ACCELERATOR STRUCTURES FOR LINEAR COLLIDERS*

R. H. Miller, SLAC, Stanford, CA94309

Abstract

Linear Colliders require high gradient (to reduce length dependent costs), high shunt impedance (to reduce power dependent costs) accelerator structures in which the long range dipole wakefields have been reduced by 2 to 3 orders of magnitude. The precise dipole wake reduction factor required depends on many factors. These include the beam intensity and time structure, the accelerator aperture and accelerating gradient, the strength of the focussing, the alignment precision and position stability of accelerator structures and the focussing elements, the effectiveness of the tuning procedure and feedback system, the pulse to pulse stability of the injected beam, and the required emittance at the final focus. For the purposes of this paper we accept that large reduction factors are required and discuss various approaches to achieving them. There are basically two approaches: detuning and damping, which are often used in combination. Damping can be accomplished either by introducing loss selectively into the accelerating structure cells, or by coupling the dipole modes out of the accelerating region and absorbing them in external loads or lossy materials. Detuning can greatly reduce the amount of damping required by causing the dipole modes to decohere, and it is possible to achieve destructive interference so that the wakefields from different cells in the accelerator cancel each other as seen by the bunches of electrons travelling through the linac. Several different approaches have achieved large reduction in the long-range dipole wakefields so that they pose no restriction on the length of the bunch trains which can be used for linear colliders.

1 INTRODUCTION

There are two primary issues for accelerator structures for linear colliders: achieving high gradients and controlling dipole wakefields which destroy the beam brightness and hence the luminosity of the collider. A decade ago there was little concern that the high frequency structures could operate at their design gradients (but rf power sources were a serious concern), and a lot of concern about controlling the degradation of beam quality by dipole wakefields. Today, the reverse is the case: accelerator physicists are fairly comfortable that they can control the effects of dipole wakes, but very worried about physical degradation of the accelerator structures when operated at high gradients

In this paper I will say more about the successful wakefield mitigation techniques, but will engage in a little speculation about structure damage from arcs and pulse heating. I will consider only room temperature structures, not because super conducting structures are not serious contenders for future linear colliders, but because others are much more able to discuss the important issues for super conducting structures.

2 HIGH GRADIENT ISSUES

2.1 RF Breakdown

There are a number reasons for the reversal of concern from wakefields to structure damage. 1) The success of the strategies for mitigating long range wakefield effects has led to collider design parameters with longer and longer beam and rf pulse length which exacerbates the problems of structure damage from rf breakdown and from pulse heating. 2) High gradient breakdown tests in resonant single cavities and short structures by J.W. Wang and others suggested that X-Band structures could be operated at accelerating gradients approaching 200 MeV/m, and that the breakdown threshold varied approximately as the square root of the frequency. 3) Because limited rf power, early tests with travelling wave structures were done with short, low-group- velocity structures, and indicated that an accelerating gradient of 100 MeV/m at X-band was comfortable. 4) The early tests did not look quantitatively at the damage to the structures, i.e. did not look at the change in phase advance per period which is critical for a long accelerator structure. 5) Last and most important is that tests at both SLAC [1] and CERN [2] have indicated significant damage to structures running near or below the design gradient. Using resonant or short, low-group-velocity structures for high gradient tests tends to reduce the damage done to the structure by arcs by reducing the energy that gets deposited at the arc site.

At SLAC tests were done on structures which were considered similar to the final design in the relevant features (same range of group velocity, same range of iris size and thickness, same ratio of peak surface to average accelerating field). These tests showed significant changes in phase velocity after relatively short periods of running at or below the design gradient with shorter pulse lengths than the design. This was what might be called a startling "reality check". The damage was concentrated in the upstream third of the structures. The NLC structures are approximately constant gradient as a result of the Gaussian detuning we use to suppress wakefields. This fact combined with the successful early tests on short, low-group-velocity structures suggests the first reason for the damage being concentrated at the input end. Low group velocity is good[1] for two reasons. 1) The power and therefore the energy in the rf pulse required to achieve a given gradient varies linearly with group velocity. 2) The reflection coefficient created by an arc of fixed resistance varies inversely with group velocity. For these 2 reasons the energy deposited at the site of the arc should vary approximately as the square of the group velocity. I think

* Work supported by the Department of Energy, contract DE-AC03-76SFO0515

there are two other reasons the damage was concentrated at the input end. 1) When an arc occurs it causes a mismatch which can double the fields upstream. This must cause arcs upstream to occur almost instantly. The reflection from an arc propagates with the group velocity, so the arcs may propagate upstream almost with the group velocity depending how long it takes an arc to occur when the fields double. 2) There is a region in the first third of the SLAC structure where the surface fields are 5% higher than the downstream half of the structure. KEK and SLAC have embarked on an urgent program to measure damage as a function of gradient in a number of constant gradient structures of various lengths and initial group velocities. I anticipate that we will find that the probability per pulse of an arc varies roughly linearly with the length of the structure and therefore that the damage will increase with increasing length and will be concentrated toward the input end. Furthermore, we will probably find that lower group velocity also reduces damage. If true this will suggest going to shorter lower group velocity structure which will cause a moderate increase in the linac cost due to increased power distribution costs and the increase in the number of couplers. The shorter structures naturally go with lower group velocities, since we want the attenuation through the structure to be in the range between 0.5 and 0.6 nepers for reasonable efficiency and to keep the beam loading derivative small.

If you scale the cell dimensions of a structure linearly with wavelength, you find that to keep the attenuation through the structure constant, you need to scale the number of cells as $\lambda^{1/2}$. If you use this scaling and scale the pulse length with the filling time ($\lambda^{3/2}$), you find that the field at which an arc should cause a given phase shift scales as $\omega^{1/2}$. However, because the small apertures exacerbate the dipole wake problems and because couplers and high power RF distribution waveguides are a significant part of the cost of a linac we chose a structure length almost 5 times as long as the value scaled from the SLAC S-band linac. This choice means that the energy in the RF pulse is more than five times larger. In addition the group velocity is, of course, five times larger, so an arc can vaporize 25 times more copper than if we had kept the average group velocity constant and scaled the length with the filling time. Evidently we were much less conservative than we thought we were being.

We almost certainly will go to shorter structures with lower group velocities, but we don't yet know how far we need to go or what method we will use to reduce the group velocity. We don't want to reduce the iris diameters because that increases the short range dipole wakes. We may add magnetic coupling slots to cancel part of the electric coupling, increase the phase advance per cell, go to thicker disks, or use a combination of these. We probably will even take another look at standing wave structures.

2.2 Pulse Heating

Measurements by Pritzkau[3] indicate that pulse heating of 120^o C. for 250 hours at 60 PPS degrades the copper surface of a cavity. This is comfortable for the current generation of proposed collider linacs except for regions close to coupling holes. It does however place important restrictions on the coupling of higher order modes. If the holes for coupling the higher order modes out of the structure lower the Q of the accelerator mode by much more than 1%, there will probably be a problem with pulse heating in the vicinity of the coupling irises. I believe the Shintake Choke-Mode structure is an exception to this rule, because the dipole modes couple out through an axisymmetric slot which does not concentrate the surface currents. The threshold gradient for pulse heating problems scales as $\omega^{1/8}$ for linearly scaled structures with pulse length scaled as $\lambda^{3/2}$ (i.e. constant attenuation) and the thermal diffusion length therefore scaled as $\lambda^{3/4}$. The surface resistivity scales as $\omega^{1/2}$.

3 DIPOLE WAKEFIELDS

3.1 Resonant suppression

If the resonant or synchronous frequency of the dipole mode is arranged to be

$$d_{ipole} = (n/2))f_{bunch},$$

where f_{bunch} is the bunching frequency then all the bunches will fall at the zero crossings of the dipole deflecting force. Therefore none of the bunches will be deflected by the dipole fields. This approach works best for n odd so that each bunch tends to cancel the field left by the previous bunch.

A similar approach is to use a rectangular distribution of dipole mode frequency. In this case the wakefield, W(t) is a sinc function and the bunches can be arranged to come at the nodes of the sinc function by choosing the width of the rectangular distribution to be an integer multiple of the bunch frequency.

3.2 Gaussian Density Distribution

Both NLC and CLIC have chosen to use a Gaussian density distribution of the dipole impedance in the frequency domain. This was chosen because the wakefield, which is the Fourier transform of the spectral function is a Gaussian function of time which falls off rapidly. The Gaussian has several advantages.

1) It is non-resonant and therefore does not freeze collider operation to any particular bunch spacing other than some minimum spacing.
2) The wakefield decreases rapidly and monotonically for about 2 orders of magnitude.
3) It permits error function interpolation of parameters with very sparse calculated date.

The Gaussian density detuning has the disadvantage that it is not limited, and so must be truncated. This causes a sinc function like wake to stop the rapid Gaussian fall at a level dependent on the truncation point. Relying on detuning alone has the disadvantage that there are a finite number of discrete modes. This causes a partial recoherence of the wakefield starting at a time $t \approx 1/\delta f_{max}$ where δf_{max} is the widest frequency separation between adjacent modes. With no damping the wakefield then increases until a time $t \approx$

$1/\delta f_{min}$ where δf_{min} is the minimum spacing between modes at the highest density point of the distribution. The natural solution to this problem is to combine detuning with light to moderate damping.

3.3 Sinc^4 ($\delta\omega t$) Wakefield

Jones[4] has pointed out that if you convolve two rectangular frequency distributions you get a triangle. If you convolve two triangular distributions you get a bell-shaped curve which looks somewhat Gaussian, but which is limited and the function and its first derivative are both continuous at the point where it reaches zero. Its Fourier transform and therefore the wake is $W(t) = \text{sinc}^4 (\delta\omega t)$, where $\delta\omega$ is the width of the bell shaped curve. As a continuous function the Fourier transform of this function is clearly superior to the transform of a truncated Gaussian, but when discrete modes are used its superiority is less striking, but may be significant.

3.4 Manifold Damping

Figure 1: Two cells of NLC Structure RDDS1.

Manifold damping as implemented for the NLC structure consists of 4 TE11 round waveguides positioned at 90^0 azimuthal intervals around the accelerator structure, Fig.1, two in the horizontal plane which couple to the vertically deflecting dipole modes, and 2 in the vertical plane coupling to the horizontally deflecting modes. In the absence of coupling to the accelerator cells, these damping manifolds have dispersion curves which are hyperbolas asymptotically approaching the velocity of light line as the frequency goes to infinity. Hence the phase velocity is always greater than the velocity of light. Each damping manifold is coupled to each accelerator cell by electric field coupling through a longitudinal slot which connects the manifold to the accelerator cells and runs almost the full length of the accelerator structure, stopping a few cells short of each end to avoid mechanical interference problems. Coupling between 2 waveguides joined by a slot (or many closely spaced holes) depends on both waveguides having the same phase velocity. However the beam excites a wave with a phase velocity equal to c, while the damping manifold is a waveguide with a phase velocity greater than c. The coupling occurs because the accelerator structure is detuned, i.e. tapered. In the NLC accelerator structure each frequency with significant dipole impedance propagates in the tapered structure until it runs into a stop band at π mode at one end and 0 mode at the other end. The phase velocity of the dipole mode is equal to c at about $5\pi/6$ phase advance per period near the input end of the structure, and close to π mode near the output end. Thus the phase velocity varies in each dipole mode in the structure from about c near the π mode end to infinite at the 0-mode end. The phase velocity is equal to the phase velocity of the damping manifold somewhere in between and this is the location of the avoided crossing where the coupling of dipole fields in the accelerator to the damping manifolds occurs. As a velocity of light bunch moves through an accelerator structure in each region it excites a dipole field at a frequency which has phase velocity equal to c in that region.

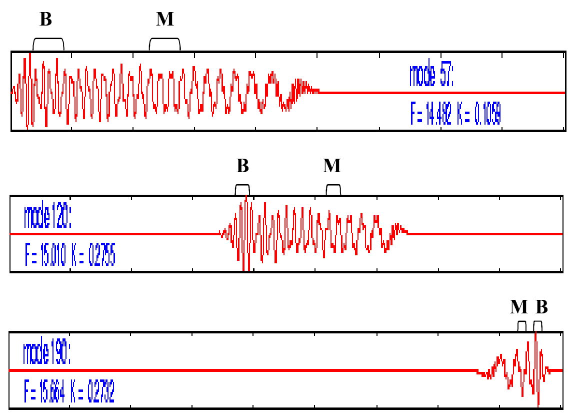

Figure 2: Modes 57, 120, and 190 of Detuned Structure showing beam coupling region (**B**), and manifold coupling region (**M**).

The energy then propagates in the direction of the group velocity for the dipole wave in that region (downstream if a forward wave, upstream if it's a backward wave) until it reaches the region where its phase velocity is synchronous with the manifolds where it couples into the manifold (see Fig. 2). The damping is optimized when the coupling between each cell and the manifold is adjusted so that the coupling region for a particular dipole frequency is a 100% coupler of power from the accelerator cells to the damping manifolds. In this condition what had been a resonant dipole mode becomes a pure travelling wave if the manifold is perfectly terminated, since all the power is coupled into the manifold on the first pass. Thus, no resonant build up occurs and the resonant structure in the Gaussian density distribution should disappear. If the cell to manifold coupling is either increased or decreased from this optimum value, the effect is almost the same. If you reduce the coupling you won't reach 100% coupling, if you increase it you reach 100% coupling and then begin to couple some of the wave from the manifold back into the accelerator cells. The only difference is a π phase shift in the wave in the accelerator, which lets the designer know when he has passed through the optimum 100% directional coupler value.

While it is possible in principal to achieve an optimum coupling which gives a smooth spectral function, a surprisingly small mismatch (VSWR ≈ 1.05) at the output end of the manifolds significantly degrades the long range wakefield suppression. Such a small mismatch doubles the size of the ripples in the middle region of the spectral function for RDDS1 and doubles the average value of the wakefield in the region from about 10 to 30 meters behind the driving bunch. This effect is understandable when one considers that when the coupling is adjusted to couple all the dipole energy out of the structure, it also couples all of the wave reflected from the manifold loads back into the accelerator cells. Thus, new resonances are created. One solution for this problem is to design very good terminations for the manifolds. An alternative is to introduce significant distributed loss (perhaps 10 dB) into the downstream end of the manifolds which would reduce the reflection coefficient by an order of magnitude.

While with ideal cell to manifold coupling and perfect manifold terminations the long range dipole impedance can be a smooth function of frequency there is still a characteristic decay time (a function of frequency) during which the dipole fields radiated from successive bunches can build up. These time constants are the result of the travelling wave filling (or unfilling) time of the velocity of light region for each frequency. These time constants are of the order of 20 to 40 ns in the NLC structure; longer than we would like them to be. These time constants vary inversely with the square root of the dipole group velocity and the dipole taper derivative, dω/dz. Thus, they depend on parameters like the iris apertures of the accelerator structure and the detuning function chosen, which have other constraints on them. This is true for any manifold damping design where the phase velocities in the manifolds for the dipole frequencies of interest are significantly greater than the velocity of light. For this case the dipole energy at any frequency must propagate longitudinally out the region where it interacts with the beam in order to reach a region where it couples into the manifold. The solution is to couple out of the accelerator cells transversely, i.e. couple the dipole mode energy which the beam deposits in each cell directly out into loads as is done in the CERN TDS structure discussed below. Another possibility is velocity of light manifolds that would couple dipole energy from the structure in the same cells where the beam deposited it. The most obvious way to make velocity of light manifolds is to make them coaxial lines. The principal problem with this is that the manifolds would propagate the frequency of the accelerator fundamental mode. The field configuration of a coax would make it difficult not to couple the fundamental. Perhaps the only way to avoid coupling the accelerator fundamental into the coax would be to couple each cell of the accelerator to each manifold through a short section of waveguide which is cut off for the fundamental frequency of the accelerator.

The damping manifolds serve functions in addition to damping. First, they act as vacuum manifolds in parallel with the accelerator structure, increasing the vacuum conductance by about a factor of four. Secondly, they act as position monitors for the structures to verify that the beam is going through the center of each structure. The synchronous frequency of the structure varies smoothly and monotonically from about 14.3 GHz at the input end to about 16.0 GHz at the output end. Thus, measuring the spectrum of the power from the manifolds generated by a single bunch reveals the straightness of the structure and whether its axis is parallel with the beam axis. If diurnal temperature variations or tidal or long term earth movement or any other stresses cause the structures to move, bend, or squirm, a measurement of the frequency spectrum from the manifold can reveal it with a resolution of a few micrometers.

Finally, the equivalent circuit analysis is not well suited for looking at higher order dipole modes (above 17 GHz)[5]. The manifold damping was designed for the lowest band and probably fails to damp the higher bands. ASSET measurements indicate the presence of these higher bands at levels comparable to the damped detuned lowest band. The combination of the lowest band with the higher bands is a difficult beam dynamics problem and requires further study.

3.5 The CERN Tapered Damped Structure

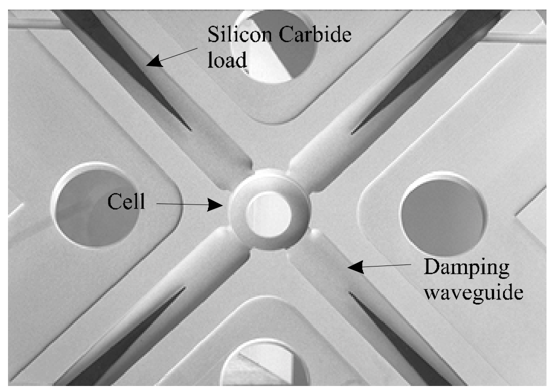

Figure 3: Tapered Damped Structure cell, with the four damping waveguides and silicon carbide loads.

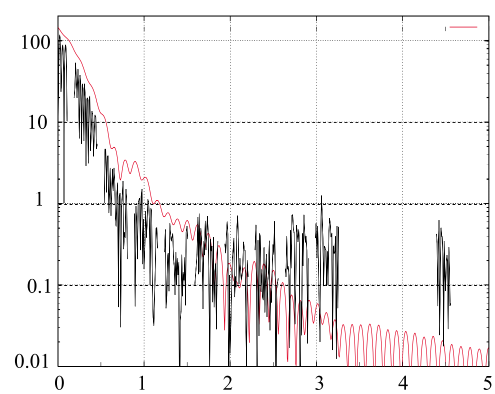

Figure 4: Plot of the measured and computed (smooth) transverse wakefields, V/(pC*mm) vs. nanosec.

The CLIC study group has chosen to combine detuning with much heavier damping than that used for the NLC structure. In their present design the dipole mode Q is about 20, so that the dipole fields damp by a factor of about 200 per nanosecond, but the detuning also makes a not negligible contribution to the initial fall off of the fields. This low Q is accomplished by magnetically coupling (using azimuthal slots) each cell to four rectangular waveguides running radially outward and each terminated in a tapered, lossy ceramic load, Fig. 3. The suppression of the long-range dipole modes in this structure is excellent as shown in the ASSET[6] test results[7] shown in Fig. 4. The present design has two closely related problems: 1) the coupling slots lower the Q and consequently the shunt impedance per unit length by about 20%; 2) The pulse heating of the structure around the coupling slots is too high. The second problem is a direct result of the first. If the damping slots lower the Q by 20%, it means that 20% of the power dissipated by the fundamental mode is dissipated in a tiny region around the damping irises. For a structure intended to be run at high gradients, this is sure to be a problem. The 20% loss in shunt impedance will require 10 to 15% increase in the power to the structure to achieve the same beam loaded energy gain. The increase is less than the 20% loss in shunt impedance because while the copper losses increase by 20%, the power delivered to the beam remains constant and the power delivered to the output loads actually decreases.

3.6 The Shintake Choke-Mode Structure

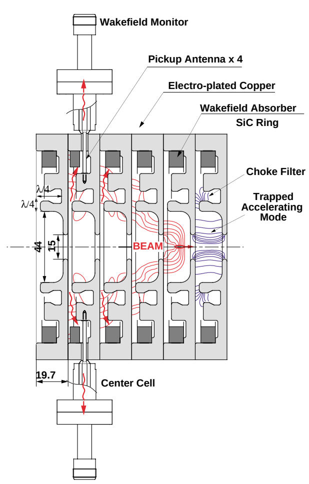

Figure 5: Shintake Choke-Mode Structure.

Shintake[8] has designed a C-band heavily damped structure, Fig. 5, which relies solely on damping for mitigation of the long-range dipole modes. It uses a choke joint to confine the accelerating mode, and thus lets the dipole modes which have higher frequencies radiate outwards where they are absorbed in lossy ceramic. Like the CERN structure, this structure performed well in the ASSET test, as seen in Fig. 6.

Also like the CERN structure, the damping loses about 20 or 25% in shunt impedance, but in this case I believe the loss does not create a pulse heating problem. The reason is that the loss in shunt impedance is primarily because of lower r/Q rather than lower Q. The Q is also lowered somewhat by the large surface to volume ratio of the choke joint rather than by raised surface current densities.

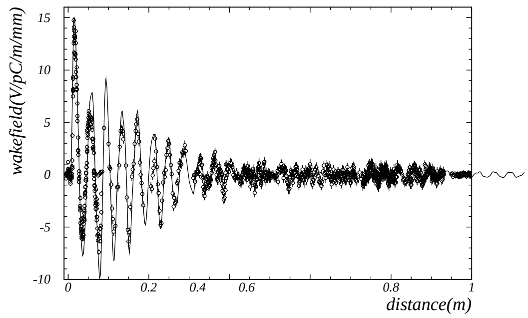

Figure 6: ASSET measurement (dots) and theory (curve).

3.7 The DESY S-band Collider

Although this is no longer an active project, it is still of interest because the solutions found there may be applicable to other projects. This linac mitigated the dipole wakefields by combining linear detuning of the dipole modes with internal damping. Plating the inner edge of the disks with lossy steel selectively damped the dipole mode. Because this was a region of high surface currents for the dipole mode but low surface currents for the monopole mode it was effective for selectively damping the dipole mode. It lowered the Q of the dipole to about 2000, but only lowered the monopole mode Q by a few percent.

4 CONCLUSIONS

There are a number of techniques for dealing with long range dipole wakefields. One can confidently say that dipole wakefields do not prevent linear colliders from being designed with long bunch trains required for high luminosity and rather high rf to beam efficiency. However, it is clear that both rf breakdown and pulse heating need to be studied carefully to achieve the design gradients in the proposed high frequency linear colliders.

ACKNOWLEDGEMENTS

As I hope is evident, this paper is a mixture of information and ideas gathered from the worldwide community of physicists and engineers working on room temperature accelerator structures for linear colliders and opinions and conclusions for which I am responsible.

REFERENCES

[1] C. Adolphsen, et al, TUE01, this conference.
[2] I. Wilson, W Wuensch, TUA17, this conference.
[3] D Pritzkau, R. Siemann, TUA18, this conference.
[4] R.M. Jones, et al, TUA10, this conference.
[5] C. Adolphsen, et al, TUA18, this conference.
[6] C. Adolphsen, et al, p.481, Proc. LINAC94.
[7] I. Wilson, et al, EPAC99.
[8] T. Shintake, et al, pp. 3411-3413, Proc. PAC99.