VERTEX DETECTORS

Vera Lüth
Superconducting Super Collider Laboratory, Dallas, Texas*

ABSTRACT

Different methods and technologies for vertex detection will be introduced, starting from nuclear emulsions and bubble chambers, to drift chambers, scintillating fibers and solid state devices.

TABLE OF CONTENTS

*On leave from the Stanford Linear Accelerator Center, Stanford, California.

© V. Lüth 1993

1. INTRODUCTION

The purpose of a vertex detector is to measure position and angles of charged particle tracks to sufficient precision so as to be able to separate tracks originating from decay vertices from those produced at the interaction vertex. Such measurements are interesting because they permit the detection of weakly decaying particles with lifetimes down to 10^{-13} s, among them the τ lepton and charm and beauty hadrons.

These two lectures intend to introduce the reader to the different techniques for the detection of secondary vertices that have been developed over the past decades. The first lecture includes a brief introduction to the methods used to detect secondary vertices and to estimate particle lifetimes. It describes the traditional technologies, based on photographic recording in emulsions and on film of bubble chambers, and introduces fast electronic registration of signals derived from scintillating fibers, drift chambers and gaseous micro-strip chambers.

The second lecture is devoted to solid state detectors. It begins with a brief introduction into semiconductor devices, and then describes the application of large arrays of strip and pixel diodes for charged particle tracking. These lectures can only serve as an introduction to the topic of vertex detectors. Time and space do not allow for an in-depth coverage of many of the interesting aspects of vertex detector design and operation. More complete reviews of the older techniques and lifetime measurements can be found in scientific journals [ref 1.1]. An excellent introduction to silicon vertex detectors is presented in the lecture notes by C. Damerell [ref 1.2].

Since the discovery of the J/Ψ particle almost two decades ago, the study of hadrons carrying heavy flavor quantum numbers has become a fashionable topic, for theorists and experimentalists alike.

In the framework of the standard electro-weak theory, flavor changing transitions among quarks and leptons are described by their coupling to the charged weak boson, W±. The parton diagrams for heavy flavor decays are identical to μ decay, and thus apart from the differences in the coupling of the W± to the various partons, the semileptonic decay rates of heavy flavor quarks Q are calculable and closely related to the muon decay rate. The differences in the couplings of heavy quarks to the W± are given by the Cabibbo-Kobayashi Maskawa (CKM) matrix elements V_{qQ}. Precise measurements of decay rates of heavy flavor mesons and baryons, combined with studies of their branching ratios and decay spectra, can therefore be used to determine the elements of the CKM mixing matrix, and to test and formulate theoretical models.

Heavy quarks also couple to the flavor conserving neutral weak current. The study of Z^0 decay rates and charge asymmetries at LEP has allowed for sensitive tests of the predicted couplings.

Photon-gluon fusion is one of the principal QCD processes that leads to the production of heavy flavor particles and the measurement of heavy flavor production in *ep* interaction can be used to measure the gluon structure function over a wide kinematical range. Such measurements are very important to our understanding of many QCD processes.

Heavy flavor particles also play a major role in the search for new particles, hadrons and leptons. In the Standard Model, the hitherto undiscovered *t* quark is expected to decay as $t \rightarrow W^+b$, resulting in a high transverse momentum lepton and two jets, one of which is associated with the b quark. Non-standard decays may result in transitions involving a Higgs particle and a b-jet. Many searches for the elusive Higgs particle can be assisted by precision vertex detection. In particular, the detection of signals for new phenomena will be aided by information about the presence of secondary vertices, most importantly the differentiation of prompt and secondary leptons.

Among the heavy flavor particles, those carrying the beauty quantum number are expected to carry the key to our understanding of CP violation, a phenomenon that so far has only been observed in the decays of neutral kaons. In the framework of the Standard Model, a number of quite sizable CP violating effects are predicted in rare decays of B mesons. As in K^0 decay, the CP violating effects are expected to be dependent on the B decay times. Thus experiments that are designed to search for and study CP violation in the B^0 system will require precision vertex measurements.

In summary, the detection of heavy flavor production and decay in a wide range of different processes has played and will continue to play a prominent role in experiments that test our understanding of weak and strong interactions.

2. HEAVY FLAVOR DETECTION AND LIFETIME MEASUREMENTS

From the point of view of an experimental physicist, the detection of heavy flavor particles and the measurements of their lifetimes are attractive because they represent a major challenge to the design and operation of detectors and to the data analysis. The principle difficulties experimentalists face in the detection of heavy flavor particles are their small production rates in hadron and photon interactions, and their short decay lengths.

2.1 Signatures for Heavy Flavor Particles

Signatures for heavy flavor particles can be derived from their relatively high mass and the weak nature of their decay. Masses of 2 GeV/c^2 or more give rise to relatively large transverse momenta of the decay secondaries. Unfortunately, they also allow for a large variety of different decay modes and thus lead to rather small branching ratios for any particular decay. The lowest lying charm and beauty particles decay via the flavor changing charged weak current resulting in relatively long lifetimes. The CKM enhancement leads to sequential decays, i.e., the decay chains $b \rightarrow c$ and $c \rightarrow s$. Thus strange particles can be used to enhance a sample of charm particles, and charm particles can be used to enhance a sample of beauty candidates. The weak coupling also causes the emission of leptons ($e^{\pm}$, $\mu^{\pm}$, $\tau^{\pm}$, and neutrinos).

In Table 2.1, the lifetimes of common particles are listed together with their average decay length $c\tau$. The proton, electron and photon are stable particles. The neutron and muon are also considered stable, because for momenta of more that 1 GeV/c their average decay length far exceeds the dimension of any detector. Particles which decay via electromagnetic or strong interactions have lifetimes of less than of 10^{-16} s and are considered short-lived, because their decay length are too short to allow for direct experimental observation. These particles are observable only via their decay products as peaks in effective mass distributions, or in formation experiments via the energy dependence of their production rate. Particles which decay via weak interaction include the charged pions, and the strange particles, like the charged and neutral kaons and hyperons. Their decay lengths $c\tau$ range from a few cm to several meters. Particles carrying the charm or beauty quantum numbers have lifetimes of the order of 1 ps, resulting in decay lengths $c\tau$ of less than 1 mm. Thus these heavy flavor particles can be distinguished from stable and very short lived particles by their decay vertex which is displaced from the primary production vertex. Figure 2.1 shows typical vertex topologies for $B\overline{B}$ production in both colliding beam and fixed target experiments.

In summary, an experiment with good sensitivity to heavy flavor particles requires a detector with large acceptance, excellent momentum resolution and good particle identification, preferentially both for hadrons and leptons, and a vertex detector with superb resolution and granularity. Furthermore, large samples of charm and beauty particles are not possible without selective and efficient triggers, or at least the possibility of a fast off-line event filter.

Table 2.1. Lifetimes and Decay Lengths of Selected Particles.

Stable Particles	$\tau > 10^{-6}$ s	
p	$\tau_p > 1.6 \times^{-23}$ s	
n	$\tau_n = 888.6 \pm 3.5$ s	$c\tau_n = 2.66 \pm 10^8$ km
e,γ	$\tau_e > 2 \times 10^{22}$ y	
$\mu^{\pm}$	$\tau_\mu = 2.19703 \pm 0.00004$ μs	$c\tau_\mu = 658$ m
Very Long-Lived Particles	$\tau > 10^{-10}$ s	
$\pi^{\pm}$, $K^{\pm}$ $K_L{}^o$	$\tau_\pi = 2.6030 \pm 0.0024 \times 10^{-8}$ s	$c\tau_\pi = 7.8$ m
$K_S{}^o$, Λ^o, $\Sigma^{\pm}$	$\tau_\Lambda = 2.632 \pm 0.020 \times 10^{-10}$ s	$c\tau_\Lambda = 7.9$ cm
Long-Lived Particles	$\tau > 10^{-13}$ s	
$\tau^{\pm}$	$\tau_\tau = 0.303 \pm 0.008 \times 10^{-12}$ s	$c\tau_\tau = 91$ μm
D^+, $D_s{}^+$, $\Lambda_c{}^+$	$\tau_{D\pm} = 1.062 \pm 0.028 \times 10^{-12}$ s	$c\tau_D = 318$ μm
$B_d{}^o$, $B_s{}^o$, Λ_b	$\tau_B = 1.18 \pm 0.011 \times 10^{-12}$ s	$c\tau_B = 350$ μm
Short-Lived Particles	$\tau < 10^{-16}$ s	
π^o, η^o	$\tau_{\pi^o} = 8.4 \pm 0.6 \times 10^{-17}$ s	$c\tau = 0.025$ μm
ρ, ω, K*, Δ, Σ	$\tau_\rho = 4 \times 10^{-24}$ s $^{-17}$ s	$c\tau \leq 10^{-9}$ μm
D*, B*	$\tau_{D^*} = 1.6 \times 10^{-23}$ s	$c\tau < 10^{-8}$ μm

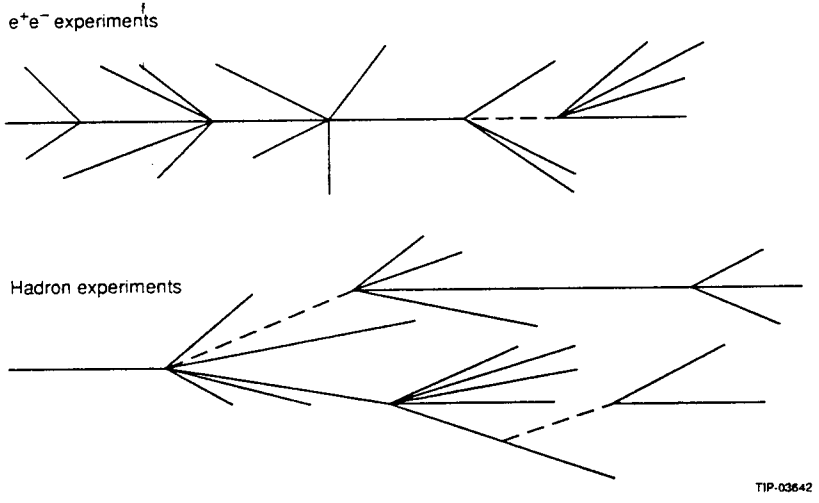


Figure 2.1 Vertex topologies for: (a) e^+e^- experiments, and (b) hadron experiments.

2.2 Heavy Flavor Production

The principal difficulty in detecting heavy flavor particles arises from their small production rate. Experimenters have chosen a variety of different beams, at energies ranging from threshold to the highest available. In high energy neutrino experiments, roughly 10% of the total hadron production by charged current interactions involves charm particles, the total cross section at 100 GeV is, however, of the order of 1 pb. In photoproduction, the total charm cross section is of the order of a few μb, and represents roughly 1% of the inelastic cross section. Both experiments at HERA are planning to use the measurement of charm mesons to derive the gluon structure function. The highest cross sections for charm and beauty particles have been measured in hadron beams, of the order of 10-100 μb/nucleon for charm, representing less than 1/1000 of the total inelastic cross section. Both charm and beauty cross sections are expected to rise substantially with energy. Calculations indicate that at the SSC energy of 40 TeV, the beauty cross section may be as high as 1 mb, constituting as much as 1% of the total cross section. These very favorable rates may make studies of rare decays of beauty particles feasible, and may allow for the observation and measurement of CP violation.

Experiments at e^+e^- storage rings have the advantage that heavy quark production makes up more that 30% of the total cross section. Experiments at CESR and LEP have accumulated large samples of events containing charm and beauty particles, many of which can be selected on the basis of kinematics alone, without the use of a special trigger and avoiding losses due to cuts on minimum decay distances. The total data rate is limited by the luminosity, given by the current and cross sections of the colliding beams.

In hadron machines, candidate events are selected on the basis of the event characteristics described above, namely the transverse energy, leptons of high transverse momentum, and multiple charged kaons. More recently, several experiments have attempted to detect decay vertices on-line. So far the results have not been convincing, but tests are underway to improve these techniques and the use of fast processors will enhance the chances for success. In fixed target experiments, active targets made of nuclear emulsion or silicon have been employed to maximize the sensitivity to very short decay distances. The beam intensity and target density are usually adjusted to match the data acquisition capability of the detector.

In colliding beam experiments, detectors are usually placed on the outside of the thin-walled vacuum at a radius of several cm. and charged particle tracks have to be extrapolated over this distance. This is a serious disadvantage at lower energies, where the decay length are only of the order of 100 μm, and where multiple scattering contributes significantly to the measurements errors. Figure 2.2 shows an example of a topology for an event $Z^o \rightarrow b\bar{b}$ + *anything*, without and with measurement errors as they occur typically in a LEP experiment. It is quite evident that the distinction between the particles from B decay is not straightforward.

2.3 Decay Length Measurement

The standard method to determine the particle decay times t is to measure the particle momentum p and decay path L, and thus requires an accurate determination of the production and decay vertices,

$$L = \beta \gamma c t = p/m \; c t.$$

If the decay products are not fully detected, the momentum is often estimated from an unconstrained kinematic fit or from the measured effective mass and momentum sum of the measured decay tracks.

A more model independent estimate is based on the measurement of the decay length in the plane transverse to the beam,

$$L_t = p_t /m \; c t,$$

where p_t and L_t are the transverse momentum and transverse decay length, respectively. This method uses the fact that the transverse momentum distributions are much better understood. The problem is that the transverse momentum components are generally much smaller than the longitudinal components, and thus for short lifetimes, the error on the length L_t can become significant.

The measurement of the transverse decay length in a colliding beam experiment is illustrated in Figure 2.3. The beam-beam interaction region is of elliptical shape and can be measured and monitored using Bhabha scattering events. Its size varies from machine to machine, typically dimensions are 300 μm x 40 μm. The beam position and size are usually stable for many hours of operation. One can either estimate the

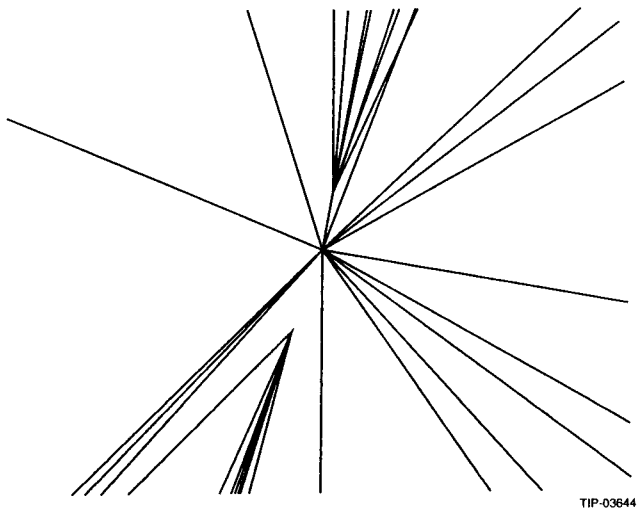


Figure 2.2 Track pattern of a Monte Carlo generated event $e^+e^- \rightarrow Z^o \rightarrow b\bar{b}$ + *anything* projected onto the plane transverse to the beam; (a) without measurement errors; (b) with typical measurement errors of a vertex detector located outside the vacuum pipe.

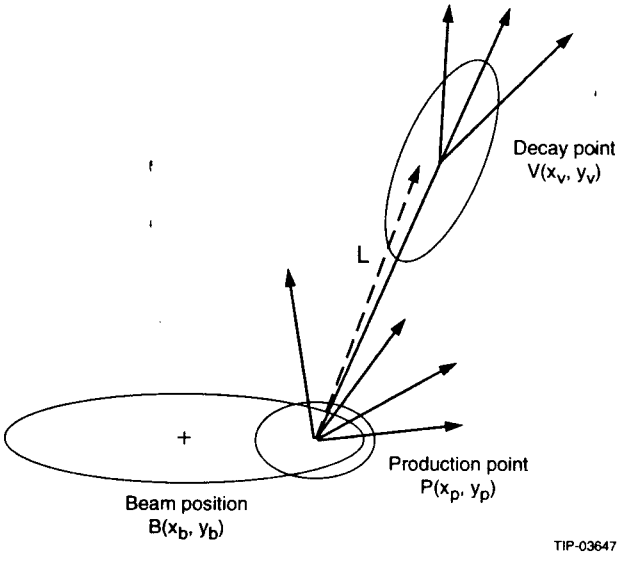


Figure 2.3 Decay length measurement for a particle produced in a colliding beam experiment.

production point from the knowledge of the position $B(x_b,y_b)$ and size averaged over many events, or from the tracks of the individual event. In most cases the beam spot is larger than the position error of the individual tracks. In a typical $Z^0 \rightarrow b\bar{b}$ event, on average about 11 of the 20 charged tracks originate from B or D decays. Finding the vertex $P(x_p,y_p)$ from the remaining 11 tracks is not easy, because many are of low momentum. The decay point $V(x_v,y_v)$ is determined using the secondary tracks. The direction of flight (t_x,t_y) is usually approximated by the jet axis, thus compensating for unmeasured neutral decay particles. The transverse decay length L_t and its error σ_L can be written as [ref 2.1]

$$L_T = \frac{x_v\sigma_{yy}t_x + y_v\sigma_{xx}t_y - \sigma_{xy}\left(x_vt_y + y_vt_x\right)}{D}$$

$$\text{and} \quad \sigma_L^2 = \frac{\sigma_{xx}\,\sigma_{yy} - \sigma_{xy}^{\,2}}{D}$$

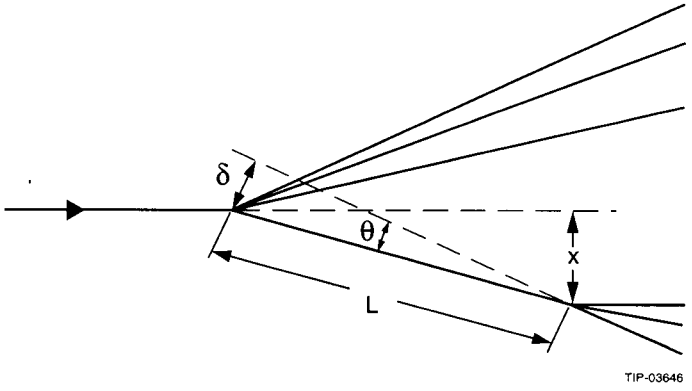
with $D = t_y^{\,2}\,\sigma_{xx} + t_x^{\,2}\,\sigma_{yy} - 2t_x\,t_y\,\sigma_{xy}$.

2.4 Impact Parameter Measurement

A very common method relies on the measurement of the so-called impact parameter b which is defined as the distance of closest approach of a track to the interaction point (IP) (see Figure 2.4). The impact parameter is positive if the intersection of the track with the reconstructed flight path of the decaying particle corresponds to a positive decay length; it is negative otherwise [ref 2.2]. b is proportional to the product of the decay length and the decay angle Ψ of the track,

$$b = \frac{\beta c\tau \cos\Psi^* \sin\theta^*}{\sqrt{\left(\beta/\beta^* + \cos\theta^*\right)^2 + \gamma^{-2}\sin^2\theta^*}} \propto ct$$

where Θ is the polar angle in the center-of-mass system. In the relativistic limit, b becomes insensitive to the momentum of the decaying particle. This is not surprising since there is a trivial correlation between the decay angle Ψ and the momentum $\beta\gamma$. As the momentum increases the angle Ψ decreases. The clear advantage of this estimator is that it does not require a fully reconstructed decay or estimate of momentum. Individual tracks from hadronic or semileptonic decays can be used, thus



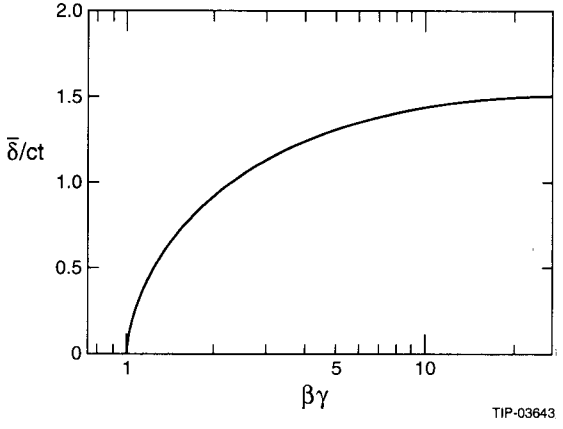


Figure 2.4 Impact parameter δ: (a) definition; (b) dependence of the normalized impact parameter δ/cτ on the track momentum βγ.

avoiding losses due to small branching ratios and limited detector acceptance. Monte Carlo simulations are needed to relate the impact parameter to the decay time; this can be done to an accuracy of about 10%.

Impact parameter measurements are not only used for lifetime measurements, they can also serve as an indicator for the presence of one or more secondary vertices and thus for heavy flavor particles in the event. It is hoped that in the future fast on-line determination of track impact parameters and their errors can be used as a trigger signal for beauty particles.

The error on the impact parameter b depends on the detector geometry, resolution and the total mount of material. For a simple vertex detector with only two measurements with the resolution σ_x at distances R_1 and R_2 from the production point, the resolution in the impact parameter is given by

$$\sigma_b^2 = a^2 + b^2/p^2 + c^2,$$

with

$$a = \frac{R_1\sigma_x}{R_2 - R_1} + \frac{R_2\sigma_x}{R_2 - R_1}$$

$$b = 0.015 \sum R_i \sqrt{\frac{X}{X_o}} \sin^{-3/2}\theta \ .$$

Here the first term represents the intrinsic resolution of the detector and is determined by the detector geometry and resolution, the second term accounts for the multiple scattering in the first detector layer plus any other material before the first measurement. The third term accounts for all other effect that influence the resolution, like the accuracy of the relative alignment of the two detectors and their stability. As an example, Figure 2.5 shows the calculated resolution in the impact parameter measured relative to the beam in the center of a two-layer cylindrical silicon detector as a function of the intrinsic resolution σ_x, the radius of the first layer, and the thickness of the vacuum pipe which is placed just inside the first detector layer. It is evident that the intrinsic resolution is only relevant for high momentum tracks, while the distance of extrapolation given by the placement of the first layer is important for all momenta. Thus the design of a vertex detector should adhere to the following criteria:

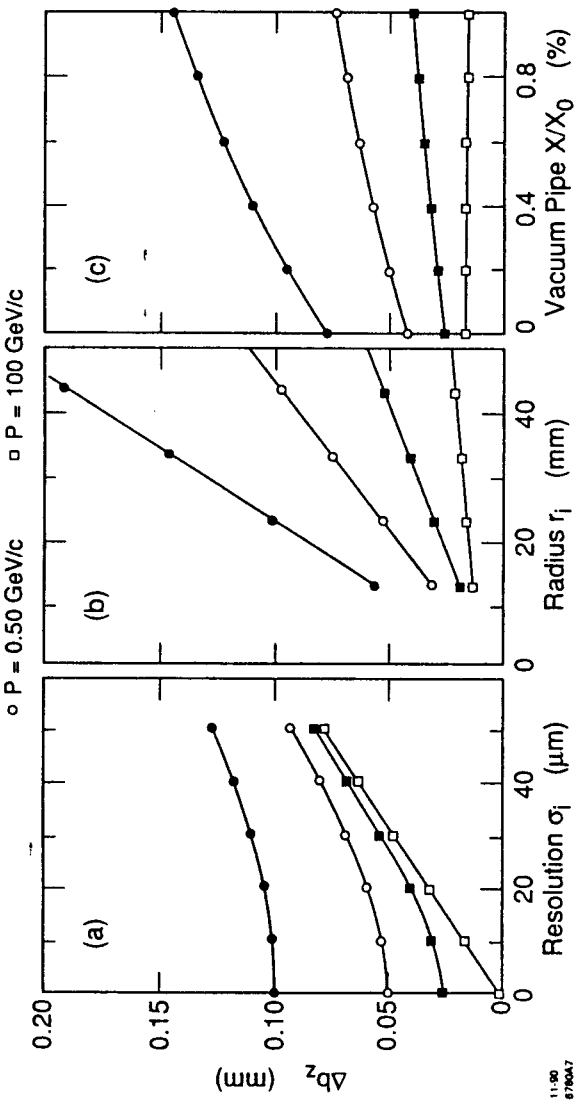


Figure 2.5 Calculated impact parameter resolution for tracks of different momentum as a function of (a) the intrinsic position resolution, (b) the distance of the first measurement from the beam, and (c) the thickness of the vacuum pipe right inside of the first detector layer. The detector is a two-layer cylindrical array of detectors with 10 μm resolution, with 300 μm thick detectors placed at 20 mm and 50 mm from the beam.

- a minimum distance between the interaction point and the first detector layer (in collider experiments, this is determined by the size of the vacuum pipe),
- a large distance between the first and the last measurement, $\Delta R = R_2 - R_1$,
- good position resolution, especially in the first layer,
- as little material between the interaction point and the first measurement to reduce multiple scattering (in collider experiments, this means a thin vacuum pipe), and
- good relative stability and accurate alignment of the detectors.

2.5 Lifetime Measurements

For an exponential decay distribution

$$P(t) = 1/\tau\, e^{-t/\tau},$$

where τ is the average lifetime and t is the measure decay time, the logarithmic likelihood function is

$$\mathcal{L}(\tau) = \sum_{i=1}^{N} \{-t_i / t - ln\, \tau\}$$

where the summation extends over all N measurements. The mean lifetime is determined from a fit that maximizes the likelihood with respect to τ, resulting in

$$\bar{\tau} = \Sigma t_i / N.$$

Thus the best estimate of the average lifetime τ is the mean of the observed decay times. In most experiments, decays can only be observed in a limited range $t_{min} < t < t_{max}$, and thus the decay distribution needs to be modified,

$$P(t) = 1/\tau\, e^{-t/\tau} \left\{ e^{-t_{min}/\tau} - e^{-t_{max}/\tau} \right\}^{-1}.$$

The logarithm of the likelihood changes to

$$\mathcal{L}(\tau) = \left\{ -t_i / \tau - ln\, \tau - ln\, (e^{-t_{min}/\tau} - e^{-t_{max}/\tau}) \right\},$$

where t_{min} and t_{max} may vary for each decay, for instance $t_{min} = \sigma_i$. Specifically, for $t_{max}=\infty$, the best estimate for the average lifetime becomes

$$\bar{\tau} = \Sigma(t_i - t_{min}) / N,$$

i.e. one does not need to measure decay times from t = 0, but relative to a cut-off t_{min}, which may in some cases vary from event to event. If, on the other hand the cut-off at large times is finite, the simple summation does not hold, and the decay rate provides a better estimator.

The error on the lifetime estimate is usually estimated from the second derivative of the likelihood function at the maximum,

$$\sigma_\tau = \left(-\frac{\partial^2 \mathcal{L}}{\partial \tau^2} \right)^{-1/2} = \bar{\tau} / N .$$

The upper and lower error limits can be derived as

$$\mathcal{L}(\bar{\tau} + \sigma_i) = \mathcal{L}(\bar{\tau} - \sigma_i) = \mathcal{L}(\bar{\tau}) - \tfrac{1}{2} .$$

In most lifetime measurements, the sample is not free of background events, and the measured distribution needs to be treated as a sum of signal and background, where the background time distribution has to be derived independently from the data.

3. NON-SOLID STATE DETECTORS

The very first tracking detectors for ionizing particles were photographic plates, that were soon followed by nuclear emulsions. At that time, however, the superb spatial resolution achievable with these devices was not considered necessary and the emergence of bubble chambers and electronic detectors lead to the decline of these techniques. By the early 1970s, drift chambers had been developed and were used as electronic tracking devices, both in fixed target and colliding beam experiments. Their precision of typically 200 μm was considered adequate for charged particle tracking and momentum measurements, as well as K_S^0 and Λ^0 detection.

The discovery of J/ψ particle in November of 1974 and the τ lepton in 1975 changed this situation dramatically. Measurements of particles lifetimes of the order of 10^{-13}s demanded much high precision for particle tracking near the production vertex. In fact, there existed hints for decays of short-lived particles in cosmic rays recorded in emulsions, but the decays were not fully reconstructable. The identification of these particles and the measurement of their lifetimes required precision vertex detectors that were embedded into large and complex spectrometers operating at accelerators and with capabilities for momentum measurement and full particle identification.

3.1 Nuclear Emulsions

Nuclear emulsions consist of thin layers (25 μm - 2000 μm) of gelatin into which silver bromide crystals are embedded. Like in photographic plates, in nuclear emulsions ionizing particles generate changes in the silver bromide grains such that the Ag^{++} ions can be reduced to Ag atoms during the development process. Thus ionizing particles are registered by the location and density of silver grains, furbishing information about the track position and direction as well as energy. The spatial resolution is governed by the grain size, 0.1 μm to 0.6 μm, and the grain density, typically 300 grains/mm for minimum ionizing tracks. Both the grain density and the development process are specifically tailored to each application.

Stacks of nuclear emulsions have been used as active targets in which both the production and decay vertices were observed down to distances of less than 1 μm. The major shortcomings of the emulsion technique are the following:

- The transparent emulsions need to be scanned and measured under the microscope, a very laborious and time consuming process.
- Since tracks are often crossing several layers of emulsion, the scanning process is very difficult and results in efficiency losses that increase rapidly at distances of a few mm. This problem is particularly severe for neutral particles that do not leave a track that can be followed through the different layers. Precise fiducials are required to index the relative position of the layers.
- Emulsions are continuously sensitive and the total exposure is limited by the total number of tracks. As a result, neutral beams, neutrons or neutrinos, have been used frequently. The total lack of time resolution not only excludes the use of a trigger but also leads to difficulties in linking emulsion tracks to those observed in the downstream spectrometer. As a consequence, emulsion experiments can only record a limited number of rare events.

- The relatively high density of emulsion targets can lead to secondary interactions but due to the excellent resolutions for the low energy nuclear recoils, such interactions can usually be distinguished from decays in flight.

Nuclear emulsions have a long history of discovery of elementary particles: the decays $\pi^- \rightarrow \mu^- \rightarrow e^-$ and the $K^+ \rightarrow \pi^+\pi^+\pi^-$ were first observed in emulsions, and in 1971 the first candidates for charm particle decays were observed in cosmic rays interacting in an emulsion chamber [ref 3.1]. Since then, nuclear emulsions have been revived as active targets for lifetime measurements because of their superb spatial resolution and granularity. Figure 3.1 illustrates the precision that can be achieved in emulsions, of the order of 0.5μm, corresponding to decay times as small as 10^{-14} s [ref 3.2]. In 1985, the first $B\overline{B}$ event with measurable lifetimes was recorded at CERN [ref 3.3]. With the addition of high precision tracking external to the emulsion, computer-aided scanning substantially enhanced the detection efficiency and data rate capability of this technique. Several experiments have reported results on charm particles, most recently E 653 at Fermilab [ref 3.4].

Figure 3.2 shows the layout of the E 653 spectrometer that was operated in a 600 GeV π^- beam at Fermilab. Each active target module consisted of 0.84 ℓ emulsion composed of sheets of 330 μm thickness, placed both parallel and transverse to the beam direction. To ensure uniform exposure during the 20 s beam spill, the target was moved transverse to the beam, with the relative position known to 10 μm; 49 target modules were used during the run. Silicon strip detectors were placed upstream and downstream of the target to measure the incoming beam tracks and the secondary track from the interaction. The vertices in the target were extrapolated from the silicon microstrip detectors with a resolution of $\sigma_x=\sigma_y=15-30$ μm and $\sigma_z=300$ μm. This lead to a reduction in the emulsion scanning time by an order of magnitude and also increased the track finding efficiency, reaching over 90%. The downstream spectrometer consisted of a dipole magnet and precision drift chambers for momentum and angle measurement, Time-Of-Flight (TOF) counters and calorimeters for electron identification, followed by a muon detector made of magnetized iron and drift chambers.

The readout was triggered by the presence of a muon of more than 8 GeV/c total and 1 GeV/c transverse momentum. This trigger was designed to select muons from the semi-leptonic decays of charm and beauty particles. A sample of $8 \cdot 10^6$ events out of a total of $2.5 \cdot 10^8$ interactions was recorded. Almost 700 charm meson decays were

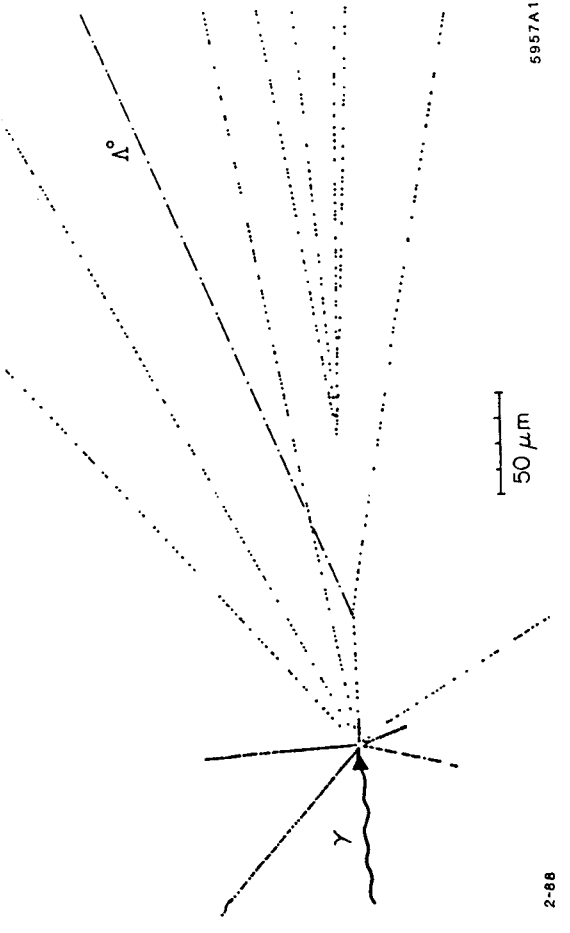


Figure 3.1 Reconstructed event from the CERN WA-58 emulsion experiment for the reaction $\gamma p \rightarrow \overline{D}^0 \Lambda_c^+ +$ *anything* showing the decays $\Lambda_c^+ \rightarrow \Lambda^0$ (undetected) π^+ *at* 50 ± 1 μ*m* and $\overline{D}^0 \rightarrow K^+\pi^+\pi^-\pi^-$ *at* 124 ± 2 μm from the interaction point.

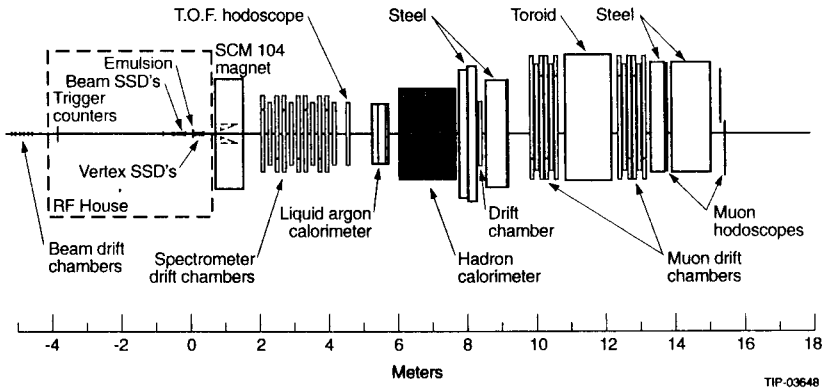


Figure 3.2 Layout of the E 653 spectrometer at Fermilab which employed an emulsion target backed by a silicon strip detector and followed by a dipole magnet, drift chambers, e.m. and hadron calorimeter and a muon detector.

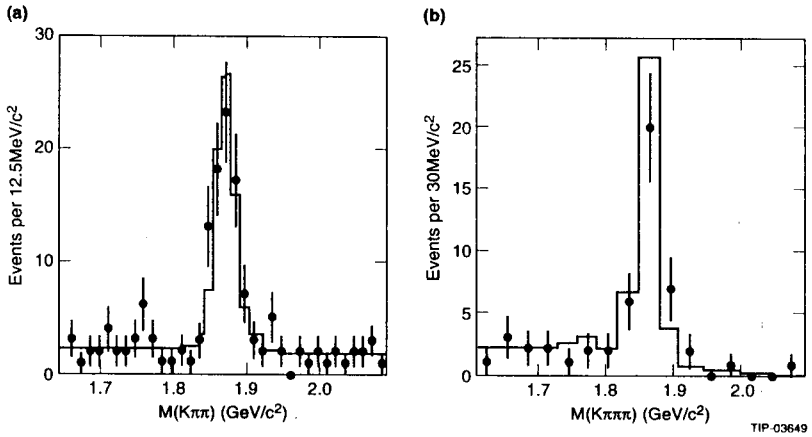


Figure 3.3 Fermilab E 653: Effective mass distribution for exclusive D decays a) $D^+ \rightarrow K^-\pi^+\pi^+$and, b) $D^0 \rightarrow K^-\pi^-\pi^+\pi^+$.

detected and reconstructed. Examples of effective mass distributions for exclusive D decays are given in Figure 3.3, the summary of the cross section measurement is given in Table 3.1.

Table 3.1. $D/\overline{D}$ Cross Sections as Measured by the Fermilab E653 Experiment .

	Decay Mode	Events $(x_F > 0)$	Estimated Background	$\sigma(D/\overline{D})$ for $x_F > 0$ (μb/nucleon)
D^+	$K^-\pi^+\pi^+$	73	8	$10.07 \pm 1.04 \pm 3.10$
	muonic 3-prong	278	7	$8.32 \pm 0.51 \pm 1.98$
D^o	$K^-\pi^+$	34	1	$15.05 \pm 2.92 \pm 4.93$
	$K^-\pi^-\pi^+\pi^+$	27	1	$19.05 \pm 4.08 \pm 5.82$
	muonic 2-prong	264	11	$22.71 \pm 1.43 \pm 5.15$
D^{+*}	$D^0\pi^+$	133	47	$7.78 \pm 1.46 \pm 1.99$

3.2 Bubble Chambers

In the overheated liquid of a bubble chamber, ionizing particles form traces of gas bubbles along their path. The bubbles are recorded in two or more projections on photographic film. With the help of coordinate measuring equipment the projected track images are used to reconstruct three-dimensional images of the tracks in the chamber. Chambers varying in size from a 150 cm^3 to 3 m^3 have been operated with H_2, D_2, and heavier cryogenic liquids. Their spatial resolution is determined by the bubble diameter, as small as 20 μm in small chambers, and bubble density, up to 5 bubbles/mm. Impact parameters have been measured to an accuracy of a few μm.

The primary advantage of bubble chambers is their complete solid angle coverage and their excellent pattern recognition capability for complex multi-vertex event topologies. In hydrogen (density ρ=0.071 g/cm^3 and radiation length X_0=865 cm) there is no ambiguity between decays and interactions of secondary particles.

The major disadvantage of this technique is the limitation in rate due to a number of different factors:

- The maximum exposure rate is 30 pictures/s due to requirements for film transport and flash light operation.
- The beam intensity is limited to about 30 charged beam tracks/picture.
- The only way to select events is to suppress the film transport and flash lamp operation unless an inelastic interaction has been detected by an external trigger system.
- The scan of the recorded film is very time consuming, though semi-automatic procedures have been employed.

At CERN a small, rapid cycling bubble chamber LEBC was used as a target and vertex detector for the European Hybrid Spectrometer for several years [ref 3.5]. LEBC was built from a single piece of Lexan (to avoid spurious boiling points along joints) and had an 12 cm long active volume, 5 cm high and 2.5 cm wide. The chamber was cycled at 30 Hz, illuminated by laser light and operated with two cameras. The position resolution was pushed to a few μm by optimizing the operating conditions for small bubble size and high bubble density.

Following event selection by two independent scans, the tracks were measured by a High Precision Measuring Device (HPD). Figure 3.4 shows a candidate for charm pair production as digitized by the HPD. A typical track of 5 cm length has roughly 400 bubbles of less than 20 μm diameter.

Though the achieved resolution is comparable to nuclear emulsion, the visual scanning process applied to select charm decays required at least one track with an impact parameter of more than 50 μm, severely limiting the ability to recognize very short-lived particles. Nevertheless, the NA-27 group measured the lifetimes of D^0 and D^+ mesons based on about 150 decays each, extracted from a total of 2 million photographs. The fact that only a few Λ_c and not a single D_s decay were found may be indicative of the scanning losses at short lifetimes.

More recently holographic techniques have been developed to increase the resolution of larger chambers suitable for operation in a neutrino beam [ref 3.6]. To obtain a hologram, the chamber is illuminated by a laser beam such that it interferes on the film with the diffracted light from the bubbles. For scanning purposes, the hologram has to be reproduced from the film. To date, this technique has only been applied in a pilot experiment to demonstrate its feasibility, and at this time there are no plans for further applications.

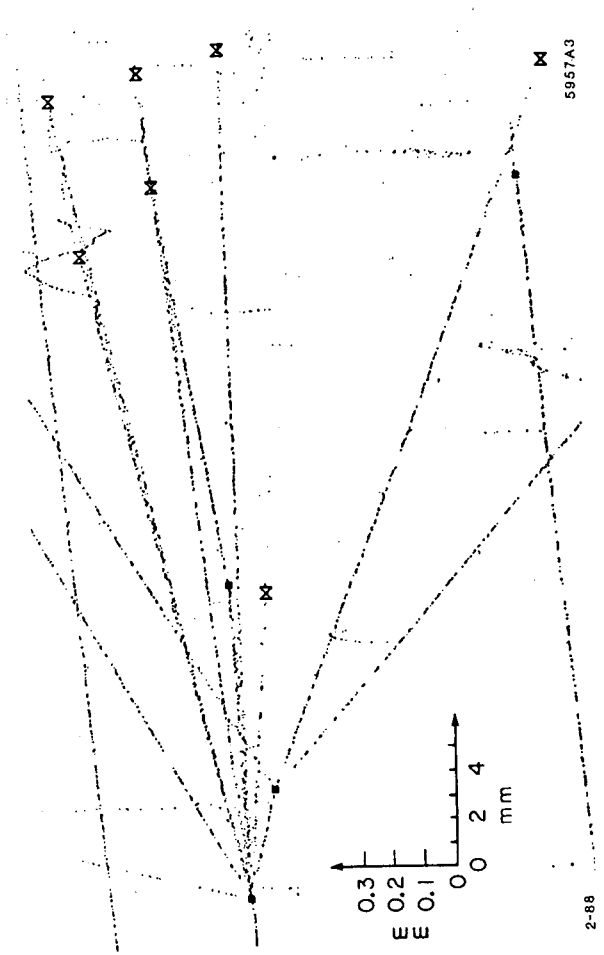


Figure 3.4 Reconstructed event with two charm decay vertices from the EHS spectrometer at CERN. The figure shows the digitization by the HPD of a photo recorded by the rapid cycling bubble chamber (LEBC) which served as an active target.

3.3 Scintillating Fibers

In scintillators, ionizing particles excite molecular levels and lead to the emission of light. In organic scintillators, the scintillation light arises from the transitions by the free electrons in the benzene ring structures of hydrocarbon compounds. The scintillation process involves two steps: The primary process is the de-excitation of excited molecular states leading to the emission of UV light. The secondary process is the absorption and re-emission of this energy at a longer wavelength. This re-emission is extremely fast, it occurs in less than 10 ps. Standard scintillators emit in the blue-violet region of the visible spectrum, the spectral region that is readily affected by radiation-induced damage in the polymer, leading to a decrease in light yield. Recently new fluorescent compounds emitting in the green/yellow region have been shown to be more resistant for damage by ionizing radiation [ref 3.7].

Plastic scintillators are most widely used because of their fast signal (decay constants of less than 1 ns) and their high light output. Plastic scintillators are polymers doped with scintillating compounds. They are relatively inexpensive and they are produced commercially in a wide variety of sizes and shapes (from thin films and fibers to large sheets, blocks and cylinders). Scintillating light can be channeled by total reflection to a photon detector, the most popular among them are photo-multipliers and vacuum photo-diodes.

Following the pioneering work by Ruchti et al. [ref 3.8], scintillating fibers of small cross section, round plastics or square shaped glass, are being considered as tracking and vertex detectors. Position sensitive photon detectors have been developed, for instance multi-anode solid state photomultipliers and most recently visible light photon counters (VLPC).

Two different applications of fibers have been proposed:

- Active targets made of a stack of densely packed and accurately placed fibers of 40 μm in diameter was designed for an experiment at CERN to study heavy flavor production [ref 3.9]. A schematic view of the target arrangement and magnified view of the cross section of the fiber target are shown in Figure 3.5. There are typically five fibers per mm of track, the position resolution is of the order of 20 μm, and the light yield was measured to be 40 photoelectrons per mm. Results reported so far are based on a rather elaborate readout system,

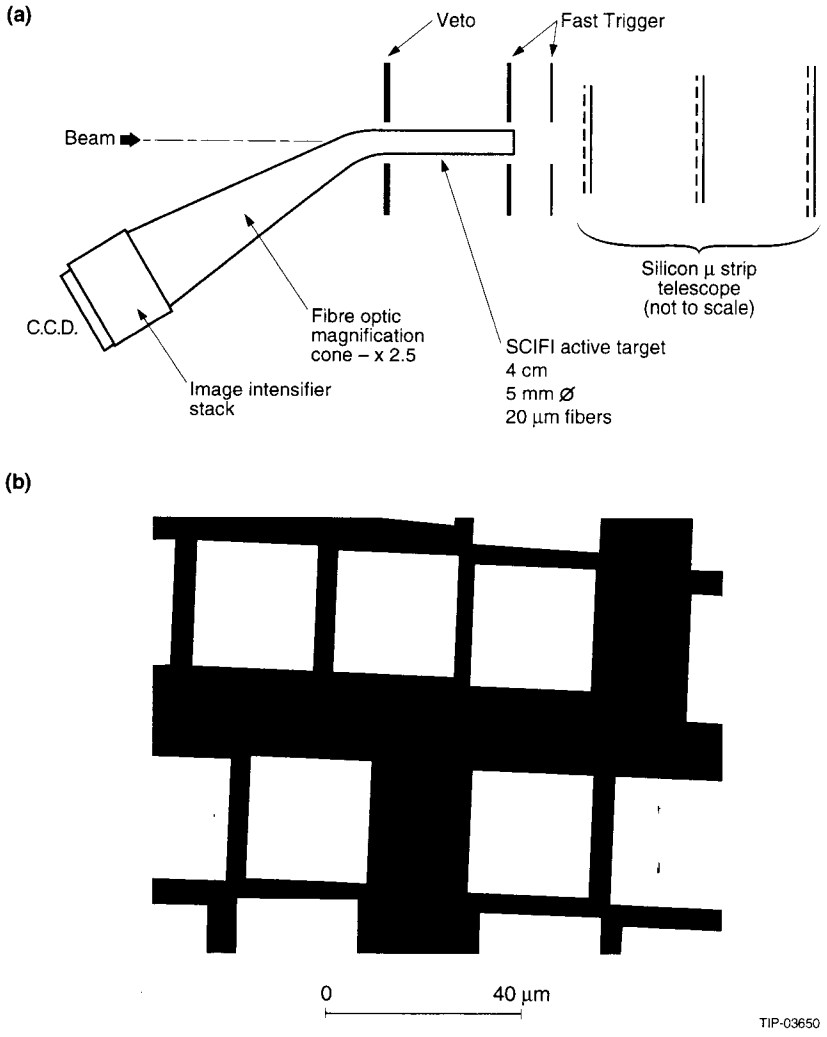


Figure 3.5 Scintillating Fibers: (a) schematic illustration of the target and readout; (b) high magnification photograph of a scintillating fiber target. The dimension of the fiber core is 25 μm and the center-to-center spacing is typically 31μm.

consisting of a photo-cathode proximity-focused on to a multi-channel plate image intensifier, followed by a CCD or photographic film.

- At LHC and SSC large cylindrical arrays of small diameter (0.3 mm to 1 mm) scintillating fibers of 3-4 m length are considered as charged particle tracking devices [ref 3.10]. Their high rate capability and fine segmentation have major advantages over gaseous tracking chambers. In a typical experiment of this kind, the scintillating fibers are coupled to optical fibers for signal transport to the VLPC read-out that is placed outside the sensitive volume. The survivability of the fibers in radiation levels of about 1 MRad, their attenuation length, and their VLPC readout remain under study.

3.4 Precision Drift Chambers

In drift chambers, the primary ionization generated by a charged particle is localized by a measurement of the time difference, $\Delta t = t_1 - t_0$, between the passage of the particle and the arrival of the pulse on the nearest anode wire. If the drift velocity v_D is known, then the distance from the sensing wire to the origin of the primary electrons is

$$x = \int_{t_0}^{t_1} v_D(t)dt \equiv v_D \cdot \Delta t \, .$$

For a typical drift velocity of 5μm / ns, a time resolution of 3 ns results in a spatial resolution of 15 μm. Electric field configurations can be shaped by the choice of cathode and anode wire placement and voltage. It is generally desirable to have an approximately constant electric field over most of the drift distance so as to obtain a linear relation between drift time and distance. In practice, the spatial resolution is limited by three principal factors:

- the knowledge of the wire location,
- the knowledge of the drift time and the space coordinate, and
- the amount of diffusion suffered by the electrons as they drift.

The diffusion depends on the total drift time t,

$$\sigma_x = \sqrt{2Dt} = \sqrt{2Dx / \mu E}$$

where x is the total drift path, D the diffusion coefficient, and μ the mobility. D and μ depend on many parameters; for instance the gas composition, the pressure, and the electric field.

Operation in a high rate environment can cause voltage breakdown and aging effects that will limit the performance.

Drift chambers are attractive tracking devices because of their good spatial resolution, their rate capability, and their fast electronic read-out. They also are rather low mass detectors, the main contribution to multiple scattering comes from the potential wires. Recently, low density gases like helium and low density wire materials, like aluminum, have been employed to reduce multiple scattering. The position measurement is usually restricted to the coordinate transverse to the wire direction. However, the coordinate along the wire can be obtained either by cathode strip readout or charge division on the wires. The latter method relies on the fact that the charge collected on either end of the resistive wire is divided in proportion to the length of the wire from the point at which the charge was injected. Accuracy of up to 0.5% of the wire length have been obtained. Cathode strip readout exploits the fact that the signals induced on the cathode plane are inversely proportional to the distance from the individual cathode strips. The track coordinate can be derived as the pulse height weighted average coordinate of adjacent strips. Resolutions of the order of 100 μm have been obtained.

Much effort has been devoted to enhance the resolution by improved cell design, by precision wire placement, and by the choice of chamber gases with low drift velocity and low diffusion. Chambers have been operated at high pressure to reduce diffusion and increase primary ionization, and point resolutions as small as 40 μm have been achieved over most of the drift cell for tracks of normal incidence. More typically, resolutions of 100 μm per point are obtained, leading to an impact parameter resolution of similar magnitude. Thus drift chambers have little chance of distinguishing individual tracks from primary and secondary vertices. Nevertheless, small cylindrical drift chambers have been built for colliding beam experiments to fit between the vacuum pipe and the main central tracking chamber. A recent example of such a chamber is the Mark II Vertex Drift Chamber [ref 3.11] that was installed at the Stanford Linear Collider.

The Mark II Vertex Drift Chamber was filled with a low diffusion gas (a 92%/8% mixture of CO_2/C_2H_5) and was operated at a pressure of 2 bar. This chamber, shown in Figure 3.6, is divided into 10 axial drift cells. Individual cells are tilted with respect to the radial direction to resolve left-right ambiguity, and to ensure that most tracks cross a cell boundary. In each cell, there are 38 active anode wires, 55 cm long, and placed at radii between 5 cm and 17 cm. Grid wires on both sides of the sense wire plane improve the charge collection isochronism. The anode signals are readout via a fast integrating preamplifier and transmitted to a 6 bit 100 MHz Flash ADC. The anode wires were placed to an accuracy of better than 3 μm relative to the wire plane which in turn was positioned with an accuracy of 20 μm. The operating conditions of the chamber were closely monitored and controlled. The voltages were set to 0.1% and the cathode supply voltages were kept stable to 0.03%. Temperature excursions were limited to less than 0.15°C, and the gas pressure was held stable to about 0.03%. The typical drift field was 0.73 kV/cm/bar, resulting in a drift velocity of about 5.6 μm/ns.

The spatial resolution of the vertex drift chamber was carefully monitored. Over most of the cell, the spatial resolution as a function of drift distance d can be approximated in the form $\sigma_x^2(\mu m)=15^2 + 37^2\, d(cm)$, where the second term accounts for the diffusion of the drifting electrons. Near the anode, the resolution is worse because of the large, rapidly varying drift velocity, and because the random spacing between the primary ionization electrons becomes significant (Figure 3.7a). The resolution in the impact parameter of the tracks extrapolated to the center of the beam inside the vacuum pipe has been measured using cosmic rays. Figure 3.7b shows the miss distance for the two halves of a cosmic ray track crossing the chamber. The rms width of 44 μm for the measured distance translates to a 31 μm resolution in the impact parameter of a single track.

3.5 Microstrip Gas Chambers

Microstrip Gas Chambers (MSGC) combine feature of a gaseous proportional chamber with those of a solid state detector. They are position sensitive detectors operating in the proportional mode, with anodes and cathodes realized as thin metal strips on insulating substrates [ref 3.12]. Silicon processing techniques are being applied in the fabrication, i.e., UV lithography and plasma etching combined with thin film deposition and ion implantation.

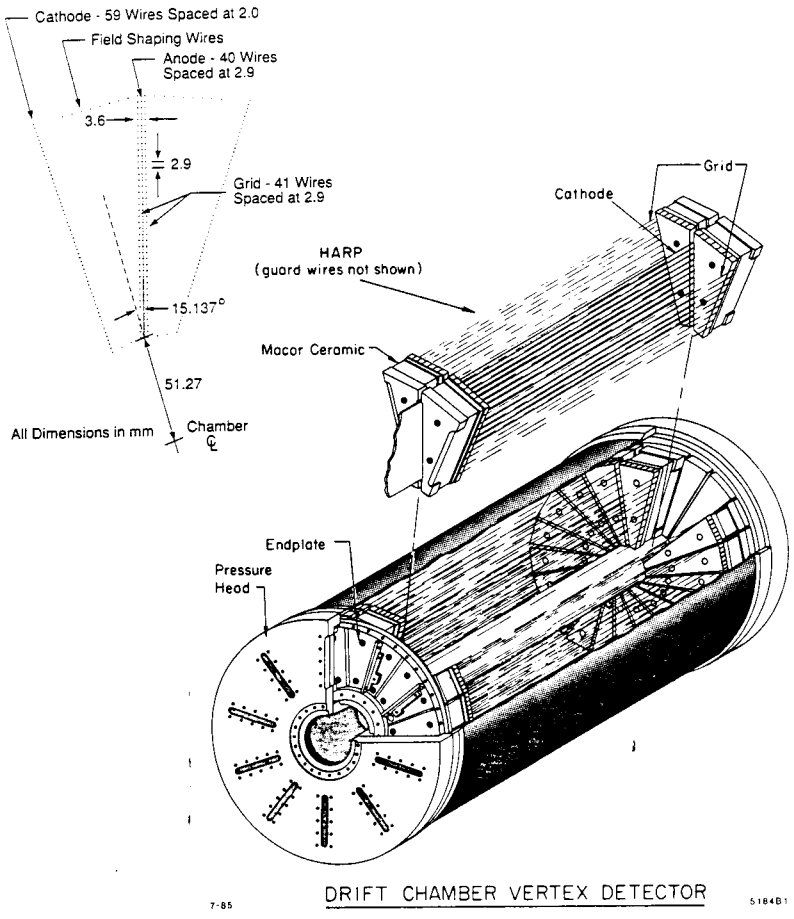


Figure 3.6 Mark II Vertex Drift Chamber showing the modular construction of the jet cell, the wire placement in a unit cell, and the pressure vessel.

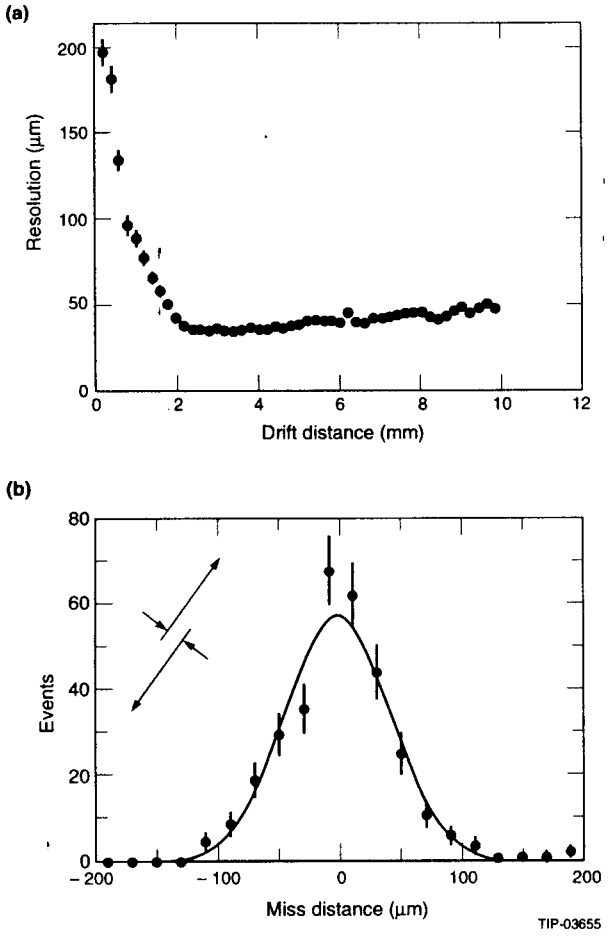


Figure 3.7 Performance of the Mark II Vertex Drift Chamber: (a) measured resolution versus drift distance near the anode wire; (b) the measured 'miss-distance' for the two segments of cosmic ray tracks extrapolated to the center of the detector, inside the aluminum vacuum pipe.

With very small spacing between electrodes, these devices provide excellent granularity as required at high luminosity accelerators. Good spatial resolution is due to the precise placement and dimensions of the strips (up to 0.1 μm accuracy) and the absence of the effects of electrostatic forces that lead to wire displacements in wire chambers. The small drift space leads to short charge collection times (< 9 ns).

MSGC were first employed in the NA-12 magnetic spectrometer at CERN [ref 3.13]. Figure 3.8 shows the cross section of such a detector together with the electric field map. The detectors had an active area of 25.6 mm x 25.6 mm with 128 anode (9 μm wide) and cathode (70 μm wide) aluminum strips placed at a pitch of 200 μm. The 2 μm thick aluminum strips were produced by UV lithography and plasma etching on a quartz substrate. The devices were operated with a Xenon (60%) - Di-Methyl-Ether (40%) gas mixture to enhance the primary ionization and achieve a high drift velocity in an electric field of 8 kV/cm. The typical gas gain was 10^4. Figure 3.9 shows performance data. The extrapolation error of two track segments is measured in the MSGCs, the width of this distribution translates to a position resolution of 50 μm per point. The prompt signal from groups of 16 cathode strips has an rms width of 9 ns, and thus can be used for triggering purposes.

While the precision fabrication techniques are attractive from the point of resolution, they impose a limit on the maximum size of these devices and incur substantial cost. The major shortcomings of MSGCs are connected with the insulating substrate: the accumulation of charge between the strips can lead to deformation of the field lines, change the gas gain, and ultimately cause surface breakdown. This problem can be avoided if the surface conductivity can be enhanced. Several approaches are being tried: NA–12 used ion-implanted glass as a substrate, a rather expensive solution which also leads to a substantial increase in the amount of material. Several groups [ref 3.14] have since investigated the use of thin (100 μm) plastic foils, that have lower resistivity, are flexible and allow non-planar designs. The only materials that were found to be suitable, with mechanical stability and sufficient adherence to metal layers, are Kapton and Tedlar with bulk resistivities of $10^{17}\ \Omega$ cm and $10^{14}\ \Omega$ cm, respectively. Figure 3.10 shows results from a test with a collimated x-ray tube with variable flux. For these tests a small area (1 mm^2) of the chamber was irradiated continuously and the gain was measured as a function of rate at certain time intervals. While the Kapton substrate shows substantial gain changes even at relatively low rates, chambers with Tedlar substrates show an initial drop in gain immediately after

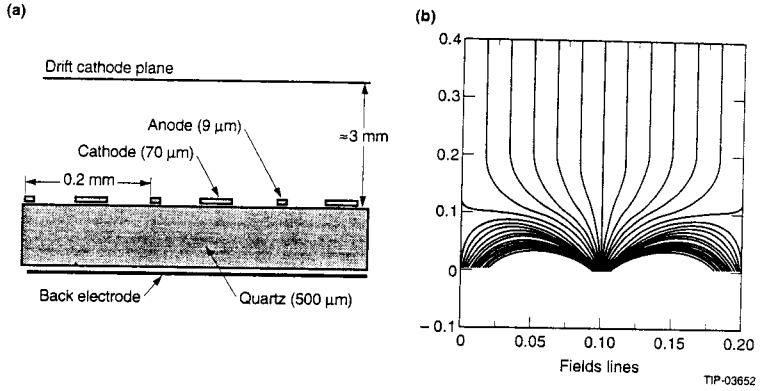


Figure 3.8 Gaseous Microstrip Detectors: (a) schematic layout; (b) electric field lines for a single cell with the anode in the center.

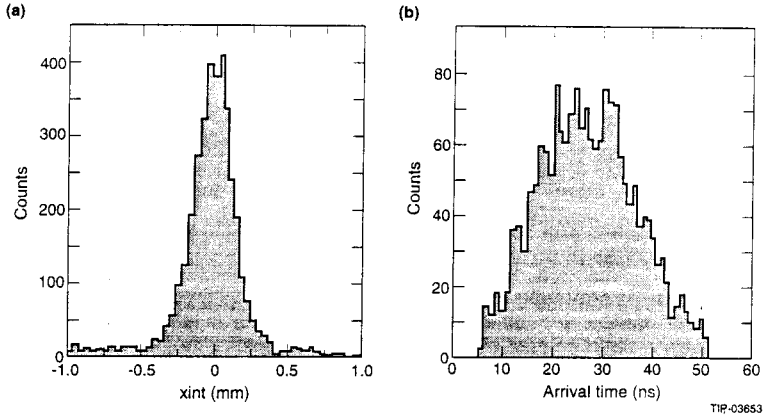


Figure 3.9 Performance of gaseous microstrip detectors measured in the CERN NA-12 spectrometer. (a) Distribution of residuals: the rms width of 140 μm translates to a position resolution of 50 μm in a single plane. (b) Distribution of signal arrival times with an rms width of 9 ns.

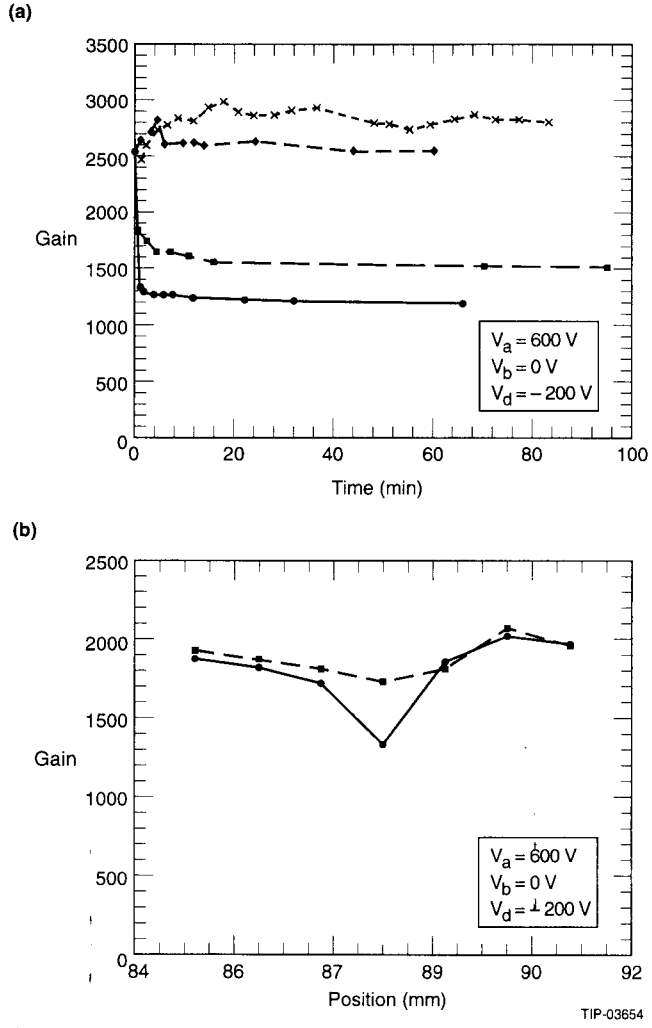


Figure 3.10 Effect of a one-hour-long irradiation by an x-ray source on the gain of gaseous microstrip chambers built on a Tedlar substrate. (a) As a function of time after the beginning of operation, starting 48 hours after the exposure. (b) As a function of distance from the position of the x-ray source, for a rate of 10^4 Hz/mm^2, immediately after the irradiation and with a 60 minute delay.

turn-on of the voltage, but level off (typically after 10 minutes of operation), even at rates of more than 10^4 Hz/mm^2. However, the gain changes are a local and temporary phenomenon; the gain is fully recovered 60 minutes after the irradiation.

These results are encouraging, and it is hoped that further investigations will lead to the development of gaseous detectors with thin substrates and good resolution and high rate capabilities. Such devices could find applications as general purpose tracking chambers, as position detectors in preshower counters, and transition radiation detectors at SSC and LHC, and potentially in charm and beauty experiments.

4. SILICON DETECTORS

Given the limitations of nuclear emulsions and bubble chambers, e.g., long sensitivity, limited data rate capability, and difficult scanning procedures, a major effort has been devoted in recent years to the development of silicon detectors with finely segmented electrodes for charged collection. Such detectors represent an outgrowth of the semiconductor counters that have been used in nuclear physics for more than twenty years. In high energy physics, the use of silicon as an active target and as a high resolution tracking device was pioneered by two groups at CERN, NA-1 [ref 4.1] and NA-11 [ref 4.2]. Since then, many other experimenters have built on this experience and have employed commercially available detectors with strip or pixel readout.

Before we describe these devices in detail, we will briefly review the basic properties of semiconductor materials [ref 4.3] which are important for their application as tracking detectors and we will study the fundamental limits to the precision of charge location in silicon.

4.1 Principle of Operation

Solid state materials can be grouped in three classes - insulators, semiconductors, and conductors. They differ by their electrical conductivity σ (and corresponding resistivity $\rho=1/\sigma$). Insulators such as glass have conductivities of the order of 10^{-18} S/cm, while conductors have typical values from 10^4 to 10^6 S/cm, and semiconductors range from 10^3 to 10^{-8} S/cm. The electrical conductivity is directly related to the atomic energy level structure of the materials. Semiconductors are crystalline materials in which the outer shell atomic levels form an energy band structure. Figure 4.1 schematically shows this band structure consisting of a valence band, a "forbidden" energy gap, and a conduction band. The band structure for conductors and insulators is also shown for comparison.

The energy bands are regions in which the many discrete levels of the atoms forming the crystal lattice are so closely spaced that they can be considered a continuum. The upper energy band is the conduction band. Electrons in this band are detached from the lattice atoms and are free to move. The electrons in the valence band are more tightly bound to the lattice atoms. The two bands are separated by a region in which there are no available energy levels. In insulators this gap is large, while in conductors the energy bands overlap and the gap does not exist. In insulators, all levels in the valence band are occupied by electrons and the conduction band is empty. The energy gap is so large that electrons cannot be transferred to the conduction band. In a semiconductor, the energy gap is substantially smaller, and therefore thermally excited electrons can jump into the conduction band, leaving a hole in the valence band. At 0°K, the band gap is 1.17 eV for silicon and 1.52 eV for gallium arsenide. The width of the band gap decreases with increasing temperature. At temperatures above 0°K, electron-hole pairs are constantly generated by thermal excitation in the semiconductor, while at the same time a certain number of electrons and holes recombine. Holes act like positive charges. The concentration of charge carriers, (electrons or holes) n_i depends on the temperature T, namely

$$n_i = AT^{3/2} e^{-E_g/2kT},$$

where E_g is the energy gap at 0°K and k the Boltzmann constant, and A is a constant independent of temperature. For silicon at 300°K the typical charge density is $1.5 \cdot 10^{13}$ / cm^3, which means that the concentration of charge carriers is very low, only 1 in 10^{12} silicon atoms is ionized. If an external electric field E is applied to a semiconductor, an electric current arises from two sources: the movement of the free electrons in the conduction band and the movement of the holes in the valence band. The current density is $J=\rho\, \upsilon$, where ρ is the charge density and υ the velocity of the charge, explicitly

$$J = en_i(\mu_e + \mu_h)E = \sigma E,$$

where μ_e and μ_h represent the mobilities of the electrons and holes, respectively, and σ is the conductivity. For a given material the mobilities depend on the field E and the temperature T. Some of the physical properties of silicon and germanium are given in Table 4.1.

Table 4.1. Some Physical Properties of Silicon and Germanium.

	Si	Ge
Atomic Number Z	14	32
Atomic Weight A	28.1	72.6
Density (g/cm^3)	2.33	5.32
Dielectric Constant (Relative)	12	16
Intrinsic Resistivity (300° K) (Ω cm)	230000	45
Energy Gap (300°K) [eV]	1.1	0.7
Electron Mobility (300° K) [cm^2/Vs]	1350	3900
Hole Mobility (300° K) [cm^2/Vs]	480	1900
Band Gap [eV]	1.12	0.66
Energy for e-hole Pair Creation (eV)	3.55	2.85
Intrinsic Carrier Concentration [cm^{-3}]	$1.45 \cdot 10^{10}$	$2.5 \cdot 10^{13}$

In pure semiconductors, the number of holes in the valence band equals the number of electrons in the conduction band. This balance can be changed by small impurities of atoms that have either one less or one more electron in their outer shell. Silicon and germanium have four valence electrons forming four covalent bonds in the crystal. Typical dopants for these two materials are boron and gallium (tri-valent) and phosphor and antimony (penta-valent). The effect of the dopants is illustrated in Figure 4.2. The addition of phosphor atoms to silicon creates an energy level in the gap very close to the conduction band. Thus the extra electrons can easily be excited to the conduction band and thereby increase the conductivity. They can also fill up holes and decrease the normal hole concentration. Thus in phosphor-doped semiconductors, the current is mainly due to electrons (often referred to as majority carriers) and the material is called an n-type semiconductor. Analogously, the addition of boron atoms leads to an energy state in the gap near the valence band,

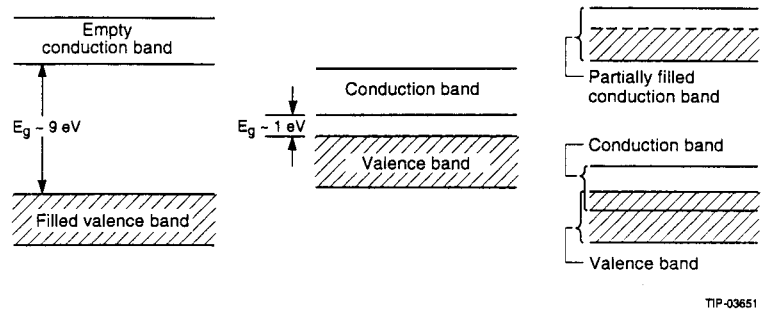


Figure 4.1 Schematic representation of the energy brands of (a) an insulator, (b) a semiconductor, and (c) a conductor.

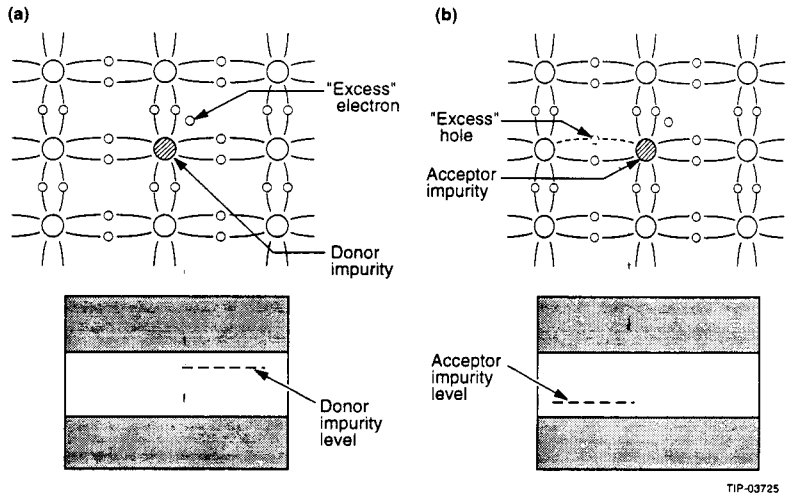


Figure 4.2 Effect of dopants in the silicon crystal structure (a) addition of donors like phosphor form n-type silicon by creating an energy level just below the conduction brand; (b) addition of acceptors like boron form p-type silicon by creating an energy level close the valence band.

electrons from the valence band are easily excited to this level, increasing the hole concentration. At the same time, the excess of holes leads to a reduction of electrons in the conduction band. Thus in boron-doped semiconductors the holes become the majority carriers and the material is referred to as a p-type semiconductor.

Doped semiconductors are commonly used in semiconductor devices. Typical concentrations of impurities are of the order of 10^{13} atoms/cm^3, i.e., only a few parts per billion. Heavy dopant concentrations of 10^{20} atoms/cm^3 are commonly used as electrical contacts.

All semiconductor detectors are made of single-crystal silicon or germanium containing both p- and n-type regions that form a so-called p-n junction. Such a junction is shown schematically in Figure 4.3. Due to the difference in the carrier concentrations charges diffuse across the junction: holes from the p-side diffuse into the n-side and electrons from the n-side diffuse into the p-side. As a result, charges build up on both sides of the junction, leading to an electric field across the junction which eventually halts the diffusion process. The region of changing potential is known as the depletion region, because it is devoid of any mobile charges. An ionizing particle crossing this region will generate electron-hole pairs which will be swept out by the electric field. It is this property of the p-n junction that makes it a good ionization detector.

The width of the depletion region can be extended by applying a reverse bias voltage across the junction, i.e., a negative voltage to the p-side as shown in Figure 4.4. The total width d of the depletion zone is then given by

$$d[\mu m] = \sqrt{2\varepsilon\rho\,\mu_e(V_o + V_b)} \;\; = \;\; 0.53\sqrt{\rho V_b}$$

where ρ is the resistivity, typically ρ=10 kΩcm for detector grade n-type silicon, ε=12 is the relative dielectric constant, μ_e is the electron mobility, V_o =1V is the contact potential and V_b is the bias voltage. For $V_b = 0$, one obtains $d = 75\ \mu m$ which is too small for practical applications. On the other hand, a typical n-type silicon detector with $D = 300\ \mu m$ can be fully depleted, i.e., $d > D$, by a reverse bias voltage of $V_b = 36$ V or larger. The maximum voltage which can be applied is limited by the resistance (i.e., purity) of the semiconductor. With extremely high

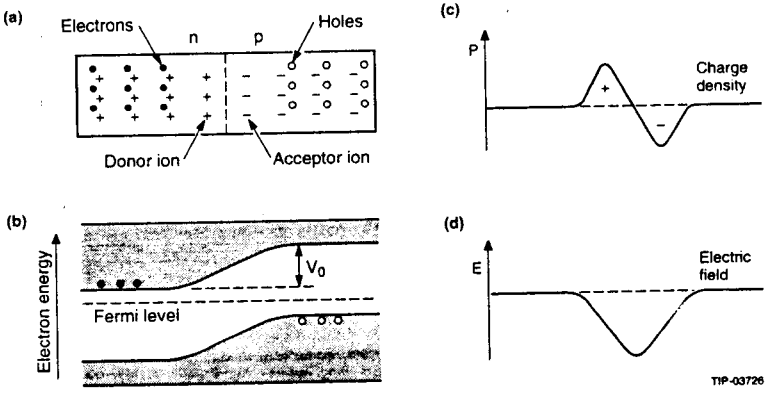


Figure 4.3 A p-n junction, (a) schematic diagram, (b) diagram of electron energy levels, (c) charge density, and (d) electric field strength.

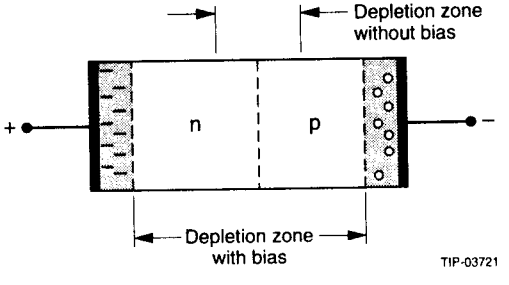


Figure 4.4 Schematics of a reversed bias p-n junction for a minimum ionizing particle.

purity silicon, depletion depths of several mm have been obtained before voltage breakdown occurred. The capacitance C is

$$C\left[pF/mm^2\right]=\frac{\epsilon}{D}A=2.2\sqrt{\rho V_b}$$

thus for a D=300 μm thick detector, C=350 pF/cm^2.

4.2 Limits to Spatial Precision

The principal reason for the high spatial precision of semiconductor detectors can be traced back to the confinement of the electrons from the primary ionization. This can be explained in a very simple, crude model of the ionization process. According to Rutherford, the probability for a collision of a particle (mass m, velocity βc, and charge e) with a free electron in the lattice to impact the kinetic energy T is [ref 4.4]

$$\frac{d^2N}{dTdx}=\frac{2\pi e^4}{m_e\beta^2c^2}\frac{NZ}{A}\cdot\frac{1}{T^2}\left\{1-\beta^2\frac{T}{T_{max}}\right\},$$

where N is Avogadro's number, A and Z are the atomic weight and atomic number of the material. The term $1/T^2$ underlines the fact that collisions with large energy transfers occur at a rather low rate. Since $1/T^2$ is proportional to the impact parameter b, this means that collisions with large impact parameters are dominant. For instance, for an extremely low transfer energy of T = 1 eV, we find (for β=1) b = 10^{-6} μm, thus the limit to precision will certainly not be set by the primary ionization which is dominated by very low energy electrons.

However, this approximation does not take into account fluctuations in the energy loss due to a few, very high energy transfers, so-called δ electrons, that give rise to the familiar Landau tail in the energy loss distribution. The maximum kinematically allowed energy transfer is

$$T_{max}=\frac{2m_e\beta^2\gamma^2c^2}{1+2\gamma m_e/m}.$$

Over the last fifty years our understanding of the energy loss processes in various material has substantially improved. While Landau [ref 4.5] ignored the effect of atomic binding and distant collisions, more recent calculations [ref 4.6] account for atomic binding and use Monte Carlo methods to include measured photo-absorption effects to take into account soft, nearly on-shell virtual photons generated in distant collisions. In Figure 4.5 the measured energy loss distribution for relativistic pions traversing a 20 μm thick CCD detector is compared to theoretical calculations, illustrating the improvement in our understanding of the complicated ionization process. Due to the long tail of the Landau distribution, the mean energy loss of a minimum ionizing particle is significantly higher than the most probable energy loss. This difference decreases with sample thickness. For a typical silicon detector of 300 μm thickness, the mean energy loss is 400 eV/μm, while the most probable is 290 eV/μm. With 3.6 eV per electron-hole pair, this corresponds to 110 and 80 electron hole pairs per μm, respectively.

The effect of the δ electrons is summarized in Figure 4.6 which shows both the probability for a minimum ionizing particle to produce a δ electron with an energy exceeding a value T, and the range of an electron of energy T. For example, in a 300 μm thick silicon detector, the probability to produce an electron of more than 20 keV is 30% and the range of a 20 keV electron is of the order of 5 μm. This 20 keV δ electron will generate secondary ionization, resulting in about 5500 electrons distributed along its path in addition to the 24,000 primary electrons. In general, this secondary ionization will distort the spatial distribution of the collected charge. Since the position of the particle traversing a silicon detector is derived from the centroid of the charge distribution, these secondary electrons will affect the position measurement. In Figure 4.7, the probability that the measured centroid is displaced by 1μm and 5μm is shown for tracks of normal incidence as a function of the detector thickness [ref 4.7]. Although the primary ionization increases linearly with thickness, the effect of secondary ionization due to δ electrons becomes more serious, because their range increases approximately as the square of their energy.

The effect of the Landau fluctuations on the centroid of the collected charge is particularly serious for tracks of non-normal incidence. The advantage of thin detectors becomes evident: in a 30 μm thick detector, a 10% of all tracks will have a δ electron of 8 keV or more which will result in a centroid shift of about 8 μm. In a 300 μm thick detector by comparison, 10% of the tracks produce a 100 keV δ electron that can result in a centroid shift of more than 60 μm. In practice, the advantage of thinner detectors must be weighed against the smaller primary

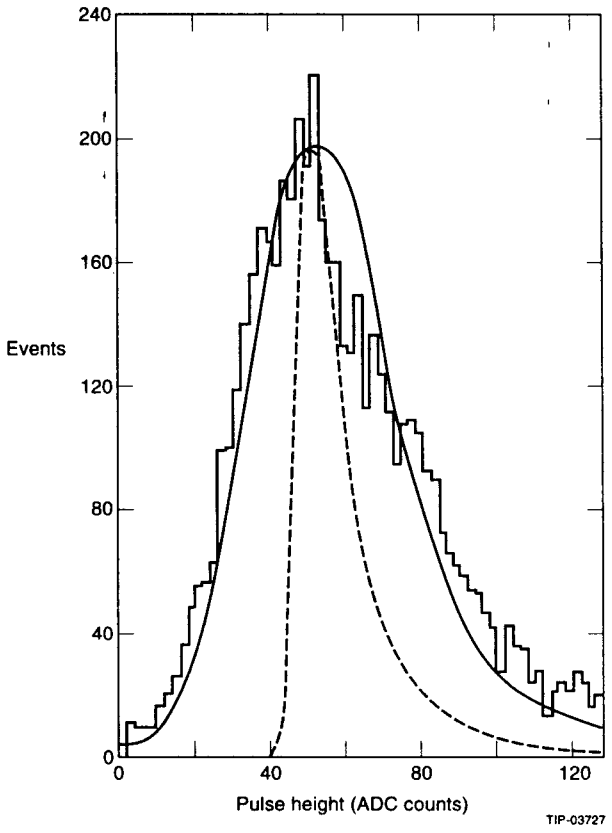


Figure 4.5 Energy loss distribution in a CCD (depletion depth of 16 μm) compared to a Landau distribution assuming free electrons (dashed line) and the calculation including atomic binding and soft virtual photons (solid line) [ref 4.7].

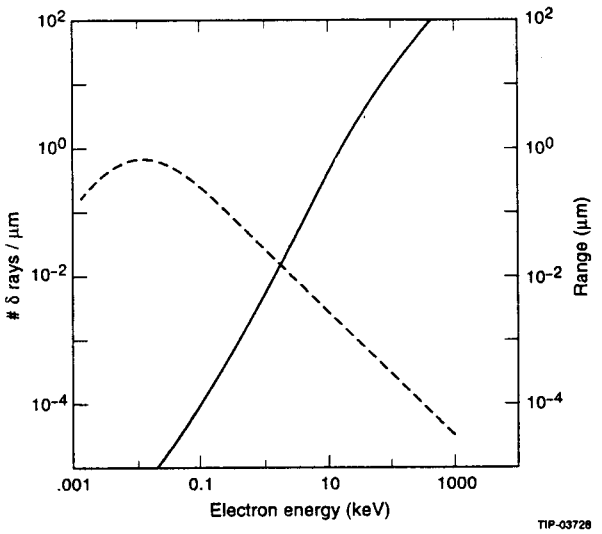


Figure 4.6 The effect of δ electrons in silicon as a function of its energy T: the number of electrons with a kinetic energy greater than T produced by a minimum ionizing particle per 1 μm path length (dashed line) and the range in silicon as a function of T (solid line).

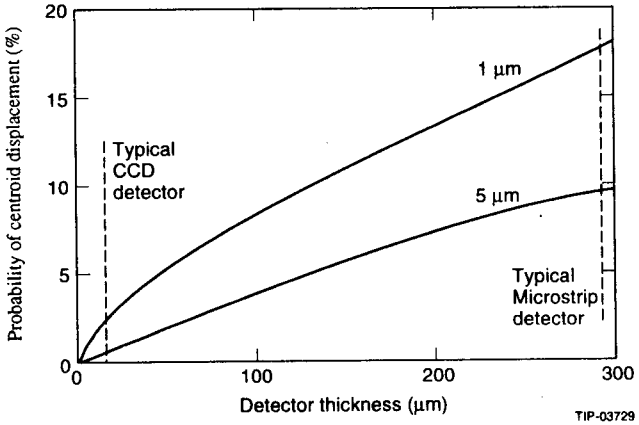


Figure 4.7 The probability for the centroid of the ionization charge to be shifted by ≥ 1 μm and ≥ 5 μm due to δ electrons produced in a silicon detector of variable thickness d [ref 1.2].

ionization and the difficulty of fabrication. Measurements of the spatial resolution of inclined tracks will be presented below.

The charge generated by an ionizing particle drifts under the influence of the electric field and undergoes diffusion. The spread of the charge distribution is proportional to the drift distance d

$$\Delta x = \sqrt{2D_e \Delta t} = d\sqrt{2kT/eV} = d\sqrt{0.026/V},$$

where $D_e = kT\,\mu_e/e$ is the diffusion constant which depends on the temperature T and the mobility of the charge carrier. At room temperature, $D_e = 34.6\ cm^2/s$ and thus for a maximum drift distance of 300 μm, the charge spreads on average by $\Delta x = 4.8$ μm.

Many silicon detectors are operated in magnetic fields. A magnetic field component transverse to the electric drift field will result in a small deflection of the charge and thus lead to a shift in the centroid of the collected charge. For a Tesla field, the deflection causes a shift in the measured coordinate by 4 μm, which can be readily corrected for.

In summary, the low ionization potential leading to a very large ionization density, the precise localization of the primary ionization, and the small electron range allow in principle for very high spatial precision in silicon detectors. The design of silicon detectors and associated electronics has to assure that these properties can be fully exploited.

4.3 Silicon Micro-Strip Detectors

Ion implanted silicon strip detectors for use in high energy physics were first developed by J. Kemmer [ref 4.8] and are now commercially available. Figure 4.8 shows the basic configuration of such a detector. The detector is made from high purity mono-crystal silicon with a resistivity ρ of the order of 2000 Ωcm. It consists of a large number of reversed bias microstrip diode junctions formed in a substrate of n-type silicon with a dopant concentration of about $10^{12}/cm^3$. These strip diodes are 8 μm to 10 μm wide and are spaced typically 20 μm apart. The bulk material is sandwiched between a thin layer of highly doped n-type silicon on the bottom and strips of p^+-implants with aluminum contacts on the top. A positive voltage is

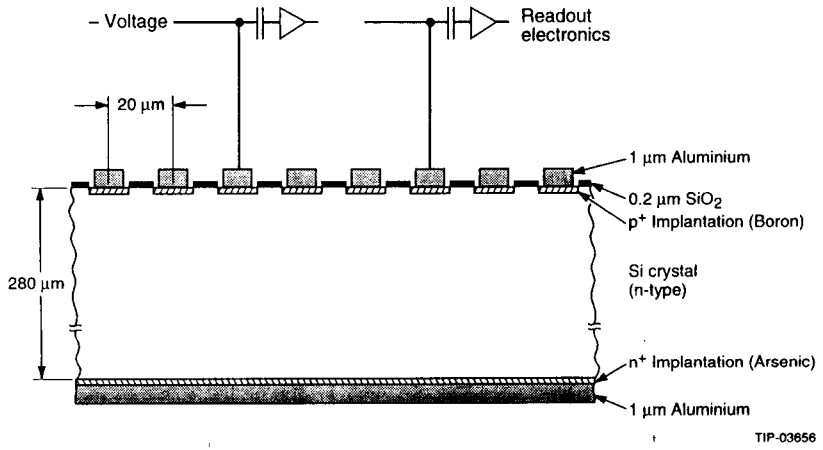


Figure 4.8 Cross section of a microstrip detector.

applied to a metal electrode covering the n^+-implant on the bottom. At the detector ends, the strips are connected to pads for wire bonding to the readout electronics.

In recent years, custom-designed VLSI circuits have been developed to accommodate the high density of the readout channels. A typical VLSI chip contains 64 or 128 channels, each comprised of a low noise charge sensitive amplifier and discriminator, and in many cases a sample-and-hold circuitry and a multiplexed digitizer. In some applications, only every nth strip (n = 2–6) is connected to readout electronics and capacitive charge division between the floating and connected strip is used to derive the position information. This readout scheme has proven to be quite efficient. For example, the precision of a 20 μm pitch detector with readout on every third strip was shown to be 4.5 μm, compared to 2.5 μm for a similar detector with readout on every channel. On the other hand, this scheme leads to a substantial degradation of the two-track resolution, and in case of a high track density to detection losses.

A minimum ionizing particle crossing the silicon detector will deposit 290 eV/μm and will create about 80 electron-hole pairs per μm. Since the energy gap is about 1 eV wide, only one third of the energy deposited is actually spent towards the generation of electron-hole pairs, the rest goes into lattice excitation. For a typical detector of d = 300 μm thickness, the total charge collected on the electrodes is

$$Q_s = 80 \cdot d = 24{,}000e^- = 3.8 \text{ fCb} .$$

The expected signal is

$$V_s = Q_s / C_{det} = 4 \text{ fCb} / 10 \text{ pF} = 4 \text{ mV} .$$

Thus to detect a signal from the detector, the capacitance needs to be kept small. A typical strip detector with 100 mm long strips and a strip pitch of 20 μm has a capacitance of less than 10 pF. The lower limit on the thickness of the detector is set by the need for a good signal-to-noise ratio for minimum ionizing particles. A thinner detector has less primary ionization and larger capacitance. It also is more fragile. The factor limiting the efficiency is the noise from the leakage current in the detector and the associated electronics. In general, the leakage current across the semiconductor junction is quite small, of the order of several nA in a high quality detector. It has several sources:

1) minority carriers, i.e., holes from the n-side which drift across the junction,
2) charge trapping centers in the depletion region that can serve as intermediate states in the energy gap, and
3) leakage through the surface channels, by far the largest source in most detectors.

As in gaseous detectors, the electrical signal on the electrodes arises from the induction caused by the moving charges, rather than the actual collection of charge itself. The pulse shape and rise time depend on the trajectory of the track, the fluctuations in ionization, the electric field, etc. The pulse length Δt can be estimated roughly from the applied voltage V_b and the depletion depth d. For example, for V_b=100 V, the average electric field is $E = V_b / d = 3.3$ kV/cm and the drift velocity is $\upsilon = \mu_e E = 1500 \text{cm}^2/\text{Vs} \cdot 3.3 \text{kV/cm} = 50 \mu\text{m/ns}$, and

$$\Delta t = d / \upsilon = d / \mu_e E = 6 \text{ ns}.$$

Thus silicon detectors are intrinsically very fast devices, permitting their application in high rate experiments, and allowing for the inclusion of signals from these devices in trigger decisions.

The passage of a charged particle is detected as a signal above threshold on one strip or several adjacent strips. The threshold per strip is typically set at two to three times the rms of the Gaussian noise distribution. The point of impact of a particle traversing a silicon detector, measured at the detector mid-plane, is derived from the centroid of the collected charge, commonly defined as the pulse height weighted average strip position of the cluster of adjacent strips;

$$\bar{x} = \frac{\sum x_i Q_i}{\sum Q_i} .$$

In many cases the readout is digital, i.e., the strips with a signal above threshold are registered but pulse height information is not retained. In this case, the centroid is simply calculated as the average coordinate of all adjacent strips with a signal above threshold. Figure 4.9 [ref 4.9] shows the pedestal corrected pulse heights versus strip number for three tracks crossing a silicon strip detector at normal incidence. The strip pitch is p = 25 μm, the detector thickness is d = 300 μm. The tracks are separated by

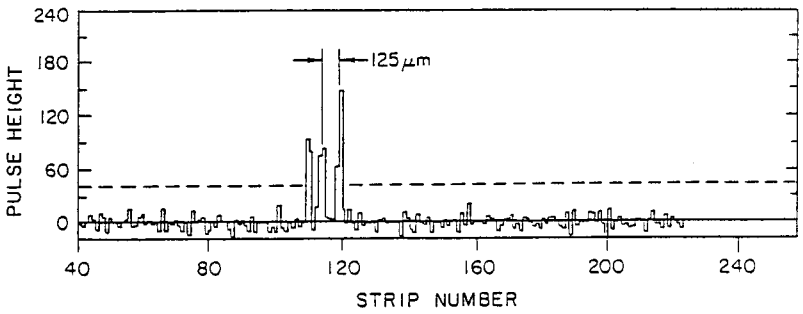


Figure 4.9 Corrected pulse height versus strip number for a microstrip detector, with three minimum ionizing particles. The dashed line indicates the pulse height threshold for a track.

about 125 μm and a study has shown that two particles can be separated down to distances of twice the pitch without substantial loss in efficiency or resolution. On average, there are 1.9 strips per track. The measured pulse height per track is shown in Figure 4.10. The most probable pulse height per track is 139 (in ADC units) compared to a single channel noise of 8. The position resolution of (3.3 ± 0.3) μm was derived from the distribution of residuals measured for high energy tracks crossing a set of three detectors.

In fixed target experiments, tracking detectors are placed transverse to the beam such that most particles have normal incidence. At colliders, a large fraction of the particles are emitted at small angles to the beam and therefore may traverse cylindrical arrays of detectors concentric with the beam at much smaller angles. For strip detectors, the ratio of the depletion depth d to the strip pitch p is generally large, and as the angle of incidence α increases the signal will spread over more and more strips. This is illustrated in Figure 4.11. With increasing angle α, the total pulse height increases in proportion to $d/\cos\alpha$, while the width of the cluster increases as $N_{st} = A + B\, d \tan\alpha / p$ (A and B are constants). This results in a smaller signal/noise ratio per strip. More serious than the reduced signal-to-noise ratio is the fact that for track angles larger than 30^o, Landau fluctuations along the track can seriously affect the centroid calculation [Figure 4.12]. These effects have been studied by several groups [ref 4.10] and different centroid algorithms have been developed to lessen the sensitivity to such fluctuations. For instance, an algorithm was introduced that defines the centroid as

$$\bar{x} = 0.5\{x_2 + x_1\}$$

where x_1 and x_2 with $N_{st} = x_2 - x_1$ determine the position and width of the cluster. The dependence of the position resolution on the angle of incidence is shown in Figure 4.13. For smaller angles, the two methods give similar results, for larger angles the simple pulse height weighting is clearly inferior.

While the original silicon strip detectors were single-sided, i.e., they had parallel diode strips measuring one coordinate, major efforts have since been devoted towards the development of double-sided detectors, i.e., detectors with strips on both sides of the bulk material providing measurements of two coordinates. A schematic cross section of a double-sided detector is shown in Figure 4.14. On both sides, the signals

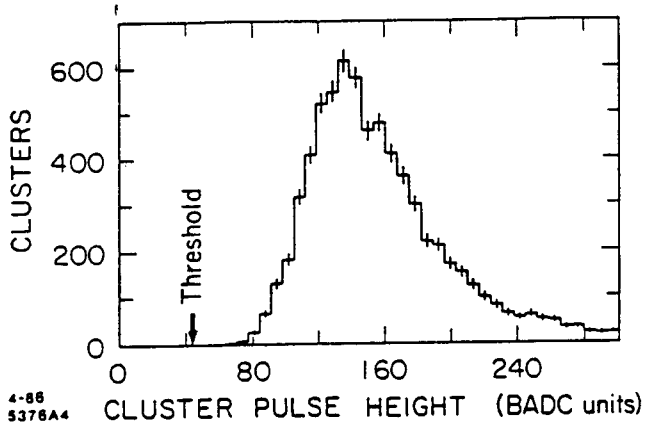


Figure 4.10 The pulse height (measured in ADC units) for tracks crossing a 300 μm thick silicon strip detector.

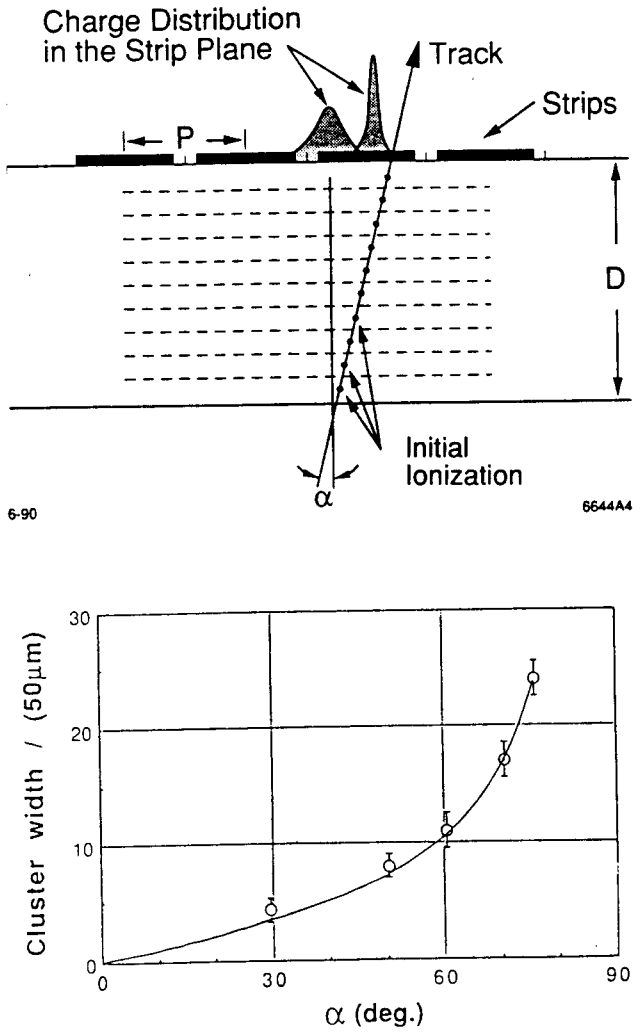


Figure 4.11 Charge collection for inclined tracks: (a) schematic illustration; (b) the width of the measured charge cluster as a function of the angle of incidence of a minimum ionizing track.

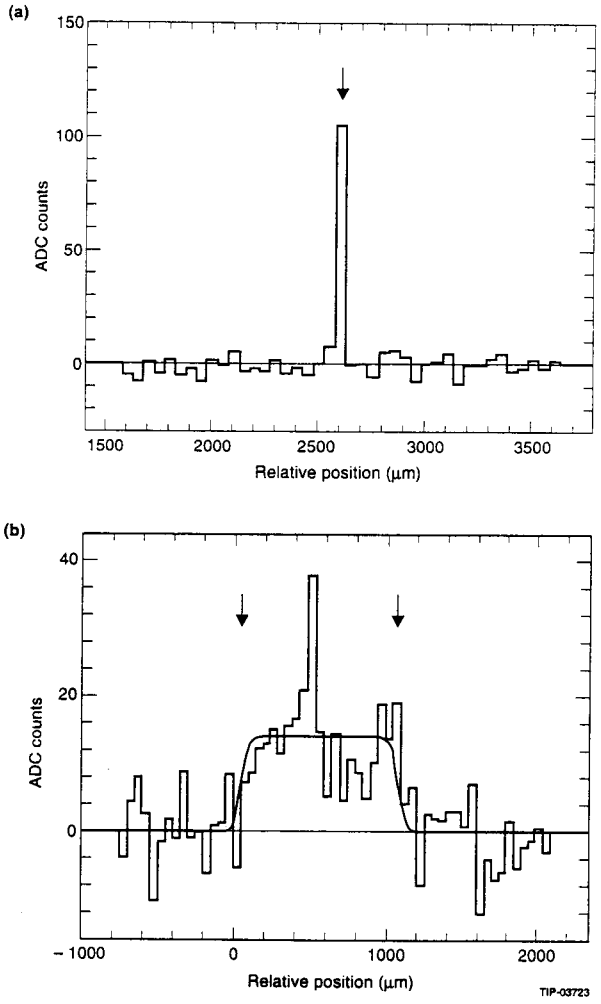


Figure 4.12 Typical charge distributions for a 300 μm thick silicon strip detector with 50 μm pitch for a minimum ionizing track crossing at an angle of (a) 0° or (b) 75° relative to the normal to the detector plane. The arrows indicate the position where the track enters and exists the detector. The curve gives the result of a fit to the charge distribution. The effect of δ rays are seen as large fluctuations in the pulse height.

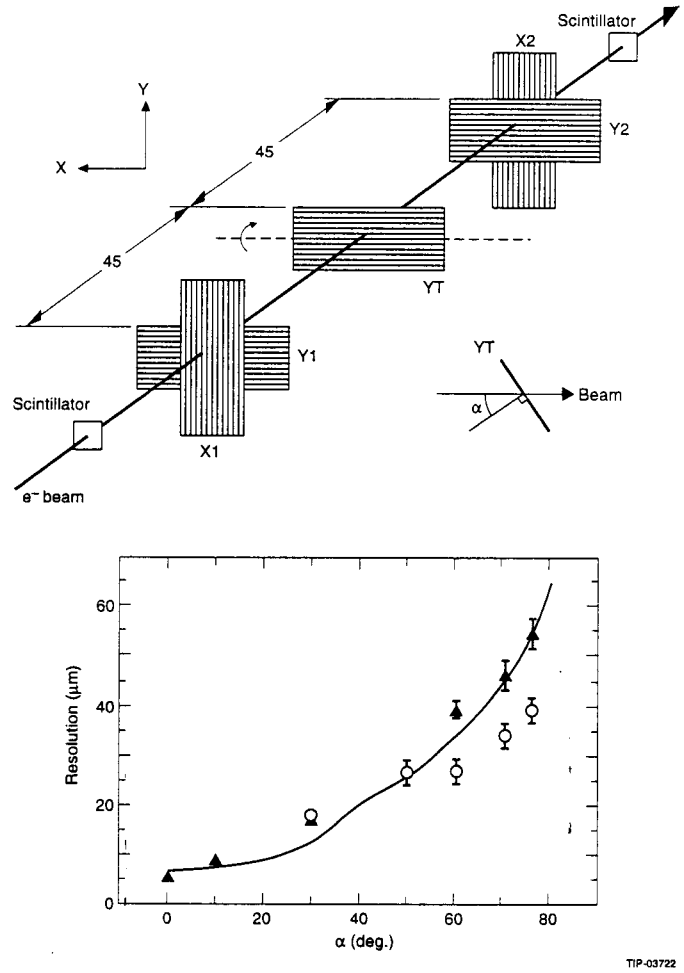


Figure 4.13 Measurement of inclined tracks in microstrip detectors: (a) beam test set-up; (b) position resolution as a function of track angle for different centroid calculation; (c) pulse height weighted average (solid triangles) and total width fit (open circles).

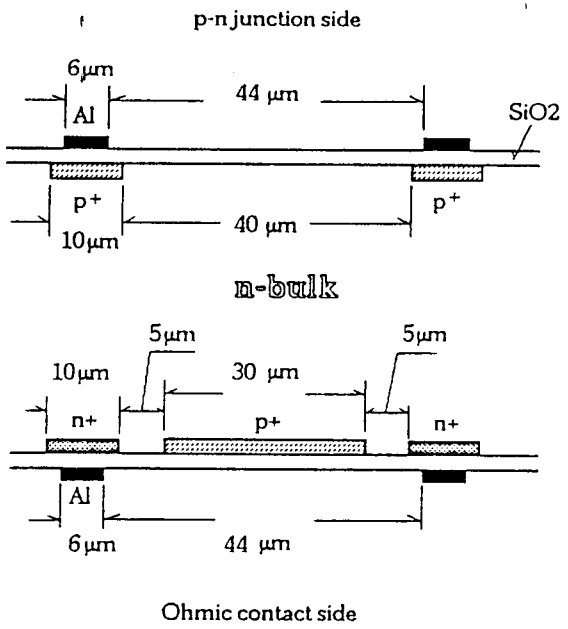


Figure 4.14 Schematic cross section of a double-sided silicon strip detector.

are coupled to the read-out through an integrated MOS capacitor consisting of the strip implant, the oxide layer and the aluminium strip on the surface. This AC coupling of the signals prevents baseline shifts and saturation of the amplifier due to large leakage currents. On the junction side, boron strips (p^+) are implanted, and on the ohmic side, phosphor implanted (n^+) strips are interleaved with p^+ strips to prevent electron accumulations at the oxide interface. The bias voltage is applied via resistors produced by an n-inversion layer at the edge of the bottom side defined by the p-implantations between the n^+ strips. The strips are orthogonal and thus allow for two coordinate readout. The first double-sided detectors strip detectors were developed for the ALEPH experiment at LEP [ref 4.11]. Several other groups have since improved the design and such detectors are now being developed commercially. In a recent test,the DELPHI group demonstated a position resolution of 8.8 μm on the p side and 11.6 μm on the n side for a 50 μm readout pitch.

Tracking systems for the SSC and LHC detectors will rely heavily on double-sided detectors and prototype tests are presently under way [ref 4.12]. These detector systems are several meters long, employ more than 10 m^2 of silicon, and have many million channels of electronics, including on-chip sparsification. Evaporative cooling systems using butane at 0°C are being designed and tested. Thin kapton cables are being developed so as to keep the total multiple scattering as low as practical. The support structure is to be fabricated using materials of high rigidity and low coefficients of thermal expansion.

The principal advantage of double-sided detectors is the read-out of two coordinates. Though these devices are not fully two-dimensional, it is hoped that the correlation between the pulse heights recorded on either sides of the detector can be used to disentangle ambiguities due to multiple hits. By combining two measurements on one piece of bulk silicon, the total material is substantially reduced. On the other hand, double-sided detectors are more difficult to fabricate because they involved many more processing steps. Up to now the yields have been substantially lower than for single-sided devices. Similarly, the assembly of the two-sided objects is far more difficult.

4.4 Silicon Drift Chambers

Gatti and Rehak [ref 4.13] proposed to reduce the number of readout channels by creating a drift field in the center plane of an n-type silicon detector. The principle is illustrated in Figure 4.15. On both sides of n-type high purity silicon, parallel p^+ strip diodes are implanted, and at the edge of the detector a single n^+ contact or a row of n^+ contacts is placed. A positive reverse bias voltage is applied to the n^+ electrode(s) creating a potential inside the bulk that is parabolic, with the minimum along the central plane. Electrons generated inside the detector by a traversing high energy particle will follow the field lines and drift towards the n^+ electrode(s). The drift time of the electrons inside the silicon measures the distance of the incident particle from the anode. If the collection electrode is divided into individual pads of several 100 μm length, then charge division among neighboring anode pads can be used to measure the second coordinate with a precision of about 4% of the anode pitch. Thus, a multi-anode drift detector can provide unambiguous two-dimensional position information. The main advantage of this concept is that fewer electronic channels are needed and that the anode capacitance is very small compared to a strip detector, and consequently, the noise performance is very good. Prototype tests have demonstrated that a position resolution of better than 5 μm can be obtained [ref 4.14]. One of the disadvantages is that the drift velocity in silicon is quite sensitive to temperature, such that a temperature change of 0.1° C measurably influences the drift times. Since it is very difficult to stabilize the overall temperature to this level, continuous calibration of the drift velocity may be needed. This can be accomplished by the injection of charge into the device from a special calibration electrode. The drift velocity is also very sensitive to the electric field and extremely high demands must be placed on the fabrication and silicon purity. The sensitivity of the drift devices to potential radiation damage has not been fully evaluated.

4.5 Silicon Pixel Detectors

While double-sided strip detectors measure two coordinates, they generally do not provide unambiguous two-dimensional information. For this reason, several groups are independently working on the development of so-called pixel detectors. They provide extremely fine segmentation and consequently have superior capabilities for

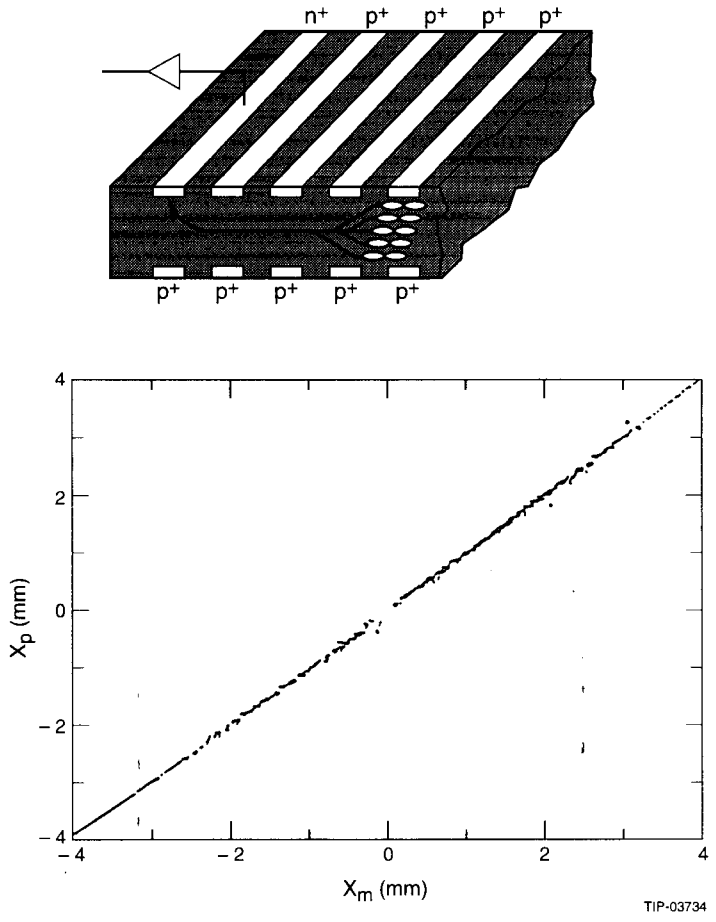


Figure 4.15 Silicon drift chamber: (a) schematic cross section, and (b) measured versus predicted position of particles crossing the detector.

pattern recognition in congested events or in high rate environments. The finer segmentation also results in smaller capacitances and thereby larger signals and less noise. This also means that the detector thickness could be reduced. Smart, selective readout electronics is required to deal with the enormous number of channels.

At present, three different types of pixel devices are under study and will be briefly described:

1) Charged Coupled Devices (CCD),
2) hybrid PIN diode arrays, and
3) monolithic PIN diode arrays.

4.5.1 Charge Coupled Devices (CCD)

Charge coupled devices were invented in 1970 at Bell Laboratories [ref 4.15] and they have since found a multitude of applications, primarily in two domains, signal processing and optical imaging. Imaging CCDs have applications ranging from TV cameras for day and night vision to detectors for optical and x-ray astronomy. Detailed information on the design and fabrication of CDDs can be found in the literature [ref 4.16]. CCDs consists of a matrix of potential wells just below the surface of a p^+ substrate. Each well constitutes one detector element and can store charge. The readout of large arrays of pixels is performed by sequential transfer of the stored charge to a single on chip amplifier, row by row and column by column. Details of the CCD structure are given schematically in Figure 4.16. Typical pixel sizes are 20 μm x 20 μm in arrays of several cm^2 area. The depletion region is very shallow, of the order of 10 μm thick. Thus only 1000 charge carriers per minimum ionizing particle contribute to the signal (electron hole pairs in the bulk material recombine readily) and this makes cooling to liquid nitrogen temperatures mandatory. On the other hand, the shallow depletion depth prevents a degradation of the resolution for tracks of non-normal incidence by δ electrons which are known to limit the resolution of much thicker strip devices.

CCDs are usually conceived as slow devices because of the sequential readout. Charge transfer rates of up to 10 MHz have been achieved while maintaining inefficiencies to the level of 10^{-6}. Thus an array of 1 cm^2 could be read out in 25 ms. With an on-chip amplifier and discriminator circuit of typically 40 MHz bandwidth and a remote main amplifier and sampling circuit rms noise levels on individual pixels of approximately 50 electrons equivalent charge have been achieved [ref 4.17]. It is expected that over the next few years both the speed and the noise performance of CCDs can be improved by another order of magnitude.

4.5.2 Hybrid PIN Diode Arrays

Hybrid pixel detectors have been successfully employed as infra-red sensors. A hybrid consists of a two-dimensional array of PIN diodes that are connected via indium bump bonds to a matched array of readout circuits, as shown schematically in Figure 4.17. An array of 256 x 256 "dumb" pixels of 30 μm x 30 μ m was recently tested in a high energy beam. A signal-to-noise ratio for a single pixel of the order of 50 was obtained and a position resolution of better than 5 μm was demonstrated for both dimensions [ref 4.18]. These pixel arrays are referred to as "dumb" because all pixels were read-out sequentially into a processor at a rate of 2.4 μsec per pixel, without any discrimination against channels without pulse height information.

For future application in high rate experiments at colliders or in fixed target spectrometers, a "smart" readout is required which allows the selection of pixels containing information and the readout of their coordinate and pulse height [ref 4.19] Thus, when a particle crosses the device the charge will be amplified and stored in the pixel and a circuit on the periphery of the chip will record the row and column as well as the time associated with the hit. Following an external trigger, the row and column information for the relevant time slot will be used to select the pixel and read out the analog signal. Design goals for applications at the SSC include a 16 ns time resolution, a readout time of 1 μs, a noise level at the order of 200 electrons and a power consumption of less than 20 μW/pixel. Radiation hard fabrication processes exist up to 10 MRad, but tests with the actual circuit have not yet been performed.

The principal advantage of the hybrid design over other pixel devices is that the detector and the readout circuit are fabricated on separate silicon wafers and therefore they can be optimized and fabricated independently. On the other hand, the bump bonds and the readout circuit add to the total material, though tests have shown that the readout drips can be reduced in thickness to 50-100 μm.

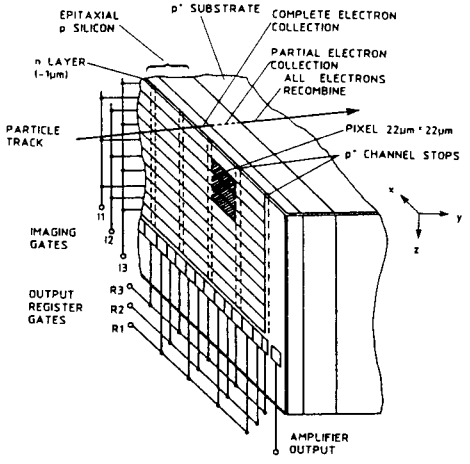


Figure 4.16 Schematics illustrating the operation of a charged coupled device.

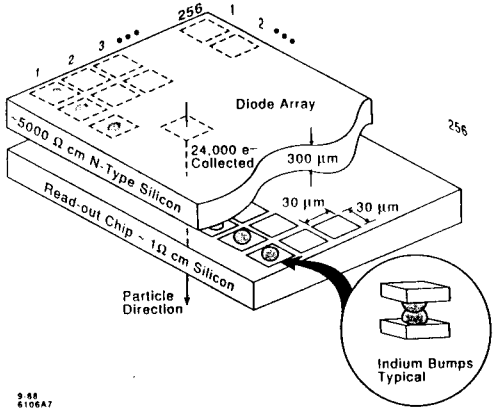


Figure 4.17 Schematic representation of a hybrid silicon PIN diode detector.

4.5.3 Monolithic Pixel Detectors

In recent years, a group from Stanford and Hawaii [ref 4.20] has pursued the development of a pixel detector with the amplifier, storage and discriminator integrated on the same silicon wafer. Figure 4.18 shows a simplified cross section of such a device. The design is based on a PIN diode with the junction formed by phosphor doped (n^+) polysilicon on the bottom of the device. On the top, an array of ohmic contacts serves as collection electrodes. The high purity p-type silicon substrate is depleted by a reverse bias voltage applied to the diode junction. Each of the contacts is surrounded by a large n-well which contains the PMOS readout electronics. This well shapes the electric field lines such that all charges in the depletion region can drift to the collection electrodes. The n-wells also shield the drift field from the switching transients of the readout circuits. The key feature of the design is that the collected ionization charge is not transferred to the edge of the chip but remains on a MOS transistor built into the pixel array. In future designs, additional electronic circuits for readout control and data compaction will be placed along the edges of the array.

A set of four small pixel arrays was recently tested in a high energy muon beam. The detectors were 300 μm thick, they contained 30 x 10 pixels of 34 μm x 125 μm dimension. The observed signal for minimum ionizing particles was on average 23 mV. The measured signal-to-noise ratio of a single channel was 55, substantially smaller than the bench test result of 143 which is still smaller than the expected performance, given the extremely small capacitance of the individual pixels.

The beam test confirmed that the signal charge can be efficiently collected over the whole area of the pixel. Of the 2665 tracks in the sample, all of them were detected in all four detectors. The spatial resolution was determined from the centroid of the charge clusters generated by the beam particles. The measured position errors for the smaller pixel dimension are shown in Figure 4.19 for tracks with pulse heights less than 2.2 time minimum ionizing. The measured distribution fits a Gaussian error function with a width of 2.2 μm, believed to be the best result to date for a semiconductor based particle detector [ref 4.21]. This precision can be attributed to the low noise performance allowing for a more accurate measurement of the charge spread over several pixels.

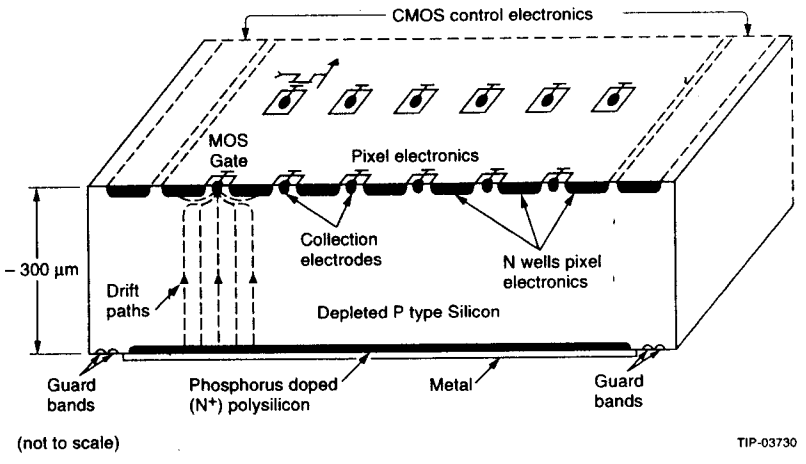


Figure 4.18 Schematic representation of the monolithic silicon pixel detector.

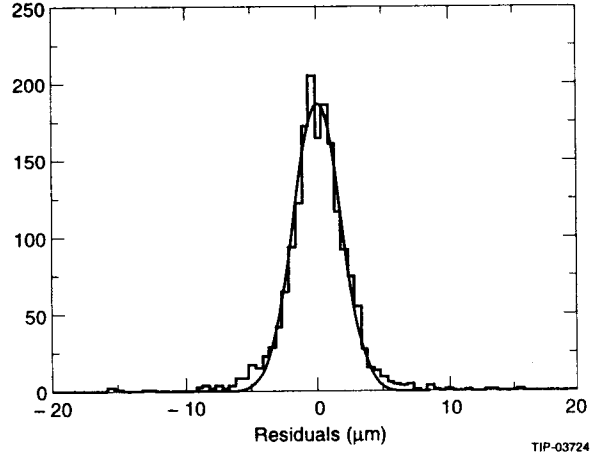


Figure 4.19 Position resolution for a monolithic silicon pixel detector with 34 μm readout pitch. The Gaussian width is 22 μm.

Future efforts will be devoted to the development of larger pixel arrays with selective readout and to the study of radiation hard design and fabrication.

4.6 Application of Silicon Detectors

Segmented Silicon Targets. As mentioned above, the application of silicon devices as active targets and precision tracking devices was pioneered by the NA-1 and NA-32 experiments at CERN. The NA-1 spectrometer [ref 4.1] was operated in a photon beam. The target was made of 40 silicon wafers (with no segmentation of the electrodes), 300 μm thick and spaced by 100 μm. It was designed to detect multiple vertices in an event by means of an increase in the ionization in the downstream wafers. The technique is illustrated in Figure 4.20. This approach was well-suited for coherent charm production in which the highly ionizing disintegration products of the target were avoided, and for relatively low energies where the primary multiplicities were moderate such that there was sensitivity to small changes in the ionization. Still, there were problems due to Landau fluctuations, secondary interactions, and photon conversions. Also, since there were two charm decays per event, the association of the decay length and the charged decay secondaries often remained ambiguous.

Microstrip Detector. Some of these difficulties were overcome by using highly segmented detectors both as active targets as well as downstream as tracking devices. This was done by the NA-32 group who operated a spectrometer in a charged particle beam at the CERN SPS. The group concluded that the active target information could not be efficiently used to select charm events at the trigger level due to problems with interactions and multiple scattering. The group subsequently decided to improve the resolution and granularity of the silicon vertex detector by adding two CCDs at distances of 10 mm and 20 mm from a 2.5 mm thick copper target [ref 4.22]. With this set-up, the tracks extrapolated to the beam interaction point were measured with a precision (in μm) of $\sigma^2 = 5^2 + (18/p)^2$, where the first term is the intrinsic resolution of the set-up, and the second term represents the contribution from multiple scattering. Typically the primary vertex was determined with a precision of $\sigma_x = \sigma_y = 2$ μm in the transverse plane and $\sigma_z = 60$ μm in the direction of the beam. This remarkable performance lead to the clean separation of secondary vertices, and to the first precise measurements of the lifetimes of the D_s meson and the Λ_c baryon. The Λ_c measurements are illustrated in Figure 4.21.

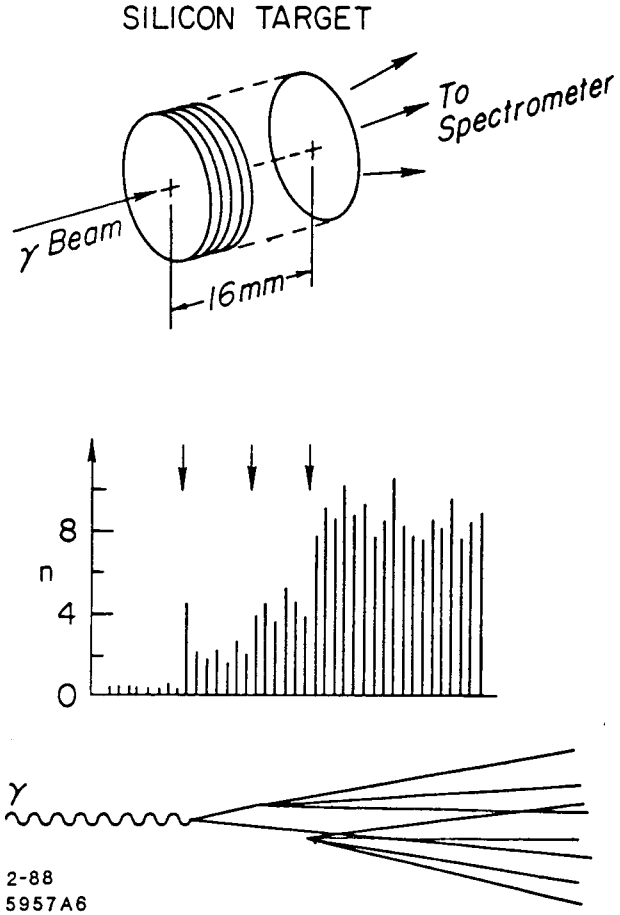


Figure 4.20 The use of silicon wafers as an active target (a) layout, and (b) pulse height versus wafer number for an event with the decay of two charm particles in the target.

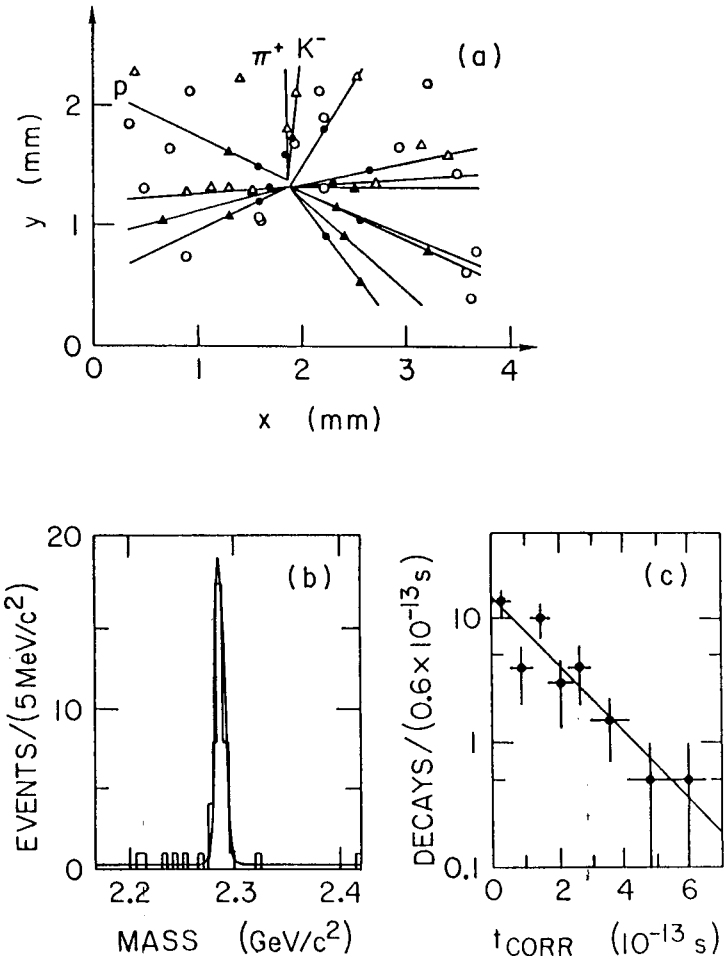


Figure 4.21 Measurement of the Λ_c lifetime by experiment NA-32 at CERN. (a) Display of a single event, detected by a pair of CCDs in the projection transverse to the beam. Hits in the first and second CCD are marked as triangles and circles, respectively. The open symbols mark hits that are not associated with this event, they are due to beam tracks passing through the CCDs during their sensitive time. (b) Effective mass of $pk^-\pi^+$, and (c) decay time distribution.

To date, the most precise measurements of the charm particle lifetimes have been reported by two experiments operating in high energy photon beams at Fermilab, E 691 [ref 4.23] and E 687 [ref 4.24]. These two groups built on the experience of the earlier experiments. They combined the high rate capability of large spectrometers with the excellent resolution of by then commercially available silicon microstrip detectors. The E 691 spectrometer is shown in Figure 4.22. Kaons were identified by two threshold Cerenkov counters, and electrons and muons were detected downstream in an electromagnetic calorimeter and behind a thick steel absorber, respectively. Three triplets of single-sided microstrip detectors with 50 μm readout pitch and digital readout were installed behind the beryllium target in the field free region. Each of the triplets had detectors with three strip orientations, X(90°), Y(90°), and V(-20.5°). The angular acceptance was about 100 mrad. Large drift chambers measured the particle trajectories before and after the dipole magnet.

E 691 recorded 10^8 inelastic interactions in a three months run in 1985. The off-line event selection was remarkably straight forward and was designed to minimize the systematic errors in the determination of the charm lifetimes:

(1) Tracks from the decay of a charm particle were required to form a good secondary vertex, all other tracks were used to fit the interaction vertex.

(2) The impact parameter of the reconstructed charm candidate relative to the primary vertex was not to exceed 80 μm (this criterion safeguarded against poorly reconstructed tracks and tracks that had undergone scattering).

(3) The assigned particle masses had to be consistent with the associated Cerenkov counter pulse height.

(4) To further reduce the non-charm background, only candidates were retained that decayed at least some distance z_{min} downstream of the primary vertex. This distance z_{min} was chosen to be 6 - 10 times the resolution σ_z, depending on the decay mode under study. The ratio z/σ_z and the proper time resolution were found to be largely independent of the charm particle momentum.

The resulting four D meson samples obtained by E 691 are presented in Figure 4.23. The proper time t was calculated using the measured momentum and the decay distance l corrected for the cut on z_{min},

$$t = l / \gamma\beta c(l - l_{min}),$$

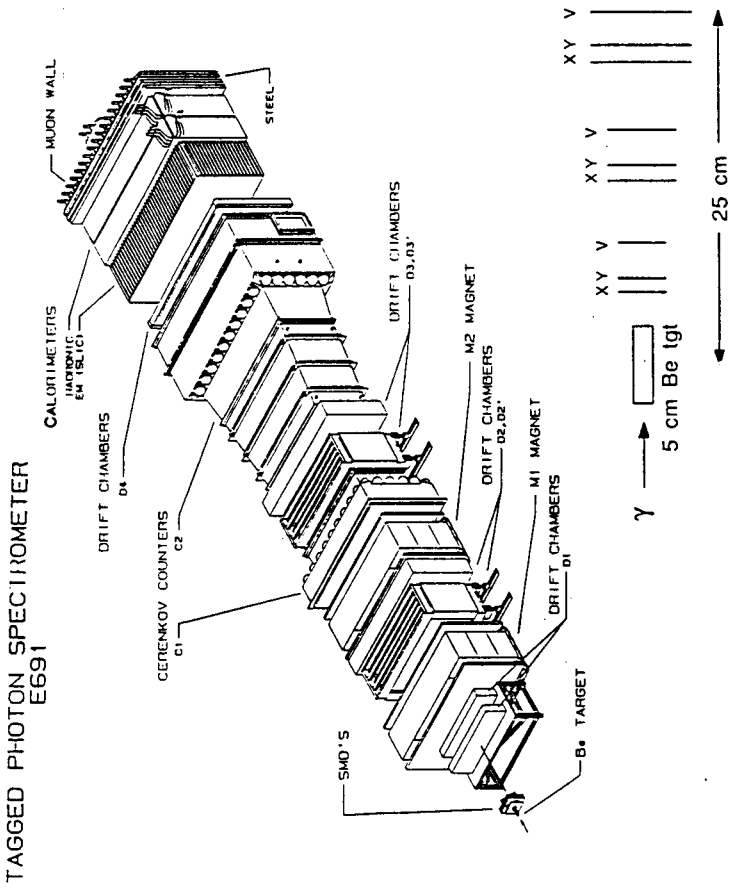


Figure 4.22 View of the Tagged Photon Experiment at Fermilab.

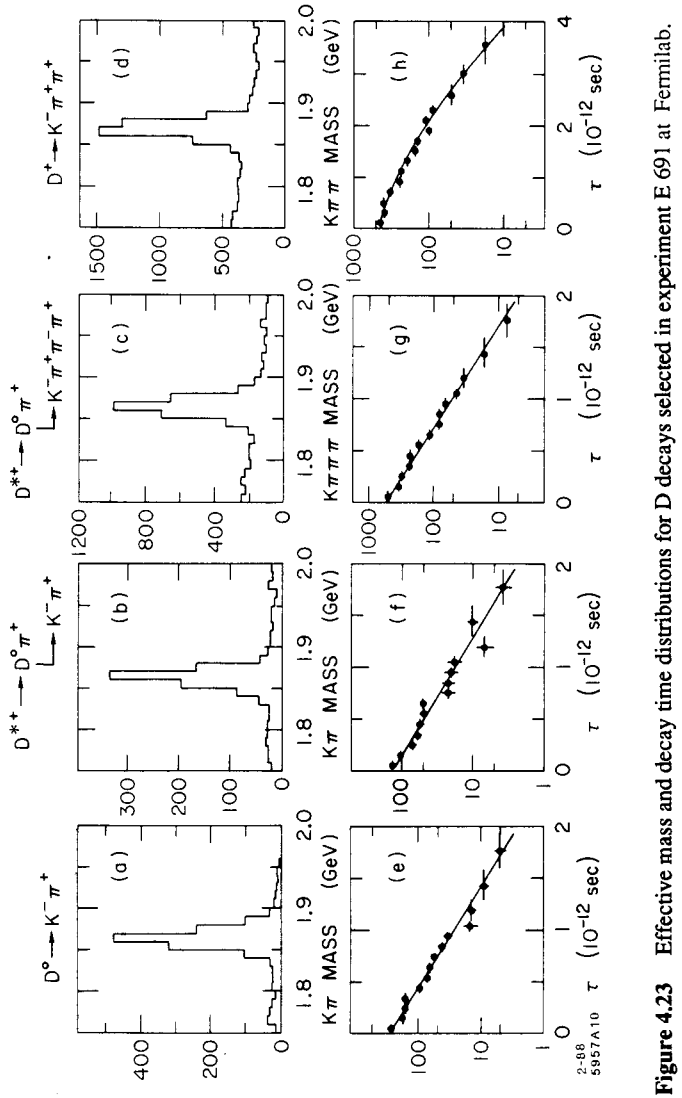


Figure 4.23 Effective mass and decay time distributions for D decays selected in experiment E 691 at Fermilab.

here l_{min} is the distance in the flight direction corresponding to z_{min}. The lifetimes τ were determined by maximum likelihood fits to the sum of the signal and background

$$N(t) = N_0 f(t) e^{-t/\tau} + b(t).$$

Here b(t) is the proper time distribution of the background, as determined from events below and above the signal in the effective mass plot. The function f(t) is obtained by Monte Carlo simulation, it corrects for the effects of acceptance, resolution, and efficiency. The free parameters in the fit are N_0, the total number of decays, and τ, the lifetime. The results of the measurements of the charm particle lifetimes are presented in Table 4.2 for both E 691 and E 687.

Table 4.2. Measurements of Charm Mean Lifetimes ($\times 10^{-13}$sec)

Experiment	D^+	D^0	D_s^+	Λ_c^+
E691	$10.90\pm0.30\pm0.25$	$4.22\pm0.08\pm0.10$	$4.7\pm0.4\pm0.2$	$2.2\pm0.3\pm0.2$
E687	$10.75\pm0.40\pm018$	$4.24\pm0.11\pm0.07$	$4.7\pm0.2\pm0.1$	$2.15\pm0.16\pm0.10$

In recent years, silicon microstrip detectors have been employed as vertex detectors at the SLC and LEP e^+e^- colliders, and at the Tevatron pp collider. All of these detectors are made of detector modules that are arranged in cylindrical layers centered on the beam. Space limitations rule out the use of conventional electronics and of the detectors employ custom designed VLSI amplifier and readout circuits. A schematic layout of the Mark II silicon strip vertex detector that was operated at the SLC is shown in Figure 4.24 [ref 4.25]. There were three layers with 12 detectors each, and 512 strips per detector, resulting in a total of 18,432 strips. The strips ran parallel to the beam axis, thus the detector measured the azimuthal angle. The signals were read out from both ends by two pairs of VLSI chips with 128 channels each [ref 4.26]. Alternate strips at each end were wire bonded to the readout chips which received power and timing signals from a hybrid circuit. A single thin cable connected the two detector ends and the hybrid to an external microprocessor-controlled data acquisition system. The individual detectors were assembled in modules that were designed to operate as independent units with their own power supply and readout. The individual modules were inserted into a support structure that was mounted on the central section of the SLC vacuum pipe. The detectors were

surveyed by a collimated x-ray beam prior to installation [ref 4.27] and tracks were used to measure the relative position of the modules to the tracking drift chambers on the outside.

Similar vertex detectors with multiple layers are presently operating in the DELPHI and OPAL detectors at LEP and in CDF at the Tevatron. To cover most of the solid angle, several detectors are connected in series by wire bonds to form strips with an effective length of up to 25 cm. So far, the only experiment that employs double-sided silicon strip detectors is ALEPH. While the resolution of the ALEPH detector has been shown to be comparable to others with similar strip spacing, occasional beam losses have led to breakdown of coupling capacitors of the p^+ strips on the ohmic side. This effect has been traced to a design flaw and can easily be prevented in future designs.

Up to now, all of these vertex detectors have been operated outside the vacuum pipe. An interesting test was recently performed at one of the interaction regions of the CERN SPS collider [ref 4.28]. In this test, silicon strip detectors were operated inside a so-called Roman Pot, an insertion in the vacuum pipe that can be placed at variable distance from the beam center. Several detectors were mounted at right angles to the beam, separated from the beam by a thin aluminized mylar foil. The test showed that silicon detectors can be operated at distances of 4 mm from the stored proton beam and that tracks from the beam-beam interactions can be cleanly recorded. A similar set-up at the SSC or LHC collider could constitute a very powerful vertex detector for particles emitted in the forward direction.

Charged Coupled Devices. The SLD vertex detector is the first device to employ CCDs on a large scale in a high energy physics experiment [ref 4.29]. It consists of 480 CCDs, with a total of 120 million pixels. Each pixel functions as an independent detector element, providing a space point measurement with a resolution of about 5 μm in each coordinate. The CCDs are arranged in four concentric cylindrical layers surrounding the vacuum pipe. This detector assembly is the result of a six year long R&D program during which the following novel construction and operating procedures were developed:

(1) the construction of a mechanically stable, low mass support system that also provides the necessary low-mass signal paths for the CCD operation;

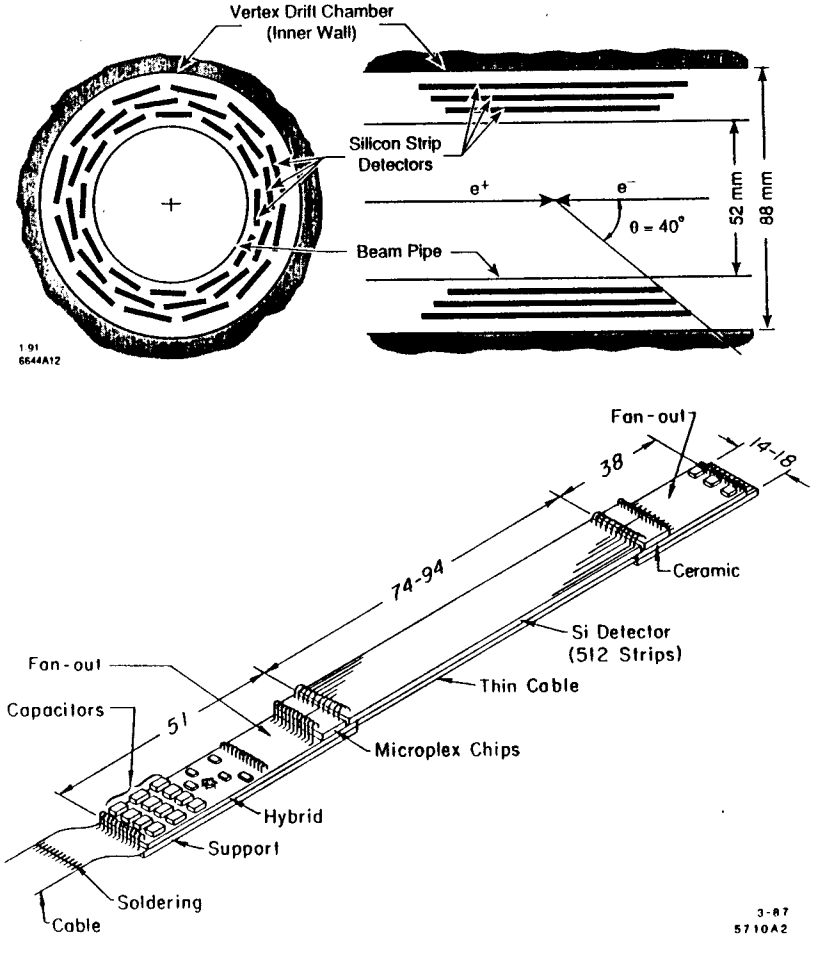


Figure 4.24 The Mark II silicon strip vertex detector. (a) A schematic layout with the vacuum pipe on the inside and the vertex drift chamber on the outside. (b) Vertex detector module.

(2) the operation at cryogenic temperature (-80°C) in order to suppress dark current and to enhance the charge transfer efficiency;
(3) the reduction of the power dissipation to 0.1 μW per pixel and 12 W total to allow for a simple gas cooling system;
(4) the modular assembly that allowed for the optical survey of the all detector elements prior to installation;
(5) the operation of a device with a 20 μm thick epitaxial silicon layer requiring extremely low noise performance of the amplifier and readout circuits; and
(6) the high speed of the readout, signal processing, and data compaction.

The operation of this detector at the SLC is a major challenge because of high background from the beam halo. Due to the high granularity of the detector and the thin depletion region, the detector was able to operate under conditions in which all wire chambers failed, and in which a strip detector would be saturated. The spatial resolution was measured for tracks registered in three or more layers, the resulting residuals are shown in Figure 4.25. Fits to these distributions imply position resolution for individual layers of 5 μm in the plane transverse to the beam and 7.8 μm in the direction parallel to the beam [ref 4.30]. Since at this time the internal geometry is defined solely by the optical survey, further improvements are expected after more precise correction for misalignment have been applied. The recognition of heavy flavor particles is expected to be greatly improved and inclusive measurements of charm and beauty production will undoubtedly lead to very interesting results.

4.7 Radiation Damage

In high energy application, silicon detectors and the associated electronic circuits will undergo damage due to radiation. Two basic mechanisms are responsible for the damage:

1) the displacement of atoms from the lattice sites, so-called bulk damage, and
2) the creation of electrons in insulation layers by ionizing radiation, leading to charge trapping at the interface between the oxide and the silicon.

At e^+e^- colliders, ionizing effects dominate and lead to interface charges, in hadron beams and at $p\bar{p}$ colliders, displacement damage dominates. The effects of ionization damage are difficult to predict and they critically depend on the details of the device fabrication. Ionization effects are proportional to the total dose and to s^2, where s is

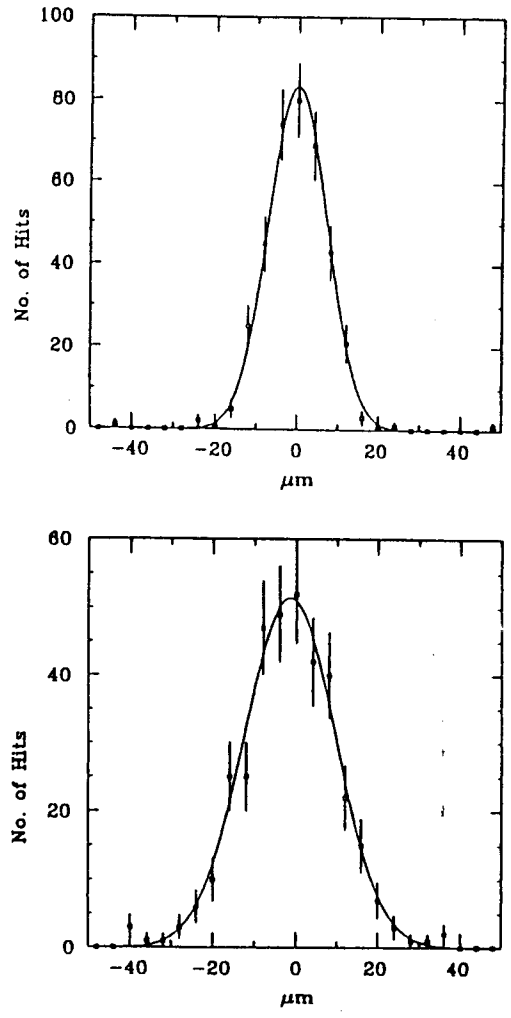


Figure 4.25 The SLD CCD Vertex Detector performance: distribution of residuals for tracks reconstructed using two of the four layers for a) transverse coordinates and b) longitudinal coordinates.

the thickness of the insulator, and thus they can be reduced by keeping the oxide thickness as small as practical.

Bulk damage manifests itself primarily in three phenomena:

a) increased leakage current,
b) change in resistivity resulting in a change of the depletion voltage, and
c) the increased trapping or recombination of signal charge.

The creation of defect states in silicon under irradiation by neutron and protons has been studied for many years.

The first effect that is observed when detectors are exposed to high levels of radiation is an increase in leakage current, $I_D = I_O + \alpha \Phi$, where I_D is the dark current per cm^3. Typical values for α are 10^{-8} nA/cm^3, e.g. for a flux of 10^{14} protons/cm^2 this would translate to an increase of 36 μA/cm^2 in a 300 μm thick detector, or a dark current of 14 μA/MRad/cm^2 in detectors of this thickness.

Two effects appear to influence the doping concentration in the bulk material: the removal of phosphor donors and the creation of negatively charged vacancy states that act as acceptors. Thus, the effective doping concentration under hadron irradiation can be parameterized as

$$N_{eff} = N_o e^{-c\Phi} - \beta\Phi,$$

where N_0 is the initial doping concentration, c has the dimension of an area and can be interpreted as cross section for the donor removal, and β represents the probability per unit path length in silicon to create an acceptor state. Measurements of the depletion voltage can be used to derive the effective doping concentration, $V_{dep} = N_{eff}\, d^2/2\varepsilon$ (here d refers to the detector thickness). Recently measurements have been extended to extremely high proton fluences [ref 4.31]. Results of these measurements are given in Figure 4.26. For a total proton flux of $1.5 \cdot 10^{14}$/cm^2, the bulk converts from n-type to p-type and the depletion voltage increases. The data fit the expected behavior of the effective doping concentration with $c = (5.5 \pm 1.1) 10^{-14}$ cm^2 and $\beta = (0.031 \pm 0.006)$ /cm.

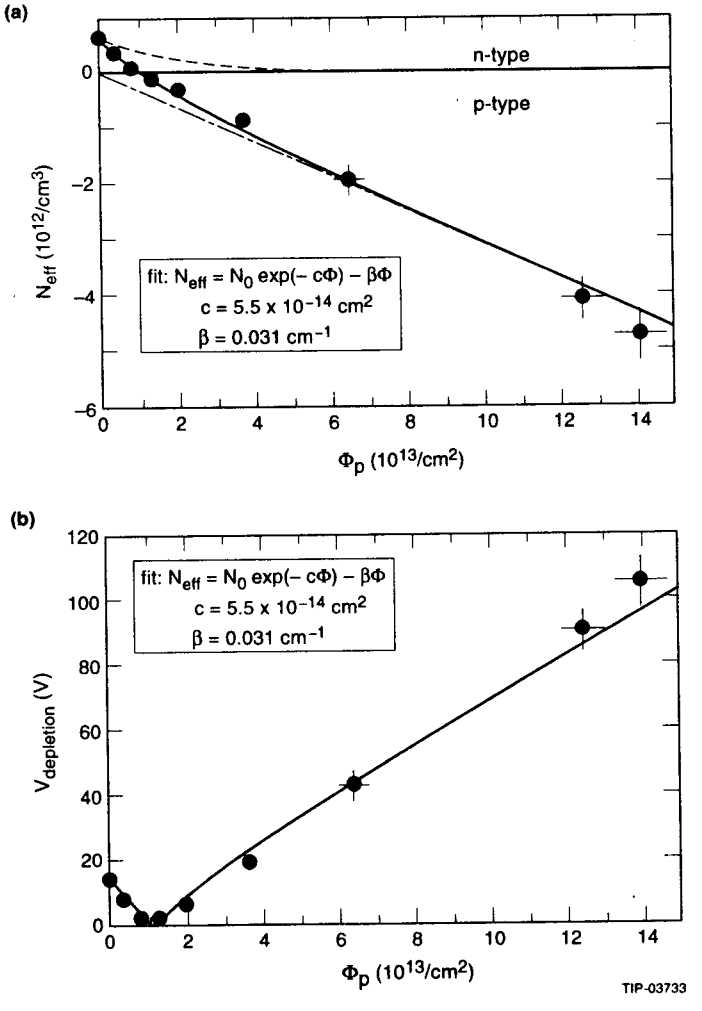


Figure 4.26 Radiation damage in silicon photodiodes by proton irradiation: (a) depletion voltage, type inversion occurs at a total flux of $\Phi = 1.1 \cdot 10^{13}$ / cm^2; (b) effective doping concentration. The dashed and dashed-dotted lines show the contributions from donor removal and acceptor creation, respectively.

The measurements indicate that silicon detectors survive at the fluence of hadrons expected at the SSC collider, provided they are placed no closer that 10 cm from the beam.

5.0 CONCLUSIONS AND OUTLOOK

Table 5.1 lists some of the critical properties of various vertex detectors that have been presented in these lectures. In fixed target experiments, active targets have lead to rather impressive results in terms of vertex resolution. Nuclear emulsions and bubble chambers have been replaced in recent years by scintillating fibers and segmented silicon targets, so as to enhance the rate capability and avoid difficult and time consuming scanning procedures. In practice, these active targets have not quite produced the results they appeared to promise. The principle reason being the multiple scattering and secondary interactions that occur in the target and lead to enhanced angular errors and confusion with decay vertices.

Table 5.1. Vertex Detectors

	Active Target	Ionization Energy (eV)	Sensitive Time (ms)	Position Resolution (mm)
Nuclear Emulsions	x	400	∞	0.5
Bubble Chambers			70-200	2.5-50
Scintillating Fibers	x	400-1000	0.002	
Drift Chambers		30	0.5	40-150
Gaseous Microstrips		30	0.01	30-50
Silicon Strip Detectors	x	3	0.02-0.20	3-15
Silicon Pixel Detectors		3	0.01	3-5

Drift chambers have been employed in a large variety of configurations as tracking detectors. They are relatively inexpensive devices of moderate resolution and granularity. They can provide a large number of position measurements with relatively little multiple scattering, thus allowing for excellent pattern recognition in one dimension. Their precision is limited by diffusion of the ionization charge, the placement and stability of the wires, the fluctuations in the drift velocity due to variations in temperature and pressure, and the inhomogenuity of the drift field. Gaseous devices operate at high voltage which may break down in high rate environments.

Silicon devices exploit the high density and the precise localization of primary ionization, the relatively low diffusion and high mobility of the charge carriers. Fabrication on the basis of VLSI technology leads to high precision and high granularity. The devices operate at low voltage with stable gain. They are also mechanically stable. The main disadvantage of semiconductor detectors is their high density leading to enhances multiple scattering. They are also relatively expensive due to custom processing on high purity materials, limited in size and thus requiring an elaborate support structure. The high density readout dissipates substantial amounts of heat so that cooling becomes mandatory in larger detector assemblies. Silicon detectors and the associated electronics are susceptibility to radiation damage.

Silicon detectors have started to have a major impact on charged particle tracking and the detection of secondary vertices. They have substantially improved the resolution and data rate capability of vertex detectors and have lead to very precise measurements of charm particle lifetimes. In the future, we expect exciting results in the area of beauty physics, among them lifetime measurements for charged and neutral beauty mesons and baryons. Enhanced capabilities for the detection of separate vertices are very important, both for the study of beauty particles but also in searches for new particles like the Higgs. Special emphasis on vertex detection will be placed by tests of CP violation in the decay of beauty particles, where the lifetime dependence of the expected asymmetries needs to be established. Such experiments are being planned both for e^+e^- B factories as well as for the large proton colliders.

In the future, two-dimensional readout will become more important for applications at colliding beam machines as well as at fixed target experiments. At higher energies, higher multiplicities and more tightly collimated jets will place more stringent demands on granularity. Major challenges for the design of larger detectors will remain to be the fast read-out with on-chip data compaction, the mechanical design of the support and cooling system, and the interconnections and transport of signals, clock pulses, and power to the individual detector elements. With increased complexity and higher degrees of multiplexing of signals and readout, performance reliability and testing attain more and more importance.

Fast on-line selection of events with displaced vertices due to weak decays will remain a challenge. Fast processors with restricted instruction sets may allow for more sophisticated decisions, combining vertex information with momentum measurements and particle identification for several tracks in a given event, and thus avoiding the relatively high losses incurred by high transverse momentum leptonic triggers. Neural network techniques are presently being tested at Fermilab.

In typical colliding beam experiments, the active area of the detectors cannot be separated from the readout, cooling, and mechanical support structure. Higher segmentation and more sophisticated electronic readout will add more material to the detector elements. Thus, future applications will place a premium on low mass materials to avoid a further increase in multiple scattering. We can therefore expect to use thin film circuits and very low mass cables with aluminium or beryllium conductors in place of the customary copper or gold. Composite carbon materials and hard, machinable foams with excellent thermal properties are likely to replace the aluminium or ceramic structures of today.

It is fair to say, that while there is an increasing number of groups participating in these developments, the efforts are not yet matched to the scale of the problems and the physics potential.

The future of vertex detection remains to be of challenge and promise!

ACKNOWLEDGMENTS

I should like to thank many of my friends and colleagues who contributed many years to the development of precision vertex detection, and who taught me many of the interesting techniques that have helped us to explore the charm and beauty of particle physics. I should also like to thank my colleagues at SLAC for the opportunity to present this material, and I should like to apologized for the short comings of this write-up. Last not least, I owe many thanks to Elberto Branson and Diane Compton for their assistance in the preparation of this document.

REFERENCES

1.1 R. Sidwell, N. W. Reay, and N. R. Stanton, Ann. Rev. Nucl. and Part. Science, **33**, 539 (1983);
C. Caso and M. C. Touboul, Rev. Nuovo Cim. **9-10**, 1 (1986).

1.2 C. Damerell, Lectures presented at the *Advanced Study on Techniques and Concepts of High Energy Physics*, St. Croix (1986).

2.1 D. Amidei *et al.*, Phys. Rev. **D37**, 1750 (1988).

2.2 S. Petrera and G. Romano, Nucl. Instr. Meth. **174**, 61 (1980).

3.1 K. Niu, E. Mikumo, and Y. Maeda, Progress Theor. Phys. **46**, 1644 (1971).

3.2 A. Adamovich *et al.*, Phys. Lett. **99B**, 271 (1981).

3.3 J. P. Albanese *et al.*, Phys. Lett. **158B**, 186 (1985).

3.4 K. Kodama *et al.*, Nucl. Instr. Meth. **A289**, 146 (1990); Phys. Lett. **284B**, 461 (1992).

3.5 M. Aguilar-Benitez *et al.*, Zeitsch. Physik **C3**, 491 (1986); Phys. Lett. **193B**, 140 (1987).

3.6 H. Bingham *et al.*, Nucl. Instr. Meth. **A297**, 364 (1990).

3.7 A. D. Bross *et al.*, Nucl. Instr. Meth. **A307**, 35 (1991).

3.8 R. Ruchti *et al.*, IEEE Trans. Nucl. Science **33**, 151 (1986).

3.9 M. Atkinson *et al.*, Nucl. Instr. Meth. **A254**, 500 (1987).

3.10 SDC Technical Design Report, April 1992, SSCL-SR-1215.

3.11 J. P. Alexander *et al.*, Nucl. Instr. Meth. **A283**, 519 (1989).

3.12 A. Oed *et al.*, Nucl. Instr. Meth. **A263**, 351 (1988).

3.13 F. Angelini *et al.*, CERN PPE-92-122 and 181; Nucl. Inst. Meth. **314A**, 450 (1992).

3.14 R. Bouclier *et al.*, CERN PPE-92-108 published in *Proceedings of 5th Pisa Meeting on Advanced Detectors*, (1991).

4.1 G. Bellini, L. Foa, and M. A. Giogir, Physics Reports **84**, 9 (1982).

4.2 E. Belau *et al.*, Nucl, Instr. Meth. **217**, 224 (1983).

4.3 S. H. Sze, Semiconductor Devices, Wiley (1985).

4.4 B. Rossi, Ionization Chambers and Counters, McGraw Hill (1949).

4.5 L. Landau, J. Physics USSR **8**, 201 (1944).

4.6 W. Allison and J. Cobb, Ann. Rev. Nucl. Science **30**, 253 (1980);

G. Hall, Nucl. Instr. Meth. **213**, 356 (1984).
4.7 R. Bailey *et al.*, Nucl. Instr. Meth. **213**, 201 (1983).
4.8 J. Kemmer, Nucl. Instr. Meth. **161**, 499 (1980).
4.9 C. Adolphsen *et al.*, Nucl. Instr. Meth. **253A**, 444 (1987).
4.10 H. Hanai, Nucl. Inst. Meth. **314A**, 455 (1992).
4.11 G. Batignani *et al.*, Nucl. Phys. Suppl, **23B**, 291 (1991).
4.12 P. Weilhammer *et al.*, CERN R&D Proposal RD-20 (1992).
4.13 E. Gatti and P. Rehak, Nucl. Inst. Meth. **225**, 608 (1984).
4.14 A. Vacchi, *et al.*, Nucl. Inst. Meth. **306A**, 187 (1991).
4.15 W. S. Boyle and G. E. Smith, Bell System Technical Journal **49**, 587 (1970).
4.16 M. J. Howes and D.V. Morgan, Charged Coupled Devices and Systems, John Wiley (1979).
4.17 C. J. S. Damerell *et al.*, IEEE Trans. Nucl. Science **37**, 305 (1990).
4.18 J. G. Jernigan *et al.*, SLAC-PUB 5925 (1992) unpublished.
4.19 G. Kramer *et al.*, SLAC-PUB 5518 (1991) published in *Proceedings of 3rd Annual IISSC Symposium,* Atlanta (1991).
4.20 W. Snoeys *et al.*, IEEE Trans. Electronics Devias (1992).
4.21 C. Kenney *et al.*, U. Hawaii, Preprint 511-741-92 (1992), IEEE Trans. Nucl. Science **34**, (1992).
4.22 S. Barlag *et al.*, Phys. Lett. **184B**, 283 (1987).
4.23 J. Raab *et al.*, Phys. Rev. **D37**, 2391 (1988).
4.24 J. Cumalat, presentation at 1992 DPF Meeting, Fermilab (1992).
4.25 C. Adolphsen *et al.*, Nucl. Inst. Meth. **310A**, 1 (1992).
4.26 J. T. Walker *et al.*, Nucl. Inst. Meth. **226**, 200 (1984).
4.27 C. Adolphsen *et al.*, Nucl. Instr. Meth., **288A**, 257 (1990).
4.28 J. Ellet *et al.*, Nucl. Inst. Meth., **317A**, 28 (1992).
4.29 C. Damerell *et al.*, Nucl. Inst. Meth. **288A**, 236 (1990).
4.30 G. D. Agnew *et al.*, contribution to the XXVI Inst. Conf. on Higher Energy Physics, Dallas (1992).
4.31 D. Prizl *et al.*, Nucl. Instr. Meth. **311A**, 98 (1992).