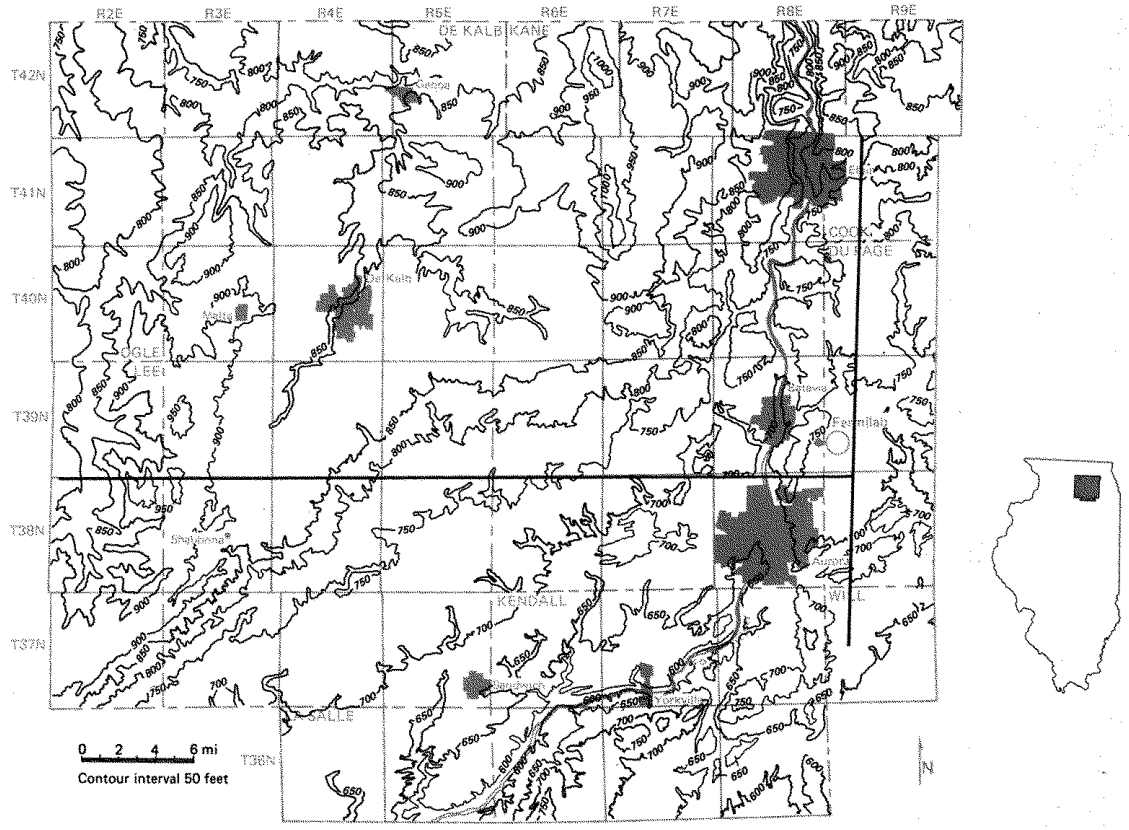


Figure 3.10.12: *Contour Map of the Surface Topography*

construction properties. The dolomite intermixed with the shale provides cement bonding. The unconfined compressive strength is 4400-9000 psi, whereas the Silurian and Galena - Platteville dolomites are 10,000-16,000 psi. This is to be compared with typical concrete 1500-3500 and with the Texas Eagleford Shale of 200-400 psi and Austin Chalk of 2000-3000 psi.

Fig. 3.10.12, Fig. 3.10.13, Fig. 3.10.14 and Fig. 3.10.15 show the topology of drift and bedrock layers in the area.

For tunnels located in the dolomite only spot rockbolting and grouting of cracks would be expected to be required. For tunnel in the shales more rockbolting and some steel straps may be required. These tunnel sections in shale would be covered with shotcrete. Halls if located in shale could be sited with shale at the foot but with the chamber roof in dolomite.

In the Illinois SSC proposal two alternative tunnel elevations were cost estimated. One in the Galena Platteville at an average depth of 430 ft. The other in a combination

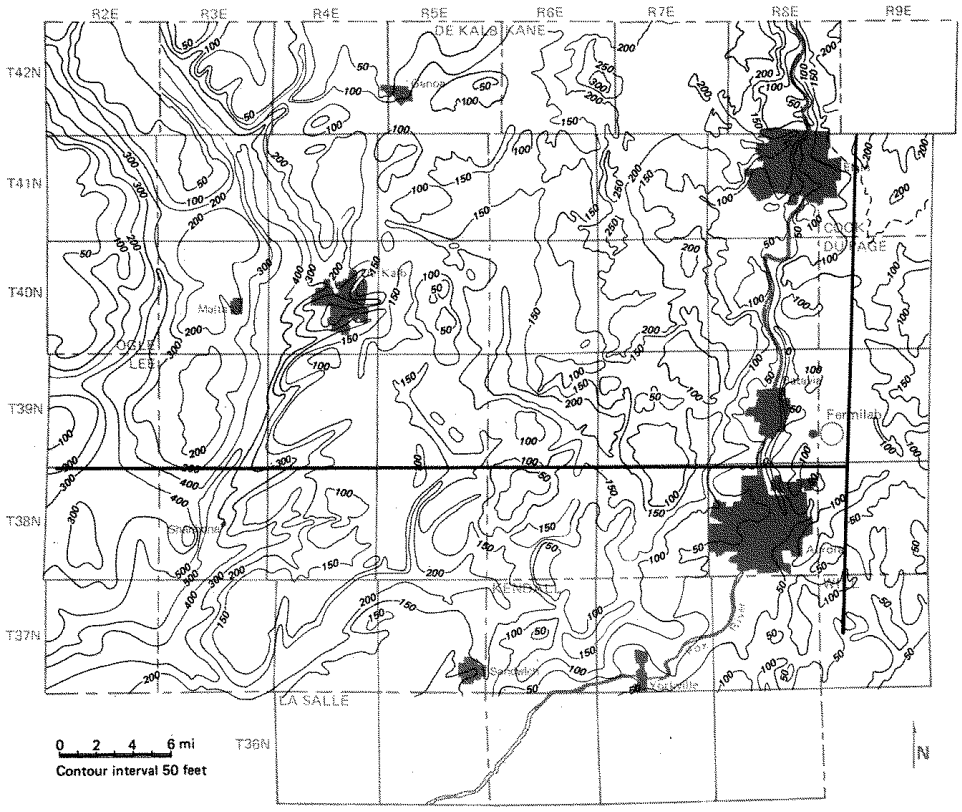


Figure 3.10.13: *Contour Map of the Thickness of the glacial Drift*

of shale and dolomite or limestone (25% shale) at an average depth of 270 ft.

Extensive tunneling has been done in the Chicago area as part of a water retention project (TARP). Over 100 miles of 8 to 32 ft diameter tunnel and over 250 shafts have been bored in the Silurian dolomite rock. Many areas of the TARP tunnels do not require tunnel linings. There are well understood and documented costs for this local tunneling.

The dolomite rock is suitable for cavern spans of 70-120 ft and heights of 130 ft. Tarp Main Pumphouse caverns in Chicago provide comparison with what would be possibly for collider experimental halls. Two of these caverns are 274 ft long, 96 ft high, 63 ft wide and 360 ft below ground.

Joints or cracks in the rock have a preferred 45 deg orientation to the 4 points of the compass-N,S,E,W and are nearly all vertical. Spacing is typically of order 100 ft. The preferred joint orientation implies that halls laid along N-S, E-W axis are more suitable for construction and reduce the need of rock bolts.

The preferred direction for the long dimension of the halls is north - south. This is because the bedrock in the area has high compressive stresses roughly in the east west direction. These stresses can help hold up a roof arch in the east west direction. Analysis has been made for spans up to 125 ft in the east west direction, but would

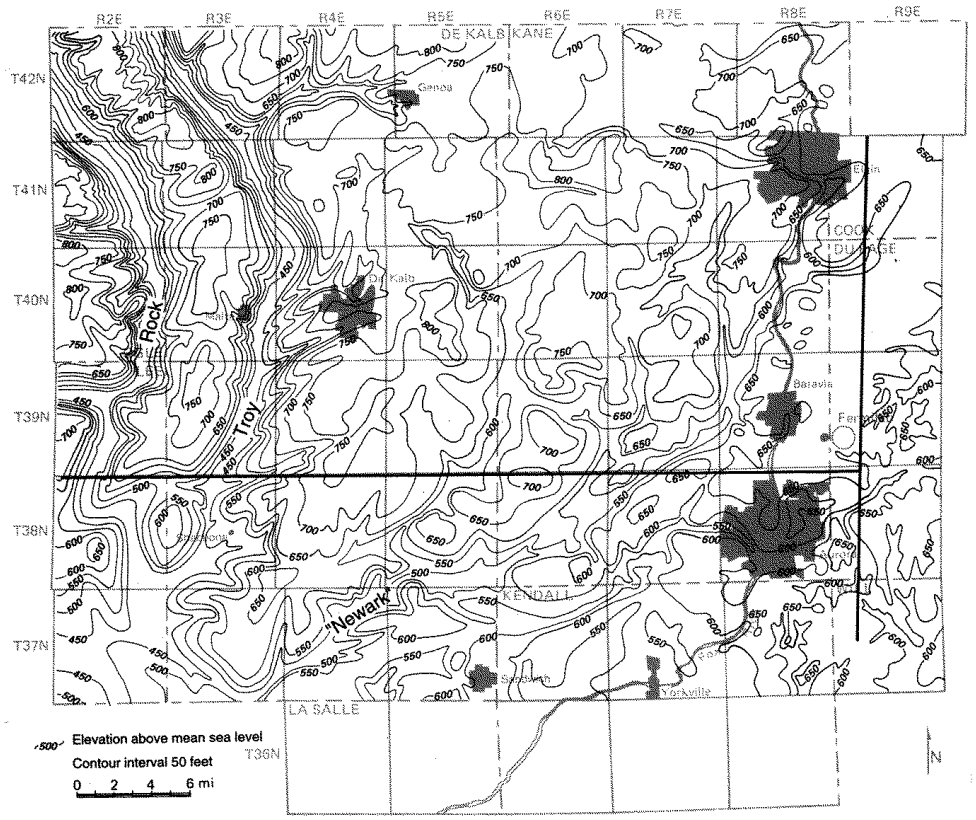


Figure 3.10.14: *Contour Map of the Bedrock Surface showing the Newark and Troy Bedrock Valleys*

need to be reevaluated for north south arch orientation. It is likely this orientation would be viable, but might need more support and have a cost implication.

Water permeabilities (hydraulic conductivity) in the bedrock basically range from 10^{-4} to 10^{-6} cm/sec except at top of Silurian bedrock where aquifers with 10^{-2} cm/sec can be found. The lower Ancell sandstone is an aquifer with $3 \cdot 10^{-3}$ cm/sec.

Initial water inflow rates during tunnel boring of up to 500 gpm/mile drop to 100 gpm/mile without grouting. Final rates of 10-50 gpm/mile can be expected with grouting. (This is lower than what has been measured in the Fermilab tunnels.) The TARP pump station caverns have been measured at only 16-30 gpm.

3.10.11.4 Vibration and Sound Velocity Measurements

Information on vibration and slow drift or diffusion of the tunnel in the bedrock is limited and will need further measurements. There are measurements of the Fermilab tunnel but that tunnel is near the surface in the glacial drift. Measurements were carried out of vibration associated with interstate highway and rail traffic very close to the highway and tracks. These measurements can only represent worst case conditions.

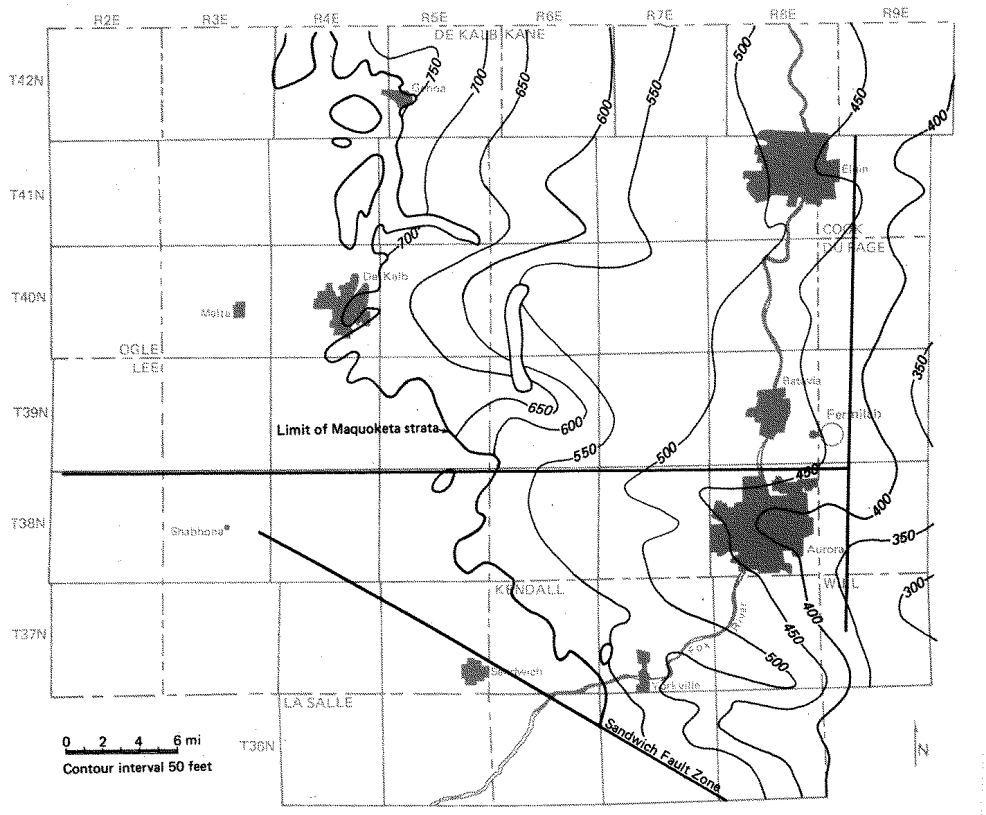


Figure 3.10.15: *Contour Map showing the Elevation of the top of the Galena dolomite*

Additional information on sound velocity in rock has been obtained as part of seismic reflection and refraction survey studies.

Fermilab Tunnel Ground motion vibration measurements have been made in the Fermilab Main Ring tunnel. These measurements were made under noisy conditions with accelerator utility systems in operation, e.g. the cryogenic system and water system. Thus they represent in some sense worst case conditions.

Over a relatively narrow frequency range of 1-25 Hz ground motion power spectra levels $[(2\pi \cdot f)^2 \cdot P(f)]$ of about 0.3 to 0.01 $(\mu m/sec)^2/Hz$ were observed. Also resonances as high as $3 \cdot (\mu m/sec)^2/Hz$ were observed at the Central Helium Liquifier compressor frequency. Higher levels were observed at the ground outside the tunnel location nearest the helium plant.

These levels are similar to those observed at other active accelerator installations [12, 13, 14] (APS, HERA, VEPP) and can be two to three orders of magnitude higher than measurements made under quiet conditions. This only points out that the design of accelerator infrastructure needs to take into account the vibration reduction requirements that are needed. Fig. 3.10.16 and Fig. 3.10.17 show the typical measured spectra.

It should be noted that the measurements carried out in the Fermilab tunnel do

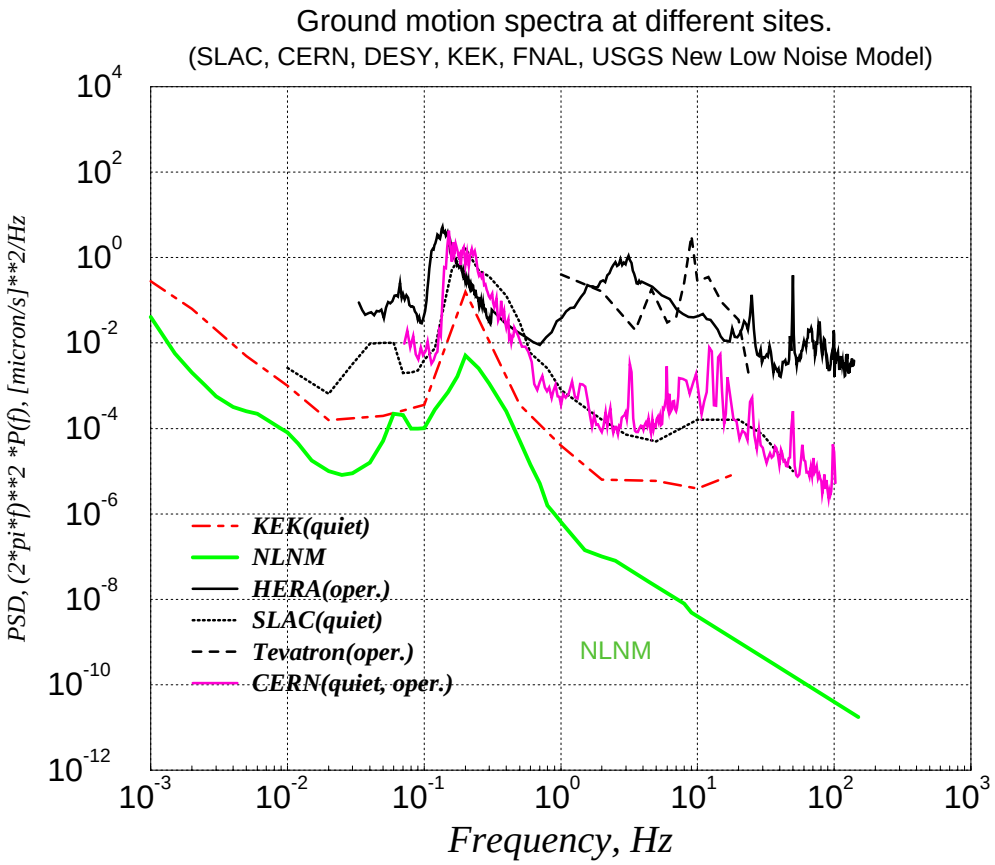


Figure 3.10.16: *Vibration Power spectrum* $(\mu m^2/Hz) \cdot (2\pi \cdot f)^2$

not have direct bearing on vibration spectra which would exist in bedrock tunnels and halls. Further measurements will need to be performed in order to determine these spectra.

Highway and Train Measurements Highway and train vibration measurements have been carried out for the Ill. Geological Survey. These measurements were performed in a rock quarry within 150 ft (horizontal) of an Interstate highway overpass bridge (I88) and 210 ft from rail tracks (and 330 ft from rail road crossing). Vertical distance from the source was about 70 and 50 ft respectively. Maximum zero to peak amplitudes of 2-3 micro inches (0.05-0.07 μm) at about 14Hz were observed from heavy (18 wheel) trucks traveling at speed over the bridge and from the train with 50 cars traveling at 8 miles/hr. At ground surface displacements 60-200 times greater

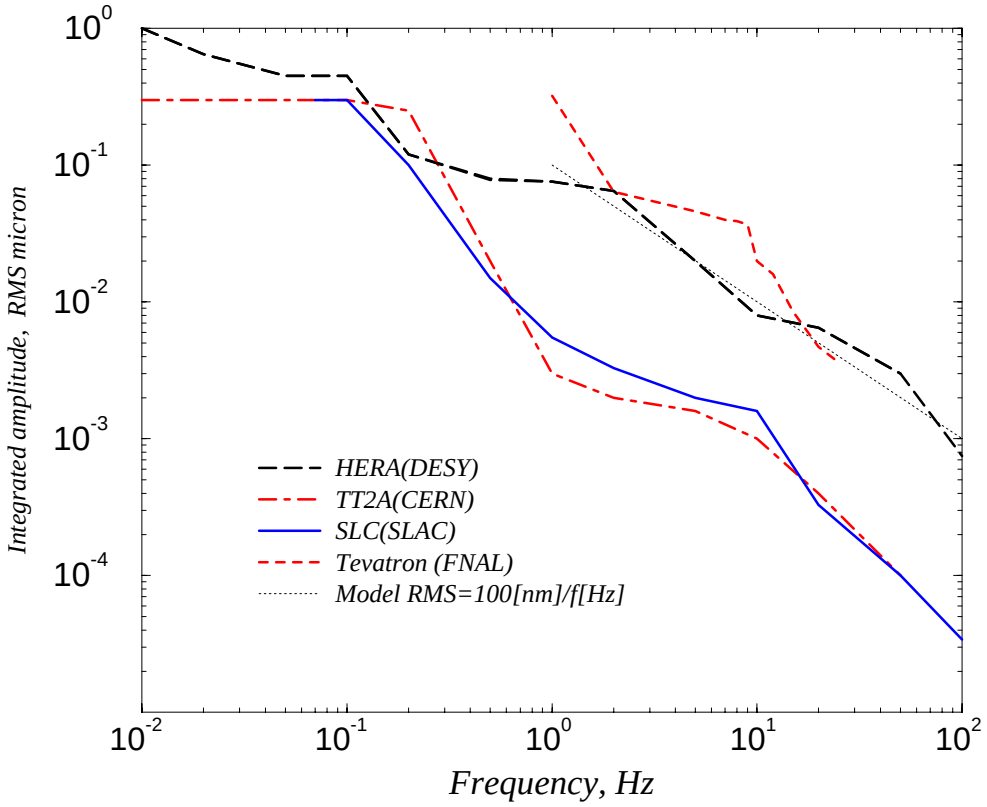


Figure 3.10.17: *Integrated Vibration Amplitude*

were observed. It appears from the measurements of greater vertical displacement that body waves (shear waves) are more dominant than surface waves at the transducer. These observed displacement magnitudes are not unlike those observed at HERA.

Motions of 0.05 μm are about an order of magnitude greater than typical integrated amplitudes from a "noisy site" (100 nm/f[Hz]). They clearly represent worst case and transient conditions. The TESLA tolerances of 100 nm at frequencies above 1 MHz and 1 μm at low frequencies clearly should not be a problem.

Velocity measurements Velocities of the compression and shear waves have been measured in the various rock groups. Typical compression velocities are from 8,700-12,300 ft/sec in the Maquoketa, and 15,700-18,100 ft/sec in the dolomites. Shear velocities are 5,300-7,100 and 8,800- 10,100 ft/sec respectively. Coherence lengths have

not been measured.

3.10.11.5 Hydrology

Water supplies will be required to provide the make up water necessary to cool the (nominal) 180 MW of power used. This amount of power will require about 2,400 gpm (3.5 mgd-million gallons per day) of make up water.

It is likely the most economical water cooling system will be separate systems located at the typically 9 service areas (2 end, 1 IR, 6 cryo areas). Each area would have its own water supply source, and cooling towers or ponds. It can be expected that each area may require from 200 to 400 gpm depending on its specific function.

The northern Illinois area is rich in water resources and there are a number of ways to meet the required make up demands. The choice will be mainly economic. The Ill. Dept. of ENR in a study carried out for the SSC came to the conclusion that generally use of water from wells was most economic especially when both potable and cooling water requirements were included and when the cost of piping transport over any substantial distance was considered. Where available without too great piping distance, municipal sources are viable and economic.

Local wells provide most municipal industrial and private sources in the region. In recent years the western suburbs of Chicago have begun to use water from Lake Michigan. Towns along the Fox River have begun to use the river as a source of city water as the water quality has improved. Even so the four Townships along the Fox River which contain Aurora, Batavia, Geneva and St. Charles use about 23 mgd (million gallons per day) of ground water. Thus the Linear collider usage would represent a small but not negligible increase. It would be distributed over a different and wider area.

Wells Wells in the area are footed in the sand and gravel of the glacial drift, at the top of the dolomite surface of the Silurian, or in the sandstones (St. Peters and Ironton- Galesville) below the dolomite. The Maquoketa and the Galena Platteville are non water bearing.

Wells in the sands and gravels yield typically 100 to 1000 gpm. In the Newark bedrock valley to the west of Fermilab test wells have yielded 1200 gpm and this region is estimated to have a potential yield of 10 mgd.

In the region east of the Fox River where Silurian dolomite is uppermost of the bedrocks, wells yield typically 500 gpm from the top of this layer. Fermilab has one such well that has been tested to 500 gpm and two others that are rated at 200 gpm. Presently one well is used for potable water at about 0.1 mgd.

The St. Peter sandstone, highest of the sandstones under the dolomite yield typically 200 gpm.

The most economical water sources have been estimated by the Ill Dept. ENR to be the Silurian in the east, the St. Peters and Newark Valley sand and gravel to the west. The expense for the development of the necessary water resources is not expected to be a large fraction of either construction or operation of the collider.

Tunnel Inflow Tunnel inflow is estimated to be 10 to 50 gpm/mile after grouting of cracks and joints in the tunnel. This rate is less than experienced in the present Fermilab tunnels. The rate will be dependent on both rock quality and the specific hydraulic head in the area of the specific aquifer. The potentiometric surfaces show considerable variation especially that associated with the deep sandstone aquifers. Thus though rates of 50 gpm/mile would be almost enough to supply half the make up water requirements, their variability would suggest that they could be used only as supplemental water sources.

Rivers The Fox River is the largest in the area with an average flow of 1100 cfs (cu ft per sec, 7.5 gal/cuft, 1 cfs = 450 gpm). The mean annual high flows are about 5000 cfs, and lows 200 cfs. Reliable sole source usage can be based on the 7 day low flow expected in 10 years. For the Fox this is about 150 cfs. In comparison the total estimated collider water need of 2400 gpm amounts to 5 cfs. Fermilab has a present capability of pumping 1000 gpm from the Fox.

Other smaller streams in the area have average flow rates of 20-50 cfs. They would not be suitable for sole source usage but can be used for discharge of tunnel or cooling tower blowdown water if proper environmental standards are met.

Fermi water system The existing Fermilab water system of wells, ponds (and the connection to the Fox River) are good sources for makeup water and drinking water. It could easily supply in excess of 1000 gpm.

Fermilab has an extensive pond system. There are about 230 acres of pond surface on the 6800 acre site. This area is sufficient to cool the total 180 MW at slightly over 1 acre/MW. However the water reservoir or pond depth would not be sufficient for long drought periods without makeup rain water (nominal 82 days) or supplemental well or Fox River sources. Dredging the existing ponds or additional deep retention ponds could be constructed if cooling of more than about 33-40 % of the total nominal 180 MW load were required from rain water makeup alone. 40 % fractional load requires that about 1600 acre-ft/yr. (1.4 mgd, 1000 gpm) of rain water be collected . This represents about 8 % of the site wide annual precipitation (34 in/yr.). As discussed above, additional potential sources of make up water are the Fox River with a maximum capacity to site of 1000 gpm providing the river has sufficient flow, and a number of wells (11), three of which are known to be able to supply 900 gpm total. At present, Fermilab uses wells only for potable water.

Though the combination of Fermilab ponds supplied by rain water, the Fox River, and existing Fermilab wells a large fraction of the total cooling water demands can be met by what presently exists at Fermilab. However to be cost effective it is likely that many of the distributed service areas may be best served by local well, pond or tower cooling systems.

3.10.11.6 Power

The power utility requirements of a Linear Collider (nominally about 180 MW) should present no difficulty for Commonwealth Edison, one of the largest electric generating systems in the country. Peak load in 1995 was 18,600 MW with a reserve margin of 18.6% of the over 22,000 MW installed. Thus a 180 MW load would be less than 1% of the total installed. The 345 kV lines typically have 1400 MVA capacity. Two points of service could provide power from separate grids so that essential loads could be maintained if either line is out of service.

Fermilab has an estimated usage of 400 GWH at present, or an average load of 45 MW. Prior to Tevatron operation power usage of 60 MW during peak operation periods was not uncommon. Line capacity for a new 345 kV line to site is 900 MW.

3.10.11.7 Conclusions

Fermilab is a natural focus point from which to base a future High Energy Physics project. The area in north-east Illinois around Fermilab is well suited as a potential site for a linear collider. Much successful experience in deep tunnel construction exists in the area.

Two possible configurations have been looked at and further analysis and cost comparison of different tunnel locations and depth could be carried out.

The possible ramifications of slight horizontal deviations from a completely linear configuration should be explored so as to understand the possible flexibilities in the collider footprint.

Further measurements of vibration spectra and their attenuation, and coherence length should be performed in the bedrock environment. These measurements should focus in particular on vibrations associated with highways and trains as the most convenient siting may follow these public easements. Vibration may not be a particular issue for the TESLA design, but it may have significant ramifications for other collider designs.

Bibliography

[1] *Proposal to Site the SSC in Illinois*, Submitted by the State of Illinois, Prepared under the direction of the Ill Dept of Energy and Natural Resources, August 1987.

[2] A. M. Graese et al, *Geological Geotechnical Studies for Siting the SSC in Illinois*, Env Geology Notes 123, Dept of ENR, Il. St. Geological Survey, 1988.

[3] Environmental Geology Notes 111, 1985 Il. Dept. of Energy and Natural Resources.

[4] *Geotechnical Summary to the Proposal to Site the SSC in Illinois*, Prepared by Harza Eng Co with assistance of Il. St. Geological Survey, 1988.

[5] *Rock Motion Measurements beneath truck and train traffic*, Wiss, Janney, Elstner, Assoc, WJE No. 870660, for Il St Geological Survey, Oct 1987.

[6] P. C. Heigold, *Seismic Reflection and Refraction Surveying in NE Illinois*, Env Geology 136, Il. St. Geological Survey, 1990.

[7] K. P. Singh et al, *Water Resources of the SSC Site in Northeastern Illinois*, Il. St. Water Survey Div, SWS Contract Report 465, June 1989.

[8] F. E. Dalton et al, *TARP Tunnel Boring Machine Preformance Chicago*, Proceed. Rapid Excavation and Tunneling Conf, SME of Amer Inst of Mining, Metallurgical, and Petroleum Eng, Inc, Littelton, Colorado, 1993.

[9] T. Budd and R. Rautenberg, *Calumet Tunnel Project- A case History*, Rapid Excavation and Tunneling Conf, Society for Mining, Metallurgy, and Exploration, Inc, Littelton, Colorado, 1995.

[10] J. E. Monsees, *Design of SSC Collider Structures*, PB/MK Team, 04279303.DOC, Dallas Texas.

[11] P. H. Gilbert, *Civil Eng. Contributions to the SSC*, ASCE Conv, Dallas, Texas, 1993.

[12] C. D. Moore, *Vibration Analysis of Tevatron Quadrupoles*, Fermilab, Batavia,Il.

[13] V.D. Shiltsev, *Alignment and Stability of Future Accelerators*, Fermilab-Conf-96/131, EPAC Conf, Barcelona, Spain, June 1996.

[14] V. D. Shiltsev, *Pipetron Beam Dynamics with Noise*, FNAL-TM- 1987, Fermilab, Oct, 1996.

[15] Prelim Report on Utilization of the Fermilab Site for a Future Accelerator, TM-1975 (draft), Fermilab 1996.

3.11 Radiation Safety

This section describes the shielding requirements to guarantee radiation safety at the TESLA collider. The most important considerations here are, of course, the impact of the collider operation on the environment, i.e. conceivable radiation levels outside of the accelerator housing in areas accessible to the public and possible activation of soil and ground water.

The beam power in each of the two colliding beams (8 MW) is such that great care must be taken to avoid an uncontrolled loss of the beams. Such loss could immediately lead to a local destruction of the accelerator components and with this to an automatic cessation of operation. For equipment safety alone the monitoring of beam current along the machine is an absolute necessity, such that in case of unexpected significant loss ($> 1\,\%$) pulsing of the accelerator is interrupted instantly. Moreover, radiation monitoring in the accelerator housing will detect losses of much smaller magnitude such that corrective action can be taken in time.

The worst conceivable scenario is the failure of the beam loss interrupt system and the local deposition of a large amount of beam power, but just not large enough to significantly damage the accelerator and thereby terminate operation. Shielding of the produced radiation in such a case must be sufficient to avoid any significant radiation in the outside world.

The case of total beam loss in a small area will lead to immediate destruction of part of the accelerator and is harmless in comparison, as far as the maximum radiation dose is concerned. Particular attemtion must be paid to the beam dumps and the collimators for the spent beam close to the interaction point, because here all the beam power will be deposited during routine operation.

3.11.1 Radiation Levels on the Earth Surface Above the Tunnel

The most penetrating particles above the position of beam loss are high energy neutrons. For a point loss or a loss which occurs over only a few meters along the accelerator, the dose equivalent on the earth surface can be calculated with a simple formula given in [1], a report which also summarizes experimental and theoretical results. The thickness of the earth cover on top of the accelerator tunnel has an average value of 14 m. The smallest coverage will be 8 m. If a local beam loss would be as large as 0.1 % and would remain undetected and uncorrected, the radiation level at the surface would be $5{\cdot}10^{-8}$ Sv/h (0.005 mrem/h or 40 % of the natural background). With 5000 hours of operation per year, such radiation could, if left uncorrected, lead to an annual dose of 250 μSv (17 % of the legally permitted local radiation level). But such a continuous loss as assumed here would be produced by a continuous loss of 8 kW of high energy radiation in the tunnel, a huge signal for the above mentioned radiation monitors. Such a situation would be immediately corrected!

The beam dumps will be positioned below DESY ground and have an earth coverage of at least 11 m (each additional thickness of 1.3 m reduces the rate from high energy

neutrons by a factor of 10). With such shielding thickness the dose rate at the the surface can be made smaller than the natural background.

Another dose component at the earth surface is due to muons produced by beam losses along the accelerator. Muons are only created within a small cone in the beam direction, but they have a very large range. The doses were estimated by means of data given in [2]. First calculations of the conceivable dose rates show them to be small compared to the natural background. Muons from the beam dumps can be avoided at the surface by bending the beams downwards before they enter the dumps.

3.11.2 Activation of the Main Beam Dump

It was shown in section 3.7.9 that a dump made essentially of graphite and backed by aluminium is capable of handling an 8 MW beam. Such a dump has the additional advantage that the number of produced neutrons and its residual radioactivity are smaller than with any other material. In the following the activity of a pure carbon dump is estimated together with the dose rate near it after a long period of operation.

We assume an electromagnetic cascade completely developed in carbon, then the yield Y of a nuclide per incoming primary electron (energy E_0) is

$$Y(E_0) = \frac{L\rho}{A} \int_0^{E_0} \sigma(k) l\prime(k) dk \tag{3.11.1}$$

Here $\sigma(k)$ is the cross section of the nuclear reaction as a function of photon energy, $l\prime(k)$ the differential track length, and L, ρ, and Δ are Avogadro's number, density and atomic weight of the material. Approximation of cascade theory is especially valid at high energies, therefore we can write

$$l\prime(k) \approx 0.572 \frac{X_0 E_0}{k^2} \tag{3.11.2}$$

where X_0 is the radiation length.

For carbon only three nuclides have to be considered: ^{11}C (Half life 20 min), ^{7}Be (54 d), and ^{3}H (12.3 a). Their saturation activity in Bq is calculated which equals the number of produced nuclei per second. A full power operation of 8 MW is assumed to produce a saturation activity of ^{11}C, a mean power of 2 MW for the production of ^{7}Be and a mean power of 0.5 MW averaged over 25 years for calculating this magnitude for ^{3}H. The photon cross sections are mainly taken from [3, 4, 5] ; a small tail is added to the cross section curves to cover the photon energy range up to 200 MeV. The resulting saturation activities are given in Table 3.11.1.

A tritium activity of $7 \cdot 10^{12}$ Bq or 0.02 g produced in 25 years of operation is easily absorbed by the graphite (together with 1 g of stable hydrogen produced by nuclear processes), its absorption capacity is about $0.3 N cm^3/g$ at room temperature. If part of it is exhaled due to elevated temperatures during operation it will be removed by the tunnel ventilation and presents no radiological hazard. The dose rate due to the residual activities of Table 3.11.1 is calculated taking roughly into account the extension of the source and the shielding effect of the surrounding aluminium. Immediately after

Nuclide	Saturation Activity [Bq]
^{11}C	$5 \cdot 10^{15}$
^{7}Be	$7 \cdot 10^{13}$
^{3}H	$7 \cdot 10^{12}$

Table 3.11.1: *Saturation activities in the Carbon beam dump.*

stopping operation the ^{11}C gives a very high dose rate of 25 Sv/h at a distance of 1 m. After a cooling period of 3 h the dose rate is determined by the ^{7}Be which amounts to 16 mSv/h and decreases with the half life of 54 d. After a very long cooling period the dose rate is determined at a much lower level by the ^{22}Na activity produced in the backing aluminium (half life 2 a). The dose rate can be reduced by lead shields if necessary, depending on the position of the dump.

3.11.3 Activations Outside the Tunnel

The carbon/aluminium dumps described in this report are designed to absorb the total beam power of the two beams. In actual collider operation a significant amount (about 10 %) has to be absorbed by collimators in the spent beam capture system for the positron source. These locations will be under DESY land and will have special shielding added. But except for these deliberate scraping losses at these collimators all beam power is dissipated in the dumps. Energetic neutrons produced in these dumps might penetrate the dump, lead shields and the concrete walls of the structures which house the dumps. These neutrons can then in turn activate soil and ground water. The resulting saturation activations under the assumption of no additional shielding have been calulated (see [6]. The results can be summarized as follows.

Saturation activity is essentially given by the product of neutron fluence and cross section, integrated over the neutron spectrum. The neutron fluence spectrum generated in a thick target is well known by measurements and calculations for primary protons. For high energy primary electrons no experimental results are available. Therefore we studied the shape of the neutron spectrum by preliminary calculations using the MC code FLUKA 92 for a thick target and as a function of concrete side shieldings. It turned out that the spectrum shape is independent of the shield for concrete thickness larger than 60 cm and that it is very similar to the neutron spectrum produced by primary protons. This can be expected. Neutrons in the relevant range up to 200 MeV are produced by intranuclear cascades (initiated by particles below 0.5 - 1 GeV) and by low energy nuclear reactions in the thick target and in the concrete or sand shield, rather independent of the high energy process. The spectrum was normalized in such a way that its total dose equivalent equals the dose of electron- produced neutrons given in [1] (which in turn is based on measurements and calculations).

When one considers nuclear reactions in the soil and ground water initiated by fast neutrons from an *unshielded* beam dump, radioactive nuclides with very short life time are of less interest, because of the time necessary to come in possible contact with people. Nuclides with very long life times are also of less interest, because of the

low dose rates associated with long life times. The cross sections for the producing reactions of the interesting nuclides are not all accurately known. But since neutron induced reactions are known to be similar to proton induced reactions, the cross sections of the latter can be reasonably substituted.

The resulting saturation activity, assuming an unrealistic continuous 8 MW operation throughout the year, in the soil near the beam dump is, ^{7}Be: 6400 Bq/g, ^{22}Na: 12000 Bq/g, ^{55}Fe: 3000 Bq/g, ^{45}Ca: 300 Bq/g and ^{54}Mn: 150 Bq/g. Apparently these numbers present no radiological problem: The nuclides are permanently attached to solid soil and are neither accessible nor movable. With a half life for all of them smaller than 3 years, 40 years after final shutdown of the accelerator all remaining activities will be smaller than the natural activity of sand.

Accelerator-produced nuclides could also be expected in the ground water due to spallation of oxygen and leaching of nuclides produced in the soil. The somewhat complex situation is described and analyzed in [6]. As a result, only ^{22}Na and ^{3}H might be found in ground water. This has been confirmed by experiment [7]. If we assume that the slowly moving ground water will be close the beam dump for 100 days, the calculated activity concentration will be 320 Bq/g and 1400 Bq/g, respectively. These numbers would be of concern only, if the drinking water for people is taken without dilution from a well in the immediate vicinity of the beam dump (they would be 5000 times as large as legally permitted for general consumption). This is an extremely unlikely scenario, since the beam dumps are under DESY land and there will be no wells in their immediate vicinity. A large dilution factor could therefore be reasonably assumed in practice. But it would also be very simple to add concrete shielding around the beam dump and reduce the water activity concentration there by a factor of 5000: One meter of heavy concrete gives a reduction factor of 30 for fast neutrons. Therefore 2.5 m of added shielding would already suffice.

Bibliography

[1] K. Tesch, Rad. Prot. Dosimetry 22 (1988) 27.

[2] E. Bräuer, K. Tesch, Internal Report DESY D3-62 (1987) (derived from W. R. Nelson, Nucl. Instr. Meth. 66 (1968) 293).

[3] B. C. Cook et al., Phys. Rev. 143 (1966) 724.

[4] H. Artus, Zeitschrift f. Physik 189 (1966) 355 V. di Napoli et al., Notas de Fisicas XIII (1967) No. 3 (Centro Brasil. Pesquisas Fisicas, Rio de Janeiro).

[5] V. V. Balashov et al., Nucl. Phys. 27 (1061) 337.

[6] K. Tesch, Internal Report DESY D3-86 (1997).

[7] S. Baker et al., SSC Dallas, SSCL-Preprint 538, 1994.