# The Photoelectric Effect: an activity

M. Swartz
Johns Hopkins University

Introduction

The photoelectric effect was first observed in the late 19th century. Understanding it began the quantum revolution in physics. This week, we will build a simple apparatus that demonstrates quantum ideas in the classroom!
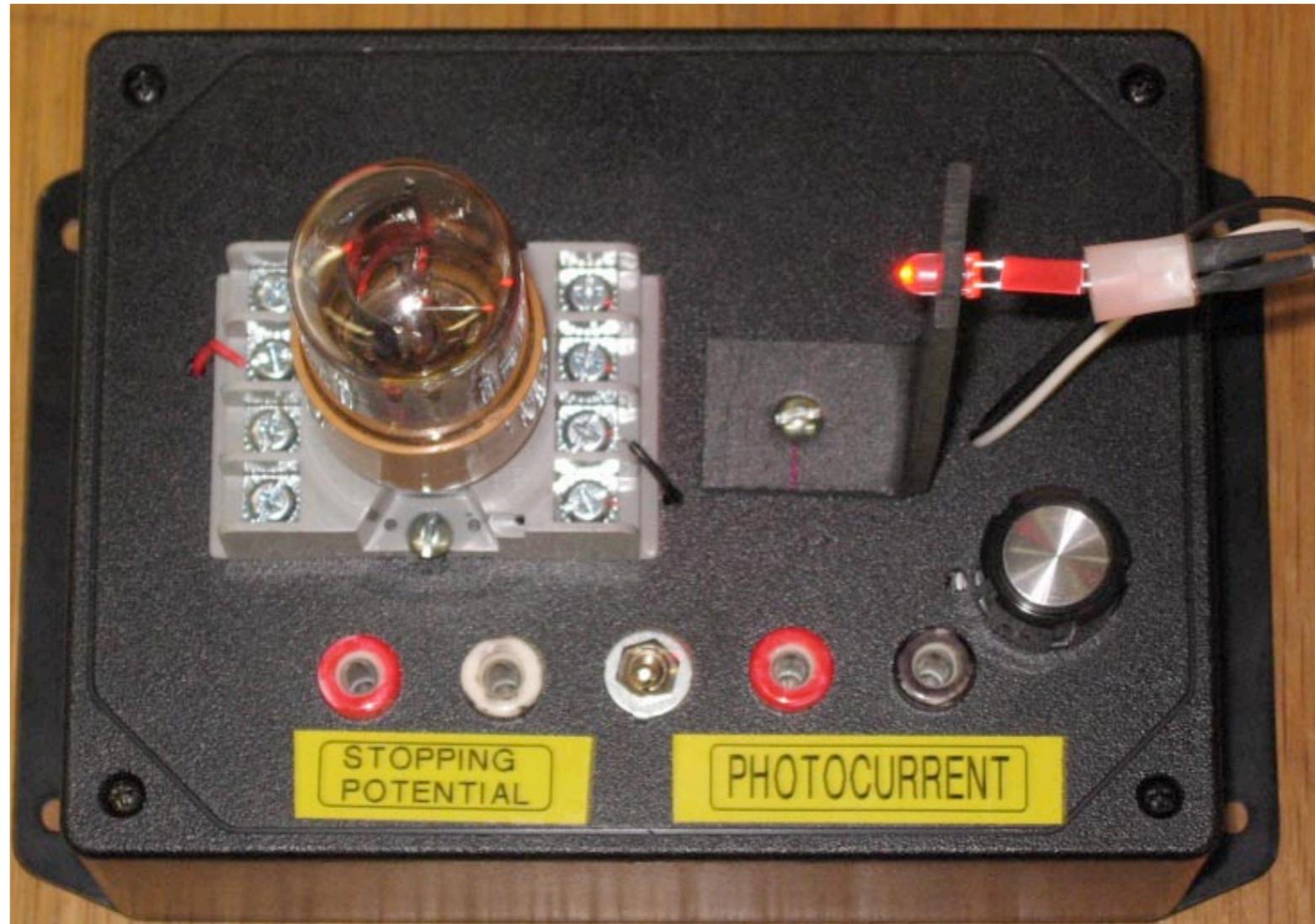
Experimental History

The discovery of the photoelectric effect and the early attempts to understand it were closely connected to research on electromagnetic waves

- First evidence observed by Heinrich Hertz in 1887

  ‣ he observed that the sparks in his spark gap detector were shorter in the dark

  ‣ phenomenon was related to presence of uv light
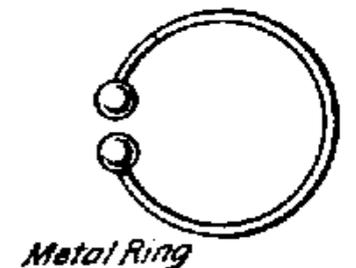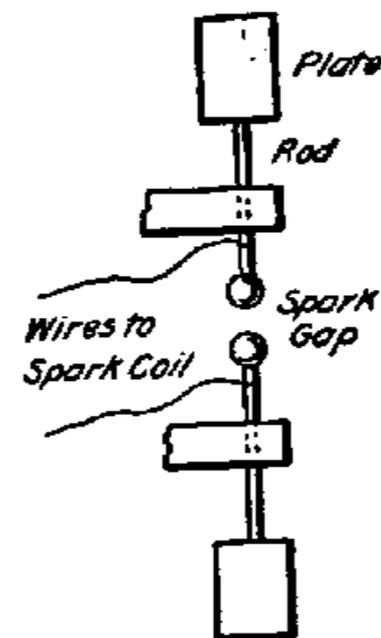
There were many related results in the next few years. The most important ones were:

- In 1888, Wilhelm Hallwachs observed that metal plates exposed to UV light became positively charged

- Also in 1888, similar work by Aleksandr Stoletov demonstrated that the rate of charging was proportional to the intensity of the light

 ‣ Stoletov's Law (first law of photoelectric effect)
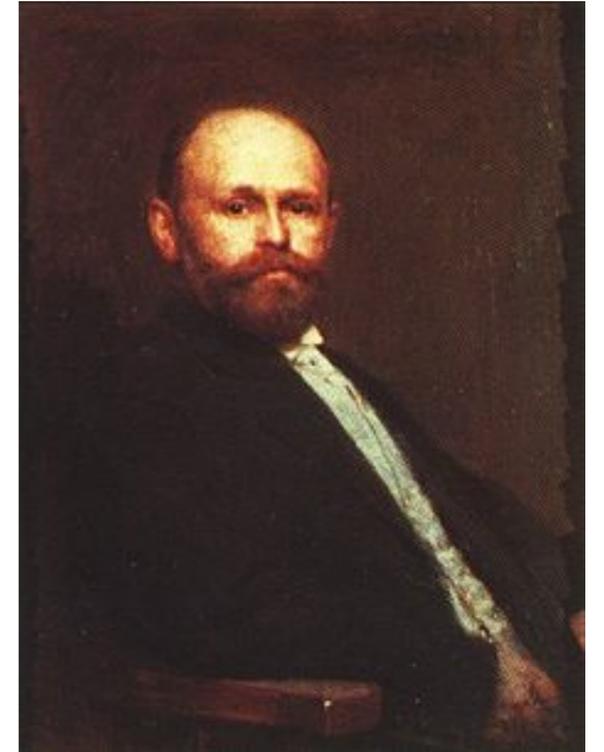
- In 1899, J.J. Thomson showed that metals exposed to UV light emitted particles identical to cathode rays (electrons)

- In 1902, Philipp Lenard studied the intensity and wavelength dependence of the energies of the emitted electrons

  ‣ energies are independent of the intensity [surprising since there are bigger E-fields in higher intensity light]

  ‣ energies decrease with increasing wavelength to 0 at a maximum wavelength

  ‣ hypothesized that the light "triggers" energetic electrons already present?

Theoretical History

- In 1900, Max Planck successfully described the spectrum of a "black body" by assuming that radiation in a cavity is in thermal equilibrium with harmonic oscillators in walls

  ‣ energies of oscillators are $E=h\nu$ [ν = freq.]

    ✳ h = Planck's constant = 6.626×10^{-34} J-s

  ‣ he hated statistical mechanics, but used it

  ‣ it worked!

The "Ultraviolet Catastrophe"

5000K

Radiation

Visible

Rayleigh-Jeans Law

Planck Radiation Formula

1000   2000   3000

Wavelength of radiation in nm

6
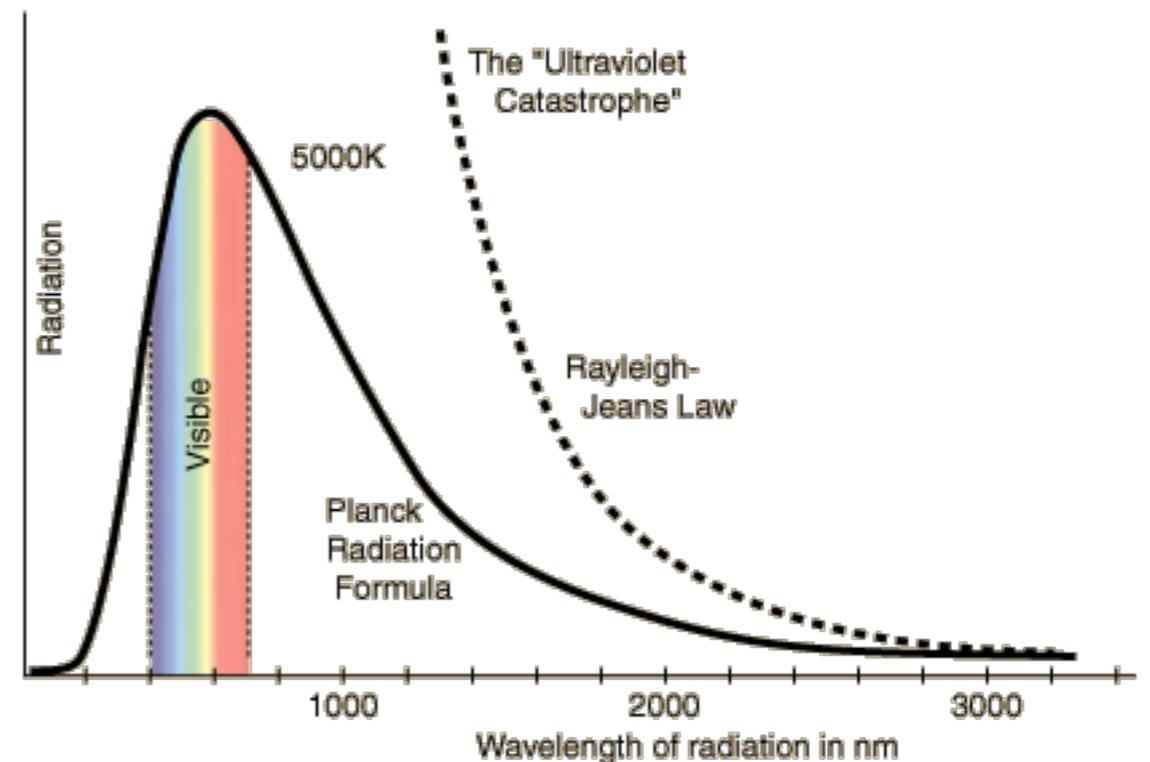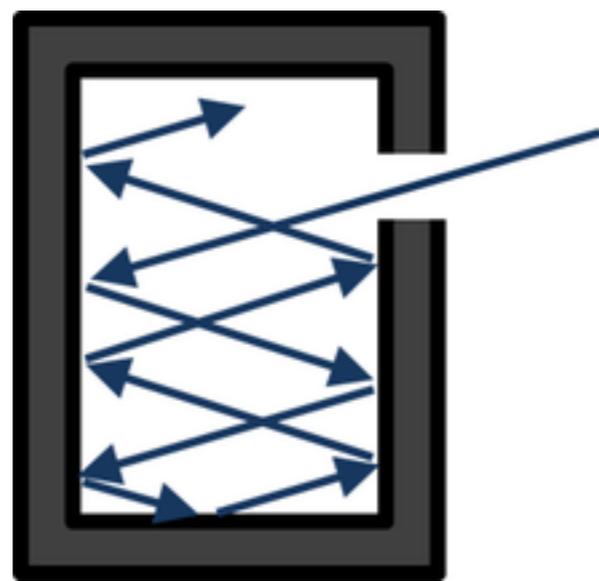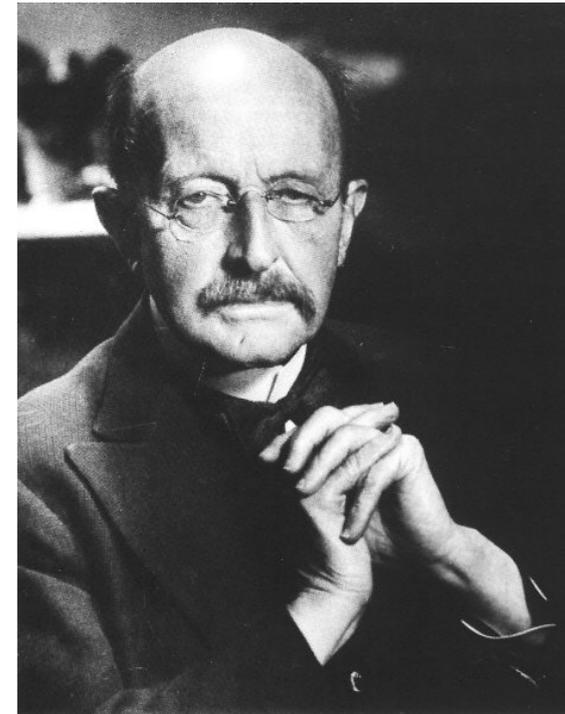
- In 1905, Albert Einstein had a very good year

  ‣ invents Special Relativity

  ‣ successfully describes Brownian Motion

  ‣ successfully describes Photoelectric Effect by building on Planck's blackbody work:

    * assumes that light is "corpuscular" and that the energy of each corpuscle [photon] is $E = h\nu$ [first explicit quantum]

    * light intensity is proportional to the number of photons

    * maximum kinetic energy of ejected electrons is the photon energy less the work function $T_{max} = h\nu - W$ [W is binding energy to the surface]

The classical Photoelectric Effect experiment is conceptually quite simple,
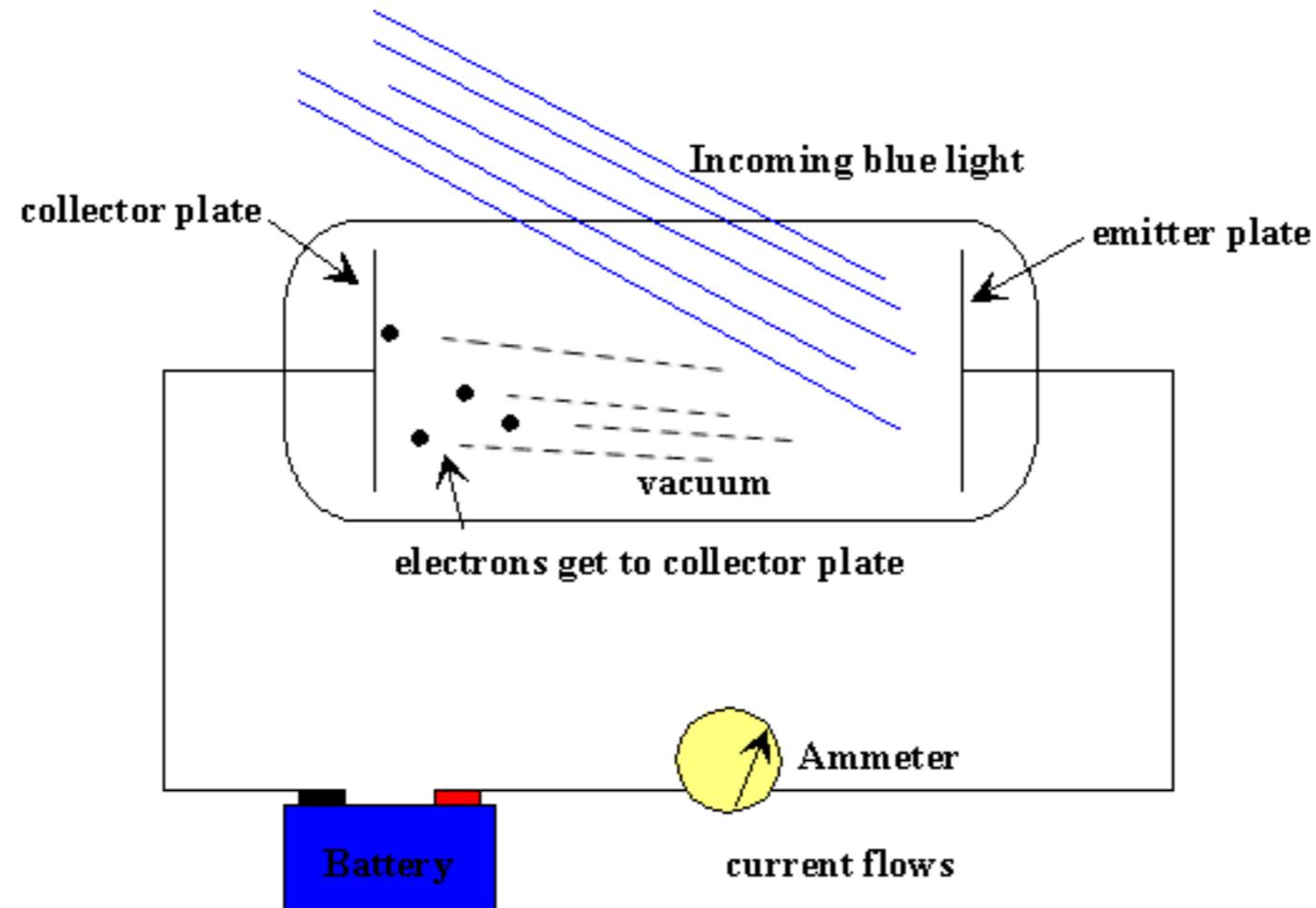
- Uses monochromatic light of frequency $\nu = c/\lambda$, λ is the wavelength

- A bias voltage V is applied to decelerate the ejected e-

  ‣ current stops flowing when $eV_s = h\nu - W = hc/\lambda - W$

Taking data at several frequencies should show a linear relationship with stopping potential,
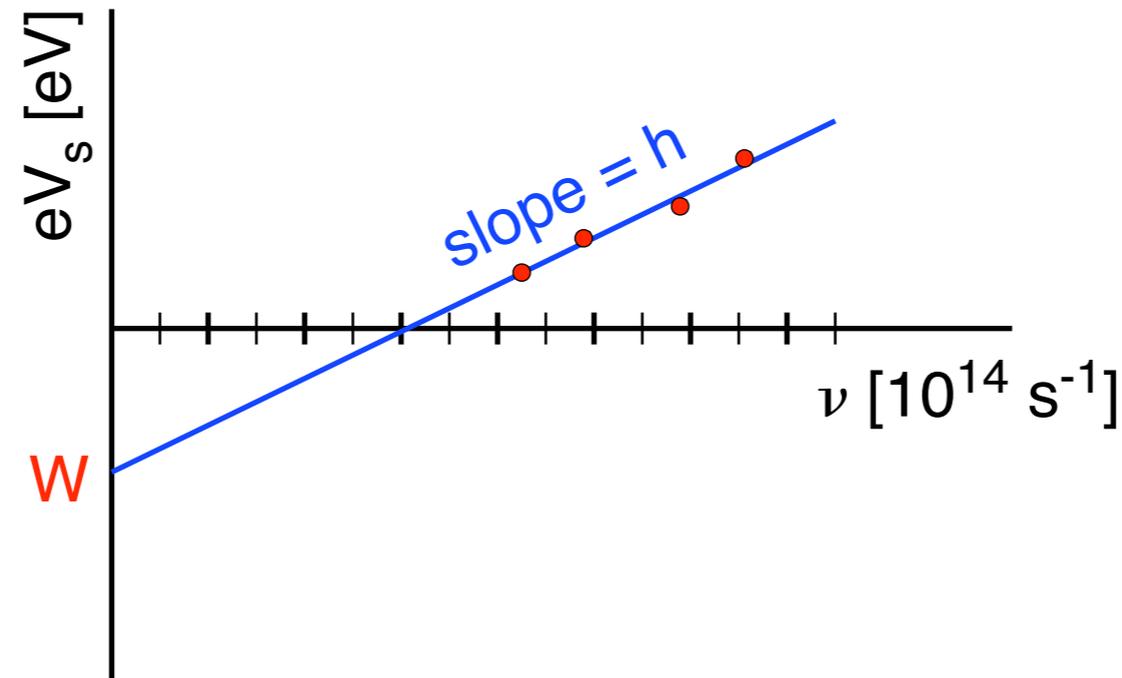

- the slope gives a measurement of Planck's constant h

- the x-intercept gives the minimum frequency [maximum wavelength] for the PE

- the y-intercept gives the work function W of the surface

- the photocurrent for $V < V_s$ should be linear in the light intensity

The Real Experiment

The key elements of the real experiment are a unit-gain phototube [1P39 or 939] and several led light sources. Both elements have "characteristics" that can complicate the interpretation of data.

- Phototube: our's is a very simple device with 2 parts

  ‣ photosensitive Sn-Cs semicircular screen: Cs lowers work functions of everything to make quantum efficiency large

 ‣ tubular anode to collect the emitted electrons
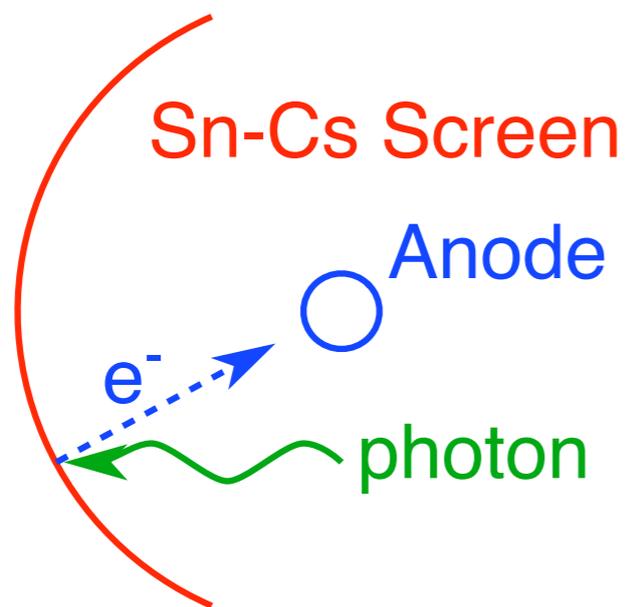
The Sb-Cs photocathode [S-4] has a very non-uniform response vs wavelength (frequency).

- If the light source has a narrow wavelength [frequency] distribution, this wouldn't matter

  ‣ led's have "broad" spectra

  ‣ need response corrections?

- The red light points will have small signals

  ‣ need bright red leds

Spectral response of common photomultiplier cathodes.
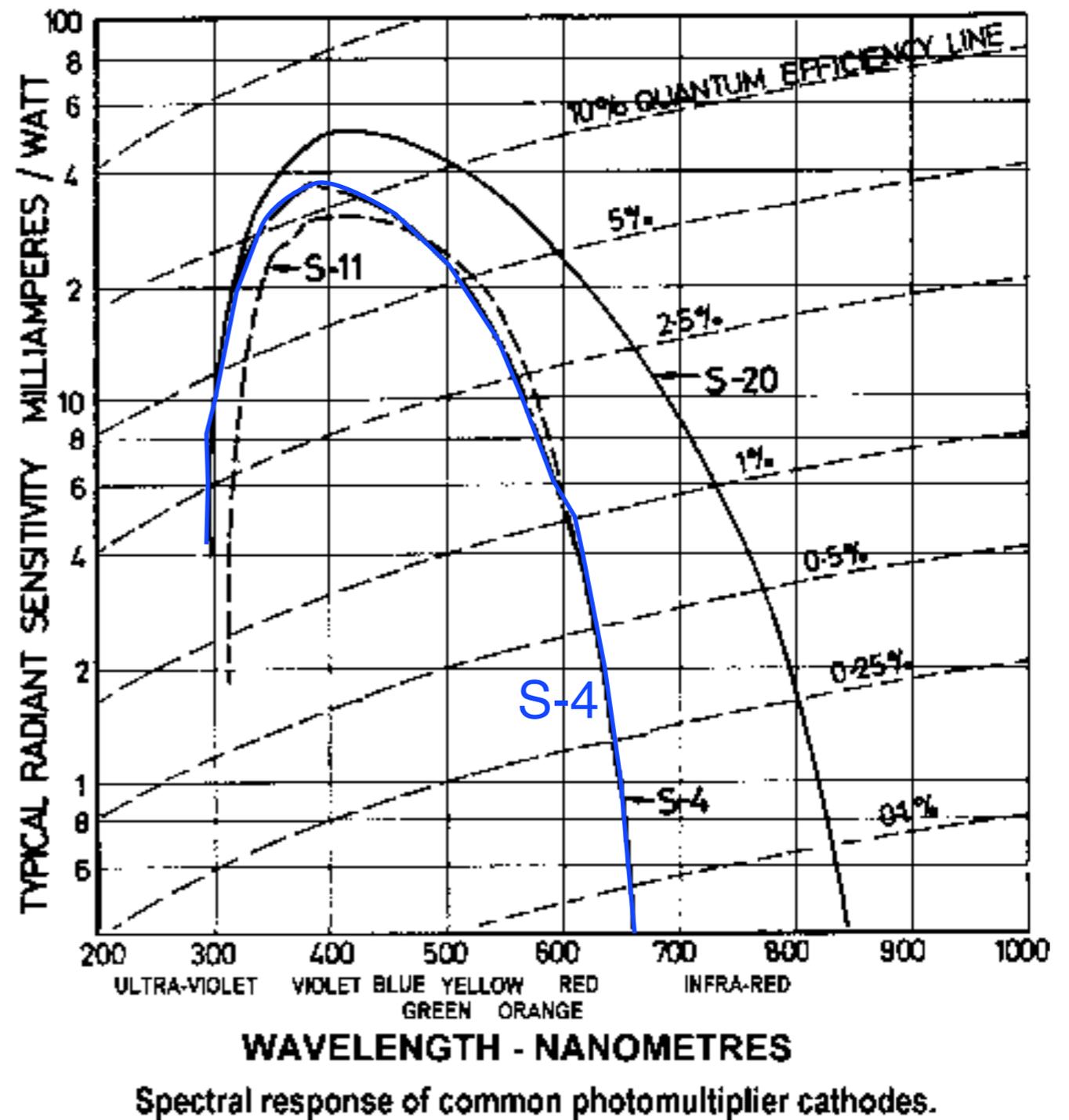
- Light Emitting Diodes [LEDs]: semiconductor pn junctions

  ‣ shorter wavelength, higher frequencies only developed during last 20 years

  ‣ makes this experiment relatively inexpensive

 ‣ unfortunately, they are not really monochromatic,
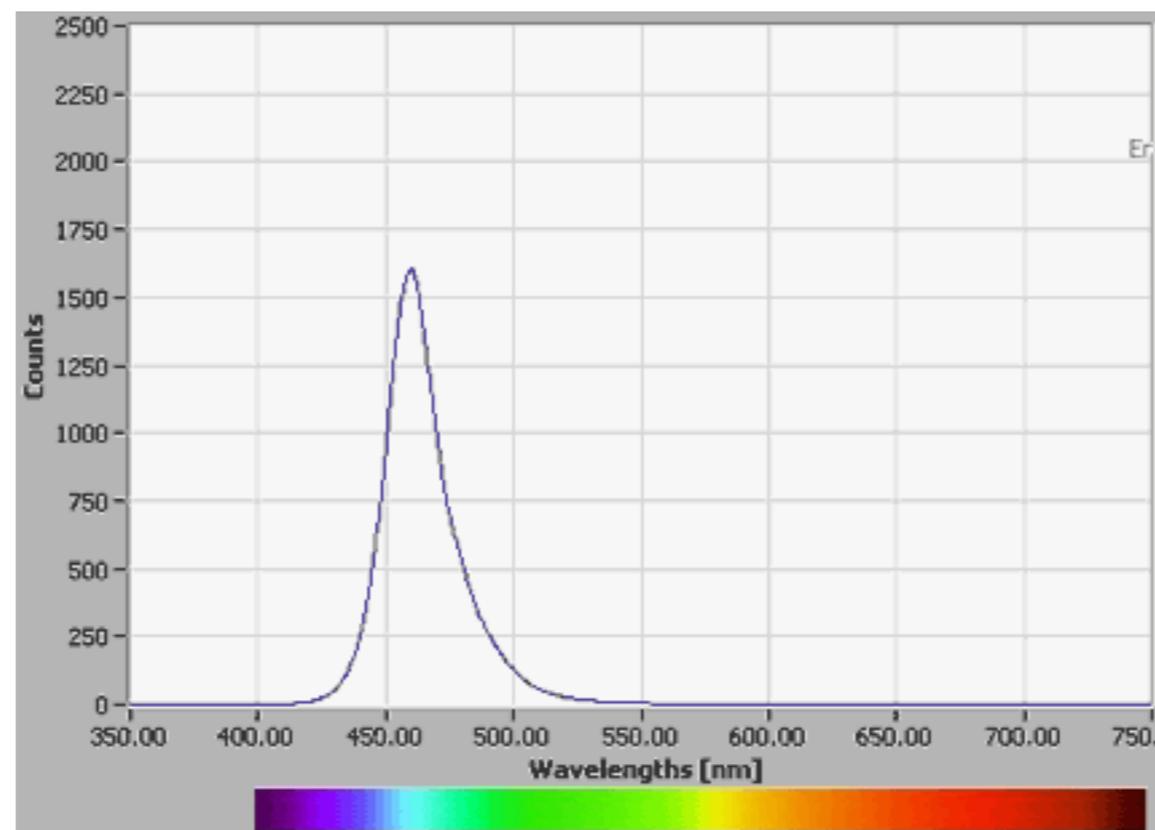
pn-Junction Diodes

Our Light-Emitting-Diodes [LEDs] are pn junctions. Since pn junctions are also at the heart of tracking in particle physics and imaging in astronomy, maybe we should spend a few slides reviewing what they are?

When many atoms are assembled in a lattice, their quantum states become grouped into several bands of closely spaced states. The higher energy bands can be "delocalized": electrons can move from one atom to another,

- Electrons in the valence band are bound to their atoms
  - ‣ missing electrons [holes] in the valence band can move from one site to the next
- Conduction band e- can move as almost free particles through crystal
- Valence band holes can also move as almost free particles through crystal
- There can be an energy gap between valence + conduction
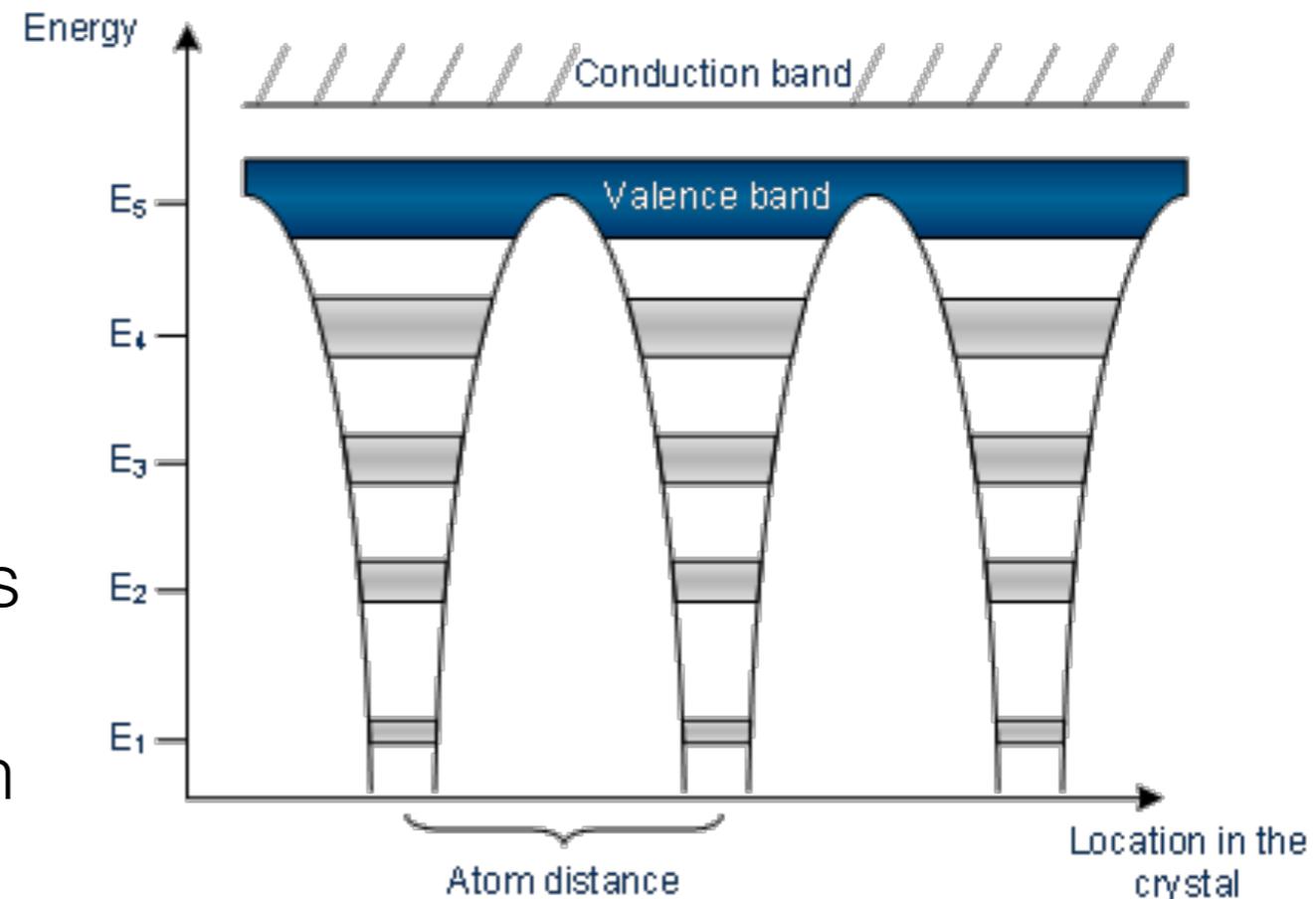
In simplest approximation, there the size of the band gap determines the properties of the crystal,
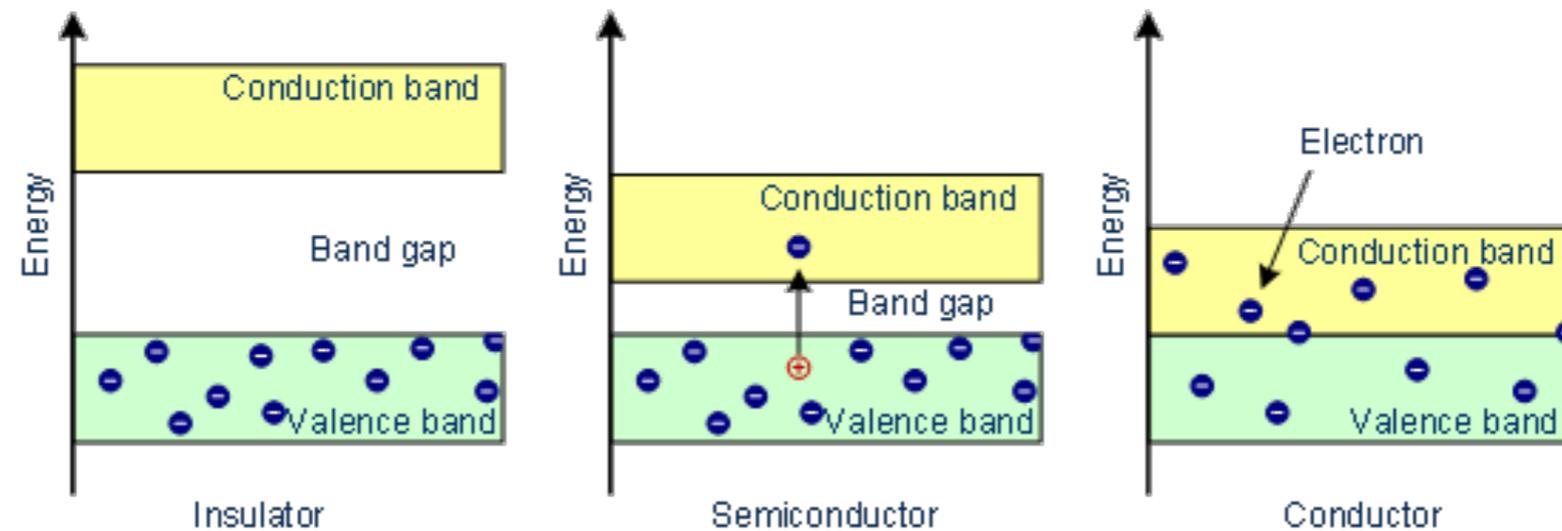

- If the gap is much larger than kT [typical thermal energy], all of the electrons remain in the valence band, bound to atoms
    ‣ material is an insulator
- If the gap is small or if there are unfilled states at the top of the valence band, a large fraction of the electrons are delocalized
    ‣ material is a conductor
- If the gap is only ~50 times larger than kT, then some of the electrons become delocalized and leave behind delocalized holes
    ‣ material is a semiconductor

The density of e+h [carriers] in a pure semiconductor is much smaller than the density of electrons [in Si, 2×10^{-14} times smaller]. What this means is that the properties are very sensitive to impurities: deliberately introduced [dopants] or not. There are two kinds of dopants,
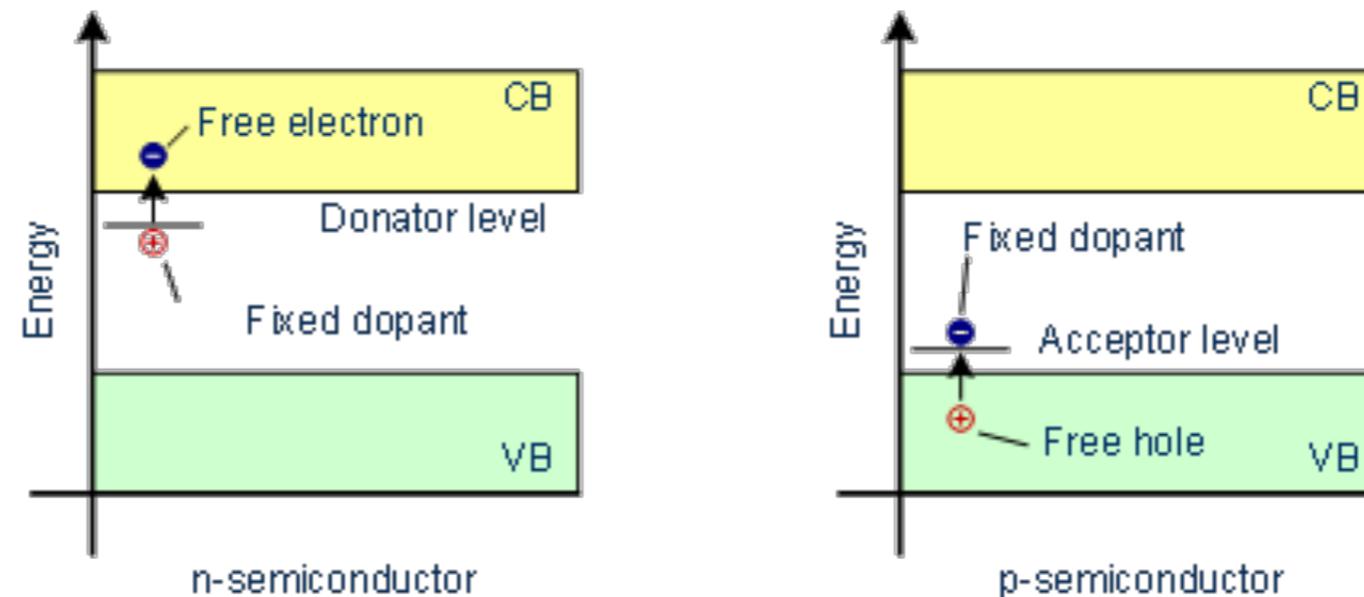

- A donor state has an energy near the conduction band and donates an electron to the CB leaving a positively charged ion
  ‣ n-type material
- An acceptor state has an energy near the valence band and captures an electron producing a free hole and leaving a negative ion
  ‣ p-type material
- Both types of material are electrically neutral until they are combined

# The pn junction contains both types of material in contact

How to connect the pn junction?

Reverse condition

Forward condition

n-crystal   p-crystal

	fixed negative charge (boron ion)		negative charge carrier (electron)
+	fixed positive charge (phosphorus ion)	⊕	positive charge carrier (hole)
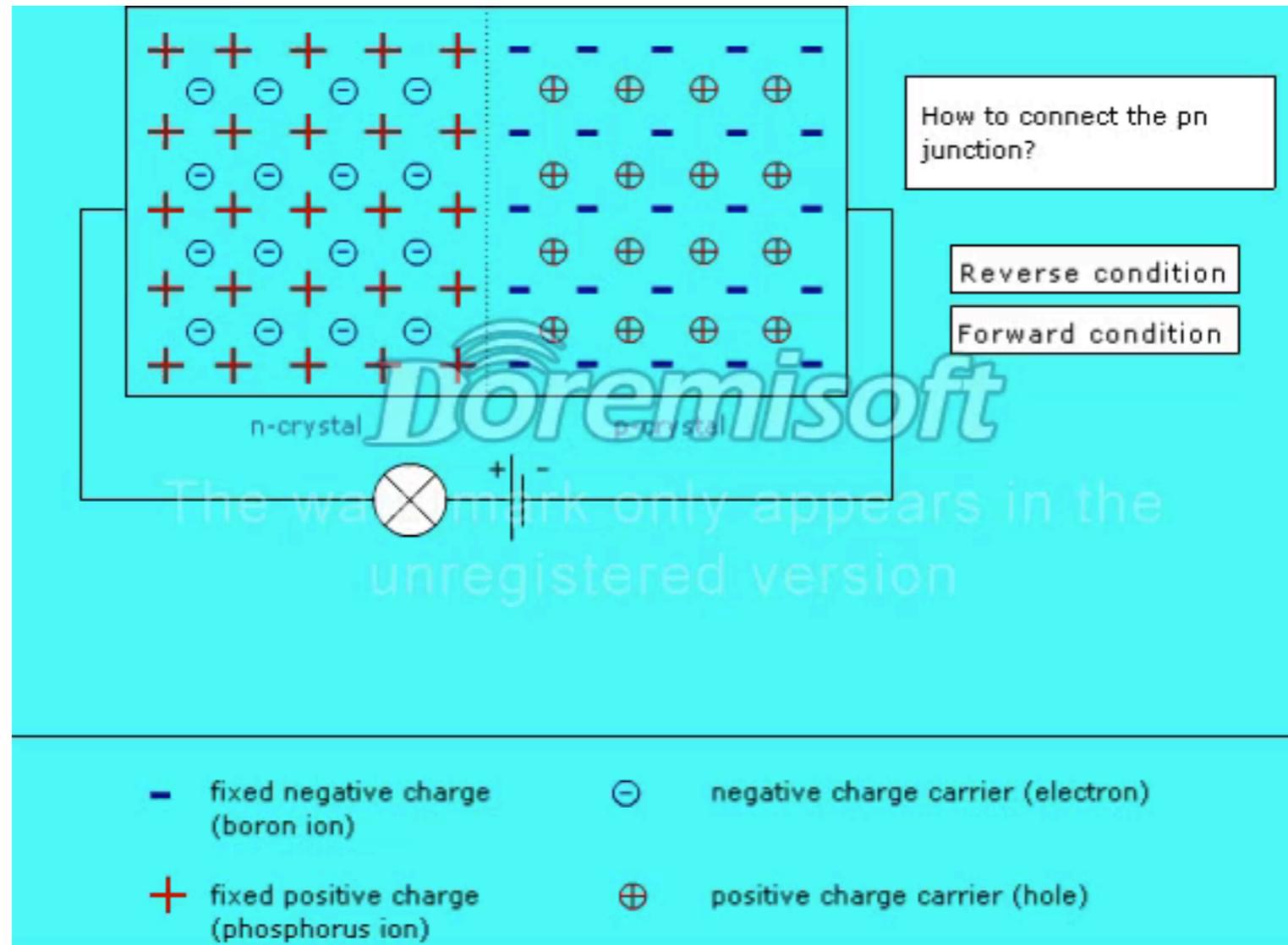
- With no applied potential get a "depleted" zone with non-zero E-field
- With reverse bias [+ to n, - to p], depleted zone grows, no current flow
  - ‣ perfect for tracking detectors and photo-sensitive devices
- With forward bias [- to n, + to p], current flows with low resistance
  - ‣ perfect for light-emitting diodes [lots of e-hole recombination]

16

Much particle physics detector technology is based upon the reverse-biased diode:
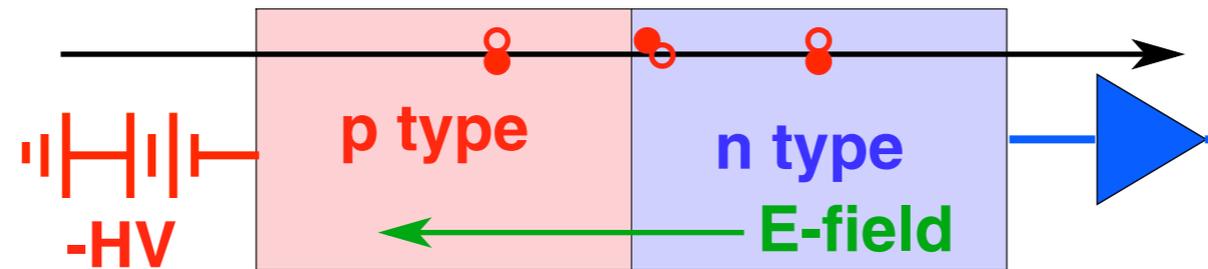
- Large depleted thickness is traversed by charged particles
  ‣ ionization produces large signals: **22,000** eh pairs in **300** μ**m**
- Supports a large internal field to sweep the carriers to electrodes
- Small leakage currents for electronics to cope with

  ‣ good signal to noise
- High precision tracking (approaching few μ**m**) possible
- Affordable: sort of ....
- Can be made quite radiation hard compared with older tracking technologies (still not enough)

Silicon used for tracking detectors is an "indirect" semiconductor, the materials used for LEDs must be "direct" semiconductors to have a good efficiency,

Silicon

E

Electrons

1.12eV

Momentum

Holes

GaAs

E

Electrons

1.43eV

Momentum

Holes
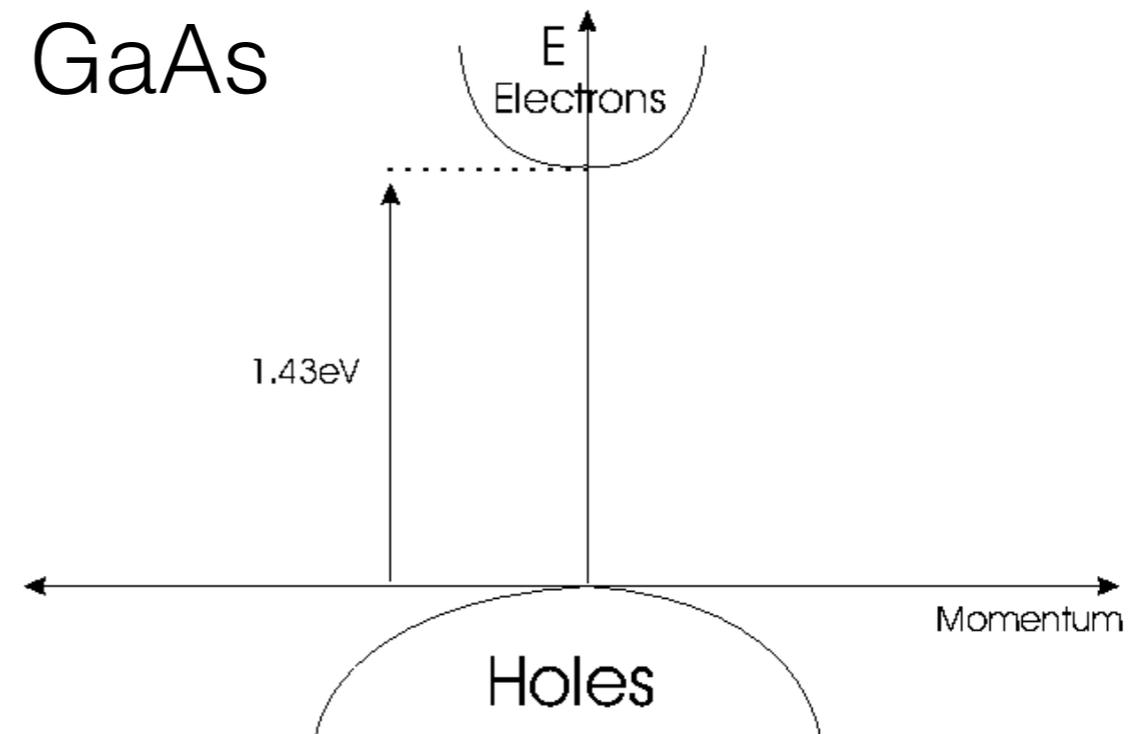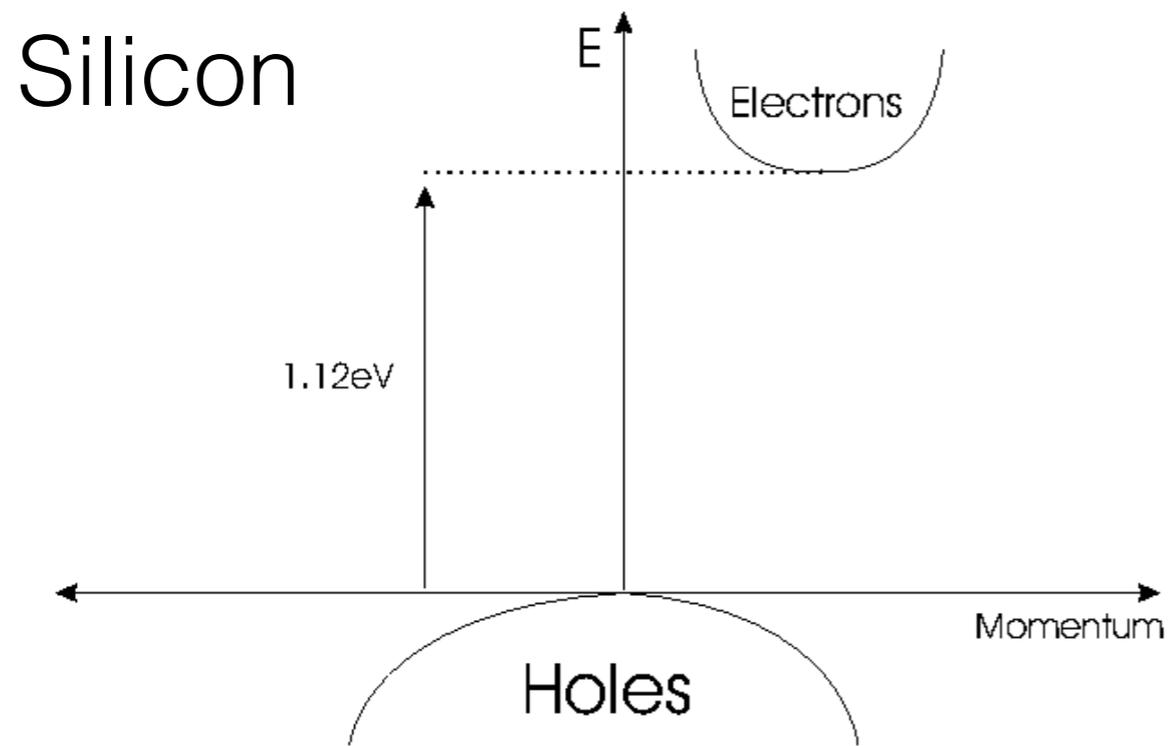
- For efficient e-h recombination, need equal e-h momenta

  ‣ direct semiconductors have "aligned" momentum vs energy

  ‣ indirect semiconductors have an offset, need "phonons" to interact and balance momenta

- LEDs are made from forward-biased direct semiconductors

  ‣ wavelength of an LED is determined by the bandgap of the material

  ‣ brightness/width is determined by the density of available states

  ‣ Some [white/purple] are not monochromatic!

|	Color	Wavelength [nm]	Voltage drop [ΔV]	Semiconductor material
	Infrared	$\lambda > 760$	$\Delta V < 1.63$	Gallium arsenide (GaAs) Aluminium gallium arsenide (AlGaAs)
	Red	$610 < \lambda < 760$	$1.63 < \Delta V < 2.03$	Aluminium gallium arsenide (AlGaAs) Gallium arsenide phosphide (GaAsP)
	Orange	$590 < \lambda < 610$	$2.03 < \Delta V < 2.10$	Gallium arsenide phosphide (GaAsP) Aluminium gallium indium phosphide (AlGaInP)
	Yellow	$570 < \lambda < 590$	$2.10 < \Delta V < 2.18$	Gallium arsenide phosphide (GaAsP) Aluminium gallium indium phosphide (AlGaInP)
	Green	$500 < \lambda < 570$	$1.9[66] < \Delta V < 4.0$	**Traditional green:** Gallium(III) phosphide (GaP)
	Blue	$450 < \lambda < 500$	$2.48 < \Delta V < 3.7$	Zinc selenide (ZnSe) Indium gallium nitride (InGaN)
	Violet	$400 < \lambda < 450$	$2.76 < \Delta V < 4.0$	Indium gallium nitride (InGaN)
	Purple	multiple types	$2.48 < \Delta V < 3.7$	Dual blue/red LEDs, blue with red phosphor,
	Ultraviolet	$\lambda < 400$	$3.1 < \Delta V < 4.4$	Diamond (235 nm)[67] Boron nitride (215 nm)[68][69]
	Pink	multiple types	$\Delta V \sim 3.3[72]$	Blue with one or two phosphor layers: yellow with red, orange or pink phosphor added afterwards,
	White	Broad spectrum	$\Delta V = 3.5$	Blue/UV diode with yellow phosphor
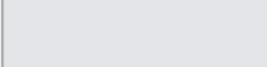
Our Experiment

The apparatus/circuit that we are building this week is VERY simple.

- A 9V battery and variable resistor are used to generate the stopping potential
  - ‣ measured across 2 terminals
- The photocurrent flows through a **100kΩ** resistor: 10nA = 1mV
  - ‣ measured across 2 terminals
- The LEDs are powered by a 6V battery pack and **180Ω** series resistor
  - ‣ Adding a rotary switch to select one of several permanently wired diodes
  - ‣ Adding variable resistor to vary LED intensity
- Adding a light shield over the tube
  - ‣ mask anode with tape?
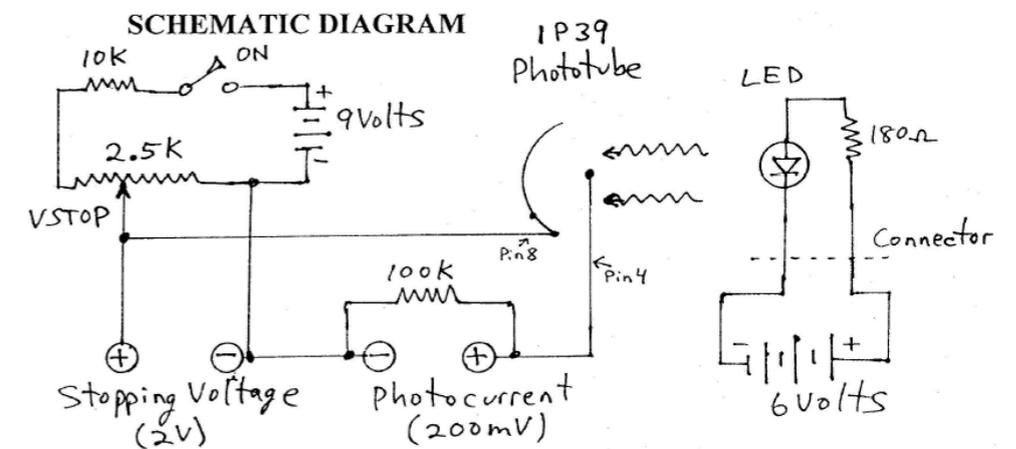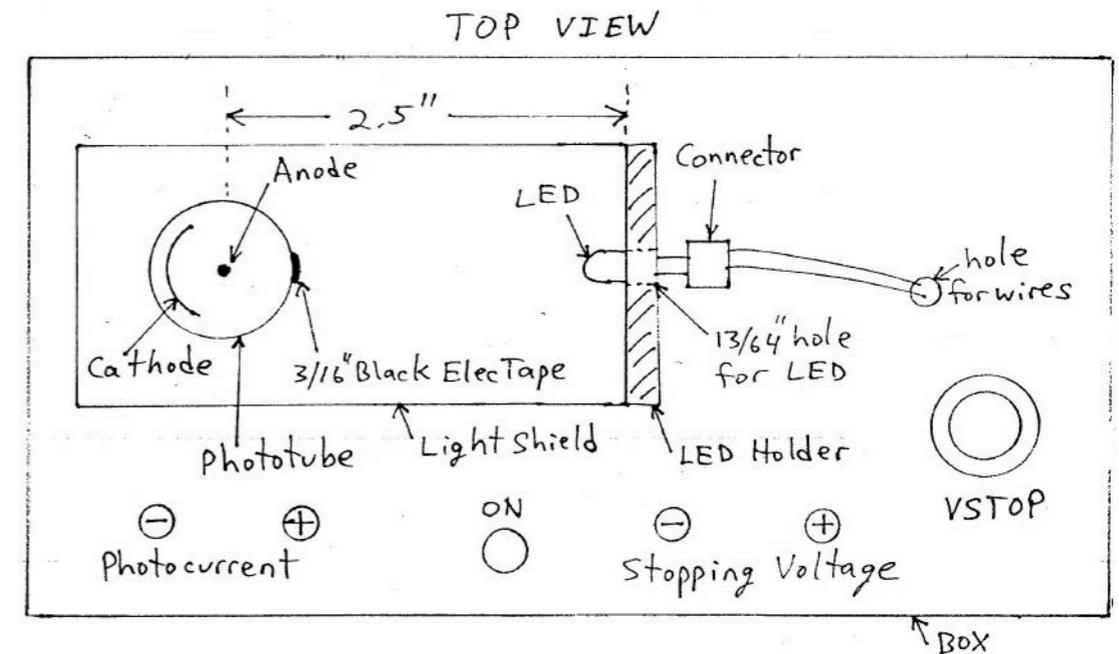
Procedure

The classical procedure has always been to increase the stopping potential until the photocurrent goes to zero. This is because the photons are absorbed in a layer near the surface of the photocathode. The more deeply penetrating ones produce electrons that scatter and lose energy before emerging from the surface,
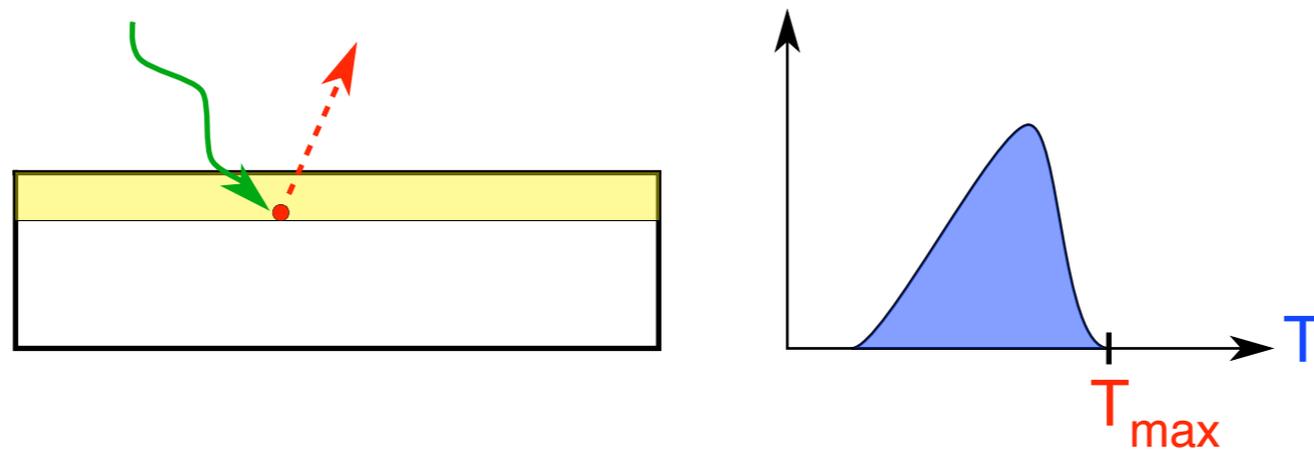
- Electrons produced at the surface don't scatter and can carry the maximum energy $T_{max}=h\nu - W$

  ‣ T_{max} is the endpoint of the electron energy spectrum

- Our photons have a spectrum of energies, this means a spectrum of T_{max}

  ‣ is zeroing the photocurrent the right thing to do?

  ‣ do we plot the data at the average photon frequency?

A second question is how does the phototube response modify the distribution of photon energies detected by the photoube? The response is quite non-uniform,
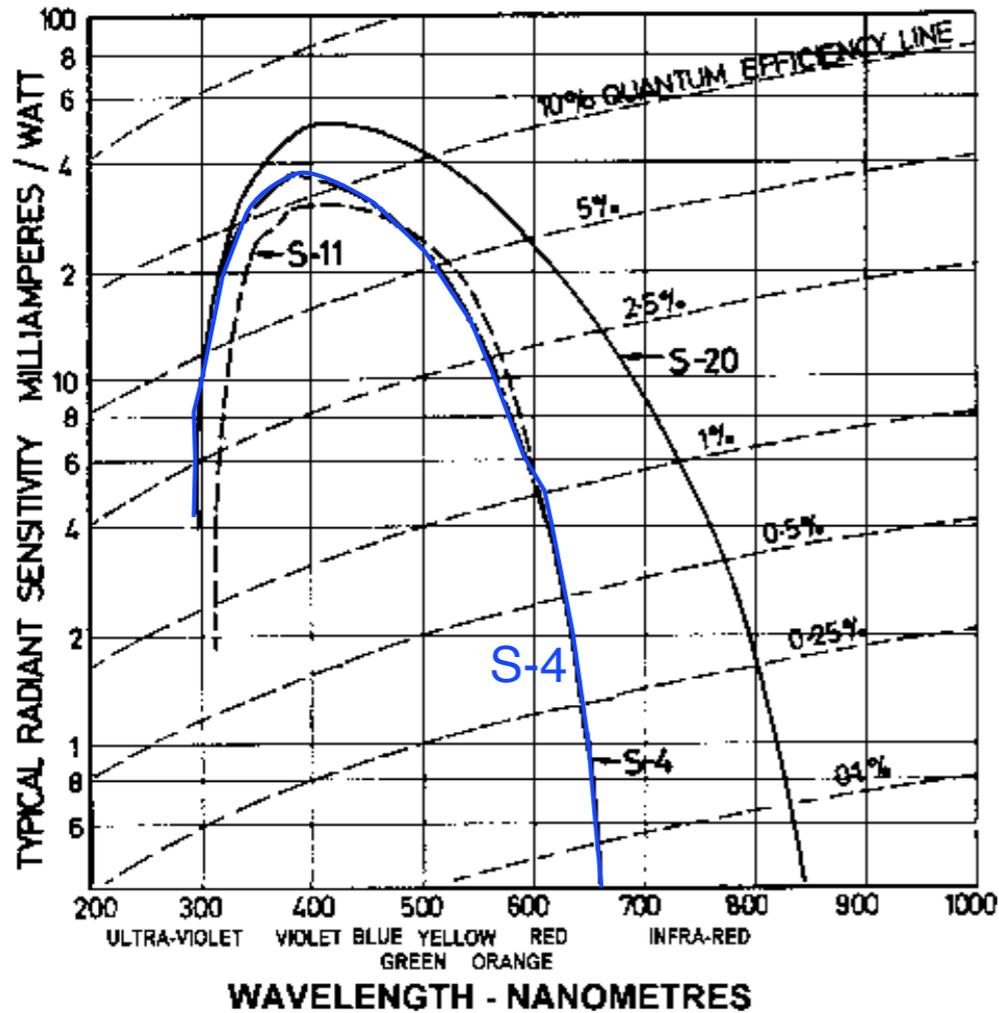
Spectral response of common photomultiplier cathodes.
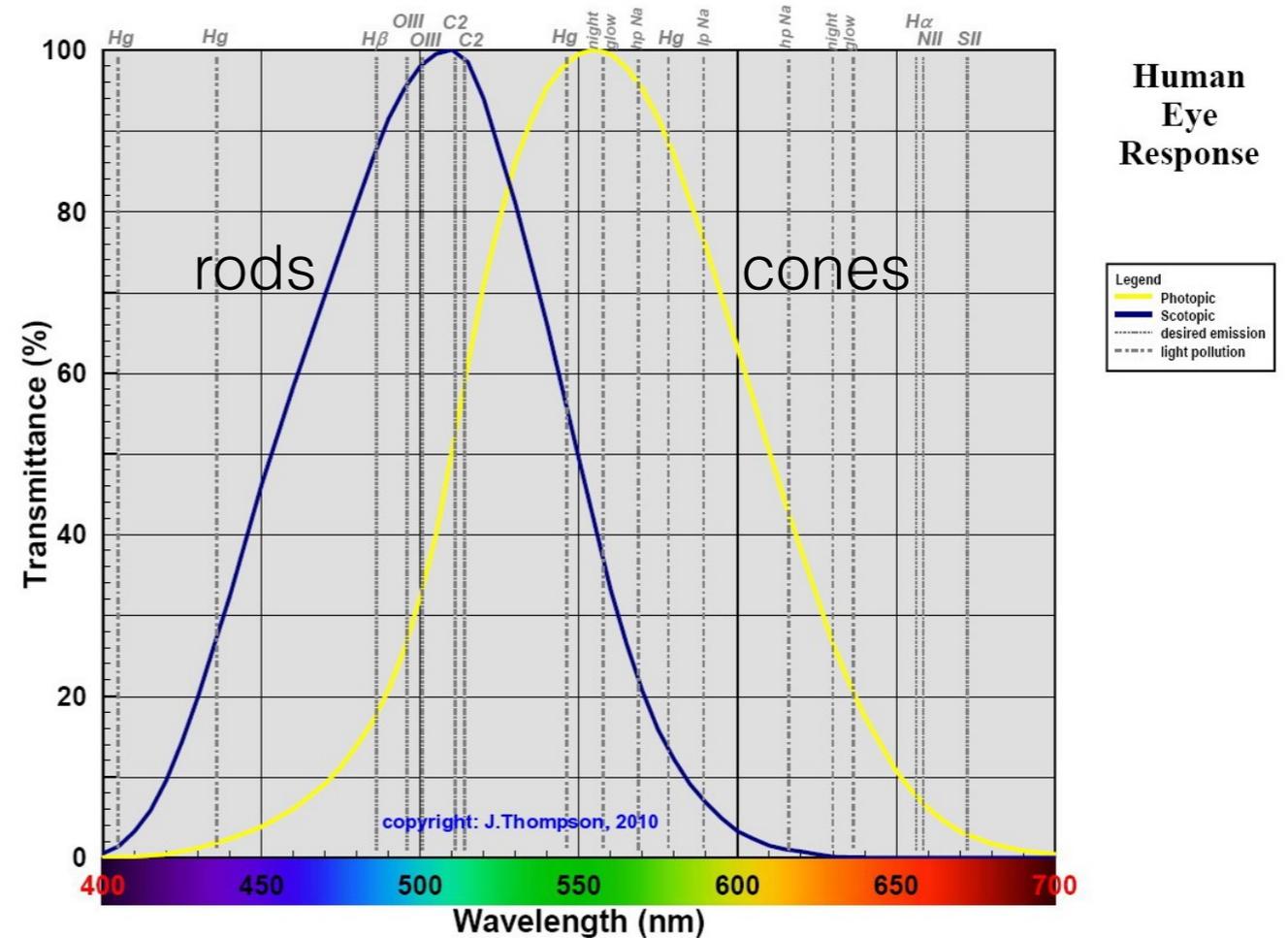
- Phototube response peaks at 400nm and drops to zero around 660nm
- Bright light [phototropic] response of the human eye also stops around 660 nm
  ‣ What does than LED manufacturer actually specify?

Let's digitize the response function for the 624nm LED [blue curve]. We can also fold it with the phototube response function [green dots],

- LED peaks at 631.7nm
- LED*PT peaks at 631.7nm
- LED*PT average at 625.5nm
  - This is close to the quoted value
  - Does the spec include human eye response?
- What number do we use?
  - the average?
  - some fraction of the area under the red dashed curve evaluated from the small λ side?
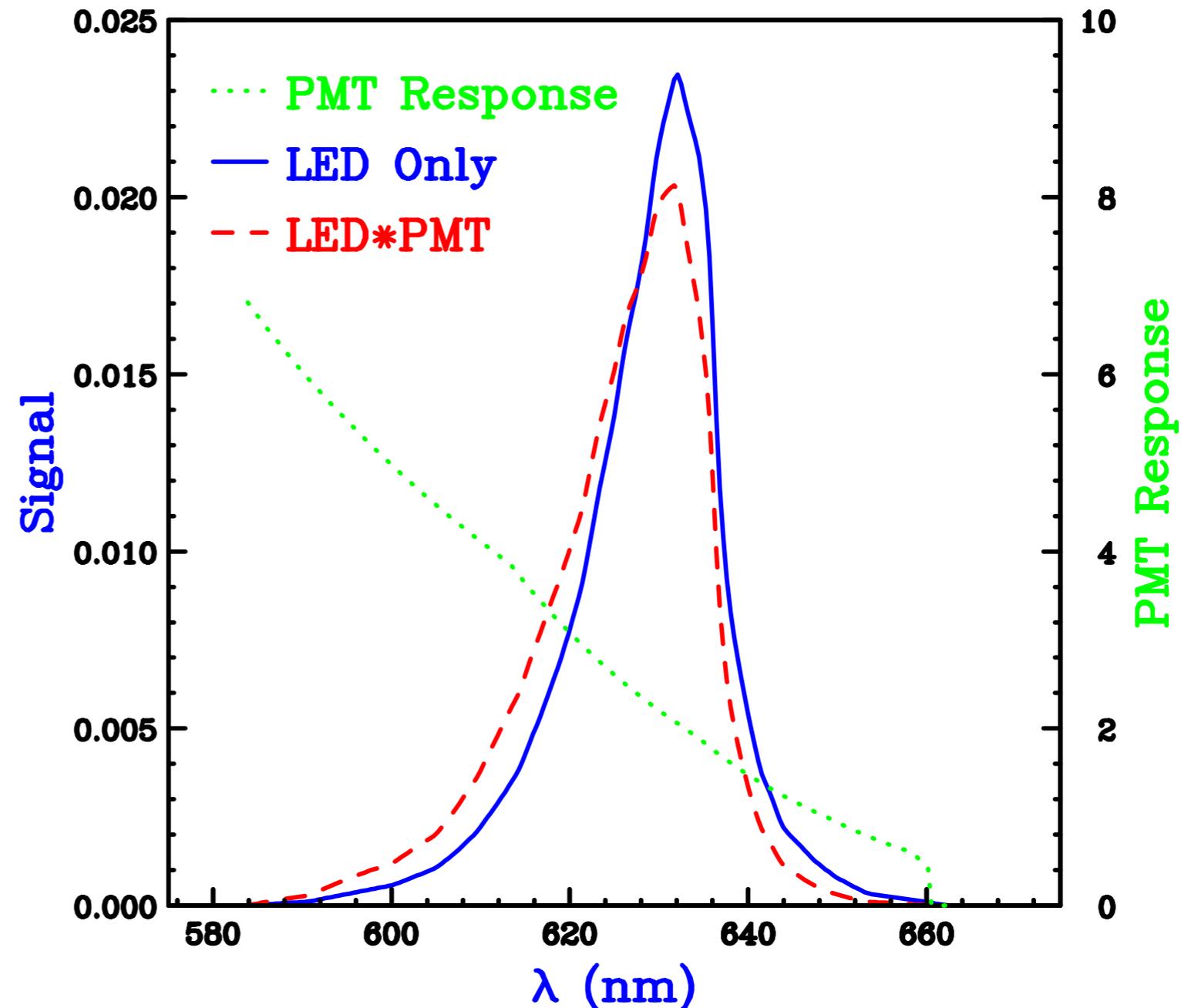
The PT effect is smaller at smaller wavelengths because the response is less steep. For the yellow diode at 590nm, it looks like this

- LED peaks at 591.8nm
- LED*PT peaks at 590.9nm
- LED*PT average at 589.3nm
- What number do we use?
  - ‣ the average?
  - ‣ some fraction of the area under the red dashed curve evaluated from the small λ side?
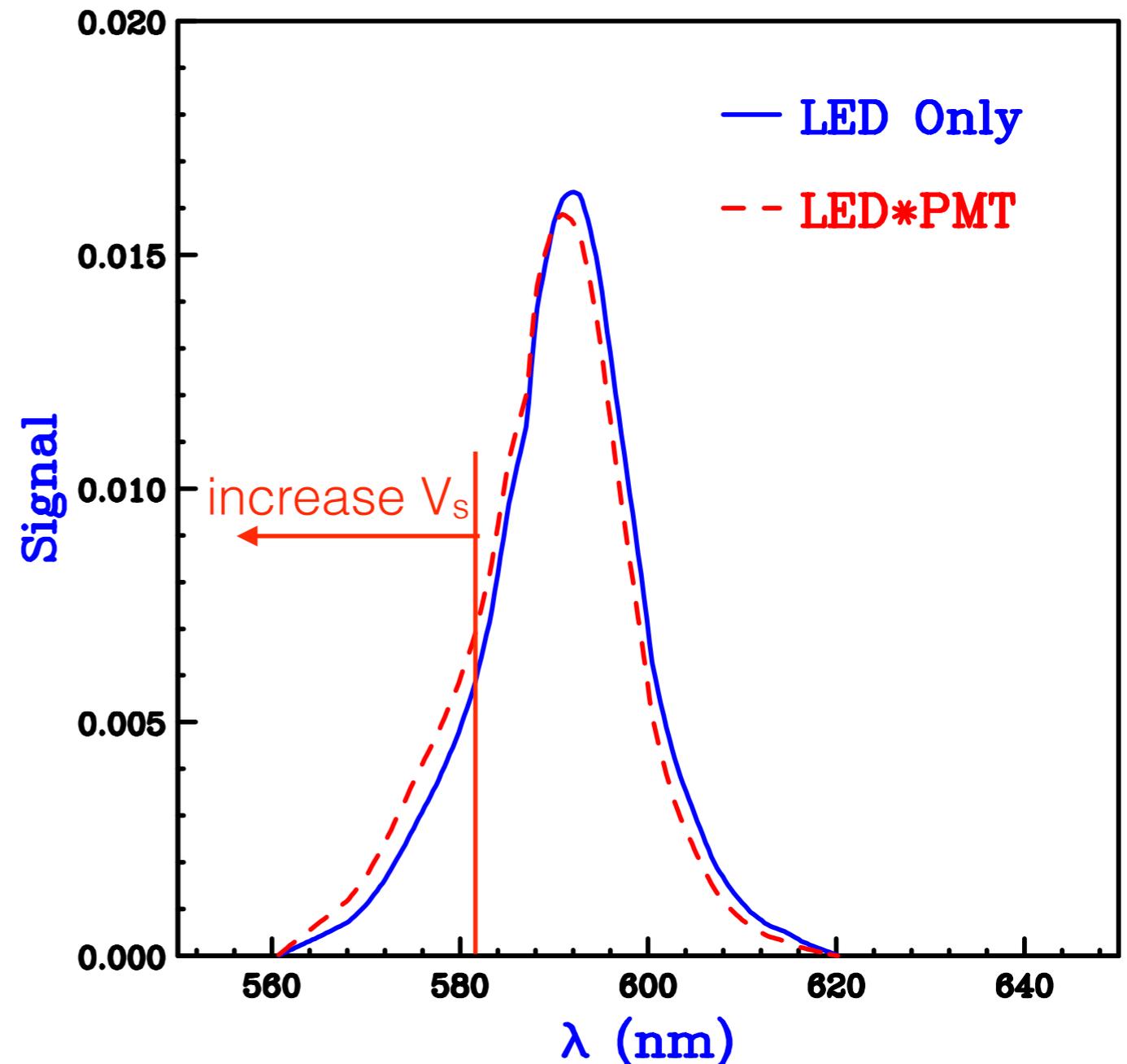
Note that there is a blue diode at 470nm and a violet diode at 460nm?
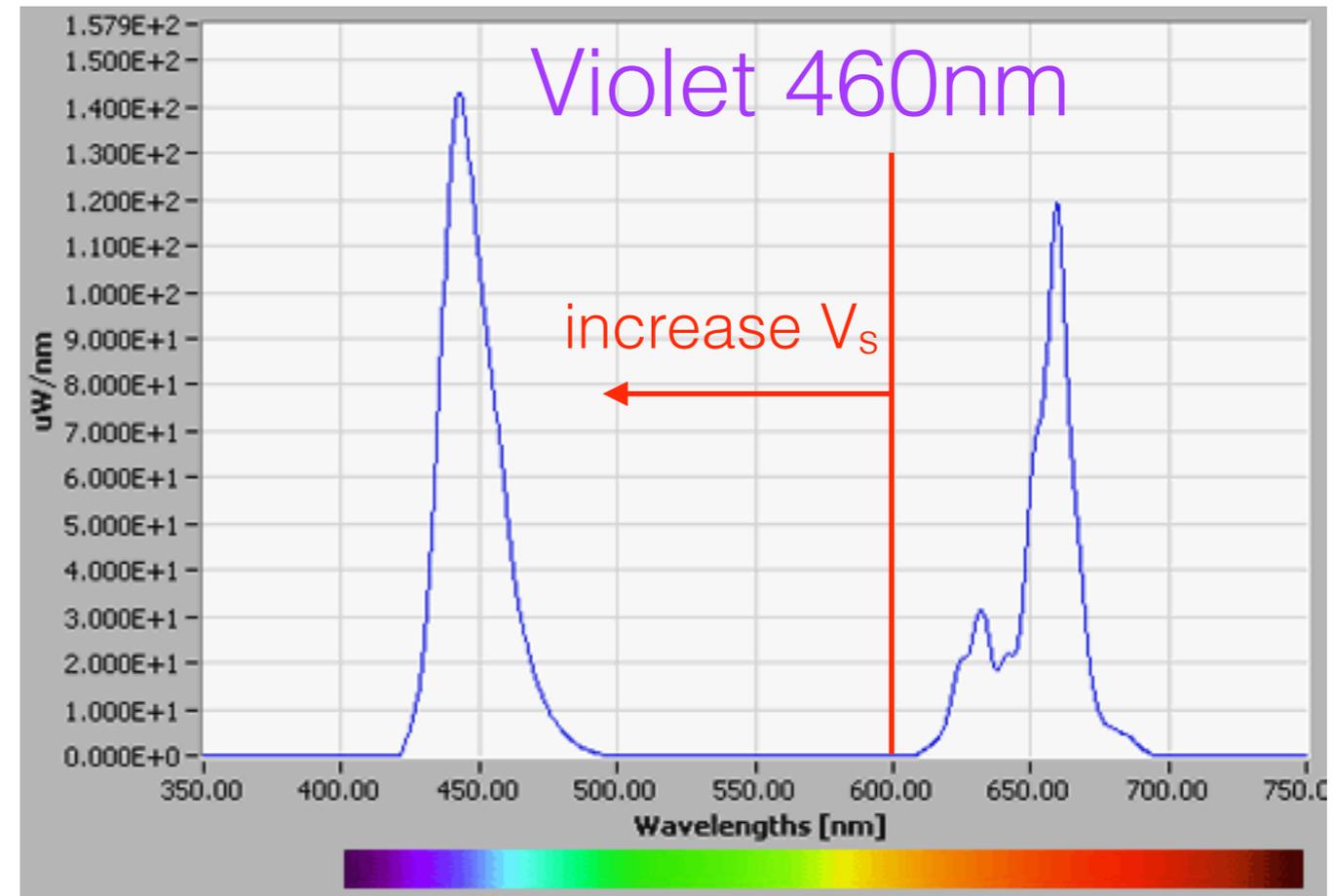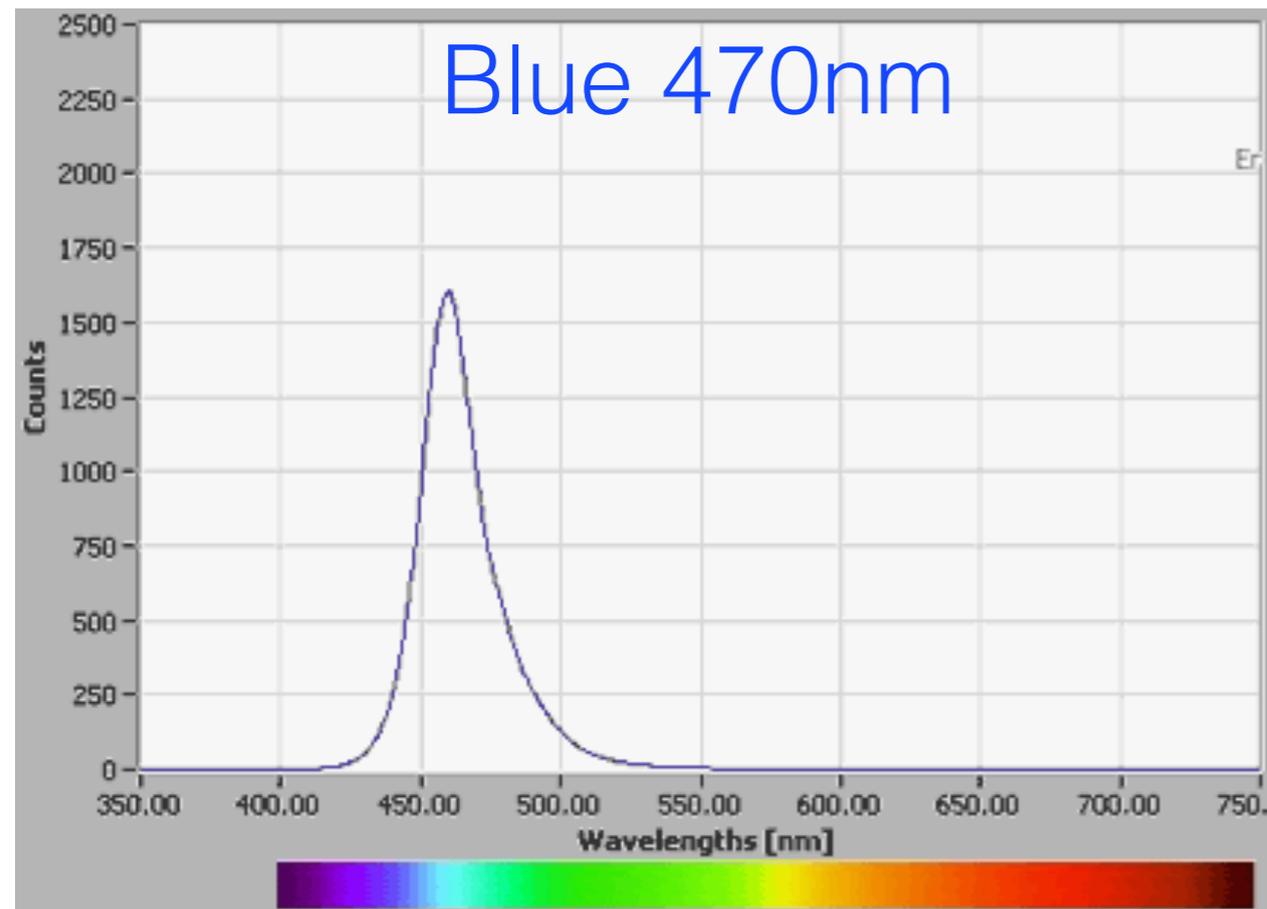
- The violet diode has peaks around 440nm and 660nm.
- The quoted value is obviously an a weighted average [human eye]?
- Our stopping potential would entirely remove the right hand peak from the photocurrent [poor PT response there also].
- We should NOT use 460nm for that diode!!

Digitizing the left-hand peak of the 460nm diode,

- LED peaks at 442.5nm

- LED*PT peaks at 442.5nm

- LED*PT average at 445.5nm

- What number do we use?

  ‣ the average?

 ‣ some fraction of the area under the red dashed curve evaluated from the small λ side?
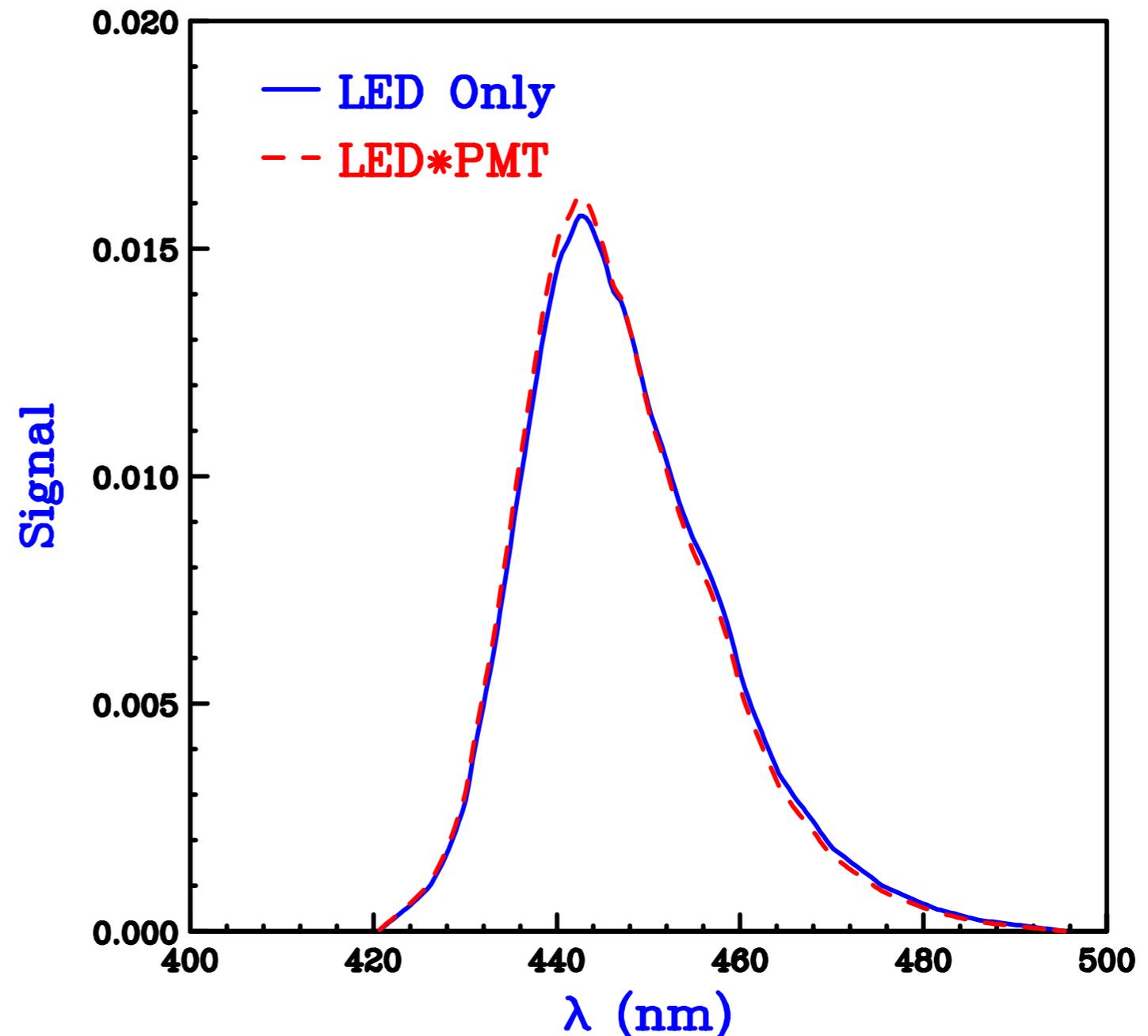
Things to investigate

- Is it important to mask the anode?
- How do we handle the finite widths of the LED signals?
  - ‣ try a more monochromatic source of the same λ?
    - ✳ try a diode laser?
    - ✳ concern about damage to the photocathode?
- Superbright LEDs does not have spectra for the UV [or IR] diodes
  - ‣ maybe we should try using only the ones that we can calibrate?
- Is result stable under intensity variations?
- More items??

Let's have fun this week!!

Extra Slides

An Introduction to Silicon Detectors

M. Swartz

Much particle physics detector technology is based upon the reverse-biased diode:
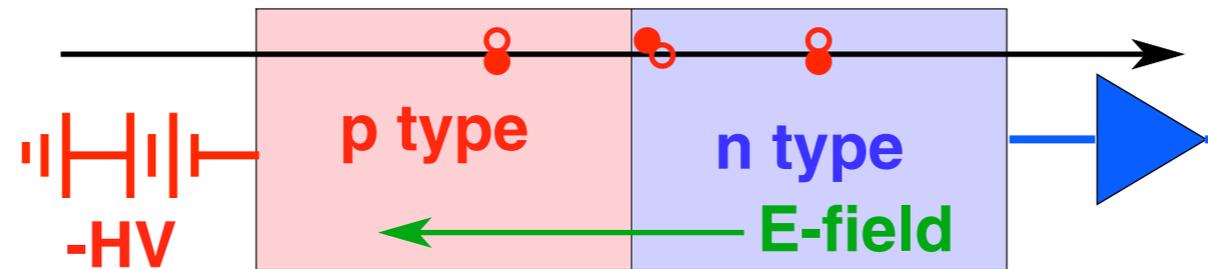
Why do we use them?

- Large signals: 22,000 e-h pairs produced in 300 μm thickness
- Supports a large internal field to sweep the carriers to electrodes
- Small leakage currents for electronics to cope with
- High precision tracking (approaching few μm) possible
- Affordable: sort of ….
- Can be made quite radiation hard compared with older tracking technologies (still not enough)

In pure silicon, the carrier densities are quite small:

$$pn = n_i^2 = N_c N_v e^{(-E_g/kT)}$$

- $n_i = 1.45 \times 10^{10}$ cm^{-3} at room temperature
  - compare with atomic number density = 3.51×10^{22} cm^{-3}
  - < 1 carrier/10^{12} atoms
  - properties of Si are totally dominated by impurities!!
- After the addition of donor or acceptor impurities, the same relationship holds
  - n-type silicon, $n \sim N_D \sim 1 \times 10^{12}$ cm^{-3} (low doping, high resistivity) or $n \sim N_D \sim 1 \times 10^{18}$ cm^{-3} (high doping, low resistivity ... n+)
  - p-type silicon, $p \sim N_A \sim 1 \times 10^{12}$ cm^{-3} (low doping, high resistivity) or $p \sim N_A \sim 1 \times 10^{18}$ cm^{-3} (high doping, low resistivity ... p+)

What actually happens at a pn junction?

Electrons diffuse into the hole-rich p-region and holes into the electron-rich n-region:

- Generates non-zero field near junction
  - depletes majority carriers
  - $x_p N_A = x_n N_D$ - thicker depletion layer for smaller doping density
- reverse biasing the junction increases the thickness of the depletion zone

$$x_n + x_p = \left[\frac{2\varepsilon_s}{e}(\phi_i + V_B)\left(\frac{1}{N_D} + \frac{1}{N_A}\right)\right]^{1/2}$$

- ϕ_i is the intrinsic potential (few V)
- dominated by smaller doping
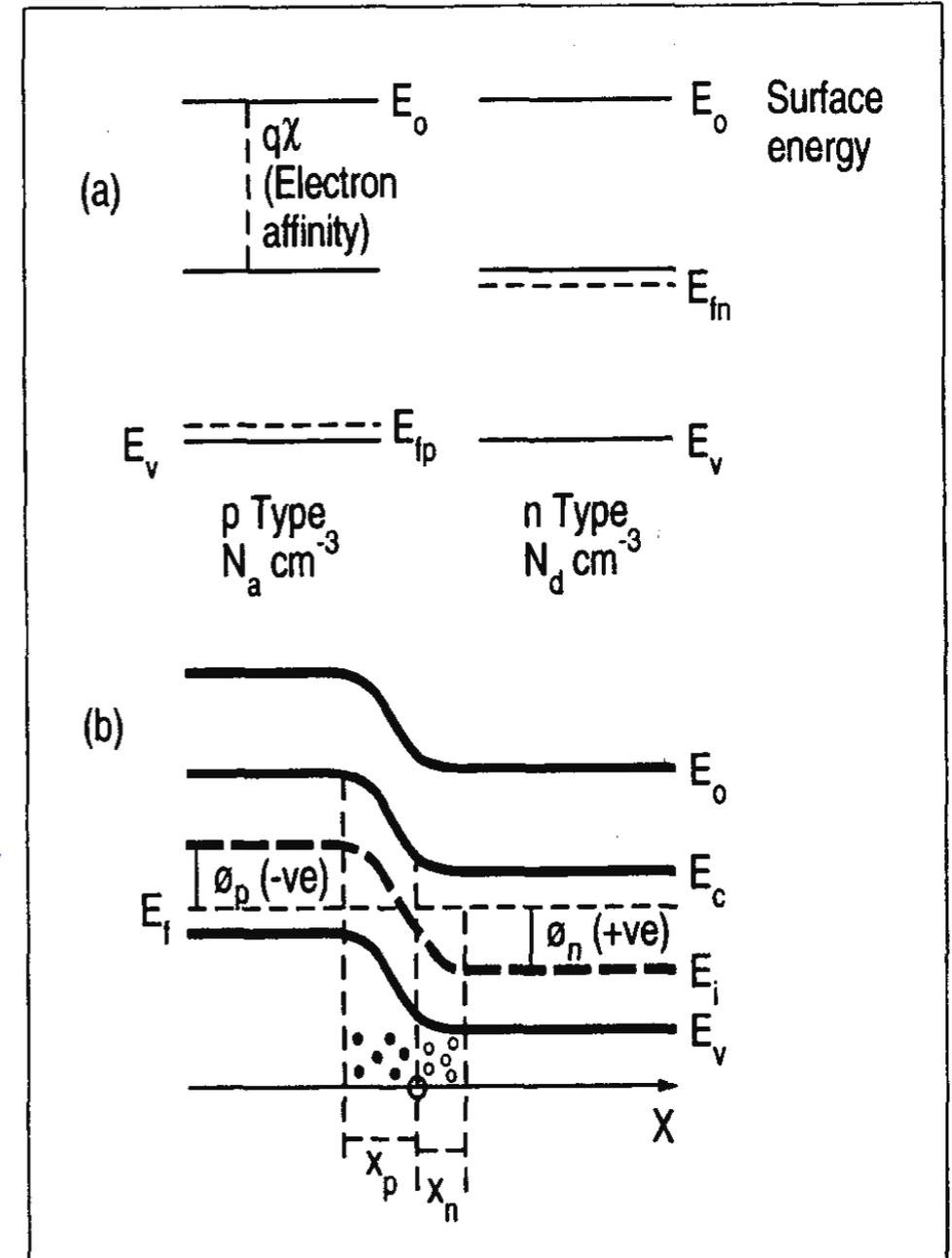
Type-Inversion

n- bulk detectors are widely believed to "type invert" after irradiation:

- pn junction moves from p+ side to n+ side
- fields vary linearly in z before and after irradiation
  - consequence of Poisson Eq: $dE_z/dz = \rho_{const} = N_{eff}$
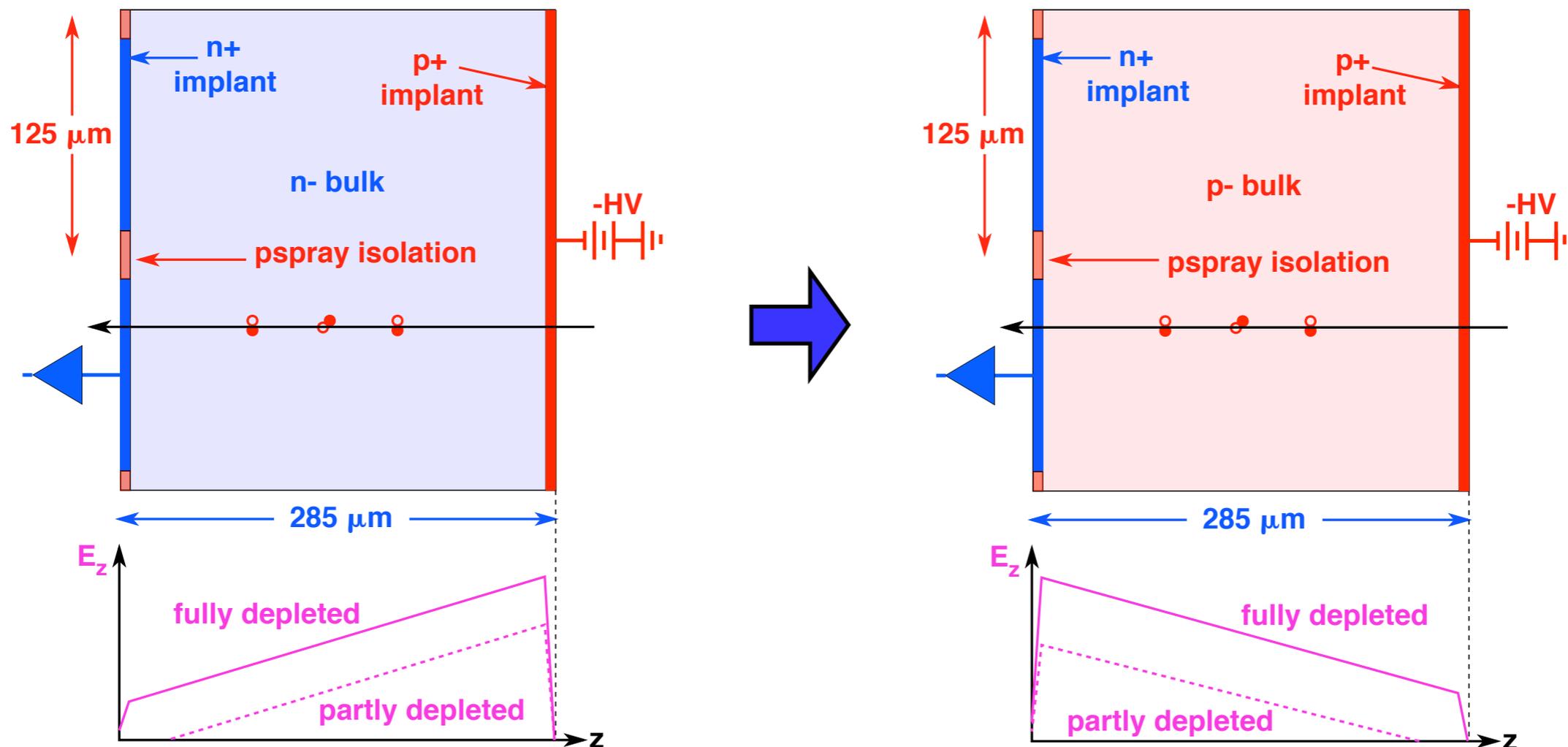

Why do we believe this?

The depletion voltage and resistivity are related to the net (uniform) density of ionized impurities N_{eff}

$$V_{\text{dep}} \simeq \frac{e}{2\varepsilon_s} N_{\text{eff}} \ell^2 \qquad \rho = \frac{6 \times 10^{15}~\Omega/\text{cm}^2}{N_{\text{eff}}}$$

description works well for unirradiated detectors

Standard picture suggests that irradiation induces acceptor defects which "invert" the n- substrate,

- $V_{\text{max-Q}} \sim V_{\text{min-C}} \sim V_{\text{dep}}$ has a minimum at a modest fluence
- interpreted to mean that substrate is now p-doped
- starting with n-substrate increases detector lifetime

[RD48-NIMA 465(2001) 60]

▲ Carbon-enriched (P503)
■ Standard (P51)
O-diffusion 24 hours (P52)
O-diffusion 48 hours (P54)
▲ O-diffusion 72 hours (P56)

Carbonated

type inversion

Standard

n-type   "p"-type

Oxygenated

$|N_{\text{eff}}|~[10^{12}\text{cm}^{-3}]$

$V_{\text{dep}}~[\text{V}]~(300~\mu\text{m})$

$\Phi_{24~\text{GeV/c proton}}~[10^{14}~\text{cm}^{-2}]$
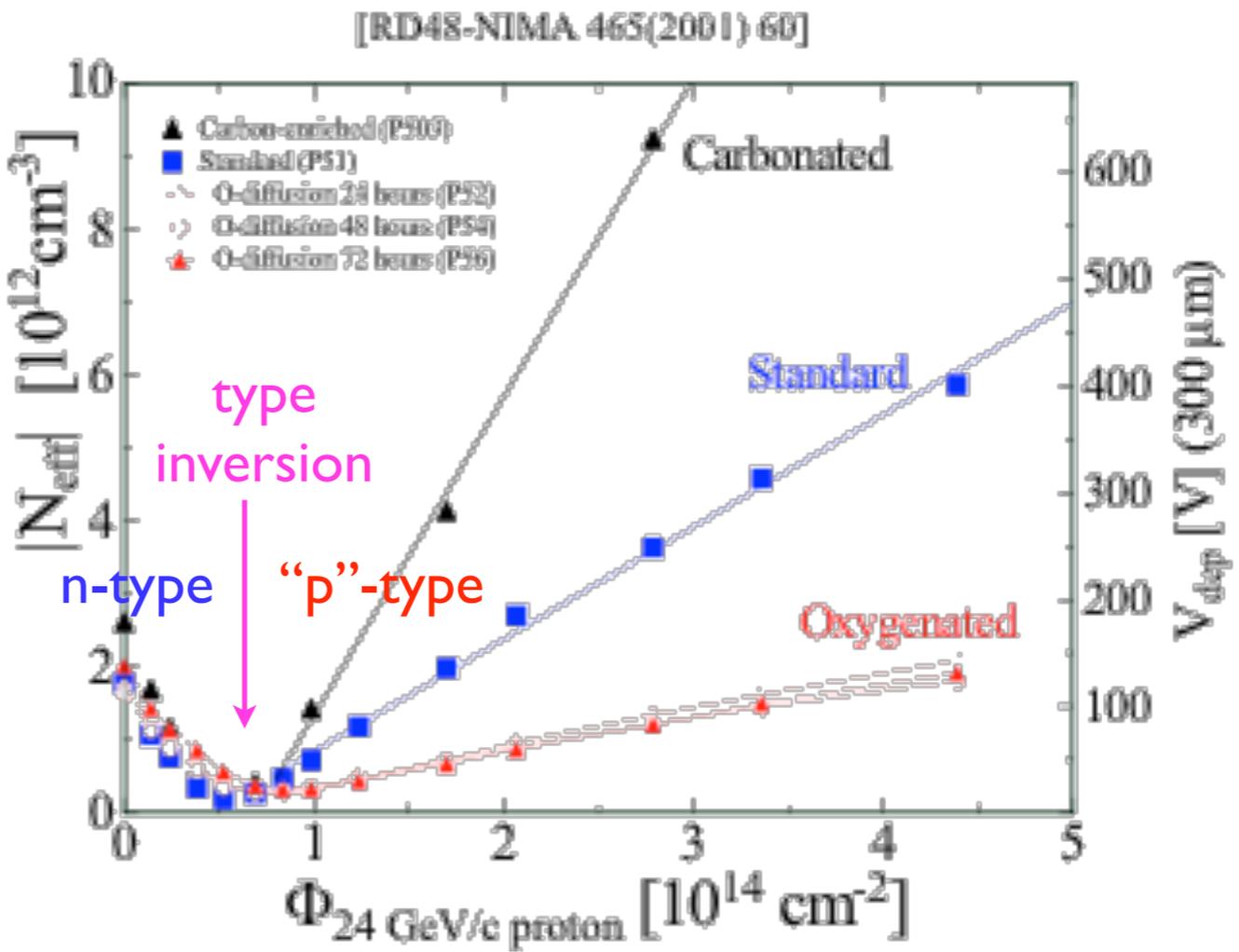
Nice, simple picture ... is this really what happens?

Radiation Damage – I. Radiation Induced Defects

particle $\longrightarrow$ Si_s $\longrightarrow$ **Frenkel pair**
Vacancy + Interstitial

$E_K > 25$ eV $\qquad$ $E_K > 5$ keV

Point Defects (V-V, V-O ..) $\qquad$ clusters
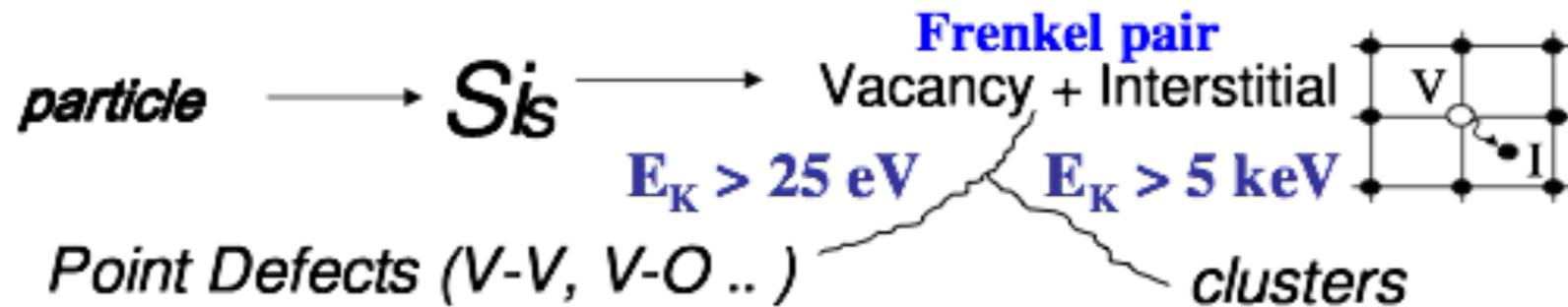
Initial distribution of vacancies after 10^{14} particles/cm^2

10 MeV protons $\qquad$ **24 GeV/c protons** $\qquad$ **1 MeV neutrons**
36824 vacancies $\qquad$ 4145 vacancies $\qquad$ 8870 vacancies
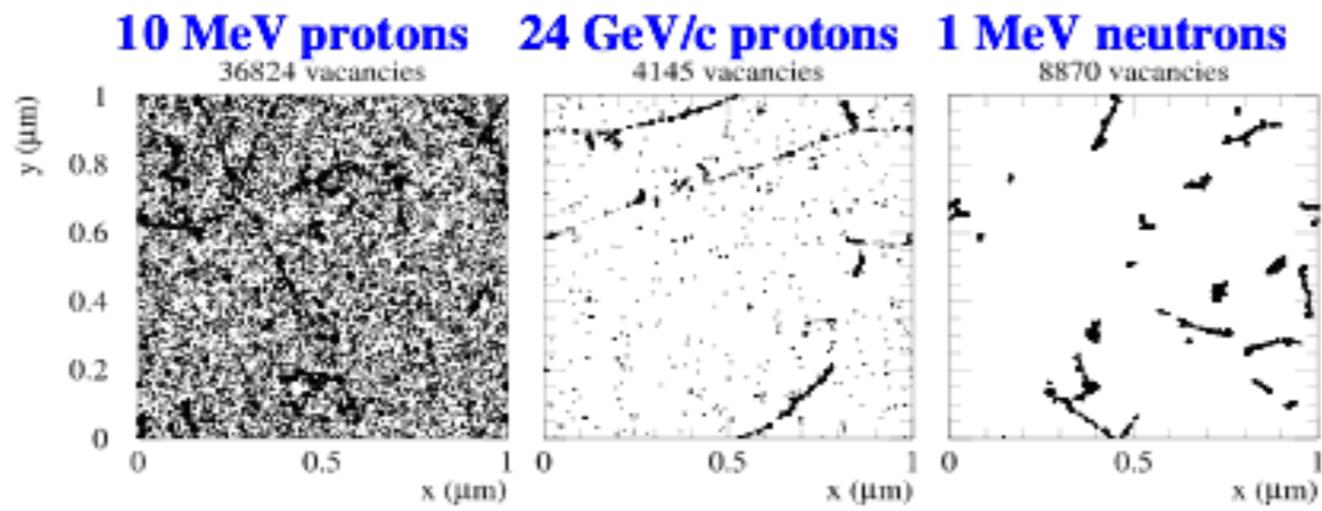
More point defects $\qquad$ **Mainly clusters**

[Mika Huhtinen NIMA 491(2002) 194]

Main radiation induced defects in Si

E_c

$\overline{V\text{-}O}$ $\quad$ $S\overline{D/TD}$ $\quad$ $\underline{V_2}$

$\underline{C_i\text{-}O_i}$ $\quad$ $\overline{V_2\text{-}O}$ $\qquad\qquad$ E_i

E_v

Point -Defects $\qquad$ **Clusters**

6

M. Bruzzi

Radiation Damage – II. Leakage Current

◼ Change of Leakage Current (after hadron irradiation)
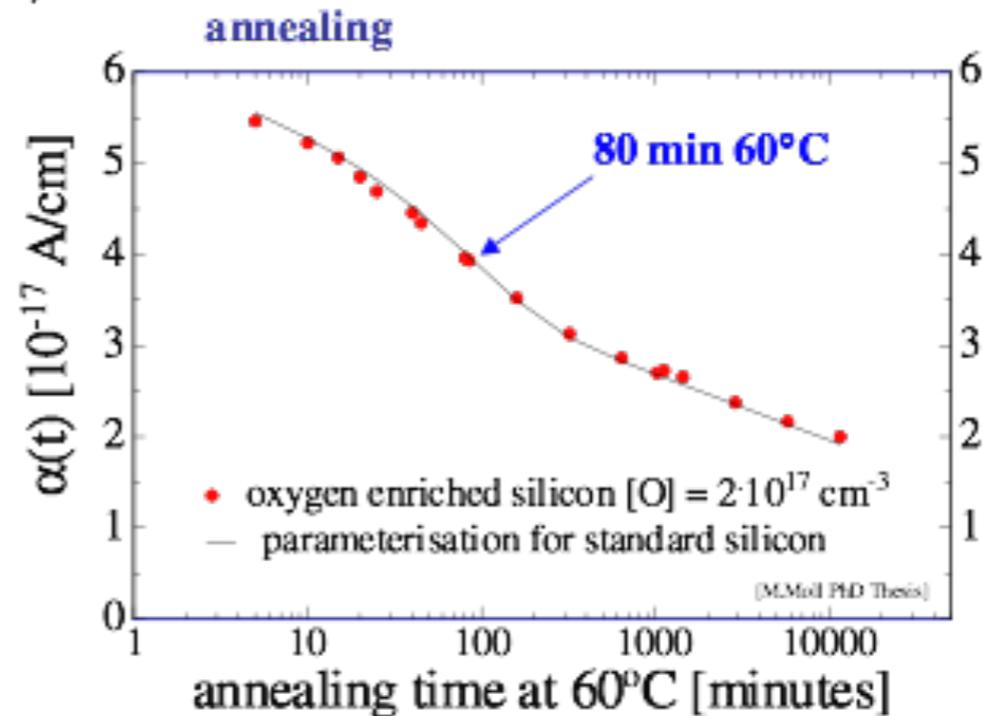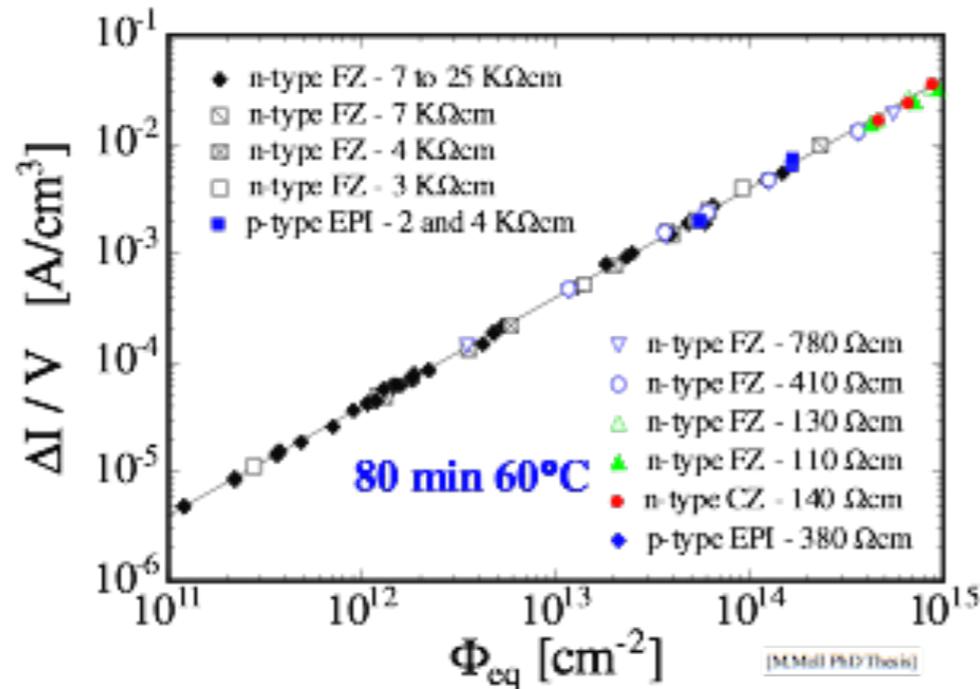
- **Damage parameter α (slope in figure)**

$$\alpha = \frac{\Delta I}{V \cdot \Phi_{eq}}$$ **Leakage current per unit volume and particle fluence**

$$\alpha(60°C, 80\,\text{min}) = (3.99 \pm 0.03) \times 10^{-17}\, \frac{A}{cm}$$

- **α is constant over several orders of fluence and independent of impurity concentration in Si**
  ⇨ **can be used for fluence measurement**

- **Leakage current decreasing in time (depending on temperature)**
- **Strong temperature dependence:**

$$I(T) \; \alpha \; T^2 \exp\left(-\frac{E_g}{KT}\right)$$

Consequence:
   Cool detectors during operation!

7

M. Bruzzi

Radiation Damage – III. Effective doping concentration

■ **Change of Depletion Voltage V_{dep} (N_{eff}) in high resistivity n-type FZ Si**
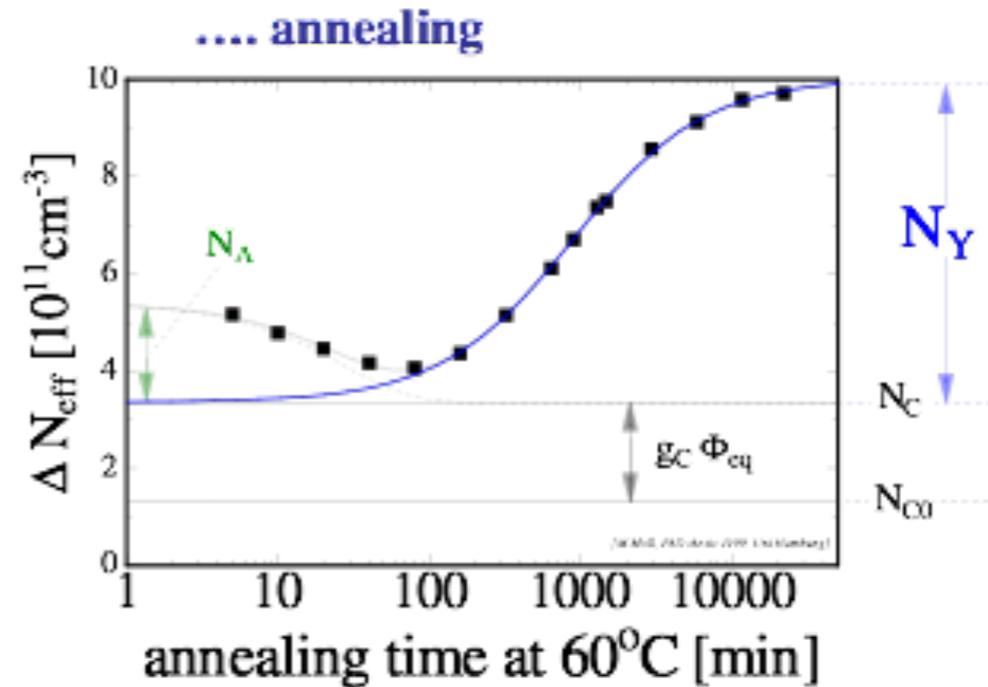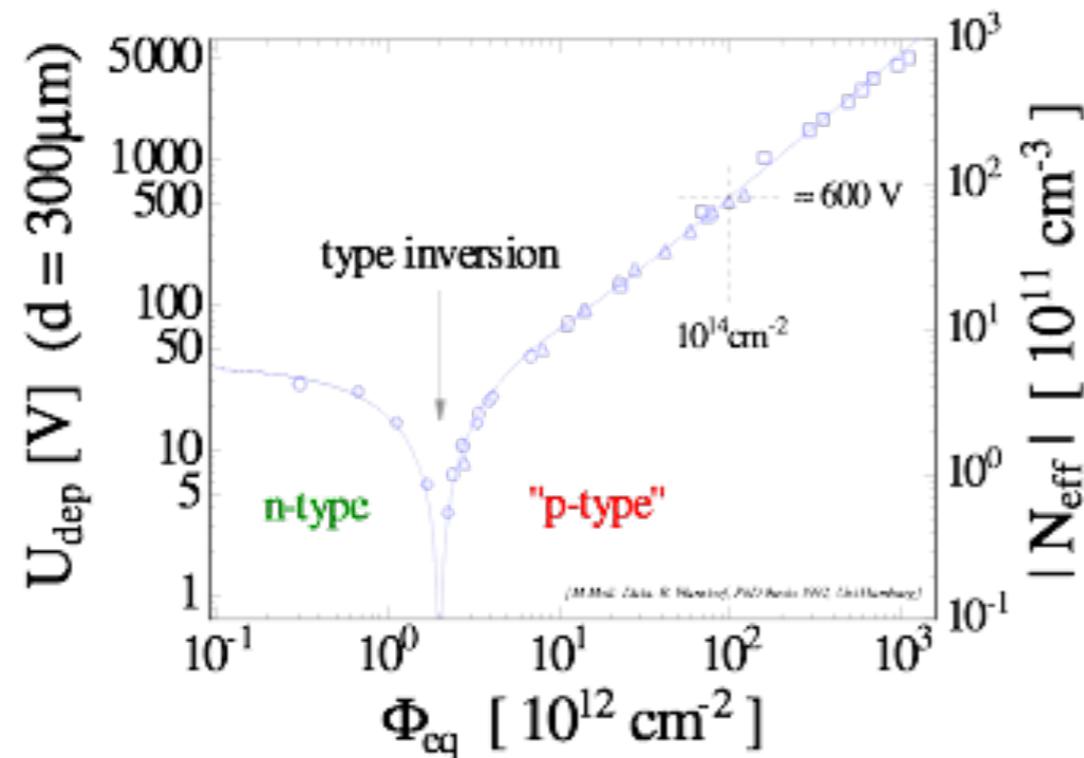
➢ Full depletion voltage too high to fully deplete detectors at very high fluences

➢ **"Type inversion"**: N_{eff} changes from positive to negative (**Space Charge Sign Inversion**)

Short term: "Beneficial annealing"
Long term: "Reverse annealing"
- time constant depends on temperature:
  - ~ 500 years (-10°C)
  - ~ 500 days (20°C)
  - ~ 21 hours (60°C)
- **Consequence:** Detectors must be cooled even when the experiment is not running!

8

M. Bruzzi

Radiation Damage – V. Trapping

■ Deterioration of Charge Collection Efficiency (CCE) by trapping

Trapping is characterized by an effective trapping time τ_{eff} for electrons and holes:

$$Q_{e,h}(t) = Q_{0e,h} \exp\left(-\frac{1}{\tau_{eff\ e,h}} \cdot t\right) \qquad \text{where} \qquad \frac{1}{\tau_{eff\ e,h}} \propto N_{defects}$$

Increase of inverse trapping time $(1/\tau)$ with fluence ..... and change with time (annealing):

10

M. Bruzzi

38