Silicon Tracking Detectors: Lecture 1

Sally Seidel
University of New Mexico

Sonora Workshop on Medical and High Energy Physics

22 May 2024

The role of the tracking detector is to image with the highest possible precision the trajectories of charged particles.

Their many applications in a particle physics experiment:

1. **Reconstruct vertices**: identify cases where 2 or more tracks emerge from a common point. The primary vertex indicates the initial hard interaction. The secondary vertex signals that a particle decay occurred at that vertex – typically a heavy particle: c, b, τ.

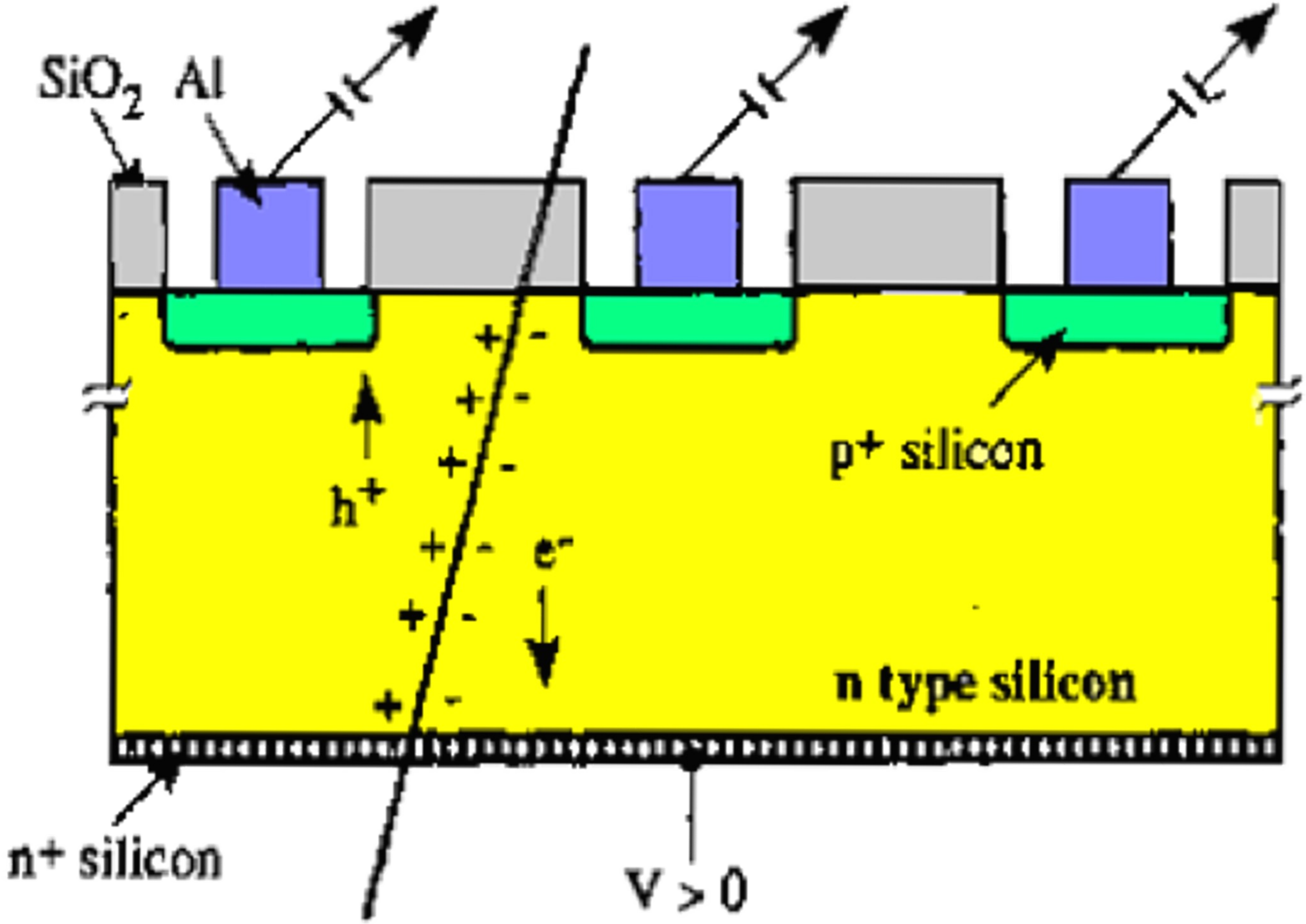
2. **Reconstruct the curvature of tracks** in a known magnetic field – to infer their momentum p. The sign of the curvature gives the particle’s electric charge.

3. **Measure track impact parameter:** gaps between vertices, or events where the momentum emerging from a secondary vertex does not point to the primary vertex – these indicate intermediate particles.

4. **Measure finite particle lifetimes** – to identify production of b-hadrons or tau leptons.

5. **Provide a trigger** for events of special interest.

The concept: ionization produced **along the track** of a particle as it traverses the detector, **drift to electrodes** where it is **integrated, digitized, read out to provide stored timing and pulse height information**.



Challenges to position resolution

The value of the tracker is in the precision it can provide. Achieving this involves overcoming **design challenges**. The variety of solutions to those challenges have led to the variety of silicon tracking detectors in use today and envisioned for the future.

Challenge #1: Seeking signs of new physics, researchers probe **increasingly rare processes.** Generating rare processes requires production of **large amounts of data**. We work to make the **rate of collisions at a facility like LHC as high as possible**.

Rate of collisions: “instantaneous luminosity” – interactions/cm^2/sec

But the possibly interesting particles produced in each collision ***also damage the detectors*** *that are placed to characterize them.*

High luminosity : new discoveries but radiation-damaged detectors

Challenge #2: **The silicon detectors are positioned as close as possible to the point of particle production** (in colliders, the primary vertex or bunch crossing point).

The motivations:

- **precision** tracking: minimum distance = minimum extrapolation error
- input to **real-time triggering** that examines tracks where their curvature is negligible, simplifying algorithms, speeding decisions

The challenges:

- **radiation damage**: fluence $\sim 1/r^n$ $(1.5 < n < 2)$
- **jets of particles are most compact near the production point**, so individual tracks are hardest to resolve at short distances – demands high segmentation

Demand for high granularity motivates the use of solid state sensors.

Solid state sensors produce electrons and holes (compare gaseous detectors: electrons and positrons).

- **Can be patterned with thousands of independent sensing elements per cm^2**, reduced range of secondary electrons in the dense substrate leads to position resolution on the order of a few microns.

- **Energy needed to ionize is low** (3.6 eV in Si, compare gas chambers at 30 eV) – so sensitivity is high. Energy resolution is good: $\frac{\Delta E}{E} \propto \frac{1}{\sqrt{N}}$

- **Signal speed is high** (~20 ns, compare gas chambers' speed, nanoseconds to microseconds).

- With very small sensing structures,

  Advantages:
  - low geometrical capacitance leads to **initially low noise**
  - high channel density leads to **low occupancy per channel, reduces event buffering needs**

  Challenges: the higher the density of the circuit, the harder it is to **read out, cool, and handle mechanically.**

Challenge #3: **minimizing multiple scattering** of the detected tracks.

- The tracker is the first detection system the particle encounters. **Anything that comes between the primary vertex and the tracker needs to be specially engineered for minimum mass**
    - this is: detector mechanical support, readout, cooling, cables and interconnects
    - and: the beampipe itself
- **It is critical to minimize the tracking detector's material itself.** Silicon has a low-Z advantage (Z = 14, compare germanium at Z = 32).

A few pages of history, and some things that were learned at each step along the way

1943 – a crystal of AgCl (4 mm thick, 4 cm diameter) **voided of free charge by cooling** in liquid air, demonstrated ionization by beta and alpha particles (P.J. Van Heerden, Utrecht)

>> the insights:

- An insulator may produce an ionization signal
- Free charge in the crystal will annihilate the signal, so suppress free charge

1950 – a germanium detector produced, **structured with a p-n junction**, depleted by reverse-bias, collected electrons and holes from alpha bombardment (K.G. McKay, Bell Labs)

>> the insights:

- Creating a p-n junction at the surface of the crystal, and applying reverse bias, can deplete the sensing volume of free charge, providing a free path for ionization produced by through-going particles.
- The applied bias across the crystal produces an electric field. The ionization particles drift along the field lines to electrodes on the crystal surface.

1955-65 – **Silicon detector development starts** (Oak Ridge, Chalk River, Harwell, CEA).
1961 – patterned, multiple rectangular diodes on one Si substrate.

>> **insight:** if the electrodes on the crystal surface are segmented, signals arriving at different segments at different times can be combined to reconstruct the path of the particle that produced the ionization.

1971 – Si **strip-like sensors** (few gold 3 mm wide strips spaced by 0.2 mm produced by a wire mask, operated at -20° C over-depleted (Karlsruhe, similar at Argonne, Fermilab, Southampton)

>> **insight**: maximize precision of track reconstruction by maximizing the number of segments – make segments as narrow and close as possible. The segments are strips that connect to amplifiers adjacent to the sensor's edge.

1973 – unstructured Si used to make **a telescope** of 5 depleted active targets (active area 300 mm^2, thicknesses 200 microns or 1 mm). Production experiment $\pi^- + \mathrm{Si} \rightarrow \pi^+\pi^-\pi^- + \mathrm{Si}$ with a 16 GeV π^- beam (CERN)

>> **insight**: the ionization can be related to the energy of the incident particle.

1978-82 NA-11 Experiment
1982-86 NA-32 Experiment
First use of silicon microstrip detectors in particle physics, motivated by goals to measure charmed particle masses and lifetimes (0.1 ps, cτ~30 microns at the CERN SPS). The 8 detectors used in NA-11 were n-doped, 2-in. diameter, 280 microns thick, strip pitch 20 microns, achieved resolution 4.5 microns, placed before and after target. (J. Kemmer, NIM A 169 (1980) 499-502; E.H.M. Heijne et al., NIM A 178 (1980) 331.)

1984: birth of the pixel concept, implemented in WA94 (OMEGA) at CERN (1990-93). This begins to shift technology challenges toward readout.

1989 - 2000's DELPHI, L3, ALEPH, OPAL, Mark-II – **silicon mini-strip and pixel detectors arranged in barrels around the interaction points** of the collider. Single-sided and eventually double-sided planar Si detectors for tracking.

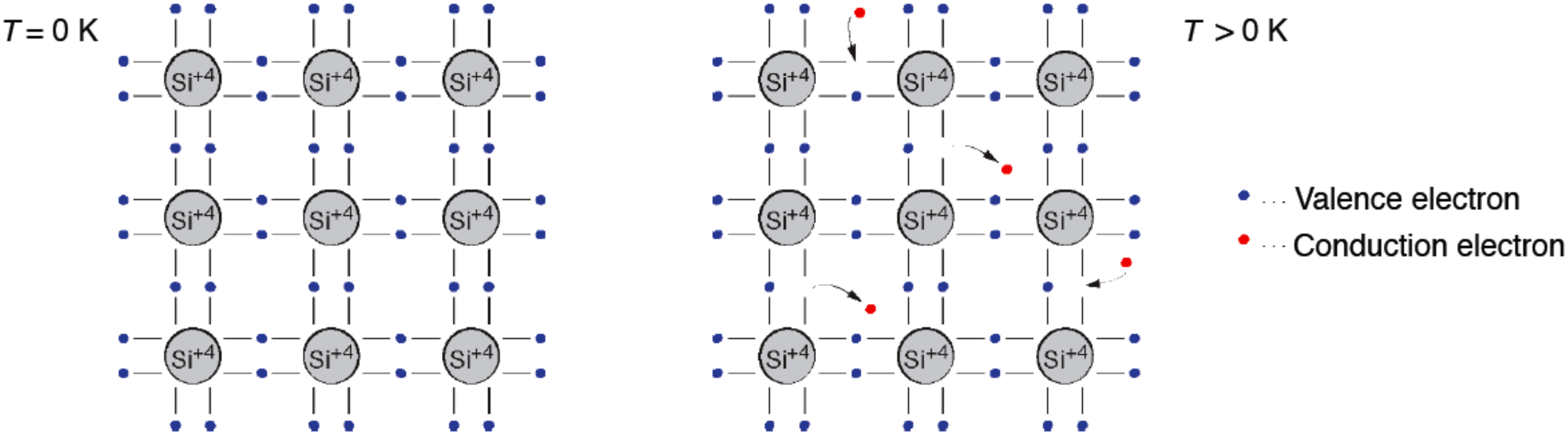
1992 – onward: CDF and D0 **Si strip detectors for vertexing, Level 2 triggering** based on impact parameters. The extended proton bunch length and short crossing time drive development of specialized readout electronics.

2005 – onward: LHC experiments implement **very large scale** pixel pattern and strip detectors

And in space: AMS, GLAST multilayer silicon strip detectors

A little bit about foundations: semiconductors and p-n junctions

Compare 2 silicon crystals at different temperatures:

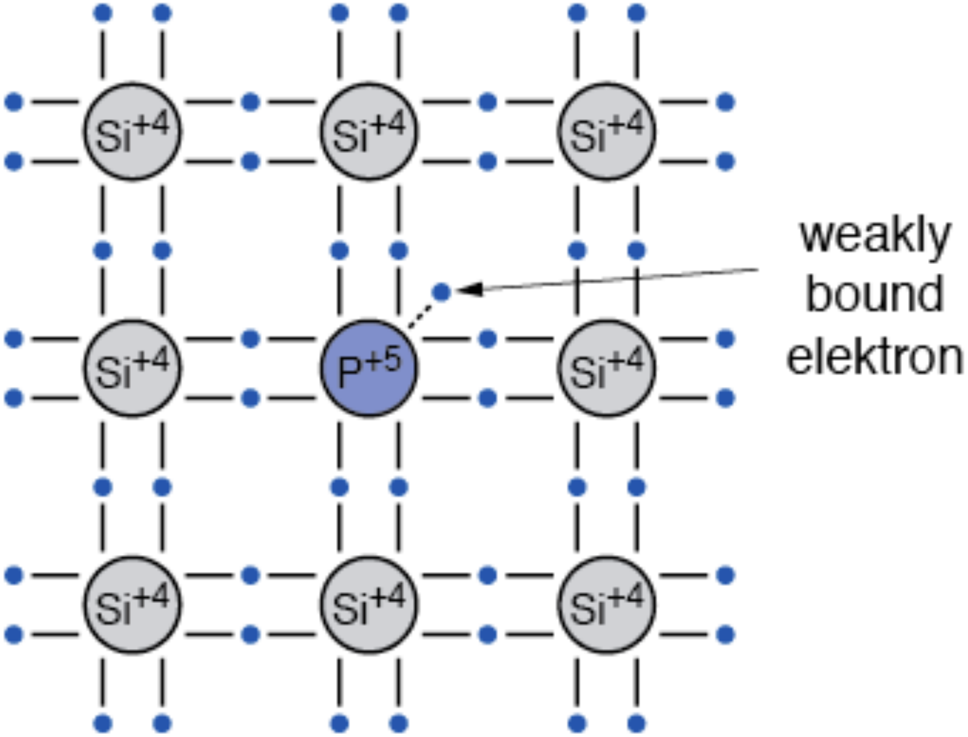


- **An electron that leaves its bond provides conduction. What remains is a "hole." When a different electron restores the bond** ("fills the hole") it leaves a new hole somewhere else ("**the hole migrates**").

- The experimenter wants free electrons to be produced only by ionization from through-going external particles. Some **ways to suppress the production of unwanted thermally-produced free electrons:**
  - lower the temperature (this strategy was used with the AgCl in 1943 and is still used for germanium)
  - choose a crystal with very high bond strength (the strategy behind diamond detectors)
  - sweep away free charge by applying a voltage across the crystal

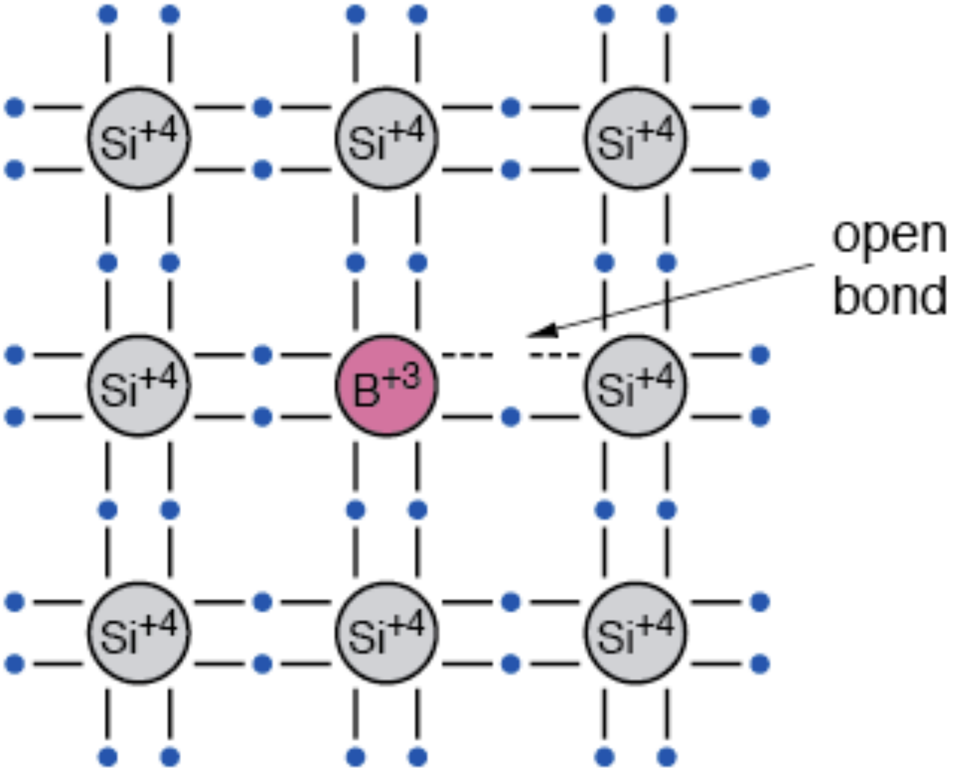
The detector starts as a pn junction. *Assembling a pn junction:*

Silicon is a Type-IV atom. Its outermost shell has 4 electrons. Replace a small percentage of the silicon atoms in a crystal with Type-V atoms (As, P), and the resulting silicon is called "n-type." *The crystal remains electrically neutral*, but one of the dopant's electrons is only weakly bound. A small perturbation frees it for conduction, leaving a hole. The dopant is called a *"donor."*



Doping is done via ion implantation + heat cure, or thermal diffusion

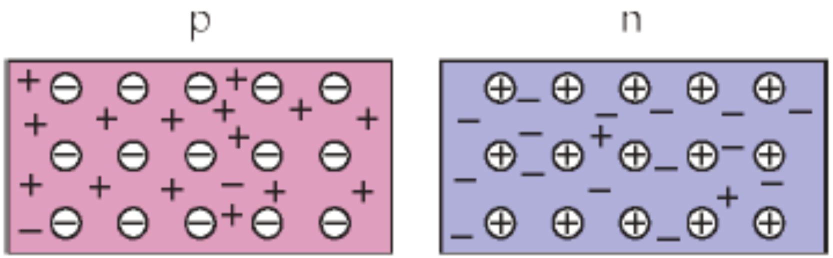
Similarly build a p-type silicon wafer by replacing a small fraction of the silicon atoms with Type-III atoms (Al, B) called "*acceptors*." Still the whole *wafer remains electrically neutral, but one bond per dopant has a hole.* It can receive an electron from another atom in the lattice. As the electron fills it, the hole migrates.

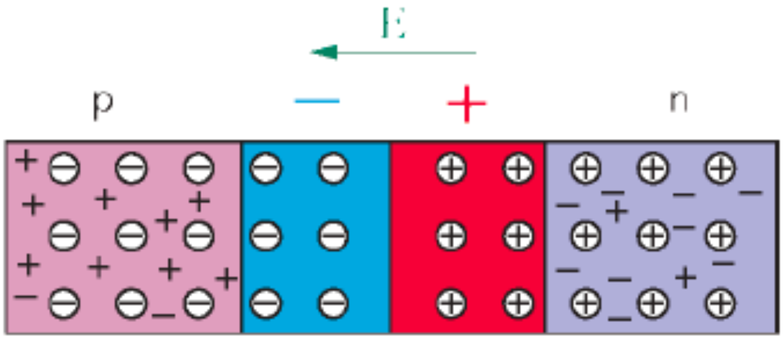


Now build the pn junction:

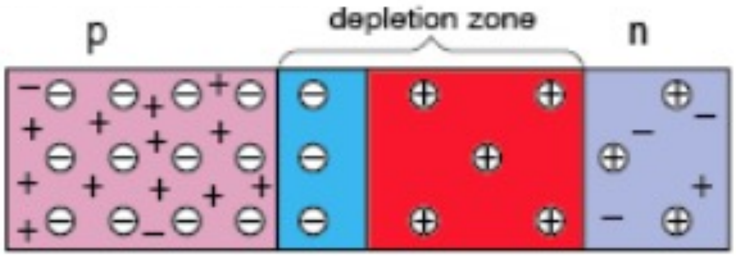
Interface the n-type ($N_{donors}=10^{12}/cm^3$) with the p-type ($n_{acceptors}=10^{19}/cm^3$) silicon. Electrons from the n-type side and holes from the p-type side diffuse across the interface until thermal equilibrium is reached. This establishes a **small electric potential across the region of the interface, called the built-in potential. It blocks further diffusion.**

Apply an external potential with – to the p-side and + to the n-side (*"reverse bias"*) **to sweep free charge out and grow the depletion zone.**
Width of the depletion zone is greatest on the n-side because of the dopant density imbalance:





Define $N_{eff} = N_{donors} - N_{acceptors}$



That depleted n-type zone is the tracking sensor.

When a charged particle crosses the depletion zone and ionizes atoms along its track, there is no free charge present to extinguish the liberated electrons and holes. They drift along the external electric field to electrodes on the opposing surfaces.

The width of the depletion zone depends on the applied voltage through the Poisson Equation:

$$\vec{\nabla}\cdot\vec{E} = \frac{\rho}{\varepsilon\varepsilon_0}$$

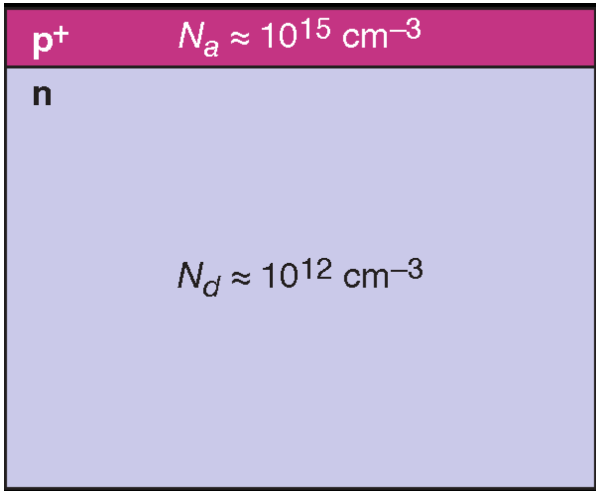
But $\vec{E} = -\vec{\nabla}V$ and $\rho = e\cdot N_{eff}$

So $$-\frac{d^2V(x)}{dx^2} = \frac{e\cdot N_{eff}}{\varepsilon\varepsilon_0}$$

E is maximized at the junction and zero at the opposite edge "w" of the depleted region. Then,

$$V(x) = -\frac{e\cdot N_{eff}}{\varepsilon\varepsilon_0}\cdot\frac{(w-x)^2}{2}$$

$$w = \sqrt{\frac{-2\cdot V\cdot\varepsilon\varepsilon_0}{e\cdot N_{eff}}} \Rightarrow \sqrt{\frac{2\cdot V\cdot\varepsilon\varepsilon_0}{e\cdot\left|N_{eff}\right|}}$$



When w = the physical width of the sensor, then V = V$_{\text{dep}}$, "depletion voltage"

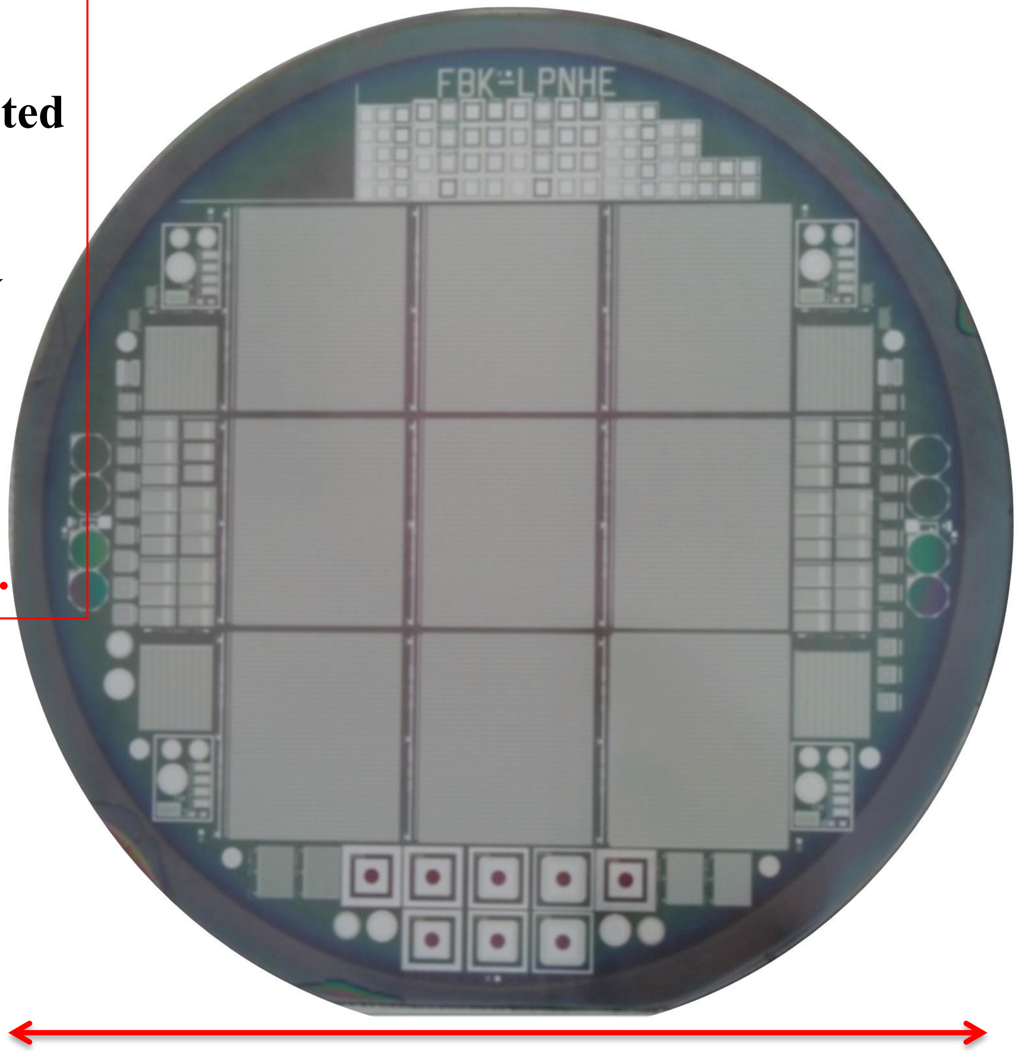
A paradigm of silicon sensor geometry, and implementation choices

The sensor is fabricated on a thin silicon wafer whose type (n or p) will define the depleted bulk.

The pn junction is formed when the other type (resp. p or n) is ion-implanted on the wafer surface.

Patterns and structure are defined by photolithography followed by wet etching or room-temperature implantation + thermal annealing.

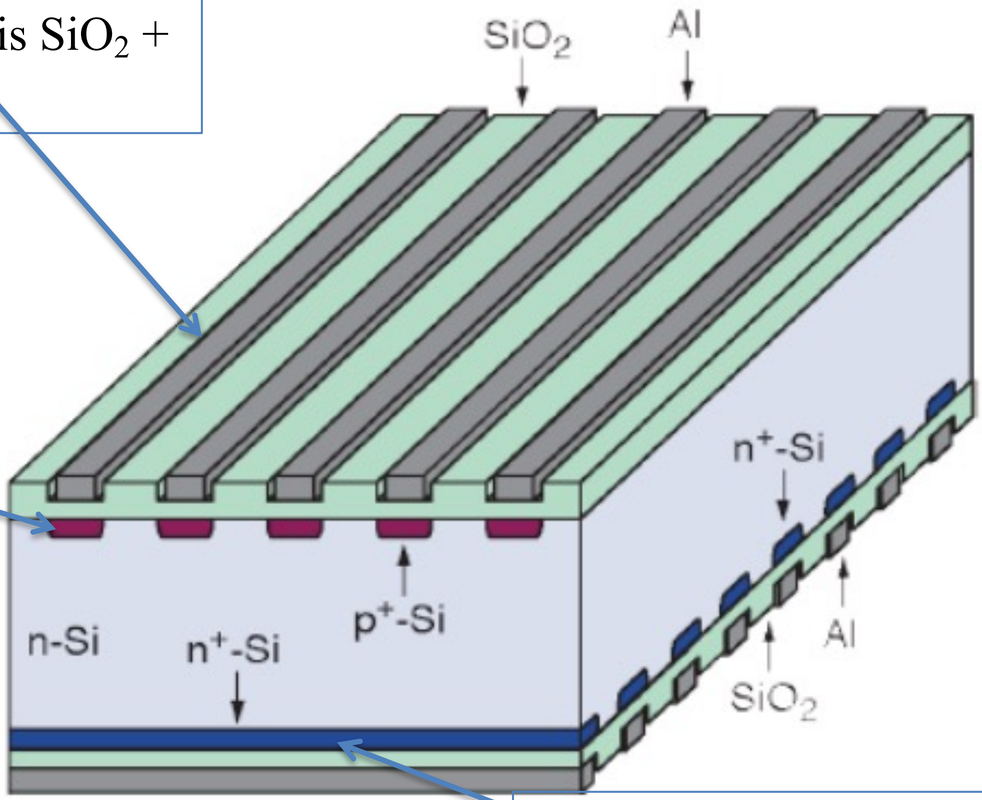
This is the basis of the planar process.



6 inches

In modern detectors, typically signal routed to read-out electronics is capacitively induced on **metal electrodes**. The capacitor dielectric is SiO_2 + Si_3N_4.

SiO_2 grows naturally on wafer surface, and electrically **isolates channels.**



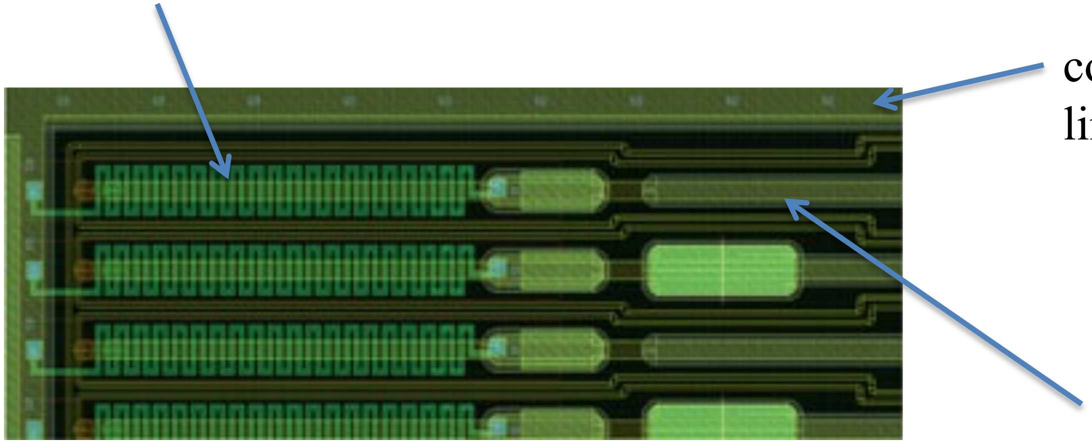
The pn junction is at the interface of the bulk with these **implanted strips** (“p^+” means $10^{18} \geq n_{dopant}/cm^3 >> 10^{14}$). Under reverse bias, the region depleted of free carriers grows from the junction toward the n^+ side (“back side”).

The **back side also takes an implant, which** can be segmented (in “double sided detectors”) or not.

Implementation issue #1:

The strips need to be isolated from each other to provide distinct charge collections
-but-
they must all be connected to a source of bias voltage.

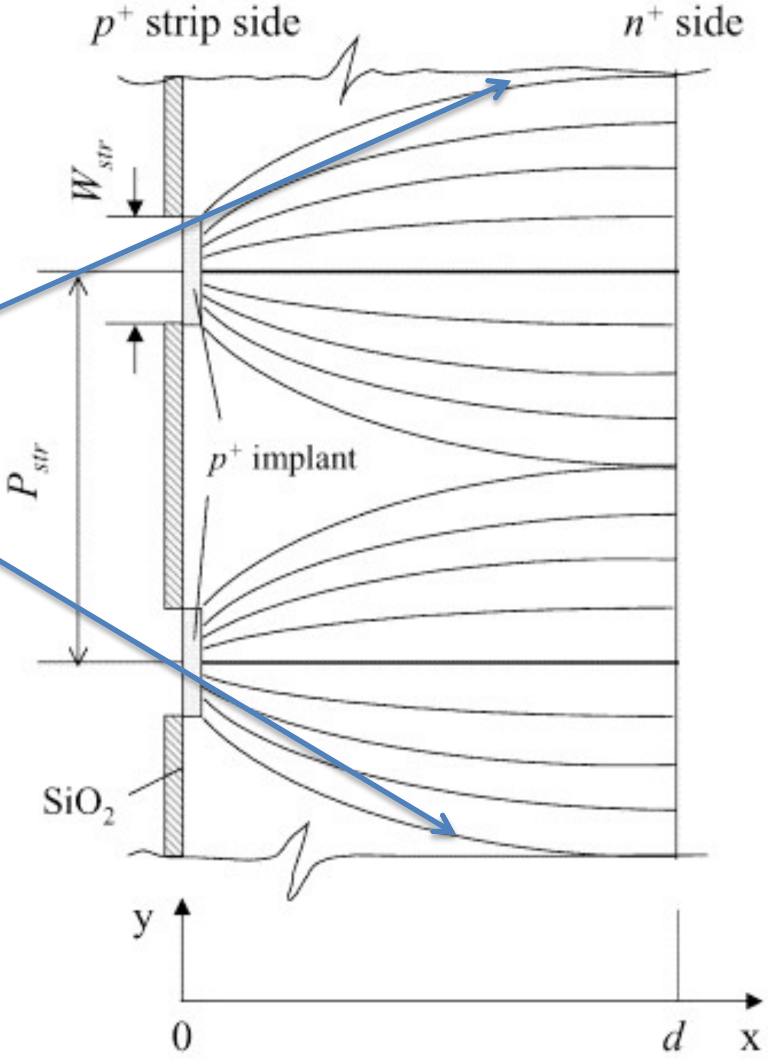
Solution: apply the bias potential to the strip through a very high resistivity (>1 MΩ) resistor (“**the polysilicon bias resistor**”)



Implementation issue #2

The process of laser cutting the sensor from the wafer produces micro-cracks and dangling bonds.

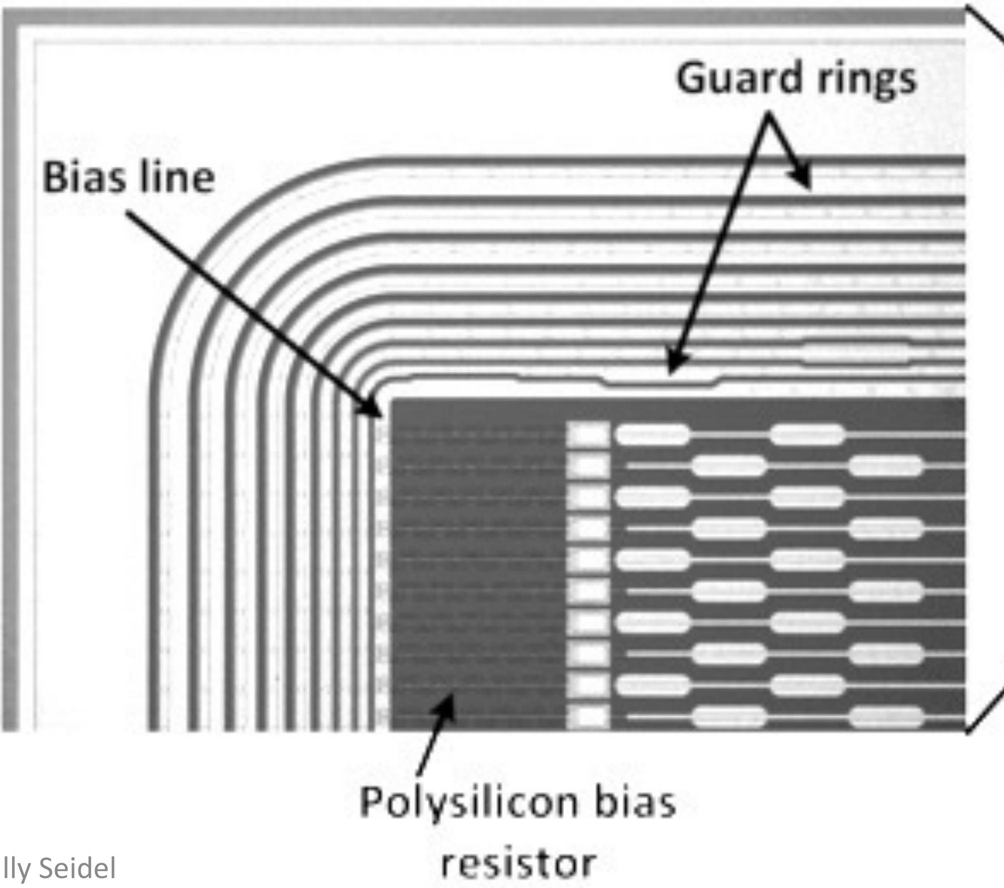
As the depletion region develops, it expands toward the cut edge, which is conductive, producing a condition unstable and sensitive to environmental changes: ***need to manage the boundaries of the electric field.***



3 solutions:

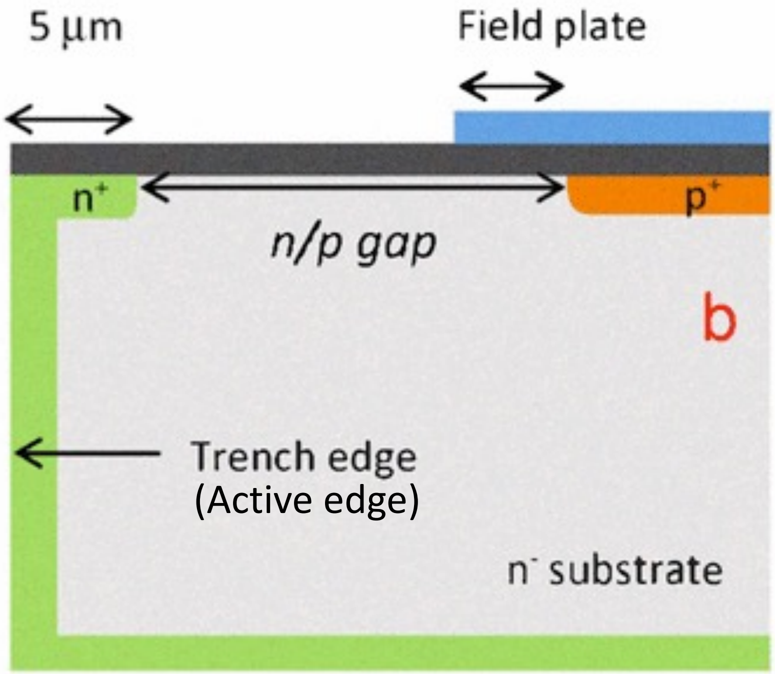
(1) On planar detectors - guard rings: **metal lines atop the oxide, and one or more ring-shaped p-n junctions that surround the sensor array but are not contacted or biased directly.**

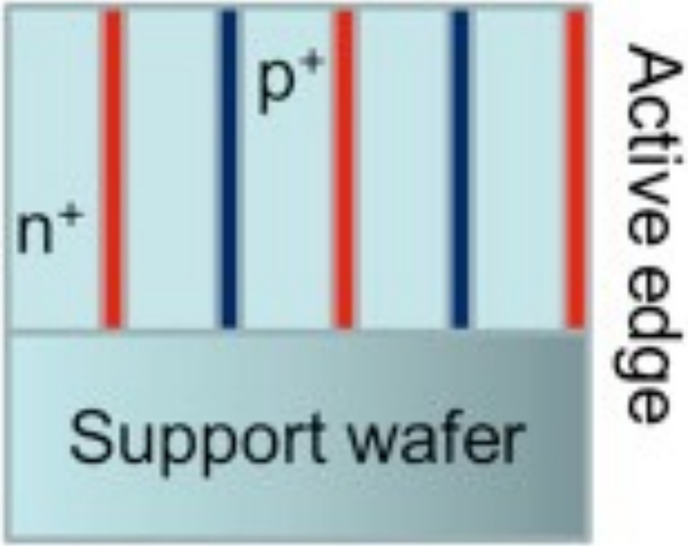
For p-implants in n-bulk: **bias the n-side, ground the active area and innermost guard.** As bias voltage rises, **depletion region expands.** When it contacts the first floating ring, that **guard charges up. Increasing V biases all the rings sequentially.** Each ring's V depends on: bulk doping concentration, inter-ring distance, and oxide charge. **The rings distribute the diode's field beyond the diode's perimeter, reducing the gradient of V at every surface point.**



(2) Active edges: a broad implant at the edge of the sensor cut face, of the same polarity as the back side doping, shapes the field and prevents it from reaching the sidewalls and providing conductive link for leakage current.

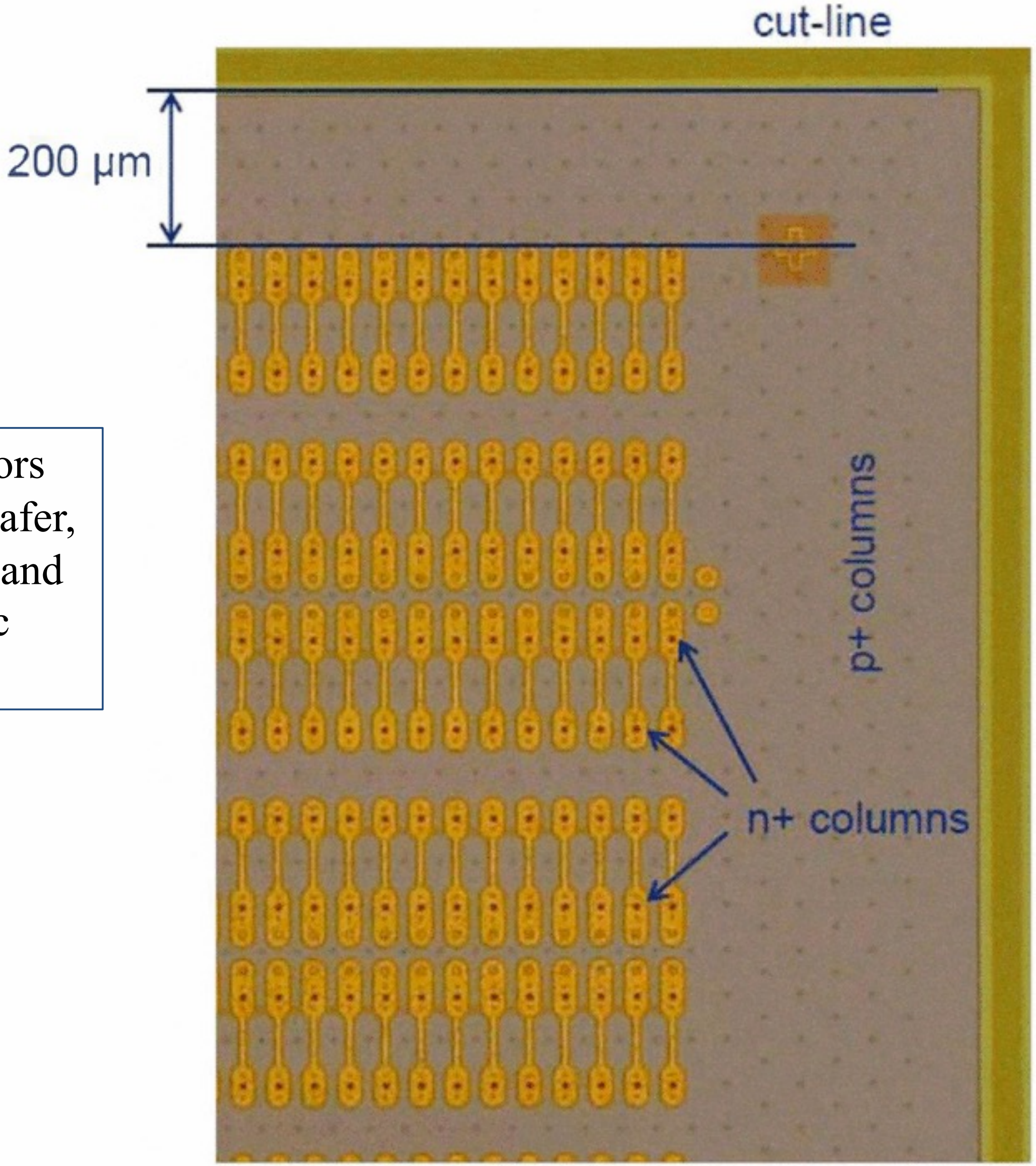
Traditional planar processing, with the electrodes parallel to wafer surfaces:





3D geometry – to be explained further – electrodes are perpendicular to wafer surfaces.

(3) Slim edges – on 3D detectors fabricated without a support wafer, a "fence" of junction columns and ohmic columns drains parasitic current coming from the edge.



Characteristics of the operating sensor

The detector records ionizations:

- A minimum ionizing particle traversing the Si loses energy at rate ***dE/dx* = 3.87 MeV/cm**.
- The mean ionization energy for silicon is **E_0 = 3.62 eV.**

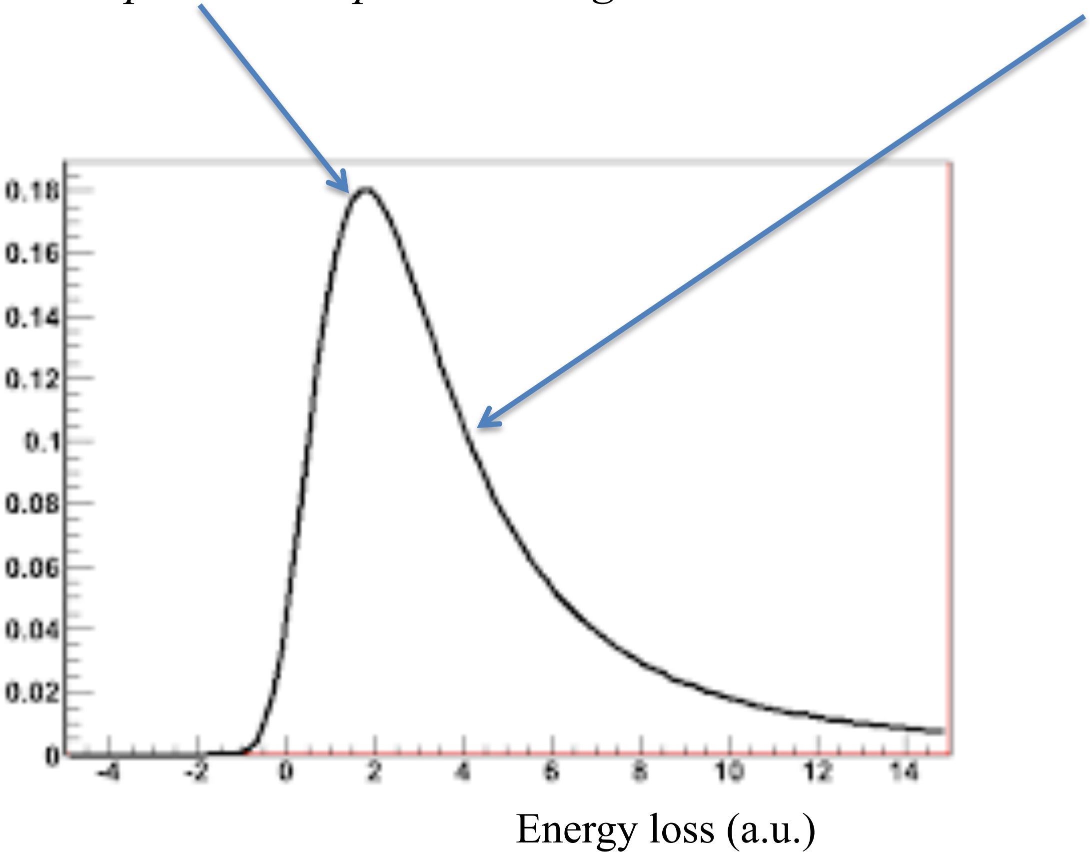
Thus: A detector of area A = 1 cm^2 and thickness *d* = 300 microns records a mean **signal**:

$$\boxed{\frac{\frac{dE}{dx}\cdot d}{E_0} = 3.2\times 10^4 \text{ electron-hole pairs signal}}$$

Notice why depleting the detector is critical for silicon:

In an undepleted, undoped ("intrinsic") semiconductor, the densities of holes and electrons are equal. In silicon at temperature 300K they are ***both* $1.45\times 10^{10}/\mathrm{cm}^3$**. Scale that to the thickness *d* = 300 microns to predict **4.35×10^8 thermal (noise) e-h pairs in undepleted silicon. Without depletion, they will swamp the signal**.

But the ionization is a statistical process, leading to a distribution in deposited charge: The *most probable deposited charge is not the same as the mean*.



This **Landau distribution** combines #collisions in a finite medium (Poisson distribution) with energy transfer per scatter (includes "straggling function" for high-energy delta-electron transfer).

For a MIP, most probable # pairs = 76/micron; mean # pairs = 108/micron.

The detector produces leakage current:

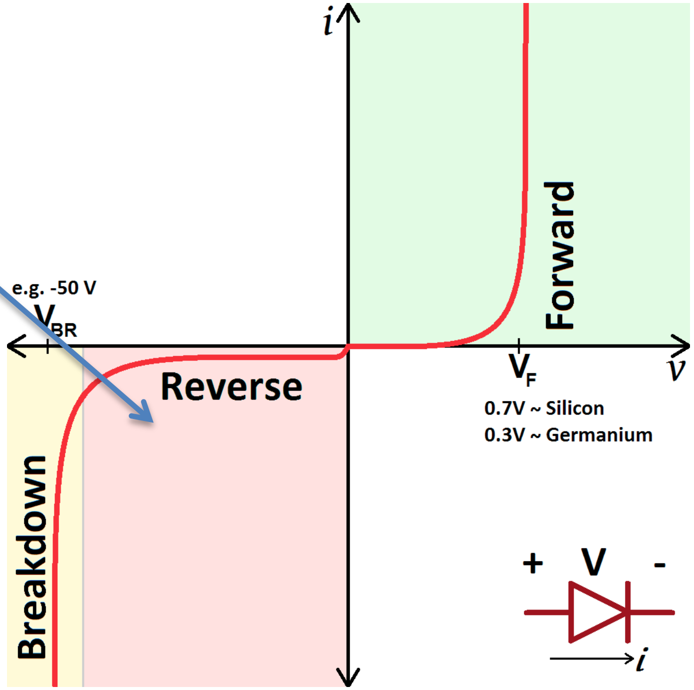
- thermal generation in the depleted region:

$$i \sim T^{3/2} \cdot \exp\left(E_{gap} / 2kT\right)$$

- and diffusion inward from the undepleted region

This current is both a diagnostic of the quality of the crystal and a source of noise.

The current versus voltage (“IV”) characteristic of a detector is typically the first thing you will measure to check that it is operating properly and to find the range of safe bias voltages.



The sensor presents a capacitance to the preamplifier:

The capacitance depends on the depth *w* of the depletion region:

$$C = \frac{dQ}{dV} = \frac{dQ}{dw}\frac{dw}{dV}$$

Recall (Slide 18): $w(V) = \sqrt{\frac{2\varepsilon\varepsilon_0 V}{e|N_{eff}|}}$, so $\frac{dw}{dV} = \sqrt{\frac{\varepsilon\varepsilon_0}{2e|N_{eff}|V}}$

For a capacitor of area A and thickness w storing charge Q,

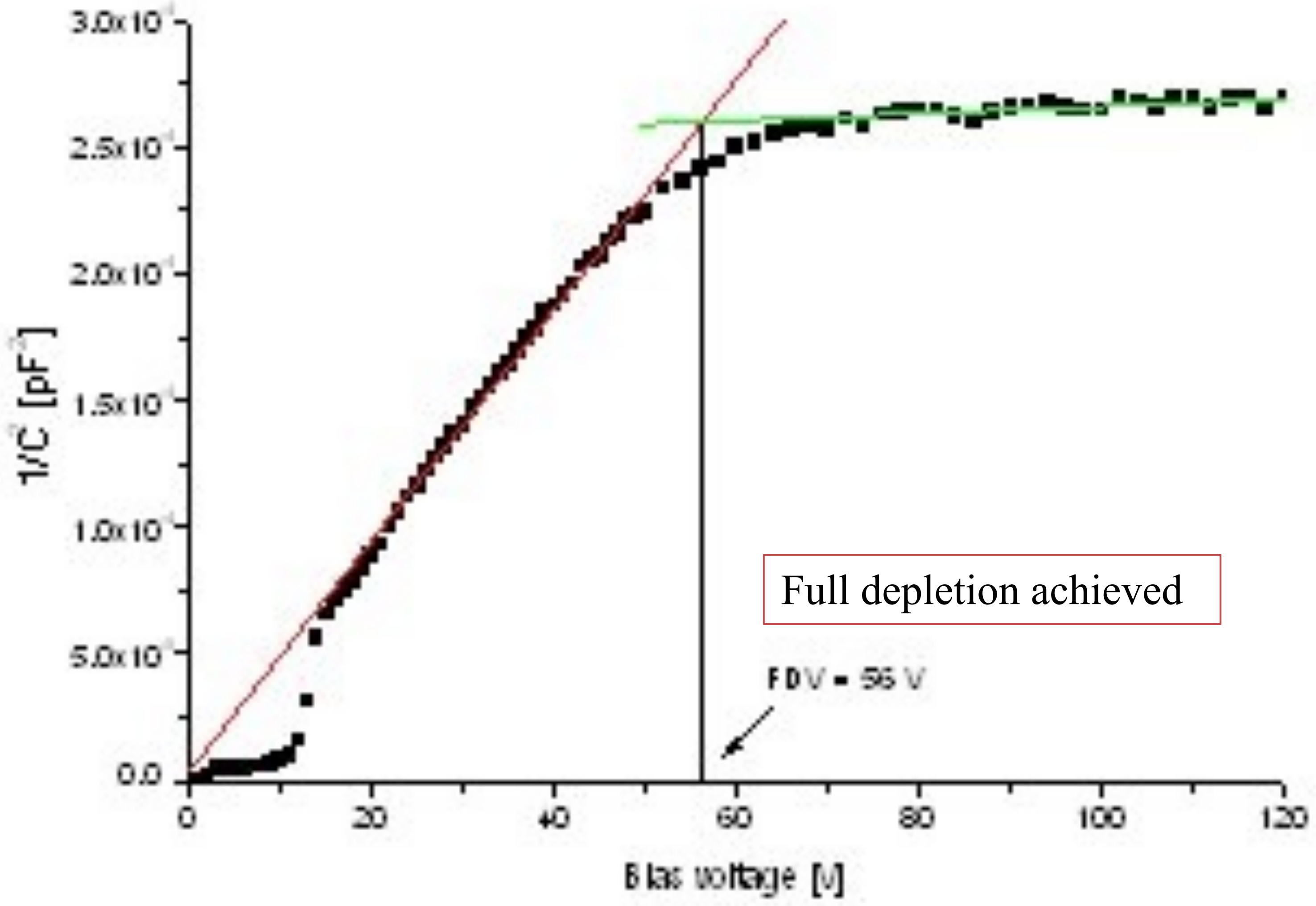
$Q = e|N_{eff}|Aw$, so $\frac{dQ}{dw} = e|N_{eff}|A$

Then $C = e|N_{eff}|A \cdot \sqrt{\frac{\varepsilon\varepsilon_0}{2e|N_{eff}|V}} = A\sqrt{\frac{\varepsilon\varepsilon_0 e|N_{eff}|}{2V}}$

Note: $\mathbf{C \sim V^{-1/2}}$

We measure the capacitance as a function of applied voltage to determine when the sensor's depletion zone has been extended to the full physical volume of the crystal ("the sensor has been fully depleted.")

$C \sim V^{-1/2}$ so, prior to full depletion, $1/C^2 \sim V$

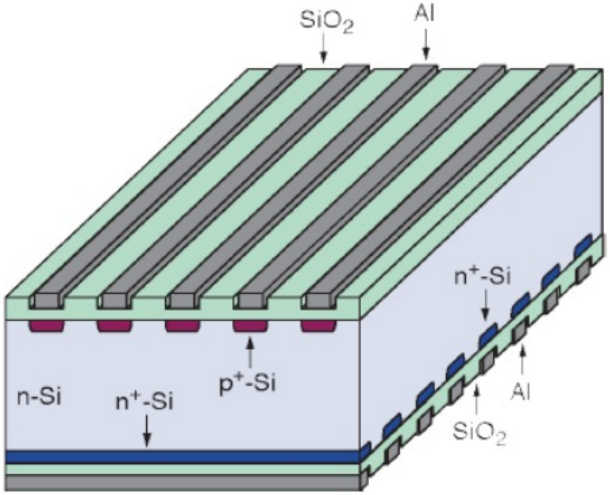


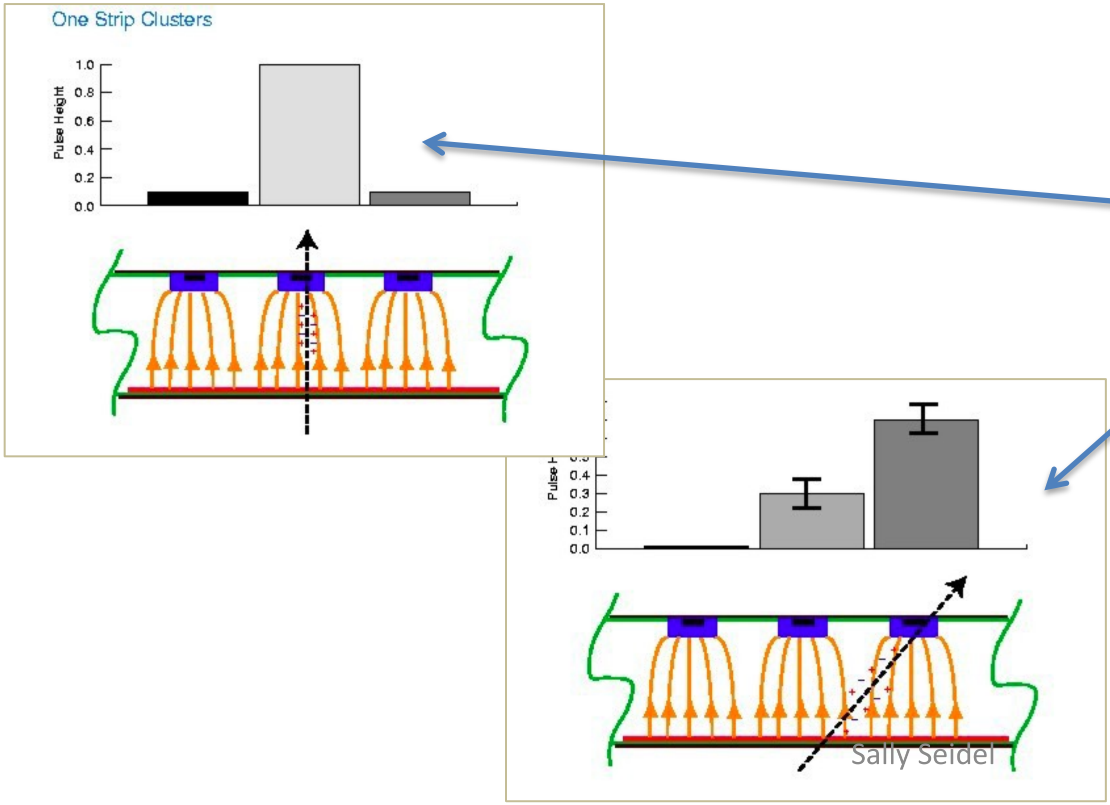
Full depletion achieved

Some useful numbers:

Modern typical **thickness is about 300 microns or less.**

For “high resistivity” silicon ($N_{donors} \sim 2.2 \times 10^{12}/cm^3$) the pre-irradiation $V_{dep} \sim 150$ V





A typical strip **pitch is p = 50 µm**.

- For binary charge readout on a single strip, the position resolution is $\sigma = p/\sqrt{12}$.
- If the charge is shared over multiple strips, with analog readout, resolution improves to $\sigma \cong p$/(signal-to-noise ratio)

The sensor can be designed to collect the electrons, the holes, or both.

The mobility of a carrier is given by $\mu = \frac{e\tau}{m}$ where m = effective mass, τ = mean time between collisions, e = electron charge.

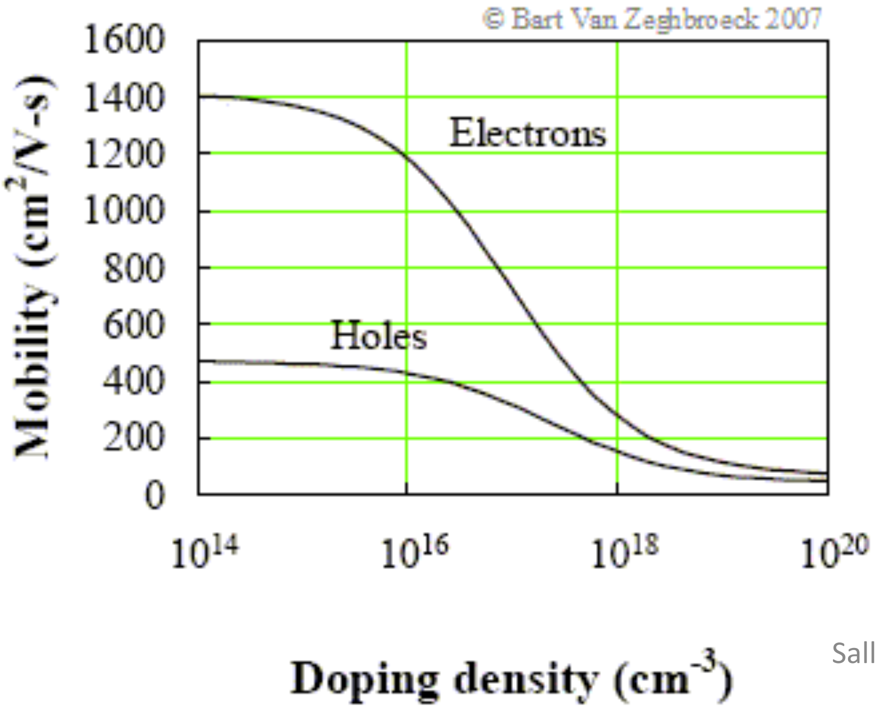
Electrons have higher mobility:

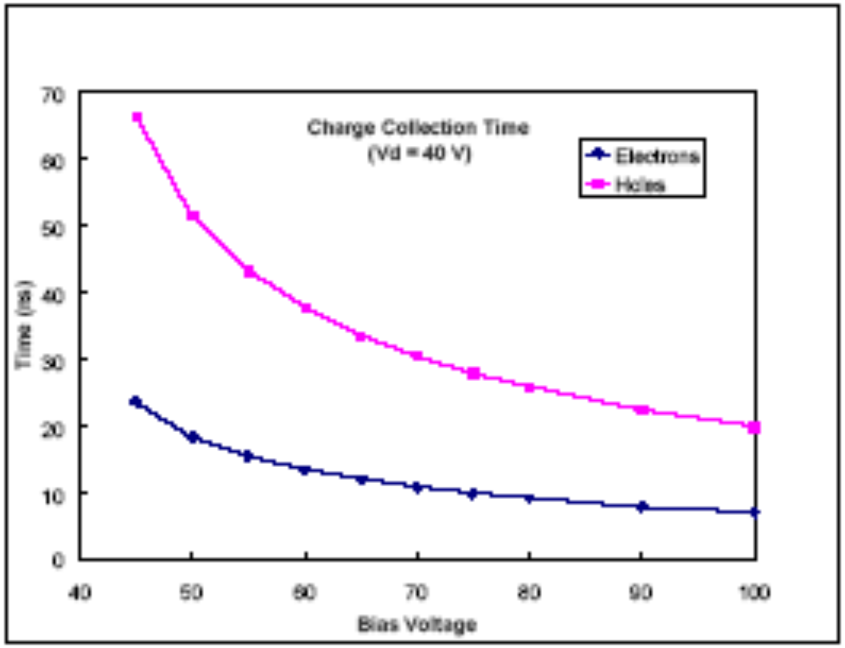
$\mu_e = 1400$ cm^2/Vs

$\mu_h = 450$ cm^2/Vs

so they are collected faster.

Their velocities depend simply on the applied field E: $v_{e,h} = \mu_{e,h} E$





Experimental goals and their impact on sensor optimization

The experimenter wants to maximize signal/noise.

Typical strip detector signal to noise ratio before irradiation: 15.

There are several **sources of noise in this sensor:**

- Capacitance C (interstrip, bulk, coupling...). Equivalent noise charge $ENC \sim C$
- Leakage current. $ENC \sim \sqrt{I_{leakage}}$
- Thermal noise in the bias resistor of resistance R: $ENC \sim \sqrt{\frac{kT}{R}}$
- Series resistance in the aluminum traces and connection to the amplifier: $ENC \sim \sqrt{R_{series}}$

To optimize S/N in strip detectors: ***minimize capacitance, minimize leakage current, maximize bias resistance, minimize the resistive connection to the amplifier.*** The contribution from capacitance typically goes inversely as the pre-amplifier integration time, so ***use long integration time*** (but this is restricted by the accelerator beam structure).

For pixel detectors the critical parameter is Signal/Threshold. Typical contemporary front end electronics threshold ~ 2000 e.

The experimenter also wants to optimize position resolution. Contributors to this:

- The shape of the Landau distribution indicates that the ionization deposited in the detector includes statistical variation. (E.g., a delta-ray production due to a hard collision with an electron will redirect the track.)

- The ionization electrons and holes *drift* along the electric field to the electrodes, but at the same time they *diffuse*. Diffusion broadens the cluster.

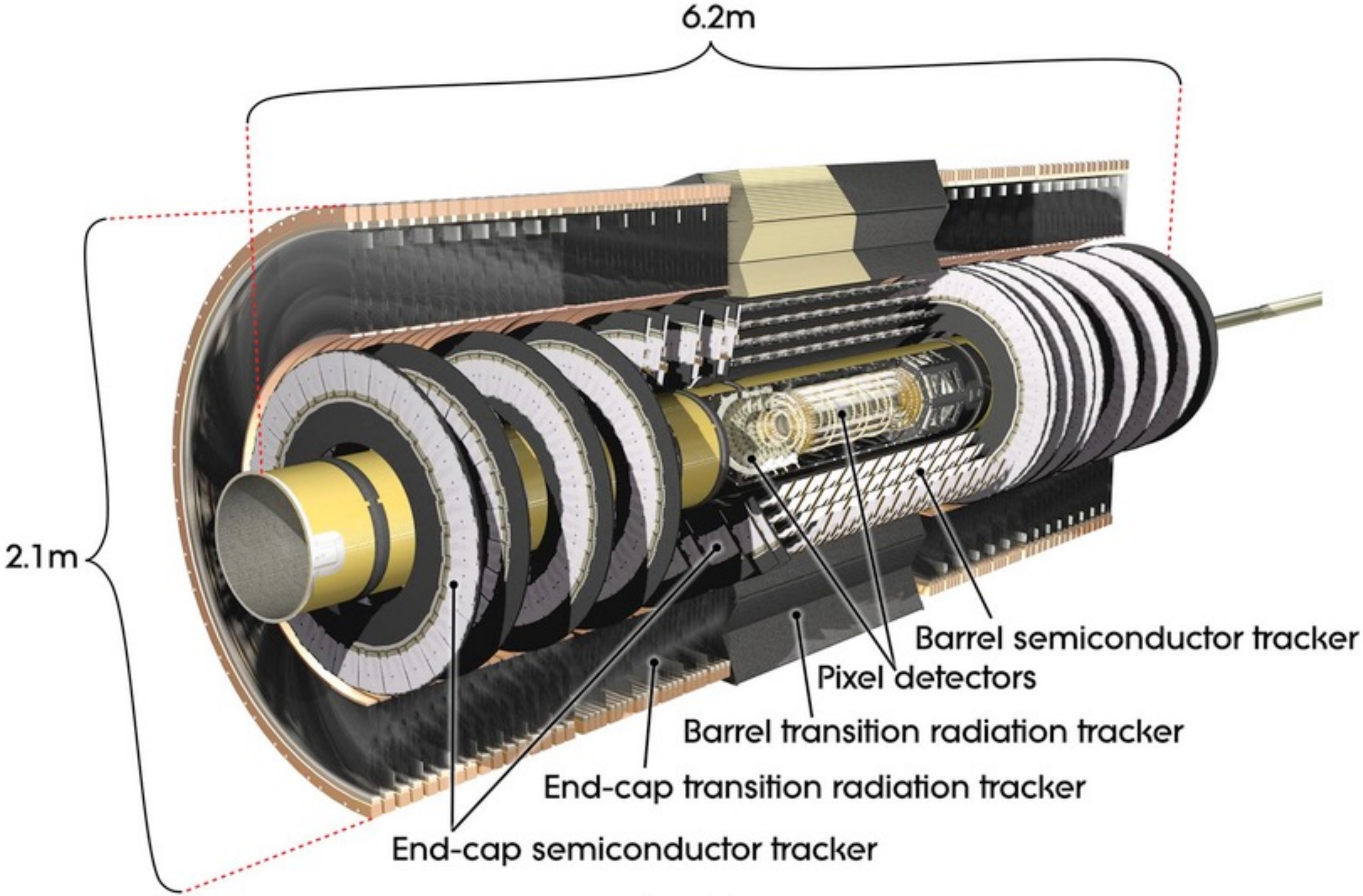
For drift time t, temperature T, mobility μ:

Diffusion coefficient $D = \frac{kT\mu}{e}$

Width of the ion distribution $\sigma = \sqrt{2Dt}$

If multiple channels are hit due to diffusion, and analog readout is employed, resolution is improved.

- Material in the detector: minimize thickness to reduce multiple scattering.



Conclusions from Lecture 1

Silicon detectors offer precision measurements at distance scales not previously resolved in particle physics experiments.

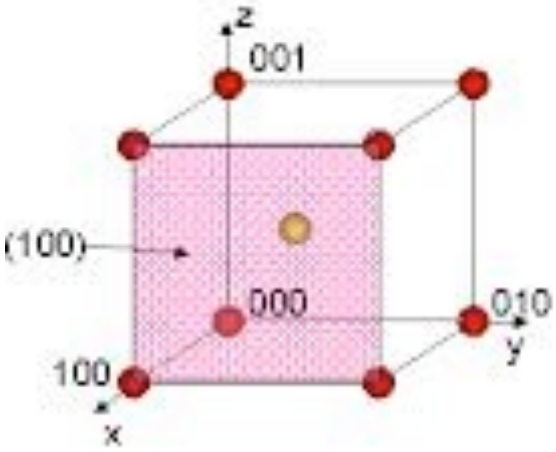
Although this technology has been developing for decades, there is no limit yet in sight to its applicability and potential for further improvement.

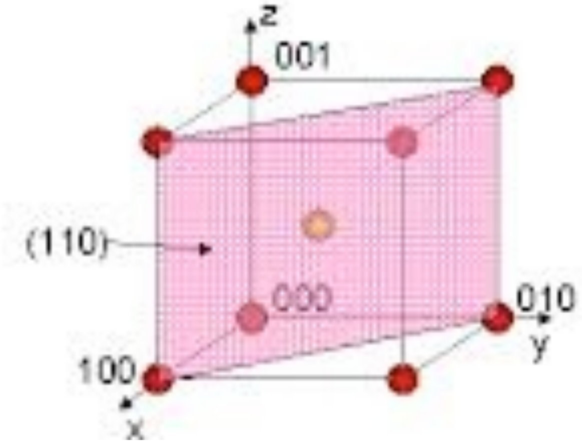
But: operation in radiation fields presents an extreme challenge, and this drives much contemporary development. That's where we will begin Lecture 2.

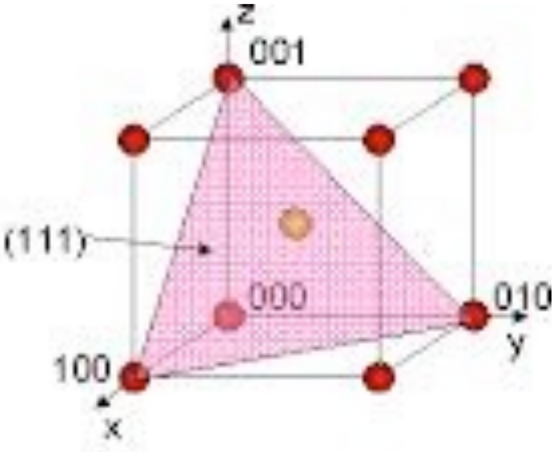
Backup

- Crystal orientation

Sensors with features registered along different crystal planes (denoted by Miller indices) may have different electrical characteristics before irradiation but are comparable after irradiation







The crystal orientation is specified by the location of the flat on the silicon wafer.

<111>:
- chosen by CDF SVX, ATLAS SCT

<100>
- chosen by CDF L00, CMS pixels, ATLAS pixels, LHCb

Use of 6-inch wafers is now standard.

Summary on readout noise, from “Pixel Detectors” by L. Rossi et al. (Springer, 2006)

$$\mathrm{ENC} = \frac{\text{noise output voltage (rms)}}{\text{signal output voltage for the input charge of } 1\mathrm{e}^-}$$

$$ENC_{tot}^2 = ENC_{shot}^2 + ENC_{therm}^2 + ENC_{1/f}^2$$

$$ENC_{\text{shot}} = \sqrt{\frac{I_{\text{leak}}}{2q}\tau_f} = 56e^- \times \sqrt{\frac{I_{\text{leak}}}{\mathrm{nA}}\frac{\tau_f}{\mu s}}$$

$$ENC_{\text{therm}} = \frac{C_f}{q}\sqrt{\langle v_{\text{therm}}^2\rangle} = \sqrt{\frac{kT}{q}\frac{2C_D}{3q}\frac{C_f}{C_{load}}} = 104e^- \times \sqrt{\frac{C_D}{100\,\mathrm{fF}}\frac{C_f}{C_{load}}}$$

$$ENC_{1/\mathrm{f}} \approx \frac{C_D}{q}\sqrt{\frac{K_f}{C_{ox}WL}}\sqrt{\ln\left(\tau_f\frac{g_m}{C_{load}}\frac{C_f}{C_D}\right)} = 9e^- \times \frac{C_D}{100\,\mathrm{fF}}(\text{for NMOS trans.})$$

W, L = width and length of trans. gate

K_f = 1/f noise coefficient

C_{ox} = gate oxide capacitance

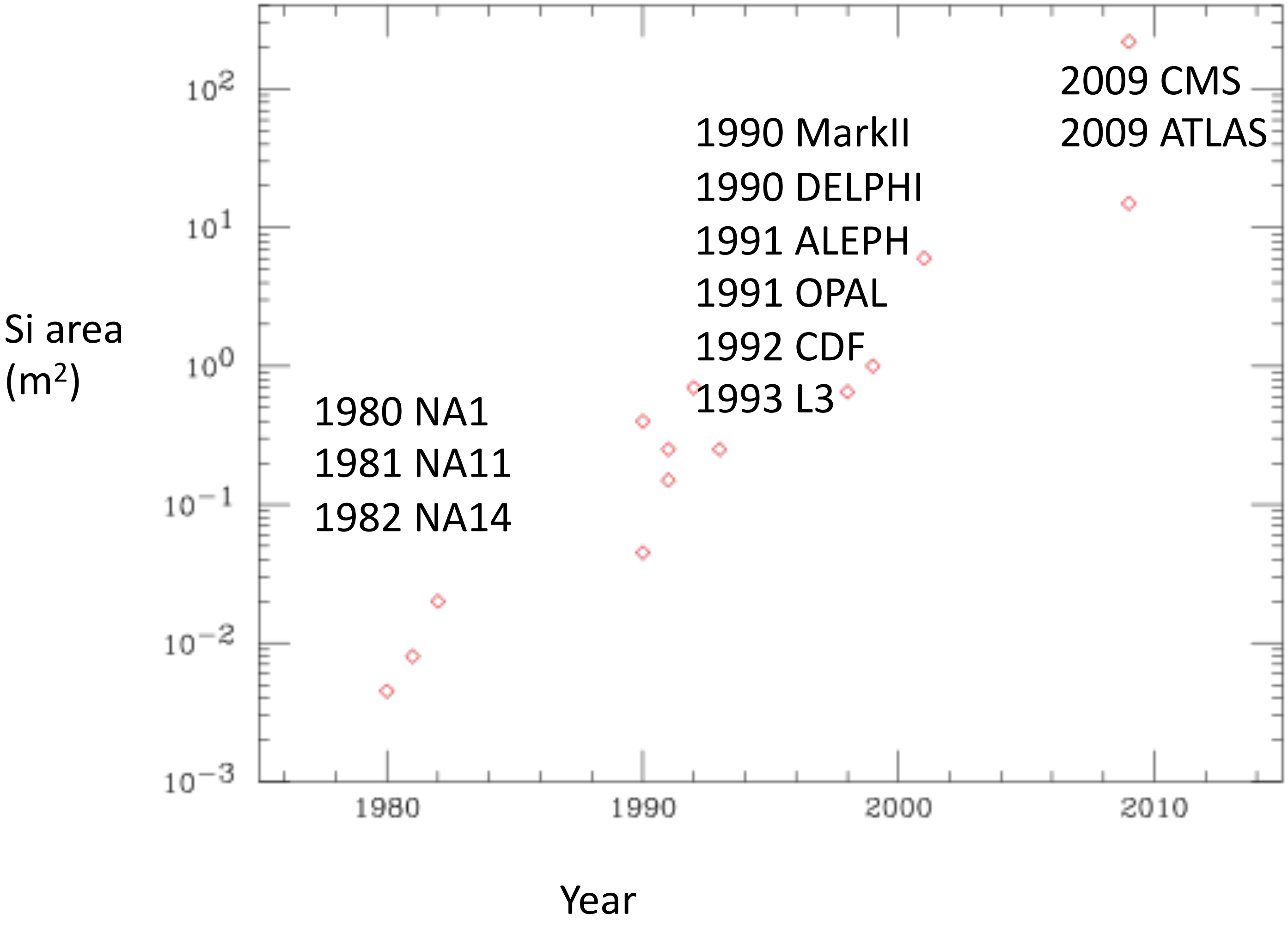
C_f = feedback capacitance

C_{load} = load capacitance

C_D = detector capacitance

τ_f = feedback time constant

Evolution of the scale of silicon detector coverage in HEP experiments



1980 NA1
1981 NA11
1982 NA14

1990 MarkII
1990 DELPHI
1991 ALEPH
1991 OPAL
1992 CDF
1993 L3

2009 CMS
2009 ATLAS

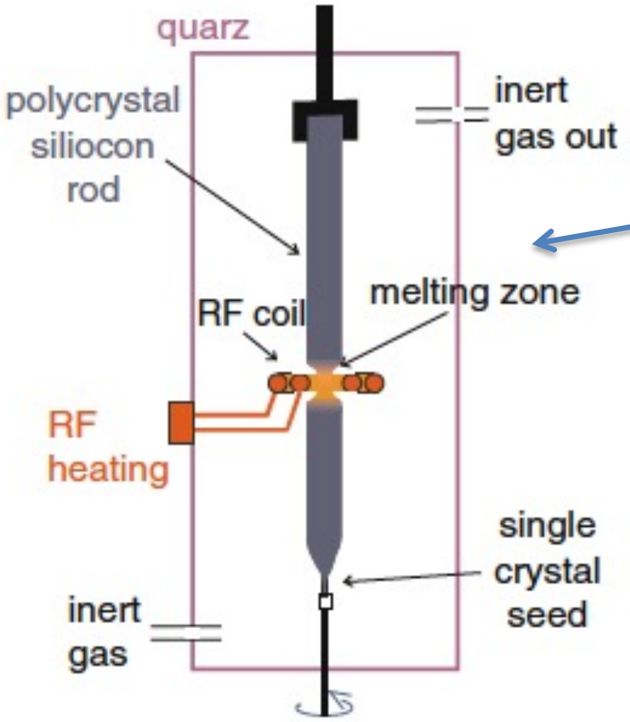
Binary resolution of strip sensor for pitch p

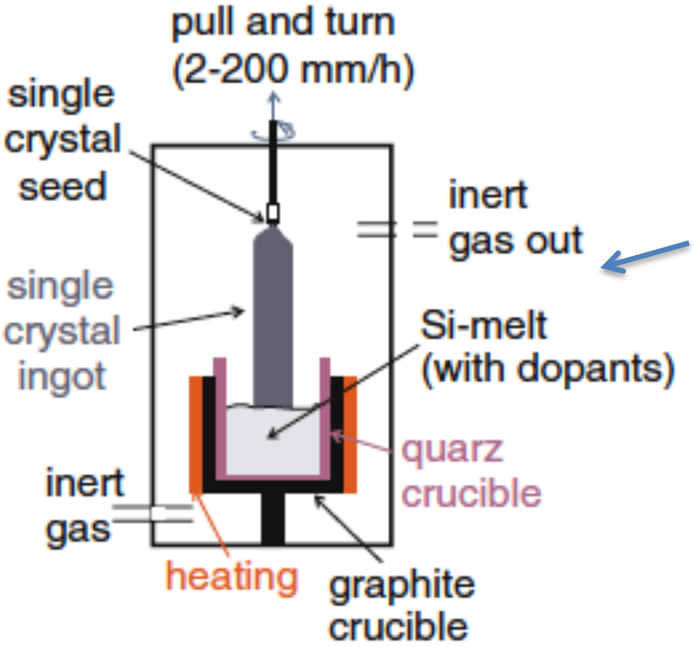
$$\sigma^2 = \frac{\int_{-\frac{p}{2}}^{\frac{p}{2}} (x_r - x_m)^2 D(x_r)\, dx_r}{\int_{-\frac{p}{2}}^{\frac{p}{2}} D(x_r)\, dx_r} = \frac{p^2}{12}$$

$D(x) = 1$ uniform distribution of tracks

$X_m = 0$ pixel centre

Silicon wafer fabrication





Float Zone

- Using a single Si crystal seed, melt the vertically oriented rod onto the seed using RF power and “pull” the monocrystalline ingot
- Can be oxygenated by diffusion at high T

Epitaxial silicon

- Chemical-Vapor Deposition (CVD) of Silicon
- CZ silicon substrate used a diffusion of oxygen
- Growth rate about 1mm/min
- Excellent homogeneity of resistivity
- 150 mm thick layers produced (thicker is possible)
- price depending on thickness of epi-layer but not exceeding ~ 3 x price of FZ wafer

Czochralski silicon

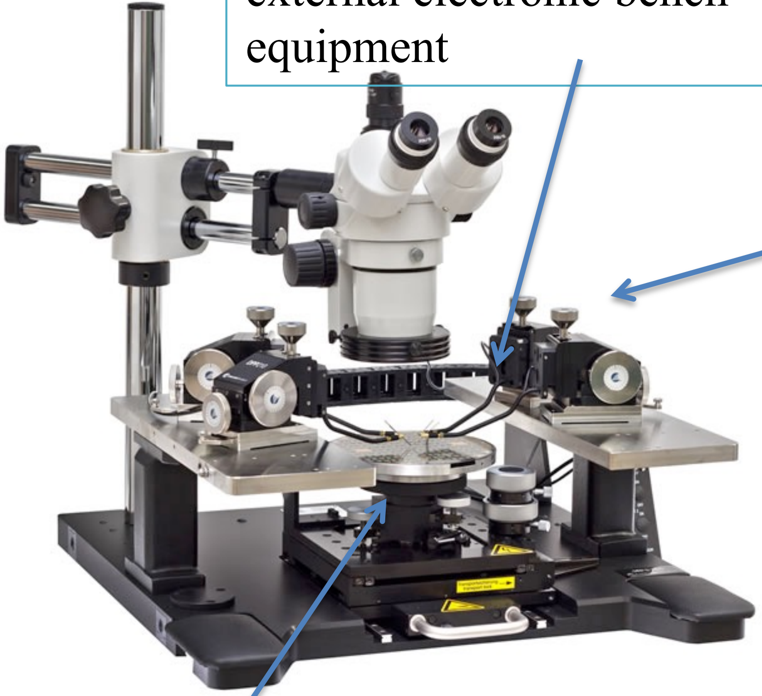
- Pull Si-crystal from a Si-melt contained in a silica crucible while rotating.
- Silica crucible is dissolving oxygen into the melt at high concentration
- Material used by IC industry (cheap). Also available in high purity for use as particle detector (MCz)

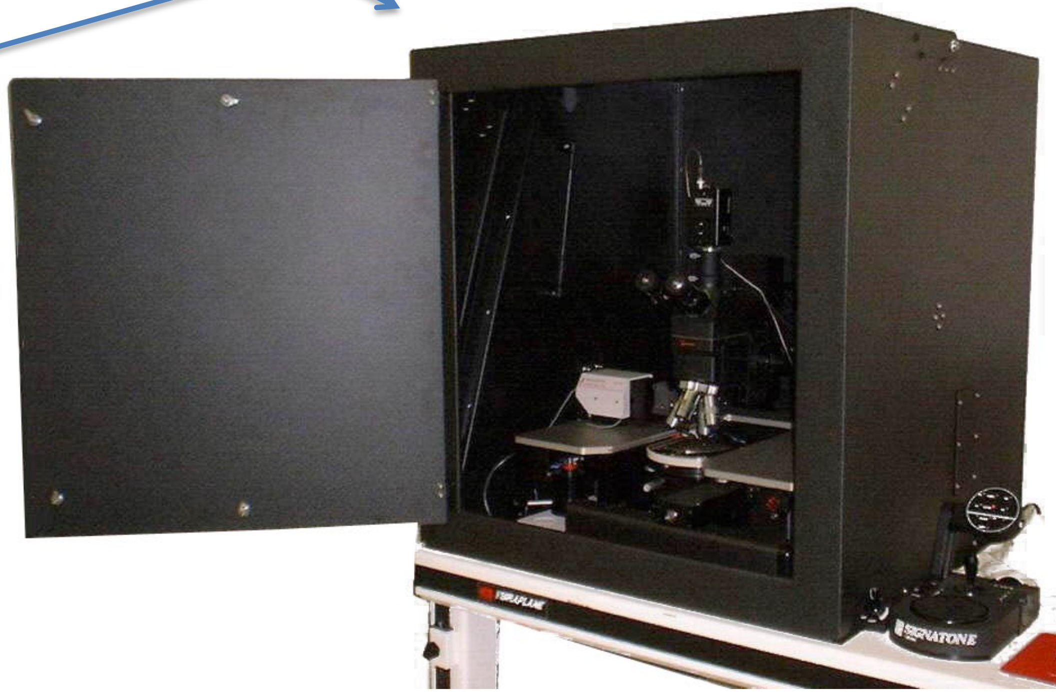
The environment for the experimenter characterizing the sensors

Silicon detectors can operate in air and at room temperature – but **as radiation damage progresses, cooling becomes essential.**

Probes connect individual structures on the sensor to external electronic bench equipment

Silicon is light-sensitive: characterization measurements require a dark enclosure. The standard laboratory facility is **a cleanroom-based probe station in a dark box:**





Vacuum chuck for secure positioning of sensor, can be cooled